

N° d'ordre : 3027

THÈSE

présentée

DEVANT L'UNIVERSITÉ DE RENNES I

pour obtenir

le grade de : ***DOCTEUR DE L'UNIVERSITÉ DE RENNES I***

Mention : Traitement du signal et télécommunications

par

Elise Arnaud

Équipe d'accueil : VISTA (IRISA, RENNES)

École doctorale : Mathématiques, Informatique, Signal, Électronique et
Télécommunications

Composante Universitaire : IFSIC, Université de Rennes 1

Titre de la thèse :

**Méthodes de filtrage pour du suivi dans des séquences d'images -
Application au suivi de points caractéristiques**

Soutenue le 24 novembre 2004, devant la commission d'examen

Composition du jury :

M.	François	LE GLAND	Président
M.	Michel	DHOME	Rapporteurs
M.	Arnaud	DOUCET	
M.	Bruno	CERNUSCHI-FRÍAS	Examineurs
M.	Etienne	MÉMIN	
M.	Patrick	PÉREZ	

Remerciements

Je tiens à remercier François Le Gland pour avoir accepté de présider ce jury de thèse. Je remercie également Arnaud Doucet et Michel Dhome pour avoir apporté leurs points de vue sur ces travaux en acceptant d'en être les rapporteurs. Je suis également reconnaissante à Patrick Pérez d'avoir accepté d'examiner ce travail et de s'y être penché avec tant d'attention.

Je suis particulièrement reconnaissante à Bruno Cernuschi-Frías pour son soutien, son enthousiasme et le grand intérêt qu'il a porté à mes travaux.

Un grand Merci à Etienne Mémin. Tout simplement, je pense que j'aurais difficilement pu être mieux encadrée, à la fois scientifiquement et humainement.

Merci à tous ceux qui m'ont entouré et aidé durant ces trois années. Merci en particulier à François Rousseau pour m'avoir fait rire aux éclats tous les jours pendant 2 ans et demi, à Anthony Remazeilles pour avoir été mon "compagnon de route" durant cette thèse et à Anne Cuzol et Thomas Bréhard pour tous les moments passés ensemble (du matin au soir, du lundi au dimanche ...).

Enfin, merci à Tangi pour avoir patienté avant que je ne m'aperçoive de quelque chose.

Table des matières

Introduction générale	9
I Filtrage bayésien pour les modèles de chaînes de Markov cachées	13
Introduction et modélisation du problème	15
1 Méthodes non aléatoires	21
1.1 Modèle linéaire avec bruits additifs gaussiens : filtre de Kalman	21
1.2 Extensions au cas non-linéaire	24
1.2.1 Filtre de Kalman étendu	25
1.2.2 Filtre de Kalman unscented	26
1.2.3 Filtre par mélange de lois gaussiennes	28
1.3 Méthodes numériques par maillage de l'espace d'état	29
1.4 Synthèse	30
2 Méthodes séquentielles de Monte Carlo	33
2.1 Méthodes d'échantillonnage par Monte Carlo	34
2.1.1 Le principe de Monte Carlo	34
2.1.2 Algorithme d'acceptation rejet	35
2.1.3 Échantillonnage pondéré	36
2.2 Méthodes séquentielles de Monte Carlo	37
2.2.1 Échantillonnage pondéré séquentiel	38
2.2.2 Problème de dégénérescence et taille efficace du N-échantillon	40
2.2.3 Échantillonnage pondéré séquentiel avec rééchantillonnage	41
2.3 Stratégies d'échantillonnage	42
2.3.1 Filtre bootstrap	42
2.3.2 Fonction d'importance optimale	46
2.3.3 Autres stratégies, quand la fonction d'importance optimale n'est pas disponible	48
2.4 Stratégies de rééchantillonnage	51
2.4.1 Sélection des particules	52
2.4.2 Perturbation des particules	53
2.5 Réduction de variance par rao-blackwellisation	55
2.6 Synthèse	59

3 Filtrage mono-objet en environnement bruité	61
3.1 Fenêtre de validation des mesures	62
3.2 Hypothèses d'association et modèle de fausses alarmes	64
3.3 Modèle linéaire gaussien : filtre à association de données probabiliste	65
3.4 Modèle non linéaire : utilisation de filtres particuliers	70
3.5 Synthèse	72
Conclusion	73
 II Modèles pour la trajectographie dans des séquences d'images	 75
Introduction et problématique	77
 4 Approches existantes	 81
4.1 Représentation de l'objet	82
4.1.1 Quand la nature de l'objet est connue	82
4.1.2 Quand la nature de l'objet est inconnue	84
4.2 Caractérisation de l'objet	86
4.3 Stratégies algorithmiques	89
4.3.1 Approche <i>Top-Down</i>	90
4.3.2 Approche <i>Bottom-Up</i>	91
4.3.3 Choix de la fonction d'importance	92
4.4 Synthèse et discussion	94
 5 Modélisation proposée : prise en compte explicite des données images	 97
5.1 Du modèle classique au modèle conditionnel	97
5.2 Filtres conditionnels	100
5.2.1 Filtre linéaire conditionnel	100
5.2.2 Filtre non linéaire conditionnel	105
5.3 Modèles conditionnels non linéaires gaussiens : deux cas particuliers	108
5.3.1 Vraisemblance monomodale	109
5.3.2 Vraisemblance multimodale	110
5.3.3 Fenêtre de validation	113
5.3.4 Récapitulatif	117
5.4 Synthèse	117
Conclusion	119
 III Application au suivi de points caractéristiques et objets plans	 121
Introduction	123
 6 Approches existantes	 125
6.1 Suivi de points d'intérêt	125
6.1.1 Détection - sélection des points d'intérêts	126
6.1.2 Méthodes de corrélation	127

6.1.3	Méthodes différentielles	130
6.2	Suivi de motifs plans	133
6.3	Synthèse	136
7	Algorithmes proposés pour la trajectographie de points	139
7.1	Modèle dynamique estimé par une approche différentielle	140
7.1.1	Estimation robuste d'un mouvement paramétrique	140
7.1.2	Équation dynamique linéaire ou non linéaire ?	141
7.1.3	Choix de la covariance du bruit d'état	142
7.2	Modèle d'observation utilisant une mesure de similarité	145
7.2.1	Estimation de la covariance du bruit de mesure	145
7.2.2	Mise à jour de la caractérisation de référence	151
7.3	Étapes complémentaires et algorithme final	153
7.4	Extension à un algorithme robuste aux ambiguïtés	155
7.5	Résultats expérimentaux	159
7.5.1	Résultats du TFLC	159
7.5.2	Résultats du TFNLC	163
7.5.3	Robustesse aux déformations géométriques et aux occlusions	167
7.5.4	Validation de l'algorithme dédié au traitement des ambiguïtés	170
7.6	Application au suivi de visages	175
7.6.1	Algorithme de détection de visages	176
7.6.2	Utilisation d'une méthode différentielle	177
7.6.3	Algorithme de suivi des yeux	178
7.6.4	Résultats expérimentaux	179
7.7	Synthèse	180
8	Extension à la trajectographie de motifs plans	183
8.1	De la trajectographie d'un point vers le suivi d'un groupe de points	183
8.2	Modèle pour la trajectographie d'un nuage de points	184
8.2.1	Modèle dynamique en deux étapes	184
8.2.2	Modèle d'observation	186
8.2.3	Algorithme final	186
8.3	Résultats expérimentaux	187
8.4	Synthèse	191
	Conclusion	195
	Conclusion générale et perspectives	197
	Annexes	201
A	Estimation de la loi de filtrage $p(\mathbf{x}_k \mathbf{z}_{1:k})$	203

B	Estimateurs de variance minimale	205
B.1	Estimateur de variance minimale	205
B.2	Estimateur linéaire non biaisé de variance minimale	205
B.3	Estimateur linéaire conditionnel de variance minimale	206
C	Filtre linéaire conditionnel : demonstrations	211
C.1	Description du modèle et notations	211
C.2	Quelques propriétés utiles	212
C.3	Expression récursive de $\hat{\mathbf{x}}_{k+1 k}$	213
C.4	Expression récursive de la matrice de covariance conditionnelle de l'erreur d'estimation	215
C.5	Récapitulatif : équations du filtre linéaire conditionnel	216
D	Modèles conditionnels non linéaires gaussiens : deux cas particuliers	219
D.1	Vraisemblance monomodale	219
D.2	Vraisemblance multimodale	220
D.3	Fenêtre de validation	222
	Liste des algorithmes	224
	Liste des figures	225
	Bibliographie	231
	Publications	249

Introduction générale

L'utilisation de capteurs vidéo offre un moyen peu coûteux, flexible et non-invasif d'obtenir une grande quantité d'informations spatio-temporelles sur la scène observée. Ces capteurs sont aujourd'hui largement employés dans des applications diverses, comme l'analyse de phénomènes physiques ou biologiques (études des flux en mécanique des fluides expérimentale, vidéo-microscopie, examens chirurgicaux par endoscopie, etc.), ou la construction de systèmes robotiques indépendants.

Parmi la quantité d'informations transportées par une séquence d'images, l'étude de la composante cinématique a une importance cruciale pour la compréhension de l'environnement. Elle donne accès, entre autres, à une meilleure connaissance de la structure 3D de la scène, et permet de détecter les zones mobiles et leurs déformations. Une étude plus locale de l'information de mouvement et son intégration temporelle amènent à considérer les problèmes de suivi dans une séquence d'images.

Réaliser un suivi robuste d'une entité quelconque est un des enjeux de la vision par ordinateur. En raison de nombreuses difficultés, une simple intégration temporelle de l'information cinématique n'est pas suffisante. Celles-ci peuvent être liées d'une part au processus d'acquisition (introduction de bruits, de flous de bougé, variations d'illumination), et d'autre part au fait que la scène tridimensionnelle n'est perçue que par sa projection dans l'image (situations d'occultations, changements d'aspect de l'objet). De plus, les seules données exploitables sont des mesures d'intensité lumineuse, qui peuvent se révéler très similaires dans différentes régions de l'image. Pour contourner ces difficultés, la mise en œuvre d'un algorithme de suivi nécessite la définition de trois éléments principaux :

1. Une caractérisation de l'entité. Construite à partir de données issues de l'image (contour, modèle de luminance, histogramme de couleur, etc.), la caractérisation doit permettre de reconnaître ou de détecter l'objet. Elle doit être à la fois persistante tout au long de la séquence et spécifique pour permettre de différencier l'objet suivi des autres éléments de la scène.
2. Une contrainte pour limiter la zone de recherche de la nouvelle position de l'objet. Cette contrainte peut se traduire par un modèle de mouvement décrivant grossièrement l'évolution de l'entité.
3. Une stratégie de recherche performante. À chaque pas de temps, l'espace des positions possibles doit être exploré de manière efficace, prenant en compte la position à l'instant précédent et la contrainte de mouvement.

La modélisation du problème de trajectographie de façon probabiliste par un système dynamique, dont l'estimation est apportée par un algorithme de filtrage, permet de réunir ces trois éléments. Ces algorithmes sont connus pour leur robustesse aux minima locaux qui mettent en défaut les méthodes classiques déterministes. Ils reposent sur une description du système par une chaîne de Markov cachée, décrite par une loi dynamique (aussi appelée modèle d'évolution) et une vraisemblance des données. La vraisemblance décrit le lien entre la position recherchée et la caractérisation de l'objet, alors que la loi dynamique décrit son modèle d'évolution. Cette dernière est généralement porteuse d'un *a priori* qui rend compte d'une modélisation du mouvement de l'objet. Or, une telle modélisation n'est plus possible lors de la construction d'un algorithme de suivi général, où aucune information n'est disponible sur l'entité. Dans ce cas, une solution intéressante repose sur l'estimation en ligne de la dynamique à partir d'informations extraites des images.

Contributions

L'étude décrite dans cette thèse se situe dans ce contexte. Nous nous sommes intéressés à la modélisation du problème de trajectographie dans une séquence d'images, par un système dynamique, dont le modèle d'évolution est estimé à chaque pas de temps sur les images. L'étude de ces systèmes a soulevé un problème théorique. En effet, le choix d'un tel modèle d'évolution met en évidence les limitations de la représentation par un modèle simple de chaîne de Markov cachée, qui ne modélise pas la dépendance des éléments du système aux images. Partant de ce constat, les principales contributions de cette thèse s'organisent en deux parties.

Reformulation du modèle de chaîne de Markov cachée prenant en compte de manière explicite les données images

Dans un premier temps, une reformulation du modèle de chaîne de Markov cachée est proposée. Celle-ci rend la séquence d'images explicite dans la modélisation. Elle a l'avantage d'autoriser l'estimation en ligne de la loi dynamique, et permet de construire des algorithmes ne nécessitant pas d'information *a priori*. Les algorithmes de filtrage permettant de résoudre cette nouvelle représentation d'un système dynamique sont dérivés. Deux instances de ces filtres se distinguent : la première se rapporte au cas où les équations du système sont linéaires, et la seconde au cas où les équations sont non linéaires. Enfin, l'étude de quelques systèmes plus spécifiques met en évidence l'intérêt, au niveau théorique, du choix d'une loi dynamique informative, estimée sur la séquence.

Validation du modèle proposé dans le cadre d'applications de suivi de points caractéristiques et de motifs plans

Dans un second temps, nous nous sommes intéressés à la validation pratique de la modélisation proposée. En raison de la difficulté de poser un modèle de mouvement *a priori* d'un point de l'image, la construction d'algorithmes pour la trajectographie de points caractéristiques nous a semblé être un cadre approprié. Le système mis en œuvre est entièrement estimé sur la séquence. Il repose sur l'hypothèse la plus couramment utilisée pour définir

les méthodes de suivi, à savoir l'hypothèse de conservation de la luminance du point et de son voisinage le long de la trajectoire. Plus précisément, nous proposons d'associer des mesures fournies par une technique de mise en correspondance à un modèle dynamique défini à partir d'un mouvement estimé selon la contrainte de flot optique. Les résultats très satisfaisants obtenus ont amené à étendre cette approche au suivi d'un motif plan.

Description du document

Ce document est organisé en trois parties selon le plan suivant.

Partie I : Filtrage bayésien pour les modèles de chaînes de Markov cachées

La première partie dresse un panorama des algorithmes de filtrage pour un objet unique. Elle est composée de trois chapitres. Le premier traite des algorithmes dits "non aléatoires", à savoir le filtre de Kalman et ses extensions, ainsi que les méthodes numériques par maillage d'état. Après une présentation générale du principe de Monte Carlo, les méthodes séquentielles de Monte Carlo, ou filtres particuliers, sont présentées en détails dans le second chapitre. Nous insisterons sur les deux étapes essentielles de ces algorithmes qui sont la phase d'échantillonnage et la phase de sélection. Enfin, le troisième chapitre est dédié au problème du filtrage d'un objet unique dans un environnement bruité. Le problème d'association des mesures et les moyens de le résoudre y sont décrits.

Partie II : Modèles pour la trajectographie dans des séquences d'images

La seconde partie est consacrée à la construction des modèles dans le contexte de notre étude, à savoir le suivi dans des séquences d'images. Elle débute par un chapitre d'état de l'art, où les principales modélisations de la vraisemblance et de la dynamique dans ce contexte sont données. Les différentes constructions du modèle d'évolution et les choix de caractérisations y sont décrits. Cet état de l'art met en évidence certaines limitations du modèle simple de chaîne de Markov cachée. C'est pourquoi, dans le deuxième chapitre, une nouvelle formulation du problème de trajectographie prenant en compte de manière explicite les données images est proposée. Les avantages de cette formulation ainsi que la description des filtres associés sont donnés.

Partie III : Application au suivi de points caractéristiques et objets plans

La troisième et dernière partie est dédiée aux applications en trajectographie de points caractéristiques et de motifs plans. Elle se décompose en trois chapitres. Dans le premier, les méthodes existantes sont décrites. Elles peuvent schématiquement être classées en deux catégories, appelées méthodes de corrélation et méthodes différentielles. Afin de profiter des avantages de ces deux approches pour le suivi de points, nous proposons de combiner ces deux formulations complémentaires dans une modélisation probabiliste. Pour cela, le système mis en place combine une équation d'évolution inférée par un mouvement estimé par le biais d'une méthode différentielle et des mesures apportées par une approche de corrélation. Ce système, les trois algorithmes proposés qui en découlent,

et leurs résultats sur des séquences réelles sont décrits dans le second chapitre. L'extension de cette approche au suivi d'un motif plan est expliquée dans le dernier chapitre. Elle permet d'introduire une cohérence spatiale en utilisant un modèle homographique.

Enfin, la conclusion générale présente une synthèse de cette étude et quelques perspectives liées à ces travaux.

Première partie

Filtrage bayésien pour les modèles de chaînes de Markov cachées

Introduction et modélisation du problème

L'estimation d'un système dynamique à partir d'observations bruitées, et éventuellement incomplètes est un problème central dans de nombreuses applications. Parmi celles-ci, on peut citer le suivi d'entités en vision par ordinateur, l'analyse de signaux radars, la localisation d'objet dans un espace 3D en robotique, la modélisation de données financières, le traitement de la parole, l'étude de séries temporelles, la déconvolution de signaux numériques, ou le diagnostic médical.

De façon générale, le système est modélisé par une chaîne de Markov cachée, à temps discret. On cherche donc à estimer la valeur du processus aléatoire markovien (processus d'état) à partir des réalisations obtenues à différents pas de temps du processus d'observation (ou processus de mesure). Le système est décrit par (a) la distribution du processus d'état à l'instant initial, (b) un modèle d'évolution de l'état, fixé *a priori* et (c) un modèle de vraisemblance qui relie l'observation à l'état.

L'ajout d'une connaissance *a priori* permet l'utilisation de méthodes bayésiennes. Dans ce cadre, la résolution du problème d'inférence passe par la connaissance de la loi *a posteriori*, i.e. la loi du processus d'état connaissant les observations, qui transporte l'ensemble des informations statistiques disponibles sur le processus à estimer. À partir de la connaissance ou d'une approximation de cette distribution, il est possible de construire un estimateur de l'état, optimal selon un certain critère. Les estimateurs les plus couramment utilisés sont la moyenne ou le mode *a posteriori*. De plus, une mesure de confiance de cet estimateur est disponible, par un calcul de covariance ou d'entropie.

Selon la nature de l'application, on peut distinguer trois problèmes. Le problème de **filtrage** consiste à rechercher la distribution de l'état à l'instant courant connaissant l'ensemble des observations passées et présentes. Le **lissage** utilise en plus les mesures futures. Il correspond à une analyse "batch" de la série temporelle et est utile lorsque la résolution du problème d'estimation est différée par rapport à l'acquisition des observations. Enfin, la distribution recherchée pour le problème de **prédiction** est la loi d'un état futur connaissant l'ensemble des observations passées et présentes.

Dans ce document, nous nous intéressons au suivi dans des séquence d'images. Un grand nombre d'applications en vision par ordinateur utilise le résultat d'un algorithme de suivi (en particulier suivi de points caractéristiques) pour des calculs de plus haut niveau (reconstruction, localisation, etc.). L'étape de suivi se doit donc de consommer peu de temps de calcul et de ressources. Pour répondre à de tels besoins, l'algorithme doit

être construit en supposant observer la séquence d'images séquentiellement et la position de l'entité suivie doit être estimée en-ligne. Nous nous limiterons donc dans cette étude au problème de filtrage. D'autre part, dans le but de limiter à la fois le temps de calcul entre chaque itération et la quantité de données stockées, il est essentiel de construire des algorithmes récursifs pour calculer la distribution de filtrage *a posteriori*.

Une itération d'un tel algorithme de filtrage récursif se décompose de façon schématique en deux étapes : l'étape de prédiction et l'étape de correction. L'étape de prédiction utilise le modèle d'évolution *a priori* afin d'approcher la distribution de l'état futur conditionnellement aux mesures disponibles. L'étape de correction permet de mettre à jour la distribution prédite à l'aide de la nouvelle observation et du modèle de vraisemblance associé. Suivant les caractéristiques des modèles d'évolution et de vraisemblance (linéaires / non-linéaires, gaussiens / non-gaussiens), différents algorithmes ont été proposés pour résoudre le problème d'estimation de la distribution de filtrage. Après une formalisation du problème, nous les décrirons succinctement dans cette introduction. Ils seront détaillés dans les chapitres suivants.

Formalisation du problème et filtre bayésien optimal

Par souci de simplicité, nous ferons les abus de notations suivants dans l'ensemble du document. Une variable aléatoire sera notée de la même façon que sa réalisation. Ainsi, dans le cas continu, la distribution de probabilité $\Pr(\mathbf{X} \in d\mathbf{x})$ sera notée $p(d\mathbf{x})$. Dans le cas discret, la distribution de probabilité $\Pr(\mathbf{X} = \mathbf{x})$ sera notée $p(\mathbf{x})$. De plus, si ces distributions admettent des densités par rapport à une mesure de probabilité, celles-ci seront également notées $p(\mathbf{x})$.

Le problème qui nous intéresse peut alors se formuler comme suit. On considère un système dynamique composé de deux processus aléatoires à temps discret $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$, $\mathbf{x}_k \in \mathbb{R}^{n_x}$ et $\{\mathbf{z}_k\}_{k \in \mathbb{N}^*}$, $\mathbf{z}_k \in \mathbb{R}^{n_z}$.

Le processus $\{\mathbf{x}_k\}$ est appelé *processus d'état* ou *signal*. Il est inconnu et supposé markovien de distribution initiale $p(\mathbf{x}_0)$ et de loi de transition $p(\mathbf{x}_k | \mathbf{x}_{k-1})$. Cette loi de transition décrit le modèle d'évolution de l'état. L'équation associée reliant $\mathbf{x}_{k-1}$ à $\mathbf{x}_k$ sera appelée *équation d'état* ou *équation dynamique*.

Le signal est indirectement observé par le processus $\{\mathbf{z}_k\}$, appelé *processus d'observation* ou *processus de mesure*. Les variables $\mathbf{z}_k$ sont supposées indépendantes conditionnellement au processus d'état et leur distribution ne dépend que de l'état au même instant. Le processus d'observation sera donc décrit par la distribution $p(\mathbf{z}_k | \mathbf{x}_k)$, aussi appelée *vraisemblance*. On appellera *équation de mesure* l'équation associée qui relie l'observation à l'état.

Enfin, conditionnellement à l'état à l'instant k , la mesure au même instant et l'état suivant sont indépendants :

$$p(\mathbf{z}_k, \mathbf{x}_{k+1} | \mathbf{x}_k) = p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_{k+1} | \mathbf{x}_k). \quad (0.1)$$

La propriété markovienne du processus d'état ainsi que l'indépendance conditionnelle du processus d'observation amènent à modéliser le système dynamique par une chaîne de

Markov cachée, représentée par le graphe d'indépendance de la figure 0.1. Il est entièrement décrit par les distributions suivantes :

$$p(\mathbf{x}_0) \quad (0.2)$$

$$p(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k-1}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}) \quad (0.3)$$

$$p(\mathbf{z}_k | \mathbf{x}_{0:k}, \mathbf{z}_{1:k-1}) = p(\mathbf{z}_k | \mathbf{x}_k), \quad (0.4)$$

où $\mathbf{x}_{0:k} \triangleq \{\mathbf{x}_0, \dots, \mathbf{x}_k\}$ représente le signal jusqu'au temps k et $\mathbf{z}_{1:k} \triangleq \{\mathbf{z}_1, \dots, \mathbf{z}_k\}$ désigne l'ensemble des observations jusqu'au temps k .



FIG. 0.1: Graphe de dépendance orienté de la chaîne de Markov cachée constituée des processus $\mathbf{x}_{0:k}$ et $\mathbf{z}_{1:k}$

La résolution du problème de filtrage consiste alors à construire un algorithme itératif pour l'estimation de la distribution *a posteriori* de l'état à l'instant courant conditionnellement à l'ensemble des mesures passées et présentes $p(\mathbf{x}_k | \mathbf{z}_{1:k})$. Cette distribution est appelée *distribution de filtrage*. Le **filtre bayésien optimal** nous fournit une solution exacte en deux étapes.

- En supposant $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ connue, l'*étape de prédiction* permet de connaître la distribution prédite. L'écriture de cette étape est permise par la caractère markovien du signal (0.3).

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) &= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_{1:k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \\ &= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}. \end{aligned} \quad (0.5)$$

- Durant l'*étape de correction*, l'introduction de la nouvelle observation permet de mettre à jour la prédiction. Le théorème de Bayes nous permet d'obtenir une expression de la distribution de filtrage recherchée. Cette expression est simplifiée par la propriété d'indépendance conditionnelle des mesures (0.4).

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{z}_{1:k-1}) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})} = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{\int p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) d\mathbf{x}_k}. \quad (0.6)$$

Remarquons qu'il peut être également intéressant de pouvoir accéder de façon récursive à $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$. Une expression récursive du même type que (0.5,0.6) peut être construite. Enfin, à partir d'une estimation précise de la distribution *a posteriori*, il sera possible de calculer toute quantité de la forme $\mathbb{E}_{p(\cdot | \mathbf{z}_{1:k})}[\phi(\mathbf{x}_k)]$ où ϕ est une fonction intégrable de $\mathbb{R}^{n_x}$.

Aperçu des méthodes existantes

De par la très grande dimension des intégrales mises en jeu par le filtre bayésien optimal, un calcul direct de ces expressions n'est pas possible dans le cas général. Dans le cas d'un modèle linéaire gaussien, le filtre de Kalman fournit la solution optimale sous la forme d'une expression récursive des deux premiers moments de la distribution gaussienne $p(\mathbf{x}_k|\mathbf{z}_{1:k})$. Une telle expression est déduite d'un estimateur de variance minimale.

Si l'hypothèse gaussienne est relâchée, mais que les propriétés linéaires du modèle sont conservées, l'utilisation d'un estimateur linéaire de variance minimale de l'état sachant l'ensemble des observations disponibles est un choix raisonnable. Celui-ci conduit finalement aux mêmes équations que le filtre de Kalman. Néanmoins, cet estimateur ne fournit que les deux premiers moments de $p(\mathbf{x}_k|\mathbf{z}_{1:k})$.

De façon similaire, dans le cas non linéaire, le filtre de Kalman étendu n'apporte que les statistiques des deux premiers ordres de la distribution recherchée. Cette solution non optimale est déduite d'une linéarisation au premier ordre des équations du système dynamique, qui n'est pas satisfaisante en cas de multi-modalité du système. D'autres extensions du filtre de Kalman ont été proposées pour résoudre le problème non linéaire, tels que le filtre de Kalman "unscented"¹, ou les approximations par mélange de lois gaussiennes. Bien qu'améliorant les performances du filtre de Kalman étendu, ces algorithmes restent fortement limités du fait de l'hypothèse sur le caractère gaussien de la loi *a posteriori*.

D'autres approches, appelées méthodes par grille, proposent la construction d'un maillage déterministe de l'espace d'état. Elles permettent l'obtention d'estimations numériques des intégrales du filtre optimal. La solution exacte peut être atteinte si l'espace d'état est discret, constitué d'un nombre fini (et faible) d'états possibles et si le maillage choisi correspond à ces états. L'utilisation de méthodes séquentielles de Monte Carlo constitue une alternative à ces algorithmes de grande complexité. L'intérêt de ces approches est leur facilité de mise en œuvre. La distribution de filtrage $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ est alors approchée par une somme finie pondérée de mesures de Dirac, centrées en des éléments de l'espace d'état appelés particules. Les poids associés sont choisis en utilisant le principe de l'échantillonnage pondéré. Le nuage de particules est mis à jour récursivement.

Enfin, outre les non-linéarités du système, le problème de filtrage peut être perturbé par la présence de fausses alarmes. Les fausses alarmes sont des mesures observées non issues de l'entité suivie. Ce problème est très présent en imagerie où l'arrière plan constitue une zone de forte ambiguïté et peut générer plusieurs observations par objet suivi. Dans ce cas, la résolution du *problème d'association des mesures* ("data association") permet de lever les ambiguïtés. En se restreignant à la trajectographie d'une seule entité, une première solution consiste à ne prendre en compte qu'une seule observation, à savoir la plus proche de la mesure prédite, ou celle de probabilité maximale. Évidemment, cette solution est trop restrictive et peut amener à la perte de l'entité suivie. Il est également possible de conserver l'ensemble des mesures à un instant donné en modélisant les probabilités d'association des observations à des fausses alarmes ou à l'entité suivie. La loi de filtrage peut alors s'écrire comme un mélange de lois. Deux types de filtres permettent de résoudre ce nouveau problème. D'une part le filtre à association de données probabiliste (PDAF) qui est une extension du filtre de Kalman ; d'autre part, les méthodes

¹Par souci de simplicité, la terminologie anglaise a été conservée

séquentielles de Monte Carlo qui permettent de gérer simplement les multi-modalités du système. La résolution du problème d'association de données se complexifie avec le nombre d'entités suivies. Notons que dans ce document, le suivi multi-objets ne sera pas abordé. Pour une étude détaillée de ce problème, nous renvoyons au travail de Hue [Hue 03].

L'ensemble de ces algorithmes seront décrits en détail dans cette partie. Le premier chapitre porte sur les méthodes non aléatoires (filtre de Kalman et extensions au cas non linéaire, méthodes par grille), alors que le deuxième chapitre est dédié aux méthodes séquentielles de Monte Carlo. Le problème spécifique de l'association de mesures, et les algorithmes associés seront étudiés dans le troisième chapitre.

La lecture de quelques articles et livres permettent d'appréhender le problème de filtrage dans son ensemble, et de bien comprendre les algorithmes existants : l'ouvrage [Anderson 79] fait état de référence pour la description du problème de filtrage linéaire par filtre de Kalman et ses extensions au cas non linéaire. Le livre [Doucet 01a] est une compilation d'articles qui traitent à la fois des travaux théoriques et applicatifs sur les méthodes séquentielles de Monte Carlo. Celles-ci sont également très bien décrites dans les articles [Liu 98a] et [Doucet 00b], et l'article [Andrieu 03] permet de situer ces algorithmes par rapport aux méthodes de Monte Carlo par chaînes de Markov. En ce qui concerne le suivi en environnement bruité et le problème d'association de mesures, l'ouvrage [BarShalom 88] est une bonne introduction. Enfin, le tutoriel [Arulampalam 02] ainsi que la thèse de doctorat (en français) [Hue 03] apportent un bon état de l'art du problème et des algorithmes de filtrage.

Chapitre 1

Méthodes non aléatoires

Dans ce chapitre sont décrites les méthodes non aléatoires¹ existantes pour la résolution du problème de filtrage. Dans un premier temps, le filtre de Kalman est détaillé. Ensuite, quelques extensions de ce filtre sont présentées (filtre de Kalman étendu, filtre de Kalman unscented, filtre par mélange de lois gaussiennes). Enfin, les méthodes numériques par maillage d'état sont décrites.

1.1 Modèle linéaire avec bruits additifs gaussiens : filtre de Kalman

Le **filtre de Kalman** [Kalman 60] et son extension au temps continu appelé **filtre de Kalman-Bucy** [Kalman 61] permettent de résoudre de façon optimale le problème du filtrage linéaire, quand les bruits du système sont additifs et gaussiens. Depuis 1960, le filtre de Kalman a été intensivement utilisé dans des applications diverses, et a donné lieu à de nombreuses recherches. Une étude détaillée du filtrage linéaire est décrite dans [Gelb 74] et [Anderson 79]. Enfin, [Rotella 03] propose une introduction claire en français.

On considère un système linéaire à bruits additifs :

$$\begin{cases} \mathbf{x}_k = F_k \mathbf{x}_{k-1} + \mathbf{b}_k + \mathbf{w}_k \\ \mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{d}_k + \mathbf{v}_k, \end{cases} \quad (1.1)$$

où les hypothèses suivantes sont faites :

- l'état initial $\mathbf{x}_0$ est gaussien, d'espérance $\hat{\mathbf{x}}_0$ et de covariance Σ_0 ;
- les matrices F_k et H_k sont déterministes, appelées respectivement matrices d'état et de mesure. $\mathbf{b}_k$ et $\mathbf{d}_k$ sont des entrées du système connues ;
- les bruits d'état $\mathbf{w}_k$ et de mesure $\mathbf{v}_k$ sont des bruits blancs gaussiens de moyennes nulles et de covariances respectives Q_k et R_k connues. Ils sont supposés mutuellement indépendants et indépendants de la condition initiale.

¹On préférera le terme “non aléatoire” au terme “déterministe” pour éviter les confusions possibles avec les méthodes entièrement déterministes où aucune modélisation probabiliste du problème n'est utilisée.

Le système peut alors s'écrire de la façon suivante :

$$\begin{cases} \mathbf{x}_0 \sim \mathcal{N}(\mathbf{x}_0; \hat{\mathbf{x}}_0, \Sigma_0) \\ \mathbf{x}_k | \mathbf{x}_{k-1} \sim \mathcal{N}(\mathbf{x}_k; F_k \mathbf{x}_{k-1} + \mathbf{b}_k, Q_k) \\ \mathbf{z}_k | \mathbf{x}_k \sim \mathcal{N}(\mathbf{z}_k; H_k \mathbf{x}_k + \mathbf{d}_k, R_k). \end{cases} \quad (1.2)$$

La notation $\mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \Sigma)$ désigne la loi normale de variable $\mathbf{x}$, d'espérance $\hat{\mathbf{x}}$ et de covariance Σ . Par linéarité des équations d'état et de mesure, le processus $\{\mathbf{x}_k, \mathbf{z}_k\}$ est gaussien et la loi de filtrage recherchée $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ est gaussienne. Le problème est alors de dimension finie. La loi de filtrage est entièrement décrite par son espérance, notée $\hat{\mathbf{x}}_{k|k} = \mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}]$ et sa covariance $\Sigma_{k|k} = \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^t]$. Ces deux premiers moments sont calculés récursivement par le filtre de Kalman en deux étapes :

- À partir de la distribution de filtrage au temps $k-1$, l'étape de *prédiction* donne les équations pour le calcul de la distribution prédite.

$$\mathbf{x}_k | \mathbf{z}_{1:k-1} \sim \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \Sigma_{k|k-1}) \quad (1.3)$$

$$\hat{\mathbf{x}}_{k|k-1} = F_k \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}_k \quad (1.4)$$

$$\Sigma_{k|k-1} = F_k \Sigma_{k-1|k-1} F_k^t + Q_k \quad (1.5)$$

- L'étape de *correction* permet le calcul de la distribution de filtrage au temps k .

$$\mathbf{x}_k | \mathbf{z}_{1:k} \sim \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \Sigma_{k|k}) \quad (1.6)$$

$$K_k = \Sigma_{k|k-1} H_k^t (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1} \quad (1.7)$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k [\mathbf{z}_k - (H_k \hat{\mathbf{x}}_{k|k-1} + \mathbf{d}_k)] \quad (1.8)$$

$$\Sigma_{k|k} = (Id - K_k H_k) \Sigma_{k|k-1} \quad (1.9)$$

La matrice K_k est appelée le gain du filtre de Kalman. Durant la phase de correction, cette matrice est multipliée par l'innovation $\tilde{\mathbf{z}}_k = \mathbf{z}_k - (H_k \hat{\mathbf{x}}_{k|k-1} + \mathbf{d}_k)$ (qui correspond à la différence entre la mesure observée et la mesure prédite). Ce terme permet de corriger l'estimée prédite. Cette étape peut s'interpréter facilement : une grande confiance dans les précédentes estimations ($\Sigma_{k|k-1}$ "faible") et un doute dans les mesures (R_k "élevé") impliquent un gain faible. La prédiction est alors faiblement corrigée. Un doute sur les précédentes estimations et une confiance en la nouvelle mesure entraînent un gain fort. La mesure prend donc de l'importance dans la valeur finale de l'estimée.

On peut remarquer que les matrices de covariance et le gain du filtre ne dépendent pas de la mesure courante. Pour une réduction du temps de calcul, ces matrices pourront être calculées au préalable dans le cas où les matrices du système F_k, H_k, R_k, Q_k sont indépendantes du temps.

Les équations du filtre de Kalman peuvent être obtenues de plusieurs façons. Une première approche consiste à calculer explicitement les deux étapes du filtre optimal (0.5,0.6).

Cela est possible du fait des propriétés des lois gaussiennes. Cette approche n'est plus valable lorsque l'hypothèse gaussienne est relâchée.

Une deuxième approche repose sur la théorie de l'estimation dans un cadre non bayésien. En effet, l'estimateur de variance minimale $\hat{\mathbf{x}}_{k|k} = \mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}]$, qui dans le cas gaussien est équivalent à l'estimateur linéaire non biaisé de variance minimale $E^*[\mathbf{x}_k | \mathbf{z}_{1:k}]$ peut être utilisé (c.f. annexe B pour un rappel sur les estimateurs de variance minimale, et leurs propriétés utilisées pour le développement du filtre de Kalman). On se ramène alors à un problème d'estimation au sens des moindres carrés. Sous la contrainte que l'estimateur est une combinaison linéaire de l'innovation et de l'estimateur prédit, on recherche le gain qui minimise la covariance de l'erreur d'estimation [Gelb 74]. Les équations obtenues sont celles décrites précédemment.

Enfin, toujours en se plaçant dans un cadre d'estimation au sens du minimum de variance, une troisième approche utilise les propriétés d'indépendance de la séquence d'innovation. Une projection du processus état sur l'espace engendré par les innovations permet de calculer l'estimateur $\hat{\mathbf{x}}_{k|k}$ [Anderson 79].

Dans le cas gaussien, ces trois approches sont équivalentes. L'algorithme 1 décrit en résumé le filtre de Kalman pour le problème de filtrage linéaire en temps discret.

Algorithme 1 Filtre de Kalman pour la résolution du système (1.2)

- **initialisation** : $\hat{\mathbf{x}}_{0|0} = \hat{\mathbf{x}}_0$, $\Sigma_{0|0} = \Sigma_0$

pour $k = 1, 2, \dots$

- **prédiction** :

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= F_k \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}_k \\ \Sigma_{k|k-1} &= F_k \Sigma_{k-1|k-1} F_k^t + Q_k\end{aligned}$$

- **correction** :

$$\begin{aligned}K_k &= \Sigma_{k|k-1} H_k^t (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + K_k [\mathbf{z}_k - (H_k \hat{\mathbf{x}}_{k|k-1} + \mathbf{d}_k)] \\ \Sigma_{k|k} &= (Id - K_k H_k) \Sigma_{k|k-1}\end{aligned}$$

Relâchement de l'hypothèse gaussienne

Dans de nombreuses applications de vision par ordinateur, les séquences d'observations, qui résultent de mesures sur les images sont rarement gaussiennes. Il est alors intéressant de considérer une formulation du filtre de Kalman adaptée à des mesures non gaussiennes perturbées par des données aberrantes. On parle de **filtre de Kalman robuste** [Tarel 02], où l'estimation au sens des moindres carrés est remplacée par un algorithme de moindres carrés pondérés itérés.

Plus généralement, si les processus de bruit ne sont plus supposés gaussiens, mais que

l'on dispose de leur deux premiers moments, les équations du filtre de Kalman classique peuvent être utilisées. Elles permettent le calcul du meilleur estimateur linéaire au sens du minimum de la variance d'estimation. La connaissance de la loi de filtrage obtenue est alors incomplète et réduite aux deux premiers moments d'une approximation gaussienne. Ces deux moments sont calculés de façon exacte [Anderson 79]

Dépendance entre l'état et la mesure

Dans le cas où les matrices et les vecteurs du système F_{k+1} , $\mathbf{b}_{k+1}$, Q_{k+1} , H_k , $\mathbf{d}_k$ et R_k dépendent des observations passées $\mathbf{z}_{k-1}$, les équations classiques sont valides. En effet, bien que le couple $\{\mathbf{x}_k, \mathbf{z}_k\}$ ne soit plus un processus gaussien, la loi de filtrage $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ reste gaussienne. Ses deux premiers moments sont toujours fournis par le filtre de Kalman décrit par l'algorithme 1. Par contre, il n'est plus possible de calculer les covariances et le gain à l'avance.

Une dépendance entre l'état et la mesure peut aussi apparaître à travers les bruits du modèle, lorsque ceux-ci sont corrélés. Dans ce cas, une façon rapide d'établir les équations du filtre de Kalman correspondantes consiste à se ramener au cas d'un modèle à bruits décorrélés par la construction d'un modèle équivalent. Un nouveau vecteur de bruit dynamique est défini. Le nouveau filtre est dérivé à partir du filtre classique. On retrouve les équations de prédiction et de mise à jour, qui ont été modifiées par ajout d'un terme correcteur. Comme dans le cas précédent, les covariances et les gains doivent être calculés en-ligne.

1.2 Extensions au cas non-linéaire

Lorsque le système à résoudre n'est pas linéaire, mais que les hypothèses sur les bruits sont conservés (bruits blancs additifs, gaussiens, indépendants entre eux et indépendants de la condition initiale), il s'écrit :

$$\begin{cases} \mathbf{x}_k = f_k(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ \mathbf{z}_k = h_k(\mathbf{x}_k) + \mathbf{v}_k, \end{cases} \quad (1.10)$$

où les fonctions f_k et h_k sont des fonctions non linéaires, supposées dérivables, à valeurs dans $\mathbb{R}^{n_x}$ et $\mathbb{R}^{n_z}$ respectivement. On note Q_k la covariance du bruit d'état et R_k celle du bruit de mesure. Ce modèle peut être étendu au système suivant, où l'état et la mesure dépendent non linéairement des bruits du modèle :

$$\begin{cases} \mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{w}_k) \\ \mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{v}_k). \end{cases} \quad (1.11)$$

Enfin, il peut être généralisé à un système non linéaire non gaussien. Il s'écrit alors de façon similaire au système (1.11), mais les hypothèses sur le caractère gaussien des bruits sont relâchées.

Dans ces trois cas, la solution optimale ne peut plus s'écrire analytiquement. Le filtre de Kalman n'est plus valide et la construction d'un filtre récursif requiert des approximations. En particulier, la loi de filtrage $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ peut être approchée par une gaussienne

dont les moments sont obtenus par linéarisation (**filtre de Kalman étendu**) ou qui est décrite par un ensemble de points (**filtre de Kalman unscented**). On trouvera dans [Lefebvre 04] une description et une comparaison des filtres qui supposent une forme gaussienne de la loi de filtrage. Dans le cas de multi-modalité de la loi de filtrage, l'extension des filtres précédents à l'estimation par un mélange de densités gaussiennes a conduit aux développements d'algorithmes de **filtres par mélange de lois**.

1.2.1 Filtre de Kalman étendu

On considère le système (1.10). Il peut être approché par un modèle linéaire gaussien par linéarisation des équations d'état et de mesure. L'extension la plus simple est obtenue par développement au premier ordre (autour de l'estimée précédente pour l'équation d'état et de la prédiction pour l'équation de mesure) :

$$\begin{aligned} f_k(\mathbf{x}_{k-1}) &\simeq f_k(\hat{\mathbf{x}}_{k-1|k-1}) + F_k (\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}) \quad \text{avec} \quad F_k = \nabla f_k(\hat{\mathbf{x}}_{k-1|k-1}) \\ h_k(\mathbf{x}_k) &\simeq h_k(\hat{\mathbf{x}}_{k|k-1}) + H_k (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \quad \text{avec} \quad H_k = \nabla h_k(\hat{\mathbf{x}}_{k|k-1}) \end{aligned} \quad (1.12)$$

Le système linéaire déduit est le suivant :

$$\begin{cases} \mathbf{x}_k = F_k \mathbf{x}_{k-1} + \mathbf{b}_k + \mathbf{w}_k, & \text{avec} \quad \mathbf{b}_k = f_k(\hat{\mathbf{x}}_{k-1|k-1}) + F_k \hat{\mathbf{x}}_{k-1|k-1} \\ \mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{d}_k + \mathbf{v}_k, & \text{avec} \quad \mathbf{d}_k = h_k(\hat{\mathbf{x}}_{k|k-1}) + H_k \hat{\mathbf{x}}_{k|k-1} \end{cases} \quad (1.13)$$

L'utilisation du filtre de Kalman pour ce système fournit les équations du **filtre de Kalman étendu** au premier ordre (résumé dans l'algorithme 2). Dans le cas où les bruits ne sont plus additifs mais liés de façon non linéaire à l'état et à la mesure (système (1.11)), on procède de la même manière. Les fonctions des bruits sont linéarisées autour de 0 afin d'obtenir un système qui peut être résolu par un filtre de Kalman. Des filtres équivalents peuvent être déduits de linéarisation à des ordres supérieurs, mais ils ne sont pas utilisés en pratique.

Le filtre de Kalman étendu repose donc sur une approximation de la loi de filtrage $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ par une loi gaussienne. Il permet de calculer les deux premiers moments de cette approximation $\hat{\mathbf{x}}_{k|k}$ et $\Sigma_{k|k}$ par linéarisation. Bien qu'intensivement utilisé en ingénierie pour traiter les problèmes non linéaires, aucun travaux théoriques n'ont été réalisés justifiant son utilisation dans le cas général. Néanmoins, des études valident son emploi pour la résolution de systèmes avec un bruit d'observation faible [Picard 91] [Milheiro 94]. Dans ce cas, l'erreur commise par le filtre est négligeable par rapport à la précision du filtre optimal. Mais dans le cas général, l'erreur commise n'est pas connue. Il est possible de mesurer l'adéquation entre le système non linéaire initial et le système linéaire approché en évaluant les quantités $\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\|^2$ et $\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\|^2$, dont une approximation peut être obtenue grâce aux traces des matrices de covariance.

Si les non-linéarités sont trop importantes, le filtre de Kalman étendu peut diverger. Cela est d'autant plus vrai que les équations du filtre ne prennent pas en compte l'erreur de linéarisation. Pour remédier à ce problème, quelques heuristiques ont été proposées. D'une part, il est possible d'augmenter la variance des bruits Q_k et R_k afin de prendre en compte cette erreur. D'autre part, une augmentation artificielle de la covariance $\Sigma_{k|k}$ permet de limiter la confiance dans les estimées passées. Ces méthodes nécessitent un ajustement précis et difficile des paramètres.

Algorithme 2 Filtre de Kalman étendu pour la résolution du système (1.10)

- **initialisation** : $\hat{\mathbf{x}}_{0|0} = \hat{\mathbf{x}}_0$, $\Sigma_{0|0} = \Sigma_0$

pour $k = 1, 2, \dots$

- **prédiction** :

$$\begin{aligned} F_k &= \nabla f_k(\hat{\mathbf{x}}_{k-1|k-1}) \\ \hat{\mathbf{x}}_{k|k-1} &= f_k(\hat{\mathbf{x}}_{k-1|k-1}) \\ \Sigma_{k|k-1} &= F_k \Sigma_{k-1|k-1} F_k^t + Q_k \end{aligned}$$

- **correction** :

$$\begin{aligned} H_k &= \nabla h_k(\hat{\mathbf{x}}_{k|k-1}) \\ K_k &= \Sigma_{k|k-1} H_k^t (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + K_k [\mathbf{z}_k - h_k(\hat{\mathbf{x}}_{k|k-1})] \\ \Sigma_{k|k} &= (Id - K_k H_k) \Sigma_{k|k-1} \end{aligned}$$

Pour améliorer la linéarisation du modèle de mesure, le **filtre de Kalman étendu itéré** [Jazwinski 70] ajoute un processus itératif dans l'étape de correction. Alors que dans le filtre de Kalman étendu, la fonction d'observation h_k est linéarisée autour de la prédiction (cf. (1.12)), la version itérée utilise un développement autour de la nouvelle estimée :

$$h_k(\mathbf{x}_k) \simeq h_k(\hat{\mathbf{x}}_{k|k}) + H_k (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) \quad \text{avec} \quad H_k = \nabla h_k(\hat{\mathbf{x}}_{k|k}). \quad (1.14)$$

Cela est obtenu par étapes successives : le filtre linéarise autour de la prédiction, puis calcule une estimée, autour de laquelle un développement de Taylor est écrit. Une nouvelle estimée est obtenue. Le processus est itéré jusqu'à ce que la différence entre deux estimées successives soit inférieure à un seuil. Cette étape améliore sensiblement les performances du filtre de Kalman étendu.

Enfin, pour une étude plus approfondie, [Rong Li 01] décrit en détail différentes méthodes de linéarisation qui améliorent les performances du filtre et éventuellement le simplifient. Certaines de ces méthodes incluent une réduction de l'erreur de linéarisation.

1.2.2 Filtre de Kalman unscented

Toujours dans le cadre de l'estimation de systèmes non linéaires, le **filtre de Kalman unscented** [Julier 97] [Wan 00] propose une alternative au filtre de Kalman étendu. Le principe d'approximation de la loi de filtrage $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ par une loi gaussienne est conservé. Cependant, il n'y a pas de linéarisation des équations d'état et de mesure. Cela permet d'éviter les deux principaux inconvénients du filtre de Kalman étendu, à savoir les erreurs de linéarisation et les difficultés d'implémentation des matrices jacobiniennes.

A chaque pas de temps, l'approximation gaussienne est représentée par un ensemble de points choisis $\{\mathbf{x}_{k|k}^{(i)}\}_{i=1\dots N}$, auxquels sont associés des poids $\{w_{k|k}^{(i)}\}_{i=1\dots N}$. Ce nuage

pondéré capte la moyenne $\hat{\mathbf{x}}_{k|k}$ et la covariance $\Sigma_{k|k}$ de l'approximation. Ces points sont choisis de façon déterministe. Ils sont propagés à l'aide des modèles non linéaires, et les paramètres de l'approximation gaussienne sont estimés à nouveau. Le choix des points représentant la loi de filtrage est décrit dans [Julier 97]. Ils sont définis par la transformation "unscented" qui nécessite la connaissance des deux premiers moments de l'approximation. L'algorithme 3 décrit le filtre de Kalman unscented pour la résolution du système (1.10). Tout comme le filtre de Kalman étendu, cet algorithme est facilement adaptable à la résolution du système (1.11) où les états et mesures sont des fonctions non linéaires des bruits. Dans ce cas, il est appliqué à l'état augmenté du vecteur de bruit.

Algorithme 3 Filtre de Kalman unscented pour la résolution du système (1.10)

- **initialisation** : $\hat{\mathbf{x}}_{0|0} = \hat{\mathbf{x}}_0$, $\Sigma_{0|0} = \Sigma_0$

pour $k = 1, 2, \dots$

- Calcul du nuage pondéré de points $\{\mathbf{x}_{k-1|k-1}^{(i)}, w_{k-1|k-1}^{(i)}\}_{i=1\dots N}$ à partir de l'estimation précédente $\hat{\mathbf{x}}_{k-1|k-1}$, $\Sigma_{k-1|k-1}$.

- **prédiction** :

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1}^{(i)} &= f_k(\hat{\mathbf{x}}_{k-1|k-1}^{(i)}) \\ \hat{\mathbf{x}}_{k|k-1} &= \sum_{i=1}^N w_{k-1|k-1}^{(i)} \mathbf{x}_{k|k-1}^{(i)} \\ \Sigma_{k|k-1} &= \sum_{i=1}^N w_{k-1|k-1}^{(i)} [(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})^t] + Q_k \\ \hat{\mathbf{z}}_{k|k-1}^{(i)} &= h_k(\hat{\mathbf{x}}_{k|k-1}^{(i)}) \\ \hat{\mathbf{z}}_{k|k-1} &= \sum_{i=1}^N w_{k-1|k-1}^{(i)} \mathbf{z}_{k|k-1}^{(i)}\end{aligned}$$

- **correction** :

$$\begin{aligned}\Sigma_{\mathbf{zz}} &= \sum_{i=1}^N w_{k-1|k-1}^{(i)} [(\mathbf{z}_{k|k-1}^{(i)} - \hat{\mathbf{z}}_{k|k-1})(\mathbf{z}_{k|k-1}^{(i)} - \hat{\mathbf{z}}_{k|k-1})^t] + R_k \\ \Sigma_{\mathbf{zx}} &= \sum_{i=1}^N w_{k-1|k-1}^{(i)} [(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_{k|k-1}^{(i)} - \hat{\mathbf{z}}_{k|k-1})^t] \\ K_k &= \Sigma_{\mathbf{zx}} \Sigma_{\mathbf{zz}}^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + K_k[\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}] \\ \Sigma_{k|k} &= \Sigma_{k|k-1} - K_k \Sigma_{\mathbf{zz}} K_k^t\end{aligned}$$

Alors que le filtre de Kalman étendu capte la moyenne et la covariance de l'approximation gaussienne au premier ordre (au sens d'un développement de Taylor), il peut être montré que la version unscented donne une estimation précise au second ordre, quelles que soient les non linéarités du système. Dans le cas de bruits additifs gaussiens, l'approximation est précise jusqu'au troisième ordre. Ce filtre améliore donc les performances, du fait de la meilleure estimation des moments. Quelques problèmes numériques subsistent. En

particulier des instabilités sont observées dans la phase de choix des points. Celle-ci repose sur une décomposition de la matrice de covariance. Une amélioration du filtre unscented permettant de remédier à ces problèmes est proposée dans [Merwe 01].

Plus généralement, le filtre de Kalman unscented fait partie d'une famille commune de filtres appelée **sigma-points Kalman filter** [Merwe 03] ou **linear regression Kalman filter** [Lefebvre 04]. En font également partie les filtres **central difference Kalman filter** et **divided difference Kalman filter**. Tous ces filtres utilisent une approche d'échantillonnage déterministe pour représenter une approximation gaussienne de la loi de filtrage. Les différences résident essentiellement dans le choix du nuage de points.

1.2.3 Filtre par mélange de lois gaussiennes

Les algorithmes de Kalman étendu et unscented se fondent sur l'hypothèse forte que la loi de filtrage peut être bien approchée par une loi normale. Dans de nombreux cas, notamment lorsque cette loi est multi-modale, ces algorithmes sont mis en défaut. Il est alors naturel de considérer des approximations par mélange de lois.

On se place dans le cas général où le système à résoudre est décrit par les densités $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ et $p(\mathbf{z}_k|\mathbf{x}_k)$ sur lesquelles aucune hypothèse n'est faite. Le **filtre par mélange de lois gaussiennes** ("Gaussian sum filter") approche la loi de filtrage $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ par un mélange de densités gaussiennes [Sorenson 71][Alspach 72]. À chacune des densités est attribué un filtre de Kalman étendu pour le calcul récursif de sa moyenne et covariance. L'estimateur utilisé correspond donc à un banc de filtres de Kalman étendus dont le résultat sera une somme pondérée des sorties des filtres. Bien que donnant des résultats de meilleure qualité que les filtres reposant sur l'approximation par une seule loi, le filtre par mélange souffre de deux problèmes majeurs :

- *Augmentation exponentielle du nombre de composantes*

Le premier inconvénient de ce filtre est la croissance exponentielle du coût de calcul, due à l'augmentation du nombre de composantes du mélange au cours du temps. Cela est d'autant plus vrai lorsque les bruits ne sont pas gaussiens et que leur loi est elle-même approchée par un mélange de lois normales. En pratique, le nombre de modes doit être réduit, soit par élagage, soit par fusion. Cette étape est très difficile à automatiser.

- *Manque de précision du filtre de Kalman étendu*

Le deuxième inconvénient est lié à l'utilisation du filtre de Kalman étendu pour suivre l'évolution des composantes gaussiennes. En particulier, l'accumulation des erreurs de linéarisation se répercute sur la qualité de l'estimation. Pour que son utilisation soit valide, les covariances des lois gaussiennes doivent être petites. Dans le cas contraire, une ré-initialisation du banc de filtres est nécessaire.

1.3 Méthodes numériques par maillage de l'espace d'état

Dans de nombreux cas, faire une approximation gaussienne, à une ou plusieurs composantes, de la loi de probabilité $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ est une mauvaise hypothèse. Une alternative consiste à décrire la loi recherchée par un nuage pondéré de points.

Les **méthodes numériques par maillage de l'espace d'état** sont fondées sur ce principe [Kitagawa 87] [Sorenson 88] [Kramer 88]. Elles sont communément appelées méthodes HMM pour *hidden Markov model*. Un maillage déterministe de l'espace d'état est effectué, afin d'évaluer la loi de filtrage en ces points. Ces points peuvent servir de base à une approximation continue de la loi par des fonctions splines ou linéaires par morceaux. Alternativement, une solution plus simple approche la loi par une somme pondérée de mesures de Dirac en ces points. En notant $\{\mathbf{x}_k^{(i)}\}_{i=1\dots N}$ les points de la grille au temps k , $\{w_{k|k}^{(i)}\}_{i=1\dots N}$ les poids associés et $\delta_{\mathbf{x}}(\cdot)$ la mesure de Dirac prise au point $\mathbf{x}$, la relation suivante est obtenue :

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \simeq \sum_i w_{k|k}^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k) \quad (1.15)$$

Cette approximation permet de simplifier le calcul des intégrales des équations du filtre optimal (0.5, 0.6). Les étapes de prédiction et correction sont décrites dans l'algorithme 4.

Algorithme 4 Méthode numérique par maillage de l'espace d'état (approximation par somme de mesures de Dirac)

- **initialisation** : fixer la grille $\{\mathbf{x}_0^{(i)}\}_{i=1\dots N}$, $w_{0|0}^{(i)} = 1/N \quad \forall i$

pour $k = 1, 2, \dots$

- **prédiction** : fixer la grille $\{\mathbf{x}_k^{(i)}\}_{i=1\dots N}$,

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) \simeq \sum_{i=1}^N w_{k|k-1}^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k)$$

$$w_{k|k-1}^{(i)} = \sum_{j=1}^N w_{k-1|k-1}^{(j)} p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(j)})$$

- **correction** :

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \simeq \sum_{i=1}^N w_{k|k}^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k)$$

$$w_{k|k}^{(i)} = \frac{w_{k|k-1}^{(i)} p(\mathbf{z}_k|\mathbf{x}_k^{(i)})}{\sum_{j=1}^N w_{k|k-1}^{(j)} p(\mathbf{z}_k|\mathbf{x}_k^{(j)})}$$

Lorsque l'espace d'état est discret et fini, la solution donnée par cette méthode est exacte si les points de la grille sont choisis tels qu'ils correspondent à l'ensemble des états

discrets. Dans le cas continu, l'espace d'état est découpé en cellules et les points de la grille sont choisis au centre de ces cellules. En pratique, la grille doit être suffisamment dense pour que l'approximation soit correcte. Elle peut être déterminée par une technique de maillage fixe, ou par une technique de maillage adaptatif [Cai 95], plus souple, permettant de diminuer le pas de maillage dans les zones de forte probabilité. Dans tous les cas, le choix des points rend l'algorithme difficile à mettre en œuvre, d'autant plus que le nombre de points nécessaires augmente exponentiellement avec la dimension de l'espace d'état. Enfin, les potentielles difficultés de calcul en chaque point rendent ces méthodes très gourmandes en temps de calcul. Ces algorithmes deviennent ainsi inutilisables dès lors que la dimension de l'espace d'état est trop grande (en pratique 3 ou 4).

1.4 Synthèse

Dans ce chapitre, nous nous sommes intéressés à la description d'algorithmes non aléatoires pour la résolution du problème de filtrage. Ce problème est modélisé par une chaîne de Markov cachée, décrite par un système de deux équations (équation dynamique et équation de mesure).

Lorsque le système est linéaire gaussien, le filtre de Kalman fournit une solution au problème de filtrage qui est récursive, optimale et très simple à mettre en œuvre. Dès que le modèle à résoudre présente des non linéarités, l'utilisation de ces équations n'est plus valide. L'approche la plus simple consiste à linéariser le système par des développements de Taylor. Les équations de Kalman peuvent alors être appliquées au modèle linéaire gaussien déduit. Le filtre obtenu est appelé filtre de Kalman étendu. Il fournit une approximation gaussienne de la densité recherchée. Son principal inconvénient résulte des accumulations d'erreurs de linéarisation, qui ne sont jamais prises en compte dans le filtre et peuvent amener à sa divergence. Afin d'éviter ce problème, le filtre de Kalman unscented propose de décrire l'approximation gaussienne par un nuage de points pondérés, mis à jour récursivement en utilisant les modèles non linéaires. Le calcul des deux premiers moments de l'approximation est fait à partir de ce nuage et ne nécessite plus de linéarisation.

Bien que donnant de meilleurs résultats que le filtre de Kalman étendu, ses performances sont mauvaises dès que la densité de filtrage est multimodale. En effet, les versions étendues et unscented du filtre de Kalman supposent une approximation gaussienne de la densité recherchée. Cette hypothèse est rapidement mise en défaut dans le cas général. Cela a amené à la dérivation de nouveaux filtres, qui approchent la densité de filtrage par un mélange de lois gaussiennes. Chaque loi est mise à jour récursivement par un filtre de Kalman étendu. Ces algorithmes permettent théoriquement la description de densités hautement multimodales. En pratique, leur utilisation est difficile. Le nombre de composantes du mélange doit être contrôlé, afin d'éviter une augmentation exponentielle du coût de calcul.

Une dernière classe d'algorithmes discrétise l'espace d'état pour évaluer la loi de filtrage en des points d'une grille. La loi recherchée est décrite comme une somme pondérée de mesures de Dirac prises en ces points. Ces méthodes numériques sont exactes lorsque l'espace d'état est discret, constitué d'un nombre fini (et faible) d'états, et que les points de la grille correspondent à ces états. Les principaux inconvénients de ces méthodes sont l'augmentation exponentielle du coût de calcul avec la taille de l'espace d'état, la difficulté de mise en œuvre, et surtout le choix de la grille. Ce dernier point est essentiel pour

obtenir une approximation précise.

Bien que performants sous certaines hypothèses, ces algorithmes souffrent de désavantages importants. D'une part, le filtre de Kalman et ses versions étendues et unscented sont valides sur des modèles spécifiques, avec des hypothèses fortes sur les densités mises en jeu, rarement rencontrées en pratique. D'autre part, bien que les filtres par mélange de lois gaussiennes et les méthodes numériques par maillage soient valides sur les modèles généraux, ils sont difficilement utilisables en pratique à cause d'un coût de calcul élevé et de leur complexité de mise en œuvre. Récemment, l'essor des méthodes de Monte Carlo a permis la dérivation de nouveaux filtres appelés filtres particuliers (ou méthodes séquentielles de Monte Carlo). Ces filtres apportent une nouvelle solution au problème de filtrage non linéaire. Leur performance dépassent celles des algorithmes non aléatoires dans le cas général. Ces méthodes sont décrites dans le chapitre suivant.

Chapitre 2

Méthodes séquentielles de Monte Carlo

Lorsque le modèle décrivant le système dynamique à estimer est hautement non linéaire et/ou non gaussien, les méthodes non aléatoires (filtre de Kalman et extensions, méthodes numériques par maillage) échouent à donner une approximation précise de la loi de filtrage, avec un temps de calcul acceptable. L'utilisation de simulations par Monte Carlo pour la construction de cette loi est une alternative dont les performances se sont révélées encourageantes. Les méthodes séquentielles de Monte Carlo (MSMC) sont fondées sur ce principe.

Le premier document publié relatif aux méthodes de Monte Carlo est attribué à Metropolis et Ulam [Metropolis 49]. Dans les années 50-70, se sont développées indépendamment les méthodes de Monte Carlo par chaînes de Markov dans la communauté de la physique statistique [Metropolis 53] [Hastings 70], et les méthodes d'intégration séquentielles par simulation dans la communauté de l'automatique [Handschin 70]. Au début des années 90, les méthodes de Monte Carlo ont eu un impact important dans la communauté du traitement de signal, en grande partie grâce à l'augmentation des capacités de calcul. Cet essor a donné lieu à la construction d'algorithmes récurrents pour la résolution du problème de filtrage non linéaire. Ces algorithmes ont été appelés filtre bootstrap (**bootstrap filter**) [Gordon 93], filtre de Monte Carlo (**Monte Carlo filter**) [Kitagawa 96], algorithme de condensation (**condensation algorithm** pour conditional density propagation) [Isard 98a], filtre particulaire avec interaction (**interacting particle filter**) [Del Moral 96], ou encore **survival of the fittest** [Kanazawa 95].

Depuis, les méthodes séquentielles de Monte Carlo ont été largement exploitées dans de nombreuses applications. Citons par exemple le traitement du signal (pistage par mesure d'angles, navigation, etc.) [Gustafsson 02] [Hue 02] [Ristic 04a] [Bréhard 04], le traitement de la parole [Vermaak 02a], les problèmes de classification [Freitas 01], le suivi d'entités en vision par ordinateur [Isard 98a] [Pérez 04], le recalage en imagerie médicale [Ma 04], la robotique mobile [Kwok 04], le diagnostic médical [Berzuini 97], la modélisation de données financières [Pitt 01], la déconvolution de signaux digitaux [Liu 95]. Notons également que ces algorithmes sont utilisés dans les domaines de la bio-technologie (génétique des populations moléculaires) [Chen 03], de la physique (physique quantique) et de la chimie (science des polymères) [Iba 01].

Pour résoudre le problème de filtrage non linéaire, ces filtres (que nous appellerons par la suite MSMC ou filtres particuliers) approchent la densité *a posteriori* par un ensemble de points de l'espace d'état appelés *particules*. À chaque particule est associé un poids, et la *distribution objectif* est approchée par une somme pondérée discrète, dont le support correspond aux particules. La probabilité de chaque particule est proportionnelle à son poids. Le nuage est mis à jour récursivement à chaque pas de temps : chaque particule évolue en utilisant le principe de l'**échantillonnage pondéré séquentiel**, et certaines sont sélectionnées par une étape de **rééchantillonnage**. La connaissance des particules et de leur poids permet d'accéder aux moments de la loi *a posteriori*.

Remarquons que des similarités entre les filtres particuliers et les méthodes numériques par maillage existent. Ces deux méthodes utilisent un ensemble de points (ou particules) de l'espace d'état, qui sert de support à une approximation discrète de la loi de filtrage. Cependant, ces points sont choisis de façon arbitraire dans le cas des méthodes par grille, alors qu'ils sont générés aléatoirement selon les équations du modèle dans le cas des MSMC. Un maillage probabiliste permet de mieux guider les particules dans les zones d'intérêt de l'espace d'état.

Après une présentation générale du principe de Monte Carlo et de deux méthodes d'échantillonnage par simulation, nous verrons comment le principe d'échantillonnage pondéré s'applique au cadre bayésien pour le filtrage non linéaire. Cela nous amène à la définition d'un algorithme générique de filtrage particulier. Ensuite, le choix de la fonction utilisée pour l'évolution des particules et le choix d'une méthode de sélection seront discutés. Enfin, la méthode de réduction de variance par rao-blackwellisation sera présentée.

Le contenu de ce chapitre repose en grande partie sur les articles [Andrieu 03] [Liu 98a] [Doucet 00b], ainsi que la thèse [Hue 03]. Pour une étude plus approfondie, les états de l'art des thèses [Fearnhead 98] [Bergman 99] [Clapp 00] [Schön 03a] sont intéressants. Citons également le livre [Doucet 01a] qui présente une compilation d'articles traitant à la fois de problèmes théoriques et applicatifs des MSMC. Des travaux récents sont aussi disponibles dans un numéro spécial de la revue *Proceedings of the IEEE* édité en mars 2004 [Haykin 04]. Enfin, pour une étude des résultats de convergence des MSMC que nous détaillerons pas ici, nous renvoyons aux travaux de Del Moral [Del Moral 00] et Crisan [Crisan 01] [Crisan 02].

2.1 Méthodes d'échantillonnage par Monte Carlo

Les paragraphes suivants présentent des notions importantes pour une bonne compréhension des méthodes séquentielles de Monte Carlo. Le principe de Monte Carlo et l'estimation d'intégrales par l'**algorithme d'acceptation rejet** et la méthode d'**échantillonnage pondéré** y sont décrits.

2.1.1 Le principe de Monte Carlo

Le principe de Monte Carlo consiste à utiliser un ensemble d'échantillons discrets obtenus par simulation pour approcher une distribution de probabilité $p(\mathbf{x})$, dont on ne connaît pas l'expression analytique. À partir de cet ensemble, des intégrales de grandes dimensions pourront être calculées. Soit $\{\mathbf{x}^{(i)}\}_{i=1\dots N}$ un ensemble de réalisations indépendantes et

identiquement distribuées selon la densité $p(\mathbf{x})$ définie sur un espace $\mathcal{X}$. Cette densité peut être approchée par la loi discrète $p_N(\mathbf{x})$:

$$p_N(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{x}^{(i)}}(\mathbf{x}), \quad (2.1)$$

où $\delta_{\mathbf{x}^{(i)}}(\mathbf{x})$ désigne la mesure de Dirac prise en la valeur $\mathbf{x}^{(i)}$. À partir de cette densité empirique, il est possible d'approcher des intégrales $I(\phi(\mathbf{x})) = \mathbb{E}_{p(\mathbf{x})}[\phi(\mathbf{x})]$ par des sommes $I_N(\phi(\mathbf{x}))$, telles que :

$$I_N(\phi(\mathbf{x})) = \frac{1}{N} \sum_{i=1}^N \phi(\mathbf{x}^{(i)}) \xrightarrow[N \rightarrow \infty]{p.s.} I(\phi(\mathbf{x})) = \int_{\mathcal{X}} \phi(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}, \quad (2.2)$$

L'estimateur $I_N(\phi(\mathbf{x}))$ est non biaisé et la convergence presque sûre est démontrée par le loi forte des grands nombres. Si de plus la variance de $\phi(\mathbf{x})$ notée σ_ϕ^2 est finie, alors la variance de l'estimateur $I_N(\phi(\mathbf{x}))$ est égale à σ_ϕ^2/N , et le théorème central limite nous assure la convergence en loi de l'erreur d'estimation :

$$\sqrt{N}(I_N(\phi(\mathbf{x})) - I(\phi(\mathbf{x}))) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sigma_\phi^2). \quad (2.3)$$

À partir des échantillons $\{\mathbf{x}^{(i)}\}_{i=1}^N$, il sera ainsi possible d'estimer toute quantité de la forme $\mathbb{E}_{p(\mathbf{x})}[\phi(\mathbf{x})]$, la vitesse de convergence de l'estimateur ne dépendant que de N . Contrairement aux méthodes d'intégration numériques déterministes, la vitesse de convergence ne dépend pas de la taille de l'espace $\mathcal{X}$.

Lorsque la loi objectif $p(\mathbf{x})$ n'est pas connue, et qu'il n'est pas possible d'en tirer des échantillons, des solutions sont apportées par les méthodes de Monte Carlo [Andrieu 03] [Robert 99]. Les algorithmes d'acceptation rejet et d'échantillonnage pondéré sont décrits dans les paragraphes suivants. Nous omettons volontairement les méthodes de Monte Carlo par chaîne de Markov (algorithme de Métropolis-Hastings, échantillonneur de Gibbs, etc.) car elles ne seront pas utilisées dans la suite du document.

2.1.2 Algorithme d'acceptation rejet

Lorsque l'on ne sait pas échantillonner selon la loi $p(\mathbf{x})$, mais qu'il est possible d'évaluer cette loi en tout point à une constante près, un ensemble de réalisations de $p(\mathbf{x})$ peut être obtenu par l'algorithme d'acceptation rejet ("rejection sampling"). Pour cela, il est nécessaire de savoir échantillonner selon une distribution $\pi(\mathbf{x})$, relativement proche de $p(\mathbf{x})$. Cette distribution est appelée *fonction d'importance* ou *loi de proposition*. On suppose également qu'il existe une constante $c < \infty$ telle que $p(\mathbf{x}) \leq c \pi(\mathbf{x})$. La méthode, décrite par l'algorithme 5, consiste à générer un échantillon $\mathbf{x}^{(i)}$ de la loi $\pi(\mathbf{x})$, et de l'accepter avec la probabilité $\frac{p(\mathbf{x}^{(i)})}{c \pi(\mathbf{x}^{(i)})}$. Il peut être démontré que la loi des réalisations obtenues en sortie de l'algorithme est bien la loi de densité $p(\mathbf{x})$ [Robert 99].

Bien que cet algorithme soit très simple, il est parfois difficile à mettre en œuvre. En effet, il n'est pas toujours possible trouver un majorant au rapport $p(\mathbf{x}^{(i)})/\pi(\mathbf{x}^{(i)})$ sur l'ensemble de l'espace $\mathcal{X}$. De plus, la probabilité d'acceptation d'un échantillon étant de $1/c$, la méthode devient impraticable en grande dimension lorsque c est grand.

Algorithme 5 Algorithme d'acceptation rejet

fixer $i = 0$

tant que $i \neq N$ faire

1. générer une réalisation $\mathbf{x}^{(i)}$ de la loi $\pi(\mathbf{x})$
 2. générer une réalisation u de la loi uniforme $\mathcal{U}_{(0,1)}$
 3. si $u < \frac{p(\mathbf{x}^{(i)})}{c \pi(\mathbf{x}^{(i)})}$ alors accepter $\mathbf{x}^{(i)}$ et incrémenter le compteur i de 1 ;
sinon, rejeter l'échantillon.
-

2.1.3 Échantillonnage pondéré

L'échantillonnage pondéré ("importance sampling") est une alternative à l'algorithme d'acceptation rejet. L'hypothèse sur l'existence de la borne c est relâchée. On suppose savoir tirer des échantillons selon une fonction d'importance $\pi(\mathbf{x})$ dont le support inclut le support de $p(\mathbf{x})$ (i.e. $p(\mathbf{x}) > 0 \Rightarrow \pi(\mathbf{x}) > 0$). Sous ces hypothèses, on peut écrire la relation suivante :

$$I(\phi(\mathbf{x})) = \int_{\mathcal{X}} \phi(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{X}} \phi(\mathbf{x}) \frac{p(\mathbf{x})}{\pi(\mathbf{x})} \pi(\mathbf{x}) d\mathbf{x}. \quad (2.4)$$

S'il est possible de tirer N échantillons i.i.d. $\{\mathbf{x}^{(i)}\}_{i=1\dots N}$ de la loi de proposition $\pi(\mathbf{x})$, et d'évaluer les *poids d'importance* associés $\{w^{(i)}\}_{i=1\dots N}$, $w^{(i)} = \frac{p(\mathbf{x}^{(i)})}{\pi(\mathbf{x}^{(i)})}$, alors un estimateur de $I(\phi(\mathbf{x}))$ est donné par :

$$\hat{I}_N(\phi(\mathbf{x})) = \frac{1}{N} \sum_{i=1}^N \phi(\mathbf{x}^{(i)}) w^{(i)}. \quad (2.5)$$

Cet estimateur est non biaisé, et de variance plus faible que l'estimateur obtenu par l'algorithme d'acceptation rejet. Il satisfait la loi forte des grands nombres ainsi que le théorème central limite. Dans le cas où la constante de normalisation de la loi objectif $p(\mathbf{x})$ n'est pas connue, les poids d'importance ne peuvent être évalués qu'à une constante près. Il est donc nécessaire de les normaliser. Cela conduit à estimer l'intégrale par :

$$I_N(\phi(\mathbf{x})) = \sum_{i=1}^N \phi(\mathbf{x}^{(i)}) \tilde{w}^{(i)}, \quad (2.6)$$

où les $\tilde{w}^{(i)}$ sont les *poids d'importance normalisés* :

$$\tilde{w}^{(i)} = \frac{w^{(i)}}{\sum_{j=1}^N w^{(j)}} \quad \text{avec} \quad w^{(i)} = \frac{p(\mathbf{x}^{(i)})}{\pi(\mathbf{x}^{(i)})}. \quad (2.7)$$

Ce nouvel estimateur converge presque sûrement vers $I(\phi(\mathbf{x}))$ quand N tend vers l'infini.

Cette méthode d'intégration peut être interprétée comme une méthode d'échantillonnage décrite par l'algorithme 6. Quand N tend vers l'infini, la densité objectif $p(\mathbf{x})$ est approchée par :

$$p_N(\mathbf{x}) = \sum_{i=1}^N \tilde{w}^{(i)} \delta_{\mathbf{x}^{(i)}}(\mathbf{x}). \quad (2.8)$$

Algorithme 6 Algorithme d'échantillonnage pondéré

1. générer N réalisations $\{\mathbf{x}^{(i)}\}_{i=1\dots N}$ de la loi $\pi(\mathbf{x})$
 2. affecter à chacune de ces réalisations le poids normalisé $\tilde{w}^{(i)} = \frac{w^{(i)}}{\sum_{j=1}^N w^{(j)}}$
- où $w^{(i)} = \frac{p(\mathbf{x}^{(i)})}{\pi(\mathbf{x}^{(i)})}$

L'application de l'échantillonnage pondéré dans le cadre bayésien où la loi objectif est la loi $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$, nous amène aux méthodes de Monte Carlo séquentielles. Les extensions de cet algorithme vers les versions séquentielle (**échantillonnage pondéré séquentiel**, “sequential importance sampling”) et avec rééchantillonnage (**échantillonnage pondéré avec rééchantillonnage**, “sampling importance resampling”), ainsi que la notion de taille efficace de l'ensemble d'échantillons seront expliquées dans le paragraphe suivant. Nous verrons également l'importance du choix de la fonction de proposition sur la variance des poids d'importance.

2.2 Méthodes séquentielles de Monte Carlo

On considère de nouveau le problème de filtrage non linéaire, où le système est décrit par les densités $p(\mathbf{x}_0)$, $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ et $p(\mathbf{z}_k|\mathbf{x}_k)$. En particulier, les processus d'état et de mesure pourront être définis par :

$$\begin{cases} \mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{w}_k) \\ \mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{v}_k), \end{cases} \quad (2.9)$$

où aucune hypothèse de linéarité sur les fonctions f_k , h_k n'est faite et où $\mathbf{w}_k$, $\mathbf{v}_k$ sont des bruits blancs indépendants, éventuellement non gaussiens. Dans un premier temps, nous nous intéressons à la loi complète $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$, et à l'estimation d'intégrales de la forme suivante :

$$\mathbb{E}_{p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})}[\phi(\mathbf{x}_{0:k})] = I(\phi(\mathbf{x}_{0:k})) = \int \phi(\mathbf{x}_{0:k}) p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) d\mathbf{x}_{0:k} \quad (2.10)$$

Cette intégrale est impossible à calculer en pratique du fait de sa trop grande dimension. L'application de l'algorithme d'échantillonnage pondéré, avec pour loi objectif $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$, nous permet d'en obtenir un estimateur.

On suppose tirer N échantillons (ou *particules*) $\{\mathbf{x}_{0:k}^{(i)}\}$ selon la *fonction d'importance* $\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ (dont le support contient le support de $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$). En appliquant directement l'algorithme 6, on obtient l'estimateur de l'intégrale suivant :

$$I_N(\phi(\mathbf{x}_{0:k})) = \sum_{i=1}^N \phi(\mathbf{x}_{0:k}^{(i)}) \tilde{w}_k^{(i)}, \quad (2.11)$$

où les *poids d'importance normalisés* s'écrivent :

$$\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}} \quad \text{avec} \quad w_k^{(i)} = \frac{p(\mathbf{x}_{0:k}^{(i)}|\mathbf{z}_{1:k})}{\pi(\mathbf{x}_{0:k}^{(i)}|\mathbf{z}_{1:k})}. \quad (2.12)$$

De la même manière, on peut approcher la loi objectif $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ par la somme pondérée :

$$p_N(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) = \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_{0:k}^{(i)}}(\mathbf{x}_{0:k}). \quad (2.13)$$

2.2.1 Échantillonnage pondéré séquentiel

Pour le moment, l'estimation de l'intégrale décrite est une procédure non séquentielle. Or, le but est de construire une solution itérative, qui permet l'estimation de la loi *a posteriori* $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ à partir du nuage pondéré de particules approchant $p(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1})$ et de la nouvelle observation. Cela est rendu possible en supposant que la fonction d'importance est causale :

$$\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:t}) = \pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) \quad \forall t \geq k, \quad (2.14)$$

et que son expression est de la forme récursive suivante :

$$\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) = \pi(\mathbf{x}_0) \prod_{t=1}^k \pi(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{z}_{1:t}), \quad (2.15)$$

ce qui implique :

$$\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) = \pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}) \pi(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1}). \quad (2.16)$$

Chaque nouvelle particule $\mathbf{x}_{0:k}^{(i)} \sim \pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ est alors définie comme étant la particule à l'instant précédent $\mathbf{x}_{0:k-1}^{(i)} \sim \pi(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1})$ augmentée d'un nouvel état $\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k})$.

Afin d'obtenir une expression pour la mise à jour des poids, la loi $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ est exprimée en fonction de $p(\mathbf{x}_k|\mathbf{x}_{k-1})$, $p(\mathbf{z}_k|\mathbf{x}_k)$ et $p(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1})$ en utilisant le théorème de Bayes ainsi que les propriétés d'indépendance définies en (0.3, 0.4) p.17 :

$$\begin{aligned} p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) &= p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k-1}, \mathbf{z}_k) = \frac{p(\mathbf{x}_{0:k}, \mathbf{z}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} \\ &\propto p(\mathbf{z}_k|\mathbf{x}_{0:k}, \mathbf{z}_{1:k-1}) p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k-1}) \\ &\propto p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k-1}) p(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1}) \\ &\propto p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1}). \end{aligned} \quad (2.17)$$

En insérant cette relation et (2.16) dans l'expression des poids d'importance, on obtient la formule de récursivité des poids :

$$\begin{aligned} w_k^{(i)} &= \frac{p(\mathbf{x}_{0:k}^{(i)}|\mathbf{z}_{1:k})}{\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})} \\ &\propto \frac{p(\mathbf{z}_k|\mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}) p(\mathbf{x}_{0:k-1}^{(i)}|\mathbf{z}_{1:k-1})}{\pi(\mathbf{x}_k^{(i)}|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}) \pi(\mathbf{x}_{0:k-1}^{(i)}|\mathbf{z}_{1:k-1})} \\ &\propto w_{k-1}^{(i)} \frac{p(\mathbf{z}_k|\mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)}|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}. \end{aligned} \quad (2.18)$$

En résumé, la construction récursive du nuage approchant la loi *a posteriori* au temps k se fait par une procédure d'**échantillonnage pondéré séquentiel**. Tout d'abord chaque particule $\mathbf{x}_{0:k-1}^{(i)}$ est augmentée d'un nouvel état $\mathbf{x}_k^{(i)}$ tiré selon la fonction d'importance $\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})$. Puis les nouveaux poids sont calculés par (2.18) et normalisés. L'algorithme 7 décrit cette procédure. Il a l'avantage d'être parallélisable.

Il est important de noter que le nuage pondéré obtenu permet une estimation de la loi de filtrage $p(\mathbf{x}_k|\mathbf{z}_{1:k})$, et de toute quantité de la forme $\mathbb{E}_{p(\mathbf{x}_k|\mathbf{z}_{1:k})}[\phi(\mathbf{x}_k)]$ pour toute fonction ϕ intégrable par rapport à $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ (cf annexe A).

Algorithme 7 Échantillonnage pondéré séquentiel dans le cadre bayésien

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{x}_0^{(i)} \sim \pi(\mathbf{x}_0)$, et fixer $w_0^{(i)} = 1/N$

pour $k = 1, 2, \dots$

- **échantillonnage pondéré séquentiel :**

1. échantillonnage :

pour $i = 1 \dots N$, générer $\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})$ et poser $\mathbf{x}_{0:k}^{(i)} \triangleq (\mathbf{x}_{0:k-1}^{(i)}, \mathbf{x}_k^{(i)})$.

2. mise à jour des poids d'importance :

pour $i = 1 \dots N$, calculer $w_k^{(i)} = w_{k-1}^{(i)} \frac{p(\mathbf{z}_k|\mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)}|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}$.

3. normalisation des poids :

pour $i = 1 \dots N$, calculer $\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}}$

- **estimations de Monte Carlo :**

le nuage $\{\mathbf{x}_{0:k}^{(i)}, \tilde{w}_k^{(i)}\}_{i=1 \dots N}$ permet d'approcher la loi *a posteriori*

$$p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) \simeq \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_{0:k}^{(i)}}(\mathbf{x}_{0:k})$$

et pour toute fonction ϕ intégrable par rapport à $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$

$$\mathbb{E}_{p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})}[\phi(\mathbf{x}_{0:k})] \simeq \sum_{i=1}^N \phi(\mathbf{x}_{0:k}^{(i)}) \tilde{w}_k^{(i)}$$

de plus, la loi de filtrage est approchée par

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \simeq \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k)$$

2.2.2 Problème de dégénérescence et taille efficace du N-échantillon

Choisir des fonctions d'importance de la forme (2.16) conduit à une augmentation de la variance des poids dans le temps [Kong 94]. En pratique, cela a pour effet de faire décroître rapidement le nombre de particules significatives. Ce *problème de dégénérescence* conduit à une divergence du nuage de particules, et à une détérioration de l'estimation en sortie de l'algorithme d'échantillonnage pondéré séquentiel. Cette dégénérescence est d'autant plus grande que la fonction d'importance est différente de $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$.

Une mesure de la dégénérescence de l'algorithme, décrite dans [Kong 94] [Liu 95], est apportée par la *taille efficace du N-échantillon* ("effective sample size") notée ESS . Ce critère est défini en utilisant le ratio entre la variance de l'estimée obtenue par l'échantillonnage pondéré selon $\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ et la variance de l'estimée obtenue par un échantillonnage parfait selon $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$. Ce ratio reflète l'efficacité de la procédure d'échantillonnage pondéré. En remarquant que :

$$\frac{var_{\pi(\cdot|\mathbf{z}_{1:k})}[I_N(\phi(\mathbf{x}_{0:k}))]}{var_{p(\cdot|\mathbf{z}_{1:k})}[I_N(\phi(\mathbf{x}_{0:k}))]} \simeq 1 + var_{\pi(\cdot|\mathbf{z}_{1:k})}[w_k] = \mathbb{E}_{\pi(\cdot|\mathbf{z}_{1:k})}[(w_k)^2], \quad (2.19)$$

la taille efficace du N-échantillon est définie telle que

$$ESS = \frac{N}{\mathbb{E}_{\pi(\cdot|\mathbf{z}_{1:k})}[(w_k)^2]}. \quad (2.20)$$

Bien que cette expression ne puisse pas être calculée directement, il est possible de l'approcher par l'estimateur

$$ESS_N = \frac{N}{\sum_{i=1}^N (\tilde{w}_k^{(i)})^2} \quad (2.21)$$

où les $\tilde{w}_k^{(i)}$ sont les poids d'importance normalisés. Une valeur faible de ESS_N indique une forte dégénérescence.

Carpenter et al. [Carpenter 99] proposent de calculer une approximation de la taille efficace du N-échantillon par une méthode de Monte Carlo. La procédure consiste à lancer le filtre particulaire L fois sur un ensemble de données identiques avec N particules, pour obtenir L estimées indépendantes $I_N^l(\phi)$ d'une fonction d'intérêt ϕ . Ces estimées sont calculées par (2.11). Pour chaque réalisation l du filtre, une *covariance intra-réalisation* P_l est calculée, qui correspond à la covariance de $I_N^l(\phi)$.

Les méthodes séquentielles de Monte-Carlo étant fondées sur la simulation de variables aléatoires, les estimées $I_N^l(\phi)$ et P_l varient d'une réalisation du filtre à l'autre. Par conséquent, on peut calculer une *covariance inter-réalisations* qui est définie comme la covariance de l'ensemble des L estimées $\{I_N^l(\phi)\}_{l=1\dots L}$. Dans ce cas, la taille efficace du N-échantillon est approchée par le rapport entre la moyenne des covariances intra-réalisation et la covariance inter-réalisations (il est à remarquer que cette valeur dépend de la fonction ϕ).

En pratique, la première possibilité pour le calcul de ESS sera conservée, pour une raison de simplicité de mise en œuvre.

Enfin, un autre critère de la dégénérescence des poids a été proposé par Pham [Pham 98]. Il est fondé sur un critère d'entropie mesurant la dispersion des pondérations :

$$\log(N) + \sum_{i=1 \dots N} \tilde{w}_k^{(i)} \log \tilde{w}_k^{(i)}. \quad (2.22)$$

Cette quantité est toujours positive et vaut zéro uniquement lorsque tous les poids sont égaux à $1/N$.

2.2.3 Échantillonnage pondéré séquentiel avec rééchantillonnage

Pour limiter la dégénérescence des poids, deux méthodes ont été proposées (qui peuvent être couplées). Une première solution consiste à choisir une fonction d'importance appropriée, qui minimise la variance des poids et maximise *ESS*. Il a été démontré qu'une telle fonction de proposition optimale s'écrit $\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k)$ [Doucet 00b] (la démonstration sera donnée dans la section 2.3 p. 42 où le choix d'une loi de proposition est discuté). Malheureusement, en pratique, celle-ci n'est que très rarement accessible. Dans tous les cas, il est important d'utiliser une fonction qui guide les particules dans les zones de l'espace d'état de forte vraisemblance.

Une deuxième solution consiste à ajouter une étape de **rééchantillonnage** à l'algorithme d'échantillonnage pondéré séquentiel. La procédure complète est appelée **échantillonnage pondéré séquentiel avec rééchantillonnage** et constitue l'algorithme de filtrage particulaire. Le principe de cette étape est de supprimer les particules ayant un faible poids normalisé et d'ajouter des copies des particules associées à un poids fort. La technique classique consiste à tirer avec remise N nouvelles particules parmi l'ensemble des particules, proportionnellement à leur poids. Plusieurs autres méthodes de simulation peuvent être utilisées. Elles seront décrites dans la section 2.4.

L'introduction du rééchantillonnage permet une amélioration de la qualité des estimations de Monte Carlo en diminuant le problème de dégénérescence. Cependant, cette étape introduit des problèmes pratiques et théoriques. Du point de vue de la mise en œuvre, la possibilité de paralléliser l'algorithme est réduite. Au niveau théorique, le rééchantillonnage introduit une dépendance entre les particules. Cela rend les résultats de convergence difficile à établir. De plus, les particules ayant un fort poids seront sélectionnées plusieurs fois. Cela implique le problème connu sous le nom d'*appauvrissement des états*, qui décrit une perte de diversité parmi les particules. Ce problème est d'autant plus important que les modes de la vraisemblance sont très prononcés et que la densité de prédiction est peu informative. Dans un cas extrême, toutes les particules se trouvent à la même position de l'espace d'état. De nombreuses améliorations ont été proposées [Doucet 01a], qui portent sur le choix de la fonction d'importance ou le choix d'une étape de rééchantillonnage. Ces méthodes seront décrites dans les deux sections suivantes.

En résumé, la procédure de rééchantillonnage permet de corriger une divergence du nuage de particules. Cette divergence est inhérente à l'algorithme d'échantillonnage pondéré séquentiel. Elle rend compte de la différence entre la fonction d'importance $\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ et la densité recherchée $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$. Une étape de rééchantillonnage est essentielle, mais du fait de ces inconvénients, elle doit être réalisée uniquement quand cela est nécessaire. En

pratique, un tirage des particules est effectué lorsque la taille efficace du N-échantillon ESS_N est inférieure à un seuil donné. Bien qu'il n'existe pas de résultat de convergence pour l'échantillonnage pondéré séquentiel avec rééchantillonnage adaptatif, ces filtres donnent en pratique de meilleures approximations que ceux avec un rééchantillonnage systématique [Hue 03].

L'algorithme 8 présente le schéma générique du filtrage particulaire. Il se décompose en deux étapes principales : un échantillonnage pondéré séquentiel et une étape de sélection par rééchantillonnage. Notons qu'il est préférable de calculer les estimations de Monte Carlo avant d'effectuer le rééchantillonnage, car cette étape introduit artificiellement de l'aléa dans le nuage courant.

Remarque 1 : L'algorithme de filtrage particulaire présenté permet l'obtention d'un ensemble d'échantillons selon la loi $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ par la méthode d'échantillonnage pondéré. Une autre possibilité consiste à utiliser la méthode d'acceptation rejet [Pitt 99]. L'argument principal est que cet algorithme permet de tirer un échantillon exact de la loi objectif, qui ne nécessite pas de pondération. Cependant, son utilisation paraît limitée pour une application en-ligne, du fait des inconvénients décrits dans le paragraphe 2.1.2.

Remarque 2 : Pour limiter la divergence de l'algorithme due à l'accumulation des pondérations, une idée classique du domaine du filtrage consiste à limiter la mémoire du filtre utilisé. L'idée est de ne prendre en compte que les p observations les plus récentes. Il est possible de limiter la mémoire du filtre brutalement (filtre à mémoire limitée), ou progressivement (filtre à oubli exponentiel) [Oudjane 00]. Ces méthodes sont *a priori* frustrantes, car une partie des données accessibles n'est pas utilisée pour calculer un estimateur de l'état. Il semble plus intéressant d'utiliser des stratégies d'échantillonnage et de rééchantillonnage qui diminuent la dégénérescence des poids et l'appauvrissement des états.

2.3 Stratégies d'échantillonnage

Théoriquement, l'algorithme d'échantillonnage pondéré est valide, quel que soit le choix de la loi de proposition utilisée (sous réserve que son support inclut le support de la loi objectif). Le calcul des poids d'importance permet de compenser la différence entre les deux fonctions. Cependant, l'algorithme est d'autant plus efficace que la fonction d'importance $\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ est proche de $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$. Cela est accentué par l'hypothèse de récursivité de la loi de proposition, qui permet la construction d'un algorithme séquentiel, mais qui en pratique implique une dégénérescence des poids. Plusieurs stratégies d'échantillonnage peuvent être discutées. La qualité de l'estimation en dépend, à travers une répartition efficace des particules dans l'espace d'état.

2.3.1 Filtre bootstrap

La solution la plus répandue consiste à propager les particules selon la loi de transition *a priori* $\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$. Dans ce cas, l'équation de mise à jour des poids (2.18) devient :

$$w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}). \quad (2.23)$$

Algorithme 8 Échantillonnage pondéré séquentiel avec rééchantillonnage adaptatif

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0)$, et fixer $w_0^{(i)} = 1/N$

pour $k = 1, 2, \dots$

- **échantillonnage pondéré séquentiel :**

1. échantillonnage :

pour $i = 1 \dots N$, générer $\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})$ et poser $\mathbf{x}_{0:k}^{(i)} \triangleq (\mathbf{x}_{0:k-1}^{(i)}, \mathbf{x}_k^{(i)})$.

2. mise à jour des poids d'importance :

pour $i = 1 \dots N$, calculer $w_k^{(i)} = w_{k-1}^{(i)} \frac{p(\mathbf{z}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)} | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}$.

3. normalisation des poids :

pour $i = 1 \dots N$, calculer $\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}}$

- **estimations de Monte Carlo :**

le nuage $\{\mathbf{x}_{0:k}^{(i)}, \tilde{w}_k^{(i)}\}_{i=1 \dots N}$ permet d'approcher la loi *a posteriori*

$$p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) \simeq \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_{0:k}^{(i)}}(\mathbf{x}_{0:k})$$

et pour toute fonction ϕ intégrable par rapport à $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$

$$\mathbb{E}_{p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})}[\phi(\mathbf{x}_{0:k})] \simeq \sum_{i=1}^N \phi(\mathbf{x}_{0:k}^{(i)}) \tilde{w}_k^{(i)}$$

de plus, la loi de filtrage est approchée par

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) \simeq \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k)$$

- **rééchantillonnage adaptatif :**

1. calculer $ESS_N = \frac{N}{\sum_{i=1}^N (\tilde{w}_k^{(i)})^2}$

2. si $ESS_N < \text{seuil}$,

tirer avec remise N particules $\tilde{\mathbf{x}}_{0:k}^{(i)}$ parmi $\{\mathbf{x}_{0:k}^{(i)}\}_{i=1 \dots N}$ proportionnellement aux poids $\{\tilde{w}_k^{(i)}\}_{i=1 \dots N}$ et fixer $\tilde{w}_k^{(i)} = 1/N$, $\mathbf{x}_{0:k}^{(i)} = \tilde{\mathbf{x}}_{0:k}^{(i)}$ pour $i = 1 \dots N$.

Ce choix correspond aux premiers filtres particuliers avec rééchantillonnage qui ont été proposés [Gordon 93] [Isard 98a] [Kitagawa 96]. En particulier, le **filtre bootstrap** [Gordon 93] associe cette étape d'échantillonnage pondéré avec un rééchantillonnage systématique : à chaque pas de temps, les poids sont remis à $1/N$. On en déduit que l'étape de mise à jour des poids se réduit à :

$$w_k^{(i)} = p(\mathbf{z}_k | \mathbf{x}_k^{(i)}). \quad (2.24)$$

L'algorithme 9 décrit les étapes du filtre bootstrap. Historiquement, cet algorithme a été utilisé pour approcher la distribution de filtrage $p(\mathbf{x}_k | \mathbf{z}_{1:k})$. L'estimation de la loi $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ est facilement accessible en formant chaque particule $\mathbf{x}_{0:k}^{(i)}$ comme concaténation de la particule au temps précédent $\mathbf{x}_{0:k-1}^{(i)}$ et du nouvel état $\mathbf{x}_k^{(i)}$ (cette remarque ne se limite pas à l'algorithme bootstrap, elle est valable quelque soit la fonction d'importance utilisée)

Algorithme 9 Filtre bootstrap

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0)$, et fixer $w_0^{(i)} = 1/N$

pour $k = 1, 2, \dots$

- **échantillonnage pondéré séquentiel :**

1. échantillonnage : pour $i = 1 \dots N$, générer $\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)})$
2. mise à jour des poids d'importance : pour $i = 1 \dots N$, calculer $w_k^{(i)} = p(\mathbf{z}_k | \mathbf{x}_k^{(i)})$
3. normalisation des poids : pour $i = 1 \dots N$, calculer $\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}}$

- **estimations de Monte Carlo**

- **rééchantillonnage systématique :**

1. tirer avec remise N particules $\tilde{\mathbf{x}}_k^{(i)}$ parmi $\{\mathbf{x}_k^{(i)}\}_{i=1 \dots N}$ proportionnellement aux poids $\{\tilde{w}_k^{(i)}\}_{i=1 \dots N}$
 2. pour $i = 1 \dots N$, poser $\mathbf{x}_k^{(i)} = \tilde{\mathbf{x}}_k^{(i)}$ et $\tilde{w}_k^{(i)} = 1/N$
-

Cette méthode est très simple à mettre en œuvre. Dans le cas de l'estimation de la loi de filtrage, elle ne nécessite aucun stockage des valeurs des poids d'importance. Cependant, elle est souvent inefficace du fait de la non prise en compte de la nouvelle mesure dans la phase de diffusion des particules. L'espace d'état peut être mal exploré puisque les particules ne couvrent pas de façon optimale les zones de forte vraisemblance.

En particulier, cet algorithme est inefficace lorsqu'il y a peu de recouvrement entre la fonction d'importance et la fonction utilisée dans le calcul des poids, respectivement la densité de prédiction et la vraisemblance dans le cas du filtre Bootstrap. La figure 2.1 présente deux situations critiques.

- Lorsque la densité de prédiction est peu informative, et que la vraisemblance a des modes très prononcés (cas représenté en (a)), le phénomène d'appauvrissement des états est important. Durant les phases de rééchantillonnage systématique, la majorité des particules est supprimée et seules les quelques particules de poids fort sont dupliquées, ce qui entraîne une perte de diversité des particules.
- Lorsque la mesure est incohérente par rapport à la prédiction (cas représenté en (b)), aucune particule ne représente une hypothèse valide. Leur diffusion ne prenant pas en compte la nouvelle mesure, le nuage ne sera pas représentatif de la loi de filtrage.

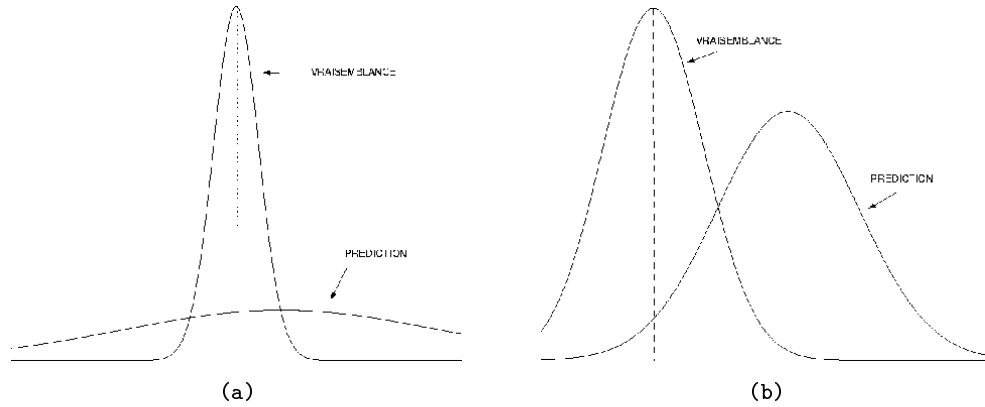


FIG. 2.1: Fonction de vraisemblance et densité prédite pour une observation précise (a) ou incohérente (b) par rapport à la prédiction (dans le cas d'un état directement observé i.e. $\mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k$) - figure tirée de [Oudjane 00].

Notons que ces situations peuvent se rencontrer dans les applications de suivi dans des séquences d'images. La situation (a) correspond à une modélisation souvent utilisée en imagerie, à savoir une densité de prédiction de l'objet suivi peu informative, et une vraisemblance hautement multimodale et porteuse d'informations. La situation (b) peut correspondre à une sortie d'occlusion, lorsque l'hypothèse d'évolution de l'objet ne correspond pas à sa dynamique pendant son occultation.

Pour mieux guider les particules dans les zones de forte vraisemblance tout en gardant le schéma du filtre bootstrap, Gordon [Gordon 93] propose d'ajouter une procédure de **prior editing**, qui augmente artificiellement le nombre de particules de poids fort. L'idée consiste à ajouter un test d'acceptation de l'échantillon en fonction de son poids d'importance, par rapport au terme de différence entre la mesure observée et la mesure prédite. Si le poids de la particule est trop faible, celle-ci aura une faible chance d'être sélectionnée. Elle est donc rejetée. Dans le cas contraire, elle est conservée. La procédure est itérée jusqu'à ce que N particules aient été acceptées.¹

Cette méthode est très similaire à l'idée de **prior boosting**, également introduite par Gordon [Gordon 95], qui consiste à générer kN particules (avec k positif) durant l'étape

¹Liu a proposé une méthode de sélection des particules appelée "**rejection control**" [Liu 98b] proche de l'idée de prior editing. Celle-ci est également fondée sur un test d'acceptation de chaque échantillon. Elle n'est pas intéressante en pratique car elle fait diminuer le nombre de particules.

d'échantillonnage. Seuls N échantillons sont sélectionnés pendant le rééchantillonnage. Ces approches sont fondées sur le fait qu'augmenter le nombre d'échantillons accroît les chances de mieux représenter la distribution *a posteriori*. Cependant, ces méthodes *ad hoc* ne sont pas toujours intéressantes, car elles induisent une trop forte augmentation du coût de calcul, pour un effort de recherche qui peut se révéler inefficace.

En résumé, se limiter à utiliser la densité de prédiction pour la phase de diffusion des particules constitue dans de nombreux cas un mauvais choix. Afin de construire un nuage représentatif selon la distribution de filtrage, des particules doivent être proposées dans les régions de forte vraisemblance. Idéalement, les particules doivent être guidées en fonction de la dynamique de l'objet et de la nouvelle mesure disponible. Cela est le cas de la fonction d'importance optimale.

2.3.2 Fonction d'importance optimale

Afin de limiter la dégénérescence des poids d'importance, il est naturel de considérer la loi de proposition qui minimise la variance des poids, conditionnellement à $\mathbf{x}_{0:k-1}^{(i)}$ et $\mathbf{z}_{1:k}$. Il est possible de démontrer que cette loi, appelée *fonction d'importance optimale*, est de la forme suivante [Doucet 00b] :

$$\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k). \quad (2.25)$$

PREUVE : _____

L'espérance des poids d'importance s'écrit :

$$\mathbb{E}_{\pi(\cdot | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}[w_k^{(i)}] = \int w_{k-1}^{(i)} p(\mathbf{z}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}) d\mathbf{x}_k = w_{k-1}^{(i)} p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}). \quad (2.26)$$

D'autre part,

$$\mathbb{E}_{\pi(\cdot | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}[(w_k^{(i)})^2] = (w_{k-1}^{(i)})^2 \int \frac{(p(\mathbf{z}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}))^2}{\pi(\mathbf{x}_k^{(i)} | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})} d\mathbf{x}_k \quad (2.27)$$

Si on choisit $\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$, cette expression devient :

$$\begin{aligned} \mathbb{E}_{\pi(\cdot | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}[(w_k^{(i)})^2] &= (w_{k-1}^{(i)})^2 (p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}))^2 \frac{1}{\int p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k) d\mathbf{x}_k} \\ &= (w_{k-1}^{(i)})^2 (p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}))^2 \end{aligned} \quad (2.28)$$

D'où :

$$\text{var}_{\pi(\cdot | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}[w_k^{(i)}] = \mathbb{E}_{\pi(\cdot | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}[(w_k^{(i)})^2] - \mathbb{E}_{\pi(\cdot | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}^2[w_k^{(i)}] = 0 \quad (2.29)$$

_____ □

Choisir cette fonction d'importance modifie le calcul des poids d'importance. En remarquant que :

$$\frac{p(\mathbf{z}_k|\mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)} = \frac{p(\mathbf{z}_k, \mathbf{x}_k^{(i)}, \mathbf{x}_{k-1}^{(i)})}{p(\mathbf{x}_{k-1}^{(i)}) p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)} = p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}), \quad (2.30)$$

l'équation de mise à jour des poids (2.18) devient :

$$w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}). \quad (2.31)$$

En pratique, il devient possible de réaliser simultanément la diffusion des particules et la mise à jour des poids.

Pour pouvoir utiliser la fonction d'importance optimale, il est nécessaire de tirer selon $p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$ et d'évaluer à une constante près

$$p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}) = \int p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}) d\mathbf{x}_k, \quad (2.32)$$

ce qui est en général impossible sauf pour quelques modèles spécifiques. On peut citer le cas où l'espace d'état est discret et constitué d'un nombre faible d'états possibles. L'intégrale (2.32) devient alors une somme qu'il est possible de calculer. Lorsque le modèle considéré implique que $p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)})$ est gaussien, cette intégrale est également calculable. Cela est le cas pour les modèles de la forme [Doucet 00b] :

$$\begin{cases} \mathbf{x}_k = f_k(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ \mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (2.33)$$

où la fonction f_k peut être non linéaire et où les bruits d'état et de mesure sont des bruits blancs gaussiens, non corrélés entre eux, de moyennes nulles et de covariances respectives Q_k et R_k . En observant que :

$$p(\mathbf{z}_k|\mathbf{x}_{k-1}) = \int p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1}) d\mathbf{x}_k \quad (2.34)$$

et

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1})}{p(\mathbf{z}_k|\mathbf{x}_{k-1})}, \quad (2.35)$$

on obtient les expressions des distributions utiles à l'utilisation de la fonction d'importance optimale :

$$p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}) = \mathcal{N}(\mathbf{z}_k; H_k f_k(\mathbf{x}_{k-1}^{(i)}), R_k + H_k Q_k H_k^t) \quad (2.36)$$

et

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_k, \Sigma_k) \quad (2.37)$$

avec

$$\Sigma_k = (Q_k^{-1} + H_k^t R_k^{-1} H_k)^{-1} \quad (2.38)$$

$$\mathbf{m}_k = \Sigma_k (Q_k^{-1} f_k(\mathbf{x}_{k-1}) + H_k^t R_k^{-1} \mathbf{z}_k). \quad (2.39)$$

Nous verrons dans la deuxième partie de ce document une extension de ce système où la vraisemblance est modélisée par un mélange de lois gaussiennes, pour lequel il est également possible d'utiliser la fonction d'importance optimale.

2.3.3 Autres stratégies, quand la fonction d'importance optimale n'est pas disponible

Lorsque la fonction d'importance optimale n'est pas disponible, différentes méthodes ont été proposées pour l'approcher, ou pour guider les particules vers les zones de forte vraisemblance.

Ajout d'une étape de Monte Carlo

Une première possibilité consiste à mettre en œuvre une deuxième étape de Monte Carlo pour approcher la fonction d'importance optimale [Doucet 98]. Pour cela, une procédure d'acceptation rejet ou d'échantillonnage pondéré peut être utilisée. Cette approche n'est pas intéressante dans un cadre d'algorithme séquentiel car elle est trop coûteuse en temps de calcul.

D'autres méthodes de Monte Carlo par chaînes de Markov (MCMC) peuvent être introduites pour une meilleure diffusion des particules, en particulier lorsque la différence entre la densité de prédiction et la vraisemblance est trop importante pour les déplacer en un pas. L'idée est de faire migrer les particules par une ou plusieurs itérations de MCMC, en introduisant des densités intermédiaires. Ces densités sont choisies de façon à introduire graduellement l'information de la vraisemblance dans la fonction d'importance.

Ces filtres hybrides, qui combinent un filtrage particulaire avec des itérations de MCMC correspondent à un développement séquentiel de l'algorithme d'échantillonnage pondéré refroidi ("**annealed importance sampling**") proposé par Neal [Neal 01]. On peut citer par exemple le filtre particulaire avec **correction progressive** [Musso 01] ou le filtre particulaire couplé à l'utilisation de "**bridging densities**" [Godsill 01]. La difficulté d'utilisation de ces algorithmes vient du choix des distributions intermédiaires (en particulier, choix de la séquence des coefficients de refroidissement), et de leur nombre. De plus, l'ajout d'itérations de MCMC augmente significativement le coût de calcul.

Filtre particulaire étendu et filtre particulaire unscented

Une idée alternative pour l'approximation de la loi de proposition optimale repose sur les développements des extensions du filtre de Kalman. On suppose une forme gaussienne de la fonction d'importance $\pi(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{(i)}, \Sigma_{k|k}^{(i)})$. À chaque particule est associé un filtre de Kalman étendu ou unscented, qui permet le calcul des moments de cette approximation soit par linéarisation [Doucet 00b], soit par estimation sur un nuage de point [Merwe 00], en utilisant la nouvelle mesure $\mathbf{z}_k$.

Les filtres construits sont appelés **filtre particulaire étendu** ("extended particle filter") et **filtre particulaire unscented** ("unscented particle filter"). Ces méthodes prennent en compte la vraisemblance lors de la propagation des particules. Elles ont donc des performances supérieures au filtre bootstrap. Enfin, nous avons vu dans la section 1.2.2 les avantages du filtre de Kalman unscented, ce qui rend préférable son utilisation par rapport au filtre de Kalman étendu.

De façon similaire au cas non aléatoire, le filtre particulaire unscented peut être étendu au **filtre particulaire unscented, par mélange de gaussiennes** [Merwe 03].

Cette méthode représente la densité *a posteriori* par une somme finie de loi gaussiennes, dont les paramètres sont déterminés par l'algorithme EM ("expectation-maximisation") sur le nuage pondéré des particules. Le nuage évolue par le principe d'échantillonnage pondéré avec rééchantillonnage, en utilisant un banc de filtre de Kalman unscented pour la propagation des particules.

Filtre particulaire auxiliaire

Le **filtre particulaire auxiliaire** ("auxiliary particle filter") a été introduit par Pitt et Shepard [Pitt 99] [Pitt 01] dans le but de réduire le problème de dégénérescence quand la vraisemblance est très informative ou que l'observation se trouve dans la queue de la densité de prédiction. Le schéma d'échantillonnage pondéré avec rééchantillonnage est conservé, en introduisant la loi de proposition $\pi(\mathbf{x}_k, i | \mathbf{z}_{1:k})$, utilisée pour tirer les couples $\{\mathbf{x}_k^{(j)}, i^j\}_{j=1\dots N}$ où i^j est la variable auxiliaire qui désigne l'indice de la particule au temps précédent.

Le filtre particulaire auxiliaire consiste à générer un ensemble d'échantillons selon la densité jointe $p(\mathbf{x}_k, i | \mathbf{z}_{1:k})$. Puis les indices sont supprimés dans les couples $\{\mathbf{x}_k^{(j)}, i^j\}_{j=1\dots N}$ pour obtenir un ensemble d'échantillons $\{\mathbf{x}_k^{(j)}\}_{j=1\dots N}$ de la loi marginale $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ (qui correspond à la loi de filtrage recherchée). En remarquant que

$$\begin{aligned} p(\mathbf{x}_k, i | \mathbf{z}_{1:k}) &\propto p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k, i | \mathbf{z}_{1:k-1}) \\ &= p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | i, \mathbf{z}_{1:k-1}) p(i | \mathbf{z}_{1:k-1}) \\ &= p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}) w_{k-1}^{(i)}, \end{aligned} \quad (2.40)$$

la fonction d'importance est naturellement choisie telle que

$$\pi(\mathbf{x}_k, i | \mathbf{z}_{1:k}) \propto p(\mathbf{z}_k | \mu_k^{(i)}) p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}) w_{k-1}^{(i)}, \quad (2.41)$$

où $\mu_k^{(i)}$ caractérise la loi conditionnelle $p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)})$ (par exemple, $\mu_k^{(i)}$ peut être défini comme la moyenne). En écrivant

$$\pi(\mathbf{x}_k, i | \mathbf{z}_{1:k}) = \pi(i | \mathbf{z}_{1:k}) \pi(\mathbf{x}_k | i, \mathbf{z}_{1:k}) \quad (2.42)$$

et en posant

$$\pi(\mathbf{x}_k | i, \mathbf{z}_{1:k}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}) \quad (2.43)$$

on obtient

$$\pi(i | \mathbf{z}_{1:k}) \propto p(\mathbf{z}_k | \mu_k^{(i)}) w_{k-1}^{(i)}. \quad (2.44)$$

L'idée est ainsi de d'abord tirer un indice i^j selon $\pi(i | \mathbf{z}_{1:k}) \propto p(\mathbf{z}_k | \mu_k^{(i)}) w_{k-1}^{(i)}$ et de faire évoluer la particule d'indice sélectionné selon la loi de propagation *a priori* $p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i^j)})$. Sur le même principe que (2.18) (p. 38), le poids d'importance attribué au couple $(\mathbf{x}_k^j, i^j)$ est

$$w_k^{(j)} = w_{k-1}^{(i^j)} \frac{p(\mathbf{z}_k | \mathbf{x}_k^{(j)}) p(\mathbf{x}_k^{(j)} | \mathbf{x}_{k-1}^{(i^j)})}{\pi(\mathbf{x}_k^{(j)}, i^j | \mathbf{z}_{1:k})} = \frac{p(\mathbf{z}_k | \mathbf{x}_k^{(j)})}{p(\mathbf{z}_k | \mu_k^{(i^j)})}. \quad (2.45)$$

Le filtre particulaire auxiliaire est résumé par l'algorithme 10. Comme cela est décrit dans [Ristic 04b], il équivaut à effectuer une étape de rééchantillonnage des particules à

l'instant $k-1$ (en utilisant la nouvelle mesure disponible au temps k), avant la propagation des particules au temps k . Ainsi, bien qu'une étape de rééchantillonnage à la fin du temps k soit décrite dans l'algorithme original, celle-ci n'est pas nécessaire. On peut également remarquer que choisir la densité de prédiction pour le tirage des particules (2.43) n'est pas obligatoire, et que n'importe quelle densité de proposition peut être utilisée.

Les performances du filtre particulaire auxiliaire sont meilleures que celles du filtre bootstrap lorsque les pics de la vraisemblance coïncident avec des zones où l'*a priori* est faible. Cela est due à une meilleure propagation des particules dont la vraisemblance prédite est élevée. Néanmoins, si les deux distributions coïncident, il n'est pas nécessaire d'utiliser cet algorithme car l'introduction de variables auxiliaires augmente la variance de l'échantillon.

Algorithme 10 Filtre particulaire auxiliaire

- **initialisation :**

pour $j = 1 \dots N$, générer $\mathbf{x}_0^{(j)} \sim p(\mathbf{x}_0)$, et fixer $w_0^{(j)} = 1/N$

- **pour $k = 1, 2, \dots$:**

1. pour $i = 1 \dots N$ calculer, $\mu_k^{(i)}$ en fonction de $p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)})$

2. échantillonnage des variables auxiliaires :

pour $j = 1 \dots N$, générer $i^j \sim \pi(i | \mathbf{z}_{1:k}) \propto p(\mathbf{z}_k | \mu_k^{(i)}) w_{k-1}^{(i)}$

3. échantillonnage des particules :

pour $j = 1 \dots N$, générer $\mathbf{x}_k^{(j)} \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i^j)})$

4. mise à jour des poids d'importance : pour $i = 1 \dots N$, calculer $w_k^{(j)} = \frac{p(\mathbf{z}_k | \mathbf{x}_k^{(j)})}{p(\mathbf{z}_k | \mu_k^{(i^j)})}$

5. normalisation des poids : pour $j = 1 \dots N$, calculer $\tilde{w}_k^{(j)} = \frac{w_k^{(j)}}{\sum_{i=1}^N w_k^{(i)}}$

Échantillonnage par la vraisemblance

Afin de prendre en compte la nouvelle observation pour la diffusion des particules, [Fox 01] propose l'utilisation d'une fonction de proposition égale à la vraisemblance du modèle (**likelihood sampling**). Cela amène à considérer un algorithme dual du filtre particulaire classique. La méthode proposée consiste à prédire les particules selon leur vraisemblance. Puis le nuage prédit est corrigé à travers un calcul des poids d'importance proportionnellement à la densité de prédiction *a priori*. Cette phase de correction revient à projeter en arrière la particule prédite pour connaître sa probabilité d'avoir un prédécesseur parmi le nuage à l'instant précédent. La difficulté du tirage selon la vraisemblance est un problème de cette méthode.

De récents développements ont étendu cette méthode vers un échantillonnage local par la vraisemblance (**local likelihood sampling**) [Torma 04]. L'idée est de déplacer les particules selon la prédiction, puis de les faire localement se rapprocher d'un mode

de la vraisemblance. Le calcul des poids d'importance prend en compte ce double échantillonnage. Cet algorithme se révèle efficace lorsque la vraisemblance est très informative, contrairement à la densité de prédiction (cas (a) de la figure (2.1)) , mais ne permet pas de résoudre le problème de dégénérescence lorsqu'il n'y a pas de recouvrement entre ces deux densités (cas (b) de la figure (2.1)).

Échantillonnage d'un bloc de variables

Les techniques présentées jusqu'ici proposent l'introduction de la dernière mesure $\mathbf{z}_k$ dans la fonction d'importance, pour une prédiction à un pas des nouvelles particules au temps k . La prédiction à un pas peut être généralisée par une prédiction à L pas, en considérant une fonction d'importance $\pi(\mathbf{x}_{k-L:k}|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k})$ permettant de tirer un bloc de variables $\mathbf{x}_{k-L:k}$. Par extension au cas $L = 1$, la loi de proposition minimisant la variance des poids utilise un bloc de mesures et s'écrit $p(\mathbf{x}_{k-L:k}|\mathbf{x}_{k-L-1}, \mathbf{z}_{k-L:k})$. Dans le cas le plus simple, on peut utiliser $p(\mathbf{x}_{k-L:k}|\mathbf{x}_{k-L-1})$.

Les résultats présentés dans [Doucet 04] montrent qu'une prédiction d'un bloc de variables donne de meilleurs résultats qu'une prédiction à un pas, pour une dégénérescence moindre des poids d'importance (la taille efficace du N-échantillon diminue moins vite avec le temps, et le nombre de phases de rééchantillonnage nécessaires diminue). Cependant, échantillonner un bloc de variables est une étape difficile et coûteuse. Elle peut être réalisée par des récursions "forward-backward", ou à l'aide d'approximations de la loi de proposition (modèles de mélanges, algorithmes MCMC, etc.).

Stratégies d'échantillonnage pour le suivi dans des séquences d'images

Dans le cadre du suivi dans des séquences d'images, les stratégies d'échantillonnage utilisées consistent en une des méthodes décrites précédemment (fonction de proposition égale à la densité de prédiction ou permettant de prendre en compte la nouvelle mesure), ou en une méthode qui combine la densité de prédiction et le résultat d'un processus réalisé sur les images. En voici deux exemples : une fonction de proposition qui autorise un grand déplacement des particules peut être définie pour rendre plus effective l'exploration de l'espace d'état. Cette définition repose sur un module de détection [Pérez 04] [Vermaak 02a]. Une descente de gradient sur l'image peut aussi être mise en place, pour la construction d'une loi qui fusionne une recherche déterministe avec une recherche guidée par le modèle de prédiction *a priori* [Sullivan 01].

Ces techniques sont décrites dans la deuxième partie, où les modèles utilisés pour le suivi dans des séquences d'images sont présentés.

2.4 Stratégies de rééchantillonnage

Comme décrit précédemment, l'utilisation du simple algorithme d'échantillonnage pondéré séquentiel ne permet pas une estimation valide de la loi *a posteriori* du fait de la dégénérescence des poids. Au cours du temps, de plus en plus de particules deviennent inefficaces et le nuage ne permet pas de couvrir efficacement l'espace d'état. L'ajout de l'étape de rééchantillonnage permet de limiter cette dégénérescence. Son principe est de

supprimer les particules ayant un faible poids normalisé et d'ajouter des copies des particules associées à un poids fort. Pour cela, plusieurs méthodes ont été proposées, qui sont décrites dans cette section dans le cadre de l'estimation de la loi de filtrage $p(\mathbf{x}_k | \mathbf{z}_{1:k})$.

2.4.1 Sélection des particules

On note N_i le nombre de descendants de la particule $\mathbf{x}_k^{(i)}$. Les méthodes suivantes assurent qu'en moyenne le nombre de descendants de chaque particule est proportionnel à son poids, i.e. $\mathbb{E}(N_i) = N \tilde{w}_k^{(i)}$. Elles diffèrent sur la variance de Monte Carlo des nombres de descendants $\text{var}(N_i)$. Une méthode de rééchantillonnage sera d'autant meilleure que cette variance est faible. En effet, une variance forte signifie l'introduction d'un aléa sur la distribution estimée, induisant une source d'erreur supplémentaire.

Rééchantillonnage multinomial

L'approche classique de rééchantillonnage obtient le nombre de descendants de chaque particule par tirage selon la loi multinomiale de paramètres N et $\{\tilde{w}_k^{(i)}\}_{i=1\dots N}$ [Liu 95]. Cela est équivalent à tirer avec remise N nouvelles particules $\tilde{\mathbf{x}}_k^{(i)}$ parmi l'ensemble $\{\mathbf{x}_k^{(i)}\}_{i=1\dots N}$ proportionnellement aux poids $\{\tilde{w}_k^{(i)}\}_{i=1\dots N}$. Les nouvelles particules correspondent à un ensemble d'échantillons i.i.d. de l'approximation discrète de $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ et leur poids associé est remis à $1/N$.

La procédure est décrite par l'algorithme 11. Notons qu'il est possible d'implémenter cette procédure en $O(N)$ opérations [Doucet 98] en tirant N variables uniformes ordonnées. La variance de l'algorithme est $\text{var}(N_i) = N \tilde{w}_k^{(i)}(1 - \tilde{w}_k^{(i)})$. De part sa simplicité, cet algorithme est le plus couramment utilisé. Cependant, il est possible d'utiliser d'autres méthodes, de variance moindre.

Algorithme 11 Algorithme de rééchantillonnage multinomial, en $O(N \log(N))$

1. calculer la suite des poids d'importance normalisés cumulés $wc_i = \sum_{j=1}^i \tilde{w}_k^{(j)}$
 2. pour $i = 1 \dots N$
 - (a) générer $u_i \sim U[0, 1]$
 - (b) calculer l'indice j tel que $wc_{j-1} < u_i \leq wc_j$
 - (c) poser $\tilde{\mathbf{x}}_k^{(i)} = \mathbf{x}_k^{(j)}$
 3. pour $i = 1 \dots N$, poser $\mathbf{x}_k^{(i)} = \tilde{\mathbf{x}}_k^{(i)}$ et $\tilde{w}_k^{(i)} = 1/N$
-

Rééchantillonnage résiduel

Le **rééchantillonnage résiduel** [Liu 98a] procède en deux étapes. Tout d'abord, chaque particule $\mathbf{x}_{0:k}^{(i)}$ est sélectionnée $\tilde{N}_i = \lfloor N \tilde{w}_k^{(i)} \rfloor$ fois (on note $\lfloor x \rfloor$ la partie entière de x). Puis, afin de conserver un nombre constant de particules, une procédure de rééchantillonnage multinomial est utilisée pour tirer les $N' = N - \sum_{i=1}^N \tilde{N}_i$ particules

restantes, proportionnellement aux nouveaux poids $\tilde{w}_k^{\prime(i)} = (N \tilde{w}_k^{(i)} - \tilde{N}_i)/N'$. La variance de cette méthode est $\text{var}(N_i) = N' \tilde{w}_k^{\prime(i)}(1 - \tilde{w}_k^{\prime(i)})$, ce qui est inférieur à la variance de la procédure classique.

Méthodes à variance minimale

La méthode de **rééchantillonnage systématique**² [Fearnhead 98] [Carpenter 99] procède de la même manière que le rééchantillonnage multinomial, en utilisant :

$$u_1 \sim U[0, 1/N] \quad u_i = u_1 + i/N \quad (2.46)$$

à la place de la suite des u_i tirés indépendamment selon $U[0, 1]$. Cela permet un gain du temps de calcul, puisqu'une seule variable doit être tirée. En utilisant cette méthode, le nombre de descendants de chaque particule $\mathbf{x}_k^{(i)}$ ne diffère pas de $N \tilde{w}_k^{(i)}$ de plus de un. La variance de Monte Carlo de cette technique est ainsi minimale.

Cela est également le cas pour les méthodes de **rééchantillonnage par mécanisme de branchement** ("branching correction") [Crisan 99] [Crisan 01]. Une de ces méthodes débute comme le rééchantillonnage résiduel, puis les particules restantes sont choisies de la façon suivante : pour chaque indice i , sélectionner la particule i avec la probabilité $N' \tilde{w}_k^{\prime(i)}$. Cette méthode a pour inconvénient de faire fluctuer le nombre de particules. D'autres algorithmes par branchement sont donnés dans [Crisan 99].

Quelle que soit la méthode de rééchantillonnage mise en œuvre, l'utiliser à chaque pas de temps de l'algorithme de filtrage particulaire est dangereux. En effet, cela conduit à l'appauvrissement des particules, puisque la majorité des particules (celles de poids faible) sera supprimée et que celles de poids fort seront dupliquées plusieurs fois. Clairement, un équilibre doit être trouvé entre rééchantillonner à chaque pas de temps et ne jamais rééchantillonner. Comme cela a été expliqué précédemment, une solution consiste à se fonder sur la taille efficace du N-échantillon à chaque itération pour prendre la décision de sélectionner les particules ou non. On trouvera dans [Hue 03] une série d'expériences - dans le cadre applicatif du pistage mono-cible par mesure d'angle seul - qui montrent la supériorité d'un rééchantillonnage adaptatif par rapport à un rééchantillonnage à chaque pas de temps.

2.4.2 Perturbation des particules

Le phénomène d'appauvrissement des états est d'autant plus important dans le cas de systèmes faiblement bruités. Cela a conduit à proposer une étape supplémentaire de perturbations des particules au moment du rééchantillonnage. La régularisation peut constituer l'étape clé de cette perturbation. Elle est fondée sur la théorie de l'estimation de densités par noyaux, et permet de remplacer l'approximation discrète fournie par les filtres particuliers en une approximation lisse. Une autre possibilité consiste à perturber

²Le terme "systematic resampling" est utilisé pour définir une méthode de rééchantillonnage, à ne pas confondre avec le fait de réaliser une étape de rééchantillonnage à chaque pas de temps.

le nuage par l'ajout d'une itération de MCMC.

Lissage par noyaux gaussiens

Une première proposition d'une étape de perturbation des particules a été définie par Gordon [Gordon 93], qui suggère l'ajout d'un bruit dans le modèle d'état de chaque particule, avant sa propagation à l'instant suivant. Cette méthode est nommée **roughening** ou **jittering**. Ainsi, même si une particule de poids fort est dupliquée plusieurs fois, ses descendantes ne correspondront pas à une position de l'espace d'état identique. Cette technique peut être interprétée comme l'utilisation d'une estimation par noyaux gaussiens pour lisser la représentation par une somme de mesures de Dirac de la distribution *a posteriori*. La largeur du noyau correspond au paramètre de variance additionnel. Elle peut être choisie en suivant les règles classiques de l'estimation par noyaux [Duda 73].

Filtre particulaire régularisé

L'idée du **filtre particulaire régularisé** [Oudjane 00] [Musso 01] est de changer l'approximation discrète de $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ en une approximation continue, comme cela est illustré par la figure 2.2. Ainsi, l'étape de rééchantillonnage est réalisée par des simulations sur

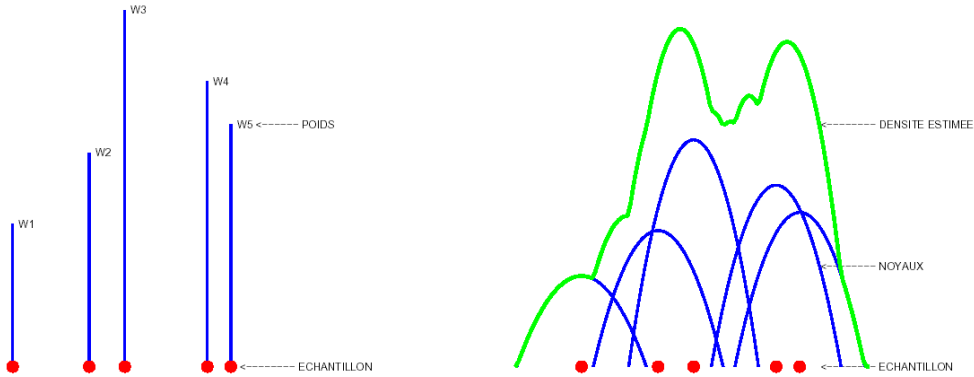


FIG. 2.2: Approximation de la loi de filtrage $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ par une somme de mesures de Dirac pondérées et approximation par une densité à noyaux - figure tirée de [Oudjane 00]

une loi continue, ce qui permet de produire un système de particules avec N positions différentes. Dans le cadre de l'estimation de la loi de filtrage, les nouvelles particules sont tirées selon

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) \simeq \sum_{i=1}^N \frac{1}{h^{n_x}} \tilde{w}_k^{(i)} K\left(\frac{\mathbf{x}_k - \mathbf{x}_k^{(i)}}{h}\right) \quad (2.47)$$

où $K(\cdot)$ désigne le noyau, h le paramètre de dilatation et n_x la dimension du vecteur d'état. La difficulté de cette méthode vient du choix du noyau de régularisation et du paramètre h . Ils sont définis tels que une estimée de l'erreur entre l'approximation (2.47) et la vraie densité *a posteriori* soit minimale. Dans le cas où tous les poids valent $1/N$, le choix optimal est le noyau d'Epanechnikov [Oudjane 00], mais il peut être remplacé par

un noyau gaussien pour une réduction de coût, sans grande influence sur l'erreur commise. Le choix du h optimal est plus problématique car son expression dépend de la densité sous-jacente inconnue et son choix est déterminant pour le comportement de l'estimée. On trouvera une discussion sur ce point dans le chapitre 4 de [Oudjane 00].

Les filtres particuliers pré- et post-régularisés ont été définis, suivant que la régularisation se fait avant ou après l'étape de correction. Notons qu'un filtre similaire au filtre pré-régularisé a été défini par Hürzeler [Hürzeler 98]. Le filtre post-régularisé est plus simple à implémenter. L'étape de rééchantillonnage modifiée est décrite par l'algorithme 12. Sa mise en œuvre nécessite le calcul de la covariance empirique du nuage pondéré de particules.

Algorithme 12 Algorithme de rééchantillonnage pour le filtre particulière post-régularisé

1. calculer la covariance empirique S_k de $\{\mathbf{x}_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1\dots N}$
 2. calculer D_k tel que $D_k D_k^t = S_k$
 3. rééchantillonner les particules selon l'algorithme 11
 4. pour $i = 1\dots N$
 - (a) générer $\epsilon^{(i)} \sim K$ (noyau gaussien ou d'Epanechnikov)
 - (b) $\mathbf{x}_k^{(i)} = \mathbf{x}_k^{(i)} + h D_k \epsilon_i$
-

Ajout d'une étape de Monte Carlo

Une alternative à la régularisation pour diversifier les particules est de les déplacer selon une méthode de Monte Carlo par chaînes de Markov [Andrieu 01] [Berzuini 01]. En effet, si les particules sont distribuées selon la loi $p(\mathbf{x}_k|\mathbf{z}_{1:k})$, et qu'on leur applique un noyau de transition markovien de distribution invariante $p(\mathbf{x}_k|\mathbf{z}_{1:k})$, alors les nouvelles particules sont toujours distribuées selon la même loi d'intérêt. Après l'étape de rééchantillonnage, chaque particule est perturbée indépendamment par le noyau de transition, ce qui contribue à améliorer l'exploration de l'espace d'état. Pour cela, toutes les méthodes de MCMC (échantillonneur de Gibbs, algorithme de Metropolis-Hastings, etc.) peuvent être utilisées. Un tel filtre hybride est par exemple l'algorithme **"resample move"** proposé par Gilks et Berzuini [Gilks 01] [Berzuini 01].

2.5 Réduction de variance par rao-blackwellisation

Bien que les méthodes de Monte Carlo soient en principe applicables en dimension quelconque, il est évidemment important de tirer parti de toutes les particularités du problème, de façon à réduire la partie de l'espace d'état où les particules seront vraiment nécessaires. S'il est possible de diviser le vecteur d'état en deux composantes, dont une partie peut être traitée par une méthode optimale, alors la méthode de réduction de la variance par rao-blackwellisation peut être utilisée [Casella 96].

Ce principe a été largement utilisé en pratique, pour des applications de robotique mobile [Doucet 00a] [Freitas 04], de suivi dans des séquences d'images [Khan 04], d'iden-

tification de fautes [Freitas 02], d'étude de séries de données financières [Andrieu 99], de communication numérique [Chen 00a] [Chen 00b] ou de suivi par mesure d'angle seul et navigation [Gustafsson 02] [Schön 03b] [Särkkä 04].

On suppose que les variables d'état $\mathbf{x}_k$ peuvent se diviser en deux groupes $\mathbf{r}_k$ et $\mathbf{y}_k$ tels que :

$$p(\mathbf{r}_k | \mathbf{r}_{k-1}, \mathbf{y}_{k-1}) = p(\mathbf{r}_k | \mathbf{r}_{k-1}), \quad (2.48)$$

ce qui implique la forme suivante de la densité de prédiction :

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}) = p(\mathbf{r}_k, \mathbf{y}_k | \mathbf{r}_{k-1}, \mathbf{y}_{k-1}) = p(\mathbf{r}_k | \mathbf{r}_{k-1}) p(\mathbf{y}_k | \mathbf{r}_{k-1:k}, \mathbf{y}_{k-1}). \quad (2.49)$$

La figure 2.3 illustre la chaîne de Markov cachée constituée des processus $\mathbf{r}_{0:k}$, $\mathbf{y}_{0:k}$ et $\mathbf{z}_{1:k}$.

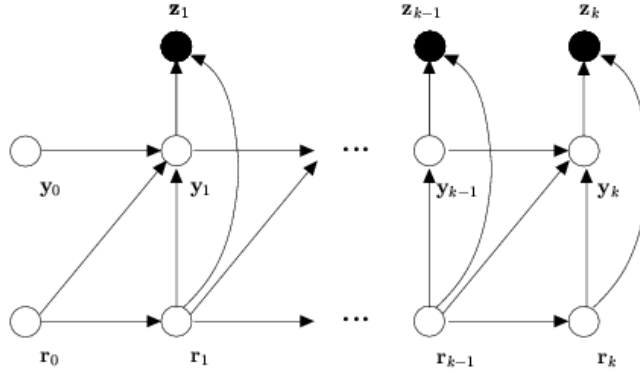


FIG. 2.3: Graphe de dépendance orienté décrivant l'évolution temporelle de la chaîne de Markov cachée constituée des processus $\mathbf{r}_{0:k}$, $\mathbf{y}_{0:k}$ et $\mathbf{z}_{1:k}$.

Si on suppose également que la distribution *a posteriori* conditionnelle $p(\mathbf{y}_{0:k} | \mathbf{r}_{0:k}, \mathbf{z}_{1:k})$ peut être estimée de façon optimale (par un filtre de Kalman, une méthode numérique par maillage d'état, etc. selon le type de modèle utilisé) ; la distribution objectif $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ pouvant s'écrire :

$$p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) = p(\mathbf{y}_{0:k} | \mathbf{r}_{0:k}, \mathbf{z}_{1:k}) p(\mathbf{r}_{0:k} | \mathbf{z}_{1:k}), \quad (2.50)$$

la difficulté se résume alors en l'estimation de $p(\mathbf{r}_{0:k} | \mathbf{z}_{1:k})$. Pour cela, un filtre particulière classique peut être utilisé. Cette densité étant située dans un espace restreint par rapport à la densité objectif, la procédure de marginalisation permet de réduire l'effort de recherche.

Si on dispose d'un nuage pondéré $\{\mathbf{r}_{0:k}^{(i)}, w_k^{(i)}\}_{i=1 \dots N}$ permettant d'estimer la densité :

$$p(\mathbf{r}_{0:k} | \mathbf{z}_{1:k}) \simeq \sum_{i=1}^N w_k^{(i)} \delta_{\mathbf{r}_{0:k}^{(i)}}(\mathbf{r}_{0:k}), \quad (2.51)$$

alors la densité marginale de $\mathbf{y}_{0:k} | \mathbf{z}_{1:k}$ est approchée par un mélange de lois calculables de

façon optimale :

$$\begin{aligned}
p(\mathbf{y}_{0:k}|\mathbf{z}_{1:k}) &= \int p(\mathbf{y}_{0:k}|\mathbf{r}_{0:k}, \mathbf{z}_{1:k}) p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k}) d\mathbf{r}_{0:k} \\
&\simeq \int p(\mathbf{y}_{0:k}|\mathbf{r}_{0:k}, \mathbf{z}_{1:k}) \sum_{i=1}^N w_k^{(i)} \delta_{\mathbf{r}_{0:k}^{(i)}}(\mathbf{r}_{0:k}) d\mathbf{r}_{0:k} \\
&\simeq \sum_{i=1}^N w_k^{(i)} p(\mathbf{y}_{0:k}|\mathbf{r}_{0:k}^{(i)}, \mathbf{z}_{1:k}).
\end{aligned} \tag{2.52}$$

De même,

$$p(\mathbf{y}_k|\mathbf{z}_{1:k}) \simeq \sum_{i=1}^N w_k^{(i)} p(\mathbf{y}_k|\mathbf{r}_{0:k}^{(i)}, \mathbf{z}_{1:k}). \tag{2.53}$$

Des résultats théoriques issus du théorème de Rao-Blackwell confirment le gain apporté par l'utilisation de cette technique de marginalisation. En particulier, en comparant les deux estimateurs suivants de $\mathbb{E}[\mathbf{y}_{0:k}|\mathbf{z}_{1:k}]$, à savoir l'estimateur standard :

$$\hat{E}_1[\mathbf{y}_{0:k}|\mathbf{z}_{1:k}] = \frac{1}{N} \sum_{i=1}^N \mathbf{y}_{0:k}^{(i)} \quad \text{où} \quad (\mathbf{y}_{0:k}^{(i)}, \mathbf{r}_{0:k}^{(i)}) \sim p(\mathbf{y}_{0:k}, \mathbf{r}_{0:k}|\mathbf{z}_{1:k}) \tag{2.54}$$

et l'estimateur rao-blackwellisé :

$$\hat{E}_2[\mathbf{y}_{0:k}|\mathbf{z}_{1:k}] = \frac{1}{N} \sum_{i=1}^N \hat{E}[\mathbf{y}_{0:k}|\mathbf{r}_{0:k}^{(i)}, \mathbf{z}_{1:k}] \quad \text{où} \quad \mathbf{r}_{0:k}^{(i)} \sim p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k}) \tag{2.55}$$

alors, on peut montrer que :

$$\text{var}[\hat{E}_2[\mathbf{y}_{0:k}|\mathbf{z}_{1:k}]] \leq \text{var}[\hat{E}_1[\mathbf{y}_{0:k}|\mathbf{z}_{1:k}]], \tag{2.56}$$

ce qui démontre la meilleure efficacité de l'estimateur rao-blackwellisé. D'autres résultats théoriques sont donnés dans [Doucet 00b] et [Doucet 01b].

L'algorithme résultant de ce processus, appelé **filtre particulaire rao-blackwellisé** ou **filtre particulaire marginalisé** se décompose en trois étapes. Tout d'abord, une étape d'échantillonnage pondéré est réalisée pour générer le nuage pondéré $\{\mathbf{r}_{0:k}^{(i)}, w_k^{(i)}\}_{i=1\dots N}$ permettant d'estimer $p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k})$. Puis, pour chaque particule $\mathbf{r}_{0:k}^{(i)}$, $p(\mathbf{y}_{0:k}|\mathbf{r}_{0:k}^{(i)}, \mathbf{z}_{1:k})$ est estimée. Enfin, une étape de rééchantillonnage est effectuée.

Dans le cas d'un modèle conditionnellement linéaire gaussien comme suit [Chen 00a] [Doucet 00b] [LeGland 03] :

$$\begin{cases} \mathbf{r}_k = f_k(\mathbf{r}_{k-1}) + \mathbf{w}_k^{\mathbf{r}} \\ \mathbf{y}_k = A_k(\mathbf{r}_{k-1:k}) \mathbf{y}_{k-1} + \mathbf{w}_k^{\mathbf{y}} \\ \mathbf{z}_k = H_k(\mathbf{r}_k) \mathbf{y}_k + \mathbf{v}_k, \end{cases} \tag{2.57}$$

la loi conditionnelle $p(\mathbf{y}_k|\mathbf{r}_{0:k}, \mathbf{z}_{1:k})$ est gaussienne et décrite par ses deux premiers moments $\hat{\mathbf{y}}_{k|k}$ et $\Sigma_{k|k}$. Le filtre particulaire rao-blackwellisé est donc équivalent à un

banc de filtres de Kalman et permet l'estimation de $\mathbb{E}_{p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k})}[\phi(\mathbf{r}_{0:k})]$, $\mathbb{E}_{p(\mathbf{y}_k|\mathbf{z}_{1:k})}[\mathbf{y}_k]$ et $\mathbb{E}_{p(\mathbf{y}_k|\mathbf{z}_{1:k})}[\mathbf{y}_k \mathbf{y}_k^t]$. Il est résumé par l'algorithme 13. Il est évident que toutes les stratégies d'échantillonnage et de rééchantillonnage décrites précédemment peuvent être incluses dans cet algorithme.

Algorithme 13 Filtre particulière rao-blackwellisé pour le système (2.57)

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{r}_0^{(i)} \sim p(\mathbf{r}_0)$, fixer $w_0^{(i)} = 1/N$, $\hat{\mathbf{y}}_{0|0}^{(i)} = \hat{\mathbf{y}}_0$ et $\Sigma_{0|0}^{(i)} = \Sigma_0$

pour $k = 1, 2, \dots$

- **échantillonnage pondéré séquentiel :**

pour $i = 1 \dots N$, générer $\mathbf{r}_k^{(i)} \sim \pi(\mathbf{r}_k | \mathbf{r}_{k-1}^{(i)}, \mathbf{z}_k)$ et poser $\mathbf{r}_{0:k}^{(i)} \triangleq (\mathbf{r}_{0:k-1}^{(i)}, \mathbf{r}_k^{(i)})$

pour $i = 1 \dots N$, calculer les poids d'importance normalisés $\tilde{w}_k^{(i)}$

- **banc de filtres de Kalman :**

pour chaque particule $\mathbf{r}_{0:k}^{(i)}$, estimer les paramètres $\hat{\mathbf{y}}_{k|k}^{(i)}$ et $\Sigma_{k|k}^{(i)}$ de la loi gaussienne

$p(\mathbf{y}_k | \mathbf{r}_{0:k}^{(i)}, \mathbf{z}_{1:k})$ par le filtre de Kalman (algorithme 1)

- **estimations de Monte Carlo :**

$$\mathbb{E}_{p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k})}[\phi(\mathbf{r}_{0:k})] \simeq \sum_{i=1}^N \phi(\mathbf{r}_{0:k}^{(i)}) \tilde{w}_k^{(i)}$$

$$\mathbb{E}_{p(\mathbf{y}_k|\mathbf{z}_{1:k})}[\mathbf{y}_k] \simeq \sum_{i=1}^N \hat{\mathbf{y}}_{k|k}^{(i)} \tilde{w}_k^{(i)} \quad \mathbb{E}_{p(\mathbf{y}_k|\mathbf{z}_{1:k})}[\mathbf{y}_k \mathbf{y}_k^t] \simeq \sum_{i=1}^N \Sigma_{k|k}^{(i)} \tilde{w}_k^{(i)}$$

- **rééchantillonnage adaptatif :**

si nécessaire, sélectionner N ensembles $\{\mathbf{r}_{0:k}^{(j)}, \hat{\mathbf{y}}_{k|k}^{(j)}, \Sigma_{k|k}^{(j)}\}_{j=1 \dots N}$ parmi

$\{\mathbf{r}_{0:k}^{(i)}, \hat{\mathbf{y}}_{k|k}^{(i)}, \Sigma_{k|k}^{(i)}\}_{i=1 \dots N}$ en fonction des poids $\tilde{w}_k^{(i)}$

Il est possible d'appliquer cette méthode de marginalisation à d'autres modèles tels que les systèmes linéaires gaussiens partiellement observés [Andrieu 02] [LeGland 03], ou les modèles dont une partie de l'état constitue une chaîne de Markov à espace d'état fini [Doucet 00b]. Dans ce dernier cas, l'algorithme optimal utilisé est un filtre numérique par maillage d'état.

2.6 Synthèse

Depuis le début des années 90, l'utilisation des méthodes séquentielles de Monte Carlo pour l'analyse de systèmes non linéaires et/ou non gaussiens fait l'objet de recherches de plus en plus actives tant en statistique théorique que dans les domaines applicatifs. Ces méthodes apportent une nouvelle solution au problème de filtrage non linéaire dont les performances dépassent celles des algorithmes non aléatoires dans le cas général. L'idée consiste à mettre à jour récursivement une approximation discrète de la loi *a posteriori* constituée d'un nuage pondéré de particules. La mise à jour du nuage se fait par des tirages aléatoires, selon des équations induites par le modèle considéré.

Dans ce chapitre, une description théorique de ces méthodes a été présentée. Elles sont fondées sur l'utilisation répétée de l'échantillonnage pondéré. L'échantillonnage pondéré est une technique d'intégration de Monte Carlo qui permet la construction d'une estimation empirique d'une loi objectif $p(\cdot)$ à partir de tirages selon une loi d'importance $\pi(\cdot)$. L'estimation est d'autant plus précise que la loi d'importance est proche de la loi objectif.

Les méthodes séquentielles de Monte Carlo consistent à appliquer cette méthode au cadre bayésien où la loi objectif est $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$. Afin d'obtenir une solution séquentielle, la loi d'importance – qui permet de faire évoluer les particules dans l'espace d'état – est supposée avoir une forme récursive. Cependant, cette forme particulière introduit des erreurs du fait de l'écart avec la loi *a posteriori* objectif menant à une dégénérescence des particules. L'introduction d'une étape de rééchantillonnage permet de limiter cette dégénérescence en sélectionnant les particules de poids forts. Néanmoins, un rééchantillonnage systématique induit une perte de diversité des particules (problème d'appauvrissement des états), et donc une mauvaise exploration de l'espace d'état.

Une itération d'un algorithme générique de filtrage particulière correspond ainsi à une étape d'échantillonnage pondéré suivie si nécessaire d'une étape de rééchantillonnage. La décision de réaliser cette étape de sélection se fait à travers le calcul de la taille efficace du N-échantillon qui est une mesure de la dégénérescence des particules.

À partir de ce schéma générique, plusieurs choix de la loi d'importance et de techniques de rééchantillonnage ont été décrits. Chacun de ces choix vise à minimiser les problèmes de dégénérescence des particules ou d'appauvrissement des états. Deux points essentiels ressortent de cette étude. D'une part, il est important de choisir une méthode de sélection des particules qui minimise l'augmentation de la variance des poids. D'autre part, la loi d'importance doit prendre en compte l'état courant ainsi que la nouvelle observation disponible pour faire migrer les particules dans les zones de fortes vraisemblances.

Enfin, la dernière partie de ce chapitre a présenté une méthode de marginalisation, souvent appelée méthode de rao-blackwellisation. Cette technique s'applique aux modèles dont une partie de l'état peut être estimée par une méthode optimale. Le filtre particulière rao-blackwellisé mis en place permet de tirer parti de cette particularité du problème, de façon à réduire la partie de l'espace d'état où les particules seront vraiment nécessaires. L'augmentation des performances par rapport à un filtre particulière classique a été démontrée du point de vue théorique et pratique.

Chapitre 3

Filtrage mono-objet en environnement bruité : problème d'association de mesures

Dans les deux chapitres précédents, nous nous sommes intéressés à l'analyse du problème de filtrage. Plusieurs algorithmes ont été décrits. Face à un problème spécifique, le choix de la méthode utilisée dépend des caractéristiques du modèle. Le principe de l'ensemble de ces algorithmes est de construire une estimation de la loi *a posteriori*. Celle-ci est mise à jour récursivement par les équations du modèle dès qu'une nouvelle mesure est disponible. Cette nouvelle mesure est supposée générée par l'état caché.

En pratique, plusieurs mesures sont souvent disponibles à chaque itération. Chacune de ces mesures peut être issue de l'entité suivie ou avoir été générée par le processus d'élaboration des mesures sans provenir de l'objet. Ces dernières sont des fausses alarmes ; leur prise en compte dans la mise à jour de la loi de filtrage perturbe la qualité du résultat et peut amener à une perte de l'entité suivie. Il apparaît essentiel de distinguer, à chaque pas de temps, les mesures informatives sur l'état à estimer des fausses alarmes. C'est le problème d'*association de mesures*, aussi appelé *association de données* ("data association").

Le problème d'association de mesures est d'autant plus important dans le cas du filtrage multi-objets. Il est alors nécessaire de déterminer si une mesure est une fausse alarme, et dans le cas contraire de déterminer de quel objet elle est issue. De nombreux travaux ont été menés sur cette problématique. Pour un état de l'art détaillé, nous renvoyons à la thèse de doctorat en français [Hue 03], aux notes de cours [Dezert 01] et à l'ouvrage [BarShalom 95]. Dans le cadre de notre étude où on se concentre sur la trajectographie d'un objet unique dans une séquences d'images, le problème d'associations de mesures se résume à la détection des fausses alarmes. Par conséquent, nous ne détaillerons pas dans ce document les algorithmes dédiés à l'association de données multi-objets.

En se restreignant à la trajectographie d'une seule entité, une première solution consiste à ne prendre en compte qu'une seule observation. Celle-ci peut correspondre à la plus proche de la mesure prédite, ce qui mène à l'algorithme **NNF** ("Nearest Neighbor Filter") [Rong Li 96a] ou à celle d'intensité la plus forte, ce qui conduit à la méthode **SNF** ("Strongest Neighbor Filter") [Rong Li 97]. Des versions améliorées de ces filtres prennent

en compte la probabilité d'association correcte de la mesure choisie [Rong Li 96b] [Song 03] [Song 04]. Bien qu'attractives car nécessitant peu de ressources informatiques, ces solutions sont trop restrictives et trop sensibles. Dans l'hypothèse où la seule observation utilisée est une fausse alarme, ces algorithmes divergent rapidement.

D'autres méthodes ont été développées permettant de prendre en compte un ensemble restreint d'observations disponibles. Les mesures sélectionnées appartiennent à une *fenêtre de validation* ("gating"), construite comme étant une région de l'espace des mesures où la probabilité d'observer la nouvelle mesure sachant les précédentes est forte. Cette procédure est décrite dans la première section de ce chapitre. Ensuite, un modèle de fausses alarmes (décrit dans le paragraphe 3.2) amène à exprimer la loi de filtrage comme un mélange de lois, où les différentes associations de mesures possibles sont modélisées. Cette loi peut alors être estimée récursivement par le **filtre à association de données probabiliste** ("probabilistic data association filter") présenté en partie 3.3. Enfin, l'émergence des méthodes particulières a conduit à de nouveaux développements pour la résolution conjointe du filtrage non linéaire et de l'association de données. Ces derniers font l'objet de la partie 3.4.

Remarque : Les méthodes présentées dans ce chapitre ont été développées dans la communauté du traitement du signal radar, où la recherche est active sur ce problème [BarShalom 88] [BarShalom 95]. En vision par ordinateur, ce dernier est rarement discuté, bien que la présence d'un arrière plan dans les images constitue une zone de forte ambiguïté et puisse générer de nombreuses fausses mesures.

3.1 Fenêtre de validation des mesures

Comme décrit en introduction, la première étape à réaliser est la définition d'une zone de validation des mesures. Cette zone correspond à une région de l'espace des mesures où l'observation future sera trouvée avec une forte probabilité. De telles régions sont généralement utilisées dans des problèmes de suivi en traitement du signal radar pour réduire le nombre de fausses alarmes [BarShalom 88]. En trajectographie dans des séquences d'images, ces régions définissent une partie de l'image où chercher l'observation future. À chaque instant, une nouvelle fenêtre de validation est définie en utilisant la distribution de probabilité de la nouvelle mesure conditionnellement aux observations passées $p(\mathbf{z}_k | \mathbf{z}_{1:k-1})$. Une observation sera validée si elle appartient à cette région.

Modèle linéaire gaussien

Dans le cas d'un modèle linéaire gaussien,

$$\begin{cases} \mathbf{x}_k | \mathbf{x}_{k-1} \sim \mathcal{N}(\mathbf{x}_k; F_k \mathbf{x}_{k-1}, Q_k) \\ \mathbf{z}_k | \mathbf{x}_k \sim \mathcal{N}(\mathbf{z}_k; H_k \mathbf{x}_k, R_k), \end{cases} \quad (3.1)$$

on a le résultat suivant :

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) d\mathbf{x}_k = \mathcal{N}(\mathbf{z}_k; \mathbf{z}_{k|k-1}, S_k) \quad (3.2)$$

$$\text{où } \mathbf{z}_{k|k-1} = H_k \hat{\mathbf{x}}_{k|k-1} \quad \text{et} \quad S_k = H_k \Sigma_{k|k-1} H_k^t + R_k. \quad (3.3)$$

On rappelle que $\hat{\mathbf{x}}_{k|k-1}$ et $\Sigma_{k|k-1}$ correspondent aux deux premiers moments de la densité prédite gaussienne $p(\mathbf{x}_k | \mathbf{z}_{1:k-1})$, dont les expressions sont données par le filtre de Kalman (cf. paragraphe 1.1 p.21). $\mathbf{z}_{k|k-1}$ désigne la mesure prédite et S_k sa matrice de covariance (S_k est aussi la matrice de covariance de l'innovation $\tilde{\mathbf{z}}_k = \mathbf{z}_k - \mathbf{z}_{k|k-1}$).

Dans le cas linéaire gaussien, la fenêtre de validation $valid_k$ est donc l'ellipsoïde de dimension n_z (dimension du vecteur de mesure) définie telle que :

$$valid_k = \{\mathbf{z}_k : \epsilon_k = (\mathbf{z}_k - \mathbf{z}_{k|k-1})^t S_k^{-1} (\mathbf{z}_k - \mathbf{z}_{k|k-1}) \leq \gamma\}. \quad (3.4)$$

Le seuil γ est choisi de telle sorte que la probabilité de trouver la mesure correcte dans la région $valid_k$ soit égale à une valeur fixée P_g . La distance ϵ_k suit une loi du chi-deux, à n_z degrés de liberté, ce qui permet de calculer γ en fonction de P_g et réciproquement (en pratique, P_g est souvent fixé à 0.99 [Dezert 01]).

Le volume V de l'ellipsoïde de validation des mesures est donné par :

$$V_k = C_{n_z} \gamma^{n_z/2} |S_k|^{1/2}, \quad (3.5)$$

où C_{n_z} est une constante qui représente le volume de l'hypersphère de rayon 1, de dimension n_z (pour $n_z = 2$, $C_{n_z} = \pi$).

Modèle non linéaire

Dans le cas non linéaire, la densité $p(\mathbf{z}_k | \mathbf{z}_{1:k-1})$ n'est pas connue analytiquement. Cependant, si l'on dispose d'une approximation discrète de $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ sous forme d'un nuage de particules pondérées $\{\mathbf{x}_{k-1}^{(i)}, \tilde{w}_{k-1}^{(i)}\}_{i=1\dots N}$ (obtenu par exemple par un algorithme de filtrage particulaire), il est possible d'approcher la densité de vraisemblance de la nouvelle observation en remarquant que [Doucet 00b] :

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) d\mathbf{x}_k \quad (3.6)$$

$$= \int p(\mathbf{z}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}. \quad (3.7)$$

À partir de l'expression (3.6), l'estimée suivante peut être construite, où les $\{\mathbf{x}_k^{*(i)}\}_{i=1\dots N}$ sont des échantillons de $p(\mathbf{x}_k | \mathbf{z}_{1:k-1})$ (obtenus par prédiction à partir du nuage approchant $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$) :

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}) \simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} p(\mathbf{z}_k | \mathbf{x}_k^{*(i)}). \quad (3.8)$$

Si une expression analytique de $p(\mathbf{z}_k | \mathbf{x}_{k-1})$ est connue, l'utilisation de l'expression (3.7) permet de considérer l'estimation suivante à partir du nuage approchant $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$:

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}) \simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}). \quad (3.9)$$

À partir de ces expressions, il est possible de calculer analytiquement ou de façon empirique les deux premiers moments des estimations de $p(\mathbf{z}_k|\mathbf{z}_{1:k-1})$, et de construire une fenêtre ellipsoïdale de validation des mesures. Une forme ellipsoïdale suppose une approximation gaussienne de la densité $p(\mathbf{z}_k|\mathbf{z}_{1:k-1})$.

Une autre solution consiste à construire numériquement ces régions de prédiction de la mesure à partir de simulations de $p(\mathbf{z}_k|\mathbf{z}_{1:k-1})$ ou d'une approximation de celle-ci [Breidt 00]. Bien que captant toutes les non linéarités et le caractère non gaussien du modèle, la construction de ces régions est très gourmande en temps de calcul.

3.2 Hypothèses d'association et modèle de fausses alarmes

L'utilisation d'une fenêtre de validation des mesures permet de classer certaines observations – celles n'appartenant pas à cette région – comme étant des fausses alarmes. Dans le cas où une seule mesure est sélectionnée, les algorithmes présentés dans les deux chapitres précédents peuvent être appliqués directement. Dans le cas contraire, une modélisation probabiliste des différentes associations possibles peut être construite. Celle-ci est fondée sur des hypothèses d'association et des modèles de fausses alarmes décrits dans ce paragraphe.

Hypothèses d'association

Plusieurs hypothèses d'association peuvent être considérées [Hue 03]. Elles définissent des règles permettant d'associer les observations validées à l'entité suivie ou à des fausses alarmes.

H1 : Une mesure peut provenir d'un objet ou être une fausse alarme.

H2 : À un instant donné, un objet peut être à l'origine d'aucune ou d'au plus une mesure.

H3 : À un instant donné, un objet peut être à l'origine d'aucune ou de plusieurs mesures.

Dans le cas du filtrage mono-objet, les solutions proposées utilisent les deux premières hypothèses. Cela induit la dépendance des associations des différentes observations. Cette restriction a du sens dans le cadre applicatif de la vision par ordinateur. En effet, un objet ne peut pas correspondre à deux parties différentes d'une image de la scène, ce qui confirme le choix de l'hypothèse **H2** par rapport à **H3**. Ces hypothèses permettent de considérer le cas où aucune observation n'est générée par l'état caché. Dans des applications de vision par ordinateur, cela correspond à des situations d'occlusions ou de très forts bruits d'image. Nous verrons dans la deuxième partie de ce document qu'il est possible de ne pas autoriser cette situation sous certaines conditions.

Au temps k , on note $\mathbf{z}_k$ le vecteur de mesures constitué des M_k observations validées $\{\mathbf{z}_{k,m}\}_{m=1\dots M_k}$. En se limitant aux hypothèses **H1** et **H2**, on note $\Phi_k \in \{0, \dots, M_k\}$ la variable aléatoire décrivant les événements d'associations :

$$\Phi_k = \begin{cases} j & \text{si } \mathbf{z}_{k,j} \text{ est générée par l'état caché,} \\ 0 & \text{si toutes les observations sont des fausses alarmes.} \end{cases} \quad (3.10)$$

Au niveau théorique, la combinaison des hypothèses **H1** et **H2** traduit le fait que les associations possibles décrites par la variable Φ_k sont exhaustives et mutuellement exclusives. Cela permet l'utilisation du théorème des probabilités totales pour exprimer la loi de filtrage recherchée comme une loi mélange :

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \sum_{j=0}^{M_k} p(\mathbf{x}_k | \Phi_k = j, \mathbf{z}_{1:k}) p(\Phi_k = j | \mathbf{z}_{1:k}) \quad (3.11)$$

où $p(\Phi_k = j | \mathbf{z}_{1:k})$ désignent les probabilités d'association *a posteriori*.

Modèle de fausses alarmes

Un modèle de fausses alarmes a été initialement défini par Bar-Shalom, qui en propose une description détaillée dans le livre [BarShalom 95]. Ce modèle est largement utilisé dans les applications de signal radar [Hue 03] et a été appliqué en trajectographie dans des séquences d'images [Isard 98a] [Hue 03].

Ce modèle est le suivant : on suppose que le nombre de fausses alarmes suit une loi de Poisson de paramètre λV_k où V_k correspond au volume de la région d'observation (si une fenêtre de validation des mesures est utilisée, V correspond au volume de cette région) et λ est un nombre réel strictement positif (cf. [BarShalom 95] pour plus de détails sur ce paramètre). En notant $\mu(l)$ la probabilité d'observer l fausses alarmes, on a :

$$\mu(l) = \exp(-\lambda V_k) \frac{(\lambda V_k)^l}{l!}. \quad (3.12)$$

Enfin, en faisant une analogie entre les fausses alarmes et du bruit homogène, celles-ci sont supposées uniformément réparties dans la fenêtre de validation et indépendantes dans le temps. Enfin, on suppose généralement qu'une mesure qui n'est pas une fausse alarme est liée à l'état par une densité gaussienne. Cela se traduit par l'équation suivante :

$$p(\mathbf{z}_{k,m} | \Phi_k = j, \mathbf{x}_k) = \begin{cases} \mathcal{N}(\mathbf{z}_{k,m}; h_k(\mathbf{x}_k), R_k) & j = m \\ V_k^{-1} & j = 0 \dots M_k ; j \neq m \end{cases} \quad (3.13)$$

Remarque : En pratique, le choix du paramètre λ peut être difficile. Dans le cas où aucune information n'est disponible, une alternative simple au modèle de Poisson est le modèle uniforme. On considère alors que la probabilité d'avoir l fausses alarmes est constante et ne dépend pas de l . Ce modèle est peu répandu.

3.3 Modèle linéaire gaussien : filtre à association de données probabiliste

Le **filtre à association de données probabiliste** (PDAF : “probabilistic data association filter”) [BarShalom 75] [BarShalom 95] [Kirubarajan 04] est une extension du filtre de Kalman dédiée au filtrage d'un objet unique évoluant dans un environnement bruité. Cet algorithme se décompose donc en deux étapes, prédiction et correction. La prise en compte des différentes hypothèses d'association est faite au travers du calcul de

l'innovation.

On se place dans le cadre linéaire gaussien :

$$\begin{cases} \mathbf{x}_k = F_k \mathbf{x}_{k-1} + \mathbf{w}_k \\ \mathbf{z}_{k,m} = H_k \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad \text{si } \mathbf{z}_{k,m} \text{ n'est pas une fausse alarme} \quad (3.14)$$

où $\mathbf{w}_k$ et $\mathbf{v}_k$ sont des bruits blancs gaussiens indépendants entre eux de matrices de covariance respectives Q_k et R_k . Pour simplifier la résolution du problème, l'information passée décrite par la loi de filtrage au temps $k-1$ est résumée par une loi gaussienne :

$$p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k-1|k-1}, \Sigma_{k-1|k-1}). \quad (3.15)$$

Résumer le passé aux seules informations disponibles à l'instant précédent représente bien une approximation puisque l'état du filtre à l'instant k dépend aussi des hypothèses d'associations formulées dans le passé. Cette formulation revient à approcher une somme de densités gaussiennes par une unique densité gaussienne.

Étape de prédiction

À partir de l'expression (3.15), la loi prédite est estimée de façon similaire au filtre de Kalman classique (c.f. paragraphe 1.1, p. 21) :

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \Sigma_{k|k-1}), \quad (3.16)$$

où les deux premiers moments sont exprimés par :

$$\hat{\mathbf{x}}_{k|k-1} = F_k \hat{\mathbf{x}}_{k-1|k-1} \quad (3.17)$$

$$\Sigma_{k|k-1} = F_k \Sigma_{k-1|k-1} F_k^t + Q_k. \quad (3.18)$$

Étape de correction

La difficulté vient de l'estimation des moments de la loi corrigée

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \Sigma_{k|k}). \quad (3.19)$$

Compte tenu du caractère exclusif et exhaustif des hypothèses d'association, sur le même principe que (3.11), l'estimateur optimal (au sens de la minimisation de la variance d'erreur d'estimation) s'écrit :

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}] = \sum_{j=0}^{M_k} p(\Phi_k = j | \mathbf{z}_{1:k}) \mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}, \Phi_k = j] \\ &= \sum_{j=0}^{M_k} \beta_{k,j} \mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}, \Phi_k = j]. \end{aligned} \quad (3.20)$$

$$\text{où } \beta_{k,j} \triangleq p(\Phi_k = j | \mathbf{z}_{1:k}) \quad j = 0 \dots M_k \quad (3.21)$$

Par analogie avec le filtre de Kalman, on peut remarquer que dans le cas où l'ensemble des observations sont des fausses alarmes, on a :

$$\mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}, \Phi_k = 0] = \hat{\mathbf{x}}_{k|k-1}, \quad (3.22)$$

et que l'estimée de l'état, conditionnellement au fait que la j ème mesure soit issue de l'état caché, s'écrit :

$$\mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}, \Phi_k = j] = \hat{\mathbf{x}}_{k|k-1} + K_k \tilde{\mathbf{z}}_{k,j} \quad \forall j = 1 \dots M_k. \quad (3.23)$$

Le gain K_k a la même expression que dans le cas classique du filtre de Kalman et $\tilde{\mathbf{z}}_{k,j} = \mathbf{z}_{k,j} - H_k \mathbf{x}_k$ désigne l'innovation associée à la mesure $\mathbf{z}_{k,j}$.

En remplaçant (3.22) et (3.23) dans (3.20), l'équation de mise à jour de l'état devient :

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k \sum_{j=1}^{M_k} \beta_{k,j} \tilde{\mathbf{z}}_{k,j}, \quad (3.24)$$

et la mise à jour de sa covariance associée se fait selon l'équation suivante (la démonstration, non triviale, est décrite dans [BarShalom 95]) :

$$\Sigma_{k|k} = \beta_{k,0} \Sigma_{k|k-1} + (1 - \beta_{k,0}) \Sigma_k^c + \tilde{\Sigma}_k. \quad (3.25)$$

La covariance de l'état corrigée par la mesure correcte est donnée par :

$$\Sigma_k^c = (Id - K_k H_k) \Sigma_{k|k-1}. \quad (3.26)$$

L'incertitude sur les mesures se traduit par une augmentation de la covariance d'estimation à travers la matrice semi définie positive $\tilde{\Sigma}_k$ dont l'expression est :

$$\tilde{\Sigma}_k = K_k \left[\sum_{j=1}^{M_k} \beta_{k,j} \tilde{\mathbf{z}}_{k,j} \tilde{\mathbf{z}}_{k,j}^t - \tilde{\mathbf{z}}_k \tilde{\mathbf{z}}_k^t \right] K_k^t. \quad (3.27)$$

On rappelle les expressions suivantes :

$$K_k = \Sigma_{k|k-1} H_k^t (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1} \quad (3.28)$$

$$\tilde{\mathbf{z}}_{k,j} = \mathbf{z}_{k,j} - H_k \mathbf{x}_k \quad (3.29)$$

$$\tilde{\mathbf{z}}_k = \sum_{j=1}^{M_k} \beta_{k,j} \tilde{\mathbf{z}}_{k,j}. \quad (3.30)$$

Calcul des probabilités d'association

Pour évaluer les probabilités *a posteriori* d'associations $\{\beta_{k,j}\}_{j=0 \dots M_k}$, le modèle de fausses alarmes décrit précédemment est utilisé. On a :

$$\beta_{k,j} \triangleq p(\Phi_k = j | \mathbf{z}_{1:k}) = p(\Phi_k = j | M_k, \mathbf{z}_{1:k}). \quad (3.31)$$

En utilisant la règle de Bayes, et l'indépendance conditionnelle des observations, cette expression devient :

$$\begin{aligned} \beta_{k,j} &= C p(\Phi_k = j | M_k, \mathbf{z}_{1:k-1}) p(\mathbf{z}_k | \Phi_k = j, M_k, \mathbf{z}_{1:k-1}) \\ &= C p(\Phi_k = j | M_k, \mathbf{z}_{1:k-1}) \prod_{m=1}^{M_k} p(\mathbf{z}_{k,m} | \Phi_k = j, M_k, \mathbf{z}_{1:k-1}) \end{aligned} \quad (3.32)$$

où C est une constante de normalisation telle que $\sum_{j=0}^{M_k} \beta_{k,j} = 1$. En utilisant le modèle décrit dans le paragraphe 3.2, et en supposant que la densité de la mesure correcte est gaussienne, centrée en la mesure prédite $\mathbf{z}_{k|k-1}$ de covariance S_k (hypothèse déduite de (3.15)), la densité jointe de l'ensemble des observations validées conditionnellement à $\Phi_k = j$ et aux mesures passées s'écrit :

$$p(\mathbf{z}_k | \Phi_k = j, M_k, \mathbf{z}_{1:k-1}) = \begin{cases} V_k^{-M_k+1} P_g^{-1} \mathcal{N}(\tilde{\mathbf{z}}_{k,j}; 0, S_k) & j = 1 \dots M_k \\ V_k^{-M_k} & j = 0 \end{cases} \quad (3.33)$$

Enfin, les probabilités d'association *a priori* s'écrivent (c.f. [BarShalom 95]) :

$$p(\Phi_k = j | M_k, \mathbf{z}_{1:k-1}) = p(\Phi_k = j | M_k) = \begin{cases} \frac{1}{D} \frac{P_d P_g}{M_k} & j = 1 \dots M_k \\ \frac{1}{D} (1 - P_d P_g) \frac{\mu(M_k)}{\mu(M_k - 1)} & j = 0 \end{cases} \quad (3.34)$$

avec

$$D = P_d P_g + (1 - P_d P_g) \frac{\mu(M_k)}{\mu(M_k - 1)}. \quad (3.35)$$

$\mu(\cdot)$ est la probabilité du nombre de fausses alarmes validées. P_d est la probabilité d'avoir détecté la mesure correcte et P_g la probabilité de fenêtrage choisie pour le test de validation des mesures.

Le filtre à association de données probabiliste en résumé

Le PDAF est résumé par l'algorithme 14. Il est dédié au problème de la trajectographie d'une entité évoluant dans un environnement bruité, modélisé par un système linéaire gaussien. Cette méthode est une extension du filtre de Kalman permettant de résoudre conjointement le problème de filtrage linéaire gaussien et le problème d'association de mesures.

Ses principaux avantages sont sa facilité de mise en œuvre et sa faible complexité (à peine supérieure à celle du Kalman classique). Cependant, il repose sur l'hypothèse forte (3.15) qui suppose qu'à chaque pas de temps la loi de filtrage est bien représentée par une unique loi gaussienne. Elle est en réalité égale à une somme de gaussiennes. Le PDAF est mis en défaut lorsque cette approximation est mauvaise, en particulier dans des situations riches en fausses alarmes.

Filtre bayésien optimal

De part l'hypothèse (3.15) qui résume la loi de filtrage par une unique loi gaussienne, le PDAF est un filtre sous-optimal. La version exacte de cet algorithme est appelée **filtre bayésien optimal** [Singer 74]. La différence essentielle réside dans le fait que la décomposition de l'estimée de l'état $\hat{\mathbf{x}}_{k|k}$ est faite sur l'ensemble des associations possibles du temps initial au temps k et non uniquement sur les associations possibles des mesures nouvellement disponibles (cf. équation (3.20)). La probabilité d'association de la séquence d'observations $\mathbf{z}_{1:k}$ doit être calculée. Dans le cas linéaire gaussien, le filtre bayésien optimal estime de façon exacte la densité de filtrage qui est un mélange de gaussiennes. Une description détaillée de cet algorithme est disponible dans [BarShalom 95], [Dezert 01], [Hue 03].

Algorithme 14 Filtre à association de données probabiliste

- **initialisation** : $\hat{\mathbf{x}}_{0|0} = \hat{\mathbf{x}}_0$, $\Sigma_{0|0} = \Sigma_0$

pour $k = 1, 2, \dots$

- **prédiction** :

$$\begin{aligned}
 \hat{\mathbf{x}}_{k|k-1} &= F_k \hat{\mathbf{x}}_{k-1|k-1} \\
 \Sigma_{k|k-1} &= F_k \Sigma_{k-1|k-1} F_k^t + Q_k \\
 \mathbf{z}_{k|k-1} &= H_k \mathbf{x}_{k|k-1} \\
 S_k &= H_k \Sigma_{k|k-1} H_k^t + R_k
 \end{aligned} \tag{3.36}$$

- **association** :

1. Calcul de l'ellipse de validation de volume V_k à partir de $p(\mathbf{z}_k | \mathbf{z}_{1:k-1}) = \mathcal{N}(\mathbf{z}_k; \mathbf{z}_{k|k-1}, S_k)$
2. Sélection des M_k mesures contenues dans l'ellipse de validation
3. Calcul des probabilités d'association $\{\beta_{k,j}\}_{j=0\dots M_k}$ avec (3.32), (3.33) et (3.34)

- **correction** :

$$\begin{aligned}
 \tilde{\mathbf{z}}_{k,j} &= \mathbf{z}_{k,j} - H_k \mathbf{x}_k \\
 \tilde{\mathbf{z}}_k &= \sum_{j=1}^{M_k} \beta_{k,j} \tilde{\mathbf{z}}_{k,j} \\
 K_k &= \Sigma_{k|k-1} H_k^t (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1} \\
 \Sigma_k^c &= (Id - K_k H_k) \Sigma_{k|k-1} \\
 \tilde{\Sigma}_k &= K_k [\sum_{j=1}^{M_k} \beta_{k,j} \tilde{\mathbf{z}}_{k,j} \tilde{\mathbf{z}}_{k,j}^t - \tilde{\mathbf{z}}_k \tilde{\mathbf{z}}_k^t] K_k^t \\
 \Sigma_{k|k} &= \beta_{k,0} \Sigma_{k|k-1} + (1 - \beta_{k,0}) \Sigma_k^c + \tilde{\Sigma}_k \\
 \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + K_k \sum_{j=1}^{M_k} \beta_{k,j} \tilde{\mathbf{z}}_{k,j}
 \end{aligned} \tag{3.37}$$

Bien qu'optimal du point de vue théorique, ce filtre est inutilisable en pratique car le nombre de modes de la loi estimée croît exponentiellement avec le temps. La mémoire et la capacité de calcul nécessaires à sa mise en œuvre augmentent donc de manière exponentielle. Il est possible d'envisager de nouveaux algorithmes limitant le nombre d'associations possibles à gérer. C'est le cas de l'algorithme **track split filter** [Smith 95] qui consiste à éliminer les séquences les moins vraisemblables, ou des techniques de réduction de mélange [Salmond 90]. Enfin, des filtres à mémoire limitée peuvent être construits comme intermédiaires entre le filtre bayésien optimal et le filtre à association de données probabiliste [Singer 74].

Remarque : D'autres algorithmes utilisables dans le cadre filtrage mono-objet en environnement bruité ont été développés, tels que le **filtre probabiliste à hypothèses**

multiples (PMHT : “probabilistic multiple hypothesis tracker”) [Streit 94] ou des extensions du PDAF [BarShalom 95]. Nous ne les détaillerons pas ici car ils s’appliquent au suivi d’objet manœuvrant qui est une situation très rarement rencontrée en trajectographie dans des séquences d’images. Notons juste que contrairement au PDAF, le PMHT utilise les hypothèses d’association **H1** et **H3**.

3.4 Modèle non linéaire : utilisation de filtres particuliers

Face à un modèle non linéaire, il est possible d’utiliser une version étendue du filtre à association de données probabiliste. Cependant, cet algorithme souffre des mêmes inconvénients que le filtre de Kalman étendu (cf. paragraphe 1.2.1, p. 25). Il est alors préférable d’utiliser les méthodes de filtrage particulière. Ces méthodes permettent de capter les non linéarités du système à travers le nuage de particules sans nécessiter d’approximations. Approchant la loi complète $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$, le nuage de particules transporte également les informations sur les associations passées.

Les difficultés essentielles de l’utilisation des méthodes séquentielles de Monte Carlo pour la résolution du problème conjoint de filtrage et d’association des données réside en :

1. la définition d’une fenêtre de validation à travers une estimation de $p(\mathbf{z}_k|\mathbf{z}_{1:k-1})$ servant à la validation des mesures,
2. la modélisation de la vraisemblance $p(\mathbf{z}_k|\mathbf{x}_k)$ où $\mathbf{z}_k = (\mathbf{z}_{k,1}, \dots, \mathbf{z}_{k,M_k})$ désigne l’ensemble des observations validées.

Une fois une approximation de ces deux densités disponibles, il est possible d’utiliser les algorithmes présentés dans le chapitre 2, en particulier le filtre bootstrap.

Le premier point a été discuté dans la partie 3.1. En ce qui concerne la modélisation de la vraisemblance, sous les hypothèses **H1** et **H2**, le théorème des probabilités totales par rapport à la variable d’associations Φ_k peut être appliqué de façon similaire à l’équation (3.11) :

$$\begin{aligned} p(\mathbf{z}_k|\mathbf{x}_k) &= \sum_{j=0}^{M_k} p(\Phi_k = j|\mathbf{x}_k) p(\mathbf{z}_k|\Phi_k = j, \mathbf{x}_k) \\ &= \sum_{j=0}^{M_k} \beta_{k,j} p(\mathbf{z}_k|\Phi_k = j, \mathbf{x}_k) \end{aligned} \tag{3.38}$$

$$\text{où } \beta_{k,j} \triangleq p(\Phi_k = j|\mathbf{x}_k) \quad j = 0 \dots M_k \tag{3.39}$$

Les deux termes mis en jeu dans cette expression ne peuvent pas être évalués de façon directe. Plusieurs approches ont été utilisées, que nous détaillons ici.

Estimation de $p(\mathbf{z}_k|\Phi_k = j, \mathbf{x}_k)$

En supposant l'indépendance conditionnelle des mesures validées, l'expression de $p(\mathbf{z}_k|\Phi_k = j, \mathbf{x}_k)$ est :

$$p(\mathbf{z}_k|\Phi_k = j, \mathbf{x}_k) = \prod_{m=1}^{M_k} p(\mathbf{z}_{k,m}|\Phi_k = j, \mathbf{x}_k). \quad (3.40)$$

Par l'équation (3.13) décrivant la vraisemblance conditionnelle des mesures, que nous rappelons ici :

$$p(\mathbf{z}_{k,m}|\Phi_k = j, \mathbf{x}_k) = \begin{cases} \mathcal{N}(\mathbf{z}_{k,m}; h_k(\mathbf{x}_k), R_k) & j = m \\ V_k^{-1} & j = 0 \dots M_k ; j \neq m, \end{cases} \quad (3.41)$$

la vraisemblance s'exprime comme un mélange de lois uniformes et gaussiennes :

$$\begin{aligned} p(\mathbf{z}_k|\mathbf{x}_k) &= \sum_{j=0}^{M_k} [p(\Phi_k = j|\mathbf{x}_k) \prod_{m=1}^{M_k} p(\mathbf{z}_{k,m}|\Phi_k = j, \mathbf{x}_k)] \\ &= p(\Phi_k = 0|\mathbf{x}_k) V_k^{-M_k} + \sum_{j=1}^{M_k} [p(\Phi_k = j|\mathbf{x}_k) V_k^{1-M_k} \mathcal{N}(\mathbf{z}_{k,j}; h_k(\mathbf{x}_k), R_k)] \end{aligned} \quad (3.42)$$

Cette formulation a été introduite dans [Gordon 97], puis reprise dans des applications de pistage radar [Marrs 02], de suivi dans des séquences d'images [Isard 98a] et de robotique [Schulz 01].

Estimation de $\beta_{k,j} \triangleq p(\Phi_k = j|\mathbf{x}_k)$

Du fait de la non connaissance du processus $\mathbf{x}_k$, l'estimation directe de $p(\Phi_k = j|\mathbf{x}_k)$ est impossible. Une première solution consiste à supposer l'indépendance des variables Φ_k et $\mathbf{x}_k$. On approche alors $p(\Phi_k = j|\mathbf{x}_k)$ par la loi *a priori* $p(\Phi_k = j)$ [Gordon 97]. Le cas le plus simple suppose l'équi-probabilité des associations, ce qui conduit à l'estimation suivante ($\beta_{k,0}$ est un paramètre à fixer) [Isard 98a] :

$$\hat{\beta}_{k,j} = \frac{1 - \beta_{k,0}}{M_k} \quad \forall j = 1 \dots M_k \quad (3.43)$$

Cependant, il peut être bénéfique d'introduire des informations disponibles sur le problème au travers de cette probabilité d'association. L'idée est alors d'utiliser les observations pour l'estimation de $\beta_{k,j}$:

$$\hat{\beta}_{k,j} = p(\Phi_k = j|\mathbf{z}_{1:k}) \quad \forall j = 0 \dots M_k \quad (3.44)$$

Si les hypothèses sur le modèle de fausses alarmes et les hypothèses d'associations décrites en 3.2 sont conservées, on retrouve l'expression de $p(\Phi_k = j|\mathbf{z}_{1:k})$ utilisée dans le PDAF (équation 3.32) [Marrs 02].

Remarque : Dans cette section, une vraisemblance permettant de modéliser les associations de mesures possibles a été proposée. Celle-ci suppose que la “vraie” mesure est liée à l’état cachée par une loi gaussienne. Cette modélisation n’est donc plus valable dans le cas où le bruit de mesure est non gaussien. Ce problème reste ouvert et constitue une perspective intéressante. De plus, à cause du terme de loi uniforme additionné au terme de mélange de lois gaussiennes, cette formulation empêche l’utilisation de la loi de proposition optimale. Nous verrons dans la partie II que cette loi optimale peut être utilisée, même dans un cadre de suivi en environnement bruité.

3.5 Synthèse

Dans ce chapitre, nous nous sommes intéressés au problème d’association de données dans le cadre du suivi mono-objet. Ce problème apparaît lorsque plusieurs mesures sont disponibles à chaque itération. Chacune de ces mesures est soit issue de l’état caché, soit une fausse alarme pouvant perturber l’estimation.

Une première étape consiste à valider ou non ces données en fonction de la densité de la nouvelle observation conditionnellement aux mesures passées. Dans le cas où une unique mesure est conservée, les algorithmes présentés dans les deux chapitres précédents peuvent être appliqués directement. Dans le cas contraire, l’introduction d’une variable d’association permet de modéliser l’incertitude sur l’origine des mesures validées. Sous l’hypothèse qu’un objet peut être à l’origine d’au plus une observation, l’introduction de cette variable permet d’exprimer la loi de filtrage recherchée comme une loi mélange dont les coefficients sont les probabilités d’association *a posteriori*. Ces probabilités peuvent être calculées en utilisant un modèle de fausses alarmes, qui associent les mauvaises mesures à du bruit homogène.

Dans le cas d’un modèle linéaire gaussien, l’estimation récursive de la loi de filtrage ainsi construite est donnée par le PDAF. Cet algorithme est une extension du filtre de Kalman associant les étapes de prédiction / correction à une étape d’association de mesures. En pratique, il donne de meilleurs résultats que les filtres qui ne considèrent qu’une unique observation. Cependant, il repose sur l’hypothèse forte que la loi de filtrage au temps précédent peut être résumée par une loi gaussienne. Une partie de l’information sur les associations passées est alors perdue, ce qui met le filtre en défaut dans des situations fortement bruitées.

Face à un système non linéaire, l’utilisation des méthodes de filtrage particulière permet d’éviter les approximations que pourraient introduire une version étendue du PDAF. Ces méthodes ont l’avantage de conserver les informations sur les associations passées, résumées par le nuage de particules. Leur mise en œuvre repose sur une modélisation de la vraisemblance qui utilise les variables d’association et le modèle de fausses alarmes. Ces nouvelles approches, basées sur des méthodes de simulation de Monte Carlo, permettent la résolution conjointe du problème de filtrage non linéaire et l’association de données. Elles ont augmenté significativement la qualité de l’estimation par rapport aux filtres classiques.

Conclusion

L'estimation d'un système dynamique à partir d'observations bruitées est un problème difficile. Son implication dans de nombreux domaines en fait un sujet de recherche actif. Lorsque les données sont acquises de manière séquentielle, l'emploi d'algorithmes itératifs pour l'estimation de la loi *a posteriori* de l'état conditionnellement aux mesures passées est nécessaire.

Le filtre bayésien optimal fournit une expression récursive exacte de cette loi. En pratique, cette expression n'est calculable que pour un nombre restreint de modèles dont font partie les systèmes linéaires à bruits additifs gaussiens. Dans ce cas, la loi de filtrage est gaussienne et le calcul explicite des équations du filtre optimal conduit au filtre de Kalman. Face à un système non linéaire non gaussien, l'utilisation des méthodes non aléatoires existantes telles que les extensions du filtre de Kalman et les méthodes numériques par maillage d'état est rarement satisfaisante. Les problèmes qui apparaissent sont soit une divergence de l'estimation due à des approximations trop fortes, soit un coût de calcul prohibitif.

L'application de méthodes d'intégration de Monte Carlo au cadre bayésien a conduit aux développements des méthodes séquentielles de Monte Carlo. Ces algorithmes apportent une nouvelle solution au problème de filtrage non linéaire dont les performances dépassent celles des algorithmes non aléatoires dans le cas général. Leur principe consiste à approcher la loi *a posteriori* par une somme pondérée de mesures de Dirac situées en des positions de l'espace d'état appelées particules. La mise à jour du nuage se fait récursivement en deux étapes : une étape d'échantillonnage pondéré faisant évoluer le nuage, suivie éventuellement d'une étape de rééchantillonnage permettant la sélection des particules pertinentes. Les filtres particuliers présentés diffèrent selon le choix de la fonction de proposition utilisée pour échantillonner les particules et selon la méthode de rééchantillonnage mise en œuvre.

Outre les non-linéarités du système, la résolution du problème de filtrage peut être perturbée par la présence de fausses alarmes, qui correspondent à des mesures observées non issues de l'entité suivie. Une phase d'association des données doit alors être mise en place pour lever l'ambiguïté sur l'origine des mesures. Cette incertitude peut être modélisée par l'introduction d'une variable d'association et d'un modèle de fausses alarmes. De nouveaux algorithmes ont ainsi été développés pour la résolution conjointe du problème de filtrage et de l'association de données. Face à un modèle linéaire gaussien, une extension du filtre de Kalman appelée filtre à association de données probabilistes est la méthode la plus courante. Dans le cas non linéaire, les méthodes séquentielles de Monte Carlo peuvent être appliquées. La modélisation du problème d'association des mesures se fait alors au travers de la vraisemblance.

Face à une application donnée, le choix de l'algorithme dépend des caractéristiques du modèle construit (linéaire ou non, gaussien ou non, présence de fausses alarmes ou non). Par conséquent, la modélisation du système dynamique constitue une étape importante pour la résolution d'un problème de trajectographie. La partie suivante est consacrée à la construction des modèles dans le contexte de notre étude, à savoir le suivi dans des séquences d'images.

Deuxième partie

Modèles pour la trajectographie dans des séquences d'images

Introduction et problématique

Réaliser un suivi robuste d'une entité quelconque dans une séquence d'images est un des enjeux de la vision par ordinateur. Les applications sont nombreuses et différent selon le type de l'entité suivie. Citons, par exemple l'utilisation du suivi de points d'intérêt pour la reconstruction 3D et son application en robotique, en réalité augmentée ou en imagerie médicale; la nécessité du suivi d'objets pour la compression vidéo, et pour l'imagerie multimédia en général; ou encore l'importance du suivi de personnes pour la construction de systèmes d'aide ou de surveillance.

Ce problème de trajectographie dans des séquences d'images peut être formalisé de manière probabiliste par un système dynamique, dont l'estimation est apportée par un algorithme de filtrage. Comme cela est décrit dans [Pérez 03], ces méthodes présentent de nombreux avantages par rapport aux techniques déterministes existantes.

À partir de la modélisation de la vraisemblance des données observées conditionnellement à l'état recherché, un algorithme déterministe fournit une approximation du maximum de vraisemblance :

$$\hat{\mathbf{x}}_k = \arg \max_{\mathbf{x}_k} p(\mathbf{z}_k | \mathbf{x}_k), \quad (3.45)$$

ou du maximum *a posteriori* instantané conditionnellement à l'estimée précédente (dans le cas où la dynamique de l'entité est modélisée) :

$$\hat{\mathbf{x}}_k = \arg \max_{\mathbf{x}_k} p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \hat{\mathbf{x}}_{k-1}). \quad (3.46)$$

Dans les deux cas, l'estimée déterministe est soit le résultat d'une recherche exhaustive dans un voisinage de l'estimation précédente $\hat{\mathbf{x}}_{k-1}$, soit le résultat d'une descente de gradient commençant en $\hat{\mathbf{x}}_{k-1}$. Dans de nombreuses situations (occlusion, grand déplacement, changement de pose de l'entité suivie, variation de luminosité de la scène, présence d'une entité similaire, images fortement bruitées), des problèmes de minima locaux peuvent conduire à une divergence de l'algorithme. Le recours aux méthodes de filtrage permet de construire des algorithmes robustes à ces situations. Cette robustesse s'explique par plusieurs points :

- *Prise en compte de l'ensemble des mesures dans l'estimation*

Un algorithme de filtrage fournit une estimation de la loi *a posteriori* $p(\mathbf{x}_k | \mathbf{z}_{1:k})$. Ainsi, contrairement aux méthodes déterministes, le résultat du suivi prend en compte l'ensemble des données passées et présente, et peut s'apparenter à l'estimation d'une trajectoire complète. De plus, la loi complète étant disponible, des

mesures de confiance de l'estimée peuvent être construites.

– *Adaptabilité de l'algorithme au travers des bruits du modèle*

La formulation probabiliste du système permet de considérer des bruits sur la vraisemblance et la dynamique. Le modèle devant rendre compte d'une réalité physique sous-jacente, ces bruits décrivent une incertitude sur la modélisation construite. Si une estimation en ligne de ces bruits est possible, l'algorithme peut s'adapter automatiquement aux situations difficiles. Par exemple, l'estimation d'un bruit important sur une mesure à un instant donné permet de ne pas la prendre en compte dans l'estimation, et de se reposer uniquement sur la prédiction et les informations passées.

– *Maintien simple de plusieurs hypothèses*

Alors que les algorithmes déterministes nécessitent des techniques complexes pour maintenir plusieurs hypothèses de suivi, l'utilisation des méthodes de filtrage permet de conserver celles-ci de manière implicite. Les multi-modalités sont modélisées par un système non linéaire. Ceci conduit à l'emploi d'une méthode particulière où chaque particule correspond à une hypothèse de trajectoire possible.

En plus d'une grande robustesse, l'utilisation des algorithmes de filtrage présente des avantages pratiques :

– *Facilité de mise en œuvre*

Les algorithmes de filtrage sont très simples à implémenter, en particulier le filtre de Kalman pour le cas linéaire gaussien, ainsi que les filtres particuliers bootstrap et avec fonction d'importance optimale dans le cas de modèles non linéaires.

– *Absence de contraintes sur la modélisation*

La capacité des méthodes particulières à gérer les non linéarités autorise l'utilisation de modèles aussi complexes que l'application le nécessite¹. Les seuls pré-requis sur la modélisation sont de permettre la construction d'une fonction de proposition dont il est possible de générer des échantillons, et l'évaluation à une constante près de la loi associée au calcul des poids.

– *Possibilité de fusionner des données de différents types*

Liée à l'aptitude des méthodes particulières à gérer les problèmes complexes, l'utilisation de plusieurs types de données ne pose pas de problème. La prise en compte des différentes modalités (couleur, mouvement, son, etc.) se fait par une modélisation de la vraisemblance sous la forme d'une loi produit, ou par la construction d'une loi de proposition en plusieurs étapes [Pérez 04].

Il apparaît donc que la mise en œuvre d'une méthode de filtrage est une approche très intéressante pour la résolution d'un problème de trajectographie en images. Celle-ci nécessite la définition d'un modèle décrivant les propriétés du système dynamique étudié. Ce modèle est composé d'une **loi dynamique** et d'une **vraisemblance** des données, et

¹tant que la complexité du modèle n'entraîne pas un coût de calcul prohibitif

correspond à une modélisation par une chaîne de Markov cachée (cf. première partie).

Le modèle dynamique décrit l'évolution *a priori* de l'entité suivie. Il est utilisé pour prédire l'espace de recherche de la nouvelle position de l'objet. La difficulté de modélisation de ce terme résulte d'un compromis concernant la dimension de l'espace de recherche. En effet, cet espace doit être suffisamment grand pour pouvoir appréhender de grands déplacements (dans le cas où ces déplacements sont peu prédictibles), tout en étant assez restreint pour réduire les situations d'ambiguïté.

La définition de cette loi repose sur la connaissance ou non de la nature de l'entité suivie. Certaines applications nécessitent le suivi d'un objet précis, tel qu'un visage ou une personne. Dans ce cas, la dynamique peut être porteuse d'un *a priori* qui rend compte d'une modélisation du mouvement de l'objet. Une telle modélisation n'est plus possible lors de la construction d'un algorithme de suivi général, où aucune information n'est disponible sur l'objet suivi. Dans cette situation, deux approches peuvent être utilisées. La première consiste à mettre en place une loi d'évolution très peu informative. La deuxième repose sur l'estimation en ligne de la dynamique à partir d'informations extraites de l'image. Cette dernière solution permet de guider avec précision la recherche, tout en étant adaptée à des algorithmes généraux.

La vraisemblance décrit le lien entre l'état caché et une caractérisation de cet état dans l'image. Une telle caractérisation est construite à partir de données issues des images, permettant de reconnaître ou détecter l'objet d'intérêt. Elle doit être à la fois persistante tout au long de la séquence et spécifique pour permettre de différencier l'objet suivi des autres éléments de la scène. Une fois un modèle d'apparence de l'objet choisi, l'observation de l'état peut schématiquement être de deux types. Prenons l'exemple où l'état recherché est la position de l'objet. Dans un premier cas, la mesure correspond à la meilleure position candidate, estimée par un module de détection. La vraisemblance modélise alors une distance entre positions. Dans un deuxième cas, l'observation est constituée d'un ensemble de parties d'images associées à des positions candidates, et la vraisemblance rend compte d'une distance entre modèles d'apparence.

On voit ainsi apparaître plusieurs stratégies algorithmiques possibles. Celles-ci dépendent du type de données extraites des images, choisies comme observations de l'état caché.

Une séquence d'images – correspondant aux données brutes disponibles – contient une grande quantité d'informations. Afin de réaliser un modèle général efficace pour la trajectographie, il apparaît donc nécessaire d'en extraire une partie pour construire la vraisemblance, mais également la dynamique. Dans cette situation, une corrélation apparaît entre l'état et la mesure au travers des données images. Un tel choix met en évidence les limitations de la représentation du système par un modèle simple de chaîne de Markov cachée, qui ne modélise pas cette dépendance aux images.

Il nous a donc paru nécessaire de reformuler le modèle de chaîne de Markov cachée en introduisant de manière explicite les données images. Cette nouvelle représentation a l'avantage d'être générale, car elle est adaptée au cas où la dynamique est estimée sur la séquence, et au cas où elle est entièrement fixée *a priori*. Les nouveaux filtres associés ont été dérivés, dans les cas linéaire et non linéaire. Les équations du filtre de Kalman et du

filtre particulière ont ainsi été validées pour cette nouvelle formulation. Enfin, ce schéma permet de considérer des modèles simples pour lesquels la fonction d'importance optimale est disponible.

Le premier chapitre de cette partie dresse un état de l'art des principales modélisations de la vraisemblance et de la dynamique dans le contexte du suivi à partir de séquences d'images. Les différentes constructions de la loi d'évolution, les choix de caractérisations possibles de l'entité, ainsi que les stratégies algorithmiques y sont décrits. Cet état de l'art met en évidence certaines limitations du modèle simple de chaîne de Markov cachée. C'est pourquoi nous proposons dans le second chapitre une nouvelle formulation du problème de la trajectographie sous la forme d'une chaîne de Markov cachée prenant en compte de manière explicite les données images. Les avantages de cette formulation sont donnés ainsi que la description des filtres associés.

Chapitre 4

Approches existantes

La mise en œuvre d'un algorithme de trajectographie suppose la description par un modèle du système dynamique à estimer. Ce modèle est constitué d'une loi d'évolution $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ et d'une vraisemblance $p(\mathbf{z}_k|\mathbf{x}_k)$. La nature du modèle dépend du type d'objet suivi (point d'intérêt ou entité non ponctuelle), et de la connaissance disponible sur cet objet. Par exemple, certaines applications nécessitent le suivi d'un objet quelconque désigné par l'utilisateur. Dans ce cas, le niveau de précision de la modélisation est plus faible que dans le cas où la nature et le type de mouvement de l'objet sont connus (par exemple une voiture, un piéton).

La construction du modèle passe par la description de l'entité suivie en deux éléments principaux : sa représentation et sa caractérisation en terme d'analyse d'images. La représentation désigne le type de résultat souhaité en sortie de l'algorithme. Par exemple, un objet sera représenté par une courbe fermée quelconque, une ellipse, ou une fenêtre englobante. La caractérisation correspond à un ensemble de données extraites des images permettant d'identifier l'objet, comme un histogramme de couleur ou un modèle de contour. La figure 4.1 illustre une telle description dans le cas du suivi d'une personne en imagerie couleur. Dans cet exemple, la représentation choisie est une fenêtre englobante et la personne est caractérisée par l'histogramme de couleur de cette fenêtre.

À partir de cette définition de l'objet, une loi d'évolution et une vraisemblance peuvent être proposées. En effet, comme cela est détaillé en section 4.1, le choix du vecteur d'état et de sa dynamique dépendent de la représentation choisie. De même, comme expliqué en section 4.2, le modèle de mesure est fondé sur la caractérisation de l'objet mise en place. Une fois ces deux points décrits, nous verrons dans la section 4.3 que face à des non linéarités du modèle, l'utilisation des méthodes particulières peut se faire selon deux stratégies distinctes, et que plusieurs formes de la fonction d'importance peuvent être utilisées. Enfin, une discussion sera menée dans la section 4.4 sur la validité du modèle simple de chaîne de Markov cachée.

Remarque 1 : Dans ce chapitre, les différentes modélisations des bruits d'état et de mesure ne seront pas détaillées. Les modèles classiques de suivi à partir d'une séquence vidéo supposent ces bruits gaussiens et stationnaires. D'autres modèles sont utilisés dans des cas très particuliers tels que le traitement d'images échographiques.

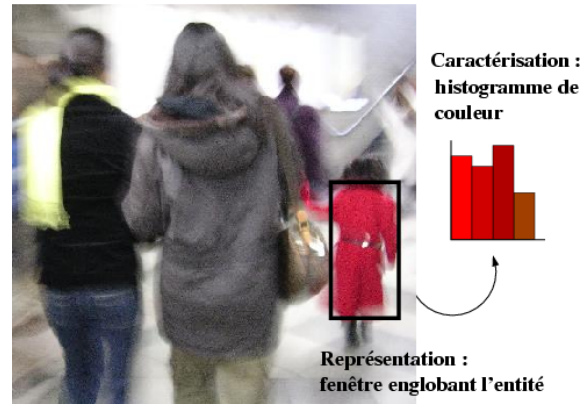


FIG. 4.1: Illustration des deux éléments nécessaires à la construction d'un modèle : représentation et caractérisation de l'entité. Exemple de suivi d'une personne en imagerie couleur. La représentation choisie est une fenêtre englobante et la personne est caractérisée par un histogramme de couleur.

Remarque 2 : À partir des modélisations de la dynamique et de la vraisemblance décrites dans les deux sections suivantes, l'application d'un algorithme de filtrage permet de suivre l'objet d'intérêt. Le suivi sera efficace si la procédure d'initialisation est précise. Une telle procédure est spécifique au choix de la représentation. Certaines applications nécessitent une initialisation automatique. Dans ce chapitre, nous ne détaillerons pas de solutions générales à ce problème difficile. Nous décrirons dans la partie suivante une procédure d'initialisation pour une application de suivi de points d'intérêt.

4.1 Représentation de l'objet : définition de l'état et de sa dynamique

Quelle que soit l'application de suivi, la définition de l'espace d'état est la première étape de la construction du modèle. La construction du vecteur d'état et de sa dynamique dépendent de la connaissance disponible sur l'entité suivie ainsi que de la représentation de l'objet utilisée. Les composantes de ce vecteur correspondent généralement à des paramètres de position et de mouvement.

4.1.1 Quand la nature de l'objet est connue

Certaines applications requièrent un suivi très précis d'une entité donnée. Par exemple, la création d'un système de formation à un poste de chargé de manœuvre peut nécessiter un algorithme de suivi des membres supérieurs d'une personne [Chateau 04]. Notons également l'utilisation d'un suivi robuste des expressions du visage et des mains pour la création d'un logiciel d'apprentissage du langage des signes.

La construction de tel algorithmes dédiés passe par le choix d'une représentation modélisant correctement l'entité suivie et sa dynamique.

Choix de l'état

Dans le cas où la nature de l'objet est connue, l'état peut rendre compte avec précision de la réalité physique. Citons par exemple le travail de Lu [Lu 04] sur le suivi de visage. Le visage y est représenté par un modèle 3d précis, dont les paramètres du mouvement 3d (rotations et translations) composent le vecteur d'état. Cette représentation est illustrée par la figure 4.2(a). Notons également les travaux de Choo [Choo 01] ou Deutscher [Deutscher 00] sur le problème difficile de la trajectographie d'une personne. Dans ces deux études, le corps humain est représenté par une structure de chaîne cinématique reproduite en figure 4.2(b). Cette représentation est reprise dans les recherches de Sidenbladh [Sidenbladh 00] [Sidenbladh 01] [Sidenbladh 02] qui modélise un individu par un modèle 3d complexe correspondant à une association de cônes tronqués, connectés par des articulations qui ont entre 1 et 3 degrés de liberté. Le vecteur d'état est alors constitué des paramètres de positions et rotations de la structure globale, ainsi que des angles de chaque articulation.

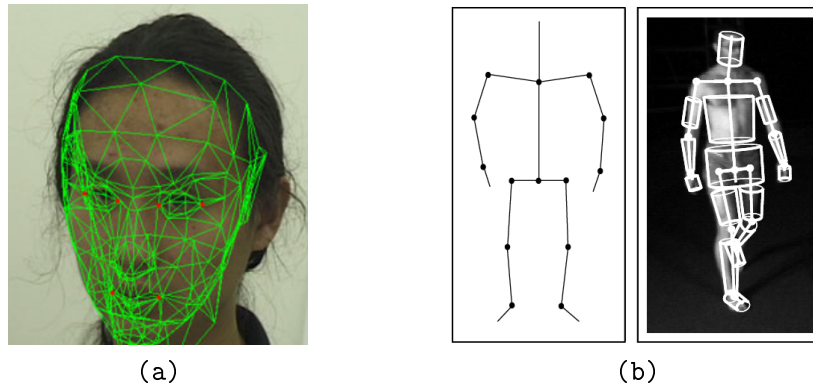


FIG. 4.2: Exemples de représentations précises d'objets. (a) représentation d'un visage par un modèle 3d - figure tirée de [Lu 04]. (b) représentation d'un corps humain par une chaîne cinématique - figure tirée de [Deutscher 00]

Choix de la dynamique

Une fois le vecteur d'état choisi, la définition de la loi dynamique permet d'introduire un *a priori* sur son évolution. Comme précédemment, cette loi peut être porteuse de plus d'informations lorsque la nature de l'objet est connue.

Lorsqu'une modélisation physique du mouvement de l'objet est disponible, la loi d'évolution est définie directement. Un exemple typique est l'utilisation d'un *a priori* de type vitesse constante dans un modèle pour la trajectographie d'une voiture. On peut également mentionner l'exemple de l'application de l'équation de Navier-Stokes pour décrire l'évolution d'une structure fluide.

Lorsque ce n'est pas le cas, il est possible d'utiliser des méthodes d'apprentissage afin de décrire cet *a priori*. Le *Visual Dynamics Research Group* (université d'Oxford) a apporté de nombreuses contributions sur ce sujet [North 98] [Blake 98] [Blake 99]. En particulier,

l'introduction d'une dynamique apprise dans l'algorithme CONDENSATION (équivalent à l'algorithme bootstrap) a été étudiée pour suivre des entités de différents types, telles qu'une main ou une feuille [Isard 98a]. L'utilisation d'une telle loi d'évolution s'est révélée intéressante pour la trajectographie d'objets dont les mouvements peuvent être décrits par un nombre restreint de classes (suivi d'une balle qui rebondit, d'une main qui dessine) [Isard 98c] [Rittscher 99] [Celeux 04]. Enfin, lorsque le nombre de classes permettant de caractériser le mouvement de l'objet est trop grand pour apprendre un modèle d'évolution, une alternative consiste à utiliser une très grande base de données d'exemples décrivant les étapes des différents mouvements possibles [Sidenbladh 02].

4.1.2 Quand la nature de l'objet est inconnue

Les méthodes présentées jusqu'ici pour définir un modèle d'état et sa dynamique reposent sur une connaissance précise de la nature de l'objet suivi. La mise en œuvre d'un algorithme de suivi plus général, permettant de reconstruire la trajectoire d'un objet quelconque, conduit à utiliser une modélisation moins précise ou adaptative.

Choix de l'état

Lorsque l'on veut conserver une forme assez précise de l'objet suivi, une représentation par un contour paramétré peut être utilisée. Deux types de paramétrisations sont souvent adoptés : la représentation analytique par courbes B-splines [Isard 98a] et la représentation fréquentielle par des contours de Fourier [Hue 01] [Hue 03]. Dans ces deux cas, le vecteur d'état est constitué des paramètres de la courbe et de son mouvement.

Un choix plus simple de représentation consiste à employer une forme géométrique englobant l'objet d'intérêt. Les formes les plus courantes sont des fenêtres rectangulaires [Pérez 02] [Odobez 03], des parallélogrammes [Zhou 04] ou des ellipses [Vermaak 02b] [Comaniciu 03] [Odobez 03] [Kwolk 04]. Ces représentations amènent à considérer un vecteur d'état comme la concaténation des paramètres d'évolution de la forme (paramètres d'échelle, rotation, translation, taille des axes de l'ellipse, etc.).

Quelques-unes de ces représentations pouvant s'adapter à un objet quelconque sont illustrées par la figure 4.3.

Choix de la dynamique

Lors de la mise en œuvre d'un algorithme de trajectographie général, il n'est pas possible d'utiliser des modèles dynamiques reposant sur une connaissance précise de la nature de l'objet. D'autres stratégies doivent être développées afin de modéliser des dynamiques permettant toutefois de guider la recherche. Deux approches ont été développées.

La méthode la plus répandue consiste à construire un système qui se contente d'un *a priori* peu informatif sur le modèle d'évolution de l'état. La dynamique choisie est alors généralement un modèle auto-régressif d'ordre 1 [Black 99] [Odobez 03] ou 2 [Isard 96] [Pérez 02]. L'instance la plus simple de ce modèle est une marche aléatoire. Ces modèles ont l'inconvénient de ne pas autoriser des déplacements de grande amplitude entre deux images successives. Un moyen de contourner cette difficulté est de modéliser la dynamique

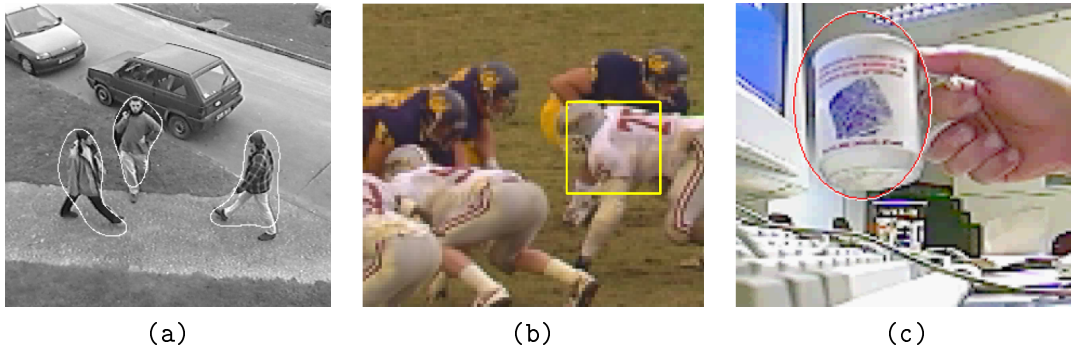


FIG. 4.3: Exemples de représentations simples d'objets. (a) représentation par un contour de Fourier - figure tirée de [Hue 01]. (b) représentation par une fenêtre rectangulaire - figure tirée de [Pérez 02]. (c) représentation par une ellipse - figure tirée de [Comaniciu 03]

comme une loi mélange de la forme [Pérez 03] [Pérez 04] :

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}) = (1 - \alpha) \mathcal{N}(\mathbf{x}_k ; F \mathbf{x}_{k-1}, Q_k) + \alpha \mathcal{U}_{\mathcal{X}}(\mathbf{x}_k), \quad (4.1)$$

où la première composante est un processus auto-régressif et la seconde une loi uniforme sur l'espace d'état $\mathcal{X}$ supposé borné. Le coefficient α ($0 < \alpha < 1$) est le poids de la composante uniforme, supposé très petit. L'ajout de cette composante uniforme permet de capter des sauts de mouvement, et éventuellement de récupérer une trajectoire après occlusion. Une manière différente de procéder consiste à utiliser une distribution de prédiction à queue plus lourde qu'une loi gaussienne comme par exemple une loi de Student [Pérez 04].

La deuxième stratégie, beaucoup moins employée, repose sur l'extraction de données à partir de la séquence d'images pour la mise en place d'un modèle d'évolution adaptatif. En particulier, des informations de mouvement estimées sur la séquence peuvent être intégrées dans la dynamique. De telles données sont déduites des images par une méthode d'estimation du mouvement [Meyer 94] [Kervrann 95]. La loi d'évolution ainsi construite s'adapte en ligne aux changements de direction de l'entité. La capacité de cette dynamique à capter les grands déplacements repose sur la précision de la méthode d'estimation du mouvement utilisée. Une estimation précise permet de prédire correctement des mouvements chaotiques et des changements brusques de vitesse.

Cette approche nous semble très intéressante pour définir une loi d'évolution lorsque la nature de l'objet et/ou de sa dynamique sont inconnues. Nous étudierons en troisième partie de ce document sa mise en œuvre dans notre application de trajectographie de point et de plan. De plus, nous verrons que dans le cas où une mesure de la qualité de l'estimation de mouvement est disponible, celle-ci peut être utilisée pour estimer en ligne le bruit d'état. Ce dernier point est un avantage important de cette méthode.

D'autres algorithmes de trajectographie utilisant une dynamique estimée sur les images ont été développés simultanément à notre étude. L'information de mouvement introduite est apportée par une méthode différentielle d'estimation (qui permet une estimation du bruit d'état) [Zhou 04] ou un filtre de Kalman [Kang 03].

Remarquons que l'utilisation de cette deuxième stratégie amène à considérer un modèle qui repose entièrement sur des données extraites de la séquence. Il est alors important de différencier les données utilisées comme observations de l'état des données utilisées pour la définition de la loi d'évolution. Cependant, l'ensemble de ces données sont déduites de la séquence d'images, et sont donc en théorie dépendantes entre elles. Cette dépendance n'est pas prise en compte dans les algorithmes existants [Meyer 94] [Kervrann 95] [Kang 03] [Zhou 04].

Dans cette section, nous avons présenté les méthodes existantes pour définir l'état et la loi d'évolution dans le cadre d'une application de suivi dans une séquence d'images. La deuxième étape de la construction du modèle consiste en une définition des observations et de leur vraisemblance. Celle-ci est décrite dans la section suivante.

4.2 Caractérisation de l'objet : définition de la mesure et de la vraisemblance

Une fois la représentation de l'entité définie, le choix de sa caractérisation est essentiel pour la mise en œuvre d'un algorithme de suivi efficace. La caractérisation de l'entité désigne un ensemble de données permettant de l'identifier, voire de la détecter dans la séquence. En général, un modèle de l'objet est défini par une ou plusieurs caractéristiques issues des images. Le choix de ces caractéristiques est crucial. En effet, ce modèle doit être à la fois persistant le long de la séquence pour ne pas perdre l'objet et discriminant pour ne pas introduire une confusion entre plusieurs objets. Les caractéristiques les plus utilisées comprennent les modèles de contour, les histogrammes de couleur, des mesures de mouvement ou utilisent des images de référence.

Modèle de contour

L'utilisation des caractéristiques de contour d'un objet a été introduite avec l'algorithme CONDENSATION [Isard 96]. Depuis, cette information a été employée dans de nombreuses méthodes de trajectographie [Isard 98b] [Vermaak 01] [Odobez 03] [Wu 04] [Chen 04].

Le modèle de contour permet de capter la forme d'un objet en utilisant les informations de gradients dans l'image. L'objet initial est représenté par une courbe paramétrée par B-spline (à p points de passage). Les particules prédites avec la loi dynamique correspondent à des courbes hypothèses. À chacune de ces hypothèses est associé un ensemble de mesures de gradient calculées le long de p lignes normales au contour hypothèse. Le poids associé à chacune des particules est calculé par la vraisemblance qui modélise la distance entre le contour hypothèse et les données de gradient de l'image les plus proches de celui-ci. Ce modèle de contour est illustré par la figure 4.4.

Bien qu'une telle caractéristique permette de capter la forme d'un objet, elle est perturbée dès que l'arrière plan de l'image contient trop de contours parasites. De plus, elle n'est pas adaptée quand il n'y a pas de classe d'objet prédéfinie à suivre, ou quand l'entité suivie n'a pas une silhouette invariante dans le temps (personne qui marche, mouvement de la main).

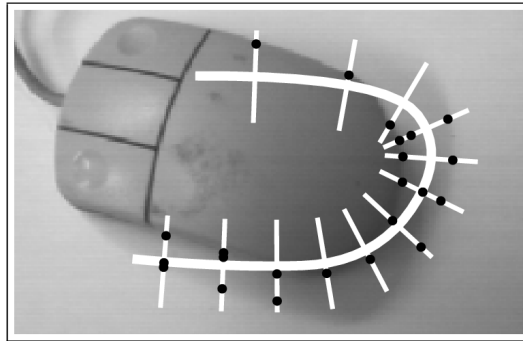


FIG. 4.4: Modèle de contour. La ligne blanche épaisse représente un contour hypothèse, les lignes fines sont les lignes sur lesquelles un algorithme de détection de contour est utilisé, l'ensemble des points noirs correspond à la mesure associée au contour hypothèse - figure tirée de [MacCormick 99]

Histogramme de couleur

L'information de couleur de l'objet est largement utilisée dans les algorithmes de suivi par filtrage [Pérez 02] [Comaniciu 03] [Nummiaro 03] [Odobez 03] [Fei 04] [Wu 04] [Kwolk 04] [Pérez 04] [Chen 04]. La couleur associée à une zone de l'image est représentée par un histogramme (le système teinte-saturation-luminosité est souvent préféré au système rouge-vert-bleu). Des mesures de distance entre histogrammes peuvent être définies en utilisant le coefficient de similarité de Bhattacharyya [Comaniciu 03], la distance de Kullback-Leibler [Liu 04] ou une mesure d'intersection [Wu 04]. En pratique, un histogramme de référence est associé à l'objet initial, généralement représenté par une fenêtre ou une ellipse. Trois approches peuvent être alors mises en place :

- La première méthode est utilisée dans [Comaniciu 03]. Elle consiste à utiliser un algorithme de détection de la fenêtre (ou ellipse) dont l'histogramme est de distance minimale avec l'histogramme de référence. Cette fenêtre est considérée comme l'observation $\mathbf{z}_k$ de l'état. Il est alors possible de construire une équation de mesure linéaire.
- La deuxième méthode, proposée dans [Pérez 02] est mise en œuvre dans le cadre des méthodes séquentielles de Monte Carlo. Il n'y a pas de module de détection d'un meilleur candidat, mais on suppose que les observations correspondent aux zones de l'image associées aux particules prédites. La vraisemblance de chaque particule est évaluée en utilisant une équation de mesure qui traduit la distance entre son propre histogramme et l'histogramme de référence (elle peut aussi éventuellement prendre en compte une modélisation de l'arrière plan de l'image [Pérez 02] [Vermaak 02b]). Cette vraisemblance est reliée aux pixels de l'image et l'équation de mesure associée est donc généralement hautement non linéaire.
- Comme la deuxième méthode, la troisième méthode s'applique au cadre des algorithmes particuliers. L'information de couleur est utilisée pour construire une loi de proposition permettant d'échantillonner efficacement les particules [Isard 98b].

Ces trois types de stratégies seront repris en détail et dans un cadre plus général dans la section suivante.

L'information de couleur modélisée par un histogramme a l'avantage d'être robuste aux changements de structure spatiale des objets dus aux modifications d'angle de vue ou aux occlusions partielles. Un histogramme est donc une caractéristique persistante. Face à des applications de suivi dans de longues séquences, l'histogramme de référence doit être mis à jour afin de s'adapter aux changements possibles d'apparence de l'objet (dus à un fort changement d'illumination, un changement de pose, etc.) [Nummiaro 03]. Cela peut être réalisé par un algorithme de type EM (Expectation-Maximisation) [Raja 98] [Vermaak 02b]. Néanmoins, les histogrammes de couleur ont le désavantage d'être assez peu spécifiques. Cette caractéristique est donc généralement combinée à d'autres informations.

Mesure de mouvement

Bien qu'intermittente, l'information de mouvement est une caractéristique très intéressante. D'une part, grâce à une recherche très active dans le domaine de l'estimation de mouvement, une grande variété d'algorithmes permettant d'extraire cette information existent. D'autre part, elle a l'avantage d'être spécifique. En particulier, il est possible de combiner les caractéristiques de couleur et de mouvement, ce qui permet de lever d'éventuelles ambiguïtés [Vermaak 02b] [Pérez 04].

L'information de mouvement peut être décrite de différentes manières, comme une carte de segmentation au sens du mouvement [Hue 01], un histogramme de mouvement lié à une zone de l'image [Pérez 04], ou une carte de différences inter-images [Sidenbladh 01] [Vermaak 02b]. Selon le type de mesure choisie, l'utilisation de l'information de mouvement peut se faire selon différentes stratégies, sur la base de principes similaires aux approches décrites dans le cas de la couleur. Le choix de la stratégie repose sur l'utilisation ou non d'un module de détection, et sur la prise en compte ou non de cette mesure dans la loi de proposition.

Image de référence ("template")

Le dernier type de caractérisation très utilisé dans des applications de suivi consiste à choisir une (ou plusieurs) image(s) de référence comme modèle d'apparence de l'objet [Hager 98] [Sullivan 01] [Nguyen 01]. Cette caractérisation s'applique en particulier à la trajectographie dans des images en niveaux de gris.

De nombreuses mesures permettant d'évaluer la similarité entre deux images de petites tailles sont disponibles [Odobez 03][Aschwandten 92], ainsi que des algorithmes permettant de trouver la partie de l'image la plus similaire à l'image de référence [Baker 04b]. Comme dans le cas de l'histogramme de couleur, la mesure désigne soit la meilleure représentation candidate au sens du minimum de distance avec l'image de référence, soit l'ensemble des parties d'images associées aux particules prédites. Les stratégies précédentes peuvent ainsi être mises en œuvre.

La caractérisation par un modèle d'image a l'avantage d'être très discriminante en cas d'ambiguïté entre plusieurs objets. Cependant, elle pose le problème important de la mise à jour de l'image de référence. Celle-ci est essentielle pour rendre l'algorithme robuste

aux changements d'apparence de l'entité considérée et éviter les dérives du suivi. Diverses méthodes pour réaliser cette mise à jour sont décrites dans [Nguyen 01], [Jepson 01] et [Matthews 04]. Une solution possible visant à contourner cette difficulté consiste à accumuler les différentes imagerie associées aux estimations. Cet ensemble forme une base d'exemples décrivant les variations d'apparence de l'objet suivi. Cette base peut être construite hors ligne [Rahimi 01] ou en ligne [Morency 03].

Autres types de mesures et fusion de données au travers de la vraisemblance

D'autres types de caractérisation de l'entité peuvent être utilisées. Citons par exemple l'utilisation de cartes de profondeur [Kwolek 04], d'images de dérivées [Sidenbladh 01], d'informations de texture [Nguyen 04], de bases d'exemples [Toyama 02] [Fei 04], de représentations dans l'espace des vecteurs propres [Black 96], ou de modèles d'apparence adaptatifs sous forme d'un mélange de composantes de l'image [Jepson 01] [Zhou 04].

Enfin, d'autres modalités non extraites de la séquence d'images peuvent être prise en compte en complément, tel que le son pour des applications de visio-conférence [Vermaak 01] [Vermaak 02b] [Pérez 04] [Chen 04].

Le fait de pouvoir fusionner simplement plusieurs modalités d'observations différentes constitue un avantage important des méthodes de filtrages. En effet, sous l'hypothèse d'indépendance des différentes observations conditionnellement à l'état, la vraisemblance globale s'écrit comme le produit des vraisemblances associées aux Γ modalités différentes $\mathbf{z}_k^\gamma$:

$$p(\mathbf{z}_k|\mathbf{x}_k) = \prod_{\gamma=1}^{\Gamma} p(\mathbf{z}_k^\gamma|\mathbf{x}_k). \quad (4.2)$$

Un modèle construit avec cette vraisemblance est alors traité à l'aide d'une méthode particulière.

Cette modélisation a permis la fusion des caractéristiques, respectivement de couleur et mouvement [Pérez 04], couleur et son [Pérez 04] [Chen 04], son et contour [Vermaak 01] [Chen 04], ainsi que contour et image de référence [Odobez 03]. Les algorithmes ainsi construits bénéficient des avantages des différents modèles d'apparence, dont la combinaison est en général à la fois persistante le long de la séquence et spécifique. Une telle fusion de données augmente la robustesse de la méthode de trajectographie, et permet de compenser une dynamique peu informative.

4.3 Stratégies algorithmiques

Comme cela a été évoqué dans la section précédente, plusieurs stratégies algorithmiques peuvent être mises en place afin de réaliser un suivi à l'aide d'une méthode de filtrage. Le schéma général de la méthode peut suivre deux approches différentes, que nous avons choisi d'appeler approches *Top-Down* et *Bottom-Up*. Il nous semble que ces appellations permettent de comprendre le principe de ces stratégies : la première utilise une méthode haut niveau pour construire un modèle pouvant être résolu de manière simple, alors que la deuxième utilise un modèle construit sur les pixels de l'image et résolu par une méthode particulière. Le choix d'un schéma particulier dépend de l'utilisation d'un module de détection

de la nouvelle observation ou non. Enfin, lorsque le modèle construit nécessite l'utilisation d'une méthode particulière, plusieurs techniques peuvent être utilisées pour construire une fonction d'importance permettant de guider de manière efficace les particules.

4.3.1 Approche *Top-Down*

La stratégie *Top-Down* consiste à utiliser un module de détection de l'entité suivie dans l'algorithme de trajectographie. Le résultat de ce module est utilisé comme observation de l'état caché. Si cette mesure $\mathbf{z}_k$ est de même nature que l'état à estimer $\mathbf{x}_k$, il est alors possible de construire une équation de mesure linéaire du type $\mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{v}_k$ où $\mathbf{v}_k$ est un bruit qui modélise l'incertitude liée au résultat du module de détection. La possibilité d'inférer une équation de mesure linéaire est un avantage important de cette approche.

Dans le cas où l'équation dynamique est elle aussi linéaire, et que les bruits sur le modèle d'évolution et sur le modèle de mesure sont additifs gaussiens, le filtre de Kalman répond de façon optimale au problème. Face à des non linéarités du modèle, une méthode particulière est utilisée. Les différentes étapes d'une itération d'une telle méthode sont illustrées par la figure 4.5. Il est important de noter que la mesure étant disponible dès le début d'une itération, celle-ci peut être prise en compte pour l'échantillonnage des particules au travers de la fonction d'importance.

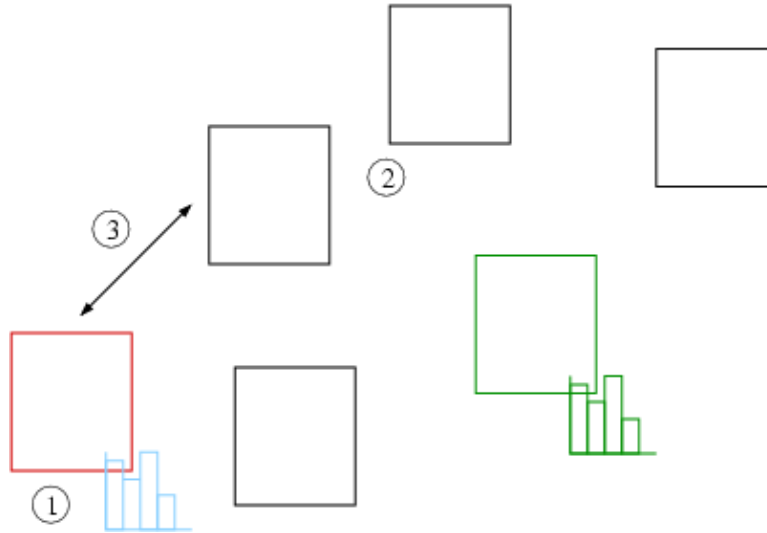


FIG. 4.5: Stratégie *Top-Down* – Exemple d'une représentation de l'entité suivie par fenêtre englobante, et d'une caractérisation par un histogramme de couleur – L'estimation précédente et l'histogramme de référence sont représentés en vert. (1) La première étape consiste à détecter la nouvelle observation (en rouge). L'observation correspond à la fenêtre qui minimise la distance entre son histogramme (en bleu) et l'histogramme de référence. (2) Puis, les nouvelles particules (en noir) sont prédites avec la fonction d'importance qui prend en compte éventuellement l'observation. (3) Enfin, le poids de chaque particule est évalué en utilisant la vraisemblance, qui modélise la distance entre sa position et la position de l'observation.

Cette approche *Top-Down* est avantageuse lorsque la phase de détection de l'observation – qui correspond à la meilleure représentation candidate de l'état – n'est ni trop complexe à mettre en œuvre, ni trop coûteuse en temps de calcul. Des exemples de son utilisation sont décrits dans [Comaniciu 03] (la mesure est une fenêtre dont l'histogramme de couleur minimise une distance avec l'histogramme de référence), [Meyer 94] (la mesure est une région obtenue par une carte de segmentation au sens du mouvement), ou [Pérez 04] (l'utilisation de mesures sonores donne une localisation de l'état caché). Nous verrons également dans la partie suivante son utilisation pour du suivi de points d'intérêt dans une séquence.

Une telle stratégie permet une modélisation simple du système à estimer. Dans ces conditions, l'estimation se révèle efficace. Elle permet de trouver un équilibre entre la complexité du modèle décrivant le système dynamique et la précision de l'algorithme permettant de traiter ce modèle. Cependant, il n'est pas toujours possible de disposer d'un module de détection, en particulier pour le suivi d'objets complexes. Dans ce cas, l'approche *Bottom-Up* constitue une alternative intéressante.

Remarque : Le module de détection nécessaire à la stratégie *Top-Down* est en général construit comme une recherche exhaustive du meilleur candidat sur une partie de l'image ou une descente de gradient initialisée à l'estimation précédente. L'utilisation de cette dernière technique sur une fonction de coût non convexe implique une dépendance entre l'observation à l'instant k et l'état à l'instant $k-1$. Or, cette dépendance viole les hypothèses du modèle classique de chaîne de Markov cachée. La construction de nouveaux modèles prenant en compte cette dépendance est une perspective à étudier.

4.3.2 Approche *Bottom-Up*

L'approche *Bottom-Up* a été introduite avec l'algorithme CONDENSATION [Isard 96]. Depuis, elle est majoritairement utilisée pour des applications de suivi par méthode de filtrage [Sidenbladh 01] [Hue 01] [Pérez 02] [Pérez 04] [Odobez 03]. Contrairement à l'approche *Top-Down*, celle-ci s'applique uniquement au cadre des méthodes particulières et ne nécessite pas l'utilisation d'un module de détection.

Historiquement, le principe de cette méthode consiste à prédire les particules selon la loi d'évolution et de les évaluer grâce à la vraisemblance. Dans ce cas, la mesure ne correspond pas à la meilleure représentation de l'état candidate, mais à l'ensemble des parties de l'image (ou l'ensemble des caractérisations de ces parties de l'image) associées aux particules prédites. La vraisemblance est donc en général complexe et hautement non linéaire car elle rend compte d'une distance évaluée directement sur les pixels de l'image. Afin d'illustrer cette approche *Bottom-Up*, la figure 4.6 décrit les différentes étapes à réaliser pour chaque itération.

Dans le chapitre 2, nous avons vu l'incidence d'une bonne répartition des particules dans l'espace d'état sur la qualité du suivi. Les particules sont d'autant mieux situées en des positions de l'espace de forte vraisemblance que la fonction d'importance repose à la fois sur le modèle d'évolution et sur la nouvelle observation. Or, dans la stratégie décrite ici, cette nouvelle observation n'est disponible qu'une fois les particules déplacées. Il est donc impossible de la prendre en compte pour la construction d'une fonction d'importance

efficace (à part de manière contournée par le filtre particulière auxiliaire). Une façon de contourner ce problème consiste à utiliser d'autres données issues des images pour définir cette fonction d'échantillonnage. La section suivante présente différents choix possibles pour définir cette fonction.

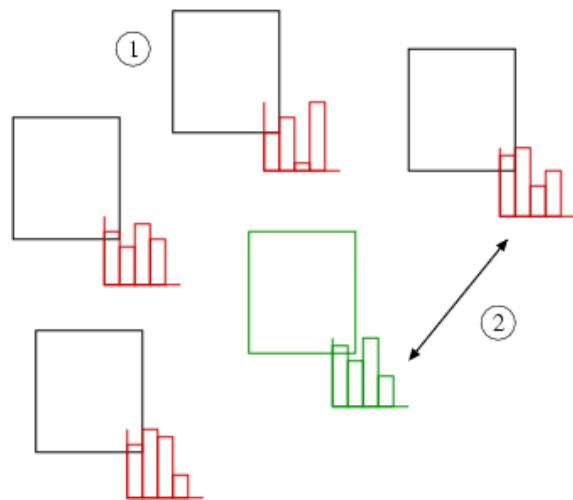


FIG. 4.6: Stratégie *Bottom-Up* – Exemple d’une représentation de l’entité suivie par fenêtre englobante, et d’une caractérisation par un histogramme de couleur – L’estimation précédente et l’histogramme de référence sont représentés en vert. (1) La première étape consiste à prédire les nouvelles particules (en noir). L’observation correspond alors à l’ensemble des histogrammes associées aux particules (en rouge). (2) La deuxième étape évalue le poids de chaque particule au travers de la vraisemblance qui modélise la distance entre son propre histogramme et l’histogramme de référence.

4.3.3 Choix de la fonction d’importance

Dans le chapitre 2, paragraphe 2.3 (p. 42), nous avons vu différentes stratégies d’échantillonnage des particules. Cette étude met en évidence l’importance de la prise en compte de l’observation courante dans la fonction d’importance. Dans le cas où la fonction d’importance optimale n’est pas disponible – ce qui est généralement le cas – différentes techniques peuvent être mise en place, dont la précision influe directement sur la qualité du filtre résultant.

Dans le cadre du suivi dans des séquences d’images, toutes ces techniques peuvent bien évidemment être employées. Citons par exemple l’utilisation du filtre particulière unscented pour un suivi de visage [Rui 01], l’ajout d’une étape de Monte Carlo pour de la trajectographie d’un corps humain [Deutscher 00] [Choo 01] [Sminchisescu 02], la mise en œuvre du filtre particulière auxiliaire pour résoudre le problème de la localisation d’un robot à partir d’acquisitions vidéo [Vlassis 02] ou l’échantillonnage par la vraisemblance pour le même problème [Wolf 02]. Les stratégies d’échantillonnage de ces méthodes utilisent le modèle d’évolution ainsi que les mesures extraites de l’image (utilisées comme observation de l’état) pour faire évoluer les particules dans l’espace d’état.

Un certain nombre de techniques, spécifiques au suivi en vision par ordinateur, utilisent d'autres données image que celles utilisées dans le modèle d'observation pour guider les particules. En particulier, dans le cadre de la trajectographie à partir de séquences, il peut être intéressant de considérer des lois de proposition qui autorisent des grands déplacements. De telles lois peuvent être utiles pour réinitialiser un suivi après un décrochage dû par exemple à une dérive du suivi ou à une situation d'occlusion. Pour cela, Vermaak et Pérez [Vermaak 02b] [Pérez 04] proposent de construire une fonction d'importance sous la forme d'une loi mélange de la densité de prédiction et de lois gaussiennes centrées en D positions $\{\mu_d\}_{d=1\dots D}$ de l'espace d'état sélectionnées par un module de détection :

$$\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}) = (1 - \nu)p(\mathbf{x}_k | \mathbf{x}_{k-1}) + \frac{\nu}{D} \sum_{d=1}^D \mathcal{N}(\mathbf{x}_k; \mu_d, \Sigma_d), \quad (4.3)$$

où ν ($0 < \nu < 1$) est le poids de la composante permettant un grand déplacement. $\{\Sigma_d\}_{d=1\dots D}$ sont les covariances estimées par le module de détection ou fixées *a priori*. Ce module de détection peut par exemple reposer sur un algorithme d'apprentissage de type Adaboost [Okuma 04] ou sur une mesure de mouvement par différence d'images [Vermaak 02a]. Notons que pour rendre l'emploi de cette fonction d'importance efficace, la densité de prédiction doit être de la forme (4.1) (p. 85) [Pérez 04]. L'idée de mettre en place un module de détection pour guider l'évolution des particules se retrouve également dans l'algorithme ICONDENSATION [Isard 98b] où l'information issue des images est une carte de segmentation par la couleur.

Une alternative à l'utilisation de l'information de mouvement au travers d'un module de détection est proposée par Hue [Hue 01]. L'idée est de faire migrer chaque particule en utilisant un vecteur de mouvement calculé en fonction de la position de celle-ci. On remarque que le filtre particulaire déduit peut être formalisé par un filtre bootstrap, où l'information de mouvement est utilisée pour définir la dynamique [Hue 03].

Toujours dans le but de rendre plus efficace l'exploration de l'espace d'état, une autre approche consiste à fusionner une évolution déterministe des particules avec un échantillonnage probabiliste [Sullivan 01]. À chaque itération, une surface déterminée à partir d'une mesure de distance est calculée (cette surface est par exemple une surface de corrélation entre l'image entière et une image de référence de l'entité suivie). Les particules suffisamment proches d'un minimum local important de cette surface évoluent selon une descente de gradient, alors que les particules situées dans une zone contenant plusieurs pics sont propagées par le modèle de prédiction *a priori*.

L'utilisation de l'échantillonnage partitionné permet également une répartition efficace des particules pour des applications de suivi de plusieurs entités [MacCormick 99] ou d'objets articulés [MacCormick 00] [Sidenbladh 01]. Cette stratégie peut être comparée à une recherche hiérarchique. Une idée similaire peut être appliquée pour fusionner plusieurs données image dès la phase d'échantillonnage [Pérez 04]. L'algorithme de filtrage particulaire est alors composé de plusieurs étapes d'échantillonnage / mise à jour des poids, chaque étape correspondant à une modalité observée. Enfin, cette fusion peut également être traduite par la construction de la fonction d'importance sous la forme d'un mélange de lois, où chaque composante est associée à une des modalités mesurées [Chen 04].

4.4 Synthèse et discussion

Dans ce chapitre, nous avons présenté les principales modélisations mises en œuvre dans le domaine du traitement d'images pour l'estimation d'un système dynamique. Dans le cadre des approches par filtrage, cette modélisation se compose d'une loi dynamique du système, et d'une vraisemblance décrivant la relation entre l'état caché et les mesures observées.

Certaines applications nécessitent la trajectographie d'un objet dont la nature est connue au préalable. Dans ce cas, la loi d'évolution peut être simplement posée *a priori*. Si l'objet a une évolution complexe, celle-ci peut être apprise sur une base d'exemples ou décrite par un nombre fini de modèles de mouvement. La construction d'algorithmes plus généraux impliquent une connaissance imprécise de l'entité suivie. Le modèle d'évolution est alors soit peu informatif, soit déduit d'une mesure de mouvement instantané calculée sur la séquence d'images. Grâce à l'existence de nombreux algorithmes d'estimation de mouvement, cette deuxième solution a l'avantage de pouvoir être utilisée quel que soit le type d'objet suivi, et de capter des mouvements chaotiques ou de grande amplitude, ou encore des changements brusques de vitesse. Dans le cas où une mesure de confiance de la mesure de mouvement obtenue est disponible, le bruit du modèle d'évolution peut être calculé automatiquement. Ce dernier point est un avantage essentiel de ce type d'approche.

Une fois la loi d'évolution décrite, la vraisemblance associée à l'observation de l'état doit être définie. Cette modélisation passe par le choix d'une caractérisation de l'objet suivi en terme d'analyse d'image. Cette caractérisation constitue un ensemble de données extraites des images permettant de le reconnaître, voire de le détecter. Les données les plus utilisées sont des informations de couleur, de contour, de mouvement, ou des mesures de similarité avec une image de référence.

Dans le cas où la caractérisation choisie permet de détecter l'objet d'intérêt, la vraisemblance rend compte d'une distance entre le résultat de cette détection et l'état caché. La stratégie algorithmique est alors de type *Top-Down*. Dans le cas où la caractérisation choisie permet uniquement d'évaluer la qualité d'une position de l'espace d'état candidate, la vraisemblance décrit une distance entre les modèles d'apparence de cette position et de l'entité recherchée. La stratégie algorithmique est alors de type *Bottom-Up*.

Lorsque le modèle construit présente des non linéarités, une méthode particulière doit être utilisée. Comme décrit dans la partie précédente, la qualité de cette méthode dépend de la stratégie d'échantillonnage des particules. Lorsque la fonction d'importance optimale n'est pas accessible, plusieurs manières d'inclure des données issues des images dans la fonction d'importance sont possibles. Ces données ne sont pas forcément utilisées dans le calcul de la vraisemblance.

En résumé, les définitions de la dynamique et de la vraisemblance peuvent s'appuyer sur des informations extraites de la séquence d'images. Celles-ci, bien que de nature différente, sont corrélées car elles sont toutes issues des images. De même, d'autres mesures provenant de la séquence peuvent entrer en jeu dans la construction de la fonction d'importance. Il est donc essentiel de considérer une modélisation par chaîne de Markov cachée

qui tienne compte de cette corrélation. Or, le modèle simple décrit en première partie (p. 16) ne décrit pas cette dépendance possible entre la vraisemblance et la dynamique au travers des données images brutes. Dans le chapitre suivant, nous proposons une nouvelle formulation qui généralise le modèle simple de chaîne de Markov classiquement employé. Cette formulation a la particularité de modéliser explicitement la série d'images, et d'être valide même lorsque la loi dynamique est estimée en ligne sur la séquence.

Chapitre 5

Modélisation proposée : prise en compte explicite des données images

Dans ce chapitre, nous décrivons une nouvelle modélisation pour le problème de l'estimation d'un système dynamique à partir de mesures bruitées. Le modèle proposé est spécifique à la trajectographie dans les séquences d'images puisqu'il modélise de façon explicite les données images. Il est appelé **modèle conditionnel**. Dans un premier temps, celui-ci est décrit après un rappel du modèle classique. Ensuite, les filtres conditionnels dédiés à sa résolution sont dérivés. Dans le cas linéaire, le **filtre conditionnel linéaire** est déduit d'un nouvel estimateur appelé **estimateur linéaire conditionnel de variance minimale**. Dans le cas non linéaire, le **filtre conditionnel non linéaire** est construit à partir d'un filtre particulière. Enfin, nous étudions dans une troisième partie des systèmes gaussiens permettant d'utiliser la fonction d'importance optimale dans le cadre particulière. Nous expliquerons comment l'utilisation de ces systèmes dans notre contexte conditionnel peut être avantageuse.

5.1 Du modèle classique au modèle conditionnel

Le problème général de filtrage consiste à estimer au cours du temps un processus d'état à partir de mesures bruitées passées et présentes. Rappelons que classiquement, ce problème est formalisé par deux processus aléatoires à temps discret, le processus d'état $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ et le processus de mesure $\{\mathbf{z}_k\}_{k \in \mathbb{N}}$. Le processus d'état est inconnu, supposé markovien. Il est observé via les mesures supposées indépendantes conditionnellement à l'état. Le système constitué de ces deux processus est alors représenté par une chaîne de Markov cachée illustrée par la figure 5.1. La chaîne est entièrement décrite par les distributions suivantes : la distribution du processus d'état à l'instant initial $p(\mathbf{x}_0)$, le modèle d'évolution du processus d'état $p(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k-1})$ et le modèle de vraisemblance $p(\mathbf{z}_k | \mathbf{x}_{0:k}, \mathbf{z}_{1:k-1})$. On rappelle les hypothèses suivantes :

$$p(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k-1}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}), \quad (5.1)$$

$$p(\mathbf{z}_k | \mathbf{x}_{0:k}, \mathbf{z}_{1:k-1}) = p(\mathbf{z}_k | \mathbf{x}_k). \quad (5.2)$$

Afin de résoudre le problème de filtrage, l'objectif est d'estimer la loi *a posteriori* $p(\mathbf{x}_k | \mathbf{z}_{1:k})$.



FIG. 5.1: Graphe de dépendance orienté de la chaîne de Markov cachée constituée des processus d'état $\mathbf{x}_{0:k}$ et de mesure $\mathbf{z}_{1:k}$.

Comme expliqué au chapitre précédent, cette formulation est limitée. En particulier elle n'est plus valide lorsque le modèle d'évolution est estimé sur les données images. Notre contribution consiste en la généralisation de ce formalisme au cadre de la trajectographie dans une séquence d'images.

On note $\mathbf{I}_k$ la variable aléatoire qui désigne une image à l'instant k . On définit la variable aléatoire notée $\mathbf{I}_{0:k}$, construite comme la concaténation des variables $\{\mathbf{I}_t\}_{t=0,\dots,k}$. $\mathbf{I}_{0:k}$ dénote une séquence d'images de l'instant initial à l'instant k , et $\mathbf{I}_{0:n}$ désigne une séquence complète. À partir de cette définition, on peut décrire la séquence d'images observée comme une réalisation de cette variable.

On souhaite modéliser la dépendance de la mesure et de l'état à cette nouvelle variable. La distribution initiale et les distributions (5.1) (5.2) décrivant le système sont modifiées en introduisant la séquence comme variable conditionnelle. 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 introduit une hypothèse de causalité par rapport à l'acquisition des images. Cette hypothèse se traduit par l'indépendance de l'état et la mesure au temps k par rapport à $\mathbf{I}_{k+1:n}$. Les distributions en jeu s'écrivent alors :

$$p(\mathbf{x}_0|\mathbf{I}_{0:n}) = p(\mathbf{x}_0|\mathbf{I}_0), \quad (5.3)$$

$$p(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k-1}, \mathbf{I}_{0:n}) = p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:n}) = p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}), \quad (5.4)$$

$$p(\mathbf{z}_k|\mathbf{x}_{0:k}, \mathbf{z}_{1:k-1}, \mathbf{I}_{0:n}) = p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:n}) = p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k}). \quad (5.5)$$

Ces distributions décrivent une modélisation du système dynamique par une chaîne de Markov cachée constituée des processus d'état, de mesure, et prenant en compte de manière explicite la séquence d'images. Cette chaîne est représentée par la figure 5.2.

Les équations du filtre optimal peuvent être appliquées à ce nouveau modèle appelé **modèle conditionnel**. La distribution *a posteriori* recherchée de l'état conditionnellement à l'ensemble des informations disponibles s'écrit alors $p(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. En supposant $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1})$ connue, les deux étapes de la solution bayésienne récursive s'écrivent naturellement :

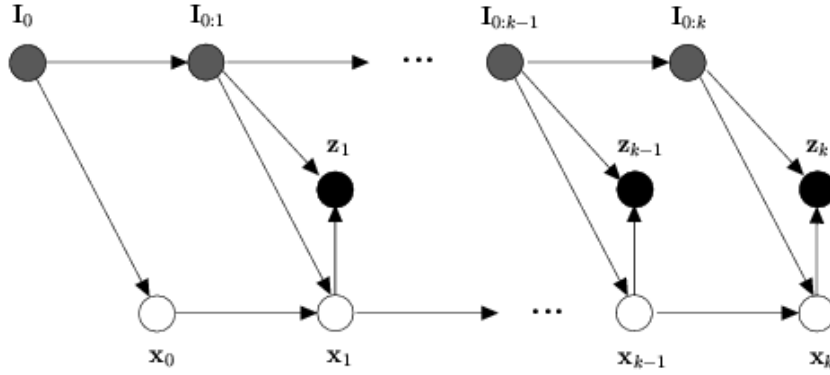


FIG. 5.2: Graphe de dépendance orienté de la chaîne de Markov cachée constituée des processus d'état $\mathbf{x}_{0:k}$, de mesure $\mathbf{z}_{1:k}$ et de la séquence d'images $\mathbf{I}_{0:k}$.

1. Étape de prédiction :

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) &= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) d\mathbf{x}_{k-1} \\ &= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1}) d\mathbf{x}_{k-1} \end{aligned} \quad (5.6)$$

2. Étape de correction :

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) &= \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})} \\ &= \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})}{\int p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) d\mathbf{x}_k} \end{aligned} \quad (5.7)$$

Une expression récursive du même type peut être construite pour la distribution *a posteriori* complète $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$.

Il est clair qu'une analogie très forte existe entre le modèle classique et le modèle conditionnel. Ce dernier constitue une généralisation qui a l'avantage d'explicitement les données images comme composantes du modèle. Ce formalisme permet de modéliser une éventuelle corrélation entre le modèle dynamique et le modèle de vraisemblance. Il constitue donc un cadre de filtrage permettant de traiter les modèles entièrement estimés sur la séquence.

De plus, ce modèle permet de clarifier la notion de mesure de l'état lors de la modélisation du problème de suivi dans une séquence d'images. En effet, comme l'ensemble des informations est contenu dans la séquence d'images, le cas idéal serait de considérer chaque image comme une mesure. Cependant, les images 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. Le fait de définir deux variables différentes, l'une pour les données images brutes, l'autre pour les données estimées sur les images utilisées comme observations de l'état permet d'éviter les confusions possibles.

Bien que le modèle conditionnel ait une structure très similaire au modèle classique, l'utilisation des filtres décrits dans la première partie pour son traitement (filtre de Kalman, filtre particulaire, etc.) n'est plus valide. En effet, ceux-ci ont été développés pour le modèle classique et ne prennent pas en compte explicitement la dépendance à la série d'images. Pour résoudre ce nouveau problème de trajectographie, nous proposons de dériver les filtres standards dans une version conditionnelle aux données image. Nous détaillerons deux instances de ce problème : une version linéaire (§5.2.1) et une version non linéaire (§5.2.2).

5.2 Filtres conditionnels

Nous proposons donc de modéliser le problème de trajectographie dans une séquence d'images par un système composé des équations dynamique et de mesure suivantes :

$$\begin{cases} \mathbf{x}_k = f_{k|\mathbf{I}}(\mathbf{x}_{k-1}, \mathbf{w}_k) \\ \mathbf{z}_k = h_{k|\mathbf{I}}(\mathbf{x}_k, \mathbf{v}_k) \end{cases} \quad (5.8)$$

L'indice $|\mathbf{I}$ indique la possible dépendance à la séquence. Ainsi, les fonctions $f_{k|\mathbf{I}}$ et $h_{k|\mathbf{I}}$ peuvent être estimées sur la séquence d'images $\mathbf{I}_{0:k}$ (de même, la fonction $f_{k+1|\mathbf{I}}$ pourra être estimée sur la séquence $\mathbf{I}_{0:k+1}$, etc.). Le bruit d'état $\mathbf{w}_k$ et le bruit de mesure $\mathbf{v}_k$ peuvent également dépendre de la réalisation observée de $\mathbf{I}_{0:k}$ (cette dépendance apparaît au travers de leur matrices de covariance notées respectivement $Q_{k|\mathbf{I}}$ et $R_{k|\mathbf{I}}$). En particulier, leur distribution est liée aux méthodes utilisées pour obtenir les mesures et estimer la dynamique.

Afin d'obtenir un calcul récursif de la distribution conditionnelle $p(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$, nous proposons de dériver les filtres standards dans une version conditionnelle aux données image. Dans une première étape, nous nous intéressons à la construction d'un **filtre linéaire conditionnel** (FLC). Celui-ci est construit à partir d'un estimateur déduit de l'estimateur linéaire de variance minimale. Dans un second temps, le cas non linéaire est étudié. Le **filtre non linéaire conditionnel** (FNLC) est défini sur le principe des méthodes particulières.

5.2.1 Filtre linéaire conditionnel

On considère un cas particulier du modèle (5.8) à savoir le système linéaire à bruits additifs suivant :

$$\begin{cases} \mathbf{x}_k = F_{k|\mathbf{I}} \mathbf{x}_{k-1} + \mathbf{b}_{k|\mathbf{I}} + \mathbf{w}_k \\ \mathbf{z}_k = H_{k|\mathbf{I}} \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (5.9)$$

Les matrices $F_{k|\mathbf{I}}$, $H_{k|\mathbf{I}}$ et le vecteur $\mathbf{b}_{k|\mathbf{I}}$ peuvent dépendre de la séquence $\mathbf{I}_{0:k}$. Ce vecteur correspond à une entrée du système. Les variables $\mathbf{w}_k$ et $\mathbf{v}_k$ sont supposées être des bruits blancs indépendants de moyenne nulle conditionnellement à $\mathbf{I}_{0:k}$ (éventuellement non gaussiens). Les variances conditionnelles correspondantes sont supposées connues (éventuellement estimées sur $\mathbf{I}_{0:k}$) et sont notées respectivement $Q_{k|\mathbf{I}}$ et $R_{k|\mathbf{I}}$.

On rappelle que dans le cas non conditionnel, la résolution d'un système linéaire à bruits additifs se fait grâce au filtre de Kalman. Une manière d'obtenir les équations de ce filtre consiste à utiliser l'estimateur linéaire non biaisé de variance minimale notée $E^* [\mathbf{x}_k | \mathbf{z}_{1:k}]$. Dans le cas gaussien, cet estimateur est équivalent à l'espérance conditionnelle $\mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}]$ de la distribution gaussienne $p(\mathbf{x}_k | \mathbf{z}_{1:k})$.

Le filtre linéaire conditionnel (FLC) est naturellement déduit d'une extension de l'estimateur linéaire non biaisé de variance minimale, adaptée à notre contexte. Cet estimateur est appelé **l'estimateur linéaire conditionnel de variance minimale**. Il permet l'obtention de l'estimation à l'ordre 2 de $p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. Cette description est optimale lorsque les bruits du modèle sont conditionnellement gaussiens, i.e. :

$$\mathbf{w}_k | \mathbf{I}_{0:k} \rightsquigarrow \mathcal{N}(\mathbf{w}_k ; 0, Q_{k|\mathbf{I}}) \quad (5.10)$$

$$\mathbf{v}_k | \mathbf{I}_{0:k} \rightsquigarrow \mathcal{N}(\mathbf{v}_k ; 0, R_{k|\mathbf{I}}) \quad (5.11)$$

Estimateur linéaire conditionnel de variance minimale

Par souci de clarté, nous rappelons la définition de l'estimateur linéaire non biaisé de variance minimale (BLUE : "best linear unbiased estimator"). Quelques-unes de ses propriétés sont données en annexe B.

Définition 1 Soit $(X, Z) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ un couple de vecteurs aléatoires. On note $E^* [X|Z]$ l'estimateur linéaire non biaisé de variance minimale. Cet estimateur est défini par :

$$E^* [X|Z] = A^* Z + b^* \quad (5.12)$$

avec A^* et b^* tels que

$$\mathbb{E}[\|X - A^* Z - b^*\|^2] \leq \mathbb{E}[\|X - AZ - b\|^2] \quad \forall A \in \mathbb{R}^{n_x \times n_z}, b \in \mathbb{R}^{n_x}. \quad (5.13)$$

Comme décrit précédemment, l'estimateur linéaire conditionnel de variance minimale est déduit du BLUE. On le définit comme suit :

Définition 2 Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires. On définit $E_W^* [X|Z]$ comme étant le meilleur estimateur (au sens de la norme L_2) de X conditionnellement à W , linéaire en Z , i.e.

$$E_W^* [X|Z] = A_W^* Z + b_W^* \quad (5.14)$$

avec A_W^* et b_W^* tels que

$$\mathbb{E}[\|X - A_W^* Z - b_W^*\|^2 | W] \leq \mathbb{E}[\|X - AZ - b\|^2 | W] \quad \forall A \in \mathbb{R}^{n_x \times n_z}, b \in \mathbb{R}^{n_x}. \quad (5.15)$$

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

En introduisant la notation suivante :

$$\Sigma_{X,Z|W} \triangleq \text{cov}(X, Z|W) = \mathbb{E}[XZ^t|W] - \mathbb{E}[X|W] \mathbb{E}[Z^t|W], \quad (5.16)$$

une expression de ce nouvel estimateur est donnée par :

$$E_W^*[X|Z] = \mathbb{E}[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]). \quad (5.17)$$

La preuve de ce résultat important n'est pas directe. Nous en donnons ici les grandes lignes. La preuve complète est décrite en annexe B.

PREUVE : _____

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

$$\begin{aligned} \mathbb{E}[\|Y\|^2|W] &= \mathbb{E}[Y^t Y|W] = \mathbb{E}[tr\{YY^t\}|W] \\ &= tr\{cov(Y, Y|W)\} + \mathbb{E}[Y|W]^t \mathbb{E}[Y|W], \end{aligned} \quad (5.18)$$

où tr signifie la trace de la matrice entre accolades, et en rappelant que pour trois vecteurs aléatoires arbitraires X , Z et W , on note :

$$\Sigma_{X,Z|W} \triangleq cov(X, Z|W) \quad (5.19)$$

$$= \mathbb{E}[(X - \mathbb{E}[X|W]) (Z - \mathbb{E}[Z|W])^t | W] \quad (5.20)$$

$$= \mathbb{E}[X Z^t | W] - \mathbb{E}[X|W] \mathbb{E}[Z|W]^t, \quad (5.21)$$

après quelques manipulations, on peut écrire :

$$\begin{aligned} &E[\|X - A Z - b\|^2 | W] \\ &= tr\{(A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1}) \Sigma_{Z,Z|W} (A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1})^t\} \\ &+ \|\mathbb{E}[X|W] - A \mathbb{E}[Z|W] - b\|^2 + tr\{\Sigma_{X,X|W} - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \Sigma_{Z,X|W}\}. \end{aligned} \quad (5.22)$$

Ces trois termes sont semi-définis positifs. $\mathbb{E}[\|X - AZ - b\|^2 | W]$ atteint donc son minimum pour :

$$\begin{aligned} A_W^* &= \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1}, \\ b_W^* &= \mathbb{E}[X|W] - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \mathbb{E}[Z|W]. \end{aligned}$$

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

$$E_W^*[X|Z] = \mathbb{E}[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]). \quad (5.23)$$

_____ □

Certaines propriétés du BLUE peuvent être étendues à l'estimateur linéaire conditionnel de variance minimale. En particulier, il vérifie les théorèmes suivants, dont les preuves sont données en annexe B. Ces propriétés seront nécessaires à la dérivation du FLC.

Propriété 1 (propriété à être conditionnellement non biaisé) Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires, alors

$$\mathbb{E}[X - E_W^*[X|Z] | W] = 0 \quad (5.24)$$

Propriété 2 (conditionnement par des vecteurs aléatoires non corrélés) Soit $(X, Z_1, Z_2, \dots, Z_k, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \dots \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un ensemble de vecteurs aléatoires avec $Z_1, Z_2, \dots, Z_k$ non corrélés conditionnellement à W (i.e. $\Sigma_{Z_i Z_j | W} = 0$ pour $i \neq j$), alors

$$\begin{aligned} E_W^*[X | Z_1, Z_2, \dots, Z_k] &= E_W^*[X | Z_1] + E_W^*[X | Z_2] + \dots \\ &\quad + E_W^*[X | Z_k] - (k-1) \mathbb{E}[X | W] \end{aligned} \quad (5.25)$$

Propriété 3 (changement de variable conditionnelle) Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires, et soit $Y = C Z + d$ (avec C non singulière), alors

$$E_W^*[X | Y] = E_W^*[X | Z] \quad (5.26)$$

Propriété 4 (principe d'orthogonalité) Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires, alors

$$\mathbb{E}[(X - E_W^*[X | Z]) Z^t | W] = 0 \quad (5.27)$$

Équations du filtre linéaire conditionnel

On s'intéresse à un modèle linéaire de la forme décrite précédemment (5.9). 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. $\hat{\mathbf{x}}_{k+1|k}$ et $\Sigma_{k+1|k}$ représentent les deux premiers moments de la distribution *a posteriori* $p(\mathbf{x}_{k+1} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k+1})$.

Une expression récursive de ces deux moments peut être obtenue sur le même principe que la dérivation du filtre de Kalman effectuée à partir de l'estimateur BLUE et des propriétés d'indépendance de la séquence d'innovation [Anderson 79]. Dans notre contexte conditionnel à la séquence, l'estimateur mis en jeu est l'estimateur linéaire conditionnel de variance minimale. Par le biais de manipulations similaires au cas classique, mais en considérant avec attention les expressions conditionnelles induites par ce nouvel estimateur, on obtient les équations du FLC. La démonstration complète est disponible en annexe C. Ces formules récursives sont :

$$\hat{\mathbf{x}}_{k+1|k} = F_{k+1|\mathbf{I}} \hat{\mathbf{x}}_{k|k-1} + \mathbf{b}_{k+1|\mathbf{I}} + \tilde{K}_k (\mathbf{z}_k - H_{k|\mathbf{I}} \hat{\mathbf{x}}_{k|k-1}) \quad (5.28)$$

$$\begin{aligned} \Sigma_{k+1|k} &= (F_{k+1|\mathbf{I}} - \tilde{K}_k H_{k|\mathbf{I}}) \Sigma_{k|k-1} (F_{k+1|\mathbf{I}} - \tilde{K}_k H_{k|\mathbf{I}})^t \\ &\quad + Q_{k+1|\mathbf{I}} + \tilde{K}_k R_{k|\mathbf{I}} \tilde{K}_k^t \end{aligned} \quad (5.29)$$

où la matrice $\tilde{K}_k$ est définie à l'aide la matrice de gain du filtre de Kalman K_k :

$$\tilde{K}_k = F_{k+1|\mathbf{I}} K_k \quad (5.30)$$

$$K_k = (\Sigma_{k|k-1} H_{k|\mathbf{I}}^t) (H_{k|\mathbf{I}} \Sigma_{k|k-1} H_{k|\mathbf{I}}^t + R_{k|\mathbf{I}})^{-1} \quad (5.31)$$

Afin de distinguer l'étape de prédiction de l'étape de correction, les équations (5.28) à (5.31) peuvent être écrites en deux étapes :

- Étape de prédiction

$$\hat{\mathbf{x}}_{k|k-1} = F_{k|\mathbf{I}} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}_{k|\mathbf{I}} \quad (5.32)$$

$$\Sigma_{k|k-1} = F_{k|\mathbf{I}} \Sigma_{k-1|k-1} F_{k|\mathbf{I}}^t + Q_{k|\mathbf{I}} \quad (5.33)$$

- Étape de correction

$$K_k = \Sigma_{k|k-1} H_{k|\mathbf{I}}^t (H_{k|\mathbf{I}} \Sigma_{k|k-1} H_{k|\mathbf{I}}^t + R_{k|\mathbf{I}})^{-1} \quad (5.34)$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k (\mathbf{z}_k - H_{k|\mathbf{I}} \hat{\mathbf{x}}_{k|k-1}) \quad (5.35)$$

$$\Sigma_{k|k} = (Id - K_k H_{k|\mathbf{I}}) \Sigma_{k|k-1} \quad (5.36)$$

Remarquons que ces équations sont similaires aux équations du filtre de Kalman (§1.1, p. 1.1). Nous avons ainsi démontré que l'utilisation en pratique de celles-ci est valide lorsque les entrées du système sont estimées sur les images à chaque itération. Cependant, au niveau théorique, une différence importante existe entre le filtre de Kalman et le filtre linéaire conditionnel : ils fournissent une estimation à l'ordre 2 de, respectivement, $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ et $p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$.

Fenêtre de validation

Afin de réduire le temps de calcul, il peut être intéressant d'utiliser une région de validation à chaque itération. Dans le cadre d'un suivi à partir d'une séquence d'images, une telle région permet de réduire la recherche de l'observation future à une partie de l'image. Rappelons que dans le cas classique décrit dans la première partie (§3.1, p. 3.1), cette région est définie par la distribution $p(\mathbf{z}_k | \mathbf{z}_{1:k-1})$. Dans notre contexte conditionnel, elle est construite à partir de la connaissance de $p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})$.

De manière similaire au cas classique, cette distribution est approchée par une loi normale (cette expression est exacte si les bruits du modèle vérifient (5.10) (5.11)) :

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; \hat{\mathbf{z}}_{k|k-1}, S_{k|\mathbf{I}}),$$

où $\hat{\mathbf{z}}_{k|k-1} = H_{k|\mathbf{I}} \hat{\mathbf{x}}_{k|k-1}$ désigne la mesure prédite et $S_{k|\mathbf{I}} = H_{k|\mathbf{I}} \Sigma_{k|k-1} H_{k|\mathbf{I}}^t + R_{k|\mathbf{I}}$ la covariance de l'innovation $\tilde{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}$ conditionnellement à la séquence. La fenêtre de validation $valid_k$ est l'ellipsoïde défini par :

$$valid_k = \{\mathbf{z}_k : (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^t S_{k|\mathbf{I}}^{-1} (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}) \leq \gamma\}. \quad (5.37)$$

Le seuil γ est choisi de telle sorte que la probabilité de trouver la mesure correcte dans la région $valid_k$ soit égale à une valeur P_g souvent fixée à 0.99. Notons que la fenêtre de validation dépend comme dans le cas classique de la covariance de l'erreur d'estimation, mais aussi de la séquence d'images au travers des matrices $R_{k|\mathbf{I}}$ et $H_{k|\mathbf{I}}$.

Le filtre linéaire conditionnel est résumé par le diagramme 5.3. Les cadres doubles représentent les étapes réalisées sur la séquence d'images. Cet algorithme est similaire au

filtre de Kalman, dont l'utilisation est maintenant justifiée face à des modèles entièrement estimés sur la séquence. Il fournit ainsi un cadre de trajectographie dans les images permettant de mettre en compétition de façon probabiliste deux processus d'estimation. Nous verrons dans la troisième partie de ce document comment cette compétition est mise en pratique pour la réalisation d'un algorithme robuste de suivi de points d'intérêt.

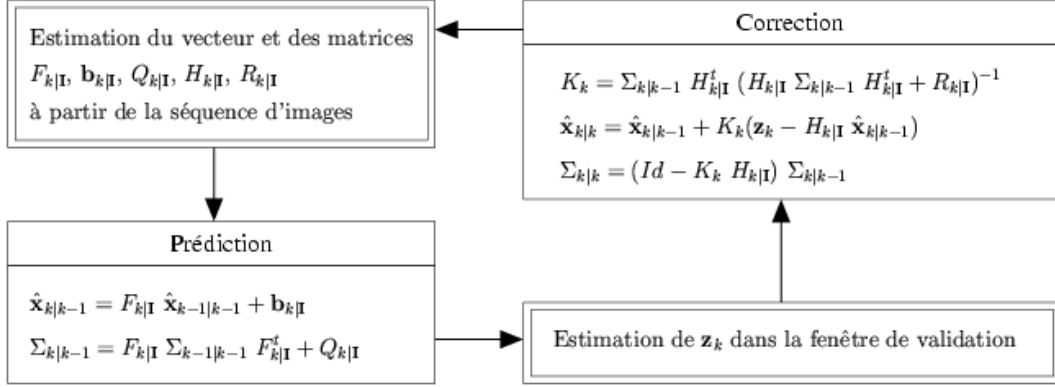


FIG. 5.3: Diagramme représentant les étapes du filtre linéaire conditionnel pour la résolution du système (5.9). Les cadres doubles représentent les étapes réalisées sur la séquence d'images.

5.2.2 Filtre non linéaire conditionnel

Lorsque le système étudié présente des non linéarités, les équations récurrentes du filtre optimal bayésien n'ont pas d'expressions calculables. Pour contourner cette difficulté, nous avons vu en première partie qu'une solution intéressante consiste en l'utilisation de méthodes séquentielles de Monte Carlo (cf. chapitre 2, p. 2). Rappelons que le principe de ces méthodes est d'approcher la distribution *a posteriori* de l'état conditionnellement à l'ensemble des données disponibles par une somme pondérée de mesures de Dirac.

Comme dans le cas linéaire, nous nous intéressons dans ce chapitre à la construction d'une méthode particulière valide dans le cas d'un modèle entièrement estimé sur la séquence d'images. Comme précédemment, un tel algorithme appelé **filtre non linéaire conditionnel** (FNLC) est déduit d'un algorithme adapté au modèle classique. La dépendance de l'ensemble du système à la séquence d'images est prise en compte via un conditionnement par rapport aux données image.

On considère le modèle conditionnel général que nous rappelons ici :

$$\begin{cases} \mathbf{x}_k = f_{k|I}(\mathbf{x}_{k-1}, \mathbf{w}_k) \\ \mathbf{z}_k = h_{k|I}(\mathbf{x}_k, \mathbf{v}_k), \end{cases} \quad (5.38)$$

où aucune hypothèse de linéarité n'est faite sur les fonctions $f_{k|I}$, $h_{k|I}$, et où $\mathbf{w}_k$, $\mathbf{v}_k$ sont des bruits blancs mutuellement indépendants, éventuellement non gaussiens, dont les covariances peuvent être estimées sur les images. On rappelle que l'indice I signifie une

possible dépendance par rapport à la séquence d'images. Dans le cas où les bruits du modèle admettent des densités, le système est décrit par les distributions $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k})$, $p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k})$ et la distribution initiale $p(\mathbf{x}_0|\mathbf{I}_0)$.

Alors que dans le cas d'un modèle classique la distribution d'intérêt est $p(\mathbf{x}_k|\mathbf{z}_{1:k})$, dans notre contexte de suivi à partir d'une séquence, nous cherchons à construire une estimation de $p(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. Dans un premier temps, nous nous intéressons à l'estimation de $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ et à l'estimation d'intégrales de la forme :

$$\mathbb{E}_{p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})}[\phi(\mathbf{x}_{0:k})] = I(\phi(\mathbf{x}_{0:k}), \mathbf{I}_{0:k}) = \int \phi(\mathbf{x}_{0:k}) p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}) d\mathbf{x}_{0:k} \quad (5.39)$$

Nous reprenons l'ensemble de la dérivation du filtre particulaire pour ce cas particulier. Comme le conditionnement par la séquence d'images n'entraîne aucune difficulté, nous ne détaillons pas l'ensemble des équations.

L'intégrale (5.39) n'a pas d'expression analytique dans le cas général. L'utilisation de l'échantillonnage pondéré permet d'approcher cette intégrale en introduisant la fonction de proposition $\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. À partir d'un ensemble $\{\mathbf{x}_{0:k}^{(i)}\}$ de N particules tirées selon la fonction de proposition, le calcul des N poids normalisés associés $\tilde{w}_k^{(i)}$ permet d'approcher l'espérance conditionnelle par une somme finie :

$$I_N(\phi(\mathbf{x}_{0:k}), \mathbf{I}_{0:k}) = \sum_{i=1}^N \tilde{w}_k^{(i)} \phi(\mathbf{x}_{0:k}^{(i)}). \quad (5.40)$$

Les poids non normalisés sont donnés par :

$$w_k^{(i)} = \frac{p(\mathbf{x}_{0:k}^{(i)}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})}{\pi(\mathbf{x}_{0:k}^{(i)}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})}. \quad (5.41)$$

Autrement dit, on peut approcher la loi objectif $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ par :

$$p_N(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}) = \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_{0:k}^{(i)}}(\mathbf{x}_{0:k}), \quad (5.42)$$

et la loi de filtrage $p(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ par :

$$p_N(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}) = \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k). \quad (5.43)$$

Afin de construire une solution itérative pour l'estimation de l'espérance conditionnelle, on suppose une forme récursive de la fonction de proposition. Comme dans le cas classique, cette formulation introduit une hypothèse de dépendance causale de l'état aux mesures. Notons qu'une hypothèse du même type est introduite sur les images.

$$\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}) = \pi(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1}) \pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k}). \quad (5.44)$$

Ainsi, à chaque itération, chaque particule $\mathbf{x}_{0:k-1}^{(i)}$ est augmentée d'un nouvel état $\mathbf{x}_k^{(i)}$ tiré selon la fonction d'importance $\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. Une telle expression conduit à une équation récursive pour le calcul des poids :

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \mathbf{I}_{0:k}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})}{\pi(\mathbf{x}_k^{(i)} | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k})}. \quad (5.45)$$

Néanmoins, de manière similaire au cas non conditionnel, cette approximation conduit à une augmentation de la variance des poids dans le temps. Il est alors nécessaire d'introduire une étape de sélection des particules pour minimiser ce phénomène. En outre, l'algorithme sera d'autant plus efficace que la fonction d'importance $\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ est proche de $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$.

Algorithme 15 Filtre non linéaire conditionnel pour la résolution du système (5.38)

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0 | \mathbf{I}_0)$, et fixer $w_0^{(i)} = 1/N$.

pour $k = 1, 2, \dots$

- **échantillonnage pondéré séquentiel :**

1. échantillonnage :

pour $i = 1 \dots N$, générer $\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ et poser $\mathbf{x}_{0:k}^{(i)} \triangleq (\mathbf{x}_{0:k-1}^{(i)}, \mathbf{x}_k^{(i)})$.

2. mise à jour des poids d'importance :

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

3. normalisation des poids $\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}}$

- **estimations de Monte Carlo :**

le nuage pondéré permet d'approcher les lois *a posteriori* :

$$p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) \simeq \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_{0:k}^{(i)}}(\mathbf{x}_{0:k}) \quad p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) \simeq \sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k)$$

- **rééchantillonnage adaptatif**

si nécessaire, sélectionner N particules parmi $\{\mathbf{x}_{0:k}^{(i)}\}$ en fonction des poids $\tilde{w}_k^{(i)}$

Les étapes du FNLC sont résumées par l'algorithme 15. Cette méthode correspond à un filtre particulière dédié à l'estimation des distributions conditionnelles à l'ensemble des

données (mesures et images) $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ et $p(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. Elle permet la résolution de systèmes de type (5.38) pouvant être entièrement estimés sur la séquence d'images.

Notons que l'ensemble des stratégies d'échantillonnage et de rééchantillonnage décrites en partie 1 peuvent être appliquées dans notre contexte de suivi dans une séquence d'images. En particulier, il peut être montré que la fonction d'importance qui minimise la variance des poids, conditionnellement à $\mathbf{x}_{0:k-1}^{(i)}$, $\mathbf{z}_{1:k}$ et $\mathbf{I}_{0:k}$ est de la forme suivante (la démonstration est similaire au cas non conditionnel) :

$$\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) = p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) \quad (5.46)$$

Le choix de cette fonction d'importance optimale implique la formule de récursivité des poids :

$$w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}). \quad (5.47)$$

Dans la section suivante, nous nous intéresserons à deux types de modèles permettant de construire des expressions analytiques des deux distributions mises en jeu dans (5.46) et (5.47).

Remarque : La méthode de réduction de variance par rao-blackwellisation (c.f. §2.5, p. 55) est applicable dans notre contexte conditionnel. En effet, en supposant que les variables d'état $\mathbf{x}_k$ peuvent se diviser en deux parties $\mathbf{r}_k$ et $\mathbf{y}_k$ telles que la distribution de prédiction soit de la forme :

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = p(\mathbf{r}_k, \mathbf{y}_k|\mathbf{r}_{k-1}, \mathbf{y}_{k-1}, \mathbf{I}_{0:k}) = p(\mathbf{r}_k|\mathbf{r}_{k-1}) p(\mathbf{y}_k|\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1}, \mathbf{I}_{0:k}), \quad (5.48)$$

et en supposant que la distribution *a posteriori* $p(\mathbf{y}_{0:k}|\mathbf{r}_{0:k}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ puisse être estimée de façon optimale, un filtre conditionnel non linéaire rao-blackwellisé peut être mise en place pour l'estimation de la distribution objectif qui s'écrit :

$$p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}) = p(\mathbf{y}_{0:k}|\mathbf{r}_{0:k}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}). \quad (5.49)$$

5.3 Modèles conditionnels non linéaires gaussiens : deux cas particuliers

Dans le chapitre précédent, nous avons étudié les principales modélisations existantes pour résoudre le problème de trajectographie dans une séquence d'images. Dans le cas d'un algorithme de suivi général, nous avons vu que l'approche la plus répandue consiste à utiliser une dynamique peu informative et une équation de mesure hautement non linéaire décrivant avec précision le lien entre l'état et la mesure observée. Ce type de modélisation repose alors essentiellement sur la qualité de la vraisemblance des données. Nécessitant un filtre particulière, elle a l'inconvénient de ne pas autoriser l'utilisation de la fonction d'importance optimale pour faire évoluer les particules.

Dans cette section, nous proposons une modélisation alternative. Cette modélisation se place dans le cadre d'une approche *Top-Down* (§4.3.1), qui nécessite un module de détection de l'entité suivie. Lorsque la mesure obtenue en sortie de ce module est du même type que le vecteur d'état, il est possible d'inférer une équation de mesure linéaire. Une telle modélisation de la vraisemblance peut paraître limitée. Cependant, nous avons vu au

début de ce chapitre qu'il est valide au niveau théorique de construire la loi de prédiction sur la séquence. Associer une telle dynamique, très informative, à une équation de mesure linéaire permet de pallier la modélisation peu précise induite par cette dernière. De plus, dans le cas où les bruits du modèle sont additifs gaussiens, il est possible d'utiliser la fonction d'importance optimale.

Dans un premier temps, nous décrivons un tel modèle constitué d'une équation de mesure linéaire, d'une dynamique non linéaire estimée sur la séquence et de bruits additifs gaussiens. Dans un second temps, nous nous intéressons à l'extension de ce modèle au cas où plusieurs mesures sont disponibles à chaque itération. Le problème d'association des données est alors résolu au travers d'une modélisation multimodale de la vraisemblance. Sous certaines conditions, le système construit permet l'utilisation de la fonction d'importance optimale. Enfin, dans un troisième temps, la définition de fenêtres de validation pour ces deux modèles est décrite.

Les modèles étudiés dans cette section reposent sur une dynamique informative, et une équation de mesure très simple. Ils sont valides théoriquement grâce au contexte conditionnel. Bien que cette modélisation ne décrive pas avec précision le lien entre l'état et les observations du système, elle en permet une estimation efficace (par le biais de la fonction d'importance optimale). En effet, nous pensons que de tels modèles sont intéressants car ils constituent un équilibre entre la complexité du modèle décrivant le système dynamique et la qualité de l'algorithme permettant de traiter ce modèle. Dans la fin de ce chapitre, nous étudierons ces modèles du point de vue théoriques. Dans la troisième partie de ce document, nous nous intéresserons à leur mise en œuvre pour du suivi de points caractéristiques et d'objets plans.

5.3.1 Vraisemblance monomodale

On considère le système suivant, composé d'une équation dynamique non linéaire avec un bruit additif conditionnellement gaussien :

$$\mathbf{x}_k = f_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + \mathbf{w}_k, \quad \mathbf{w}_k | \mathbf{I}_{0:k} \rightsquigarrow \mathcal{N}(\mathbf{w}_k ; 0, Q_{k|\mathbf{I}}) \quad (5.50)$$

et d'une équation de mesure linéaire avec un bruit additif conditionnellement gaussien :

$$\mathbf{z}_k = H_{k|\mathbf{I}} \mathbf{x}_k + \mathbf{v}_k, \quad \mathbf{v}_k | \mathbf{I}_{0:k} \rightsquigarrow \mathcal{N}(\mathbf{v}_k ; 0, R_{k|\mathbf{I}}). \quad (5.51)$$

On rappelle que l'indice $|\mathbf{I}$ indique une éventuelle estimation sur la séquence. On suppose que les bruits sont mutuellement indépendants. Ce système peut s'écrire sous la forme :

$$\begin{cases} p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k ; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}) \\ p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; H_{k|\mathbf{I}} \mathbf{x}_k, R_{k|\mathbf{I}}). \end{cases} \quad (5.52)$$

Ce modèle présente des non linéarités qui nécessitent l'utilisation de méthodes particulières. Cependant, il rend l'utilisation de la fonction d'importance optimale possible, ce qui est un avantage important. En effet, des expressions de cette fonction $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k})$, et

de la distribution nécessaire au calcul récursif des poids $p(\mathbf{z}_k|\mathbf{x}_{k-1})$ sont disponibles. De façon similaire au cas non conditionnel décrit p. 47, en remarquant que :

$$p(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \int p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) d\mathbf{x}_k \quad (5.53)$$

et

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k})}{p(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k})}, \quad (5.54)$$

on obtient les expressions des distributions utiles à l'utilisation de la fonction d'importance optimale :

$$p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I}} + H_{k|\mathbf{I}} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I}}^t) \quad (5.55)$$

et

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|\mathbf{I}}, \Sigma_{k|\mathbf{I}}) \quad (5.56)$$

avec

$$\Sigma_{k|\mathbf{I}} = (Q_{k|\mathbf{I}}^{(i)-1} + H_{k|\mathbf{I}}^t R_{k|\mathbf{I}}^{-1} H_{k|\mathbf{I}})^{-1} \quad (5.57)$$

$$\mathbf{m}_{k|\mathbf{I}} = \Sigma_{k|\mathbf{I}} (Q_{k|\mathbf{I}}^{(i)-1} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}) + H_{k|\mathbf{I}}^t R_{k|\mathbf{I}}^{-1} \mathbf{z}_k). \quad (5.58)$$

Ce résultat est identique au cas classique. Notons que dans notre contexte conditionnel, l'équation d'état est estimée à chaque itération, et pour chaque particule. Cela explique l'exposant (i) indiquant le numéro de la particule considérée sur la matrice $Q_{k|\mathbf{I}}$. Ce modèle est valide dans le cas où une unique mesure est disponible à chaque itération. Dans le paragraphe suivant, une extension de ce système dédiée au suivi en environnement bruité est décrite.

5.3.2 Vraisemblance multimodale

Considérer une unique mesure par itération peut conduire à une mauvaise estimation. En particulier, dans le cas de séquences d'images très bruitées ou avec un fouillis¹ d'arrière plan important, le module de détection peut être perturbé par les ambiguïtés générées et fournir une mauvaise observation de l'état. Un moyen de contourner ce problème est de considérer plusieurs mesures simultanément.

Nous présentons ici une extension du système conditionnel non linéaire gaussien précédent pour le problème de trajectographie dans une séquence d'images bruitées. On considère ainsi la même densité de prédiction estimée sur les images :

$$\mathbf{x}_k = f_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + \mathbf{w}_k, \quad \mathbf{w}_k|\mathbf{I}_{0:k} \sim \mathcal{N}(\mathbf{w}_k; 0, Q_{k|\mathbf{I}}). \quad (5.59)$$

Le problème d'association des mesures est pris en compte au travers de la vraisemblance. On rappelle les notations utilisées : au temps k , on note $\mathbf{z}_k$ le vecteur de mesures constitué des M_k observations disponibles $\{\mathbf{z}_{k,m}\}_{m=1\dots M_k}$. Celles-ci ont été obtenues dans une fenêtre

¹Le terme "fouillis" désigne la présence de nombreuses zones de l'arrière plan de l'image ambiguës par rapport à l'entité suivie.

de validation de volume V_k (dont nous détaillerons l'obtention dans le paragraphe suivant) ; et $\Phi_k \in \{0, \dots, M_k\}$ la variable aléatoire associée aux évènements d'associations :

$$\Phi_k = \begin{cases} j & \text{si } \mathbf{z}_{k,j} \text{ est générée par l'état caché,} \\ 0 & \text{si toutes les observations sont des fausses alarmes.} \end{cases} \quad (5.60)$$

On suppose qu'une mesure peut provenir d'un objet ou être une fausse alarme, et on suppose qu'un objet peut être à l'origine d'au plus une mesure. Ces hypothèses d'association impliquent la dépendance des associations des différentes observations. Les évènements $\{\Phi_k = j\}_{j=0 \dots M_k}$ sont alors exhaustifs et mutuellement exclusifs. Par l'utilisation du théorème des probabilités totales, la vraisemblance s'exprime :

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \sum_{j=0}^{M_k} \beta_{k,j} p(\mathbf{z}_k | \Phi_k = j, \mathbf{x}_k, \mathbf{I}_{0:k}) \quad (5.61)$$

$$\text{où } \beta_{k,j} \triangleq p(\Phi_k = j | \mathbf{x}_k, \mathbf{I}_{0:k}) \quad j = 0 \dots M_k \quad (5.62)$$

En supposant d'une part l'indépendance des mesures $\mathbf{z}_{k,m}$ conditionnellement à $\mathbf{x}_k, \mathbf{I}_{0:k}$ et à l'évènement $\Phi_k = j$, on obtient :

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \sum_{j=0}^{M_k} \beta_{k,j} \prod_{m=0}^{M_k} p(\mathbf{z}_{k,m} | \Phi_k = j, \mathbf{x}_k, \mathbf{I}_{0:k}) \quad (5.63)$$

En supposant d'autre part que la "vraie" mesure est reliée linéairement à l'état caché avec un bruit additif gaussien, et que les fausses mesures sont réparties uniformément dans la fenêtre de validation, la vraisemblance s'écrit :

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \beta_{k,0} V_k^{-M_k} + \sum_{j=1}^{M_k} [\beta_{k,j} V_k^{1-M_k} \mathcal{N}(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j})] \quad (5.64)$$

Afin d'estimer les probabilités d'association $\beta_{k,j} = p(\Phi_k = j | \mathbf{x}_k, \mathbf{I}_{0:k})$, nous proposons l'approximation suivante :

$$\beta_{k,j} \simeq p(\Phi_k = j | \mathbf{I}_{0:k}), \quad (5.65)$$

ce qui signifie que ces probabilités d'association peuvent être estimées directement sur les images. De plus, on fait l'approximation importante que la probabilité d'observer uniquement des fausses alarmes est nulle. Cela se traduit par $\beta_{k,0} = 0$, ce qui implique l'expression suivante de la vraisemblance, qui correspond alors à un mélange de lois gaussiennes :

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j}), \quad (5.66)$$

où les matrices $H_{k|\mathbf{I},j}$ et $R_{k|\mathbf{I},j}$ sont associées à la jème mesure. Ce choix diffère par rapport aux approches classiques. Il peut être problématique dans la situation où aucune bonne mesure n'est observée, ce qui est le cas lors d'occlusion de l'entité suivie. Cependant, cette déficience potentielle peut être compensée de manière avantageuse par des estimations en-ligne des covariances de bruits des mesures $R_{k|\mathbf{I},j}$. En cas d'occlusion, si ces covariances

sont fixées à $\infty \times Id$, aucune mesure ne sera prise en compte, ce qui équivaut à poser la probabilité d'avoir aucune observation valable non nulle. Enfin, ce choix permet de considérer un modèle qui autorise l'utilisation de la fonction d'importance optimale dans la phase d'échantillonnage des particules. Ce dernier point est essentiel, et représente un avantage important de notre approche.

On considère ainsi le système suivant :

$$\begin{cases} p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k ; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}) \\ p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j}). \end{cases} \quad (5.67)$$

On rappelle que l'utilisation de la fonction d'importance optimale nécessite la connaissance de la distribution $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k})$ pour échantillonner les particules et de la distribution $p(\mathbf{z}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$ pour la mise à jour des poids.

En appliquant l'égalité (5.53), on obtient une expression de la distribution utile au calcul des poids. Cette expression correspond à un mélange de lois gaussiennes (la démonstration est disponible en annexe D) :

$$\begin{aligned} p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) \\ = V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}\left(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I},j} + H_{k|\mathbf{I},j} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I},j}^t\right). \end{aligned} \quad (5.68)$$

En utilisant les expressions du système (5.67), et l'expression (5.68) dans l'égalité (5.54), on obtient la formulation suivante de la fonction d'importance optimale (la démonstration est également disponible en annexe D) :

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \sum_{j=1}^{M_k} \beta_{k,j} \frac{\alpha_{k,j}}{S_k} \mathcal{N}(\mathbf{x}_k ; \mathbf{m}_{k|\mathbf{I},j}, \Sigma_{k|\mathbf{I},j}) \quad (5.69)$$

avec

$$S_k = \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}\left(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I},j} + H_{k|\mathbf{I},j} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I},j}^t\right) \quad (5.70)$$

$$\Sigma_{k|\mathbf{I},j} = \left(Q_{k|\mathbf{I}}^{(i)}{}^{-1} + H_{k|\mathbf{I},j}^t R_{k|\mathbf{I},j}^{-1} H_{k|\mathbf{I},j}\right)^{-1} \quad (5.71)$$

$$\mathbf{m}_{k|\mathbf{I},j} = \Sigma_{k|\mathbf{I},j} \left(Q_{k|\mathbf{I}}^{(i)}{}^{-1} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}) + H_{k|\mathbf{I},j}^t R_{k|\mathbf{I},j}^{-1} \mathbf{z}_{k,j}\right) \quad (5.72)$$

$$\alpha_{k,j} = C \exp\left(-\frac{1}{2} \left[\|f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})\|_{Q_{k|\mathbf{I}}^{(i)}{}^{-1}}^2 + \|\mathbf{z}_{k,j}\|_{R_{k|\mathbf{I},j}^{-1}}^2 - \|\mathbf{m}_{k|\mathbf{I},j}\|_{\Sigma_{k|\mathbf{I},j}^{-1}}^2\right]\right) \quad (5.73)$$

$$\text{où } C = (2\pi)^{-\frac{n_z}{2}} |\Sigma_{k|\mathbf{I},j}|^{\frac{1}{2}} |Q_{k|\mathbf{I}}^{(i)}|^{-\frac{1}{2}} |R_{k|\mathbf{I},j}|^{-\frac{1}{2}} \quad (5.74)$$

On rappelle que n_z est la dimension de la mesure $\mathbf{z}_k$. L'expression (5.69) est un mélange de lois gaussiennes (la preuve est donnée en annexe D). Bien que les expressions (5.68),

(5.69) utilisées dans l'algorithme du filtre non linéaire conditionnel paraissent complexes, elles conduisent en pratique à un algorithme efficace et simple à implémenter puisque les densités mises en jeu sont uniquement des lois normales.

Le système conditionnel décrit ici est très intéressant car il permet à la fois de considérer plusieurs observations de façon simultanée, et d'autoriser l'utilisation de la fonction d'importance optimale. En pratique, cela permet de diffuser les particules de manière à les faire visiter les zones de forte vraisemblance, définies par les différentes mesures. Dans le cadre du suivi dans une séquence, un tel système permet de lever les ambiguïtés générées par un fouillis d'arrière plan par exemple.

Remarque : Pour définir la vraisemblance (5.66) prenant en compte différentes observations de l'état, nous avons fait l'hypothèse classique d'indépendance des mesures $\mathbf{z}_{k,m}$ conditionnellement à $\mathbf{x}_k$, $\Phi_k = j$, (et $\mathbf{I}_{0:k}$ dans notre contexte). Cette hypothèse d'indépendance n'est pas exacte en traitement d'image. En effet, dans ce contexte, les fausses mesures sont souvent dues à des ambiguïtés, c'est-à-dire qu'elles correspondent à des zones de l'image qui sont similaires. Une telle hypothèse peut alors apparaître abusive. L'étude d'un modèle prenant en compte cette corrélation est une perspective intéressante.

En plus de permettre l'utilisation d'une phase d'échantillonnage efficace par la fonction d'importance optimale, les deux systèmes conditionnels non linéaires gaussiens décrits dans cette section permettent d'inférer simplement une fenêtre de validation des mesures.

5.3.3 Fenêtre de validation

Tout comme dans le cas linéaire, il est important de construire une fenêtre de validation qui correspond à une zone de l'image où chercher la future observation $\mathbf{z}_k$. Cette fenêtre est définie à partir de la distribution $p(\mathbf{z}_k|\mathbf{z}_{0:k-1}, \mathbf{I}_{0:k})$. Nous avons vu que dans le cas linéaire, cette distribution a une expression analytique calculable, ce qui n'est en général pas le cas pour un système non linéaire. Les méthodes existantes dans le cas non linéaire classique ont été décrites dans la première partie (§3.1 p.62). En particulier, nous avons vu que si l'on dispose d'une approximation discrète de $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ sous forme d'un nuage de particules pondérées (obtenu par exemple par un algorithme de filtrage particulaire), il est possible d'approcher la densité de vraisemblance de la nouvelle observation. Cela est toujours valide dans notre contexte conditionnel à la séquence.

On suppose disposer d'une approximation discrète de $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1})$ sous la forme d'un nuage $\{\mathbf{x}_{k-1}^{(i)}, \tilde{w}_{k-1}^{(i)}\}_{i=1\dots N}$. On a l'expression suivante :

$$p(\mathbf{z}_k|\mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) = \int p(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1}) d\mathbf{x}_{k-1} \quad (5.75)$$

$$\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) \quad (5.76)$$

Cette équation est exploitable dans le cas où une expression analytique de $p(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k})$ est connue. Or, nous avons vu que cela est le cas pour les modèles conditionnels étudiés ici (5.52)(5.67). Ainsi, ces modèles permettent d'obtenir une approximation précise de

$p(\mathbf{z}_k | \mathbf{z}_{0:k-1}, \mathbf{I}_{0:k})$, et donc de construire une fenêtre de validation pertinente. Ce point est un avantage supplémentaire de notre approche. Afin de construire une solution non coûteuse en temps de calcul, il est raisonnable de considérer une approximation de la fenêtre de validation à l'aide de zones ellipsoïdales.

Cas du modèle construit avec la vraisemblance monomodale

Dans le cas du modèle associé à la vraisemblance monomodale, rappelé ici :

$$\begin{cases} p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k ; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}) \\ p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; H_{k|\mathbf{I}} \mathbf{x}_k, R_{k|\mathbf{I}}), \end{cases} \quad (5.77)$$

on dispose de l'expression suivante :

$$p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I}} + H_{k|\mathbf{I}} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I}}^t). \quad (5.78)$$

En observant les équations (5.76) et (5.78), il apparaît que la fenêtre de validation "exacte" est la réunion de N ellipses, où chaque ellipse est associée à une particule. La construction d'une telle région étant trop coûteuse en temps de calcul, nous proposons de l'approcher par une ellipse unique, comme cela est illustré par la figure 5.4.

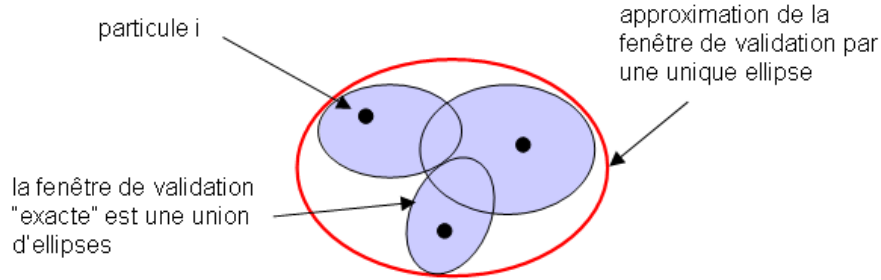


FIG. 5.4: Construction de la fenêtre de validation pour le modèle (5.77). Exemple avec trois particules. À chaque particule $\mathbf{x}_{k-1}^{(i)}$ est associée une ellipse bleue qui correspond à $p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$. La fenêtre de validation "exacte" est la réunion de ces ellipses. Elle est approchée par une unique ellipse, représentée en rouge.

Pour cela, il est nécessaire de calculer les deux premiers moments de $p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$ notés $\mathbf{m}_k^{valid}$ et Σ_k^{valid} . Des approximations empiriques de ces quantités peuvent être obtenues en utilisant (5.76) et (5.78). En effet, une approximation de l'espérance de la distribution

de la nouvelle observation est donnée par :

$$\mathbf{m}_k^{valid} = \mathbb{E}[\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] \quad (5.79)$$

$$= \int \mathbf{z}_k p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) d\mathbf{z}_k \quad (5.80)$$

$$\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} \int \mathbf{z}_k p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) \quad (5.81)$$

$$\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}). \quad (5.82)$$

En ce qui concerne le deuxième moment de la distribution, on a (la démonstration est donnée en annexe D) :

$$\Sigma_k^{valid} = \text{var}[\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] \quad (5.83)$$

$$\begin{aligned} &\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} [R_{k|\mathbf{I}} + H_{k|\mathbf{I}} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I}}^t + \|H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})\|^2] \\ &\quad - \left\| \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}) \right\|^2 \end{aligned} \quad (5.84)$$

À partir de ces expressions, on peut écrire l'approximation suivante :

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) \simeq \mathcal{N}(\mathbf{z}_k; \mathbf{m}_k^{valid}, \Sigma_k^{valid}) \quad (5.85)$$

et on peut donc définir simplement la fenêtre de validation $valid_k$ comme étant l'ellipsoïde dont l'équation est :

$$valid_k = \{\mathbf{z}_k : (\mathbf{z}_k - \mathbf{m}_k^{valid})^t \Sigma_k^{valid}^{-1} (\mathbf{z}_k - \mathbf{m}_k^{valid}) \leq \gamma\}. \quad (5.86)$$

Ainsi, en plus de permettre l'utilisation de la fonction d'importance optimale, le modèle (5.77) présente l'avantage de permettre une définition simple d'une fenêtre de validation.

Cas du modèle construit avec la vraisemblance multimodale

Dans le cas du modèle associé à la vraisemblance multimodale, rappelé ici :

$$\begin{cases} p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}) \\ p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j}; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j}). \end{cases} \quad (5.87)$$

on dispose de l'expression suivante :

$$\begin{aligned} &p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) \\ &= V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}\left(\mathbf{z}_{k,j}; H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I},j} + H_{k|\mathbf{I},j} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I},j}^t\right). \end{aligned} \quad (5.88)$$

Une fois encore, il est possible d'appliquer la relation (5.76). On remarque alors que la fenêtre de validation "exacte" correspond à l'union de $N \times M_k$ ellipses, où N est le nombre de particules et M_k le nombre de mesures disponibles au temps k . De manière similaire au cas précédent, afin de réduire le temps de calcul, nous ne souhaitons pas considérer directement cette région "exacte". La solution proposée est une extension directe de la solution étudiée dans le cas d'une vraisemblance monomodale. L'idée consiste à approcher la région de validation par la réunion de M_k ellipses (ou "modes"). Elle est illustrée par la figure 5.5.

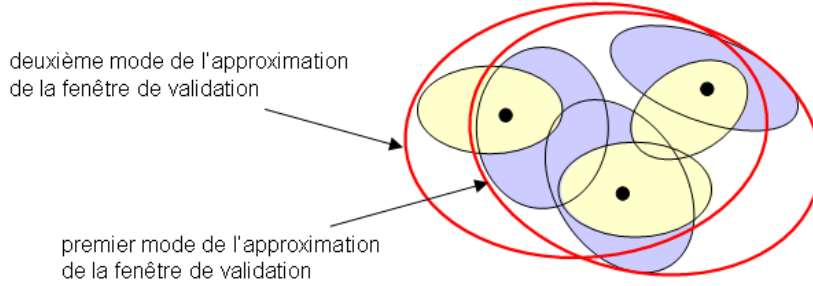


FIG. 5.5: Construction de la fenêtre de validation pour le modèle (5.87). Exemple avec trois particules et deux mesures. À chaque particule $\mathbf{x}_{k-1}^{(i)}$ est associée une ellipse bleue qui correspond à $p(\mathbf{z}_{k,1}|\mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$, et donc à la première mesure ; et une ellipse jaune qui correspond à $p(\mathbf{z}_{k,2}|\mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$, et donc à la deuxième mesure. La fenêtre de validation "exacte" est la réunion de ces ellipses. Elle est approchée par deux ellipses, représentées en rouge. Ces deux ellipses rouges correspondent à une approximation de l'union des ellipses respectivement bleues et jaunes, et sont associées aux mesures respectivement 1 et 2.

On obtient alors la fenêtre de validation $valid_k$ définie par

$$valid_k = \bigcup_{j=1}^{M_k} \Psi_j \quad (5.89)$$

où chaque ellipse Ψ_j correspond à la mesure j et est décrite par :

$$\Psi_j = \{\mathbf{z}_{k,j} : (\mathbf{z}_{k,j} - \mathbf{m}_{k,j}^{valid})^t \Sigma_{k,j}^{valid}^{-1} (\mathbf{z}_{k,j} - \mathbf{m}_{k,j}^{valid}) \leq \gamma\}. \quad (5.90)$$

dont les expressions des deux paramètres $\mathbf{m}_{k,j}^{valid}$ et $\Sigma_{k,j}^{valid}$ sont déduites du cas monomodale :

$$\mathbf{m}_{k,j}^{valid} \simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), \quad (5.91)$$

$$\begin{aligned} \Sigma_{k,j}^{valid} &\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} [R_{k|\mathbf{I},j} + H_{k|\mathbf{I},j} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I},j}^t + \|H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})\|^2] \\ &\quad - \left\| \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}) \right\|^2. \end{aligned} \quad (5.92)$$

On en conclut que les modèles de type (5.87) permettent à la fois de considérer plusieurs mesures, d'échantillonner les particules selon la fonction d'importance optimale, et de construire simplement une région de validation pertinente.

Remarque 1 : Dans le cas du modèle avec la vraisemblance multimodale, il est nécessaire de connaître le nombre de mesures disponibles au temps k , avant même de construire la fenêtre de validation ou celles-ci seront recherchées. Cela peut paraître contradictoire. Cependant, alors que dans le cadre de suivi radar, le nombre de fausses alarmes n'est pas connu à l'avance, celui-ci peut raisonnablement être fixé pour des applications de trajectographie dans des séquences d'images. En effet, nous avons vu que dans ce cadre, pour des stratégies *Top-Down*, l'observation est obtenue par un module de détection. S'il est nécessaire de traiter plusieurs mesures pour lever des ambiguïtés possibles, les observations disponibles correspondent à un nombre fixé de meilleurs candidats résultant du module de détection.

Remarque 2 : Dans le même esprit que la remarque précédente, on peut noter que pour les deux modèles décrits ici, il est nécessaire de connaître les matrices $H_{k|\mathbf{I}}$ et $R_{k|\mathbf{I}}$ ($H_{k|\mathbf{I},j}$ et $R_{k|\mathbf{I},j}$ dans le cas multimodal) pour construire la fenêtre de recherche de la mesure au temps k . Or, ces matrices décrivent le lien entre l'état et la (les) mesure(s) au temps k . Dans le cas où elles sont estimées en fonction de $\mathbf{z}_k$, et donc ne sont pas disponibles suffisamment tôt, une approche raisonnable consiste à les remplacer par $H_{k-1|\mathbf{I}}$ et $R_{k-1|\mathbf{I}}$ ($H_{k-1|\mathbf{I},j}$ et $R_{k-1|\mathbf{I},j}$ dans le cas multimodal) pour la phase de validation.

5.3.4 Récapitulatif

Pour modéliser un problème de trajectographie par les modèles conditionnels gaussiens partiellement linéaires décrits dans cette section, il est nécessaire d'utiliser une équation de mesure linéaire, et des bruits additifs gaussiens. Nous pensons qu'un tel choix – qui consiste à se restreindre à un modèle de mesure simpliste – peut être compensé par l'utilisation d'une équation dynamique précise. Cette approche se situe en quelque sorte à l'opposé des techniques classiques de suivi utilisées en vision, mais constitue une alternative intéressante. En effet, les propriétés de ces modèles permettent la mise en œuvre d'un algorithme simple et efficace décrit par le diagramme 5.6. Au niveau théorique, ces modèles sont valides dans notre cadre de filtrage conditionnel qui autorise l'estimation des fonctions et matrices du système sur les données image. Au niveau pratique, nous pensons que l'association d'une loi d'évolution porteuse d'information, d'une estimation pertinente de la covariance de bruit de mesure, et de l'utilisation de la fonction d'importance optimale permet de contrebalancer une vraisemblance informative et complexe, même face à des environnements bruités. Dans la dernière partie de ce document, ces modèles seront validés expérimentalement pour du suivi de points et de plans dans des séquences d'images réelles.

5.4 Synthèse

Le chapitre 4 met en évidence les lacunes de la modélisation du problème de trajectographie dans une séquence d'images par le modèle simple de chaîne de Markov cachée. Cette modélisation apparaît limitée notamment lorsque le peu d'informations disponibles

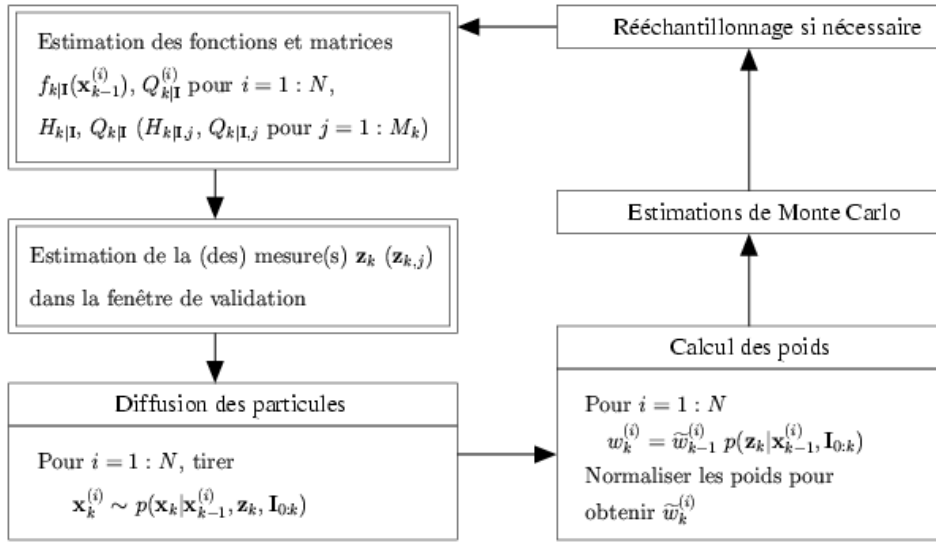


FIG. 5.6: Diagramme représentant les étapes du filtre non linéaire conditionnel pour la résolution des systèmes (5.52) (5.67). Les cadres doubles représentent les étapes réalisées sur la séquence d'images.

sur le système à estimer implique l'estimation du modèle complet sur la séquence. Nous avons donc proposé dans ce chapitre une nouvelle modélisation de ce problème sous la forme d'une chaîne de Markov cachée prenant en compte de manière explicite les données image. Ce formalisme généralise le modèle classiquement employé vers un nouveau modèle appelé **modèle conditionnel**.

Dans le but d'estimer la loi *a posteriori* de l'état conditionnellement aux mesures et aux images, de nouveaux filtres ont été définis. Ceux-ci ont été dérivés sur le principe des filtres classiques en considérant un conditionnement par rapport à la séquence d'images. Deux instances de ce problème ont été étudiées. Dans le cas où les équations du modèle conditionnel sont linéaires, il est résolu par le **filtre linéaire conditionnel** qui est déduit du filtre de Kalman. Cet algorithme est construit à partir de l'**estimateur linéaire conditionnel de variance minimale**, qui est une extension de l'estimateur BLUE. Face à des non linéarités, il est nécessaire d'utiliser le **filtre non linéaire conditionnel**, qui consiste en un filtre particulaire incluant la séquence d'images comme variable conditionnelle.

Enfin, nous avons étudié plus en détail deux cas particuliers du modèle conditionnel non linéaire : un modèle gaussien constitué d'une équation dynamique non linéaire et d'une équation de mesure linéaire ; et son extension au cas où le système est perturbé par des fausses alarmes. Ces modèles sont intéressants dans le cas où la fonction d'évolution et les covariances de bruit de mesure peuvent être estimées de manière pertinente sur la séquence d'images. Ils ont l'avantage de permettre l'échantillonnage des particules par la fonction d'importance optimale et la définition précise d'une fenêtre de validation.

Les contributions présentées dans ce chapitre ont été publiées dans les articles de conférence internationale [Arnaud 04a] [Arnaud 04c] et dans la revue IEEE Transactions on Image Processing [Arnaud 04b].

Conclusion

Dans cette partie, nous nous sommes intéressés à la modélisation d'un problème de suivi dans une séquence d'images par un système dynamique. Nous avons vu que ce choix de modélisation permet la mise en œuvre de méthodes simples à implémenter et robustes à des situations pouvant entraîner des problèmes de minima locaux. Il nécessite la définition d'un modèle décrivant les propriétés du système dynamique étudié, composé d'une équation dynamique et d'une équation de vraisemblance des données.

Un tel modèle est difficile à construire lorsque la nature de l'entité suivie n'est pas connue à l'avance. Dans ce cas, la majorité des approches existantes propose de décrire le système par une équation de vraisemblance hautement non linéaire, et une équation d'évolution très peu informative. Une autre solution, moins répandue, consiste à mettre en place une loi d'évolution déduite d'une mesure de mouvement estimée sur la séquence. Cette deuxième approche nous semble intéressante car elle permet d'augmenter la robustesse du suivi, notamment face à des mouvements chaotiques ou de grande amplitude.

Lorsque cette équation dynamique est estimée sur les images, une corrélation apparaît entre l'état et la mesure au travers des données image. Or, cette corrélation n'est pas décrite par la modélisation classique d'un système dynamique par le modèle simple de chaîne de Markov cachée. Afin de pallier ce manque, nous avons généralisé ce modèle classique vers un modèle conditionnel qui autorise l'estimation de ces deux équations sur la séquence d'images. Cette généralisation est adaptée au suivi pour des applications de vision par ordinateur car elle rend les données images explicites dans le modèle de chaîne de Markov cachée.

Les filtres associés à ce formalisme ont été dérivés à partir des filtres classiques, mais en considérant un conditionnement par rapport à la séquence d'images. On distingue deux algorithmes : le filtre linéaire conditionnel spécifique aux modèles linéaires, et le filtre non linéaire conditionnel pour résoudre les non linéarités. En pratique, ces filtres sont équivalents au filtre de Kalman et au filtre particulaire. L'utilisation des équations de ces méthodes traditionnelles a ainsi été justifiée pour le traitement de systèmes dont les fonctions et matrices sont estimées sur la séquence.

Enfin, lorsqu'une telle estimation des fonctions et matrices est réalisable, il est possible de considérer un modèle aux propriétés très intéressantes. Ce modèle est gaussien, constitué d'une équation dynamique non linéaire et d'une équation de mesure linéaire. Il permet de diffuser les particules avec la fonction d'importance optimale et d'inférer la fenêtre de validation de manière pertinente. Mais surtout, ces propriétés restent valides pour l'extension de ce modèle au cas où plusieurs observations de l'entité suivie sont

disponibles. Cette approche se situe à l'opposé des techniques classiques en suivi dans une séquence d'image, mais constitue une alternative intéressante.

Dans la partie suivante, nos modèles conditionnels et les filtres associés sont utilisés et validés dans le cadre d'une application de suivi de points d'intérêt et d'une application de suivi d'objets plans.

Troisième partie

Application au suivi de points caractéristiques et objets plans

Introduction

Le suivi de points d'intérêt dans une séquence d'images constitue un problème basique mais essentiel dans de nombreuses applications en vision par ordinateur. Un grand nombre de tâches de haut niveau peuvent en dépendre directement, par exemple la reconstruction 3D [Sturm 02], ou le contrôle robotique [Mezouar 01] [Remazeilles 04].

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. Plus précisément, suivre un point le long d'une séquence implique la construction d'une caractéristique locale invariante le long de la trajectoire. On peut également noter qu'il est très difficile de construire au préalable un modèle dynamique du point, 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 la plus couramment utilisée est la conservation de la luminance du point et de son voisinage le long de la trajectoire. La caractéristique locale considérée invariante est alors une petite fenêtre autour du point d'intérêt. Cette hypothèse a conduit à la définition de deux familles de méthodes.

Les premières sont des méthodes intuitives basées sur des critères de similarité [Aschwandten 92]. Ces méthodes, appelées **méthodes de corrélation**, 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 [Schmetz 93] ou les écoulements fluides en imagerie PIV [Adrian 91]. Malgré le fait que les critères de similarité utilisés ne soient pas invariants aux transformations géométriques (changement d'échelle, rotation, distorsion perspective) et photométriques de l'image, ces méthodes restent très employées pour leur simplicité. Nécessitant une recherche exhaustive sur une région d'intérêt, elles peuvent être coûteuses en temps de calcul. Néanmoins, elles permettent de capter d'éventuels grands déplacements du point suivi.

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 Kanade-Lucas-Tomasi (KLT) [Lucas 81] [Shi 94], dérivé à partir de l'hypothèse de conservation de luminance dans un voisinage du point. Les méthodes différentielles consistent généralement en une procédure de minimisation, initialisée au résultat précédent. Par conséquent, elles sont rapides mais peuvent être perturbées par des minima locaux. De nombreuses variantes algorithmiques existent, afin de rendre ces approches plus robustes aux changements d'illumination et

géométriques.

Afin de profiter des avantages de ces deux approches, nous proposons de combiner ces deux formulations complémentaires du problème de suivi dans une modélisation probabiliste de filtrage. Plus précisément, nous proposons d'associer des mesures fournies par une technique de corrélation à un modèle dynamique défini à partir d'estimation de mouvements. Ces estimations sont apportées par une méthode différentielle.

Dans ce cadre, trois algorithmes ont été construits pour la trajectographie de points. Ils se basent sur les filtres conditionnels décrits dans le chapitre 5, puisque le modèle proposé repose sur une dynamique estimée en ligne. Le premier algorithme est linéaire. Il est adapté aux séquences présentant un mouvement dominant. Le second est non linéaire et permet de suivre des points dont le mouvement ne peut être décrit que localement. Enfin, une extension de ce dernier pour le traitement des ambiguïtés constitue le troisième algorithme.

À partir de ces algorithmes, nous nous sommes intéressés au suivi de motifs plans. L'approche choisie consiste à décrire l'objet par un ensemble de points caractéristiques. La division de cet ensemble de primitives en deux sous-parties amène à la construction d'un modèle dont les propriétés permettent d'utiliser un filtre particulier rao-blackwellisé pour résoudre le problème de trajectographie. Le filtre résultant combine à la fois une estimation de mouvement par une méthode différentielle, et une estimation par une méthode mise en correspondance. De plus, il permet de considérer l'ajout de contraintes géométriques.

Dans le premier chapitre de cette partie, les méthodes existantes pour le suivi de points d'intérêt et d'objets plans sont décrites. Nous reviendrons en détails sur les méthodes de corrélation et les méthodes différentielles, ainsi que sur les mesures de validité de celles-ci. Dans le second chapitre, les trois algorithmes proposés seront présentés. Ils seront validés par des résultats expérimentaux sur des séquences réelles. Nous verrons que ces méthodes permettent de suivre des trajectoires qui subissent des changements abrupts de direction et de vitesse, ainsi que des occlusions. De plus, des comparaisons avec l'algorithme KLT montreront que l'association d'un modèle d'évolution estimé en ligne et de mesures de corrélations permet d'améliorer significativement les résultats. Enfin, le dernier chapitre exposera l'algorithme de suivi d'objet plan construit. Ce dernier sera également validé sur des séquences réelles.

Chapitre 6

Approches existantes

Dans ce chapitre sont décrites les deux catégories d'approches classiquement utilisées pour suivre des points caractéristiques, à savoir les méthodes de corrélation et les méthodes différentielles. Nous verrons ensuite comment ces dernières peuvent être étendues au suivi de motifs plan, et quelle contrainte géométrique est généralement introduite.

6.1 Suivi de points d'intérêt

Un point d'intérêt correspond à une singularité du signal. Dans l'image, cela se traduit par un coin, une jonction en T, ou un endroit où la texture varie fortement. L'utilité de ces points caractéristiques est établie en remarquant que le signal offre en ces points une information de localisation plus persistante qu'en des pixels correspondant à des changements unidirectionnels du signal (lignes de contours) ou à des régions homogènes [Schmid 96]. Ils sont en conséquence très largement utilisés dans les applications du traitement d'images. La chaîne algorithmique complète d'un algorithme de suivi de points d'intérêt peut être schématisée par les trois étapes suivantes :

1. **Détection - sélection des points d'intérêts :**

Cette première étape consiste soit à détecter les points d'intérêt dans l'image, soit à s'assurer que les points choisis par l'utilisateur ont un voisinage texturé. Les algorithmes sont construits de telle sorte que les points disponibles en sortie aient un voisinage porteur d'une information suffisamment stable pour être suivis.

2. **Suivi des points :**

L'étape de suivi constitue le cœur du problème. Comme décrit en introduction, les méthodes existantes reposent généralement sur l'hypothèse de conservation de la luminance du point et de son voisinage. Les deux types d'approches sont les méthodes de corrélation et les méthodes différentielles. Alors que les premières nécessitent une recherche exhaustive et sont basées sur un critère de corrélation, les secondes reposent sur une procédure de minimisation.

3. **Vérification de la qualité du suivi :**

Cette dernière étape est intrinsèquement liée au choix de la méthode de suivi. Si une méthode de corrélation est utilisée, un critère de validité du suivi peut être

construit sur la surface de corrélation. En utilisant une méthode de corrélation ou une méthode différentielle, une erreur d'estimation est disponible. Lorsque plusieurs points sont suivis simultanément, une étude sur la répartition des résidus d'estimation peut permettre de supprimer les mauvaises trajectoires. Enfin, une autre approche consiste à se baser sur des critères géométriques de la scène pour vérifier les résultats.

Ces trois points sont abordés en détail dans cette section.

6.1.1 Détection - sélection des points d'intérêts

D'après Schmid [Schmid 96], les algorithmes de détection de points d'intérêt peuvent être classés en trois catégories. La première contient les méthodes basées sur une extraction au préalable des lignes de contour dans l'image. Les points d'intérêt sont alors choisis comme étant les points particuliers tels que les points de courbure maximale, d'intersection, d'inflexion, etc.. La deuxième extrait le point caractéristique directement à partir de la fonction de luminance. Un grand nombre des techniques de cette catégorie repose sur la matrice d'auto-corrélation du signal autour du pixel. Cette matrice rend compte de la structure du voisinage du point. Enfin, la troisième concerne les méthodes fondées sur un modèle théorique du signal. Celles-ci ne sont utilisables que pour un type bien précis de points, mais elles permettent l'obtention de résultats sous-pixeliques. Pour un état de l'art détaillé, nous renvoyons à la thèse [Schmid 96] et à l'article [Schmid 00].

Dans ces deux références, ainsi que dans l'article [Tissainayagam 04], des études comparatives de différents détecteurs sont menées. Dans le travail de Schmid, la méthode d'évaluation mise en œuvre repose sur un critère de répétabilité, et les tests se font sur des images qui subissent des changements géométriques (rotation, changement d'échelle), de luminosité et de point de vue. De cette étude, il ressort que dans tous les cas, le détecteur de Harris [Harris 88] donne des résultats meilleurs ou équivalents aux quatre autres détecteurs auquel il est comparé. Dans l'article de Tissainayagam, quatre détecteurs sont comparés dans le cadre d'une application de suivi. Les critères utilisés sont la stabilité du point d'intérêt et sa localisation. Les images utilisées sont des séquences réelles où les changements du signal sont dus à du bruit d'acquisition et à des changements d'illumination. La conclusion de cette comparaison est que le détecteur de Harris et le détecteur utilisé par le KLT [Shi 94] (qui n'est pas comparé dans l'étude de Schmid) sont des méthodes adaptées à des fins de suivi. Notons que ces deux méthodes appartiennent à la deuxième catégorie citée précédemment. Nous en donnons ici quelques détails.

Détecteur KLT et détecteur de Harris

Le détecteur KLT et le détecteur de Harris reposent sur le calcul d'une matrice d'auto-corrélation, liée à la fonction du même nom et qui prend en compte les valeurs des dérivées premières du signal en niveaux de gris. Cette matrice se calcule sur une fenêtre $\mathcal{W}(\mathbf{s})$ autour de chaque point $\mathbf{s}$ testé et s'écrit :

$$M(\mathbf{s}) = g(\mathbf{s}) \otimes \sum_{\mathbf{m} \in \mathcal{W}(\mathbf{s})} \begin{bmatrix} \partial_{\mathbf{u}} \mathbf{I}(\mathbf{m})^2 & \partial_{\mathbf{u}} \mathbf{I}(\mathbf{m}) \partial_{\mathbf{v}} \mathbf{I}(\mathbf{m}) \\ \partial_{\mathbf{u}} \mathbf{I}(\mathbf{m}) \partial_{\mathbf{v}} \mathbf{I}(\mathbf{m}) & \partial_{\mathbf{v}} \mathbf{I}(\mathbf{m})^2 \end{bmatrix} \quad (6.1)$$

où $\otimes$ désigne le produit de convolution. $\partial_u \mathbf{I}$, $\partial_v \mathbf{I}$ sont les dérivées premières du signal. $g(\mathbf{s})$ désigne une fonction de fenêtrage centrée au pixel $\mathbf{s}$ et dont le rôle est de donner plus d'importance au centre du voisinage qu'à sa périphérie (typiquement, une fenêtre gaussienne). Les valeurs propres notées λ_1 et λ_2 de cette matrice sont les courbures principales de la fonction d'auto-corrélation. Elles fournissent des informations sur le profil d'intensité de la fenêtre $\mathcal{W}(\mathbf{s})$. Deux faibles valeurs sont associées à un modèle d'intensité constante, alors que deux valeurs importantes indiquent un profil de luminance qui peut être facilement suivi.

Le détecteur KLT repose sur ce principe. Un point d'intérêt est détecté lorsque la plus petite des deux valeurs propres est supérieure à un seuil λ_{seuil} . Cette valeur peut être fixée par l'utilisateur (souvent choisie entre 0.1 et 1 [Kermad 01]), ou être calculée en fonction de la répartition de l'ensemble des plus faibles valeurs propres sur l'image [Tomasi 91]. Dans tous les cas, le choix de ce seuil n'est pas critique [Tomasi 91]. Notons que cette méthode peut être utilisée pour sélectionner des points choisis par un utilisateur. Elle sera mise en œuvre dans nos applications.

Le détecteur de Harris utilise également la matrice d'auto-corrélation. Cependant, à la place du calcul des valeurs propres, ce détecteur utilise une mesure de force du point $\mathbf{s}$ reposant sur le déterminant $\det$ et la trace tr de la matrice $M(\mathbf{s})$:

$$M_f(\mathbf{s}) = \det M(\mathbf{s}) - \alpha [tr M(\mathbf{s})]^2 \quad (6.2)$$

où α est un paramètre fixé par l'utilisateur. En rappelant les relations : $\det M(\mathbf{s}) = \prod_i \lambda_i$ et $tr M(\mathbf{s}) = \sum_i \lambda_i$, on note que cette mesure rend compte de la structure du voisinage de $\mathbf{s}$. En particulier, elle est supérieure à zéro dans le cas d'un coin. Ce détecteur a récemment été généralisé pour tolérer des déformations affines dans un voisinage du point caractéristique [Mikolajczyk 02].

Une fois les points d'intérêt détectés ou sélectionnés, une méthode de suivi est utilisée pour reconstruire leur trajectoire le long de la séquence d'images. En faisant l'hypothèse de conservation de la luminance d'un point le long de sa trajectoire, deux grandes familles d'algorithmes se distinguent : les méthodes de corrélation décrites dans le paragraphe 6.1.2, et les méthodes différentielles détaillées dans le paragraphe 6.1.3.

6.1.2 Méthodes de corrélation

Les méthodes de corrélation (aussi appelées méthodes de mise en correspondance, ou par "block-matching") décrivent le point à suivre par une petite image définie autour de celui-ci (un "template"). Cette image décrit la luminance du point et de son voisinage, et constitue la caractérisation référence du point d'intérêt. Le principe de ces techniques est de rechercher dans une région d'intérêt de la nouvelle image le pixel minimisant une distance entre sa propre caractérisation en niveaux de gris et ce modèle de référence. La recherche est exhaustive. Pour chaque pixel visité, la distance est calculée, et le résultat correspond au point associé à la mesure de distance minimale. Les distances sont appelées mesures de corrélation, ou mesures de similarité. Le principe de ces méthodes est illustré par la figure 6.1.

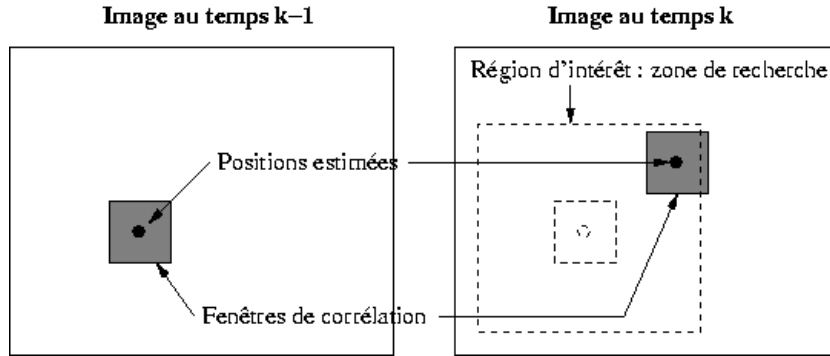


FIG. 6.1: Principe de la mise en correspondance par une méthode de corrélation

La *somme des différences au carré* (SSD) est la mesure la plus simple et la plus utilisée. Nous en donnons ici la formulation pour un pixel $\mathbf{s}$ de l'image $\mathbf{I}$, dont le modèle de référence est $\tilde{\mathbf{I}}$ (de même taille que la fenêtre $\mathcal{W}$) :

$$\text{SSD}(\mathbf{s}) = \sum_{\mathbf{m} \in \mathcal{W}} [\tilde{\mathbf{I}}(\mathbf{m}) - \mathbf{I}(\mathbf{s} + \mathbf{m})]^2. \quad (6.3)$$

D'autres mesures classiques sont la *somme des valeurs absolues* (SAD), la *corrélation croisée* (CC), la *corrélation croisée normalisée* (NCC), ainsi que les versions centrées de ces distances (ZSSD, ZSAD, ZNCC). Nous renvoyons à [Chambon 02] pour une liste détaillée des mesures de corrélation. Dans [Aschwanden 92], une étude comparative entre plusieurs mesures de similarité est présentée. Les résultats montrent que la SSD et la SAD donnent de bons résultats tant que les changements d'illuminations sont minimales. Il apparaît également que les critères ZSSD, ZSAD, ZNCC présentent un bon compromis entre vitesse et performance.

Les méthodes de corrélation sont très répandues du fait de leur simplicité. Elles sont performantes si les conditions de prise de vues des images successives sont proches. Cependant, elles souffrent de trois problèmes majeurs. Premièrement, le choix des zones d'intérêt où est effectuée la recherche relève souvent de critères expérimentaux et dépend de la séquence d'images. Pourtant, ce choix est crucial pour ne pas perdre le point d'intérêt. Deuxièmement, ces techniques reposent sur une recherche exhaustive et sont donc coûteuse en temps de calcul. Des travaux récents [Giachetti 00] proposent une succession de procédures pour réduire le temps de calcul de ces approches. Troisièmement, la plupart des mesures de similarité classiques ne sont pas robustes aux changements géométriques (en particulier rotation et changement d'échelle), aux changements d'illuminations et aux occultations partielles. Pour rendre des mesures de corrélations non sensibles aux occlusions partielles, des études proposent l'utilisation de statistiques robustes [Lan 97] [Chambon 02]. D'autres travaux s'attaquent à la robustesse aux changements d'échelles [Cahn von Seelen 97], aux rotations [Hu 94], ou encore aux changements d'illumination en modélisant une fonction décrivant le lien entre les intensités entre deux images successives [Wang 02].

Dans tous les cas, il est possible de garder une mesure très simple de type SSD si le modèle de référence du point est mis à jour. En effet, si le suivi se fait sur une longue

séquence, une référence non évolutive peut devenir obsolète en cas de changements de point de vue de la caméra, ou de changements d'illumination. Comme cela est décrit dans [Matthews 04], le choix de cette référence et de son évolution est un problème difficile. Une mise à jour naïve de ce modèle en le modifiant à chaque itération par une fenêtre autour de l'estimation trouvée peut entraîner une forte dérive des résultats due à une accumulation des erreurs. Il est important de compenser cette dérive à chaque évolution du modèle de référence. Une proposition dans ce sens est apportée dans [Nguyen 01], où le modèle de référence est mis à jour par un filtre de Kalman.

Mesure de la validité d'une méthode de corrélation

Une méthode de corrélation donne une estimation de la position du point dans la nouvelle image. Comme décrit en introduction, la validation du résultat obtenu est une étape cruciale. Une étude de différentes mesures de confiance est disponible dans [Nickels 98] et [Nickels 99]. Les travaux existants concernent essentiellement la validation d'un résultat fourni par SSD.

L'idée d'utiliser une surface de corrélation a été introduite par Anandan [Anandan 87]. Cette surface, notée SSDS (surface de SSD), est construite à partir des valeurs de la SSD obtenues par (6.3) en chaque point d'un voisinage de la position estimée. Les courbures de cette surface ont la propriété de rendre compte de la précision de l'estimation. Plus précisément, si la variation de la SSDS le long d'une direction est faible, cela signifie que le point ne peut pas être estimé avec précision le long de cette direction. Cela est illustré par la figure 6.2.

À partir de ce constat, différentes méthodes basées sur un calcul des courbures de cette surface dans quatre directions ou le long des axes principaux autour du point estimé ont été proposées [Anandan 87] [Anandan 89] [Matthies 89] [Papanikolopoulos 93]. En plus du calcul des courbures, Papanikolopoulos [Papanikolopoulos 93] propose d'étudier les propriétés statistiques de cette surface par un calcul de moyenne et d'écart-type dans les différentes directions.

Une approche qui nous semble intéressante a été introduite par Singh et Allen [Singh 92]. Le principe consiste à faire une analogie entre la distribution de probabilité d'avoir une bonne estimation et la surface de corrélation transformée en une distribution réponse construite en chaque point $\mathbf{m}$ dans un voisinage $\mathcal{W}(\mathbf{s})$ de la nouvelle position estimée $\mathbf{s}$ par :

$$\mathcal{D}(\mathbf{m}) = \exp(-c \text{ SSD}(\mathbf{m})), \quad (6.4)$$

où c est un facteur de normalisation choisi tel que le maximum prenne la valeur 0.95 [Singh 92], ou tel que $\sum_{\mathbf{m} \in \mathcal{W}(\mathbf{s})} \mathcal{D}(\mathbf{m}) = 1$ [Nickels 98] (c est alors déterminé par une procédure itérative en remarquant que $\sum_{\mathbf{m} \in \mathcal{W}(\mathbf{s})} \mathcal{D}(\mathbf{m}) - 1$ est une fonction décroissante). Cette interprétation de la SSDS comme une distribution de probabilité permet d'utiliser les méthodes d'estimation statistique pour calculer une mesure de confiance de la position obtenue $\mathbf{s}$. En particulier, il est possible de calculer de façon empirique une matrice de

variance-covariance $R(\mathbf{s})$ [Singh 92] :

$$R(\mathbf{s}) = \begin{pmatrix} \sum_{\mathbf{m} \in \mathcal{W}(\mathbf{s})} \text{SSD}(\mathbf{m})(u_{\mathbf{m}} - u_{\mathbf{s}})^2 & \sum_{\mathbf{m} \in \mathcal{W}(\mathbf{s})} \text{SSD}(\mathbf{m})(u_{\mathbf{m}} - u_{\mathbf{s}})(v_{\mathbf{m}} - v_{\mathbf{s}}) \\ \sum_{\mathbf{m} \in \mathcal{W}(\mathbf{s})} \text{SSD}(\mathbf{m})(u_{\mathbf{m}} - u_{\mathbf{s}})(v_{\mathbf{m}} - v_{\mathbf{s}}) & \sum_{\mathbf{m} \in \mathcal{W}(\mathbf{s})} \text{SSD}(\mathbf{m})(v_{\mathbf{m}} - v_{\mathbf{s}})^2 \end{pmatrix}, \quad (6.5)$$

où $\mathbf{s} = (u_{\mathbf{s}}, v_{\mathbf{s}})$ et $\mathbf{m} = (u_{\mathbf{m}}, v_{\mathbf{m}})$. Cette matrice permet la définition d'une ellipse de confiance, qui traduit une incertitude spatiale du résultat fourni par la SSD [Nickels 98].

Notons que les techniques décrites ici pour le calcul d'une mesure de confiance ont été définies pour la SSD. Elles sont cependant applicables à toute mesure de similarité du même type.

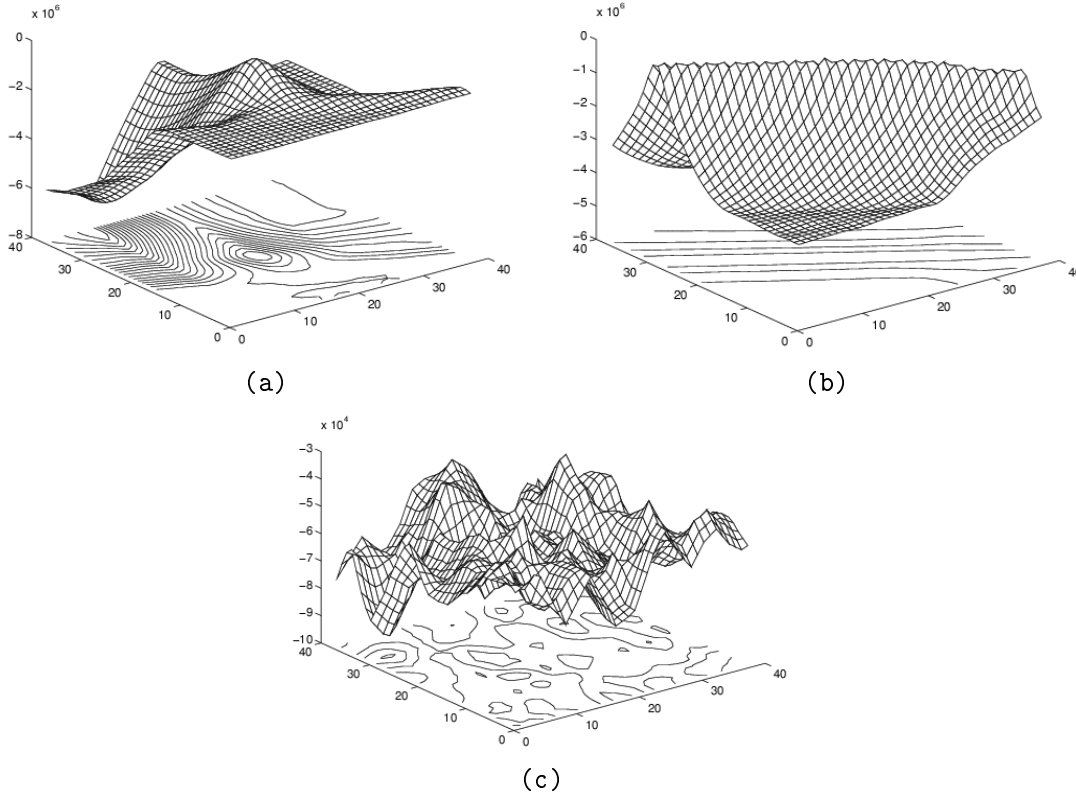


FIG. 6.2: surfaces de SSD centrées à un point placé sur (a) un coin, (b) un contour, (c) une zone homogène bruitée. Ces figures illustrent le fait que les courbures de la SSDs autour de la position estimée (au centre de chaque courbe) rendent compte de la précision de l'estimation - figures extraites de [Nickels 98].

6.1.3 Méthodes différentielles

Les méthodes différentielles reposent sur une formulation différentielle d'un critère de similarité. Ce principe a été introduit par Horn et Schunck [Horn 81] qui expriment

la conservation de la luminance d'un point le long de sa trajectoire par l'équation de contrainte du flot optique. L'algorithme de suivi de point le plus utilisé est l'algorithme de Kanade-Lucas-Tomasi (KLT) [Lucas 81] [Shi 94]. Il est dérivé à partir de la contrainte du flot optique, et d'une paramétrisation spatiale du mouvement sur un voisinage autour du point considéré. Dans le chapitre suivant, nous proposerons une comparaison entre nos algorithmes et le KLT ; c'est pourquoi nous le décrivons en détail dans ce paragraphe (cette description est inspirée de [Tommasini 98]).

Algorithme de Kanade-Lucas-Tomasi

On considère une séries d'images $\{\mathbf{I}_k\}_{k=0:n}$. Entre deux images successives, on fait l'hypothèse classique que la variation du niveaux de gris est due uniquement au mouvement de la caméra et/ou du point $\mathbf{s}$ suivi. L'hypothèse de conservation de la luminance s'écrit alors :

$$\mathbf{I}_k(\mathbf{s}) = \mathbf{I}_{k+1}(\varphi(\mathbf{s})) \quad (6.6)$$

où $\varphi(\mathbf{s})$ représente le pixel déplacé. Cette hypothèse est très forte. Elle est rarement vérifiée pour tous les pixels de l'image, notamment lorsque la séquence d'images est bruitée. Afin de la relâcher, on la modifie légèrement. Le problème est alors de chercher une fonction spatiale décrivant le déplacement qui minimise la valeur du résidu de SSD sur un voisinage :

$$\epsilon = \text{SSD}(\mathbf{s}) = \sum_{\mathcal{W}} [\mathbf{I}_{k+1}(\varphi(\mathbf{s})) - \mathbf{I}_k(\mathbf{s})]^2. \quad (6.7)$$

où $\mathcal{W}$ est une petite fenêtre. On remarque que cet algorithme choisit la caractérisation référence comme étant le modèle de luminance de l'estimation précédente. Notons que ce choix peut poser des problèmes de dérive par accumulation des erreurs d'estimation.

Dans le cas de faibles déplacements, l'hypothèse permettant d'approcher le mouvement par une simple translation (i.e. $\varphi(\mathbf{s}) = \mathbf{s} + \mathbf{d}$) est admise. Par un développement de Taylor à l'ordre 1, on peut faire l'approximation suivante :

$$\mathbf{I}_{k+1}(\mathbf{s} + \mathbf{d}) \simeq \mathbf{I}_k(\mathbf{s}) + \mathbf{g}_k(\mathbf{s})^t \mathbf{d} + \partial_k \mathbf{I}_k(\mathbf{s}), \quad (6.8)$$

où $\mathbf{g}_k(\mathbf{s})^t = [\partial_{\mathbf{u}} \mathbf{I}_k(\mathbf{s}), \partial_{\mathbf{v}} \mathbf{I}_k(\mathbf{s})]$ est le gradient spatial de $\mathbf{I}_k(\mathbf{s})$, et $\partial_k \mathbf{I}_k$ est la dérivée par rapport au temps. En supposant une unité de temps entre chaque image, l'expression du résidu (6.7) peut être réécrite sous la forme :

$$\epsilon = \sum_{\mathcal{W}} [\mathbf{g}_k \mathbf{d} - h_k]^2 \quad (6.9)$$

avec $h_k = \mathbf{I}_{k+1}(\mathbf{s}) - \mathbf{I}_k(\mathbf{s})$. Afin de minimiser ce résidu, on recherche la translation $\hat{\mathbf{d}}$ qui rend la dérivée de (6.9) par rapport à $\mathbf{d}$ nulle. On obtient le système linéaire :

$$G\mathbf{d} = \mathbf{e} \quad (6.10)$$

avec

$$G = \sum_{\mathcal{W}} \begin{pmatrix} \partial_{\mathbf{u}} \mathbf{I}_k^2 & \partial_{\mathbf{u}} \mathbf{I}_k \partial_{\mathbf{v}} \mathbf{I}_k \\ \partial_{\mathbf{u}} \mathbf{I}_k \partial_{\mathbf{v}} \mathbf{I}_k & \partial_{\mathbf{v}} \mathbf{I}_k^2 \end{pmatrix} \quad \text{et} \quad \mathbf{e} = \sum_{\mathcal{W}} h_k \begin{pmatrix} \partial_{\mathbf{u}} \mathbf{I}_k \\ \partial_{\mathbf{v}} \mathbf{I}_k \end{pmatrix}, \quad (6.11)$$

dont la solution est $\hat{\mathbf{d}} = G^{-1} \mathbf{e}$.

Le modèle de translation ne peut pas traduire certaines transformations du voisinage du point suivi, comme les rotations ou les changements d'échelle. Un modèle affine peut être utilisé [Shi 94], i.e. $\varphi(\mathbf{s}) = A\mathbf{s} + \mathbf{d}$, où $\mathbf{d}$ est le déplacement en translation et A la matrice qui rend compte des déformations de la fenêtre. Par des calculs similaires, on retombe sur une expression du type (6.10) (ce cas sera détaillé dans le paragraphe suivant dans le cadre du suivi de motifs plans). Si les déformations affines entre images successives sont négligeables, il est préférable d'utiliser le modèle translationnel. Le modèle affine peut être utilisé entre l'image courante et la première image pour contrôler la qualité du suivi.

Cette méthode est largement utilisée et donne de bons résultats de suivi. Elle peut également être employée pour l'estimation de champs denses de mouvements. L'étude comparative de Barron et al. [Barron 94] a mis en avant le bon comportement du KLT pour cette application. Notons également le travail comparatif de Baker et al. sur cet algorithme et ses extensions [Baker 04b] [Baker 03b] [Baker 03a] [Baker 04a].

Tous comme les méthodes de corrélation, un des problèmes importants des méthodes différentielles est leur sensibilité aux changements d'illuminations. En effet, un tel changement viole l'hypothèse de départ qui est la conservation de la luminance le long de la trajectoire. Pour pallier cette situation, Hager et Belhumeur [Hager 98] proposent de définir un modèle photométrique explicite. L'idée est d'utiliser une base d'images acquises hors ligne suivant différentes configurations d'illumination. Des solutions alternatives consistent à inclure des paramètres photométriques et/ou de contrastes [Jin 01], ou à prendre en compte des modèles de réflexions de l'objet sous-jacent [Gouiffès 04]. En outre, ces méthodes différentielles ne sont pas adaptées aux grands déplacements puisque leur dérivation suppose un mouvement faible. Afin de contourner ce problème, différentes techniques ont été mises en place, comme par exemple l'ajout d'un cadre multi-échelle [Odobez 95a], ou la mise en œuvre d'une approche robuste [Odobez 95a] pour rendre plus efficace la procédure de minimisation. Pour un état de l'art détaillé de ces techniques différentielles et de leurs améliorations, nous renvoyons le lecteur à l'article [Georgescu 04].

Enfin, insistons sur le principal défaut de ces approches qui, à notre avis, réside dans le choix de la caractérisation référence comme étant l'estimation précédente. Ajouté au fait que le point obtenu est pris comme initialisation de la procédure de minimisation, ce choix entraîne de nombreuses situations où le suivi décroche. Une mauvaise estimation à un instant donné entraîne inévitablement une dérive des estimées futures. Ce problème est d'autant plus important que la structure algorithmique ne permet pas de retrouver la bonne position.

Mesure de la validité d'une méthode différentielle

Le dernier module d'un algorithme de suivi est la validation des trajectoires obtenues. Dans le cadre des approches différentielles, deux idées se distinguent. La première est de travailler sur les résidus de l'estimation, alors que la deuxième considère des mesures de validation géométriques.

Dans [Shi 94], le résidu estimé par une équation de type (6.7) dans le cas d'un modèle affine est utilisé comme mesure de confiance du suivi. Lorsque plusieurs points sont suivis

dans une même séquence, l'idée est alors de rejeter les positions correspondant à des résidus nettement plus élevés que la majorité des points. Néanmoins, aucune règle précise de décision n'est proposée. Dans [Jin 01], la règle de rejet est basée sur un seuillage d'un critère de corrélation entre d'une part deux estimations successives, et d'autre part l'estimation courante et l'image initiale. Le choix des seuils n'est cependant pas détaillé.

Dans [Tommasini 98], la mesure de validité des trajectoires obtenues repose sur une approximation de la distribution des résidus. En supposant que l'intensité d'un pixel est constante le long de sa trajectoire, à un bruit gaussien près (de moyenne nulle et de variance 1), on peut écrire :

$$\mathbf{I}_k(\varphi(\mathbf{s})) - \mathbf{I}_0(\mathbf{s}) \sim \mathcal{N}(0, 1). \quad (6.12)$$

On en déduit que le résidu de SSD calculé sur une fenêtre de taille $n \times n$ entre l'image initiale et l'image courante suit une distribution du chi-deux à n^2 degrés de liberté. Quand le nombre de degrés de liberté est élevé, il est possible d'approcher la distribution du chi-deux par une distribution gaussienne. Ainsi, en considérant un support d'estimation suffisamment grand, on peut faire l'hypothèse que le résidu calculé entre l'image initiale et l'image courante d'un point correctement suivi suit une loi normale de moyenne n^2 , et de covariance $2n^2$. À chaque pas de temps, les trajectoires dont le résidu final ne suit pas cette loi sont supprimées. Une étape additionnelle en fin de séquence consiste à rejeter les trajectoires dont le résidu est supérieur à la médiane plus ou moins 3.5 fois l'écart type médian. Cette étape a l'inconvénient de supposer qu'au moins 50% des trajectoires sont valides.

Notons que lorsqu'une procédure de minimisation robuste est mise en place, il peut être intéressant d'utiliser une règle de décision basée sur le nombre de points aberrants détectés [Nguyen 01].

L'ajout de contraintes géométriques peut être également mis en place pour valider les résultats de suivi. Par exemple, la géométrie épipolaire entre deux images successives fournit des contraintes globales que les points mis en correspondance doivent vérifier [Zhang 95] [Torr 03]. Ces méthodes ont l'inconvénient de supposer un mouvement unique global de la scène. Elles sont mises en défaut dès que certains éléments ont un mouvement indépendant.

Remarquons enfin que l'ensemble de ces approches peut être également utilisé dans le cadre des méthodes de corrélations.

6.2 Suivi de motifs plans

Dans la section précédente, nous avons vu que les méthodes de suivi d'un point caractéristique consistent à suivre une petite région de l'image autour du point. Alors que les approches par corrélation ne sont pas utilisables sur de larges régions de l'image (par une trop grande simplicité du modèle translationnel), le principe des techniques différentielles a été étendu au suivi de motifs plus grands, et de façon plus spécifique aux objets plans.

Ces approches reposent sur un modèle de luminance de référence, et la région d'intérêt est décrite par l'intensité des pixels qui la composent. Le but des algorithmes de suivi consiste alors à estimer un ensemble de paramètres de mouvement qui décrivent la transformation de la région d'intérêt au cours de temps. Pour cela, une procédure d'optimisation

efficace est utilisée pour minimiser un critère de corrélation écrit sous forme différentielle. Une telle procédure permet de considérer des mouvements complexes, par exemple des mouvements affines ou homographiques (nous détaillerons le principe de l'homographie dans un paragraphe spécifique). La méthode de Hager et Belhumeur proposée en 1998 [Hager 98] appartient à cette catégorie d'approches. Son principe est très similaire à la méthode KLT avec un modèle de mouvement affine. Nous en expliquons les détails rapidement.

Méthode de Hager et Belhumeur

On note $\mathcal{R} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_p)$ la région d'intérêt composée de p pixels et $\mathbf{I}_k(\mathcal{R})$ l'ensemble des intensités de cette région dans l'image $\mathbf{I}_k$. $\mathbf{I}_0(\mathcal{R})$ est le modèle de luminance de référence. Au temps k , le déplacement d'un pixel $\mathbf{s}$ appartenant à la région d'intérêt est modélisé par un modèle de mouvement paramétrique $\varphi(\mathbf{s}, \boldsymbol{\theta}_k)$, où $\boldsymbol{\theta}_k$ désigne le vecteur de paramètres de mouvement. On note alors $\varphi(\mathcal{R}, \boldsymbol{\theta}_k)$ la nouvelle région.

Le vecteur de paramètres du mouvement $\boldsymbol{\theta}_k$ de la région d'intérêt peut être estimé en minimisant la fonction de coût suivante :

$$\phi(\boldsymbol{\theta}_k) = \|\mathbf{I}_k(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)) - \mathbf{I}_0(\mathcal{R})\|^2 \quad (6.13)$$

Afin d'obtenir une procédure itérative, il est proposé d'estimer l'incrément $\delta\boldsymbol{\theta} = \boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k$. La fonction de coût (6.13) est redéfinie et devient :

$$\phi(\delta\boldsymbol{\theta}) = \|\mathbf{I}_{k+1}(\varphi(\mathcal{R}, \boldsymbol{\theta}_k + \delta\boldsymbol{\theta})) - \mathbf{I}_0(\mathcal{R})\|^2. \quad (6.14)$$

Par une linéarisation de Taylor à l'ordre 1 de $\mathbf{I}_{k+1}(\varphi(\mathcal{R}, \boldsymbol{\theta}_k + \delta\boldsymbol{\theta}))$, on a :

$$\mathbf{I}_{k+1}(\varphi(\mathcal{R}, \boldsymbol{\theta}_k + \delta\boldsymbol{\theta})) \simeq \mathbf{I}_k(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)) + M_k \delta\boldsymbol{\theta} + \partial_k \mathbf{I}_k(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)), \quad (6.15)$$

où M_k est la matrice jacobienne de $\mathbf{I}_k$ par rapport à $\boldsymbol{\theta}_k$. Cette matrice décrit les variations de l'image $\mathbf{I}_k$ autour de la région d'intérêt en fonction d'un mouvement $\delta\boldsymbol{\theta}$ de l'objet. Il s'agit d'une matrice de taille $p \times n$, où n est le nombre de paramètres du vecteur $\boldsymbol{\theta}_k$. En faisant l'approximation suivante :

$$\partial_k \mathbf{I}_k(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)) \simeq \mathbf{I}_{k+1}(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)) - \mathbf{I}_k(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)), \quad (6.16)$$

l'expression (6.14) devient :

$$\phi(\delta\boldsymbol{\theta}) = \|M_k \delta\boldsymbol{\theta} + \mathbf{I}_{k+1}(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)) - \mathbf{I}_0(\mathcal{R})\|^2. \quad (6.17)$$

Cette expression atteint son minimum pour :

$$\delta\boldsymbol{\theta} = S_k \delta\mathbf{i} \quad (6.18)$$

où $\delta\mathbf{i} = \mathbf{I}_{k+1}(\varphi(\mathcal{R}, \boldsymbol{\theta}_k)) - \mathbf{I}_0(\mathcal{R})$ correspond à un vecteur d'erreur, et la matrice S_k est la pseudo-inverse de M_k . Un calcul direct de cette matrice amène à la solution :

$$\delta\boldsymbol{\theta} = -(M_k^t M_k)^{-1} M_k^t \delta\mathbf{i}, \quad (6.19)$$

et on obtient finalement l'expression du vecteur de paramètres de mouvement recherché :

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - (M_k^t M_k)^{-1} M_k^t \delta\mathbf{i}. \quad (6.20)$$

Le calcul direct de la matrice M_k nécessite le calcul des gradients de l'image, et doit être ré-itéré à chaque nouvelle image par une procédure coûteuse. Cependant, en supposant que l'estimée θ_k est exacte, il est possible d'exprimer M_k comme une fonction linéaire des gradients de l'image associée au motif de référence, ce qui permet d'obtenir $\delta\theta$ en quelques étapes simples.

Variantes et extensions par prise en compte de primitives visuelles

Une variante de cette approche a été proposée par Jurie et Dhome [Jurie 02]. Dans cet article, il est montré que le calcul de la pseudo-inverse S_k par l'expression $(M_k^t M_k)^{-1} M_k^t$ peut être avantageusement remplacé en utilisant une procédure d'apprentissage telle que :

$$\delta\theta = S_k \delta\mathbf{i}. \quad (6.21)$$

En effet, cette équation correspond à la modélisation de points dans un espace de dimension p par n hyperplans. En conséquence, les coefficients de la matrice S_k (qui sont les coefficients des hyperplans) peuvent être calculés par une estimation au sens des moindres carrés à partir de J couples ($J > n$) $\{\delta\mathbf{i}^j, \delta\theta^j\}_{j=1\dots J}$. Ces couples sont obtenus par une phase de simulation effectuée hors-ligne. Celle-ci consiste à déplacer J fois le motif de référence initial par un mouvement connu, et à calculer le vecteur d'image déplacée associé. Cette matrice est calculée une seule fois dans le repère du modèle de référence, à partir de l'image initiale. Au cours de la séquence, des changements de repère sont effectués de façon à toujours se trouver au voisinage de la position initiale. La procédure complète ainsi décrite permet un gain de temps de calcul significatif.

Citons également une variante récente de la méthode de Hager et Belhumeur qui utilise une minimisation du second ordre afin d'augmenter le taux de convergence de la procédure [Benhimane 04].

Afin de rendre ces méthodes plus résistantes aux occlusions et plus précises, il est possible de considérer, en plus d'une description de l'objet par son modèle de luminance, une description par un ensemble de primitives. Ces primitives sont généralement situées en des endroits de l'image à fort gradient, comme des points caractéristiques ou des lignes de contours. Elles peuvent aussi être de nature géométrique comme des droites ou des cercles qui décrivent la structure de l'objet. Une telle caractérisation est généralement employée pour du suivi d'objets non rigides [Torresani 02] [Isard 96]. Dans le cadre du suivi de plan, les méthodes différentielles ont ainsi été étendues à des méthodes hybrides. Ces dernières minimisent une fonction de coût qui dépend à la fois d'une distance entre primitives (généralement des contours) et entre modèles de luminance [Masson 04] [Pressigout 05].

Contrainte géométrique pour le suivi de motifs plans : transformation homographique

Les méthodes utilisant un modèle de luminance ont l'avantage de pouvoir capter des mouvements complexes. De plus, elles permettent d'intégrer facilement une contrainte géométrique afin de les rendre spécifiques et efficaces pour du suivi de motifs plans. Cette contrainte se traduit par exemple par la mise en œuvre d'un modèle de mouvement

homographique [Jurie 02] [Benhimane 04].

L'homographie est un principe de la géométrie projective [Faugeras 88]. On suppose disposer de deux images d'un objet plan, prises par deux caméras. La position et l'orientation de la seconde caméra par rapport à la première est définie par la matrice de rotation R et le vecteur de translation $\mathbf{t}$. Le plan $\mathcal{P}$ est défini par sa normale $\mathbf{n}$ et sa distance à l'origine d . Dans ce cas, les coordonnées 3D $\mathbf{X}_1 = (X_1, Y_1, Z_1)$ et $\mathbf{X}_2 = (X_2, Y_2, Z_2)$ d'un point $\mathbf{X}$ de l'espace dans les systèmes de coordonnées des deux caméras sont reliées par la relation :

$$\mathbf{X}_2 = R \mathbf{X}_1 + \mathbf{t}. \quad (6.22)$$

Si $\mathbf{X}$ appartient au plan $\mathcal{P}$, alors :

$$\mathbf{n}^t \mathbf{X}_1 = d, \quad (6.23)$$

et on en déduit :

$$\mathbf{X}_2 = \left(R + \frac{\mathbf{t}\mathbf{n}^t}{d} \right) \mathbf{X}_1. \quad (6.24)$$

Les projections du point $\mathbf{X}$ sur les deux images se trouvent respectivement aux coordonnées homogènes $\mathbf{x}_1 = \frac{\mathbf{X}_1}{Z_1}$ et $\mathbf{x}_2 = \frac{\mathbf{X}_2}{Z_2}$, d'où la relation suivante :

$$\mathbf{x}_2 = \frac{Z_1}{Z_2} \left(R + \frac{\mathbf{t}\mathbf{n}^t}{d} \right) \mathbf{x}_1 \quad (6.25)$$

$$\propto \mathcal{H} \mathbf{x}_1 \quad (6.26)$$

Cette relation est valide pour tous les points appartenant au plan $\mathcal{P}$. La matrice $\mathcal{H}$ est appelée la *matrice d'homographie* et est de dimension 3×3 . Elle correspond à un changement de base dans l'espace projectif, et permet de relier les points d'un plan entre deux images. Elle ne peut être déterminée qu'à un facteur d'échelle près. Elle a donc huit degrés de liberté, ce qui nécessite la connaissance d'au moins quatre couples de points mis en correspondance pour son estimation. Cette matrice est classiquement calculée par une décomposition SVD dans les cas non dégénérés.

L'utilisation de l'homographie peut être faite à travers le modèle de mouvement $\varphi(\mathbf{s}, \boldsymbol{\theta}_k)$. $\boldsymbol{\theta}_k$ est alors un vecteur de huit paramètres, qui correspondent aux entrées de la matrice d'homographie. Une idée alternative consiste à utiliser ce principe géométrique dans le cadre d'un filtre de Kalman pour prédire le mouvement du plan, qui sera corrigé par une mise en correspondance des primitives [Faugeras 88].

6.3 Synthèse

Dans la première partie de ce chapitre, nous avons présenté un état de l'art des algorithmes de suivi de points d'intérêt. Après une présentation succincte des détecteurs de points existants, nous nous sommes intéressés en particulier aux méthodes de suivi construites sur l'hypothèse de conservation de la luminance du point et de son voisinage. À partir de cette hypothèse, deux formalismes du problème sont possibles.

Les méthodes de corrélations considèrent une formulation directe de cette hypothèse sous la forme d'une mesure de distance. Leur principe est de rechercher la fenêtre de l'image courante minimisant cette distance avec le modèle de référence du point. Ces méthodes sont précises tant que les changements inter-images sont faibles. Afin de pallier le manque de robustesse aux changements importants des mesures de similarité existantes, il est intéressant de considérer des techniques de mise à jour du modèle de référence. Bien que coûteuses en temps de calcul, ces méthodes peuvent permettre de capter de grands déplacements du point suivi. Enfin, une mesure de confiance précise de ces résultats est disponible par l'étude de la surface de corrélation. Cette mesure a l'avantage d'être locale et ne dépend pas de la qualité des trajectoires reconstruites des autres points suivis.

Les méthodes différentielles sont basées sur un développement de Taylor d'une expression décrivant la conservation de la luminance du point et de son voisinage. Elles utilisent une paramétrisation du mouvement qui, de façon classique, est supposé translationnel ou affine. Elles sont rapides, et de nombreuses améliorations ont permis de les rendre moins sensibles aux changements d'illumination et géométriques, ainsi qu'à des grands déplacements. Cependant elles ne sont pas robustes aux occlusions. Les méthodes proposées pour la validation des estimées reposent sur des contraintes épipolaires ou sur l'étude de la répartition des erreurs d'estimation des différentes trajectoires reconstruites. Les premières ne sont pas valides pour des scènes réelles où des objets sont animés de mouvements indépendants. Les secondes ont l'inconvénient de supposer que la majorité des trajectoires ait été estimée avec succès. Ce problème de validation dans le cadre des méthodes différentielles reste ouvert. Il nous semble essentiel de considérer une mesure locale indépendante des autres points, telle que celle introduite par Tommasini [Tommasini 98].

Cette analyse met en évidence le besoin d'un nouvel algorithme de suivi de points caractéristiques robuste aux situations difficiles telles que des occlusions, des changements brusques de vitesse et de direction, ou des ambiguïtés engendrées par l'arrière-plan. Il apparaît également nécessaire de définir une structure algorithmique permettant de recouvrer la vraie trajectoire du pixel, même après quelques itérations de mauvaises estimations. Le cadre de filtrage nous semble approprié à ces besoins grâce à ses caractéristiques robustes décrites dans la partie précédente. De plus, l'utilisation d'une méthode de filtrage permet d'obtenir simplement une mesure de confiance sur les résultats obtenus, soit au travers de la matrice de covariance d'erreur dans le cas d'un algorithme de type filtre de Kalman, soit au travers de la répartition des particules dans le cas d'une méthode séquentielle de Monte Carlo. Afin de profiter des avantages des deux approches, nous proposons de les combiner dans ce cadre probabiliste. Les algorithmes obtenus sont décrits dans le chapitre suivant.

Dans la deuxième partie du chapitre, nous avons donné quelques éléments de l'existant pour le suivi d'objets plans. L'approche la plus courante constitue une extension des méthodes différentielles, et repose sur une description de l'objet par un modèle de luminance. À partir d'un modèle de mouvement, une procédure de minimisation est effectuée pour rechercher la nouvelle position. La difficulté de ces techniques réside dans le calcul de la matrice jacobienne qui relie les variations de l'image autour de la région d'intérêt en fonc-

tion d'un mouvement de l'objet. Plusieurs améliorations de ces approches ont été étudiées, comme par exemple l'introduction d'un modèle de mouvement de type homographie qui rend compte d'une contrainte géométrique forte, ou l'ajout d'une description de l'objet par un ensemble de primitives.

À notre connaissance, la combinaison de ces trois éléments dans un cadre probabiliste n'a jamais été étudiée. Dans le dernier chapitre, nous présenterons notre approche pour le suivi d'objets plans, développée à partir de l'étude menée pour la trajectographie de points caractéristiques. Nous verrons qu'il est possible de construire un modèle dont les entrées sont déterminées par une méthode différentielle, une mesure de distance entre primitives, et un calcul d'homographie.

Chapitre 7

Algorithmes proposés pour la trajectographie de points

Dans la partie II de ce document, nous avons 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 (appelés *filtres conditionnels*). Dans ce chapitre, nous abordons la validation de ces modèles pour du suivi de points dans une séquence. Le cadre de filtrage nous permet de construire des algorithmes robustes à des situations mettant en défaut les approches classiques décrites dans le chapitre 6.

On considère un point de la scène. Résoudre le problème de suivi de ce point dans la séquence équivaut à localiser la projection du point dans le plan image à chaque instant. Nous nous plaçons dans un contexte le plus général possible, où aucune information sur le mouvement de l'objet environnant n'est disponible. Or, comme démontré dans [Tissainayagam 01], la qualité des résultats dépend directement du choix d'un modèle de mouvement approprié. Pour contourner le problème du manque d'informations *a priori*, nous proposons de s'appuyer sur un modèle dynamique estimé sur la séquence d'images. Plus précisément, nous proposons d'associer une équation dynamique décrivant un mouvement estimé par une méthode différentielle et des observations obtenues par une procédure de mise en correspondance. Cette association nous permet de profiter des avantages des deux types de méthodes, tout en minimisant leurs inconvénients respectifs. Enfin, une estimation automatique des matrices de covariance de bruits d'état et de mesure permet de rendre adaptative l'importance de chacune d'elles dans l'estimation de la position du point, et ceci à chaque itération.

Dans les algorithmes proposés, l'état $\mathbf{x}_k$ désigne la position de la projection du point au temps k , dans l'image $\mathbf{I}_k$. $\mathbf{x}_k$ est donc un vecteur composé des coordonnées (u, v) d'un pixel dans l'image. L'état est supposé observable par une mesure $\mathbf{z}_k$, obtenue par un processus réalisé sur la séquence. Notons que comme nous considérons une équation d'évolution estimée sur les images, nous nous plaçons dans le cadre des systèmes conditionnels décrit dans le chapitre 5 (p. 97). Les algorithmes proposés sont donc construits sur la base des filtres conditionnels (§ 5.2, p. 100). Rappelons que dans ce cadre, le modèle dynamique est décrit par la distribution $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$ et le modèle de mesure par la distribution $p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k})$.

Les trois algorithmes proposés sont les suivants. Le premier est fondé sur un filtre linéaire. Il est particulièrement bien adapté aux séquences présentant un mouvement dominant. Le second repose sur un filtre non linéaire dédié au suivi de points dont le mouvement ne peut être décrit que localement. Le troisième est une extension de ce dernier adaptée au suivi dans des images comportant un fouillis important.

Dans ce chapitre, nous présenterons successivement l'estimation du modèle dynamique par une approche différentielle et le modèle d'observation qui utilise des mesures de corrélations. Puis, nous décrirons les étapes complémentaires nécessaires à la définition des deux premiers algorithmes, ainsi qu'une stratégie permettant de les associer. Ensuite, l'extension dédiée à la gestion des ambiguïtés sera expliquée, avant la présentation des résultats. Un paragraphe exposant une application de notre approche au suivi de visages terminera ce chapitre.

7.1 Modèle dynamique estimé par une approche différentielle

Dans cette section, nous décrivons le modèle dynamique d'un point représenté par la distribution de probabilité $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$. Ce modèle décrit l'évolution du pixel $\mathbf{x}_{k-1}$ entre les images successives observées $\mathbf{I}_{k-1}$ et $\mathbf{I}_k$. Dans le but d'être réactif à chaque changement de vitesse et de direction du point, nous proposons de définir cette loi au moyen d'une estimation de mouvement par une méthode différentielle. La méthode utilisée est une technique robuste d'estimation paramétrique de mouvement proposée par Odobez et Bouthemy [Odobez 95a]. À chaque nouvelle image disponible, cette technique fournit une estimation du vecteur vitesse du point caractéristique suivi.

7.1.1 Estimation robuste d'un mouvement paramétrique

La mise en œuvre d'une technique d'estimation robuste et multirésolution permet d'estimer de façon fiable un modèle paramétrique de mouvement. Ce modèle paramétrique représente le champ de vitesse dominant apparent sur un support donné $\mathcal{R}$. Le mouvement d'un point $\mathbf{s}$ de ce support entre les instants $k-1$ et k est modélisé par une fonction polynomiale des coordonnées du point :

$$\boldsymbol{\omega}_{k|\mathbf{I}}(\mathbf{s}) = P(\mathbf{s}) \boldsymbol{\theta}_{k|\mathbf{I}} \quad (7.1)$$

où $\boldsymbol{\omega}_{k|\mathbf{I}}(\mathbf{s})$ désigne le vecteur vitesse estimé du pixel $\mathbf{s} = (u_{\mathbf{s}}, v_{\mathbf{s}})^t$. $\boldsymbol{\theta}_{k|\mathbf{I}}$ est le vecteur de paramètres de mouvement qui sont les coefficients du polynôme. $P(\mathbf{s})$ est la matrice relative au modèle paramétrique choisi. Ses entrées dépendent des coordonnées $(u_{\mathbf{s}}, v_{\mathbf{s}})$ du pixel considéré. En particulier, un modèle affine à six paramètres est associé à la matrice suivante :

$$P(\mathbf{s}) = \begin{bmatrix} 1 & u_{\mathbf{s}} & v_{\mathbf{s}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & u_{\mathbf{s}} & v_{\mathbf{s}} \end{bmatrix}. \quad (7.2)$$

Le vecteur de paramètres $\boldsymbol{\theta}_{k|\mathbf{I}}$ est estimé par la minimisation d'une fonction $\phi_1(\boldsymbol{\theta})$:

$$\phi_1(\boldsymbol{\theta}) = \sum_{\mathbf{s} \in \mathcal{R}} \rho [\mathbf{I}_k(\mathbf{s} + P(\mathbf{s}) \boldsymbol{\theta}) - \mathbf{I}_{k-1}(\mathbf{s})], \quad (7.3)$$

où ρ est une fonction de coût non quadratique permettant de ne pas perturber l'estimation par des points aberrants. Ces point aberrants sont identifiés comme appartenant à des régions ne correspondant pas au mouvement estimé (régions de mouvement secondaire) ou pour lesquelles l'hypothèse de conservation de la luminance n'est pas valide. La minimisation est réalisée par une procédure de type Gauss-Newton dans un schéma multirésolution. Ce schéma permet de capter des déplacements de grandes amplitudes. Il consiste en une approche incrémentale, où des développements de Taylor successifs de ρ sont appliqués autour de l'estimation précédente $\hat{\theta}$. L'estimation de l'incrément $\delta\theta$ est alors obtenue par la minimisation de :

$$\phi_2(\delta\theta) = \sum_{s \in \mathcal{R}} \rho [\mathbf{I}_k(\mathbf{s} + P(\mathbf{s}) \hat{\theta}) - \mathbf{I}_{k-1}(\mathbf{s}) + \mathbf{g}_k(\mathbf{s} + P(\mathbf{s}) \hat{\theta})^t P(\mathbf{s}) \delta\theta]. \quad (7.4)$$

On rappelle que $\mathbf{g}_k^t = [\partial_{\mathbf{u}} \mathbf{I}_k, \partial_{\mathbf{v}} \mathbf{I}_k]$ est le gradient spatial de $\mathbf{I}_k$. La procédure de minimisation mise en place s'inscrit dans une stratégie pyramidale et est réalisée par une technique de moindres carrés pondérés itérés à chaque niveau de la pyramide.

Pour plus de détails sur cette méthode, nous renvoyons à l'article [Odobez 95a] et à la thèse [Odobez 94].

7.1.2 Équation dynamique linéaire ou non linéaire ?

Pour notre application, l'utilisation d'une telle méthode sur un support approprié $\mathcal{R}$ contenant le point d'intérêt $\mathbf{x}_{k-1}$, fournit une estimation du vecteur vitesse $\omega_{k|\mathbf{I}}(\mathbf{x}_{k-1})$ au point $\mathbf{x}_{k-1}$ entre les instants $k-1$ et k . Cela nous permet de définir l'équation dynamique suivante :

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_{k-1} + \omega_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ &= \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \theta_{k|\mathbf{I}} + \mathbf{w}_k, \end{aligned} \quad (7.5)$$

où $\mathbf{w}_k$ est un bruit blanc de moyenne nulle. Il est important de noter que le vecteur vitesse $\omega_{k|\mathbf{I}}(\mathbf{x}_{k-1})$ correspond au mouvement dominant du support d'estimation $\mathcal{R}$. Une des difficultés de cette méthode est alors de définir un support d'estimation approprié pour un point donné $\mathbf{x}_{k-1}$. Deux cas peuvent être distingués.

1. Quand le mouvement du point à suivre correspond au mouvement dominant global (par exemple, le point d'intérêt appartient à l'arrière plan de l'image), le support d'estimation est défini comme étant la grille d'image complète. Dans ce cas, l'estimation du vecteur de paramètres de mouvement $\theta_{k|\mathbf{I}}$ ne dépend pas de la position de $\mathbf{x}_{k-1}$ et l'équation dynamique est linéaire. Pour un modèle de mouvement affine, l'équation d'évolution (7.5) devient :

$$\mathbf{x}_k = F_{k|\mathbf{I}} \mathbf{x}_{k-1} + \mathbf{b}_{k|\mathbf{I}} + \mathbf{w}_k, \quad (7.6)$$

où $F_{k|\mathbf{I}}$ est la matrice relative aux rotation, divergence et cisaillement, et $\mathbf{b}_{k|\mathbf{I}}$ est le vecteur de translation. La variable de bruit $\mathbf{w}_k$ rend compte des erreurs liées au modèle *global* de vitesse. Par manque d'information, $\mathbf{w}_k$ est supposé être un bruit blanc de moyenne nulle et de covariance $Q_{k|\mathbf{I}}$ conditionnellement à $\mathbf{I}_{0:k}$. Il est essentiel de remarquer que l'estimation des paramètres de mouvement sur l'image entière permet

de considérer une information globale sur l'évolution du point. Un tel type d'information – qui, à notre connaissance n'est jamais pris en compte dans les méthodes de suivi de points classiques – a une importance cruciale lors d'un manque d'information locale (par exemple, en cas d'occlusion ou face à une séquence d'images très bruitée).

2. Quand le mouvement du point suivi ne peut être décrit que par un modèle paramétrique local (par exemple, un point appartenant à un objet mobile), le support $\mathcal{R}$ est fixé à un petit domaine centré en $\mathbf{x}_{k-1}$. En conséquence, le vecteur de paramètres estimé $\boldsymbol{\theta}_{k|\mathbf{I}}$ dépend de $\mathbf{x}_{k-1}$ et sera donc noté $\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1})$. L'équation dynamique considérée devient alors 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 conditionnellement à $\mathbf{I}_{0:k}$, i.e. $p(\mathbf{w}_k|\mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{w}_k; 0, Q_{k|\mathbf{I}})$. L'équation d'état non linéaire peut se traduire par l'expression suivante :

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k ; \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1})\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}). \quad (7.7)$$

7.1.3 Choix de la covariance du bruit d'état

Par l'utilisation d'une méthode robuste d'estimation de mouvement, il est possible d'accéder à une information révélatrice de la qualité de l'estimation. Cette information repose sur la taille du support de l'estimation. En effet, la fonction de coût robuste ρ de l'estimateur fournit en sortie de la procédure le support de validité de l'estimation. Plus précisément, chaque point du support $\mathcal{R}$ se voit attribuer un poids qui est fonction de l'importance de sa participation au mouvement global estimé. La méthode robuste donne ainsi l'ensemble des points pour lesquels le mouvement estimé est exact.

Il apparaît ainsi intéressant d'étudier la variation au cours du temps du rapport entre la taille du support de validité et la taille du support d'estimation $\mathcal{R}$ (choisi de taille fixe le long de la séquence). Ce rapport est noté τ . Un rapport proche de 1 rend compte d'une estimation parfaite. Une chute soudaine de cette valeur décrit une estimation fautive qui peut être due à une occlusion ou à un mouvement trop brusque. Afin de détecter un tel saut, nous proposons d'assimiler ce problème à celui de la détection d'une transition dans un signal. Pour cela, nous utilisons un test de Page-Hinkley qui est une technique issue de traitement du signal [Hinkley 71] [Basseville 88]. Ce test permet de signaler les sauts croissants et décroissants d'un signal. Il est décrit en détail dans un paragraphe suivant.

La figure 7.1 met en évidence l'intérêt de l'étude du rapport τ . Cette figure illustre l'évolution de τ sur une séquence très longue comportant plusieurs changements de plans. L'estimation de mouvement est ici réalisée sur l'image complète. Chaque décrochage du signal correspond à un changement de plan, et donc à une mauvaise estimation du mouvement entre les deux images successives autour de la coupure. Le test de Hinkley génère alors deux alarmes, une pour le décrochage décroissant et l'autre pour le décrochage croissant. Une deuxième illustration est donnée par la figure 7.2. La séquence étudiée représente un visage pouvant passer de face à profil et inversement, et l'estimation de mouvement est effectuée sur une fenêtre autour du visage (cadre blanc). Chacun de ces passages engendre une mauvaise estimation (du fait d'une trop grande différence entre les images successives) qui se répercute sur le rapport τ , et qui produit une alarme du test de Page-Hinkley.

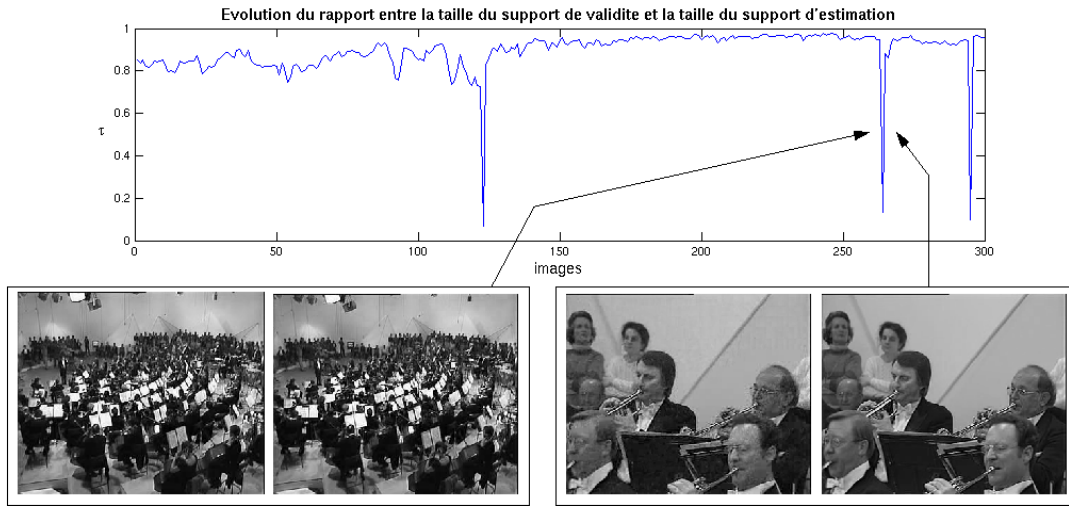


FIG. 7.1: Mise en évidence de la pertinence du rapport entre la taille du support de validité et la taille du support d'estimation pour déterminer la qualité de l'estimation de mouvement. La séquence traitée est longue et comporte plusieurs changements de plans. L'estimation du mouvement est réalisée sur l'image complète. Chaque décrochage désigne une mauvaise estimation due à une coupure. Le test de Page-Hinkley permet de détecter le décrochage du signal.

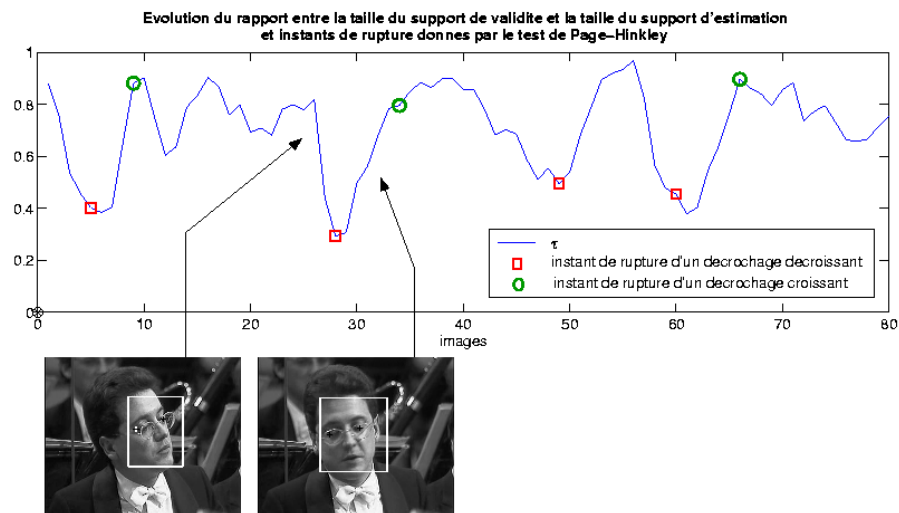


FIG. 7.2: Mise en évidence de la pertinence du rapport entre la taille du support de validité et la taille du support d'estimation et de l'utilisation du test de Page-Hinkley. La séquence utilisée représente un visage et l'estimation de mouvement est réalisée sur une fenêtre autour de celui-ci (cadre blanc). Chaque passage de face à profil et inversement entraîne une mauvaise estimation et le test de Page-Hinkley émet une alarme.

Dans le cadre de notre algorithme de suivi, nous proposons de modifier la matrice de covariance de bruit d'état en fonction de ce signal. Nous lui attribuons deux valeurs possibles. En supposant une bonne estimation entre les deux premières images, la matrice $Q_{k|\mathbf{I}}$ est initialisée à des valeurs faibles (expérimentalement, nous utilisons $Q_{k|\mathbf{I}} = 2 \times Id$). Dès qu'une alarme pour un décrochage décroissant est émise par le test, nous la modifions en posant $Q_{k|\mathbf{I}} = \infty \times Id$. De manière similaire, une alarme pour un décrochage croissant permet de re-initialiser cette matrice à $Q_{k|\mathbf{I}} = 2 \times Id$.

Test de Page-Hinkley

Le test de Page-Hinkley est un test à somme cumulée qui permet de détecter les changements dans un signal. Ce test est connu pour sa robustesse (car il permet de prendre en compte l'intégralité du passé du signal), son efficacité et son faible coût de calcul. Il permet une détection précise de la position du saut, et permet de traiter à la fois les changements abrupts et progressifs sans modification du seuil interne de détection.

Ce test repose sur le changement de la moyenne du signal pour détecter ses variations. Pour un signal discret $\tau_0, \tau_1, \dots, \tau_N$, il s'applique itérativement pour chacune des valeurs. Pour une valeur courante τ_i , deux tests sont réalisés en parallèle pour détecter les sauts décroissants ou croissants. Ces détecteurs sont définis respectivement par :

$$\text{détecteur de saut décroissant : } \begin{cases} M_0 = 0 \\ M_i = \sum_{t=1}^i (\tau_t - \mu + \frac{\nu}{2}) \\ m_i = \max_{0 \leq t \leq i} M_t \\ \text{alarme si } m_i - M_i > \lambda \end{cases} \quad (7.8)$$

$$\text{détecteur de saut croissant : } \begin{cases} U_0 = 0 \\ U_i = \sum_{t=1}^i (\tau_t - \mu - \frac{\nu}{2}) \\ u_i = \min_{0 \leq t \leq i} U_t \\ \text{alarme si } U_i - u_i > \lambda \end{cases} \quad (7.9)$$

où la moyenne du signal μ est estimée en ligne et réinitialisée après chaque saut. ν et λ sont les deux paramètres internes du test. μ peut être interprété comme l'écart toléré *a priori* à la moyenne et λ est un seuil prédéfini (afin d'être sûr de ne pas manquer les décrochages, ces seuils ont été fixés expérimentalement à $\mu = 0.3$ et $\lambda = 0.3$). Le test émet une alarme dès qu'un changement dans la moyenne du signal est détecté, et l'instant de rupture correspond au dernier indice i pour lequel $m_i = M_i$ ou $u_i = U_i$.

Remarque : En perspectives, pour un calcul plus précis de la matrice $Q_{k|\mathbf{I}}$, il serait intéressant de considérer, en plus de l'information apportée par le rapport τ , la carte des poids donnée par l'estimateur robuste. Une idée serait par exemple d'utiliser la matrice de covariance associée au vecteur de paramètres du mouvement [Gelgon 98].

7.2 Modèle d'observation utilisant une mesure de similarité

Le modèle d'observation est défini par la distribution $p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k})$. Alors que l'équation dynamique décrit l'évolution du point d'intérêt entre deux images successives, le modèle d'observation permet l'acquisition d'une information entre l'image initiale et l'image courante. Cette information pourra éviter une éventuelle dérive des résultats. Elle permettra également de recouvrer la vraie trajectoire après une mauvaise estimation, due par exemple à une occlusion du point suivi.

Au temps k , on suppose que $\mathbf{x}_k$ est observable via un processus de mise en correspondance. Ce processus fournit le point le plus similaire au point initial $\mathbf{x}_0$ dans l'image $\mathbf{I}_k$, grâce à l'utilisation d'une caractérisation de référence notée $\tilde{\mathbf{I}}_0$. Le résultat de ce processus correspond à un pic de corrélation, dont la position définit la mesure $\mathbf{z}_k$ du système. Le modèle de référence $\tilde{\mathbf{I}}_0$ est défini comme étant l'image initiale $\mathbf{I}_0$, éventuellement mise à jour par recalage en cas de déformations géométriques trop importantes autour du point à suivre le long de la séquence. La procédure de mise à jour de cette référence sera décrite dans un paragraphe spécifique (§7.2.2).

De nombreux critères de similarité peuvent être utilisés pour quantifier la similarité entre le point suivi et les points candidats. L'hypothèse de conservation de la luminance a simplement amené à considérer la mesure de SSD (ce choix sera justifié dans le paragraphe 7.2.1 décrivant l'estimation de la covariance du bruit de mesure). La mesure $\mathbf{z}_k$ est recherchée dans une zone d'intérêt notée ROI ("region of interest") et est définie par :

$$\mathbf{z}_k = \arg \min_{\mathbf{z} \in \text{ROI}} \underbrace{\sum_{\mathbf{m} \in \mathcal{W}} [\tilde{\mathbf{I}}_0(\mathbf{x}_0 + \mathbf{m}) - \mathbf{I}_k(\mathbf{z} + \mathbf{m})]^2}_{r_k(\mathbf{z})}. \quad (7.10)$$

$r_k(\mathbf{z})$ dénote le résidu calculé sur un petit voisinage $\mathcal{W}$ autour de $\mathbf{z}$, de taille $n \times n$. On suppose que cette mesure transporte suffisamment d'information pour pouvoir écrire que $\mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k$. La variable de bruit $\mathbf{v}_k$ modélise une erreur locale d'estimation et rend compte d'une déviation par rapport à la mesure idéale. Sa distribution est $p(\mathbf{v}_k | \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{v}_k; 0, R_{k|\mathbf{I}})$. L'équation d'observation est alors simplement modélisée par l'expression suivante :

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k; \mathbf{x}_k, R_{k|\mathbf{I}}). \quad (7.11)$$

Cette recherche de la mesure est effectuée à chaque nouvelle image disponible. Notons que contrairement aux méthodes classiques, notre cadre de filtrage permet de définir très simplement la zone de recherche ROI. En effet, celle-ci correspond à la fenêtre de validation de la nouvelle mesure.

7.2.1 Estimation de la covariance du bruit de mesure

Une bonne estimation de la covariance du bruit de mesure $R_{k|\mathbf{I}}$ est essentielle pour rendre la méthode de suivi plus robuste aux observations corrompues. De nombreux facteurs peuvent affecter la qualité des observations, tels que le bruit d'acquisition, des occlusions ou de grands changements géométriques et photométriques. Une telle

estimation revient à mesurer la confiance du résultat donné par la méthode de corrélation. C'est pourquoi, dans le but d'estimer la matrice $R_{k|\mathbf{I}}$, nous proposons de reprendre les principes des méthodes de validation des techniques de mise en correspondance.

Pour cela, une surface de corrélation SSDS est construite à partir des résidus $r_k(\mathbf{z})$ sur un support $\mathcal{W}'$ de taille $n' \times n'$ autour de l'observation $\mathbf{z}_k$. Notre approche reprend l'idée développée dans [Singh 92]. Rappelons que cette idée consiste à transformer la SSDS – qui correspond à une surface d'erreur – en une distribution réponse construite en chaque point $\mathbf{z}$ du support $\mathcal{W}'$ par :

$$\mathcal{D}_k(\mathbf{z}) = \exp(-c r_k(\mathbf{z})), \quad (7.12)$$

où c est un facteur de normalisation. Comme dans [Nickels 98], c est déterminé par une procédure itérative telle que $\sum_{\mathbf{z} \in \mathcal{W}'} \mathcal{D}_k(\mathbf{z}) = 1$. On fait alors l'hypothèse que cette distribution correspond à la probabilité d'avoir localisé l'observation exacte. La matrice de covariance du bruit $R_{k|\mathbf{I}}$, associée à la mesure $\mathbf{z}_k$, est estimée par un calcul empirique sur cette distribution (7.12) :

$$R_{k|\mathbf{I}} = \begin{pmatrix} \sum_{\mathbf{z} \in \mathcal{W}'} \mathcal{D}_k(\mathbf{z})(u_{\mathbf{z}} - u_{\mathbf{z}_k})^2 & \sum_{\mathbf{z} \in \mathcal{W}'} \mathcal{D}_k(\mathbf{z})(u_{\mathbf{z}} - u_{\mathbf{z}_k})(v_{\mathbf{z}} - v_{\mathbf{z}_k}) \\ \sum_{\mathbf{z} \in \mathcal{W}'} \mathcal{D}_k(\mathbf{z})(u_{\mathbf{z}} - u_{\mathbf{z}_k})(v_{\mathbf{z}} - v_{\mathbf{z}_k}) & \sum_{\mathbf{z} \in \mathcal{W}'} \mathcal{D}_k(\mathbf{z})(v_{\mathbf{z}} - v_{\mathbf{z}_k})^2 \end{pmatrix}, \quad (7.13)$$

où $\mathbf{z} = (u_{\mathbf{z}}, v_{\mathbf{z}})$ et $\mathbf{z}_k = (u_{\mathbf{z}_k}, v_{\mathbf{z}_k})$. Une telle modélisation nous permet de définir une ellipse adaptative d'incertitude autour de l'observation. Notons que nous ne recalculons pas la moyenne sur la surface. Nous supposons que cette moyenne se situe en $\mathbf{z}_k$, ce qui nous permet d'estimer la confiance en cette position.

Cette procédure est efficace dans les situations où la mise en correspondance n'est pas perturbée par des situations difficiles (bruit d'acquisition, occlusions, ambiguïtés générées par l'arrière plan, etc.). Cependant, face à de telles difficultés, se limiter au processus décrit ne permet pas une estimation précise de la covariance. Afin de rendre plus efficace la procédure, nous proposons l'ajout de deux tests statistiques, qui se révèlent utiles dans deux situations précises.

1. Tout d'abord, face à des séquences d'images très bruitées, la surface de corrélation peut présenter plusieurs pics dont l'amplitude peut être inférieure au niveau de bruit. Aucun de ces pics ne doit avoir un effet prédominant sur le calcul de la covariance. Nous proposons donc de les détecter et d'égaliser la valeur de la SSDS qui leur est associée. La détection de ces pics non significatifs repose sur une approximation de la distribution des résidus $r_k(\mathbf{z})$ (notons que l'approximation utilisée est valable uniquement si la mesure de corrélation utilisée est la SSD, ce qui justifie ce choix par rapport à des mesures de similarité plus complexes ne permettant pas la réalisation du test).

Dans [Tommasini 98], il est montré que les résidus calculés par une SSD peuvent être supposés gaussiens. Nous reprenons cette idée en la modifiant légèrement. Sous l'hypothèse de conservation de la luminance entre deux points mis en correspondance, à un bruit blanc gaussien près dû au bruit d'acquisition, on a :

$$\mathbf{I}_k(\mathbf{z}) - \mathbf{I}_0(\mathbf{x}_0) \rightsquigarrow \mathcal{N}(0, \sigma^2) \quad (7.14)$$

où σ^2 correspond à la variance du bruit d'acquisition. Cette constante peut être soit fixée *a priori*, ou estimée en suivant la méthode décrite dans [Immerkær 96] (nous la détaillerons dans un paragraphe suivant). On en déduit que pour deux régions de taille $n \times n$ mises en correspondance, $r_k(\mathbf{z})/\sigma$ est distribué selon une distribution du Chi-deux à n^2 degrés de liberté (on rappelle que $n \times n$ est la taille du support utilisé pour le calcul de $r_k(\mathbf{z})$). En considérant un support d'estimation suffisamment grand, grâce à l'approximation de Fisher [Saporta 90], on peut faire l'approximation suivante :

$$\sqrt{2 * r_k(\mathbf{z})/\sigma^2} - \sqrt{2 * n^2} \rightsquigarrow \mathcal{N}(0, 1). \quad (7.15)$$

En pratique, un test statistique est réalisé en chaque point de la surface de corrélation pour savoir s'il vérifie la loi (7.15) [Saporta 90]. Chaque point détecté comme suivant cette loi voit sa valeur de SSDs modifiée et seuillée au même niveau (i.e. à la valeur correspondant au résidu le plus faible). Cela signifie qu'il leur est attribué la même probabilité d'être la "vraie" observation.

Ainsi, cette étape nous permet d'estimer une meilleure matrice de covariance dans le cas de surfaces de corrélation présentant plusieurs pics de faibles amplitudes. Par la suite, le test de la loi (7.15) est appelé *test des résidus*. Il est réalisé en pratique au niveau 0.05 en utilisant la table de loi du Chi-deux. Une illustration du gain de précision apporté par cette procédure est présentée dans la figure 7.5 (l'ensemble des illustrations de l'estimation de la covariance du bruit de mesure est détaillé ultérieurement).

2. Le second problème intervient lorsqu'approcher la distribution réponse par une loi gaussienne 2D n'a pas de sens. Cela est le cas si la surface d'erreur présente de nombreux pics significatifs, d'amplitudes différentes. Ces pics sont révélateurs de situations ambiguës, lorsque la probabilité d'erreur de localisation de la mesure est forte. Cette situation apparaît lors d'occlusions ou de mises en correspondance de points qui se trouvent dans des zones fortement texturées. L'estimation de la covariance décrite par (7.13) n'est alors pas pertinente. En effet, la surface correspondante $\mathcal{D}_k$ peut être très lisse et mieux approchée par une distribution uniforme. Afin d'éviter une mauvaise approximation, un test d'ajustement du Chi-deux [Saporta 90] est réalisé. Un tel test permet de vérifier qu'un échantillon (l'ensemble des valeurs discrètes de la distribution réponse dans notre cas) provient ou non d'une variable aléatoire de distribution connue. Nous l'utilisons ici pour tester si la distribution d'intérêt correspond à une distribution uniforme (en pratique, nous le réalisons au niveau 0.1). Si l'hypothèse est acceptée, la matrice $R_{k|\mathbf{I}}$ est supposée diagonale, avec ses coefficients de variance fixés à l'infini. Une illustration de cette situation est présentée dans la figure 7.4. Si l'hypothèse de loi uniforme est rejetée, nous supposons que la distribution est bien approchée par une loi gaussienne, et la matrice est calculée avec (7.13).

Estimation de la variance du bruit d'acquisition

Pour l'estimation de la variance σ^2 du bruit d'une image $\mathbf{I}$ de taille $row \times col$, nous utilisons la méthode décrite dans [Immerkær 96]. Cet article présente une méthode simple

et rapide dans le cas d'un bruit d'acquisition supposé additif, gaussien, de moyenne nulle. L'expression calculée est :

$$\sigma = \sqrt{\frac{\pi}{2}} \frac{1}{6(row-2)(col-2)} \sum_{(u,v) \in \mathbf{I}} |\mathbf{I}(u,v) \otimes \mathbf{M}|, \quad (7.16)$$

où $\mathbf{M}$ est le masque suivant :

$$\mathbf{M} = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}. \quad (7.17)$$

Cette estimation est réalisée sur la première image, et est supposée valable sur la séquence entière. Rappelons que la variance du bruit intervient dans les tests de résidus.

Récapitulatif du processus d'observation

Au final, le processus d'observation consiste en la succession des 6 étapes suivantes :

1. estimation de la mesure $\mathbf{z}_k$ par (7.10) ;
2. construction de la SSDS sur un support $\mathcal{W}$ centré autour de $\mathbf{z}_k$;
3. détection des pics non significatifs par les tests de résidus qui utilisent (7.15), et seuillage de la SSDS dans le cas où les résidus sont dus à du bruit d'acquisition ;
4. construction de la distribution réponse par (7.12) à partir de la SSDS modifiée ;
5. utilisation du test d'ajustement du Chi-deux pour contrôler l'adéquation de la loi uniforme sur cette distribution (dans le cas contraire, on suppose qu'une loi gaussienne est adaptée) ;
6. estimation de la matrice de covariance du bruit de mesure $R_{k|\mathbf{I}}$.

Des illustrations des étapes nécessaires à l'estimation de la covariance du bruit de mesure sont présentées dans les figures 7.3, 7.4 et 7.5. Ces figures correspondent à des situations typiques. Dans chacune d'elles sont représentés : (a) l'image de référence, le point à mettre en correspondance, et l'observation obtenue dans l'image courante, (b) la SSDS correspondant aux deux points mis en correspondance et (c) la distribution réponse associée. Pour illustrer les tests de résidus quand ceux-ci ont un impact sur l'estimation, nous présentons également le cas échéant (d) la SSDS modifiée et (e) la distribution réponse associée sur laquelle la matrice de covariance est finalement estimée.

Le premier exemple (figure 7.3) présente un cas idéal de mise en correspondance : le point d'intérêt est bien caractérisé par une image de référence qui ne pose pas de problème d'ambiguïtés, la séquence est non bruitée et le voisinage du point n'a pas subi de larges déformations photométriques ou géométriques. Dans ce cas, la SSDS n'est pas modifiée par les tests de résidus, et la distribution réponse est bien approchée par une loi gaussienne. La matrice de covariance est alors calculée par (7.13). On obtient finalement

$$R_{k|\mathbf{I}} = \begin{bmatrix} 0.45 & -0.12 \\ -0.12 & 0.37 \end{bmatrix}.$$

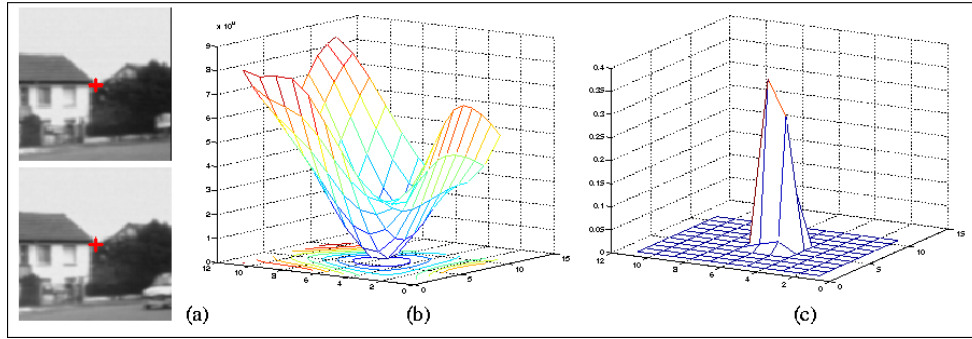


FIG. 7.3: Estimation de la covariance du bruit de mesure (niveau du bruit fixé à $\sigma = 1$) (a) image en haut : caractérisation de référence et point à mettre en correspondance (coin du toit), image en bas : image courante et position estimée ; (b) SSDs correspondant au deux points mis en correspondance ; (c) distribution réponse associée. Dans cette situation, l'hypothèse de loi uniforme est rejetée. La distribution réponse est approchée par une loi gaussienne 2D, et la matrice de covariance est estimée par (7.13).

Une situation d'occlusion est illustrée par la figure 7.4. Celle-ci montre l'intérêt du test d'ajustement du Chi-deux réalisé sur la distribution réponse. Dans cette situation, la distribution réponse comporte de nombreux résidus importants, correspondant à des possibilités d'erreurs. L'hypothèse de loi uniforme est alors validée, et la covariance du bruit de mesure est fixée à $R_{k|\mathbf{I}} = \infty \times Id$.

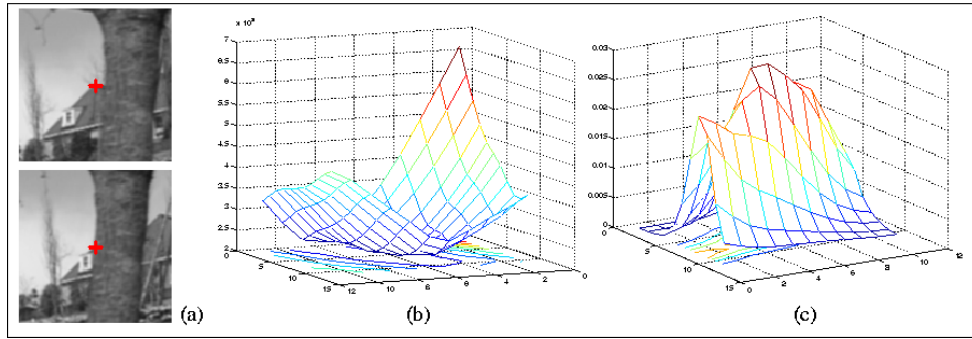


FIG. 7.4: Estimation de la covariance du bruit de mesure dans une situation d'occlusion : mise en évidence de l'intérêt du test d'ajustement (niveau du bruit fixé à $\sigma = 1$) (a) image en haut : caractérisation de référence et point à mettre en correspondance (sommet du toit), image en bas : image courante et position estimée ; (b) SSDs correspondant au deux points mis en correspondance ; (c) distribution réponse associée. Dans cette situation, l'hypothèse de loi uniforme est acceptée.

Le dernier exemple représenté par la figure 7.5 démontre l'importance de l'utilisation des tests de résidus face à des séquences d'images très bruitées, ainsi que l'influence du choix de la variance de bruit d'acquisition σ^2 . Dans cette situation, il est clair que la mise en œuvre d'une méthode de corrélation donne de mauvais résultats pour la plupart des

points de l'image à cause d'un trop fort bruit d'acquisition. Si on pose $\sigma = 1$, la SSDS n'est pas modifiée, et la distribution réponse résultante est approchée par une loi gaussienne (cf. surface 7.5 (c)). Les valeurs de la matrice de covariance sont alors relativement faibles : $R_{k|\mathbf{I}} = \begin{bmatrix} 2.01 & -0.95 \\ -0.95 & 4.34 \end{bmatrix}$. Ce résultat n'est pas pertinent puisqu'il ne retranscrit pas la forte incertitude sur l'observation obtenue. Par contre, en fixant $\sigma = 5$, la SSDS est modifiée (cf. surface 7.5 (d)), et l'hypothèse de loi uniforme pour la distribution réponse associée est validée (cf. surface 7.5 (e)). En effet, une telle surface caractérise plus précisément l'incertitude sur la mesure, et on a $R_{k|\mathbf{I}} = \infty \times Id$.

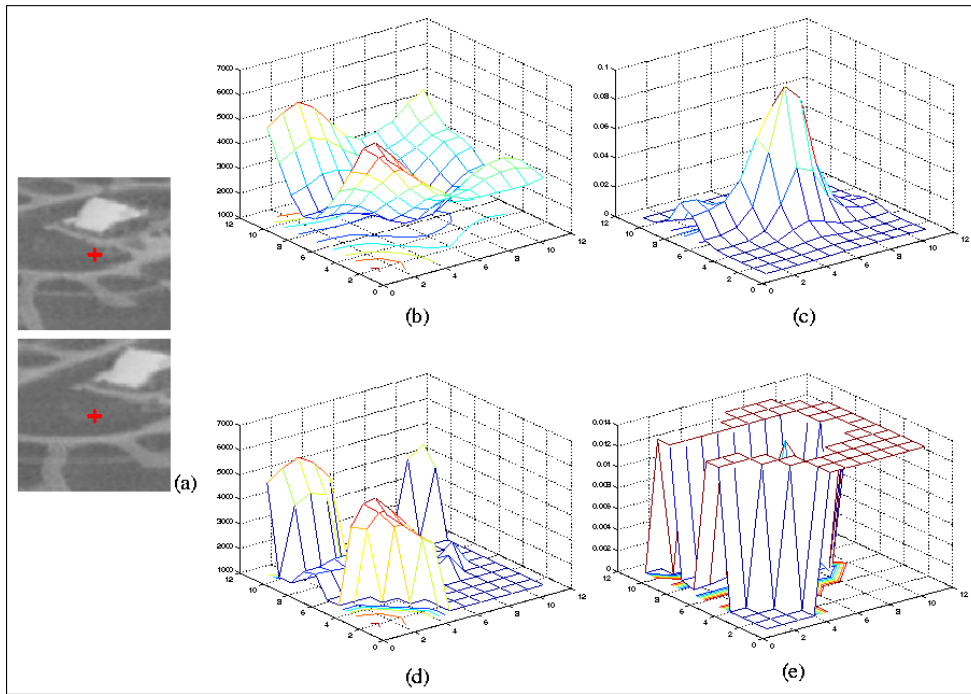


FIG. 7.5: Estimation de la covariance du bruit de mesure dans le cas de séquences bruitées : mise en évidence de l'intérêt des tests de résidus. (a) image en haut : caractérisation de référence et point à mettre en correspondance, image en bas : image courante et position estimée ; (b) SSDS correspondant au deux points mis en correspondance ; (c) la distribution réponse associée, correspondant à $\sigma = 1$ est approchée par une loi gaussienne, ce qui ne retranscrit pas l'incertitude observée ; (d) SSDS modifiée après les tests de résidus avec $\sigma = 5$; (e) la distribution réponse résultante est bien approchée par une loi uniforme.

Ces illustrations démontrent que l'utilisation des tests statistiques proposés (tests de résidus et test d'ajustement) en plus d'une modélisation gaussienne de la SSDS permet d'améliorer significativement l'estimation de la covariance du bruit de mesure. En particulier, ces tests détectent de manière robuste les occlusions et les situations ambiguës.

7.2.2 Mise à jour de la caractérisation de référence

Comme décrit dans le chapitre précédent, la mise à jour de la caractérisation de référence est une étape essentielle pour rendre l'algorithme robuste aux changements géométriques et photométriques au voisinage du point d'intérêt. Rappelons que cette image de référence est notée $\tilde{\mathbf{I}}_0$. Elle sert de modèle de luminance devant être retrouvé dans l'image courante, et est utilisée dans la phase de mise en correspondance entre le point initial $\mathbf{x}_0$ et la mesure $\mathbf{z}_k$. Le choix de $\tilde{\mathbf{I}}_0$ est d'une importance cruciale pour la détection de l'observation. En effet, se limiter à fixer $\tilde{\mathbf{I}}_0 = \mathbf{I}_0$ peut être problématique dans le cas de longues séquences où les changements géométriques et photométriques sont importants. Cela est d'autant plus vrai que la mesure de corrélation utilisée (dans notre cas la SSD) n'est pas robuste à ces situations. Cette référence doit être mise à jour afin de suivre l'évolution du voisinage du point suivi. Dans ce but, il est nécessaire de déterminer *quand* et *comment* mettre à jour le modèle de luminance.

La mise à jour de la référence doit utiliser le modèle de luminance des positions estimées. En conséquence, la règle de décision pour sa réalisation au temps k doit être fonction de la précision sur la position estimée $\hat{\mathbf{x}}_k$. Grâce au cadre de filtrage, une mesure de confiance de cette estimée est accessible. Dans le cas linéaire (qui correspond à un point appartenant au mouvement dominant), l'algorithme utilisé est un filtre linéaire conditionnel. La qualité de l'estimée est simplement apportée par la matrice de covariance de l'erreur $\Sigma_{k|k}$. La procédure est donc réalisée lorsque les valeurs propres de cette matrice sont inférieures à un seuil (fixé en pratique à 1). Dans le cas non linéaire (qui correspond à un point appartenant à un objet mobile), l'algorithme utilisé est un filtre non linéaire conditionnel. La procédure est alors identique, à la différence près que la covariance d'erreur est calculée empiriquement à partir du nuage de particules (nous reviendrons sur le choix de la structure algorithmique en fonction des points d'intérêt dans le paragraphe 7.3).

À l'initialisation, la caractérisation de référence $\tilde{\mathbf{I}}_0$ est définie comme étant une partie de l'image initiale $\mathbf{I}_0$, centrée en $\mathbf{x}_0$. La procédure de mise à jour proposée repose sur un recalage de cette image par rapport au modèle de luminance de la nouvelle position estimée. Au temps k , si l'estimée est suffisamment précise, un modèle de mouvement paramétrique est estimé entre la référence courante $\tilde{\mathbf{I}}_0$ centrée en $\mathbf{x}_0$, et une fenêtre de $\mathbf{I}_k$ centrée en $\hat{\mathbf{x}}_k$ de même dimension. On note $M_{\tilde{\mathbf{I}}_0 \rightarrow \mathbf{I}_k}$ ce modèle de mouvement. La méthode d'estimation utilisée est la même que celle utilisée pour le modèle dynamique. Elle est décrite dans [Odobez 95a]. De façon similaire, on note $M_{\mathbf{I}_0 \rightarrow \tilde{\mathbf{I}}_0}$ le modèle de mouvement qui a été utilisé pour la construction de la référence courante. La nouvelle caractérisation de référence est alors construite par un recalage d'images qui permet de compenser le mouvement. Plus précisément, la combinaison des deux modèles $M_{\tilde{\mathbf{I}}_0 \rightarrow \mathbf{I}_k} \circ M_{\mathbf{I}_0 \rightarrow \tilde{\mathbf{I}}_0}$ est appliquée à l'image initiale $\mathbf{I}_0$ afin d'obtenir la nouvelle $\tilde{\mathbf{I}}_0$.

Notons enfin que les mouvements estimés suivent un modèle affine, et sont constitués schématiquement d'une partie translationnelle et d'une partie rotationnelle. Afin d'éviter un phénomène de dérive, la partie translationnelle est ignorée. Cela a pour conséquence de conserver le point initial $\mathbf{x}_0$ au centre de la référence $\tilde{\mathbf{I}}_0$.

Afin de clarifier la procédure complète, une illustration des différentes étapes est donnée

par la figure 7.6. Les images de la ligne du bas représentent les positions estimées et leur ellipse d'incertitude associée à des instants différents. Les images de la ligne du haut représentent l'évolution de la caractérisation de référence qui est mise à jour par recalages successifs.

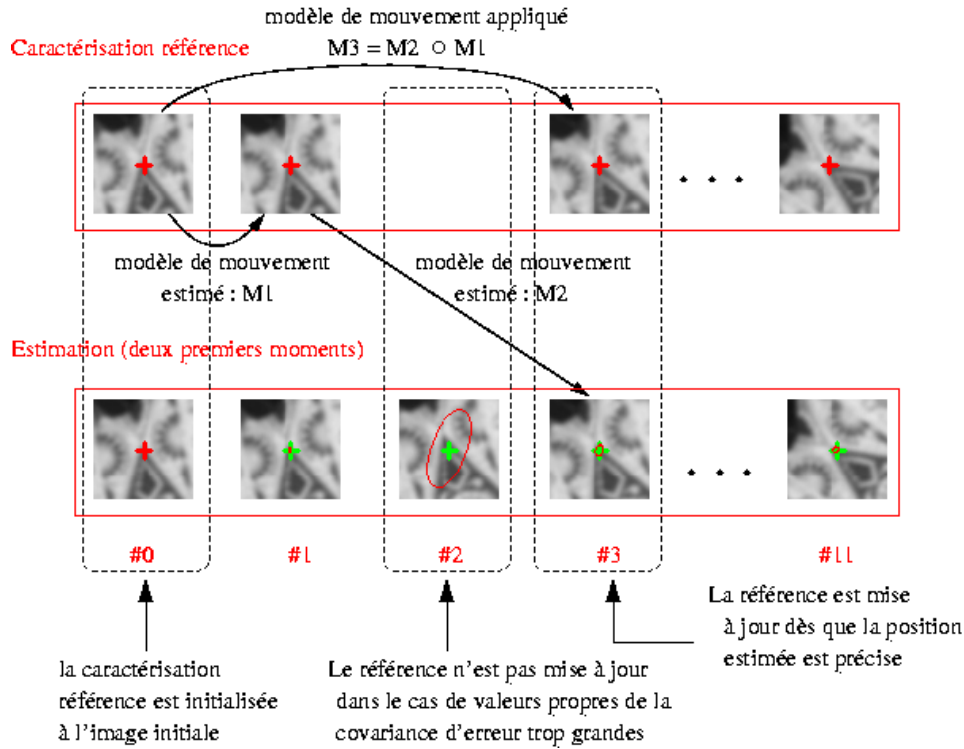


FIG. 7.6: Procédure de mise à jour de la caractérisation de référence

Remarque : La technique proposée permet de recalculer le modèle référence en fonction des évolutions géométriques du voisinage du point suivi. Pour éviter le phénomène de dérive, les valeurs photométriques de cette référence ne sont pas modifiées, ce qui peut poser problème en cas de ruptures d'illumination de la scène. En perspective, il serait intéressant de mettre en place d'autres stratégies pour cette mise à jour. En particulier, deux idées nous semblent à étudier. La première consiste à accumuler différents modèles de référence au cours du suivi pour constituer une base d'exemples décrivant les variations d'apparence du voisinage du point d'intérêt. Cette idée est similaire à celle mise en œuvre dans [Morency 03]. La difficulté réside alors en la définition d'une mesure de similarité entre une fenêtre quelconque et l'ensemble de la base. La deuxième idée repose sur la mise en place d'une méthode de filtrage pour faire évoluer la caractérisation référence [Nguyen 01]. L'approche la plus simple serait d'utiliser un modèle dynamique de type position constante et de définir les observations comme étant les caractérisations de la position estimée à chaque pas de temps. La difficulté de cette approche se situe dans l'estimation pertinente des covariances de bruit pour à la fois faire évoluer rapidement la référence, et ne pas prendre en compte l'estimation lorsque celle-ci est mauvaise.

7.3 Étapes complémentaires et algorithme final

Dans le paragraphe 7.1.2, nous avons vu que le modèle dynamique est différent selon que le point suivi appartienne au mouvement dominant ou non. Avant de décrire les deux structures algorithmiques associées, nous détaillons les deux étapes complémentaires nécessaires à la réalisation d'un algorithme global. Ces étapes sont la sélection des points à suivre, et la règle de décision permettant d'associer un point à une des deux structures algorithmiques.

Critère de sélection d'un point suivi

La première étape de l'algorithme global consiste à vérifier si le point choisi par l'utilisateur a un voisinage suffisamment texturé pour être suivi. Le critère de sélection utilisé repose sur le détecteur KLT décrit au paragraphe 6.1.1 (p. 126). Il est effectué sur la première image. Nous rappelons brièvement que le principe de cette méthode est de calculer les valeurs propres de la matrice d'auto-corrélation. Ces valeurs caractérisent le profil d'intensité du voisinage du point. Lorsque la plus faible des deux valeurs propres est supérieure à un seuil (en pratique fixé à 1), le point est conservé.

Détermination de la structure algorithmique pour le suivi

L'utilisation d'une carte de segmentation basée sur le mouvement fournit une règle de décision simple pour déterminer si le point à suivre appartient au support du mouvement dominant ou à une zone de l'image caractérisée par un modèle de mouvement local. Une telle carte est obtenue par l'utilisation conjointe d'une technique robuste d'estimation de mouvement et d'une méthode de détection du mouvement [Irani 94] [Odobez 95b]. La procédure employée est décrite en détail dans [Odobez 94]. Elle repose en particulier sur une régularisation spatiale à l'aide d'une modélisation par champs de Markov du champ des inconnues (i.e. la carte de détection).

Cette procédure est appliquée entre les deux premières images $\mathbf{I}_0$ et $\mathbf{I}_1$. Elle permet de différencier les points dont l'équation dynamique sera linéaire des points dont l'équation dynamique sera non linéaire. Une illustration de cette procédure sera donnée dans la partie exposant les résultats.

Structure algorithmique associée au suivi d'un point appartenant au support de mouvement dominant

Pour la trajectographie d'un point appartenant au support du mouvement dominant, le système mis en place est composé d'une équation dynamique linéaire à bruit additif (7.6), et d'une équation d'observation linéaire à un bruit additif gaussien (7.11). Nous le rappelons ici :

$$\begin{cases} \mathbf{x}_k = F_{k|\mathbf{I}} \mathbf{x}_{k-1} + \mathbf{b}_{k|\mathbf{I}} + \mathbf{w}_k \\ \mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (7.18)$$

Le bruit de mesure $\mathbf{v}_k$ est supposé être un bruit blanc gaussien de moyenne nulle et de covariance $R_{k|\mathbf{I}}$. Le bruit d'état $\mathbf{w}_k$ est également un bruit blanc à moyenne nulle. Son

deuxième moment est la matrice $Q_{k|\mathbf{I}}$. Bien que le bruit d'état rende compte d'une erreur d'estimation globale sur l'image (et qu'il soit vraisemblablement non gaussien), il est supposé gaussien en raison du manque de connaissance sur ses moments supérieurs. La résolution de ce système passe donc par l'utilisation d'une méthode de filtrage adaptée au cas linéaire gaussien. Puisque les paramètres du modèle dynamique sont estimés sur les images à chaque itération, le filtre linéaire conditionnel (§5.2.1, p. 100) est adapté à la résolution de ce système.

L'ensemble des étapes de la méthode de trajectographie basée sur le FLC et appelée TFLC sont résumées par l'algorithme 16. Insistons sur le fait que cette méthode de suivi permet d'associer des informations globales et locales sur le mouvement du point suivi.

Algorithme 16 TFLC : algorithme de suivi d'un point appartenant au support du mouvement dominant

• **initialisation** : $\hat{\mathbf{x}}_{0|0} = \hat{\mathbf{x}}_0$, $\Sigma_{0|0} = \Sigma_0$
pour $k = 1, 2, \dots$

• **prédiction** :

1. estimation du modèle dynamique sur $\mathbf{I}_{0:k}$: estimation de $F_{k|\mathbf{I}}$, $\mathbf{b}_{k|\mathbf{I}}$ et $Q_{k|\mathbf{I}}$ à l'aide d'une estimation différentielle et robuste du mouvement affine dominant
2. calcul de la position prédite :

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= F_{k|\mathbf{I}} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}_{k|\mathbf{I}} \\ \Sigma_{k|k-1} &= F_{k|\mathbf{I}} \Sigma_{k-1|k-1} F_{k|\mathbf{I}}^t + Q_{k|\mathbf{I}}\end{aligned}$$

• **correction** :

1. estimation du modèle d'observation sur $\mathbf{I}_{0:k}$:
calcul de $\mathbf{z}_k$ par SSD dans *valid_k*^a en utilisant $\tilde{\mathbf{I}}_0$ et estimation de $R_{k|\mathbf{I}}$
2. calcul de la position estimée :

$$\begin{aligned}K_k &= \Sigma_{k|k-1} (\Sigma_{k|k-1} + R_{k|\mathbf{I}})^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + K_k [\mathbf{z}_k - \hat{\mathbf{x}}_{k|k-1}] \\ \Sigma_{k|k} &= (Id - K_k) \Sigma_{k|k-1}\end{aligned}$$

• **mise à jour éventuelle de la caractérisation référence $\tilde{\mathbf{I}}_0$ par recalage affine**

^aLa notation *valid_k* désigne la fenêtre de validation, définie dans le chapitre 3

Structure algorithmique associée au suivi d'un point dont le mouvement ne peut être décrit que localement

Quand le point suivi n'appartient pas au support du mouvement dominant, cela signifie

que son mouvement ne peut être décrit que par une paramétrisation locale. Le problème de filtrage correspond alors à l'estimation de la distribution *a posteriori* associée au système suivant :

$$\begin{cases} \mathbf{x}_k = \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ \mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (7.19)$$

Ce système correspond au modèle dynamique (7.7) et au modèle d'observation (7.11). Rappelons que les bruits de ce système sont des bruits blancs qui suivent les lois suivantes :

$$p(\mathbf{w}_k | \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{w}_k; 0, Q_{k|\mathbf{I}}) \quad (7.20)$$

$$p(\mathbf{v}_k | \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{v}_k; 0, R_{k|\mathbf{I}}) \quad (7.21)$$

Rappelons également que l'équation dynamique modélise l'évolution du point par une fonction non linéaire, ce qui amène à considérer une structure algorithmique basée sur une méthode particulière. De plus, de façon similaire au cas précédent, l'estimation en ligne du modèle dynamique conduit à utiliser un filtre conditionnel. Ces remarques nous amènent à construire une méthode de suivi reposant sur le filtre non linéaire conditionnel (§5.2.2, p. 105). Notons enfin que le système considéré appartient aux modèles conditionnels partiellement linéaires gaussiens étudiés dans le chapitre 5 (§5.3, p. 108). Il autorise donc l'utilisation de la fonction d'importance optimale.

La méthode mise en place pour la reconstruction de la trajectoire d'un point appartenant à une zone de l'image de mouvement secondaire est appelée TFNLC. Elle est résumée par l'algorithme 17.

Algorithme final

En résumé, 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 choisi par l'utilisateur peut être suivi avec succès est réalisé via les valeurs propres de la matrice d'auto-corrélation. Ensuite, une carte de segmentation basée sur le mouvement 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 TFLC. 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 TFNLC.

7.4 Extension à un algorithme robuste aux ambiguïtés : prise en compte de plusieurs mesures

L'estimation de la covariance du bruit d'observation permet d'évaluer la qualité de la mesure fournie par le critère de SSD. Elle permet notamment de limiter l'importance de cette information dans le cas d'ambiguïtés locales autour de cette mesure. En effet, on rappelle que dans une telle situation, la surface de corrélation (SSDs) présente de nombreux pics significatifs. La covariance estimée rend compte de cette incertitude.

La gestion des éventuelles erreurs réparties globalement sur l'image – dues notamment à un fouillis d'arrière plan important – est plus problématique. Malgré l'utilisation de

Algorithme 17 TFNLC : algorithme de trajectographie d'un point dont le mouvement ne peut être décrit que localement

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0 | \mathbf{I}_0)$, et fixer $w_0^{(i)} = 1/N$

pour $k = 1, 2, \dots$

- **estimations sur $\mathbf{I}_{0:k}$:**

1. calcul de $\mathbf{z}_k$ dans $valid_k$ par SSD en utilisant $\tilde{\mathbf{I}}_0$ et estimation de $R_{k|\mathbf{I}}$
2. pour $i = 1 \dots N$, estimation de $\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})$ et $Q_{k|\mathbf{I}}^{(i)}$ à l'aide d'une estimation différentielle et robuste du mouvement affine dominant sur une région autour de $\mathbf{x}_{k-1}^{(i)}$

- **échantillonnage pondéré séquentiel :**

1. échantillonnage :
pour $i = 1 \dots N$, générer $\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k})$ avec

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|\mathbf{I}}, \Sigma_{k|\mathbf{I}})$$

$$\Sigma_{k|\mathbf{I}} = (Q_{k|\mathbf{I}}^{(i)-1} + R_{k|\mathbf{I}}^{-1})^{-1}$$

$$\mathbf{m}_{k|\mathbf{I}} = \Sigma_{k|\mathbf{I}} (Q_{k|\mathbf{I}}^{(i)-1} [\mathbf{x}_{k-1}^{(i)} + P(\mathbf{x}_{k-1}^{(i)}) \boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})] + R_{k|\mathbf{I}}^{-1} \mathbf{z}_k)$$

2. mise à jour des poids d'importance :
pour $i = 1 \dots N$, calculer $w_k^{(i)} = p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \mathbf{I}_{0:k})$, avec

$$p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k; \mathbf{x}_{k-1}^{(i)} + P(\mathbf{x}_{k-1}^{(i)}) \boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I}} + Q_{k|\mathbf{I}}^{(i)})$$

3. normalisation des poids

- **estimation de la position du point**

- **mise à jour éventuelle de la caractérisation référence $\tilde{\mathbf{I}}_0$ par recalage affine**

- **rééchantillonnage si nécessaire**

fenêtres de validation réduisant l'espace de recherche, le suivi peut être perturbé par des zones de l'image très similaires à la caractérisation de référence du point.

Afin de traiter ces situations, nous proposons de considérer plusieurs observations $\{\mathbf{z}_{k,j}\}_{j=1\dots M_k}$ du même point par itérations. Ces observations correspondent aux positions associées aux M_k meilleurs scores de corrélation. Sous les hypothèses que (a) la “vraie” mesure est reliée linéairement à l'état caché avec un bruit additif gaussien, (b) les fausses mesures sont réparties uniformément dans la fenêtre de validation et (c) la probabilité d'avoir uniquement des fausses mesures est nulle, il est possible d'utiliser la vraisemblance suivante :

$$p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k}) \propto \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; \mathbf{x}_k, R_{k|\mathbf{I},j}), \quad (7.22)$$

qui consiste en fait en une généralisation directe du cas où une unique observation est utilisée. Les coefficients $\beta_{k,j}$ se rapportent à la probabilité que la “vraie” mesure au temps k soit la j ème. Dans notre application, ces coefficients sont choisis proportionnels au score de corrélation des observations associées, de telle sorte que $\sum_{j=1}^{M_k} \beta_{k,j} = 1$.

L'hypothèse (c) peut paraître problématique en cas d'occlusions. Cependant, cette déficience potentielle peut être compensée de manière avantageuse par les estimations en ligne des covariances de bruits des mesures $R_{k|\mathbf{I},j}$. Ces matrices sont estimées de manière similaire au cas mono-mesure. Ainsi, lors d'une occlusion, ces matrices seront fixées à $\infty \times Id$. Cela correspond à ne prendre en compte aucune mesure dans l'estimation de la position du point, ce qui équivaut en pratique à poser la probabilité d'avoir aucune observation valable non nulle.

Remarque : Pour le calcul des matrices $\{R_{k|\mathbf{I},j}\}_{j=1\dots M_k}$, il serait intéressant de les évaluer en estimant les paramètres d'un mélange de M_k lois gaussiennes, centrées en les observations $\{\mathbf{z}_{k,j}\}_{j=1\dots M_k}$, sur la région de validation. Un algorithme de type EM pourrait par exemple être employé.

Dans le cas d'un point appartenant à une zone de mouvement secondaire, l'association de cette vraisemblance (7.22) avec la dynamique (7.7) amène au système décrit par les distributions suivantes :

$$\begin{cases} p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k ; \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1})\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}) \\ p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k}) \propto \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; \mathbf{x}_k, R_{k|\mathbf{I},j}) \end{cases} \quad (7.23)$$

Tout comme dans le cas mono-mesure, on remarque que ce système appartient à la famille des système conditionnels partiellement linéaires gaussiens étudiés dans le chapitre 5 (§5.3, p. 108). Il permet à la fois d'échantillonner les particules avec la fonction d'importance optimale, et de définir simplement une région de validation. L'algorithme mis en place, décrit par la méthode 18, est simplement une extension de l'algorithme 17. Par la suite, nous le nommerons algorithme TFNLC multi-mesures.

En ce qui concerne le suivi d'un point appartenant au support du mouvement dominant, l'équation dynamique utilisée est linéaire. L'algorithme mis en place correspond à une

Algorithme 18 Algorithme TFNLC multi-mesures : algorithme de suivi dédié aux situations d'ambiguïtés

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0 | \mathbf{I}_0)$, et fixer $w_0^{(i)} = 1/N$

pour $k = 1, 2, \dots$

- **estimations sur $\mathbf{I}_{0:k}$:**

1. calcul de $\{\mathbf{z}_{k,j}\}_{j=1 \dots M_k}$ par SSD dans $valid_k$ en utilisant $\tilde{\mathbf{I}}_0$ et estimation de $\{R_{k|\mathbf{I},j}\}_{j=1 \dots M_k}$
2. pour $i = 1 \dots N$, estimation de $\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})$ et $Q_{k|\mathbf{I}}^{(i)}$ à l'aide d'une estimation différentielle et robuste du mouvement affine dominant sur une région autour de $\mathbf{x}_{k-1}^{(i)}$

- **échantillonnage pondéré séquentiel :**

1. échantillonnage :

pour $i = 1 \dots N$, générer $\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k})$ avec

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \sum_{j=1}^{M_k} \beta_{k,j} \frac{\alpha_{k,j}}{S_k} \mathcal{N}(\mathbf{x}_k ; \mathbf{m}_{k|\mathbf{I},j}, \Sigma_{k|\mathbf{I},j})$$

$$\Sigma_{k|\mathbf{I},j} = (Q_{k|\mathbf{I}}^{(i)-1} + R_{k|\mathbf{I},j}^{-1})^{-1}$$

$$\mathbf{m}_{k|\mathbf{I},j} = \Sigma_{k|\mathbf{I},j} (Q_{k|\mathbf{I}}^{(i)-1} [\mathbf{x}_{k-1}^{(i)} + P(\mathbf{x}_{k-1}^{(i)}) \boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})] + R_{k|\mathbf{I},j}^{-1} \mathbf{z}_{k,j})$$

Les expressions de $\alpha_{k,j}$ et S_k correspondent respectivement à (5.73) et (5.70).

2. mise à jour des poids d'importance :

pour $i = 1 \dots N$, calculer $w_k^{(i)} = p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \mathbf{I}_{0:k})$, avec

$$p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \mathbf{I}_{0:k}) \propto \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; \mathbf{x}_{k-1}^{(i)} + P(\mathbf{x}_{k-1}^{(i)}) \boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I},j} + Q_{k|\mathbf{I}}^{(i)}).$$

3. normalisation des poids

- **estimation de la position du point**

- **mise à jour éventuelle de la caractérisation référence $\tilde{\mathbf{I}}_0$ par recalage affine**

- **rééchantillonnage si nécessaire**

version très peu modifiée de l'algorithme 18. La seule différence réside dans le choix du support de l'estimation du mouvement qui est, dans ce cas, l'image entière. Cela conduit à considérer une unique matrice $Q_{k|\mathbf{I}}$ estimée pour toutes les particules.

Remarque : Pour les points appartenant au support du mouvement dominant, le modèle permettant de gérer les ambiguïtés est composé d'une équation dynamique linéaire et de la vraisemblance multimodale (7.22). Dans ce cas, une alternative à l'utilisation de l'algorithme 18 réside en l'implémentation du filtre à association de données probabilistes (§3.3 p. 65). Malgré le fait que ce filtre repose sur une approximation forte à chaque itération – qui consiste à résumer la loi de filtrage à l'instant précédent par une unique loi gaussienne – il pourrait amener des résultats satisfaisants, avec un temps de calcul nettement moins important. Cette étude comparative constitue une de nos perspectives à court terme.

7.5 Résultats expérimentaux

Afin de valider nos algorithmes de suivi de points caractéristiques, nous les avons testés sur des séquences réelles les confrontant à des situations diverses. Nous présentons ici les résultats. Ces résultats sont comparés aux trajectoires obtenues par la méthode de Kanade-Lucas-Tomasi KLT – qui a été détaillée dans le chapitre 6 – et par une méthode différentielle robuste MDR. Cette dernière utilise l'estimation de mouvement qui nous permet d'inférer notre modèle dynamique, et ce mouvement n'est pas corrigé par des mesures de corrélation. Enfin, une vérité terrain (correspondant à des trajectoires désignées par un utilisateur) permet d'attester de la qualité des résultats. Le code couleur utilisé est le suivant : le bleu est dédié aux résultats du TFLC, le rouge au TFNLC, le jaune au MDR, le vert au KLT, et le noir à la vérité terrain.

Nous présentons successivement des résultats de l'algorithme TFLC et de l'algorithme TFNLC. Ensuite, quelques tests sur des séquences présentant des déformations géométriques et des occlusions permettent de mettre en évidence la robustesse de nos méthodes à ces situations difficiles. Enfin des illustrations des résultats fournis par l'algorithme TFNLC multi-mesures sont données sur des séquences présentant de fortes ambiguïtés.

7.5.1 Résultats du TFLC, algorithme dédié à un point appartenant au support de mouvement dominant

Nous présentons en premier lieu des résultats expérimentaux du TFLC. Ces résultats concernent trois séquences réelles ayant un mouvement dominant. La première, **Hangars**, comporte 10 images. Cette séquence est très bruitée et présente un mouvement global chaotique. Une comparaison entre les résultats des méthodes TFLC, KLT et TFNLC est présentée sur la figure 7.7. Pour cette séquence, on peut remarquer que le KLT 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, ce qui représente une déficience connue de ce type de méthodes. Au contraire, l'utilisation du TFLC conduit à une bonne reconstruction des trajectoires de l'ensemble des points, alors que les images sont très bruitées et que le mouvement est complexe. Une illustration de la difficulté du mouvement est donnée par la figure 7.8 qui présente les trajectoires estimées des points 5 et 13. Dans

cette figure, les trajectoires TFNLC correspondent à des trajectoires moyennes sur 100 réalisations.

Il est clair que pour le traitement de ce type de séquence, le fait de disposer d'une information globale de mouvement en plus d'une information locale est un point crucial pour la reconstruction des trajectoires. Comme l'information de mouvement dominant n'est pas prise en compte par l'algorithme TFNLC, les résultats de ce dernier ne sont pas satisfaisants. Bien que notre méthode de suivi reposant sur un filtre particulière puisse être utilisée dans toutes les situations, la méthode linéaire donne de meilleurs résultats pour un coût calculatoire inférieur dans le cas de séquences avec un mouvement dominant.

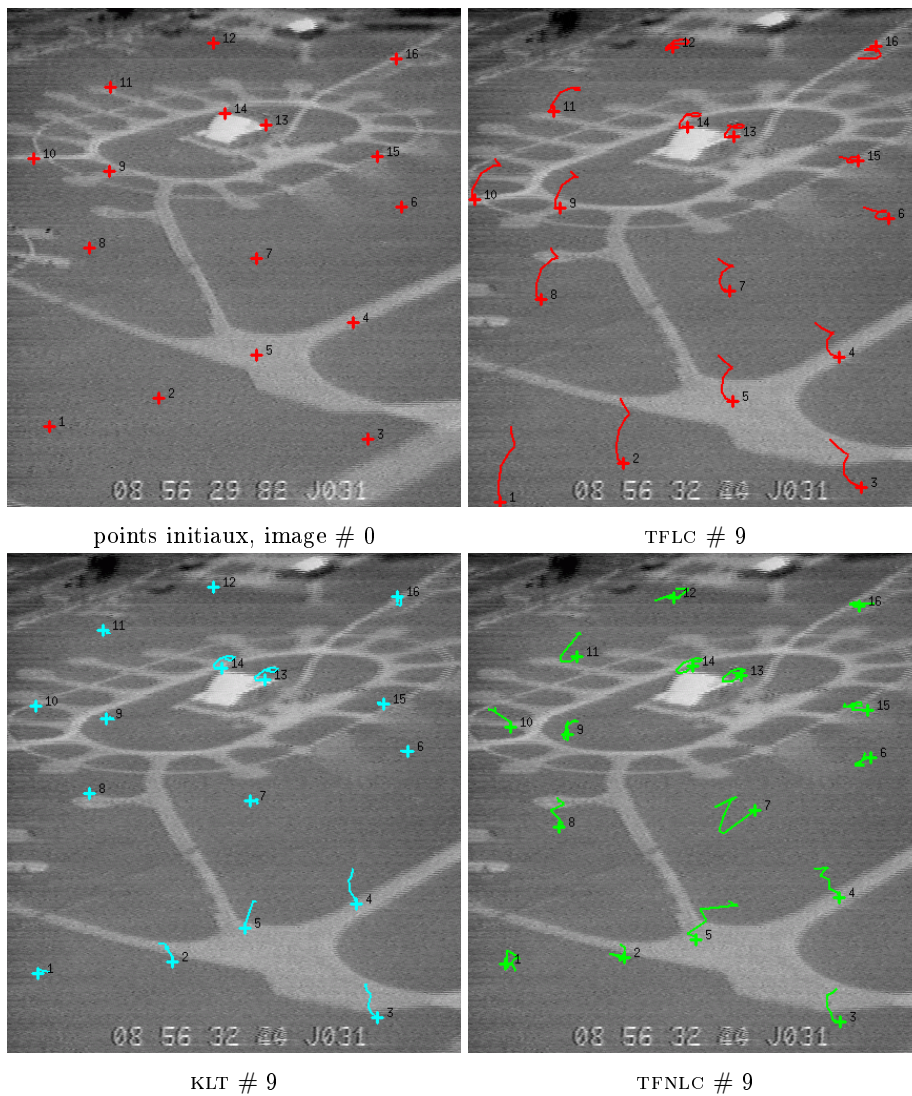


FIG. 7.7: Séquence Hangars – Trajectoires complètes estimées par les algorithmes TFLC, KLT et TFNLC

La deuxième séquence, **Corridor**, comporte 7 images. À cause des discontinuités de profondeur et des mouvements de large amplitude, elle constitue un cas limite pour une

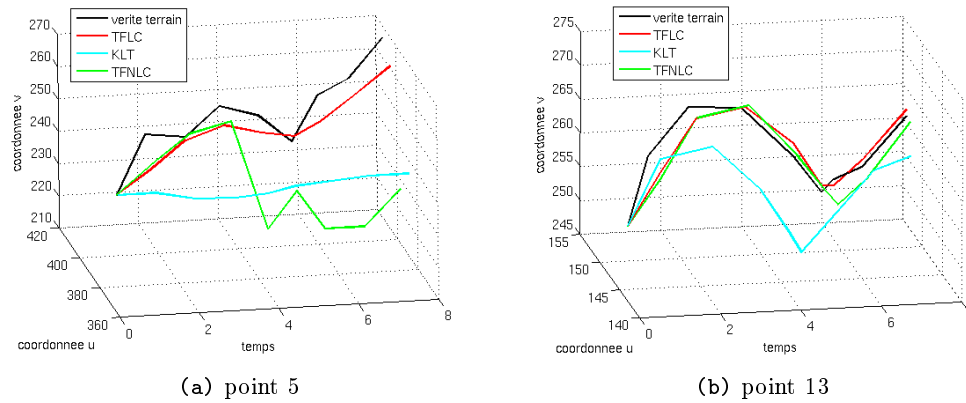


FIG. 7.8: Séquence Hangars – Comparaison entre les trajectoires estimées par les différents algorithmes et la vérité terrain. Les trajectoires TFNLC correspondent à des trajectoires moyennes sur 100 réalisations

approximation par un modèle global de mouvement affine. Les trajectoires complètes fournies par les algorithmes TFLC, KLT et MDR sont présentées dans la figure 7.9. Dans une telle séquence, nous pouvons remarquer que le KLT 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 TFLC. Nous pensons que d’une part, le modèle dynamique linéaire global utilisé fournit une assez bonne prédiction des positions (même si le modèle affine n’est pas adapté pour cette séquence), et d’autre part, les mesures de corrélation permettent de corriger une éventuelle déficience de ce modèle de mouvement global. Cela peut être vérifié en observant les trajectoires obtenues par la méthode MDR. En effet, ce résultat atteste de la non précision du modèle affine, qui constitue un modèle de mouvement grossier pour certains points (i.e. points 1,2,9).

Ces remarques sont également illustrées par la figure 7.10 qui permet de comparer les trajectoires des points 1 et 2 estimées par les différents algorithmes avec une vérité terrain. Dans ces graphiques, les trajectoires TFNLC sont des trajectoires moyennes sur 100 réalisations. On peut y observer une déviation significative entre la vérité terrain et la trajectoire MDR. Les mauvais résultats du KLT sont également visibles sur le graphique 7.10(b), où la trajectoire du KLT n’est pas visible car elle a été interrompue dès la seconde image (cet algorithme arrête le suivi quand la trajectoire est considérée mauvaise). De façon similaire à la séquence Hangars, le TFNLC donne de meilleurs résultats que le KLT, mais est moins adapté que le TFLC pour une complexité algorithmique plus importante.

La troisième séquence démontre également la robustesse du TFLC à un modèle de mouvement global peu précis. La séquence **Bâtiment** comporte 53 images et représente deux piétons marchant devant un bâtiment. Elle a été acquise par un utilisateur amateur, et le mouvement caméra est donc chaotique. Cette séquence constitue un cas difficile pour l’estimation d’un mouvement dominant affine puisque la structure 3D de la scène est composée de deux plans orthogonaux. Dans ce cas, le mouvement estimé est valide pour un des deux plans. Cependant, les résultats obtenus par le TFLC sont satisfaisants, comme cela est montré par la figure 7.11. Ces résultats permettent aussi de démontrer la robustesse

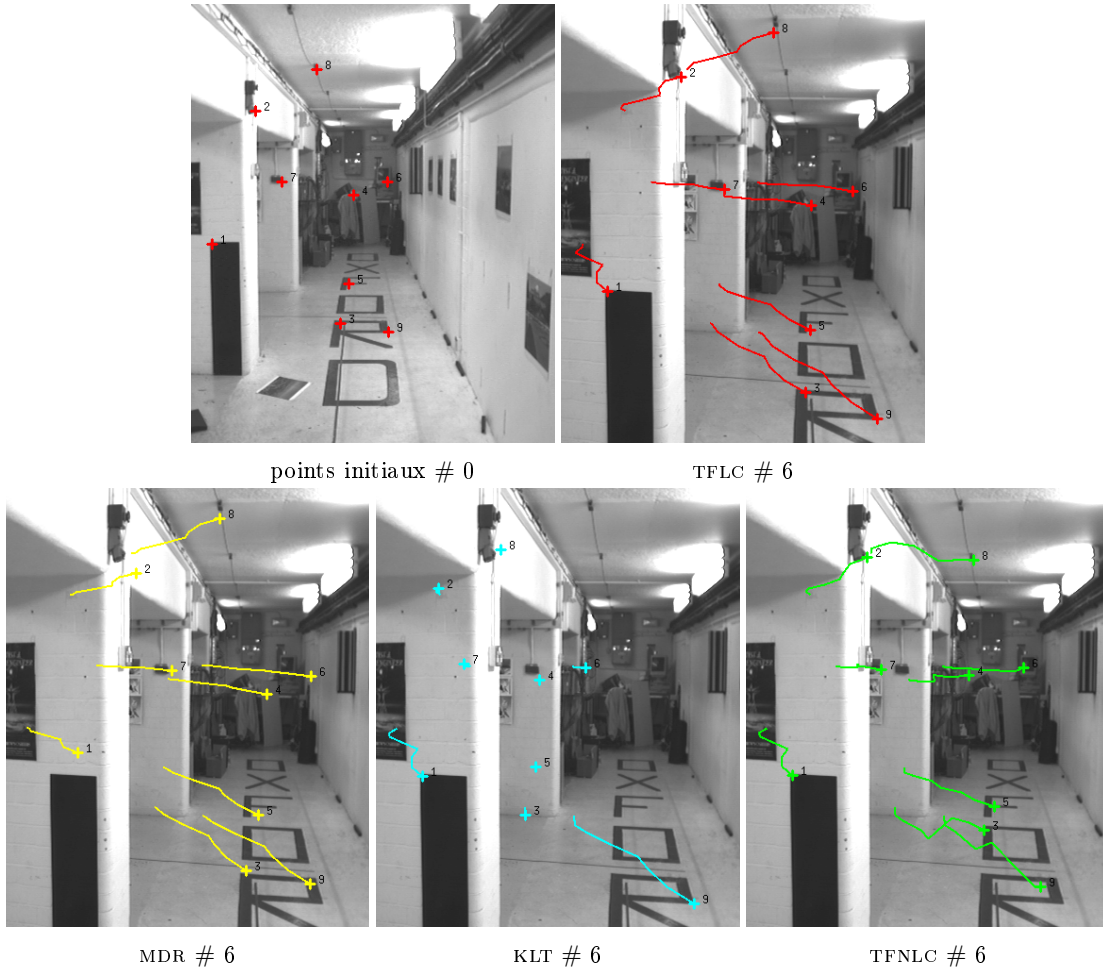


FIG. 7.9: Séquence Corridor – Trajectoires complètes estimées par les algorithmes TFLC, MDR, KLT et TFNLC

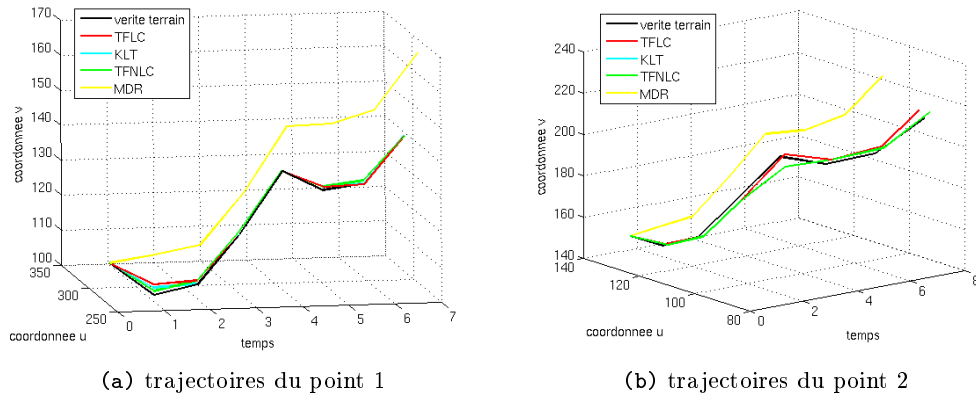


FIG. 7.10: Séquence Corridor – Comparaison entre les trajectoires estimées par les différents algorithmes et la vérité terrain. Les trajectoires TFNLC correspondent à des trajectoires moyennes sur 100 réalisations

de notre approche en présence d'occlusions. En effet, il apparaît sur cette séquence que la position des points est maintenue, même après leur occultation par un piéton (voir par exemple les points 1, 3, 4, 5, 9, 14). D'autres résultats sur des séquences montrant des occlusions seront donnés dans un paragraphe suivant.

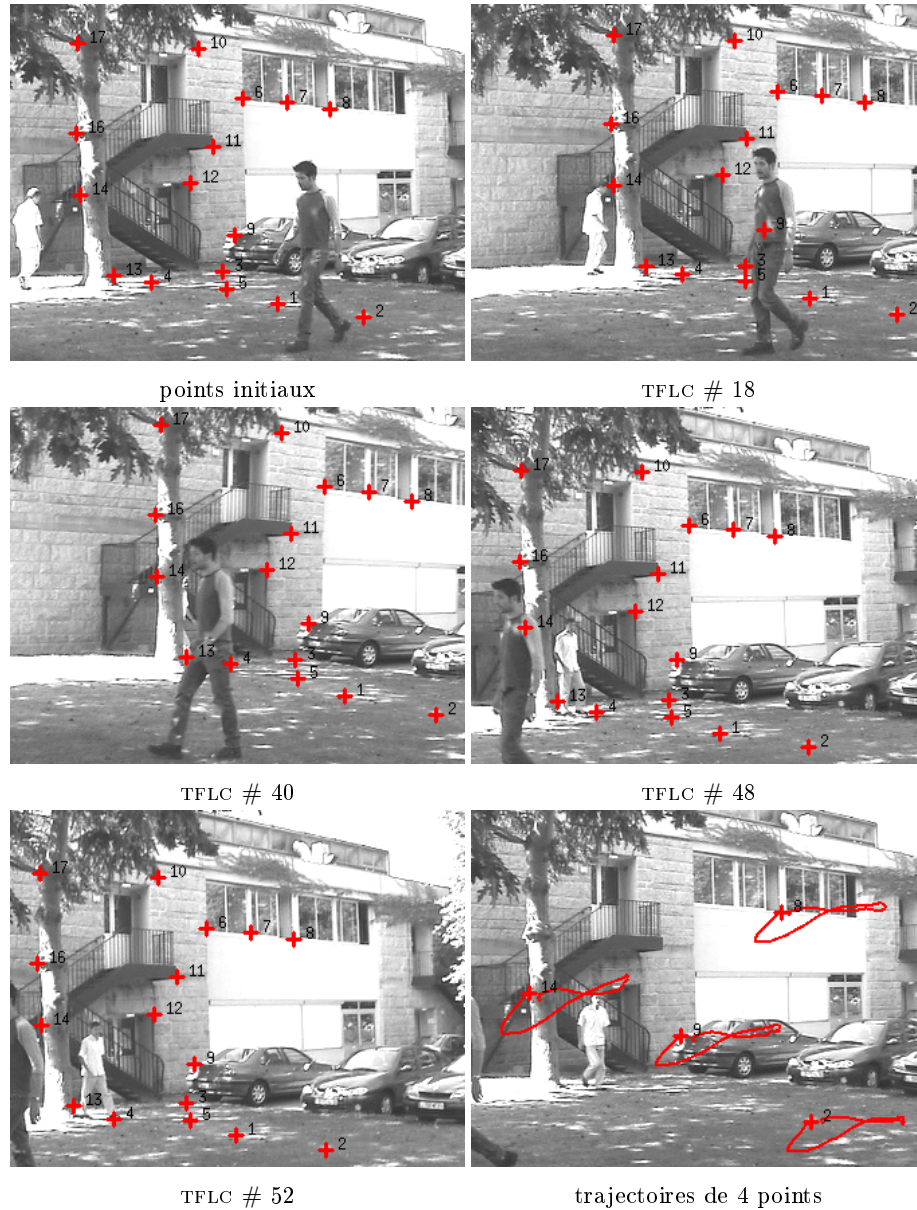


FIG. 7.11: Séquence Bâtiment – Résultat du suivi par l'algorithme TFLC

7.5.2 Résultats du TFNLC, algorithme dédié à un point dont le mouvement ne peut être décrit que localement

Un premier résultat de l'algorithme TFNLC est présenté sur la séquence **Mobile** qui est composée de 40 images. Cette séquence représente le mouvement de deux balles fixées sur

un cercle rigide en rotation, sur un arrière-plan encombré. 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 . Les points initiaux sont donnés dans la figure 7.12. Cette figure présente également une illustration de la carte de segmentation au sens du mouvement utilisée pour séparer les points suivant le mouvement dominant des autres. Comparé aux résultats apportés par le KLT et le MDR (cf. figure 7.14), le TFNLC réussit à discriminer les balles 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 (cf. figure 7.12(c)).

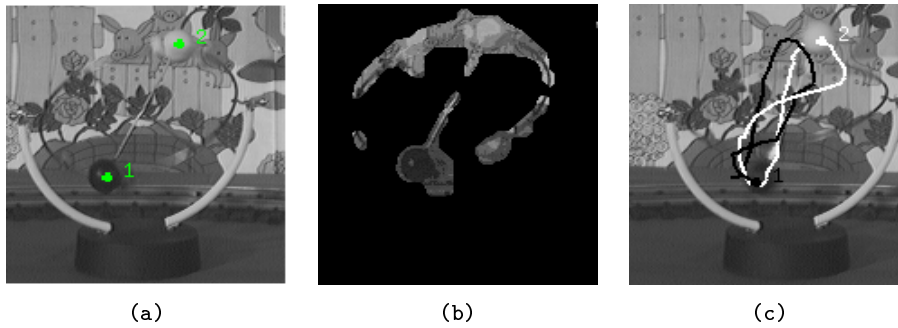


FIG. 7.12: Séquence Mobile – (a) points initiaux # 0 (b) carte de segmentation au sens du mouvement (c) trajectoires du TFNLC

Les trajectoires des balles blanches et noires, estimées par les différentes méthodes, sont comparées à une vérité terrain dans la figure 7.13. Le résultat associé au TFNLC rend compte de la moyenne sur les réalisations valides (sur 100 réalisations, 2 échecs ont été constatés pour la balle blanche et 3 pour la noire).

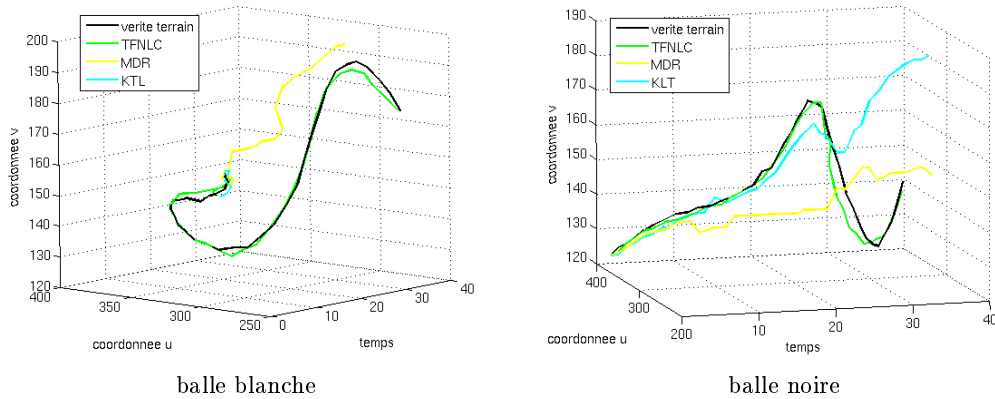
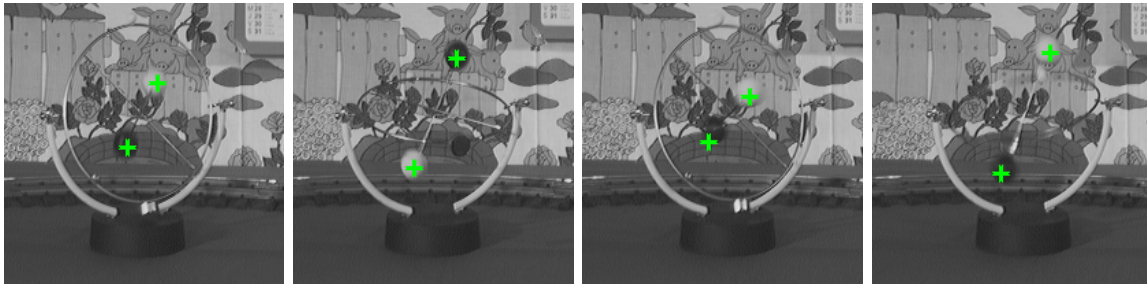
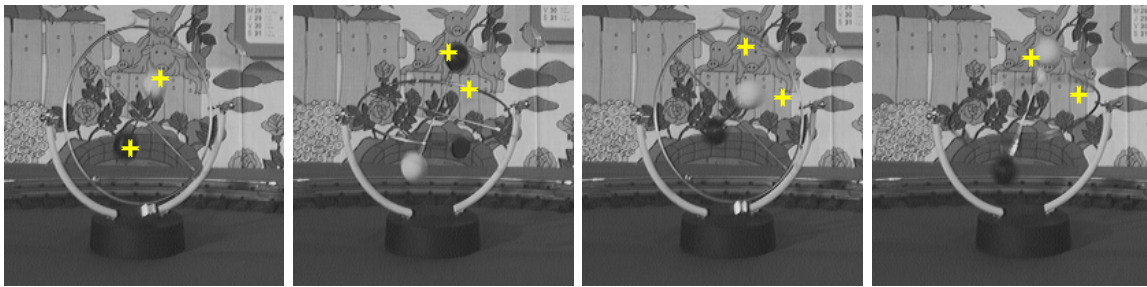


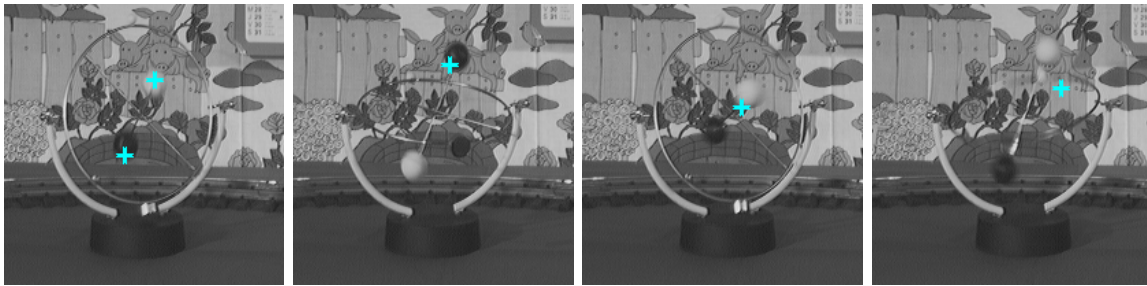
FIG. 7.13: Séquence Mobile – Comparaison entre les trajectoires estimées par les différents algorithmes et la vérité terrain. La trajectoire TFNLC correspond à une trajectoire moyenne sur les réalisations correctes (2 échecs sur 100 pour la balle blanche, 3 sur 100 pour la noire)



TFNLC # 7, # 20, # 30, # 39



MDR # 7, # 20, # 30, # 39



KLT # 7, # 20, # 30, # 39

FIG. 7.14: Séquence Mobile – Résultats de suivi des algorithmes TFNLC, MRD et KLT

La figure 7.15 montre les résultats fournis par le TFNLC sur la séquence **Examen**. Cette série de 60 images est une séquence médicale d'une angiographie. Les mouvements des vaisseaux sont caractérisés par des contractions et des dilatations. Ces résultats sont comparés aux positions estimées par le KLT. Notre algorithme réussit à reconstruire les trajectoires des points alors que le KLT en perd la moitié. Les mouvements des vaisseaux sont relativement bien captés par l'estimation locale de mouvement.

Pour retrouver une prédiction du même ordre de précision, sans utiliser une équation dynamique estimée en ligne, il aurait été nécessaire de définir au moins deux modèles d'évolution *a priori*, et donc de construire une méthode de suivi spécifique à ce type d'images. Au contraire, notre méthode permet de s'adapter à de nombreux types de mouvement et reste ainsi générale. La construction d'un nouveau modèle d'évolution à chaque itération permet de détecter des changements brusques de direction.

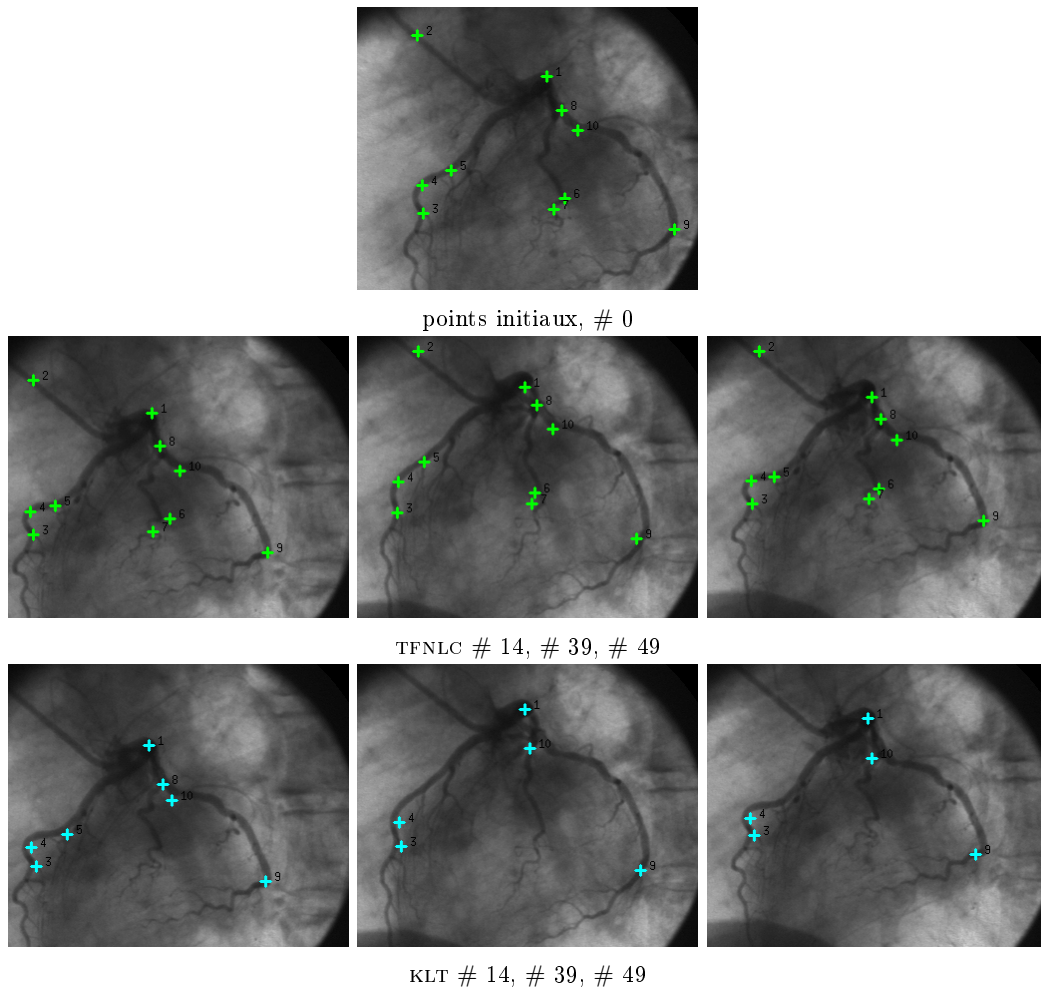


FIG. 7.15: Séquence Examen – Résultats du suivi par les algorithmes TFNLC et KLT

Enfin, le TFNLC a été testé sur une séquence météorologique de 13 images, **Fluide**, montrant l'évolution d'un nuage à basse pression. Bien que les mesures de SSD ne soient pas pertinentes pour ce genre de données, le TFNLC permet de retrouver le mouvement de vorticit , comme cela est expos  dans la figure 7.16. Il est   remarquer que les r sultats sont satisfaisants, m me si les caract ristiques de l'imagerie fluide (en particulier les changements d'illumination engendr s par des ph nom nes physiques) ne sont pas prises en compte par la m thode propos e. La pr cision des r sultats est due   une bonne estimation du mouvement local autour de chaque point, et   la d tection automatique de la mauvaise qualit  des mesures de corr lation.

Pour ces trois s quences, insistons sur le fait qu'il aurait  t  difficile de se reposer sur un  quation dynamique lin aire classique (du type mod le auto-r gressif).

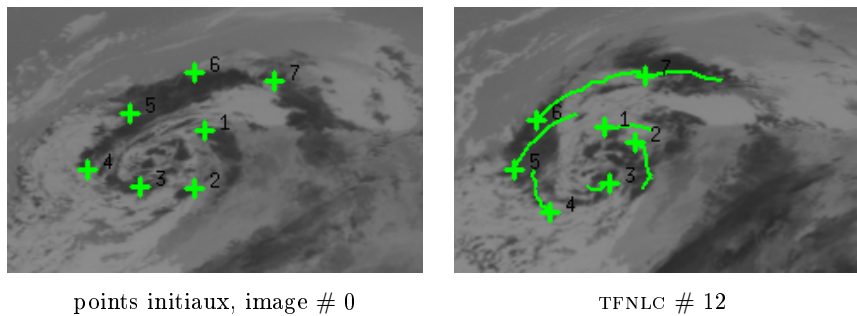


FIG. 7.16: Séquence Fluide – Trajectoires estimées par le TFNLC

7.5.3 Robustesse aux déformations géométriques et aux occlusions

Dans le but de valider nos approches face à des situations difficiles telles que des déformations géométriques et des occlusions, nous présentons dans ce paragraphe des résultats sur trois nouvelles séquences réelles. La première séquence utilisée, **Minitel**, est construite à partir de 15 images qui présentent un mouvement de rotation. Ce mouvement résulte d'une rotation de la caméra sur au moins deux axes. La figure 7.18 montre les positions estimées par l'algorithme TFLC à différents instants. La figure 7.17 indique l'évolution de la caractérisation de référence du point 1. Sans la procédure de mise à jour de la référence décrite § 7.2.2, les grandes déformations de rotation ne permettent pas l'obtention de bonnes mesures de mise en correspondance. Cela est accentué par les nombreuses possibilités d'erreurs dues aux motifs répétitifs de la nappe.

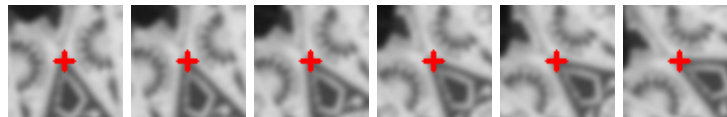


FIG. 7.17: Séquence Minitel – Évolution de la caractérisation de référence du point 1

La remarque concernant l'importance de la procédure de mise à jour de la référence est également valable pour le traitement de la séquence **Van**. Cette séquence de 75 images est une séquence routière montrant l'arrière d'une camionnette. Sur la figure 7.19, on peut noter les changements d'échelle importants. Cette figure présente la carte de segmentation au sens du mouvement et les résultats du TFLC. Les trajectoires complètes estimées par notre algorithme montrent que celui-ci permet de reconstruire des trajectoires subissant des changements abrupts de direction.

La dernière séquence, **Jardin**, comporte 27 images. Elle présente un mouvement global translationnel, avec un mouvement de parallaxe sur l'arbre en premier-plan. La figure 7.20 expose les résultats de TFNLC pour les points de l'arbre (points 2,3,4) et ceux du TFLC pour les autres (conformément à la carte de segmentation donnée par la figure 7.21(a)). Il est à noter que nous avons choisi de suivre un point de l'arrière plan (point 1) avec le TFNLC pour démontrer sa robustesse aux occlusions. La trajectoire moyenne de ce point sur les réalisations correctes du TFNLC est donnée par la figure 7.21(b). L'occultation du point par

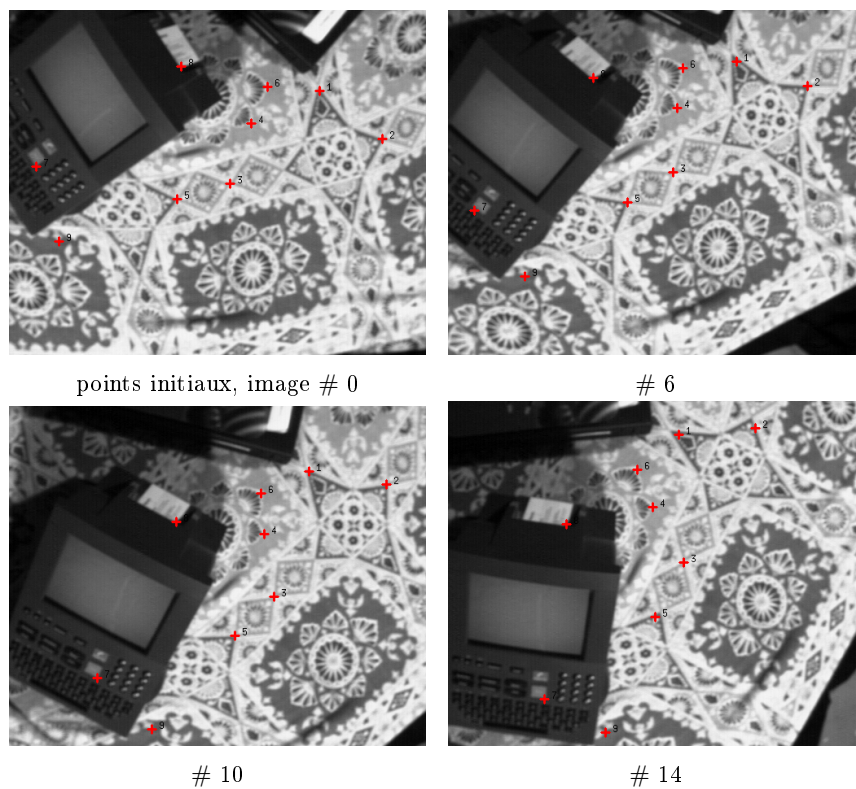


FIG. 7.18: Séquence Minitel – Résultat du suivi par l’algorithme TFLC

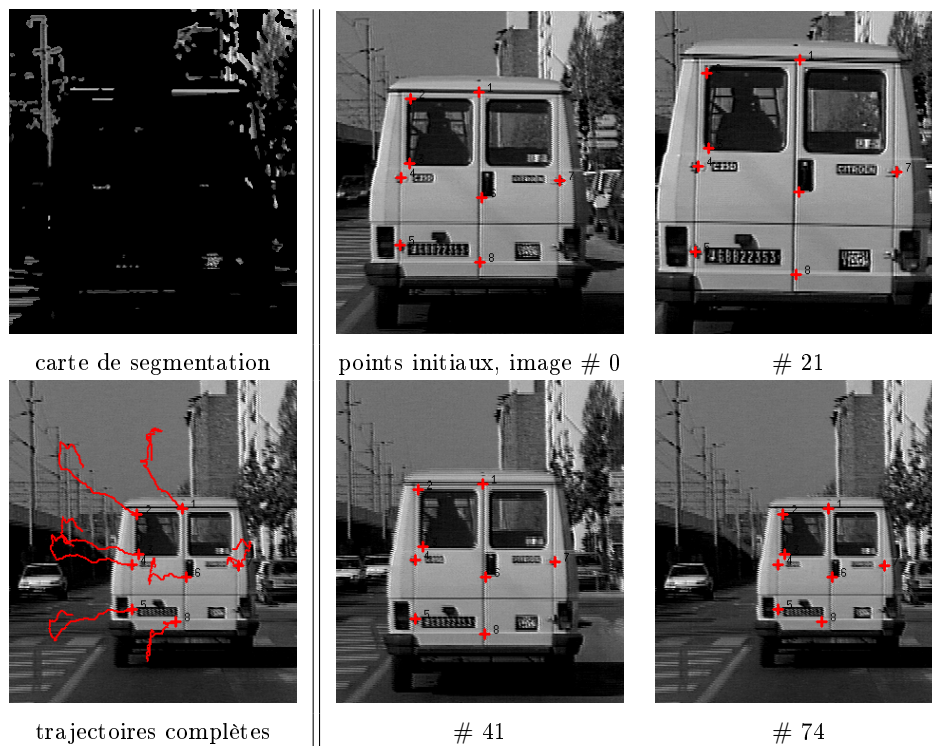


FIG. 7.19: Séquence Van – Résultat du suivi par l’algorithme TFLC

l'arbre est longue (12 images), et sur 100 essais, nous avons constaté 15 échecs. En suivant les trajectoires des points qui passent derrière l'arbre, nous pouvons remarquer que, sans spécifier de schéma particulier d'occlusion, 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, permet à nos approches d'être résistantes aux occlusions.

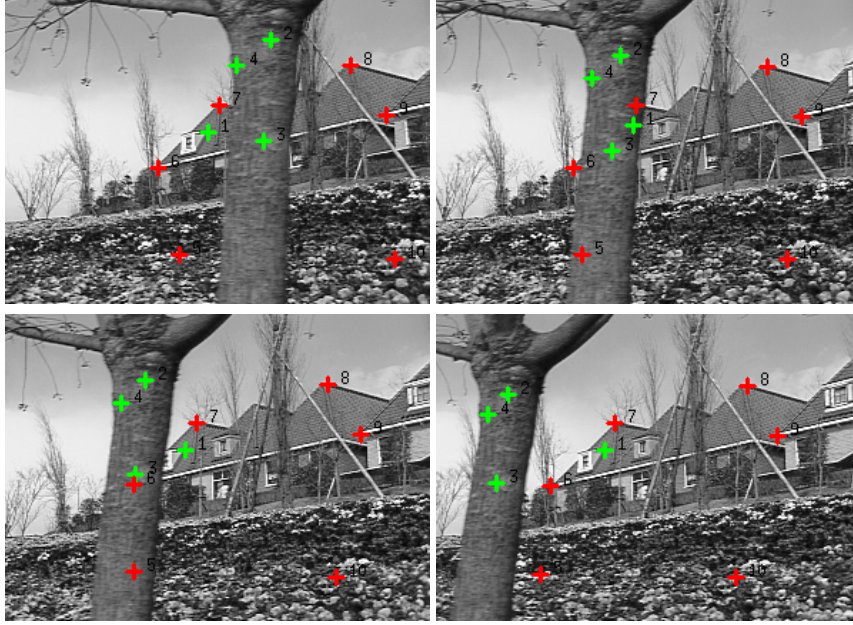


FIG. 7.20: Séquence Jardin– Résultat des algorithmes TFLC (en rouge) et TFNLC (en vert) sur les images #0,10,16,26

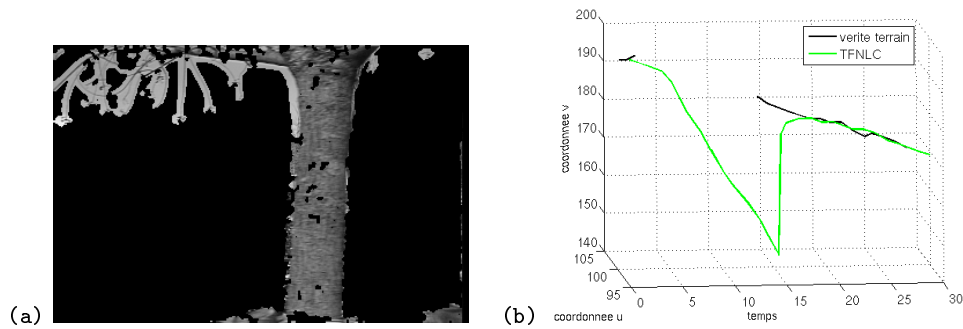


FIG. 7.21: Séquence Jardin – (a) Zone de l'image suivant un mouvement secondaire (b) Comparaison entre la trajectoire estimée du point 1 par le TFNLC et la vérité terrain. La trajectoire TFNLC correspond à une trajectoire moyenne sur les réalisations correctes (15 échecs sur 100 essais).

De telles situations d'occultation permettent de démontrer l'intérêt de l'utilisation de la fonction d'importance optimale. Pour cela, nous avons comparé notre algorithme TFNLC, à un algorithme du type CONDENSATION pour le suivi du point 1. Les deux algorithmes testés

reposent sur le même modèle, mais diffèrent par leur procédure de diffusion des particules. Le TFNLC utilise la fonction d'importance optimale qui permet de prendre en compte à chaque itération la nouvelle observation, alors que la méthode CONDENSATION utilise le modèle dynamique pour explorer l'espace d'état. La figure 7.22 présente les résultats de cette comparaison. Une bonne diffusion des particules, prenant en compte le modèle dynamique et la nouvelle observation, permet de retrouver la position du point après son occultation. Au contraire, un algorithme de type CONDENSATION n'est pas efficace dans ce type de situation, où le recouvrement entre les supports de la densité de prédiction et la vraisemblance est faible. Enfin, cette figure permet également d'illustrer l'intérêt de l'utilisation d'une fenêtre de validation des mesures. Celle-ci réduit considérablement l'espace de recherche du pic de corrélation lorsque la confiance sur la qualité de l'estimation est élevée.



TFNLC, avec la fonction d'importance optimale (# 2,6,16,30)



algorithme de type CONDENSATION, sans la fonction d'importance optimale (# 2,6,16,30)

FIG. 7.22: Séquence Jardin – Illustration de l'importance de la fonction optimale et de la région de validation – L'estimée correspond à la croix verte, l'observation à la croix bleue et la région de validation à l'ellipse bleue.

7.5.4 Validation de l'algorithme dédié au traitement des ambiguïtés

Les derniers résultats présentés permettent de valider l'algorithme TFNLC multi-mesures, construit comme une extension du TFNLC dédiée à la gestion des ambiguïtés. Pour cela, nous utilisons trois séquences réelles où les points suivis correspondent à des motifs similaires que l'on peut retrouver dans différentes parties de l'image (le code couleur utilisé précédemment n'est pas valable dans ce paragraphe).

La première séquence, **Main**, contient 20 images. Les points suivis correspondent aux intersections entre les doigts, dont les caractérisations en termes d'images sont très similaires. Afin de montrer l'intérêt de considérer plusieurs observations par points, nous avons

testé l'algorithme TFNLC (qui n'utilise qu'une mesure par point), et son extension au cas multi-mesures. Les figures présentant les résultats sont numérotées respectivement 7.23 et 7.25. On peut remarquer sur la figure 7.23 que le résultat du TFNLC n'est pas satisfaisant. En effet, le point à l'intersection entre l'index et le majeur est mal suivi à cause d'une confusion avec l'intersection entre le majeur et l'annulaire (cf. images 12 et 16). Le suivi de ce point est étudié plus en détail sur la figure 7.24. Sur cette figure sont représentées les particules (en jaune), ainsi que la position estimée (en rouge) et la mesure. L'erreur de suivi est due à une mesure erronée (cf. images 10,12,15) qui correspond cependant à un pic de corrélation important, ce qui entraîne une faible matrice de covariance de bruit de mesure, et une diffusion des particules autour de cette mauvaise observation.

Au contraire, les résultats obtenus en considérant trois observations par point sont corrects (cf. figure 7.25) : les trajectoires ne sont pas perturbées par les ambiguïtés de l'image. Comme pour la figure 7.24, la figure 7.26 permet d'étudier plus en détail le suivi du point situé entre l'index et le majeur. Sur cette figure, sont représentées les particules (en jaune), l'estimée (en rouge), l'observation correspondant au pic de corrélation le plus haut (en bleu) et les deux autres mesures (en vert). En prêtant attention aux images 10,12,15, on remarque que l'observation principale se situe à une position erronée et qu'une observation secondaire correspond à la "vraie" mesure. Cependant, on rappelle que dans ce cas, la fonction d'importance optimale est un mélange de gaussiennes, dont les coefficients sont calculés en fonction des poids des observations, et de la distance entre ces observations et une position prédite. Ce calcul permet de donner plus d'importance au pic secondaire, et donc de conserver la bonne position.

Une seconde comparaison entre le TFNLC simple et le TFNLC multi-mesures est effectuée sur la séquence **Bastille**. Cette série de 31 images montre le trafic routier sur le rond-point de la Bastille. Les difficultés de cette séquence sont les similarités entre les différentes voitures, ainsi que les changements d'illumination lorsque celles-ci passent dans la zone d'ombre à droite des images. Les résultats du TFNLC avec une observation par point sont présentés dans la figure 7.27. Les trajectoires reconstruites ne sont pas exactes. En particulier, les trajectoires des deux voitures du bas se rejoignent sur un même véhicule, et la zone d'ombre met en échec le suivi d'une autre voiture.

Le résultat d'un suivi réalisé par le TFNLC multi-mesures, avec trois observations par point est donné par la figure 7.28. L'ensemble des trajectoires sont reconstruites avec succès (à part un décrochage temporaire du point situé sur le camion), ce qui permet de valider l'utilisation de l'algorithme dédié aux traitements des ambiguïtés dans de telles situations.

Enfin, la dernière séquence, **Foot**, est un extrait d'un match de football retransmis à la télévision. Le résultat du suivi des joueurs réalisé en prenant en compte trois observations par point est illustré dans la figure 7.29. Réaliser un suivi correct sur ce type de séquence est complexe du fait du zoom sur le terrain, des fortes ressemblances entre les joueurs et de leur mouvement difficile à capter (les zones mobiles sont de taille très faible). Pourtant, le TFNLC multi-mesures donne des résultats satisfaisants. Le fait de considérer plusieurs observations permet de retrouver les joueurs lorsque les trajectoires de ceux-ci ont décroché.

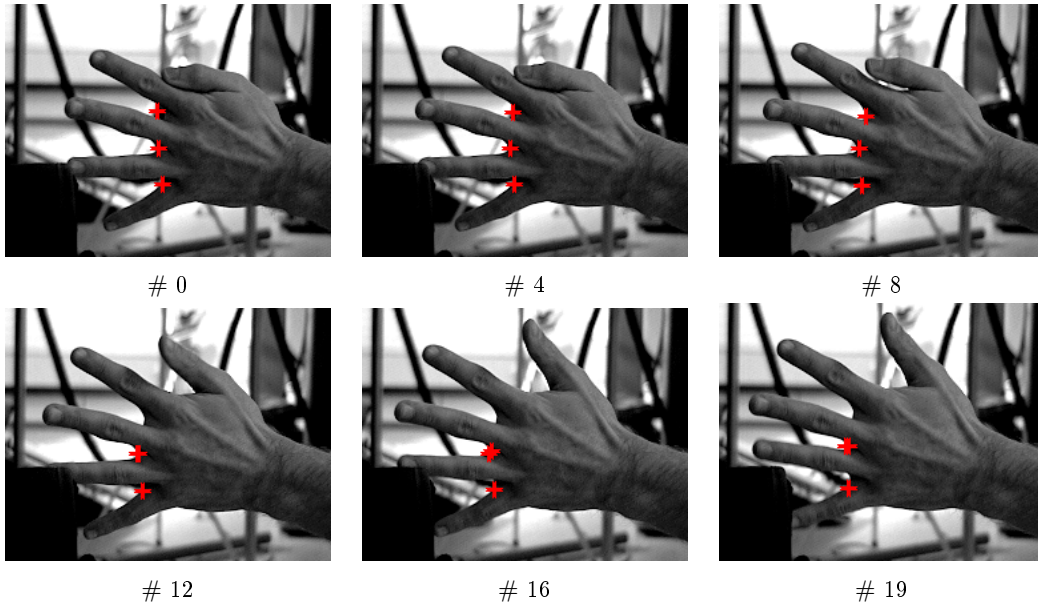


FIG. 7.23: Séquence Main – Résultat du TFNLC, avec une observation par point suivi

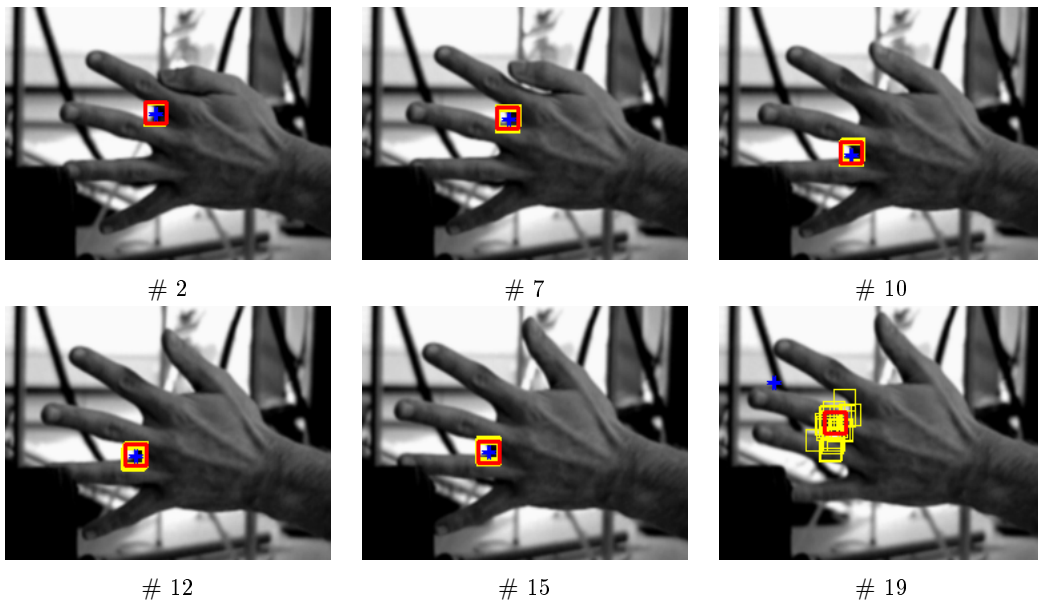


FIG. 7.24: Séquence Main – Résultat du suivi du point situé entre l'index et le majeur fourni par le TFNLC, en considérant une unique observation. Cette observation est représentée en bleu, la position estimée en rouge et les particules en jaune (100 particules sont utilisées).

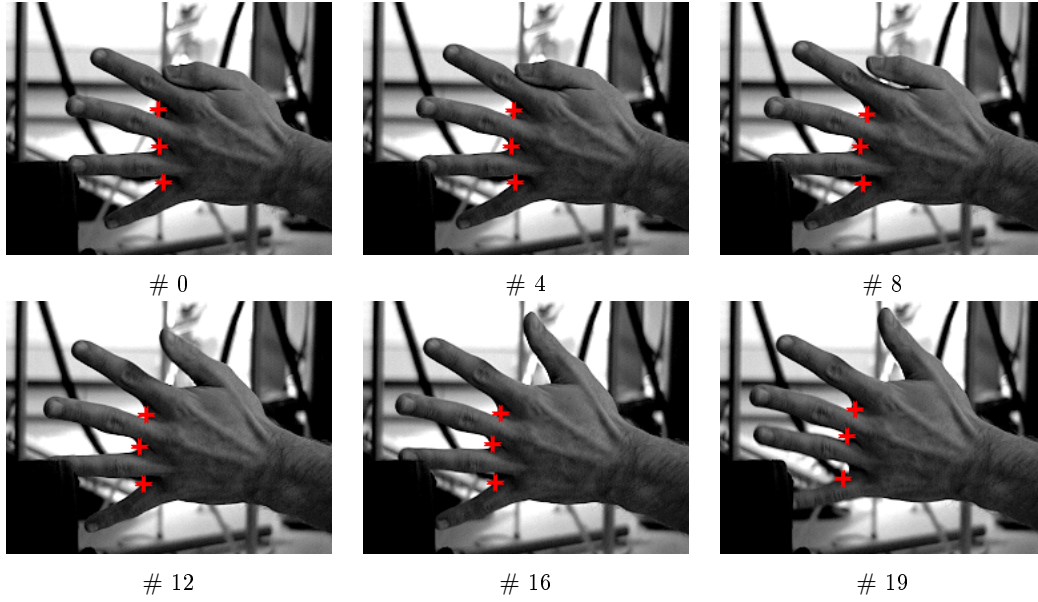


FIG. 7.25: Séquence Main – Résultat de l'algorithme TFNLC multi-mesures, avec trois observations par point suivi

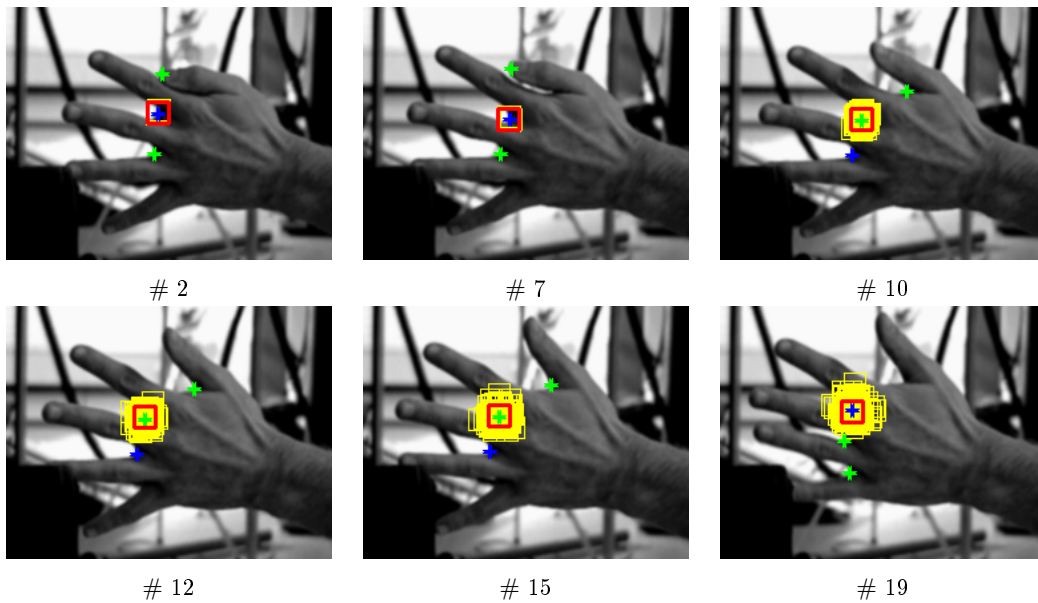


FIG. 7.26: Séquence Main – Résultat du suivi du point situé entre l'index et le majeur fourni par le TFNLC multi-mesures, en considérant trois observations. La mesure principale (qui correspond au plus haut pic de corrélation) est représentée en bleu, les deux mesures secondaires en vert, la position estimée en rouge et les particules en jaune (100 particules sont utilisées).

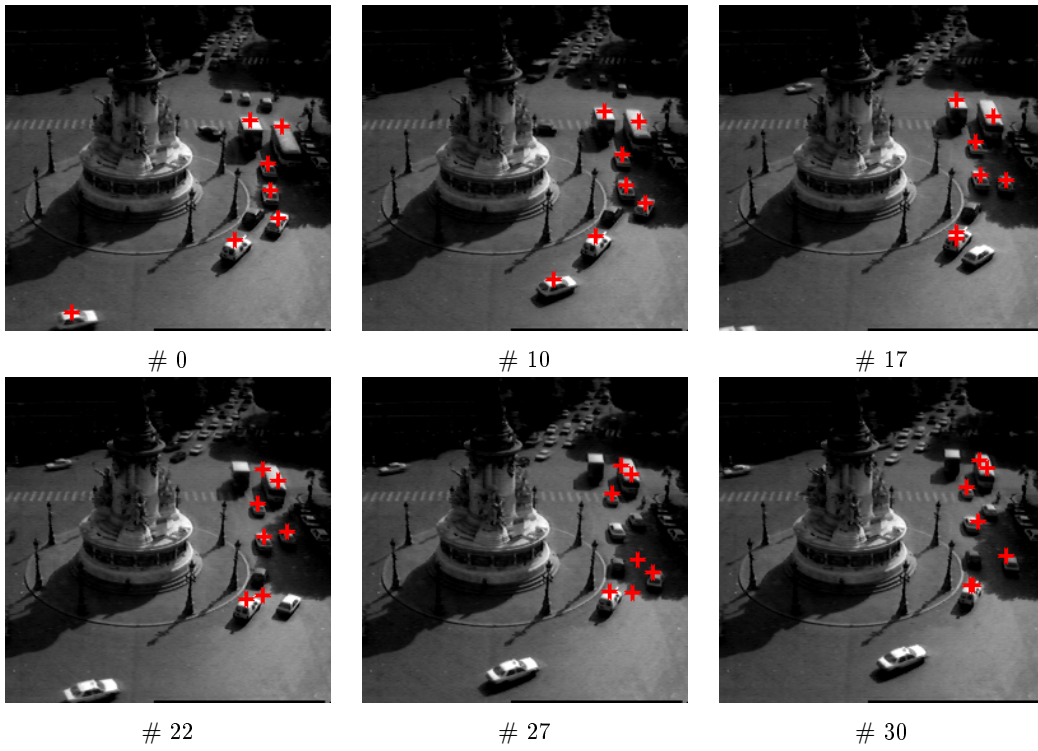


FIG. 7.27: Séquence Bastille – TFNLC, avec une observation par point suivi

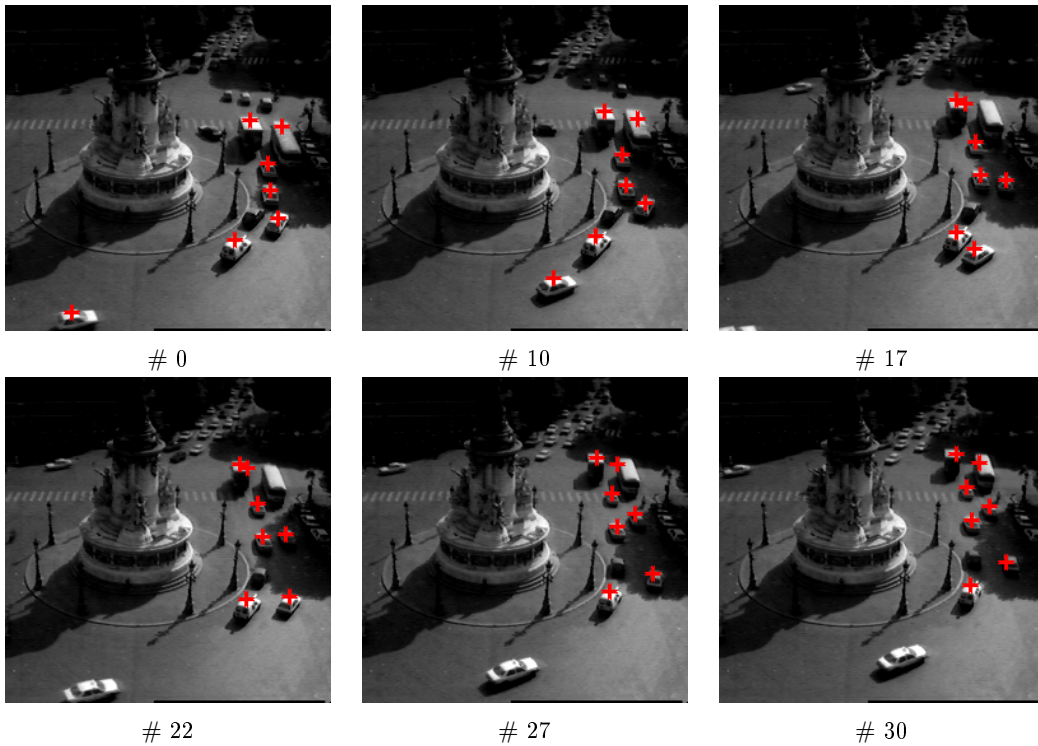


FIG. 7.28: Séquence Bastille – TFNLC multi-mesures, avec trois observations par point suivi

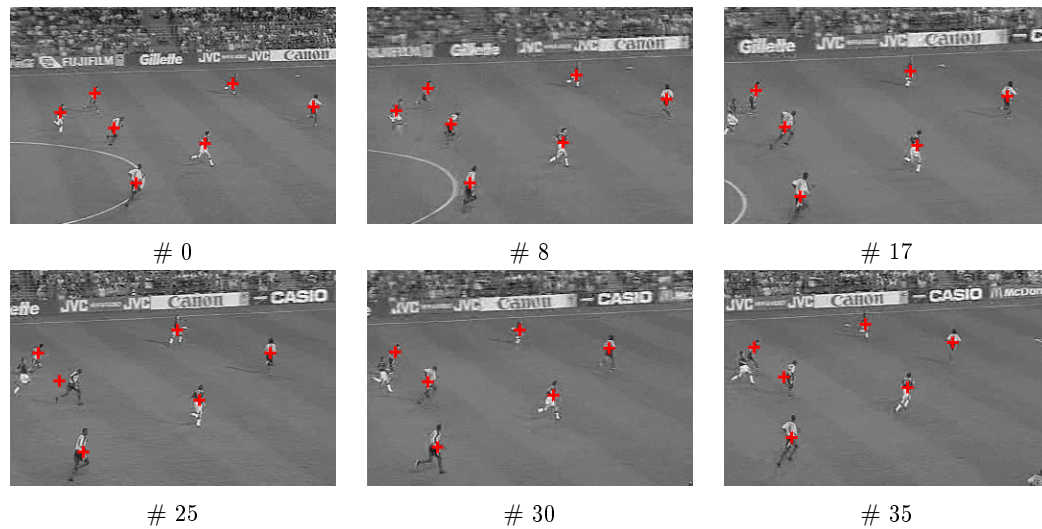


FIG. 7.29: Séquence Foot – TFLC, avec trois observations par point suivi

7.6 Application au suivi de visages

L'algorithme TFLC a été utilisé pour reconstruire la trajectoire des yeux, dans une application de suivi de visages. Cette application a été développée dans le cadre de projets multimédia regroupant des partenaires industriels et académiques. Elle a été initiée dans le cadre du projet prospectif MediaWorks¹ [Boucher 03], et continue dans le cadre du projet Feria².

Feria (Framework pour l'Expérimentation et la Réalisation Industrielle d'Applications multimédias) est un projet piloté par la SSII Communication et Systèmes (CS) et l'Institut national de l'audiovisuel INA. Son objectif est la mise en place de solutions logicielles facilitant la représentation et la manipulation des données multimédia. Ce projet est motivé par le besoin croissant d'outils permettant de manipuler les contenus des documents audiovisuels. Du point de vue technique, la manipulation des contenus nécessite la capacité de représenter et de manipuler informatiquement des données les décrivant. La classe des applications ayant besoin de manipuler de telles données est très vaste : on y trouve par exemple la gestion de fonds audiovisuels, la constitution de collections de documents, la production de vidéo enrichie, les bancs de montage virtuel, etc.

Dans ce cadre, la reconnaissance de personnages dans une vidéo est un des éléments nécessaires à la description de séquences. Afin de faciliter cette étape en évitant sa mise en œuvre à chaque image, il est intéressant d'effectuer un suivi automatique de l'ensemble des visages de la série d'images. La reconnaissance pourra alors être réalisée pour chaque trajectoire complète. En conséquence, l'objectif de l'algorithme est de fournir des résultats de suivi robustes aux mouvements rapides du personnage considéré et surtout à ses passages de profil ; ceci afin de conserver le plus longtemps possible une piste³ unique par individu présent dans la séquence.

¹Les partenaires du projet MediaWorks sont TF1, la société Aegis, le Limsi et l'INRIA.

²Les partenaires du projet Feria sont CS, l'INA, l'IRIT, Arte, NDS, l'IRISA et VECSYS.

³On appelle "piste" une trajectoire associée à un visage.

L'algorithme mis en place repose sur l'association d'une méthode de suivi de région par approche différentielle, et d'un couplage entre un module de suivi des yeux et un module de détection de visage. Dans un premier temps, la méthode de détection de visage est expliquée. Nous verrons ensuite l'utilisation de la méthode de suivi de région. Puis, nous détaillerons comment le module de détection intervient dans le suivi des yeux, et pourquoi ce dernier est nécessaire au suivi de visages.

7.6.1 Algorithme de détection de visages

La détection de visages dans une image est un problème difficile d'une grande variabilité. Les proportions et les traits diffèrent beaucoup d'un individu à l'autre, et les attributs comme des lunettes ou une moustache en changent fortement l'aspect. En plus de ces difficultés s'ajoutent les problèmes dus aux variations de luminosité et de pose.

La méthode utilisée a été développée par Garcia et Delakis [Garcia 02] [Delakis 03]. Elle repose sur une technique de filtrage⁴ rapide correspondant à une cascade d'opérations élémentaires (convolution et sous-échantillonnage). Les paramètres de ces opérations sont obtenus par apprentissage neuronal à partir d'une grande base d'exemples. Cette technique s'applique sur des images en niveaux de gris. Tout en demeurant rapide, elle est particulièrement robuste aux variations d'échelles, de pose et d'illumination, ainsi qu'aux éléments occultants, et au bruit d'acquisition. En sortie, elle renvoie la position des visages détectés dans l'image sous la forme d'une liste de rectangles, chacun contenant la partie centrale du visage qu'il représente (incluant yeux, nez, bouche). Une illustration de résultats de détection sur différentes images est donnée dans la figure 7.30.



FIG. 7.30: Résultat de la détection de visages

⁴Le terme "filtrage" doit être ici compris comme l'application de filtre ou masque

Cependant, quelques fausses alarmes ou non détections d'un visage peuvent survenir, comme cela est montré dans la figure 7.31. Du fait de ces erreurs, une méthode de suivi de visage construite uniquement sur une mise en correspondance des résultats de détection n'est pas satisfaisante.



FIG. 7.31: Résultat de la détection de visages - cas d'une fausse alarme, et d'une non détection d'un visage de profil

Dans chaque visage détecté, une procédure de localisation des yeux est réalisée. Celle-ci est effectuée à l'aide d'une méthode de "eigenfeature" [Vastel 04]. Cette approche dérive de la technique classique de détection de visage par "eigenfaces" [Moghaddam 97] [Turk 91] et est adaptée aux éléments faciaux caractéristiques (yeux, nez, bouche). Elle consiste à caractériser chaque élément par apprentissage. Face à un nouveau visage, la localisation de l'œil gauche et de l'œil droit se fait par une recherche exhaustive de la sous-partie de l'image qui correspond le plus aux caractéristiques issues de l'apprentissage. Cette méthode est intéressante pour sa capacité d'adaptation et sa robustesse aux conditions diverses pouvant apparaître dans une séquence.

7.6.2 Utilisation d'une méthode différentielle

Comme introduit précédemment, l'algorithme de suivi de visage repose en partie sur une méthode différentielle. Dans la première image, les régions de l'image signalées par le module de détection permettent d'initialiser les différentes pistes associées aux personnages de la séquence. L'évolution de chacune de ces régions le long de la séquence est régie par une méthode d'estimation de mouvement paramétrique, dont le support est la région elle-même. La méthode utilisée est une approche différentielle, identique à celle mise en œuvre pour l'estimation de notre modèle dynamique (elle est détaillée §7.1.1). L'application du mouvement estimé sur chaque fenêtre permet de retrouver la zone d'intérêt (le visage suivi) dans l'image suivante.

Ce processus est efficace tant que l'estimation du mouvement est pertinente. Or, celle-ci devient mauvaise dans des cas de mouvements trop rapides de la tête, et surtout lorsque le personnage passe de profil. Dans ces situations difficiles, il est nécessaire d'une part de détecter le problème, et d'autre part de recalculer la fenêtre sur le visage. Pour répondre à ces deux items, il est proposé de s'appuyer sur un suivi des yeux. Ce module est expliqué dans le paragraphe suivant.

Dans la suite de la séquence, de nouveaux personnages peuvent apparaître. Le critère d'ajout de piste est lié au résultat (présence ou non de visages) fourni par le module de détection. Ce module est appliqué à chaque image. Une étape de mise en correspondance (basée sur des critères géométriques de recouvrement) est réalisée afin d'associer les visages détectés aux pistes suivies. Un visage détecté mais non affecté sera utilisé comme initialisation d'une nouvelle piste.

7.6.3 Algorithme de suivi des yeux

Dans le but de construire un algorithme efficace de suivi des yeux, l'idée mise en place consiste à coupler un algorithme de suivi de point au module de détection. Ce dernier permet de détecter à chaque image les visages présents dans la scène et de localiser la position des yeux associés.

La reconstruction de la trajectoire des yeux s'appuie sur l'algorithme linéaire proposé TFLC. Rappelons que cet algorithme est dédié aux points appartenant au support du mouvement dominant et que l'estimation de mouvement est effectuée sur l'image entière. Ces hypothèses ne sont pas vérifiées lorsque les points caractéristiques suivis sont des yeux. En effet, ces points appartiennent à des personnes la plupart du temps mobiles. Bien que l'algorithme adapté à cette situation soit le TFNLC, celui-ci est remplacé par le TFLC dont le temps d'exécution est plus faible. Pour cela, il est nécessaire de faire l'approximation qui suppose linéaire l'équation dynamique construite avec le mouvement estimé sur le visage. Cette approximation est raisonnable dans ce cas puisque le support d'estimation du mouvement correspond à une grande région de l'image (cf. *Choix du support d'estimation*).

Afin de rendre spécifique notre algorithme à la trajectographie des yeux, il est nécessaire d'en redéfinir deux éléments, à savoir le choix du support d'estimation et le choix de l'image de référence. Les autres éléments de la méthode restent inchangés (modèle dynamique estimé par une méthode différentielle, observations de l'état apportées par une mesure de corrélation, estimation automatique des matrices de covariance des bruits, utilisation de la fenêtre de validation).

- *Choix du support d'estimation*

Rappelons que le support d'estimation est utilisé pour calculer le mouvement qui sert à inférer l'équation dynamique. Naturellement, le support qui semble le plus approprié est une boîte englobant le visage. Il est choisi commun aux deux yeux, et correspond à la fenêtre évolutive estimée par la méthode différentielle. On note ici une différence avec l'algorithme de base où le support d'estimation est toujours choisi comme une petite région autour du point. Ici, le support d'estimation et donc le domaine de validité du mouvement paramétrique estimé correspondant à une partie importante de l'image. Cela justifie l'approximation linéaire de l'équation dynamique.

- *Choix de l'image de référence*

La détection des visages et des yeux est réalisée à chaque image. On suppose que cette détection est toujours pertinente et correspond à un personnage intéressant (qui est en général au premier plan). L'idée est alors de mettre en correspondance

chaque visage détecté avec les pistes suivies. Les caractérisations référence des yeux d'une piste sont choisies comme étant des fenêtres autour des yeux détectés du visage associé. Il arrive que la détection des yeux ne soit pas précise et pointe sur une autre partie du visage. Cependant, ces mauvaises détections sont ponctuelles. La détection systématique permet de mettre à jour la référence très régulièrement, et de se recalculer sur les yeux. Dans le cas où il n'y a pas de visage détecté, les caractérisations référence ne sont pas mises à jour.

Dans les cas simples, lorsque le visage suivi est de face et a des mouvements raisonnables, la méthode de suivi de région ainsi que la reconstruction des trajectoires des yeux sont satisfaisantes. Les sorties de ces deux modules sont cohérentes, c'est-à-dire que les yeux restent dans la fenêtre englobante.

Les problèmes apparaissent essentiellement lors des passages de profil. Dans de telles situations, un des yeux est caché et son suivi échoue naturellement. Cet échec est détecté par le biais de la matrice de covariance de l'erreur, dont les valeurs diagonales deviennent très grandes. Dans ce cas, il est fréquent que l'œil mal suivi sorte de la fenêtre englobante. L'idée mise en place consiste à se référer à l'œil encore visible, détectable par le fait qu'il soit encore à l'intérieur de la fenêtre englobante. La qualité de son suivi peut également être vérifiée par la matrice de covariance d'erreur. Trois situations se distinguent :

1. Lorsque le passage de profil est très rapide, il n'est pas correctement pris en compte dans le processus de trajectographie de l'œil occulté. Néanmoins, la distance entre les deux positions estimées des yeux reste du même ordre de grandeur que dans une situation d'un visage de face. L'idée est alors de repositionner la fenêtre englobante suivant des critères de proportion spécifiques à la position des yeux dans un visage. La position de cette fenêtre sera pertinente dès le retour du visage de face.
2. Lorsque la position de profil dure plus longtemps, la trajectoire de l'œil occulté diverge nettement. L'écart entre les deux yeux augmente significativement. Afin de rester sur la zone du visage, la fenêtre englobante est recentrée autour de l'œil bien suivi (présent dans le cadre).
3. La situation où les deux yeux sont en dehors de la fenêtre englobante est un signe de mise en difficulté de l'estimation. Dans ce cas, un sursis de 11 images est accordé à l'algorithme pour revenir à une des situations précédentes. En effet, le suivi est considéré fiable à partir du moment où au moins un des deux yeux est dans la fenêtre. Si ce n'est pas le cas à la fin du sursis, le suivi est considéré comme incohérent et la piste est arrêtée.

Enfin, les critères d'arrêt d'une piste sont, d'une part la détection d'un suivi non pertinent après le sursis (les deux points sont hors de la fenêtre), et d'autre part une fenêtre située en bordure d'image.

7.6.4 Résultats expérimentaux

L'algorithme de suivi de visages développé est efficace et robuste aux mouvements rapides ainsi qu'aux passages de profil. Des tests sur des séquences réelles ont permis de le valider. Nous en présentons trois exemples. Sur ces exemples, les résultats affichés montrent

l'estimation de la position des yeux par des croix, et la fenêtre englobante définie par un rectangle.

La première séquence présente une personne faisant un discours sur 52 images. Elle reste de profil entre les images 13 et 49. Le résultat du suivi est donné par le figure 7.32. Ce résultat est très satisfaisant et démontre la capacité de l'algorithme à gérer les passages de profil. Pendant cette période, un seul des deux yeux est correctement suivi, mais la fenêtre englobante reste centrée sur le visage.

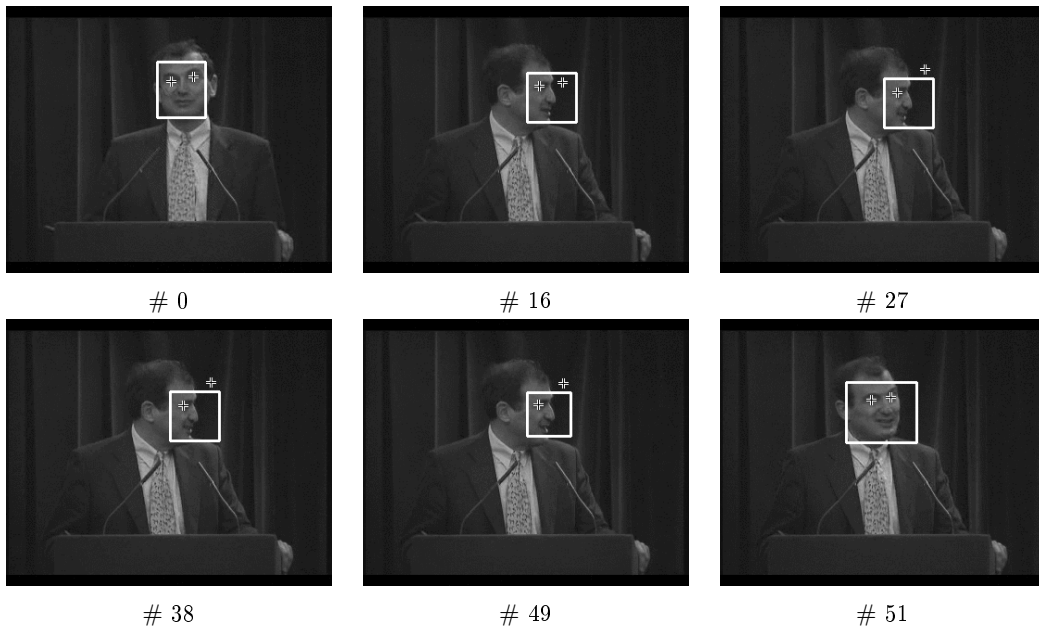


FIG. 7.32: Résultat de l'algorithme de suivi de visage - Illustration de la robustesse au passage de profil

La deuxième séquence de 31 images illustre le suivi de deux personnes. Ses difficultés sont les mouvements rapides des visages, qui se répercutent par un flou de bougé important, ainsi que l'occultation de la personne au deuxième plan par celle au premier plan. Les résultats du suivi de ces deux têtes sont montrés par la figure 7.33. Ces derniers permettent de valider l'utilisation de la méthode développée face à ces situations difficiles. Enfin, les résultats obtenus sur la troisième séquence (figure 7.34) montrent la robustesse de la méthode face à un visage dont les expressions changent beaucoup le long de la séquence.

7.7 Synthèse

Les algorithmes proposés pour la trajectographie de points caractéristiques reposent sur trois éléments essentiels :

1. Un modèle dynamique adaptatif, inféré par une mesure de mouvement. Ce mouvement est estimé grâce à une approche différentielle et apporte une information sur l'évolution du point entre deux images successives.
2. Un modèle d'observation simple. Les observations de l'état sont obtenues par une méthode de mise en correspondance par SSD et apportent une mesure de similarité

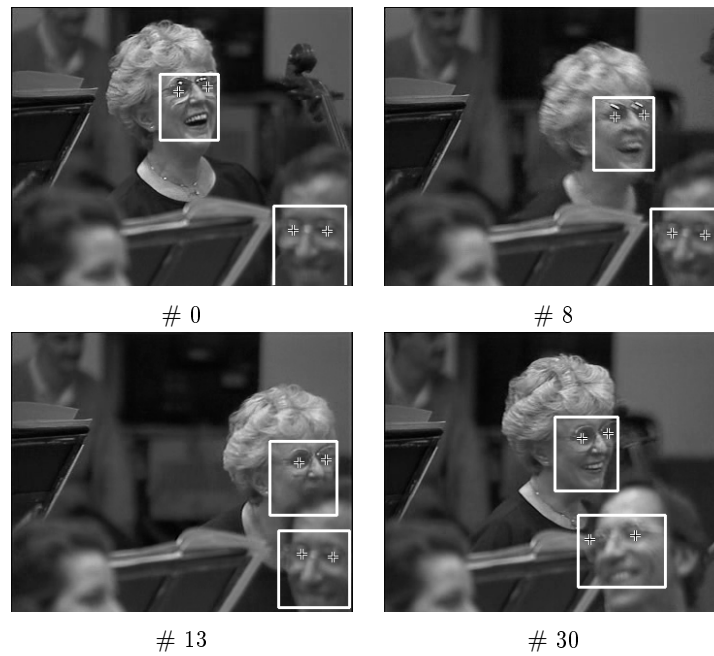


FIG. 7.33: Résultat de l'algorithme de suivi de visage - Illustration de la robustesse au flou de bougé et aux occultations partielles

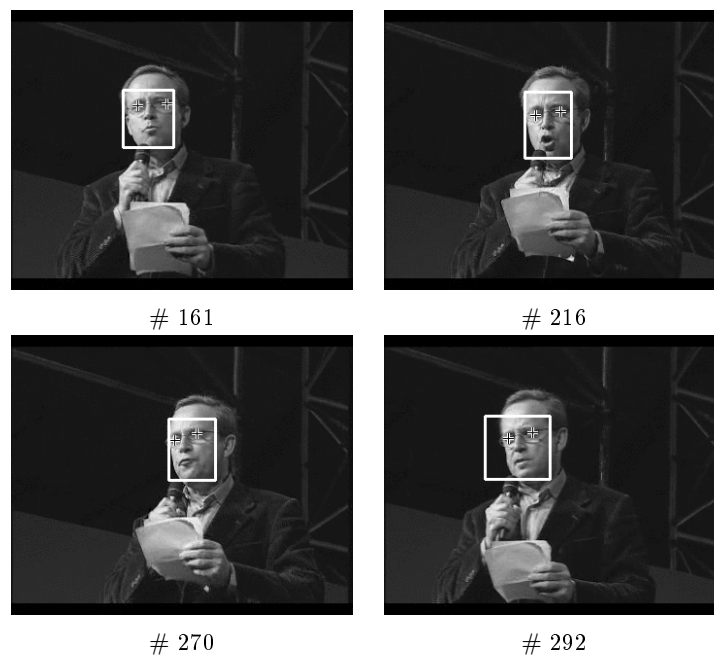


FIG. 7.34: Résultat de l'algorithme de suivi de visage - Illustration de la création de la robustesse aux changements d'expressions

entre l'instant courant et l'instant initial.

3. Une estimation à chaque itération des covariances des bruits du système. Ces matrices rendent alors compte d'une mesure de confiance des modèles dynamique et d'observation utilisés. Elles permettent d'en modérer l'importance si ceux-ci sont imprécis.

Ces éléments rendent les algorithmes proposés entièrement adaptatifs et ne nécessitant pas d'information *a priori* sur la loi d'évolution du point suivi. Ils sont théoriquement valides grâce au cadre des filtres conditionnels exposé dans la partie II.

Trois situations ont été distinguées, ce qui a permis de construire trois méthodes adaptées. La première correspond au suivi des points appartenant au support du mouvement dominant. L'algorithme associé (TFLC) repose sur le filtre linéaire conditionnel. Il présente l'avantage important de combiner une information locale fournie par la mesure de corrélation à une information globale apportée par le modèle dynamique. Celle-ci est d'une importance cruciale face à des images très bruitées. La deuxième situation se rapporte aux points appartenant à des zones de mouvement secondaire. La méthode dédiée (TFNLC) est construite sur un filtre non linéaire conditionnel. Enfin, la troisième situation correspond à un suivi dans des séquences d'images où les possibilités d'ambiguïtés sont importantes. Ces difficultés sont alors gérées par l'algorithme TFNLC multi-mesures. Ce dernier s'appuie sur une méthode particulière et permet de prendre en compte plusieurs observations par point suivi.

Les résultats expérimentaux obtenus sur de nombreuses séquences réelles ont permis de valider nos algorithmes. En effet, des trajectoires subissant des changements abrupts, des occlusions, sur des environnements bruités ont été reconstruites avec succès. Enfin, l'utilisation de notre approche pour une application de suivi de visage a montré la facilité d'adaptation de notre algorithme au suivi d'yeux, et son intérêt dans une application de plus haut niveau.

Les algorithmes de trajectographie de points présentés dans ce chapitre ont été publiés dans les articles de conférence internationale [Arnaud 04a] [Arnaud 04c] et dans la revue IEEE Transactions on Image Processing [Arnaud 04b].

Chapitre 8

Extension à la trajectographie de motifs plans

Une approche similaire à celle développée pour la trajectographie de points caractéristiques peut être mise en place pour la dérivation d'un algorithme de suivi d'un nuage de points. Un tel algorithme est très intéressant car un nuage de points permet de caractériser avec précision la texture d'un objet. Le cadre de filtrage permet d'introduire simplement des contraintes géométriques liant les primitives¹ entre elles. Nous nous sommes intéressés plus particulièrement au suivi d'objets plans en considérant une contrainte de type homographie. Cette étude est motivée par l'importance du suivi de ces objets dans de nombreuses applications, et notamment pour les systèmes basés vision en robotique.

Dans ce chapitre, nous présentons l'algorithme proposé pour le suivi d'un motif plan. Celui-ci est construit à partir d'un modèle du même type que celui utilisé par nos approches de trajectographie de points, à savoir entièrement estimé sur la séquence d'images. Le développement principal consiste en la généralisation de ce modèle vers un modèle décrivant un nuage de points, qui sera résolu au travers d'un filtre particulière rao-blackwellisé.

8.1 De la trajectographie d'un point vers le suivi d'un groupe de points

Dans le chapitre précédent, des algorithmes robustes pour le suivi d'un point caractéristique ont été définis. Ces méthodes sont construites sur la base de techniques de filtrage, et considèrent un modèle entièrement estimé sur la séquence d'images. Rappelons que ce modèle consiste en la combinaison d'une dynamique construite sur la contrainte de flot optique, de mesures fournies par une méthode de corrélation, et d'une estimation des covariances des bruits du système. Ce modèle a permis la dérivation de deux méthodes, la première est utilisée pour le suivi de points appartenant au mouvement dominant alors que la seconde est mise en œuvre pour les points dont le mouvement ne peut être décrit que localement.

Une extension logique de cette approche vers un algorithme de suivi d'un nuage de

¹Dans notre cas, une primitive correspond à un point du nuage

points consiste en l'utilisation du même type de modèle entièrement estimé sur la séquence d'images. La solution la plus simple est de considérer l'état $\mathbf{x}_k$ du système comme étant la concaténation des coordonnées des primitives à estimer ; $\mathbf{x}_k$ est alors un vecteur de taille $2p$, où p est le nombre d'éléments du groupe de points. Le vecteur d'observation $\mathbf{z}_k$, de taille $2p$, est alors composé des positions des pics de corrélation de chaque point. En ce qui concerne le vecteur de paramètres du mouvement instantané $\boldsymbol{\theta}_{k|\mathbf{I}}$, il est unique pour l'ensemble des points et est estimé sur une région $\mathcal{R}$ contenant le nuage de primitives. En raison de la dépendance de ce vecteur de mouvement instantané à la position des points, l'équation dynamique est non linéaire. Le système ainsi construit s'écrit :

$$\begin{cases} \mathbf{x}_k = \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1})\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ \mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (8.1)$$

où les bruits d'état $\mathbf{w}_k$ et de mesure $\mathbf{v}_k$ sont des vecteurs de taille $2p$, supposés blancs et gaussiens de moyenne nulle. La non linéarité de l'équation dynamique implique l'utilisation du FNLC, ce qui correspond en pratique à un filtre particulaire.

Cependant, cette approche n'est pas satisfaisante en raison de la grande dimension du vecteur d'état. En effet, la précision de la méthode de Monte Carlo mise en œuvre dépend fortement de la qualité de la phase d'échantillonnage des particules. Celle-ci peut se révéler inefficace (ou beaucoup trop onéreuse si le nombre de particules est très élevé) face à un espace de très grande dimension. Ce problème important des méthodes de Monte-Carlo est accentué dans les applications de vision par ordinateur, où le nombre de particules doit être limité pour rendre les temps de calcul acceptables.

Ce problème peut être en partie contourné si le vecteur d'état se décompose en deux composantes $\mathbf{r}_k$ et $\mathbf{y}_k$, dont une partie peut être traitée par une méthode optimale (par exemple un filtre de Kalman). Il est alors possible d'utiliser une méthode de réduction de variance par rao-blackwellisation qui consiste à marginaliser la distribution de filtrage relativement à la partie de l'état précédemment estimée. Cette procédure (décrite dans le paragraphe 2.5 de la partie I) fournit un estimateur plus efficace que le filtre particulaire classique. Elle reste applicable dans le contexte des filtres conditionnels, en introduisant un conditionnement par rapport à la séquence d'images.

Sur la base du modèle (8.1), nous proposons un nouveau système dédié au suivi d'un nuage de points, dont les propriétés rendent possible l'utilisation de la rao-blackwellisation. Celui-ci est décrit dans le paragraphe suivant.

8.2 Modèle pour la trajectographie d'un nuage de points

8.2.1 Modèle dynamique en deux étapes

Dans le chapitre 6, nous avons vu qu'un des avantages des algorithmes classiques de suivi de motifs plans est de pouvoir prendre en compte simplement un modèle de mouvement de type homographie. Rappelons que l'homographie est un principe de la géométrie projective décrivant le lien entre deux projections 2D d'un même objet plan de l'espace 3D. Plus précisément, les coordonnées d'un point quelconque du plan projeté dans deux images sont reliées linéairement par la matrice d'homographie $\mathcal{H}$.

Cette matrice est estimée à partir d'au moins quatre points projetés mis en correspondance.

Afin d'intégrer cette contrainte géométrique dans notre modèle, l'idée proposée consiste à séparer le nuage de points en deux sous-ensembles, notés $\mathbf{r}_k$ et $\mathbf{y}_k$. Le vecteur $\mathbf{r}_k$ est composé des points utilisés pour le calcul de la matrice d'homographie notés $\mathbf{r}_k(1), \mathbf{r}_k(2), \dots, \mathbf{r}_k(r)$. Il est de taille $2r$ (avec $4 \leq r$). Le vecteur $\mathbf{y}_k$ rend compte des points restants, et est de taille $3(p - r)$ (pour décrire ces points, nous utilisons les coordonnées homogènes). Le modèle dynamique décrivant l'évolution de l'état $\mathbf{x}_k = (\mathbf{r}_k, \mathbf{y}_k)^t$ est construit sur l'image en deux étapes.

La première étape concerne le déplacement des points de $\mathbf{r}_k$ appelés *points de référence*. Dans le but d'être réactif aux changements brusques de la trajectoire du plan, aucune information *a priori* n'est utilisée. De façon similaire à l'approche développée pour le suivi d'un seul point, nous proposons d'estimer un mouvement paramétrique instantané sur l'image par une méthode différentielle robuste. Les paramètres du mouvement $\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{r}_{k-1})$ sont calculés sur une région contenant l'ensemble des points de référence, ce qui assure une cohérence entre eux. En faisant le choix d'un modèle affine, le vecteur $\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{r}_{k-1})$ est composé de 6 paramètres. L'équation dynamique des points de référence peut alors s'écrire sous la forme suivante :

$$\mathbf{r}_k = \mathbf{r}_{k-1} + P(\mathbf{r}_{k-1}) \boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{r}_{k-1}) + \mathbf{w}_k^{\mathbf{r}}. \quad (8.2)$$

$P(\mathbf{r}_k)$ est une matrice de taille $2r \times 6$ dont l'expression est :

$$P(\mathbf{r}_k) = \begin{bmatrix} 1 & u_{\mathbf{r}_k(1)} & v_{\mathbf{r}_k(1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & u_{\mathbf{r}_k(1)} & v_{\mathbf{r}_k(1)} \\ \vdots & & & & & \\ 1 & u_{\mathbf{r}_k(r)} & v_{\mathbf{r}_k(r)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & u_{\mathbf{r}_k(r)} & v_{\mathbf{r}_k(r)} \end{bmatrix}, \quad (8.3)$$

et $\mathbf{w}_k^{\mathbf{r}}$ est le vecteur de bruit de taille $2r$. Ce bruit est supposé blanc, gaussien, de moyenne nulle et de matrice de covariance $Q_{k|\mathbf{I}}^{\mathbf{r}}$. Par simple extension du cas dédié au suivi d'un seul point, cette matrice est supposée diagonale, de type $c \times Id$. La constante c peut prendre deux valeurs différentes en fonction d'un test de Page-Hinkley effectué sur le rapport entre le support de validité et le support d'estimation $\mathcal{R}$ de la méthode différentielle robuste.

La seconde étape concerne la définition de la loi d'évolution des points du vecteur $\mathbf{y}_k$ que nous appellerons *points attachés*. Au temps $k - 1$, une estimation des positions des points de référence $\mathbf{r}_{k-1}$ et des points attachés $\mathbf{y}_{k-1}$ est accessible. Si de plus on dispose d'une estimation des nouveaux points de références $\mathbf{r}_k$, la matrice d'homographie décrivant le changement de base entre les deux projections peut être estimée. Nous la notons $\mathcal{H}_{k|\mathbf{I}}(\mathbf{r}_{k-1:k})$. Cette matrice permet de définir l'équation dynamique des points associés. Cette équation est linéaire et s'écrit :

$$\mathbf{y}_k = \mathcal{H}_{k|\mathbf{I}}(\mathbf{r}_{k-1:k}) \mathbf{y}_{k-1} + \mathbf{w}_k^{\mathbf{y}}. \quad (8.4)$$

Le vecteur de bruit $\mathbf{w}_k^{\mathbf{y}}$, de taille $3(p - r)$, est supposé blanc, gaussien, de moyenne nulle et de matrice de covariance $Q_{k|\mathbf{I}}^{\mathbf{y}}$. Cette matrice est diagonale et fixée *a priori*.

Remarque : La matrice $Q_{k|\mathbf{I}}^{\mathbf{y}}$ est pour le moment fixée *a priori*. Il serait intéressant de définir une procédure pour son estimation, qui permettrait de rendre compte d'une confiance sur le calcul de l'homographie. Un calcul d'incertitude sur la matrice d'homographie est par exemple donné dans [Kanatani 99].

8.2.2 Modèle d'observation

Alors que le modèle dynamique permet de considérer une cohérence entre les points du nuage à la fois au sens du mouvement et géométrique, nous considérons un modèle d'observation local pour chaque primitive. Ce choix permet de rendre l'algorithme robuste aux occlusions partielles de l'objet. Par simple extension au cas d'un unique point, le vecteur d'observation $\mathbf{z}_k$, de taille $2p$, est composé des coordonnées (cartésiennes) des pics de corrélations déterminés par un critère de mise en correspondance de type SSD. L'équation de mesure est donc linéaire et s'écrit :

$$\mathbf{z}_k = H_k \begin{bmatrix} \mathbf{r}_k \\ \mathbf{y}_k \end{bmatrix} + \mathbf{v}_k, \quad (8.5)$$

où la matrice H_k a l'expression suivante :

$$H_k = \begin{bmatrix} H_k^{\mathbf{r}} \\ H_k^{\mathbf{y}} \end{bmatrix} \quad \text{avec} \quad H_k^{\mathbf{r}} = Id \quad \text{et} \quad H_k^{\mathbf{y}} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}. \quad (8.6)$$

La matrice $H_k^{\mathbf{r}}$ est de taille $2r \times 2r$ et la matrice $H_k^{\mathbf{y}}$ est de taille $2(p-r) \times 3(p-r)$. Le bruit de mesure $\mathbf{v}_k$ est un bruit blanc, gaussien, de moyenne nulle et de covariance $R_{k|\mathbf{I}}$. Cette matrice de covariance est diagonale par blocs. Chaque bloc correspond à un point du nuage, et est estimé sur la surface SSDS associée à ce point, selon la procédure décrite dans le paragraphe 7.2.1.

Remarque : Le modèle d'observation étant construit de manière locale pour chaque point, les images de référence des différentes primitives sont également mises à jour indépendamment les unes des autres. Il serait certainement bénéfique de considérer un modèle de référence global de l'objet plan. Le problème de l'actualisation de ce modèle de luminance est alors plus difficile. En effet, de faibles variations dans l'estimation des paramètres du modèle de mouvement peuvent entraîner des fortes erreurs de recalage lorsque le motif est de grande taille.

8.2.3 Algorithme final

En résumé, le modèle construit pour le suivi d'un nuage de points appartenant à un plan est un modèle conditionnellement linéaire gaussien, qui s'écrit :

$$\begin{cases} p(\mathbf{r}_k | \mathbf{r}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{r}_k ; \mathbf{r}_{k-1} + P(\mathbf{r}_{k-1}) \boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{r}_{k-1}), Q_k^{\mathbf{r}}) \\ p(\mathbf{y}_k | \mathbf{r}_{k-1:k}, \mathbf{y}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{y}_k ; \mathcal{H}_{k|\mathbf{I}}(\mathbf{r}_{k-1:k}) \mathbf{y}_{k-1}, Q_k^{\mathbf{y}}) \\ p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; H_k \begin{bmatrix} \mathbf{r}_k \\ \mathbf{y}_k \end{bmatrix}, R_{k|\mathbf{I}}) \end{cases} \quad (8.7)$$

La distribution $p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ peut être exprimée par l'équation :

$$p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}) = p(\mathbf{y}_{0:k}|\mathbf{r}_{0:k}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}). \quad (8.8)$$

On a également :

$$p(\mathbf{r}_{0:k}, \mathbf{y}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}) = p(\mathbf{y}_k|\mathbf{r}_{0:k}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k}). \quad (8.9)$$

Dans le cas du système (8.7), la loi $p(\mathbf{r}_{0:k}|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ peut être estimée par le filtre non linéaire conditionnel, où la phase d'échantillonnage est réalisée à l'aide de la fonction d'importance optimale. En ce qui concerne la loi conditionnelle $p(\mathbf{y}_k|\mathbf{r}_{0:k}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$, elle est gaussienne et décrite par ses deux premiers moments $\hat{\mathbf{y}}_k$ et $\Sigma_{k|k}$. L'algorithme final est une méthode de suivi qui repose sur le filtre non linéaire conditionnel rao-blackwellisé (où à chaque particule est associé un filtre linéaire conditionnel). Remarquons que grâce au caractère linéaire gaussien de l'équation de mesure, et de la propriété gaussienne du bruit d'état des points de référence, la phase d'échantillonnage de cet algorithme est effectuée par la fonction d'importance optimale. La méthode de suivi proposée, notée TFNLC-RB, est résumée par l'algorithme 19.

Remarque : L'algorithme construit pour le suivi d'un motif plan décrit par un nuage de points combine les méthodes TFLC et TFNLC. Rappelons que ces méthodes considèrent une observation par entité suivie à chaque itération. Dans le but de construire une technique de suivi d'un point robuste aux ambiguïtés engendrées par le fouillis des images, le TFNLC a été étendu au TFNLC multi-mesures, qui permet de considérer plusieurs observations par itération. Les performances de ce dernier sont meilleures que celles du TFNLC lorsque les images utilisées contiennent de nombreuses régions similaires. De façon analogue, afin de mieux gérer les situations d'ambiguïtés, il serait intéressant d'étudier la même extension adaptée à notre méthode de suivi d'un nuage de points.

8.3 Résultats expérimentaux

Afin de valider notre approche pour le suivi d'un objet plan, nous avons testé l'algorithme sur des séquences réelles. Sur les résultats présentés, les points en rouge désignent les points de référence alors que les points en vert désignent les points attachés. L'ensemble des séquences utilisées ont été acquises par un utilisateur amateur. Elles présentent donc des mouvements chaotiques ainsi que des flous de bougés.

Les deux premières séquences, **BureauOcclusion** (60 images) et **BureauHoule** (169 images) représentent un bureau recouvert de pochettes et livres assez similaires, qui créent de nombreuses situations d'ambiguïtés.

La séquence **BureauOcclusion** permet de valider notre méthode lorsque l'objet plan suivi est occulté. En effet, l'objet est caché par une main qui passe au-dessus du bureau. Les résultats obtenus par le TFNLC-RB sont présentés sur la figure 8.1. Ces résultats sont très satisfaisants puisque la trajectoire du nuage de point n'est pas perturbée par les deux occlusions successives.

Les résultats obtenus sur la séquence **BureauHoule** démontrent la robustesse du TFNLC-RB aux larges mouvements de caméra, qui entraînent des changements brusques

Algorithme 19 Algorithme TFNLC-RB : méthode de suivi dédiée à la trajectographie d'un nuage de points d'un plan. Cette méthode repose sur un filtre non linéaire conditionnel rao-blackwellisé pour le système (8.7).

- **initialisation :**

pour $i = 1 \dots N$, générer $\mathbf{r}_0^{(i)} \sim p(\mathbf{r}_0)$, fixer $w_0^{(i)} = 1/N$, $\hat{\mathbf{y}}_{0|0}^{(i)} = \hat{\mathbf{y}}_0$ et $\Sigma_{0|0}^{(i)} = \Sigma_0$

pour $k = 1, 2, \dots$

- **estimations sur $\mathbf{I}_{0:k}$:**

calcul de $\mathbf{z}_k$ par SSD, et estimation de $R_{k|\mathbf{I}}$

pour $i = 1 \dots N$ estimation de $\boldsymbol{\theta}_{k|\mathbf{I}}(\mathbf{r}_{k-1}^{(i)})$ et $Q_{k|\mathbf{I}}^{\mathbf{r}}{}^{(i)}$ à l'aide d'une estimation différentielle robuste d'un mouvement affine

- **échantillonnage pondéré séquentiel :**

pour $i = 1 \dots N$, générer $\mathbf{r}_k^{(i)} \sim p(\mathbf{r}_k | \mathbf{r}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k})$ et poser $\mathbf{r}_{0:k}^{(i)} \triangleq (\mathbf{r}_{0:k-1}^{(i)}, \mathbf{r}_k^{(i)})$

pour $i = 1 \dots N$, calculer les poids d'importance normalisés $\tilde{w}_k^{(i)}$

- **estimations sur $\mathbf{I}_{0:k}$:**

pour chaque particule $\mathbf{r}_{0:k}^{(i)}$, estimation de la matrice d'homographie $\mathcal{H}_{k|\mathbf{I}}^{(i)}(\mathbf{r}_{k-1:k})$

et de $Q_{k|\mathbf{I}}^{\mathbf{y}}{}^{(i)}$

- **banc de filtres linéaires conditionnels :**

pour chaque particule $\mathbf{r}_{0:k}^{(i)}$, estimer les paramètres $\hat{\mathbf{y}}_{k|k}^{(i)}$ et $\Sigma_{k|k}^{(i)}$ de la loi gaussienne

$p(\mathbf{y}_k | \mathbf{r}_{0:k}^{(i)}, \mathbf{z}_{1:k})$ par le FLC

- **estimations de Monte Carlo :**

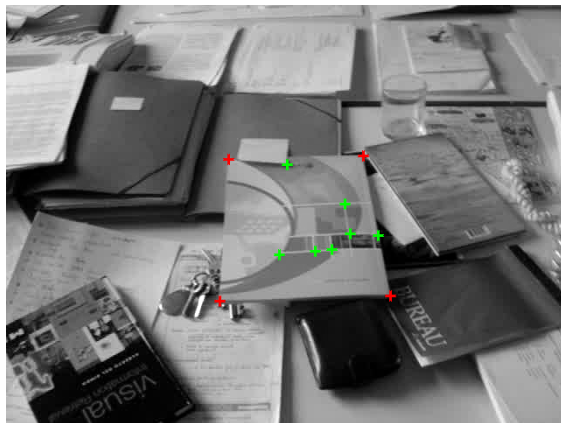
$$\mathbb{E}_{p(\mathbf{r}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})}[\phi(\mathbf{r}_{0:k})] \simeq \sum_{i=1}^N \phi(\mathbf{r}_{0:k}^{(i)}) \tilde{w}_k^{(i)}$$

$$\mathbb{E}_{p(\mathbf{y}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})}[\mathbf{y}_k] \simeq \sum_{i=1}^N \hat{\mathbf{y}}_{k|k}^{(i)} \tilde{w}_k^{(i)} \quad \mathbb{E}_{p(\mathbf{y}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})}[\mathbf{y}_k \mathbf{y}_k^t] \simeq \sum_{i=1}^N \Sigma_{k|k}^{(i)} \tilde{w}_k^{(i)}$$

- **rééchantillonnage adaptatif :**

si nécessaire, sélectionner N ensembles $\{\mathbf{r}_{0:k}^{(j)}, \hat{\mathbf{y}}_{k|k}^{(j)}, \Sigma_{k|k}^{(j)}\}_{j=1 \dots N}$ parmi

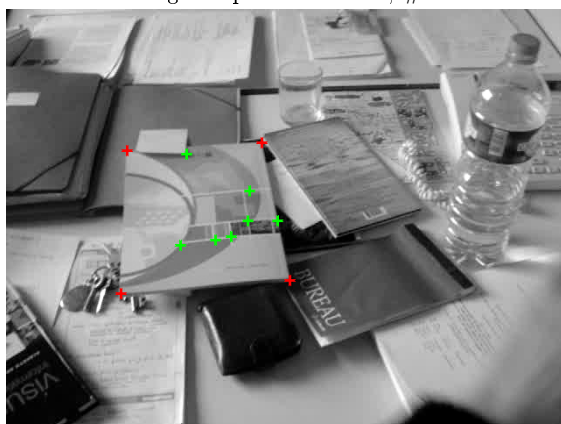
$\{\mathbf{r}_{0:k}^{(i)}, \hat{\mathbf{y}}_{k|k}^{(i)}, \Sigma_{k|k}^{(i)}\}_{i=1 \dots N}$ en fonction des poids $\tilde{w}_k^{(i)}$



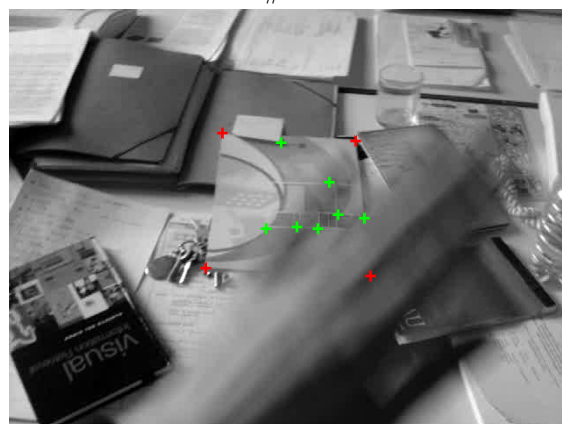
nuage de points initiaux, # 0



15



22



39



44



59

FIG. 8.1: Séquence BureauOcclusion – Résultats obtenus par le TFNLC-RB ; les points de référence sont en rouge, et les points attachés en vert.

de trajectoires. La figure 8.2 présente le suivi de deux nuages de points qui caractérisent deux objets différents. Il est important de noter que l'algorithme construit traite les situations où l'objet suivi sort partiellement de l'image de façon équivalente à une occultation partielle (sous réserve qu'il reste encore au moins quatre points de référence visibles, dans le cas contraire, le suivi est arrêté). Cela est illustré sur l'image 127 de la figure 8.2.

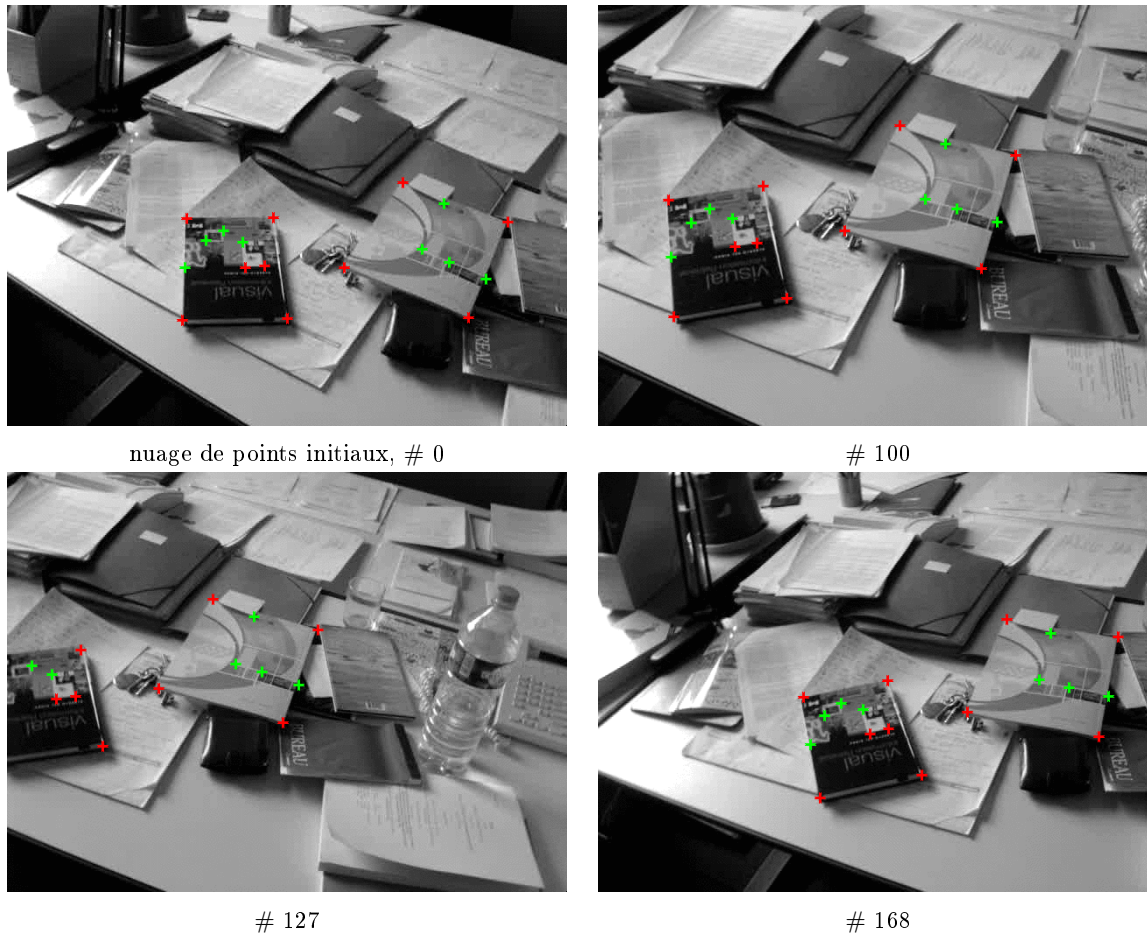


FIG. 8.2: Séquence BureauHoule – Résultats obtenus par le TFNLC-RB pour le suivi de deux objets. La situation où une partie de l'objet est hors de l'image est traitée comme une situation d'occlusion. Les points de référence sont en rouge, et les points attachés en vert.

La séquence **Visual** comporte 26 images. Sa caractéristique est de présenter des flous de bougé très importants, comme on peut le voir sur l'image 11 de la figure 8.3. De tels flous de bougé mettent en défaut les méthodes classiques de suivi. Comme on peut le constater sur la figure 8.3, le cadre du filtrage autorise la reconstruction de la trajectoire du nuage de points, sans que celle-ci soit perturbée.

Enfin, le TFNLC-RB a été testé sur la séquence **Campus** qui est composée de 113 images. Cette séquence illustre le passage de deux piétons devant des bâtiments d'un campus.

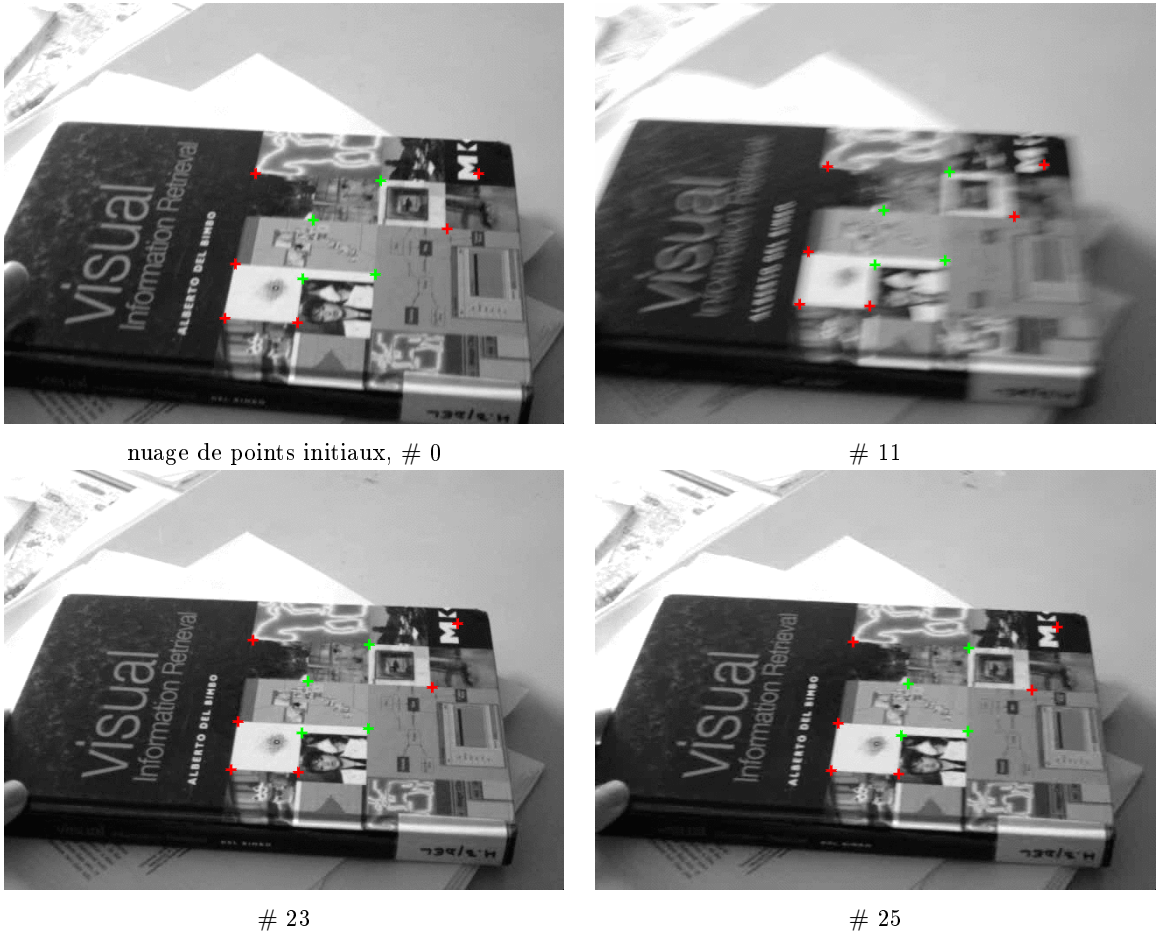


FIG. 8.3: Séquence Visual – Résultats obtenus par le TFNLC-RB. Les points de référence sont en rouge, et les points attachés en vert.

Réaliser un suivi d'une façade d'un des immeubles de cette séquence est une tâche complexe du fait des conditions d'illumination difficiles, et des occlusions par les piétons. Le résultat du suivi est présenté dans la figure 8.4. À part le décrochage d'un des points du nuage (voir l'image 34), les trajectoires reconstruites sont satisfaisantes malgré les fortes similarités entre les différents modèles de luminance des primitives du nuage. Ce décrochage fait suite à une occlusion, et le point s'accroche à une partie de l'image semblable à son motif de référence. Nous pensons qu'un tel décrochage pourrait être résolu par la prise en compte de plusieurs mesures par points, ainsi que la prise en compte d'un modèle de référence global pour l'ensemble de la façade.

8.4 Synthèse

Dans ce chapitre, nous avons présenté une extension de nos algorithmes de trajectographie de points caractéristiques vers une méthode de suivi d'un objet plan. L'objet est représenté par un nuage de points, qui constitue une bonne description de sa texture. Par l'utilisation du principe de réduction de variance par rao-blackwellisation, la méthode



FIG. 8.4: Séquence Campus – Résultats obtenus par le TFNLC-RB. Les points de référence sont en rouge, et les points attachés en vert.

construite permet de combiner les filtres FLC et FNLC. L'idée mise en œuvre est simple. Elle consiste à séparer le nuage de points en deux sous-ensembles, à savoir l'ensemble des *points de référence* et l'ensemble des *points attachés*.

Afin de construire un algorithme robuste aux occlusions, le modèle d'observation est supposé indépendant pour chaque point : le vecteur de mesure est la concaténation des pics de corrélations associés aux primitives du nuage. La cohérence spatiale entre les points apparaît aux travers du modèle d'évolution. Cette cohérence est de deux types. Premièrement, une information de mouvement est estimée sur une région englobant les points de référence. Ce mouvement est utilisé pour spécifier en ligne l'équation dynamique de ces points. Deuxièmement, une information géométrique est prise en compte. Celle-ci prend la forme d'une matrice d'homographie, estimée à partir des points de référence, et utilisée pour définir l'équation de prédiction linéaire des points attachés.

L'algorithme ainsi défini a été testé sur des séquences réelles. Ces séquences ont été acquises en amateur et présentent des mouvements chaotiques, des flous de bougé et des occultations. Les résultats obtenus ont démontré la robustesse de notre approche à ces situations difficiles. La combinaison d'informations de mouvement apportées par une méthode différentielle et une méthode de corrélation, et d'une contrainte géométrique sur la structure du nuage, permet de reconstruire avec succès les trajectoires des objets suivis.

Nous travaillons actuellement sur une version de cette méthode moins spécifique, pouvant suivre un objet plan déformable. L'idée mise en œuvre consiste à remplacer le modèle d'évolution construit sur une information de mouvement et une information géométrique par un modèle d'évolution construit sur une double estimation de mouvement. Nous pensons que celle-ci permettra de définir une zone grossière où l'objet s'est déplacé, puis de calculer un mouvement précis sur cette région.

Conclusion

La partie II de ce document a été dédiée à la modélisation d'un problème de suivi dans une séquence d'images par un système dynamique. Nous y avons étudié en particulier les systèmes construits avec une équation dynamique estimée en ligne sur la séquence. Du point de vue théorique, ces modèles se sont révélés intéressants. Nous nous sommes donc naturellement intéressés à leur validation pratique. Cette validation fait l'objet de cette dernière partie.

Dans un premier temps, nous avons étudié la dérivation d'algorithmes pour le suivi de points caractéristiques. La construction d'un modèle dynamique *a priori* de tels points est très difficile dans le cas général, où l'évolution de l'objet environnant est inconnue. La trajectographie de ces primitives constitue donc un cadre idéal pour tester nos modèles entièrement estimés sur les images. Le système mis en œuvre combine une équation dynamique construite sur la contrainte de flot optique et des observations fournies par une méthode de corrélation. L'estimation automatique des covariances de bruit du modèle permet de contrôler l'influence de ces deux informations de mouvement dans le calcul de l'estimée finale.

Trois méthodes de suivi ont été proposées. La première méthode repose sur un filtre linéaire conditionnel. Adaptée aux points appartenant au support de mouvement dominant de la séquence, elle permet de considérer à la fois une information de mouvement locale par le biais des mesures de corrélation et une information de mouvement globale de la scène au travers de la dynamique. De manière complémentaire, la seconde méthode permet de suivre des points dont le mouvement ne peut être décrit que localement. Sa structure algorithmique est le filtre non linéaire conditionnel, où la phase d'échantillonnage est réalisée par la fonction d'importance optimale. Enfin, la troisième est spécifique aux situations de fortes ambiguïtés entre différentes parties de l'image. Cette dernière s'appuie également sur une méthode particulière, mais autorise la prise en compte des pics de corrélation secondaires.

Pour ces trois méthodes, la combinaison de mesures calculées par une technique de mise en correspondance et d'un modèle d'évolution construit sur une estimation de mouvement a donné de très bons résultats. Les nombreux tests sur des séquences réelles permettent d'en attester. En effet, des trajectoires subissant des changements abrupts et/ou des occlusions dans des environnements bruités ont été reconstruites avec succès.

Dans un second temps, les résultats très satisfaisants obtenus pour le suivi de points ont amené à étendre nos algorithmes vers une méthode de trajectographie d'un motif plan. En décrivant l'objet par un nuage de points, le système proposé est de même type que précédemment, à savoir entièrement estimé sur la séquence. En plus de combiner deux

informations de mouvement, il prend en compte une information géométrique de type homographie qui permet de décrire une cohérence spatiale entre les primitives du nuage. En raison du caractère conditionnellement gaussien de ce modèle, le problème de suivi est résolu par un filtre non linéaire conditionnel rao-blackwellisé. La phase d'échantillonnage est une fois de plus réalisée avec la fonction d'importance optimale. Les résultats présentés sur des séquences réelles ont démontré la robustesse de notre approche aux occultations et aux flous de bougé.

Conclusion générale et perspectives

Dans ce document, nous nous sommes intéressés au problème de suivi dans une séquence d'images. Le choix d'une modélisation probabiliste de ce problème par un système dynamique permet la mise en œuvre d'algorithmes de filtrage. Ces algorithmes sont à la fois simples à implémenter et robustes aux situations pouvant entraîner des minima locaux. Ils nécessitent la définition d'un modèle décrivant les propriétés du système dynamique étudié. Ce modèle est composé d'une loi dynamique et d'une vraisemblance des données, et correspond à une modélisation par une chaîne de Markov cachée.

Pour une application donnée, le choix de l'algorithme dépend des caractéristiques du modèle construit. Dans le cas où les équations le définissant sont linéaires gaussiennes, le filtre de Kalman fournit une solution optimale au problème. Face à un système non linéaire non gaussien, l'utilisation des méthodes non aléatoires existantes telles que les extensions du filtre de Kalman et les méthodes numériques par maillage d'état est rarement satisfaisante. Le récent développement des méthodes séquentielles de Monte Carlo apporte une nouvelle solution au problème de filtrage non linéaire dont les performances dépassent celles des algorithmes non aléatoires dans le cas général.

L'application de ces algorithmes en vision par ordinateur implique la définition d'un modèle dynamique de l'entité suivie, de sa caractérisation en terme de données issues des images, et d'une équation décrivant le lien entre sa caractérisation et la nouvelle position recherchée. Le choix d'une dynamique estimée sur les images permet de guider avec précision la recherche, tout en étant adaptée à des algorithmes généraux. Cependant, ce choix entraîne une corrélation entre l'état et la mesure au travers des données image. Une telle corrélation n'est pas modélisée par le modèle simple de chaîne de Markov cachée, qui ne prend pas en compte la séquence d'images.

Synthèse des travaux effectués

Les travaux réalisés consistent en la **proposition d'une modélisation alternative**. Celle-ci permet de représenter explicitement la séquence d'images dans le modèle. La validation expérimentale de cette modélisation a été menée d'une part dans le cadre de la **trajectographie de points caractéristiques**, et d'autre part dans le cadre du **suivi de motifs plans**.

Proposition d'une modélisation alternative à la représentation par modèle simple de chaîne de Markov cachée

La modélisation proposée est une chaîne de Markov cachée où les données images sont introduites de manière explicite. Cette représentation originale a l'avantage d'être générale, car elle est adaptée au cas où la dynamique est estimée sur les images, et au cas où elle est fixée *a priori*. Les filtres dédiés à la résolution de ces nouveaux systèmes ont été dérivés sur la base des filtres classiques, en considérant un conditionnement par rapport à la séquence. Dans le cas où le système est constitué d'équations linéaires, il est résolu par le filtre linéaire conditionnel (FLC). Face à des non linéarités, il est nécessaire d'utiliser le filtre non linéaire conditionnel (FNLC). En pratique, la dérivation de ces filtres a permis de justifier l'emploi des équations des méthodes traditionnelles (filtre de Kalman et filtres particuliers) pour le traitement des systèmes dont les fonctions et matrices sont estimées sur la séquence.

Enfin, il a été présenté comment ce nouveau schéma permet de considérer des modèles simples, pour lesquels la fonction d'importance optimale est disponible et pour lesquels la fenêtre de validation des mesures peut être inférée de manière pertinente. En particulier, il est désormais possible de construire un modèle qui possède ces propriétés, et qui est adapté au cas où plusieurs observations de l'entité suivie sont disponibles.

Algorithmes de suivi de points caractéristiques

Les méthodes classiques de suivi de points caractéristiques peuvent être classés en deux catégories : les méthodes de corrélation et les méthodes différentielles. Afin de profiter des avantages de ces deux approches, le système proposé combine des mesures fournies par une méthode de corrélation et un modèle dynamique défini à partir d'estimations de mouvement. Le modèle considéré entre dans le cadre des systèmes aux propriétés intéressantes décrits précédemment. Afin de contrôler l'importance de ces deux informations de mouvement, un point important sur lequel nous avons travaillé est l'estimation des matrices de covariance des bruits du modèle.

Trois algorithmes complémentaires ont été construits. Le premier filtre (TFLC) est linéaire, particulièrement bien adapté aux séquences présentant un mouvement dominant. Le second filtre (TFNLC) repose sur un système non linéaire. Il permet de suivre des points dont le mouvement ne peut être décrit que localement. Enfin, le dernier (TFNLC multi-mesures) est approprié aux cas où les images ont un arrière plan encombré générant des situations d'ambiguïtés.

Algorithme de suivi d'un motif plan

Les résultats très satisfaisants obtenus pour le suivi de points ont amené à étendre le modèle développé à une application de suivi d'un motif plan. L'objet est décrit par un ensemble de points caractéristiques. La cohérence spatiale entre ces primitives est introduite par une estimation du mouvement sur la région les contenant, et par une contrainte géométrique de type homographie. La robustesse aux occultations est apportée par un modèle d'observation indépendant pour chaque point. Celui-ci s'appuie sur des mesures obtenues par une méthode de corrélation. Le système finalement construit peut être résolu par l'algorithme TFBLC-RB dont la structure est un FNLC rao-blackwellisé.

Perspectives

De nombreuses perspectives se dégagent de ces travaux, aussi bien dans l'approfondissement de l'étude réalisée que dans des axes de recherche plus généraux.

Approfondissement des travaux réalisés

Il nous semble pertinent de poursuivre l'étude effectuée dans les directions suivantes :

Problème de l'initialisation : La question de l'initialisation de nos algorithmes reste un problème ouvert. Les primitives suivies sont pour le moment choisies par un utilisateur. Dans le cadre du suivi de points caractéristiques, deux situations se distinguent. Dans un premier cas, certaines applications requièrent le suivi d'un grand nombre de points de la séquence. Un détecteur de points d'intérêt peut alors être utilisé en entrée des algorithmes. Le deuxième cas est plus problématique. Il correspond aux applications qui nécessitent le suivi de points associés à des "concepts" (un nez, des yeux, des joueurs de football, etc.). Des méthodes de détections doivent être définies. Celles-ci pourraient par exemple reposer sur des descripteurs appris à partir d'une base d'images. Mentionnons que le problème de la détection d'un motif plan est également à explorer.

Vers un algorithme de suivi d'objets déformables : Comme cela a été évoqué en conclusion du dernier chapitre, une de nos perspectives à court terme est la généralisation de l'algorithme dédié aux objets plans vers le suivi d'objets déformables. La première idée à mettre en œuvre consiste à construire un modèle d'évolution à partir d'une double estimation du mouvement. Il serait également intéressant d'intégrer des modèles de déformations plus complexes (par exemple un modèle de mouvement élastique) dans la loi dynamique *a priori* en remplacement de la contrainte homographique.

Développement d'algorithmes pouvant prendre en compte plusieurs observations : L'algorithme construit pour la gestion des ambiguïtés repose sur une méthode particulière. La résolution proposée du problème d'association des données s'appuie sur l'hypothèse que la probabilité d'observer aucune mesure valable est nulle. Cette hypothèse simplifie les équations de l'algorithme et est avantageusement compensée par une estimation en ligne des bruits de mesure. En conséquence, il nous apparaît important d'appliquer le même type d'hypothèse sur un algorithme linéaire (en particulier le filtre à association de données probabilistes), et de tester l'algorithme résultant. Dans la même direction, il serait intéressant d'étudier les avantages d'un filtre particulier rao-blackwellisé, pouvant prendre en compte plusieurs observations.

Évaluation des résultats de suivi

Dans un cadre plus général, l'évaluation d'une méthode de suivi reste problématique. Typiquement, pour les méthodes de filtrage, l'accès à une covariance d'erreur des estimées successives ne permet pas toujours la détection d'une mauvaise reconstruction de la trajectoire. Pour le moment, la validation des résultats se fait essentiellement de façon qualitative. Il conviendrait d'établir des critères automatiques d'arrêt d'un suivi. L'étude

de l'évolution de la répartition des particules dans l'espace, et de leur poids associé pourrait être une piste dans le cas des méthodes particulières.

Vers un suivi des structures caractéristiques en images de fluides

L'intégration conjointe d'une contrainte de mouvement estimée sur les images et d'une contrainte géométrique dans le modèle d'évolution a fourni des résultats intéressants pour le suivi d'objet plan. L'utilisation de cette approche dans des domaines d'imagerie, où des propriétés physiques des entités à suivre sont connues, pourrait se révéler intéressante. La trajectographie des structures caractéristiques (vortex, puits, sources) dans des images d'écoulements fluides entre dans ce cadre. Des méthodes d'estimation du mouvement spécifiques à ces images pourraient fournir une information pertinente, à combiner avec les équations physiques d'évolution de ces structures (équation de Navier-Stokes, équation de continuité).

Développement des méthodes séquentielles de Monte Carlo

Les méthodes séquentielles de Monte Carlo ont fait l'objet de nombreux travaux depuis dix ans. Pourtant, en raison de problèmes encore ouverts, elles constituent toujours une voie d'étude à explorer. Nous pouvons mentionner quelques difficultés parmi d'autres comme le choix du nombre de particules, la règle de décision pour savoir si l'étape d'échantillonnage doit être effectuée à une itération donnée ou encore l'incapacité de ces méthodes à suivre plusieurs modes de la loi de filtrage sur une longue durée.

En dernier lieu, nous pensons que la capacité de ces méthodes à explorer efficacement un espace pourrait être améliorée par l'introduction d'une coopération entre les particules. De telles idées sont utilisées dans les principes de métaheuristiques (par exemple les algorithmes de colonies de fourmis) qui permettent de résoudre des problèmes d'optimisation de fonctions non convexes.

Annexes

Annexe A

Estimation de la loi de filtrage $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ à partir d'une approximation de la loi complète $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$

Dans cette annexe, on démontre qu'il est possible d'estimer la loi de filtrage $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ à partir d'un nuage pondéré $\{\mathbf{x}_{0:k}^{(i)}, w_k^{(i)}\}_{i=1\dots N}$ approchant la loi complète $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$.

En effet, on remarquant que :

$$I(\phi(\mathbf{x}_k)) = \mathbb{E}_{p(\mathbf{x}_k | \mathbf{z}_{1:k})}[\phi(\mathbf{x}_k) | \mathbf{z}_{1:k}] \quad (\text{A.1})$$

$$= \int \phi(\mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k}) d\mathbf{x}_k \quad (\text{A.2})$$

$$= \int \phi(\mathbf{x}_k) \left[\int p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) d\mathbf{x}_{0:k-1} \right] d\mathbf{x}_k \quad (\text{A.3})$$

$$= \int \phi(\mathbf{x}_k) p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) d\mathbf{x}_{0:k} \quad (\text{A.4})$$

$$= \int \phi(\mathbf{x}_k) \frac{p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})}{\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})} \pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) d\mathbf{x}_{0:k} \quad (\text{A.5})$$

$$= \int \phi(\mathbf{x}_k) w_k \pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) d\mathbf{x}_{0:k} \quad (\text{A.6})$$

$$\text{où } w_k = \frac{p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})}{\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})}. \quad (\text{A.7})$$

On en déduit que si N échantillons i.i.d. $\{\mathbf{x}_{0:k}^{(i)}\}_{i=1\dots N}$ de la loi de proposition $\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ peuvent être tirés et que les poids d'importance normalisés associés $\{\tilde{w}_k^{(i)}\}_{i=1\dots N}$ peuvent être évalués, alors un estimateur de $I(\phi(\mathbf{x}))$ est donné par :

$$\hat{I}_N(\phi(\mathbf{x}_k)) = \sum_{i=1}^N \phi(\mathbf{x}_k^{(i)}) \tilde{w}_k^{(i)} \quad (\text{A.8})$$

$$\text{avec } \tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}} \text{ et } w_k^{(i)} = \frac{p(\mathbf{x}_{0:k}^{(i)}|\mathbf{z}_{1:k})}{\pi(\mathbf{x}_{0:k}^{(i)}|\mathbf{z}_{1:k})} \quad (\text{A.9})$$

Rappelons qu'il est simple d'obtenir $\mathbf{x}_k^{(i)}$ à partir d'un échantillon $\mathbf{x}_{0:k}^{(i)}$, puisque $\mathbf{x}_{0:k}^{(i)}$ est construit comme la concaténation des $\{\mathbf{x}_t^{(i)}\}_{t=0\dots k}$. Cela est valide même si la fonction de proposition $\pi(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$ n'a pas une forme causale.

De façon similaire, un estimateur de la densité de filtrage $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ est donné par

$$\sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k). \quad (\text{A.10})$$

Annexe B

Estimateurs de variance minimale

Cette annexe présente quelques résultats importants sur les estimateurs de variance minimale. Pour plus de détails sur les deux premiers estimateurs présentés, nous renvoyons au livre [Anderson 79].

B.1 Estimateur de variance minimale

Soit $(X, Z) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ un couple de vecteurs aléatoires. On note $\hat{X}$ l'estimateur de variance minimale de X . $\hat{X}$ est unique et correspond à l'espérance conditionnelle de X sachant Z .

$$\hat{X} = \mathbb{E}[X|Z] \quad (\text{B.1})$$

Il est à noter que $\hat{X}$ n'est pas nécessairement une fonction linéaire de Z .

B.2 Estimateur linéaire non biaisé de variance minimale

Soit $(X, Z) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ un couple de vecteurs aléatoires. On note $E^*[X|Z]$ l'estimateur linéaire non biaisé de variance minimale (BLUE : “best linear unbiased estimator”). Cet estimateur est défini par :

$$E^*[X|Z] = A^* Z + b^* \quad (\text{B.2})$$

avec A^* et b^* tels que

$$\mathbb{E}[\|X - A^* Z - b^*\|^2] \leq \mathbb{E}[\|X - AZ - b\|^2] \quad \forall A \in \mathbb{R}^{n_x \times n_z}, b \in \mathbb{R}^{n_x}. \quad (\text{B.3})$$

Il est à noter que $E^*[X|Z]$ n'est pas une espérance sauf dans un cas particulier (ce qui explique la notation traditionnelle E). Cet estimateur vérifie les théorèmes suivants (L'ensemble des démonstrations est disponible dans [Anderson 79]).

Propriété B.2.1 *Soit $(X, Z) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ un couple de vecteurs aléatoires de moyenne et covariance*

$$\begin{bmatrix} m_X \\ m_Z \end{bmatrix} \begin{bmatrix} \Sigma_{X,X} & \Sigma_{X,Z} \\ \Sigma_{Z,X} & \Sigma_{Z,Z} \end{bmatrix}, \quad (\text{B.4})$$

alors

$$E^*[X|Z] = m_X + \Sigma_{X,Z} \Sigma_{Z,Z}^{-1} (Z - m_Z). \quad (\text{B.5})$$

Propriété B.2.2 Soit $(X, Z) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ un couple de vecteurs aléatoires, de densité jointe gaussienne, alors l'estimateur linéaire de variance minimale correspond à l'estimateur de variance minimale.

$$E^*[X|Z] = \mathbb{E}[X|Z] \quad (\text{B.6})$$

Propriété B.2.3 (propriété de l'estimateur à être non biaisé) L'estimateur linéaire de variance minimale est non biaisé :

$$\mathbb{E}[X - E^*[X|Z]] = 0 \quad (\text{B.7})$$

Propriété B.2.4 (conditionnement par des vecteurs aléatoires non corrélés)

Soit $(X, Z_1, Z_2, \dots, Z_k) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \dots \times \mathbb{R}^{n_z}$ un ensemble de vecteurs aléatoires, où les vecteurs $Z_1, Z_2, \dots, Z_k$ sont non corrélés (i.e. $\Sigma_{Z_i Z_j} = 0$ pour $i \neq j$), alors

$$E^*[X|Z_1, Z_2, \dots, Z_k] = E^*[X|Z_1] + E^*[X|Z_2] + \dots + E^*[X|Z_k] - (k-1) m_X \quad (\text{B.8})$$

Propriété B.2.5 (changement de variable conditionnelle) Soit $(X, Z) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ un couple de vecteurs aléatoires, soit $Y = C Z + d$ (avec C non singulière), alors

$$E^*[X|Y] = E^*[X|Z] \quad (\text{B.9})$$

Propriété B.2.6 (principe d'orthogonalité) Soit $(X, Z) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ un couple de vecteurs aléatoires, alors

$$\mathbb{E}[(X - E^*[X|Z]) Z^t] = 0 \quad (\text{B.10})$$

B.3 Estimateur linéaire conditionnel de variance minimale

Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires. On note

$$\Sigma_{X,Z|W} = \mathbb{E}[X Z^t | W] - \mathbb{E}[X|W] \mathbb{E}[Z^t | W]. \quad (\text{B.11})$$

On définit $E_W^*[X|Z]$ comme étant le meilleur estimateur (au sens de la norme L_2) de X conditionnellement à W , linéaire en Z , i.e.

$$E_W^*[X|Z] = A_W^* Z + b_W^* \quad (\text{B.12})$$

avec A_W^* et b_W^* tels que

$$\mathbb{E}[\|X - A_W^* Z - b_W^*\|^2 | W] \leq \mathbb{E}[\|X - AZ - b\|^2 | W] \quad \forall A \in \mathbb{R}^{n_x \times n_z}, b \in \mathbb{R}^{n_x}. \quad (\text{B.13})$$

Cet estimateur est appelé estimateur linéaire conditionnel de variance minimale. Certaines propriétés de l'estimateur linéaire non biaisé de variance minimale peuvent être étendues à ce nouvel estimateur.

Propriété B.3.1 Soit $X, Z, W \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires, alors

$$E_W^*[X|Z] = \mathbb{E}[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]) \quad (\text{B.14})$$

PREUVE : _____

Pour un triplet de vecteurs aléatoires quelconques (Y_1, Y_2, V) , on a le résultat suivant :

$$\begin{aligned} \Sigma_{Y_1, Y_2|V} &\triangleq \text{cov}(Y_1, Y_2|V) \\ &= \mathbb{E}[(Y_1 - \mathbb{E}[Y_1|V])(Y_2 - \mathbb{E}[Y_2|V])^t|V] \\ &= \mathbb{E}[Y_1 Y_2^t|V] - \mathbb{E}[Y_1|V] \mathbb{E}[Y_2|V]^t - \mathbb{E}[Y_1|V] \mathbb{E}[Y_2|V]^t \\ &\quad + \mathbb{E}[Y_1|V] \mathbb{E}[Y_2|V]^t \\ &= \mathbb{E}[Y_1 Y_2^t|V] - \mathbb{E}[Y_1|V] \mathbb{E}[Y_2|V]^t. \end{aligned} \quad (\text{B.15})$$

En notant la trace d'une matrice par l'opérateur tr , on peut écrire une expression de $\mathbb{E}[\|X - A Z - b\|^2|W]$:

$$\begin{aligned} \mathbb{E}[\|X - A Z - b\|^2|W] &= \mathbb{E}[(X - AZ - b)^t (X - AZ - b)|W] \\ &= tr \{ \mathbb{E}[(X - AZ - b) (X - AZ - b)^t|W] \} \\ &= tr \{ \mathbb{E}[X X^t|W] + A \mathbb{E}[Z Z^t|W] A^t \\ &\quad + b b^t - A \mathbb{E}[Z X^t|W] - \mathbb{E}[X Z^t|W] A^t \\ &\quad - \mathbb{E}[X|W] b^t - b \mathbb{E}[X^t|W] + A \mathbb{E}[Z|W] b^t + b \mathbb{E}[Z^t|W] A^t \} \\ &= tr \{ \Sigma_{X,X|W} + A \Sigma_{Z,Z|W} A^t - \Sigma_{X,Z|W} A^t - A \Sigma_{Z,X|W} \\ &\quad + (\mathbb{E}[X|W] - A \mathbb{E}[Z|W] - b)^t (\mathbb{E}[X|W] - A \mathbb{E}[Z|W] - b) \} \\ &= tr \{ (A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1}) \Sigma_{Z,Z|W} (A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1})^t \\ &\quad + \|\mathbb{E}[X|W] - A \mathbb{E}[Z|W] - b\|^2 \\ &\quad + tr \{ \Sigma_{X,X|W} - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \Sigma_{Z,X|W} \} \} \end{aligned} \quad (\text{B.16})$$

$\Sigma_{Z,Z|W}$ est une matrice semi-définie positive. On en déduit que $H = (A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1}) \Sigma_{Z,Z|W} (A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1})^t$ est également une matrice semi-définie positive (dans le cas où le déterminant de $\Sigma_{Z,Z|W}$ est nul, la relation est valable en utilisant la pseudo-inverse).

Le minimum de $\mathbb{E}[\|X - A Z - b\|^2|W]$ est atteint pour $tr\{H\} = 0$, et pour $\|\mathbb{E}[X|W] - A \mathbb{E}[Z|W] - b\|^2 = 0$, c'est à dire pour

$$A_W^* = \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1}, \quad (\text{B.17})$$

$$b_W^* = \mathbb{E}[X|W] - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \mathbb{E}[Z|W]. \quad (\text{B.18})$$

On en déduit l'expression de $E_W^*[X|Z] = A_W^* Z + b_W^*$:

$$E_W^*[X|Z] = \mathbb{E}[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]). \quad (\text{B.19})$$

_____ □

Propriété B.3.2 (propriété à être conditionnellement non biaisé) Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires, alors

$$\mathbb{E}[X - E_W^*[X|Z] \mid W] = 0 \quad (\text{B.20})$$

PREUVE : _____

$$\begin{aligned} \mathbb{E}[X - E_W^*[X|Z] \mid W] &= \mathbb{E}[X - \mathbb{E}[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]) \mid W] \\ &= \mathbb{E}[X|W] - \mathbb{E}[X|W] - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \mathbb{E}[(Z - \mathbb{E}[Z|W]) \mid W] \\ &= 0 \end{aligned} \quad (\text{B.21})$$

_____ □

Propriété B.3.3 (conditionnement par des vecteurs aléatoires non corrélés)

Soit $(X, Z_1, Z_2, \dots, Z_k, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \dots \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un ensemble de vecteurs aléatoires avec $Z_1, Z_2, \dots, Z_k$ non corrélés conditionnellement à W (i.e. $\Sigma_{Z_i Z_j} = 0$ pour $i \neq j$), alors

$$\begin{aligned} E_W^*[X|Z_1, Z_2, \dots, Z_k] &= E_W^*[X|Z_1] + E_W^*[X|Z_2] + \dots \\ &\quad + E_W^*[X|Z_k] - (k-1) \mathbb{E}[X|W] \end{aligned} \quad (\text{B.22})$$

PREUVE : _____

On pose $Z = [Z_1^t, Z_2^t, \dots, Z_k^t]^t$ alors

$$\begin{aligned} E_W^*[X|Z] &= \mathbb{E}[X|W] + \Sigma_{XZ|W} \Sigma_{ZZ|W}^{-1} (Z - \mathbb{E}[Z|W]) \\ &= \mathbb{E}[X|W] + [\Sigma_{XZ_1|W} \dots \Sigma_{XZ_k|W}] \text{Diag}[\Sigma_{Z_i Z_i|W}^{-1}] \sum_{i=1}^k (Z_i - \mathbb{E}[Z_i|W]) \\ &= \mathbb{E}[X|W] + \sum_{i=1}^k \Sigma_{XZ_i|W} \Sigma_{Z_i Z_i|W}^{-1} (Z_i - \mathbb{E}[Z_i|W]) \\ &= E_W^*[X|Z_1] + E_W^*[X|Z_2] + \dots + E_W^*[X|Z_k] - (k-1) \mathbb{E}[X|W] \end{aligned} \quad (\text{B.23})$$

_____ □

Propriété B.3.4 (changement de variable conditionnelle) Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires, et soit $Y = C Z + d$ (avec C non singulière), alors

$$E_W^*[X|Y] = E_W^*[X|Z] \quad (\text{B.24})$$

PREUVE : _____

On a les trois relations suivantes :

$$\begin{aligned} Y - \mathbb{E}[Y|W] &= C Z + d - C \mathbb{E}[Z|W] - d \\ &= C (Z - \mathbb{E}[Z|W]) \end{aligned} \quad (\text{B.25})$$

$$\begin{aligned}
\Sigma_{X,Y|W} &= \mathbb{E}[XY^t|W] - \mathbb{E}[X|W] \mathbb{E}[Y|W]^t \\
&= \mathbb{E}[X (C Z + d)^t|W] - \mathbb{E}[X|W] \mathbb{E}[C Z + d|W]^t \\
&= \mathbb{E}[X (Z^t C^t + d^t)|W] - \mathbb{E}[X|W] (\mathbb{E}[Z|W]^t C^t + d^t) \\
&= \mathbb{E}[X Z^t|W] C^t + \mathbb{E}[X|W] d^t - \mathbb{E}[X|W] \mathbb{E}[Z|W]^t C^t - \mathbb{E}[X|W] d^t \\
&= \Sigma_{X,Z|W} C^t
\end{aligned} \tag{B.26}$$

$$\begin{aligned}
\Sigma_{Y,Y|W} &= \mathbb{E}[(CZ + d) (Z^t C^t + d^t)|W] - \mathbb{E}[CZ + d|W] \mathbb{E}[CZ + d|W]^t \\
&= C \mathbb{E}[Z Z^t|W] C^t + d d^t + C \mathbb{E}[Z|W] d^t + d \mathbb{E}[Z^t|W] C^t \\
&\quad - C \mathbb{E}[Z|W] \mathbb{E}[Z|W]^t C^t - d d^t - C \mathbb{E}[Z|W] d^t - d \mathbb{E}[Z^t|W] C^t \\
&= C \Sigma_{Z,Z|W} C^t
\end{aligned} \tag{B.27}$$

d'où

$$\begin{aligned}
E_W^*[X|Y] &= \mathbb{E}[X|W] + \Sigma_{X,Y|W} \Sigma_{Y,Y|W}^{-1} (Y - \mathbb{E}[Y|W]) \\
&= \mathbb{E}[X|W] + \Sigma_{X,Z|W} C^t C^{t-1} \Sigma_{Z,Z|W}^{-1} C^{-1} C (Z - \mathbb{E}[Z|W]) \\
&= E_W^*[X|Z]
\end{aligned} \tag{B.28}$$

□

Propriété B.3.5 (principe d'orthogonalité) Soit $(X, Z, W) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$ un triplet de vecteurs aléatoires, alors

$$\mathbb{E}[(X - E_W^*[X|Z]) Z^t | W] = 0 \tag{B.29}$$

PREUVE : _____

$$\begin{aligned}
&\mathbb{E}[(X - E_W^*[X|Z]) Z^t | W] \\
&= \mathbb{E}[(X - \mathbb{E}[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W])) Z^t | W] \\
&= \mathbb{E}[X Z^t - \mathbb{E}[X|W] Z^t - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]) Z^t | W] \\
&= \Sigma_{X,Z|W} - \mathbb{E}[\Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]) Z^t | W] \\
&= \Sigma_{X,Z|W} - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \mathbb{E}[(Z - \mathbb{E}[Z|W]) (Z - \mathbb{E}[Z|W])^t | W] \\
&= \Sigma_{X,Z|W} - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \Sigma_{Z,Z|W} \\
&= 0
\end{aligned} \tag{B.30}$$

□

Annexe C

Filtre linéaire conditionnel : demonstrations

C.1 Description du modèle et notations

On considère une séquence de $k + 1$ images décrite par la variable aléatoire $\mathbf{I}_{0:k}$. On suppose que le problème de suivi d'une entité dans cette séquence peut être modélisé par le système linéaire à bruits additifs suivant :

$$\begin{cases} \mathbf{x}_k = F_{k|\mathbf{I}} \mathbf{x}_{k-1} + \mathbf{b}_{k|\mathbf{I}} + \mathbf{w}_k \\ \mathbf{z}_k = H_{k|\mathbf{I}} \mathbf{x}_k + \mathbf{v}_k. \end{cases} \quad (\text{C.1})$$

Les matrices $F_{k|\mathbf{I}}$, $H_{k|\mathbf{I}}$ et le vecteur $\mathbf{b}_{k|\mathbf{I}}$ sont déterministes. Les variables $\mathbf{w}_k$, $\mathbf{v}_k$ sont supposées être des bruits blanc indépendants (éventuellement non gaussiens), tels que :

$$\mathbb{E}[\mathbf{w}_k | \mathbf{I}_{0:k}] = 0, \quad \Sigma_{\mathbf{w}_k, \mathbf{w}_k | \mathbf{I}_{0:k}} = Q_{k|\mathbf{I}}, \quad (\text{C.2})$$

$$\mathbb{E}[\mathbf{v}_k | \mathbf{I}_{0:k}] = 0, \quad \Sigma_{\mathbf{v}_k, \mathbf{v}_k | \mathbf{I}_{0:k}} = R_{k|\mathbf{I}}. \quad (\text{C.3})$$

où la notation $\Sigma_{X,Z|W}$ est définie dans l'annexe B, équation B.11.

Contrairement au cas classique d'un système linéaire à bruits additif résolu par un filtre de Kalman [Anderson 79] [Kalman 60], les paramètres de ce modèle conditionnel (les matrices $F_{k|\mathbf{I}}$, $H_{k|\mathbf{I}}$, $Q_{k|\mathbf{I}}$, $R_{k|\mathbf{I}}$, et le vecteur $\mathbf{b}_{k|\mathbf{I}}$) peuvent dépendre des données images. Cette éventuelle dépendance est représentée par l'indice $|\mathbf{I}$. Pour simplifier les notations, cet indice sera supprimé dans les démonstrations qui suivent. On suppose également que la distribution de l'état initial est gaussienne d'espérance $\bar{\mathbf{x}}_0$ and covariance Σ_0 .

L'observation de l'état $\mathbf{z}_k$ est obtenue à chaque pas de temps par un processus d'estimation. Enfin, une hypothèse de causalité est faite par rapport à l'acquisition des images. Cette hypothèse se traduit par l'indépendance des variables $\mathbf{x}_k$ et $\mathbf{z}_k$ aux images futures $\mathbf{I}_n$ avec $n > k$.

On introduit les notations suivantes :

$\mathbf{z}_{1:k} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k\}$: l'ensemble des observations de 0 à k ,
 $\hat{\mathbf{x}}_{k|k-1} = E_{\mathbf{I}_{0:k}}^*[\mathbf{x}_k | \mathbf{z}_{1:k-1}]$: l'estimateur de $\mathbf{x}_k$ conditionnel en $\mathbf{I}_{0:k}$, linéaire en $\mathbf{z}_{1:k}$, de variance minimale,
 $\hat{\mathbf{z}}_{k|k-1} = E_{\mathbf{I}_{0:k}}^*[\mathbf{z}_k | \mathbf{z}_{1:k-1}]$: l'estimateur de $\mathbf{z}_k$ conditionnel en $\mathbf{I}_{0:k}$, linéaire en $\mathbf{z}_{1:k-1}$, de variance minimale,
 $\tilde{\mathbf{x}}_{k|k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}$: l'erreur de prédiction de $\mathbf{x}_k$,
 $\tilde{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}$: l'erreur d'estimation de $\mathbf{z}_k$. Cette variable est un processus d'innovation conditionnellement à $\mathbf{I}_{0:k}$,
 $\Sigma_{k|k-1} = \Sigma_{\tilde{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}}$: la matrice de covariance de l'erreur d'estimation conditionnelle, $\tilde{\mathbf{z}}_{2|1:k|k-1} = \{\tilde{\mathbf{z}}_{2|1}, \dots, \tilde{\mathbf{z}}_{k|k-1}\}$: l'ensemble des processus d'innovation.

On suppose que les variables de bruits $\mathbf{v}_k$ et $\mathbf{w}_k$ sont mutuellement indépendantes et indépendantes de $\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0$ et de $\mathbf{z}_k, \mathbf{z}_{k-1}, \dots, \mathbf{z}_1$ conditionnellement à $\mathbf{I}_{0:k}$. On en déduit que ces bruits sont également indépendants de $\tilde{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{x}}_{k-1|k-2}, \dots, \tilde{\mathbf{x}}_{2|1}$ et de $\tilde{\mathbf{z}}_{k|k-1}, \tilde{\mathbf{z}}_{k-1|k-2}, \dots, \tilde{\mathbf{z}}_{2|1}$ conditionnellement à $\mathbf{I}_{0:k}$.

Le but est de construire un algorithme itératif permettant d'estimer $\mathbf{x}_{k+1}$ connaissant l'ensemble des observations $\mathbf{z}_{1:k}$ et la séquence d'images $\mathbf{I}_{0:k+1}$. En d'autres termes, on cherche une expression récursive de $\hat{\mathbf{x}}_{k+1|k} = E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \mathbf{z}_{1:k}]$. On rappelle ici l'expression de l'estimateur utilisé, à savoir l'estimateur linéaire conditionnel de variance minimale :

Soit X, Z, W un triplet de vecteurs aléatoires $\in \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_w}$, alors

$$E_W^*[X|Z] = \mathbb{E}[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - \mathbb{E}[Z|W]) \quad (\text{C.4})$$

$$\text{avec } \Sigma_{X,Z|W} = \mathbb{E}[XZ^t|W] - \mathbb{E}[X|W] \mathbb{E}[Z^t|W]. \quad (\text{C.5})$$

La définition de cet estimateur ainsi que certaines de ses propriétés utiles pour la démonstration du filtre linéaire conditionnel sont données en annexe B.

C.2 Quelques propriétés utiles

Pour construire l'expression récursive de $\hat{\mathbf{x}}_{k+1|k}$, nous commençons par étudier quelques propriétés de variables impliquées dans le problème.

Propriété de $\tilde{\mathbf{x}}_{k|k-1}$

- L'espérance conditionnelle $\mathbb{E}[\tilde{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}]$ est égale à zéro.

$$\begin{aligned} \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}] &= \mathbb{E}[\mathbf{x}_k | \mathbf{I}_{0:k}] - \mathbb{E}[\hat{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}] \\ &= \mathbb{E}[\mathbf{x}_k | \mathbf{I}_{0:k}] - \mathbb{E}[E_{\mathbf{I}_{0:k}}^*[\mathbf{x}_k | \mathbf{z}_{1:k-1}] | \mathbf{I}_{0:k}] \\ &= \mathbb{E}[\mathbf{x}_k | \mathbf{I}_{0:k}] - \mathbb{E}[\mathbf{x}_k | \mathbf{I}_{0:k}] \\ &\quad - \mathbb{E}[\Sigma_{\mathbf{x}_k, \mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}} \Sigma_{\mathbf{z}_{1:k-1}, \mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}}^{-1} (\mathbf{z}_{1:k-1} - \mathbb{E}[\mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}]) | \mathbf{I}_{0:k}] \\ &= -\Sigma_{\mathbf{x}_k, \mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}} \Sigma_{\mathbf{z}_{1:k-1}, \mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}}^{-1} \mathbb{E}[\mathbf{z}_{1:k-1} - \mathbb{E}[\mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}] | \mathbf{I}_{0:k}] \\ &= 0 \end{aligned} \quad (\text{C.6})$$

Propriétés du processus d'innovation conditionnel $\tilde{\mathbf{z}}_{k|k-1}$

- L'espérance conditionnelle $\mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k}]$ est égale à zéro.

$$\begin{aligned}
 \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k}] &= \mathbb{E}[\mathbf{z}_k | \mathbf{I}_{0:k}] - \mathbb{E}[\hat{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k}] \\
 &= \mathbb{E}[\mathbf{z}_k | \mathbf{I}_{0:k}] - \mathbb{E}[E_{\mathbf{I}_{0:k}}^*[\mathbf{z}_k | \mathbf{z}_{1:k-1}] | \mathbf{I}_{0:k}] \\
 &= \mathbb{E}[\mathbf{z}_k | \mathbf{I}_{0:k}] - \mathbb{E}[\mathbf{z}_k | \mathbf{I}_{0:k}] \\
 &\quad - \mathbb{E}[\Sigma_{\mathbf{z}_k, \mathbf{z}_{1:k-1}} | \mathbf{I}_{0:k} \Sigma_{\mathbf{z}_{1:k-1}, \mathbf{z}_{1:k-1}}^{-1} | \mathbf{I}_{0:k} (\mathbf{z}_{1:k-1} - \mathbb{E}[\mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}]) | \mathbf{I}_{0:k}] \\
 &= -\Sigma_{\mathbf{z}_k, \mathbf{z}_{1:k-1}} | \mathbf{I}_{0:k} \Sigma_{\mathbf{z}_{1:k-1}, \mathbf{z}_{1:k-1}}^{-1} | \mathbf{I}_{0:k} \mathbb{E}[(\mathbf{z}_{1:k-1} - \mathbb{E}[\mathbf{z}_{1:k-1} | \mathbf{I}_{0:k}]) | \mathbf{I}_{0:k}] \\
 &= 0
 \end{aligned} \tag{C.7}$$

- À partir de la définition $\tilde{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \mathbb{E}_{\mathbf{I}_{0:k}}^*[\mathbf{z}_k | \mathbf{z}_{1:k-1}]$, on peut remarquer que $\tilde{\mathbf{z}}_{k|k-1}$ est orthogonal à $\mathbf{z}_{1:k}$ conditionnellement à $\mathbf{I}_{0:k}$. Cette variable est donc orthogonale à $\tilde{\mathbf{z}}_{l|l-1}$ avec $l < k$, qui est une combinaison linéaire de $\mathbf{z}_{1:l}$. L'ensemble de variables $\tilde{\mathbf{z}}_{k|k-1}$ sont mutuellement indépendantes conditionnellement à $\mathbf{I}_{0:k}$.

$$\mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} \tilde{\mathbf{z}}_{l|l-1}^t | \mathbf{I}_{0:k}] = 0 \quad \forall k, l \tag{C.8}$$

C.3 Expression récursive de $\hat{\mathbf{x}}_{k+1|k}$

En remarquant que la séquence $\tilde{\mathbf{z}}_{2|1:k|k-1}$ peut être obtenue par une opération linéaire à partir de $\mathbf{z}_{1:k}$ connaissant $\mathbf{I}_{0:k}$ (et inversement), on peut écrire grâce à la propriété B.3.4 :

$$\hat{\mathbf{x}}_{k+1|k} = E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \mathbf{z}_{1:k}] = E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \tilde{\mathbf{z}}_{2|1:k|k-1}]. \tag{C.9}$$

Comme il a été montré que les processus d'innovation conditionnel sont mutuellement indépendants (équation C.8), on a par la propriété B.3.3 :

$$\hat{\mathbf{x}}_{k+1|k} = E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \tilde{\mathbf{z}}_{k|k-1}] + E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \tilde{\mathbf{z}}_{2|1:k-1|k-2}] - \mathbb{E}[\mathbf{x}_{k+1} | \mathbf{I}_{0:k+1}] \tag{C.10}$$

avec

$$\begin{aligned}
 E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \tilde{\mathbf{z}}_{k|k-1}] &= \mathbb{E}[\mathbf{x}_{k+1} | \mathbf{I}_{0:k+1}] \\
 &\quad + \Sigma_{\mathbf{x}_{k+1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}} \Sigma_{\tilde{\mathbf{z}}_{k|k-1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}}^{-1} (\tilde{\mathbf{z}}_{k|k-1} - \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}])
 \end{aligned} \tag{C.11}$$

- **Première expression pour $\hat{\mathbf{x}}_{k+1|k}$**

$$\begin{aligned}
 \hat{\mathbf{x}}_{k+1|k} &= \Sigma_{\mathbf{x}_{k+1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}} \Sigma_{\tilde{\mathbf{z}}_{k|k-1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}}^{-1} (\tilde{\mathbf{z}}_{k|k-1} - \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}]) \\
 &\quad + E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \tilde{\mathbf{z}}_{2|1:k-1|k-2}]
 \end{aligned} \tag{C.12}$$

- Calcul de $E_{\mathbf{I}_{0:k+1}}^* [\mathbf{x}_{k+1} | \tilde{\mathbf{z}}_{2|1:k-1|k-2}]$

$$\begin{aligned}
E_{\mathbf{I}_{0:k+1}}^* [\mathbf{x}_{k+1} | \tilde{\mathbf{z}}_{2|1:k-1|k-2}] &= E_{\mathbf{I}_{0:k+1}}^* [\mathbf{x}_{k+1} | \mathbf{z}_{1:k-1}] \\
&= E_{\mathbf{I}_{0:k+1}}^* [F_{k+1} \mathbf{x}_k + \mathbf{b}_{k+1} + \mathbf{w}_{k+1} | \mathbf{z}_{1:k-1}] \\
&= F_{k+1} E_{\mathbf{I}_{0:k+1}}^* [\mathbf{x}_k | \mathbf{z}_{1:k-1}] + \mathbf{b}_{k+1} \\
&= F_{k+1} E_{\mathbf{I}_{0:k+1}}^* [\mathbf{x}_k | \mathbf{z}_{1:k-1}] + \mathbf{b}_{k+1} \\
&= F_{k+1} \hat{\mathbf{x}}_{k|k-1} + \mathbf{b}_{k+1}
\end{aligned} \tag{C.13}$$

- Calcul de $\Sigma_{\tilde{\mathbf{z}}_{k|k-1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}}$

On a :

$$\tilde{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - H_k \hat{\mathbf{x}}_{k|k-1} = H_k \tilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k \tag{C.14}$$

On peut alors écrire :

$$\begin{aligned}
\Sigma_{\tilde{\mathbf{z}}_{k|k-1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}} &= \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} \tilde{\mathbf{z}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] - \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}] \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] \\
&= \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} \tilde{\mathbf{z}}_{k|k-1}^t | \mathbf{I}_{0:k}] - \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k}] \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1}^t | \mathbf{I}_{0:k}] \\
&= \mathbb{E}[(H_k \tilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k)(H_k \tilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k)^t | \mathbf{I}_{0:k}] \\
&\quad - \mathbb{E}[H_k \tilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k | \mathbf{I}_{0:k}] \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t H_k^t + \mathbf{v}_k^t | \mathbf{I}_{0:k}] \\
&= H_k \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1} \tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] H_k^t + \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^t | \mathbf{I}_{0:k}] \\
&\quad - H_k \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}] \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] H_k^t \\
&\quad - \mathbb{E}[\mathbf{v}_k | \mathbf{I}_{0:k}] \mathbb{E}[\mathbf{v}_k^t | \mathbf{I}_{0:k}] \\
&= H_k \Sigma_{k|k-1} H_k^t + R_k
\end{aligned} \tag{C.15}$$

- Calcul de $\Sigma_{\mathbf{x}_{k+1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}}$

$$\begin{aligned}
\Sigma_{\mathbf{x}_{k+1}, \tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}} &= \mathbb{E}[\mathbf{x}_{k+1} \tilde{\mathbf{z}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] - \mathbb{E}[\mathbf{x}_{k+1} | \mathbf{I}_{0:k+1}] \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] \\
&= \mathbb{E}[(F_{k+1} \mathbf{x}_k + \mathbf{b}_{k+1} + \mathbf{w}_{k+1})(H_k \tilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k)^t | \mathbf{I}_{0:k+1}] \\
&\quad - \mathbb{E}[F_{k+1} \mathbf{x}_k + \mathbf{b}_{k+1} + \mathbf{w}_{k+1} | \mathbf{I}_{0:k+1}] \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t H_k^t + \mathbf{v}_k^t | \mathbf{I}_{0:k+1}] \\
&= F_{k+1} \mathbb{E}[\mathbf{x}_k \tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] H_k^t + \mathbf{b}_{k+1} \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] H_k^t \\
&\quad - F_{k+1} \mathbb{E}[\mathbf{x}_k | \mathbf{I}_{0:k+1}] \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] H_k^t \\
&\quad - \mathbf{b}_{k+1} \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k+1}] H_k^t \\
&= F_{k+1} \mathbb{E}[\mathbf{x}_k \tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] H_k^t + \mathbf{b}_{k+1} \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] H_k^t \\
&\quad - F_{k+1} \mathbb{E}[\mathbf{x}_k | \mathbf{I}_{0:k}] \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] H_k^t \\
&\quad - \mathbf{b}_{k+1} \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] H_k^t \\
&= F_{k+1} \Sigma_{\mathbf{x}_k, \tilde{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}} H_k^t \\
&= F_{k+1} (\Sigma_{k|k-1} + \Sigma_{\hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}}) H_k^t \\
&= F_{k+1} \Sigma_{k|k-1} H_k^t
\end{aligned} \tag{C.16}$$

Le passage à la dernière ligne se fait en remarquant que :

$$\begin{aligned}
 \Sigma_{\hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}} &= \mathbb{E}[\hat{\mathbf{x}}_{k|k-1} \tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] - \mathbb{E}[\hat{\mathbf{x}}_{k|k-1} | \mathbf{I}_{0:k}] \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] \\
 &= \mathbb{E}[\hat{\mathbf{x}}_{k|k-1} \tilde{\mathbf{x}}_{k|k-1}^t | \mathbf{I}_{0:k}] \\
 &= 0
 \end{aligned} \tag{C.17}$$

Ce résultat vient du principe d'orthogonalité appliqué à l'estimateur linéaire conditionnel de variance minimale (propriété B.3.5).

- Calcul de $\tilde{\mathbf{z}}_{k|k-1} - \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}]$

$$\begin{aligned}
 \tilde{\mathbf{z}}_{k|k-1} - \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k+1}] &= \tilde{\mathbf{z}}_{k|k-1} - \mathbb{E}[\tilde{\mathbf{z}}_{k|k-1} | \mathbf{I}_{0:k}] \\
 &= \tilde{\mathbf{z}}_{k|k-1} \\
 &= \mathbf{z}_k - H_k \hat{\mathbf{x}}_{k|k-1}
 \end{aligned} \tag{C.18}$$

- Expression récursive de $\hat{\mathbf{x}}_{k+1|k}$

À partir des expressions (C.12) à (C.18), on peut écrire l'expression récursive de $\hat{\mathbf{x}}_{k+1|k}$:

$$\begin{aligned}
 \hat{\mathbf{x}}_{k+1|k} &= F_{k+1} \hat{\mathbf{x}}_{k|k-1} + \mathbf{b}_{k+1} + (F_{k+1} \Sigma_{k|k-1} H_k^t) \\
 &\quad (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1} (\mathbf{z}_k - H_k \hat{\mathbf{x}}_{k|k-1})
 \end{aligned} \tag{C.19}$$

On définit la matrice $\tilde{K}_k$ à l'aide de la matrice de gain du filtre de Kalman K_k :

$$K_k = \Sigma_{k|k-1} H_k^t (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1} \tag{C.20}$$

$$\tilde{K}_k = F_{k+1} K_k \tag{C.21}$$

C.4 Expression récursive de la matrice de covariance conditionnelle de l'erreur d'estimation

Comme on peut écrire l'expression de $\tilde{\mathbf{x}}_{k+1|k}$ à partir de :

$$\begin{aligned}
 \tilde{\mathbf{x}}_{k+1|k} &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k} \\
 &= F_{k+1} \mathbf{x}_k + \mathbf{b}_{k+1} + \mathbf{w}_{k+1} - F_{k+1} \hat{\mathbf{x}}_{k|k-1} - \mathbf{b}_{k+1} - \tilde{K}_k (\mathbf{z}_k - H_k \hat{\mathbf{x}}_{k|k-1}) \\
 &= F_{k+1} \tilde{\mathbf{x}}_{k|k-1} + \mathbf{w}_{k+1} - \tilde{K}_k (H_k \tilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k) \\
 &= (F_{k+1} - \tilde{K}_k H_k) \tilde{\mathbf{x}}_{k|k-1} + \mathbf{w}_{k+1} - \tilde{K}_k \mathbf{v}_k,
 \end{aligned} \tag{C.22}$$

une expression récursive de covariance conditionnelle de l'erreur d'estimation $\Sigma_{k+1|k}$ peut être obtenue :

$$\begin{aligned}
\Sigma_{k+1|k} &\triangleq \Sigma_{\tilde{\mathbf{x}}_{k+1|k}, \tilde{\mathbf{x}}_{k+1|k} | \mathbf{I}_{0:k+1}} \\
&= \mathbb{E}[\tilde{\mathbf{x}}_{k+1|k} \tilde{\mathbf{x}}_{k+1|k}^t | \mathbf{I}_{0:k+1}] - \mathbb{E}[\tilde{\mathbf{x}}_{k+1|k} | \mathbf{I}_{0:k+1}] \mathbb{E}[\tilde{\mathbf{x}}_{k+1|k}^t | \mathbf{I}_{0:k+1}] \\
&= \mathbb{E}[(F_{k+1} - \tilde{K}_k H_k) \tilde{\mathbf{x}}_{k|k-1} + \mathbf{w}_{k+1} - \tilde{K}_k \mathbf{v}_k \\
&\quad ((F_{k+1} - \tilde{K}_k H_k) \tilde{\mathbf{x}}_{k|k-1} + \mathbf{w}_{k+1} - \tilde{K}_k \mathbf{v}_k)^t | \mathbf{I}_{0:k+1}] \\
&\quad - \mathbb{E}[(F_{k+1} - \tilde{K}_k H_k) \tilde{\mathbf{x}}_{k|k-1} + \mathbf{w}_{k+1} - \tilde{K}_k \mathbf{v}_k | \mathbf{I}_{0:k+1}] \\
&\quad \mathbb{E}[\tilde{\mathbf{x}}_{k|k-1}^t (F_{k+1} - \tilde{K}_k H_k)^t + \mathbf{w}_{k+1}^t - \mathbf{v}_k \tilde{K}_k^t | \mathbf{I}_{0:k+1}] \\
&= (F_{k+1} - \tilde{K}_k H_k) \Sigma_{k|k-1} (F_{k+1} - \tilde{K}_k H_k)^t + \Sigma_{\mathbf{w}_{k+1}, \mathbf{w}_{k+1} | \mathbf{I}_{0:k}} \\
&\quad - \tilde{K}_k \Sigma_{\mathbf{v}_k, \mathbf{v}_k | \mathbf{I}_{0:k}} \tilde{K}_k^t \\
&= (F_{k+1} - \tilde{K}_k H_k) \Sigma_{k|k-1} (F_{k+1} - \tilde{K}_k H_k)^t + Q_{k+1} + \tilde{K}_k R_k \tilde{K}_k^t \quad (\text{C.23})
\end{aligned}$$

C.5 Récapitulatif : équations du filtre linéaire conditionnel

On considère le modèle linéaire à bruits additifs décrits par le système (C.1). Sous les conditions que les paramètres du système peuvent dépendre de la variable aléatoire $\mathbf{I}_{0:k}$ qui désigne la séquence d'images jusqu'au temps k , un estimateur $\hat{\mathbf{x}}_{k+1|k}$ du premier moment d'une approximation gaussienne de $p(\mathbf{x}_{k+1} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k+1})$ peut être obtenu par les équations du **filtre linéaire conditionnel** (FLC). Cet estimateur correspond au meilleur estimateur de l'état $\mathbf{x}_{k+1}$, linéaire en $\mathbf{z}_{1:k}$, conditionnellement à $\mathbf{I}_{0:k}$. Il est appelé l'**estimateur linéaire conditionnel de variance minimale**.

Un calcul itératif de $\hat{\mathbf{x}}_{k+1|k}$ est défini par les équations récursives suivantes, qui constituent l'algorithme FLC :

$$\hat{\mathbf{x}}_{k+1|k} = F_{k+1|\mathbf{I}} \hat{\mathbf{x}}_{k|k-1} + \mathbf{b}_{k+1|\mathbf{I}} + \tilde{K}_k (\mathbf{z}_k - H_{k|\mathbf{I}} \hat{\mathbf{x}}_{k|k-1}) \quad (\text{C.24})$$

$$\tilde{K}_k = F_{k+1|\mathbf{I}} (\Sigma_{k|k-1} H_{k|\mathbf{I}}^t) (H_{k|\mathbf{I}} \Sigma_{k|k-1} H_{k|\mathbf{I}}^t + R_{k|\mathbf{I}})^{-1} \quad (\text{C.25})$$

$$\begin{aligned}
\Sigma_{k+1|k} &= (F_{k+1|\mathbf{I}} - \tilde{K}_k H_{k|\mathbf{I}}) \Sigma_{k|k-1} (F_{k+1|\mathbf{I}} - \tilde{K}_k H_{k|\mathbf{I}})^t \\
&\quad + Q_{k+1|\mathbf{I}} + \tilde{K}_k R_{k|\mathbf{I}} \tilde{K}_k^t \quad (\text{C.26})
\end{aligned}$$

On remarque que ces équations peuvent être écrites en deux étapes, ce qui permet de distinguer l'étape de prédiction de l'étape de correction (À l'initialisation on a $\hat{\mathbf{x}}_0 \rightsquigarrow \mathcal{N}(\hat{\mathbf{x}}_0; \bar{\mathbf{x}}_0, \Sigma_{0|0})$). On obtient ainsi une estimation des deux premiers moments d'une approximation gaussienne de $p(\mathbf{x}_{k+1} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k+1})$

- Étape de prédiction

$$\hat{\mathbf{x}}_{k|k-1} = F_{k|\mathbf{I}} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}_{k|\mathbf{I}} \quad (\text{C.27})$$

$$\Sigma_{k|k-1} = F_{k|\mathbf{I}} \Sigma_{k-1|k-1} F_{k|\mathbf{I}}^t + Q_{k|\mathbf{I}} \quad (\text{C.28})$$

- Étape de correction

$$K_k = \Sigma_{k|k-1} H_{k|\mathbf{I}}^t (H_{k|\mathbf{I}} \Sigma_{k|k-1} H_{k|\mathbf{I}}^t + R_{k|\mathbf{I}})^{-1} \quad (\text{C.29})$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k (\mathbf{z}_k - H_{k|\mathbf{I}} \hat{\mathbf{x}}_{k|k-1}) \quad (\text{C.30})$$

$$\Sigma_{k|k} = (Id - K_k H_{k|\mathbf{I}}) \Sigma_{k|k-1} \quad (\text{C.31})$$

Annexe D

Modèles conditionnels non linéaires gaussiens : deux cas particuliers

Dans cette annexe, on s'intéresse à des modèles conditionnels gaussiens, dont l'équation dynamique est non linéaire. Dans un premier temps, on associe à cette équation dynamique une équation de mesure linéaire. Dans un second temps, on étend ce modèle simple au cas où plusieurs mesures sont disponibles. Dans ces deux cas, on montre que ces modèles autorisent l'utilisation de la fonction d'importance optimale. Pour cela, on démontre les expressions de cette fonction $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k})$ utile pour faire évoluer les particules et de la distribution $p(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k})$ nécessaire au calcul récursif des poids. Enfin la construction de la fenêtre de validation est décrite dans le cas du modèle simple.

D.1 Vraisemblance monomodale

On s'intéresse au système suivant :

$$\begin{cases} p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k ; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}) \\ p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; H_{k|\mathbf{I}} \mathbf{x}_k, R_{k|\mathbf{I}}). \end{cases} \quad (\text{D.1})$$

Pour ce système, on donne les expressions des deux fonctions recherchées. Les démonstrations ne sont pas détaillées car ces résultats sont connus. En particulier, il sont disponibles dans le cas non conditionnel aux images dans [Doucet 00b].

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

En remarquant que :

$$p(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \int p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) d\mathbf{x}_k \quad (\text{D.2})$$

et en utilisant dans cette égalité les équations du modèle (D.1), on obtient :

$$p(\mathbf{z}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{z}_k ; H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I}} + H_{k|\mathbf{I}} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I}}^t) \quad (\text{D.3})$$

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

En remarquant que :

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

et en utilisant dans cette égalité les équations du modèle (D.1), et l'équation (D.3), on obtient :

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|\mathbf{I}}, \Sigma_{k|\mathbf{I}}) \quad (\text{D.5})$$

avec

$$\Sigma_{k|\mathbf{I}} = (Q_{k|\mathbf{I}}^{(i)-1} + H_{k|\mathbf{I}}^t R_{k|\mathbf{I}}^{-1} H_{k|\mathbf{I}})^{-1} \quad (\text{D.6})$$

$$\mathbf{m}_{k|\mathbf{I}} = \Sigma_{k|\mathbf{I}} (Q_{k|\mathbf{I}}^{(i)-1} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}) + H_{k|\mathbf{I}}^t R_{k|\mathbf{I}}^{-1} \mathbf{z}_k). \quad (\text{D.7})$$

D.2 Vraisemblance multimodale

On considère l'extension du modèle précédent au cas où plusieurs mesures sont disponibles. La vraisemblance étudiée est alors multimodale, et correspond à un mélange de lois gaussiennes (sous les hypothèses décrites dans le corps du document §5.3.2). Le système est le suivant :

$$\begin{cases} p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}), Q_{k|\mathbf{I}}) \\ p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j}; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j}). \end{cases} \quad (\text{D.8})$$

Ce modèle est très intéressant car il permet à la fois de considérer plusieurs observations simultanément, et d'autoriser l'utilisation de la fonction d'importance optimale, comme cela est montré ici.

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

Toujours en utilisant (D.2), on peut écrire :

$$\begin{aligned} p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) &= V_k^{1-M_k} \int \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j}; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j}) \mathcal{N}(\mathbf{x}_k; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), Q_{k|\mathbf{I}}^{(i)}) d\mathbf{x}_k \\ &= V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \int \mathcal{N}(\mathbf{z}_{k,j}; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j}) \mathcal{N}(\mathbf{x}_k; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), Q_{k|\mathbf{I}}^{(i)}) d\mathbf{x}_k \end{aligned} \quad (\text{D.9})$$

Par analogie avec (D.3), on en déduit :

$$\begin{aligned} p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) &= V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}\left(\mathbf{z}_{k,j}; H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I},j} + H_{k|\mathbf{I},j} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I},j}^t\right). \end{aligned} \quad (\text{D.10})$$

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

D'après (D.4), en utilisant les expressions du système (D.1), et l'équation (D.10) on peut écrire :

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) \\ = \frac{V_k^{1-M_k} \mathcal{N}(\mathbf{x}_k ; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), Q_{k|\mathbf{I}}^{(i)}) \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j})}{V_k^{1-M_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I},j} + H_{k|\mathbf{I},j} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I},j}^t)}. \end{aligned} \quad (\text{D.11})$$

On remplace le dénominateur par la constante S_k (S_k ne dépend pas de $\mathbf{x}_k$) telle que :

$$S_k = \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), R_{k|\mathbf{I},j} + H_{k|\mathbf{I},j} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I},j}^t) \quad (\text{D.12})$$

ce qui nous donne :

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) \\ = \frac{1}{S_k} \sum_{j=1}^{M_k} \beta_{k,j} \mathcal{N}(\mathbf{z}_{k,j} ; H_{k|\mathbf{I},j} \mathbf{x}_k, R_{k|\mathbf{I},j}) \mathcal{N}(\mathbf{x}_k ; f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}), Q_{k|\mathbf{I}}^{(i)}) \end{aligned} \quad (\text{D.13})$$

On pose :

$$\begin{aligned} N &= (\mathbf{x}_k - f_{k|\mathbf{I}}(\mathbf{x}_{k-1}))^t Q_{k|\mathbf{I}}^{-1} (\mathbf{x}_k - f_{k|\mathbf{I}}(\mathbf{x}_{k-1})) \\ &\quad + (\mathbf{z}_{k,j} - H_{k|\mathbf{I},j} \mathbf{x}_k)^t R_{k|\mathbf{I},j}^{-1} (\mathbf{z}_{k,j} - H_{k|\mathbf{I},j} \mathbf{x}_k) \end{aligned} \quad (\text{D.14})$$

En remarquant que :

$$\begin{aligned} N &= \mathbf{x}_k^t \left(Q_{k|\mathbf{I}}^{-1} + H_{k|\mathbf{I},j}^t R_{k|\mathbf{I},j}^{-1} H_{k|\mathbf{I},j} \right) \mathbf{x}_k \\ &\quad - 2 \mathbf{x}_k^t \left(Q_{k|\mathbf{I}}^{-1} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + H_{k|\mathbf{I},j}^t R_{k|\mathbf{I},j}^{-1} \mathbf{z}_{k,j} \right) \\ &\quad + (f_{k|\mathbf{I}}(\mathbf{x}_{k-1}))^t Q_{k|\mathbf{I}}^{-1} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + \mathbf{z}_{k,j}^t R_{k|\mathbf{I},j}^{-1} \mathbf{z}_{k,j} \end{aligned} \quad (\text{D.15})$$

$$= (\mathbf{x}_k - \mathbf{m}_{k|\mathbf{I},j})^t \Sigma_{k|\mathbf{I},j}^{-1} (\mathbf{x}_k - \mathbf{m}_{k|\mathbf{I},j}) + K \quad (\text{D.16})$$

où

$$\Sigma_{k|\mathbf{I},j} = \left(Q_{k|\mathbf{I}}^{-1} + H_{k|\mathbf{I},j}^t R_{k|\mathbf{I},j}^{-1} H_{k|\mathbf{I},j} \right)^{-1} \quad (\text{D.17})$$

$$\mathbf{m}_{k|\mathbf{I},j} = \Sigma_{k|\mathbf{I},j} \left(Q_{k|\mathbf{I}}^{-1} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + H_{k|\mathbf{I},j}^t R_{k|\mathbf{I},j}^{-1} \mathbf{z}_{k,j} \right) \quad (\text{D.18})$$

$$K = (f_{k|\mathbf{I}}(\mathbf{x}_{k-1}))^t Q_{k|\mathbf{I}}^{-1} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}) + \mathbf{z}_{k,j}^t R_{k|\mathbf{I},j}^{-1} \mathbf{z}_{k,j} + \mathbf{m}_{k|\mathbf{I},j}^t \Sigma_{k|\mathbf{I},j}^{-1} \mathbf{m}_{k|\mathbf{I},j}, \quad (\text{D.19})$$

on peut écrire l'expression de la fonction d'importance optimale :

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \sum_{j=1}^{M_k} \beta_{k,j} \frac{\alpha_{k,j}}{S_k} \mathcal{N}(\mathbf{x}_k ; \mathbf{m}_{k|\mathbf{I},j}, \Sigma_{k|\mathbf{I},j}) \quad (\text{D.20})$$

avec

$$\alpha_{k,j} = C \exp \left(-\frac{1}{2} \left[\|f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})\|_{Q_{k|\mathbf{I}}^{(i)}}^2 + \|\mathbf{z}_{k,j}\|_{R_{k|\mathbf{I},j}^{-1}}^2 - \|\mathbf{m}_{k|\mathbf{I},j}\|_{\Sigma_{k|\mathbf{I},j}^{-1}}^2 \right] \right) \quad (\text{D.21})$$

$$\text{et } C = (2\pi)^{-\frac{n_z}{2}} |\Sigma_{k|\mathbf{I},j}|^{\frac{1}{2}} |Q_{k|\mathbf{I}}^{(i)}|^{-\frac{1}{2}} |R_{k|\mathbf{I},j}|^{-\frac{1}{2}} \quad (\text{D.22})$$

On rappelle que n_z est la dimension de la mesure $\mathbf{z}_k$. Notons que cette expression est un mélange de lois gaussiennes.

• **Preuve que $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k})$ est un mélange de lois gaussiennes**

Pour prouver que l'expression (D.20) est un mélange de lois normales, il faut prouver que la somme des coefficients du mélange vaut 1, i.e. :

$$\sum_{j=1}^{M_k} \beta_{k,j} \frac{\alpha_{k,j}}{S_k} = 1 \quad (\text{D.23})$$

Par hypothèse, on a :

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

d'où :

$$\int \sum_{j=1}^{M_k} \beta_{k,j} \frac{\alpha_{k,j}}{S_k} \mathcal{N}(\mathbf{x}_k ; \mathbf{m}_{k|\mathbf{I},j}, \Sigma_{k|\mathbf{I},j}) d\mathbf{x}_k = 1 \quad (\text{D.24})$$

$$\sum_{j=1}^{M_k} \beta_{k,j} \frac{\alpha_{k,j}}{S_k} \int \mathcal{N}(\mathbf{x}_k ; \mathbf{m}_{k|\mathbf{I},j}, \Sigma_{k|\mathbf{I},j}) d\mathbf{x}_k = 1 \quad (\text{D.25})$$

On peut donc écrire :

$$\sum_{j=1}^{M_k} \beta_{k,j} \frac{\alpha_{k,j}}{S_k} = 1 \quad (\text{D.26})$$

D.3 Fenêtre de validation

En plus de permettre l'utilisation de la fonction d'importance optimale, les deux modèles (D.1) (D.8) présentent l'avantage de permettre une définition simple d'une fenêtre de validation. Nous nous limitons ici à présenter le cas du modèle construit avec la vraisemblance monomodale. La construction d'une telle fenêtre pour l'autre modèle est

une simple extension.

• **Cas du modèle construit avec la vraisemblance monomodale**

Dans le cas du modèle avec vraisemblance monomodale (D.1), il nous semble raisonnable de construire une fenêtre ellipsoïdale. Pour cela, il est nécessaire de construire des approximations des deux premiers moments de $p(\mathbf{z}_k | \mathbf{z}_{k-1}, \mathbf{I}_{0:k})$ notés $\mathbf{m}_k^{valid}$ et Σ_k^{valid} .

On suppose disposer d'une approximation discrète de $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1})$ sous la forme d'un nuage $\{\mathbf{x}_{k-1}^{(i)}, \tilde{w}_{k-1}^{(i)}\}_{i=1\dots N}$. On a l'expression suivante :

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) = \int p(\mathbf{z}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1}) d\mathbf{x}_{k-1} \quad (\text{D.27})$$

$$\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) \quad (\text{D.28})$$

En utilisant (D.3) et (D.28) une approximation de l'espérance de la distribution de la nouvelle observation est obtenue en remarquant que :

$$\begin{aligned} \mathbf{m}_k^{valid} = \mathbb{E}[\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] &= \int \mathbf{z}_k p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) d\mathbf{z}_k \\ &\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} \int \mathbf{z}_k p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) d\mathbf{z}_k \\ &\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}). \end{aligned} \quad (\text{D.29})$$

En ce qui concerne le deuxième moment Σ_k^{valid} , on a :

$$\Sigma_k^{valid} = \text{var}[\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] = \int \|\mathbf{z}_k - \mathbf{m}_k^{valid}\|^2 p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) d\mathbf{z}_k \quad (\text{D.30})$$

Cette expression n'est pas calculable directement. On utilise alors la relation valable quelque soit le vecteur $\mathbf{a}$:

$$\text{var}[\mathbf{a}] = \mathbb{E}[\|\mathbf{a}\|^2] - \|\mathbb{E}[\mathbf{a}]\|^2. \quad (\text{D.31})$$

En utilisant (D.3), on remarque que :

$$\mathbb{E}[\|\mathbf{z}_k\|^2 | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}] = \int \|\mathbf{z}_k\|^2 p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) d\mathbf{z}_k \quad (\text{D.32})$$

$$= \text{var}[\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}] + \|\mathbb{E}[\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}]\|^2 \quad (\text{D.33})$$

$$= R_{k|\mathbf{I}} + H_{k|\mathbf{I}} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I}}^t + \|H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})\|^2, \quad (\text{D.34})$$

et on peut écrire :

$$\mathbb{E}[\|\mathbf{z}_k\|^2 | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] = \int \|\mathbf{z}_k\|^2 p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) d\mathbf{z}_k \quad (\text{D.35})$$

$$= \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} \int \|\mathbf{z}_k\|^2 p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) d\mathbf{z}_k \quad (\text{D.36})$$

$$= \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} [R_{k|\mathbf{I}} + H_{k|\mathbf{I}} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I}}^t + \|H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})\|^2]. \quad (\text{D.37})$$

On en déduit une expression de $\Sigma_k^{valid} = \text{var}[\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}]$:

$$\Sigma_k^{valid} = \text{var}[\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] \quad (\text{D.38})$$

$$= \mathbb{E}[\|\mathbf{z}_k\|^2 | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] - \|\mathbb{E}[\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}]\|^2 \quad (\text{D.39})$$

$$= \mathbb{E}[\|\mathbf{z}_k\|^2 | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}] - \|\mathbf{m}_k^{valid}\|^2 \quad (\text{D.40})$$

$$\begin{aligned} &\simeq \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} [R_{k|\mathbf{I}} + H_{k|\mathbf{I}} Q_{k|\mathbf{I}}^{(i)} H_{k|\mathbf{I}}^t + \|H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)})\|^2] \\ &\quad - \left\| \sum_{i=1}^N \tilde{w}_{k-1}^{(i)} H_{k|\mathbf{I}} f_{k|\mathbf{I}}(\mathbf{x}_{k-1}^{(i)}) \right\|^2 \end{aligned} \quad (\text{D.41})$$

Liste des algorithmes

1	Filtre de Kalman pour la résolution du système (1.2)	23
2	Filtre de Kalman étendu pour la résolution du système (1.10)	26
3	Filtre de Kalman unscented pour la résolution du système (1.10)	27
4	Méthode numérique par maillage de l'espace d'état (approximation par somme de mesures de Dirac)	29
5	Algorithme d'acceptation rejet	36
6	Algorithme d'échantillonnage pondéré	37
7	Échantillonnage pondéré séquentiel dans le cadre bayésien	39
8	Échantillonnage pondéré séquentiel avec rééchantillonnage adaptatif	43
9	Filtre bootstrap	44
10	Filtre particulaire auxiliaire	50
11	Algorithme de rééchantillonnage multinomial, en $O(N \log(N))$	52
12	Algorithme de rééchantillonnage pour le filtre particulaire post-régularisé	55
13	Filtre particulaire rao-blackwellisé pour le système (2.57)	58
14	Filtre à association de données probabiliste	69
15	Filtre non linéaire conditionnel pour la résolution du système (5.38)	107
16	TFLC : algorithme de suivi d'un point appartenant au support du mouvement dominant	154
17	TFNLC : algorithme de trajectographie d'un point dont le mouvement ne peut être décrit que localement	156
18	Algorithme TFNLC multi-mesures : algorithme de suivi dédié aux situations d'ambiguïtés	158
19	Algorithme TFNLC-RB : méthode de suivi dédiée à la trajectographie d'un nuage de points d'un plan. Cette méthode repose sur un filtre non linéaire conditionnel rao-blackwellisé pour le système (8.7).	188

Liste des figures

0.1	Graphe de dépendance orienté de la chaîne de Markov cachée constituée des processus $\mathbf{x}_{0:k}$ et $\mathbf{z}_{1:k}$	17
2.1	Fonction de vraisemblance et densité prédite pour une observation précise (a) ou incohérente (b) par rapport à la prédiction (dans le cas d'un état directement observé i.e. $\mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k$) - figure tirée de [Oudjane 00].	45
2.2	Approximation de la loi de filtrage $p(\mathbf{x}_k \mathbf{z}_{1:k})$ par une somme de mesures de Dirac pondérées et approximation par une densité à noyaux - figure tirée de [Oudjane 00]	54
2.3	Graphe de dépendance orienté décrivant l'évolution temporelle de la chaîne de Markov cachée constituée des processus $\mathbf{r}_{0:k}$, $\mathbf{y}_{0:k}$ et $\mathbf{z}_{1:k}$	56
4.1	Illustration des deux éléments nécessaires à la construction d'un modèle : représentation et caractérisation de l'entité. Exemple de suivi d'une personne en imagerie couleur. La représentation choisie est une fenêtre englobante et la personne est caractérisée par un histogramme de couleur.	82
4.2	Exemples de représentations précises d'objets. (a) représentation d'un visage par un modèle 3d - figure tirée de [Lu 04]. (b) représentation d'un corps humain par une chaîne cinématique - figure tirée de [Deutscher 00]	83
4.3	Exemples de représentations simples d'objets. (a) représentation par un contour de Fourier - figure tirée de [Hue 01]. (b) représentation par une fenêtre rectangulaire - figure tirée de [Pérez 02]. (c) représentation par une ellipse - figure tirée de [Comaniciu 03]	85
4.4	Modèle de contour. La ligne blanche épaisse représente un contour hypothèse, les lignes fines sont les lignes sur lesquelles un algorithme de détection de contour est utilisé, l'ensemble des points noirs correspond à la mesure associée au contour hypothèse - figure tirée de [MacCormick 99]	87

4.5	Stratégie <i>Top-Down</i> – Exemple d’une représentation de l’entité suivie par fenêtre englobante, et d’une caractérisation par un histogramme de couleur – L’estimation précédente et l’histogramme de référence sont représentés en vert. (1) La première étape consiste à détecter la nouvelle observation (en rouge). L’observation correspond à la fenêtre qui minimise la distance entre son histogramme (en bleu) et l’histogramme de référence. (2) Puis, les nouvelles particules (en noir) sont prédites avec la fonction d’importance qui prend en compte éventuellement l’observation. (3) Enfin, le poids de chaque particule est évalué en utilisant la vraisemblance, qui modélise la distance entre sa position et la position de l’observation.	90
4.6	Stratégie <i>Bottom-Up</i> – Exemple d’une représentation de l’entité suivie par fenêtre englobante, et d’une caractérisation par un histogramme de couleur – L’estimation précédente et l’histogramme de référence sont représentés en vert. (1) La première étape consiste à prédire les nouvelles particules (en noir). L’observation correspond alors à l’ensemble des histogrammes associées aux particules (en rouge). (2) La deuxième étape évalue le poids de chaque particule au travers de la vraisemblance qui modélise la distance entre son propre histogramme et l’histogramme de référence.	92
5.1	Graphe de dépendance orienté de la chaîne de Markov cachée constituée des processus d’état $\mathbf{x}_{0:k}$ et de mesure $\mathbf{z}_{1:k}$	98
5.2	Graphe de dépendance orienté de la chaîne de Markov cachée constituée des processus d’état $\mathbf{x}_{0:k}$, de mesure $\mathbf{z}_{1:k}$ et de la séquence d’images $\mathbf{I}_{0:k}$	99
5.3	Diagramme représentant les étapes du filtre linéaire conditionnel pour la résolution du système (5.9). Les cadres doubles représentent les étapes réalisées sur le séquence d’images.	105
5.4	Construction de la fenêtre de validation pour le modèle (5.77). Exemple avec trois particules. À chaque particule $\mathbf{x}_{k-1}^{(i)}$ est associée une ellipse bleue qui correspond à $p(\mathbf{z}_k \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$. La fenêtre de validation “exacte” est la réunion de ces ellipses. Elle est approchée par une unique ellipse, représentée en rouge.	114
5.5	Construction de la fenêtre de validation pour le modèle (5.87). Exemple avec trois particules et deux mesures. À chaque particule $\mathbf{x}_{k-1}^{(i)}$ est associée une ellipse bleue qui correspond à $p(\mathbf{z}_{k,1} \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$, et donc à la première mesure ; et une ellipse jaune qui correspond à $p(\mathbf{z}_{k,2} \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$, et donc à la deuxième mesure. La fenêtre de validation “exacte” est la réunion de ces ellipses. Elle est approchée par deux ellipses, représentées en rouge. Ces deux ellipses rouges correspondent à une approximation de l’union des ellipses respectivement bleues et jaunes, et sont associées aux mesures respectivement 1 et 2.	116
5.6	Diagramme représentant les étapes du filtre non linéaire conditionnel pour la résolution des systèmes (5.52) (5.67). Les cadres doubles représentent les étapes réalisées sur le séquence d’images.	118
6.1	Principe de la mise en correspondance par une méthode de corrélation . . .	128

6.2	surfaces de SSD centrées à un point placé sur (a) un coin, (b) un contour, (c) une zone homogène bruitée. Ces figures illustrent le fait que les courbures de la SSDS autour de la position estimée (au centre de chaque courbe) rendent compte de la précision de l'estimation - figures extraites de [Nickels 98].	130
7.1	Mise en évidence de la pertinence du rapport entre la taille du support de validité et la taille du support d'estimation pour déterminer la qualité de l'estimation de mouvement. La séquence traitée est longue et comporte plusieurs changements de plans. L'estimation du mouvement est réalisée sur l'image complète. Chaque décrochage désigne une mauvaise estimation due à une coupure. Les images sont représentatives de la séquence avant et après cette coupure. Le test de Page-Hinkley permet de détecter le décrochage du signal.	143
7.2	Mise en évidence de la pertinence du rapport entre la taille du support de validité et la taille du support d'estimation et de l'utilisation du test de Page-Hinkley. La séquence utilisée représente un visage et l'estimation de mouvement est réalisée sur une fenêtre autour de celui-ci (cadre blanc). Chaque passage de face à profil et inversement entraîne une mauvaise estimation et le test de Page-Hinkley émet une alarme.	143
7.3	Estimation de la covariance du bruit de mesure (niveau du bruit fixé à $\sigma = 1$) (a) image en haut : caractérisation de référence et point à mettre en correspondance (coin du toit), image en bas : image courante et position estimée ; (b) SSDS correspondant au deux points mis en correspondance ; (c) distribution réponse associée. Dans cette situation, l'hypothèse de loi uniforme est rejetée. La distribution réponse est approchée par une loi gaussienne 2D, et la matrice de covariance est estimée par (7.13).	149
7.4	Estimation de la covariance du bruit de mesure dans une situation d'occlusion : mise en évidence de l'intérêt du test d'ajustement (niveau du bruit fixé à $\sigma = 1$) (a) image en haut : caractérisation de référence et point à mettre en correspondance (sommet du toit), image en bas : image courante et position estimée ; (b) SSDS correspondant au deux points mis en correspondance ; (c) distribution réponse associée. Dans cette situation, l'hypothèse de loi uniforme est acceptée.	149
7.5	Estimation de la covariance du bruit de mesure dans le cas de séquences bruitées : mise en évidence de l'intérêt des tests de résidus. (a) image en haut : caractérisation de référence et point à mettre en correspondance, image en bas : image courante et position estimée ; (b) SSDS correspondant au deux points mis en correspondance ; (c) la distribution réponse associée, correspondant à $\sigma = 1$ est approchée par une loi gaussienne, ce qui ne retranscrit pas l'incertitude observée ; (d) SSDS modifiée après les tests de résidus avec $\sigma = 5$; (e) la distribution réponse résultante est bien approchée par une loi uniforme.	150
7.6	Procédure de mise à jour de la caractérisation de référence	152
7.7	Séquence Hangars – Trajectoires complètes estimées par les algorithmes TFLC, KLT et TFNLC	160

7.8	Séquence Hangars – Comparaison entre les trajectoires estimées par les différents algorithmes et la vérité terrain. Les trajectoires TFNLC correspondent à des trajectoires moyennes sur 100 réalisations	161
7.9	Séquence Corridor – Trajectoires complètes estimées par les algorithmes TFLC, MDR, KLT et TFNLC	162
7.10	Séquence Corridor – Comparaison entre les trajectoires estimées par les différents algorithmes et la vérité terrain. Les trajectoires TFNLC correspondent à des trajectoires moyennes sur 100 réalisations	162
7.11	Séquence Bâtiment – Résultat du suivi par l’algorithme TFLC	163
7.12	Séquence Mobile – (a) points initiaux $\neq 0$ (b) carte de segmentation au sens du mouvement (c) trajectoires du TFNLC	164
7.13	Séquence Mobile – Comparaison entre les trajectoires estimées par les différents algorithmes et la vérité terrain. La trajectoire TFNLC correspond à une trajectoire moyenne sur les réalisations correctes (2 échecs sur 100 pour la balle blanche, 3 sur 100 pour la noire)	164
7.14	Séquence Mobile – Résultats de suivi des algorithmes TFNLC, MRD et KLT	165
7.15	Séquence Examen – Résultats du suivi par les algorithmes TFNLC et KLT	166
7.16	Séquence Fluide – Trajectoires estimées par le TFNLC	167
7.17	Séquence Minitel – Évolution de la caractérisation de référence du point 1	167
7.18	Séquence Minitel – Résultat du suivi par l’algorithme TFLC	168
7.19	Séquence Van – Résultat du suivi par l’algorithme TFLC	168
7.20	Séquence Jardin – Résultat des algorithmes TFLC (en rouge) et TFNLC (en vert)	169
7.21	Séquence Jardin – (a) Zone de l’image suivant un mouvement secondaire (b) Comparaison entre la trajectoire estimée du point 1 par le TFNLC et la vérité terrain. La trajectoire TFNLC correspond à une trajectoire moyenne sur les réalisations correctes (15 échecs sur 100 essais).	169
7.22	Séquence Jardin – Illustration de l’importance de la fonction optimale et de la région de validation – L’estimée correspond à la croix verte, l’observation à la croix bleue et la région de validation à l’ellipse bleue.	170
7.23	Séquence Main – Résultat du TFNLC, avec une observation par point suivi	172
7.24	Séquence Main – Résultat du suivi du point situé entre l’index et le majeur fourni par le TFNLC, en considérant une unique observation. Cette observation est représentée en bleu, la position estimée en rouge et les particules en jaune (100 particules sont utilisées).	172
7.25	Séquence Main – Résultat de l’algorithme TFNLC multi-mesures, avec trois observations par point suivi	173
7.26	Séquence Main – Résultat du suivi du point situé entre l’index et le majeur fourni par le TFNLC multi-mesures, en considérant trois observations. La mesure principale (qui correspond au plus haut pic de corrélation) est représentée en bleu, les deux mesures secondaires en vert, la position estimée en rouge et les particules en jaune (100 particules sont utilisées).	173
7.27	Séquence Bastille – TFNLC, avec une observation par point suivi	174
7.28	Séquence Bastille – TFNLC multi-mesures, avec trois observations par point suivi	174
7.29	Séquence Foot – TFNLC, avec trois observations par point suivi	175

7.30	Résultat de la détection de visages	176
7.31	Résultat de la détection de visages - cas d'une fausse alarme, et d'une non détection d'un visage de profil	177
7.32	Résultat de l'algorithme de suivi de visage - Illustration de la robustesse au passage de profil	180
7.33	Résultat de l'algorithme de suivi de visage - Illustration de la robustesse au flou de bougé et aux occultations partielles	181
7.34	Résultat de l'algorithme de suivi de visage - Illustration de la création de la robustesse aux changements d'expressions	181
8.1	Séquence BureauOcclusion – Résultats obtenus par le TFNLC-RB ; les points de référence sont en rouge, et les points attachés en vert.	189
8.2	Séquence BureauHoule – Résultats obtenus par le TFNLC-RB pour le suivi de deux objets. La situation où une partie de l'objet est hors de l'image est traitée comme une situation d'occlusion. Les points de référence sont en rouge, et les points attachés en vert.	190
8.3	Séquence Visual – Résultats obtenus par le TFNLC-RB. Les points de réf- érence sont en rouge, et les points attachés en vert.	191
8.4	Séquence Campus – Résultats obtenus par le TFNLC-RB. Les points de réf- érence sont en rouge, et les points attachés en vert.	192

Bibliographie

- [Adrian 91] R. Adrian. – Particle imaging techniques for experimental fluid mechanics. *Annual Review on Fluid Mechanics*, 23 :261–304, 1991.
- [Alspach 72] D.L. Alspach, H.W. Sorenson. – Non-linear Bayesian estimation using Gaussian sum approximation. *IEEE Transactions on Automatic Control*, 17 :439–447, 1972.
- [Anandan 87] P. Anandan. – *Measuring visual motion from image sequences*. – Rapport de recherche, University of Massachusetts, 1987.
- [Anandan 89] P. Anandan. – A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2(3) :283–310, 1989.
- [Anderson 79] B.D.O. Anderson, J.B. Moore. – *Optimal Filtering*. – Englewood Cliffs, NJ : Prentice Hall, 1979.
- [Andrieu 99] C. Andrieu, N. de Freitas, A. Doucet. – *Sequential Bayesian estimation and model selection applied to neural networks*. – Rapport de Recherche nCUER/F-INFENG/TR 341, Cambridge University, Engineering department, 1999.
- [Andrieu 01] C. Andrieu, A. Doucet, E. Punskeya. – Sequential Monte Carlo methods for optimal filtering. [*Doucet 01a*], chap. 4. – Springer-Verlag, 2001.
- [Andrieu 02] C. Andrieu, A. Doucet. – Particle filtering for partially observed Gaussian state space models. *Journal of the Royal Statistical Society, Series B*, 2002.
- [Andrieu 03] C. Andrieu, N. de Freitas, A. Doucet, M. Jordan. – An introduction to mcmc for machine learning. *Machine Learning*, 50(1-2) :5–43, 2003.
- [Arnaud 04a] E. Arnaud, E. Mémmin. – Optimal importance sampling for tracking in image sequences : application to point tracking. – *European Conference on Computer Vision*, 2004.
- [Arnaud 04b] E. Arnaud, E. Mémmin, B. Cernuschi-Frías. – Conditional filters for image sequence based tracking - application to point tracking. *IEEE Transactions of Image Processing*, 2004.
- [Arnaud 04c] E. Arnaud, E. Mémmin, B. Cernuschi-Frías. – A robust stochastic filter for point tracking in image sequences. – *Asian Conference on Computer Vision*, 2004.

- [Arulampalam 02] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp. – A tutorial on particles filters for online nonlinear / non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2) :174–188, 2002.
- [Aschwanden 92] P. Aschwanden, W. Guggenbül. – Experimental results from a comparative study on correlation type registration algorithms. *Robust computer vision : Quality of Vision Algorithms*, éd. par Förstner, Ruwiedel, pp. 268–282. – Wichmann, 1992.
- [Baker 03a] S. Baker, R. Gross, I. Matthews. – *Lucas-Kanade 20 Years On : A Unifying Framework : Part 3.* – Rapport de Recherche nCMU-RI-TR-03-35, Robotics Institute, Carnegie Mellon University, 2003.
- [Baker 03b] S. Baker, R. Gross, I. Matthews, T. Ishikawa. – *Lucas-Kanade 20 Years On : A Unifying Framework : Part 2.* – Rapport de Recherche nCMU-RI-TR-03-01, Robotics Institute, Carnegie Mellon University, 2003.
- [Baker 04a] S. Baker, R. Gross, I. Matthews. – *Lucas-Kanade 20 Years On : A Unifying Framework : Part 4.* – Rapport de Recherche nCMU-RI-TR-04-14, Robotics Institute, Carnegie Mellon University, 2004.
- [Baker 04b] S. Baker, I. Matthews. – Lucas-Kanade 20 years on : a unifying framework. *International Journal of Computer Vision*, 53(3) :221–255, 2004.
- [Barron 94] J. Barron, D. Fleet, S. Beauchemin. – Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1) :43–77, 1994.
- [BarShalom 75] Y. Bar-Shalom, E. Tse. – Tracking in a cluttered environment with probabilistic data association. *Automatica*, 11(451-460), 1975.
- [BarShalom 88] Y. Bar-Shalom, T.E. Fortmann. – *Tracking and Data Association.* – Academic Press, 1988.
- [BarShalom 95] Y. Bar-Shalom, X.-R. Li. – *Multitarget-multisensor tracking : principles and techniques.* – Y. Bar-Shalom, 1995.
- [Basseville 88] M. Basseville. – Detecting changes in signals and systems - a survey. *Automatica*, 24(3) :309–326, 1988.
- [Benhimane 04] S. Benhimane, E. Malis. – Real-time image-based tracking of planes using efficient second-order minimization. – *International Conference on Intelligent Robots Systems*, 2004.
- [Bergman 99] N. Bergman. – *Recursive Bayesian estimation, navigation and tracking estimation.* – PhD. Thesis, Linköping University, Sweden, 1999.
- [Berzuini 97] C. Berzuini, N. Best, W. Gilks, C. Larizza. – Dynamic conditional independence models and Markov chain Monte Carlo methods. *Journal of the American Statistical Association*, 92(440) :1403–1412, 1997.

- [Berzuini 01] C. Berzuini, W. Gilks. – RESAMPLE-MOVE filtering with cross-model jumps. [*Doucet 01a*], chap. 6. – Springer-Verlag, 2001.
- [Black 96] M. Black, A. Jepson. – EigenTracking : Robust matching and tracking of articulated objects using a view-based representation. – *European Conference on Computer Vision*, pp. 329–342, 1996.
- [Black 99] M. Black, D. Fleet. – Probabilistic detection and tracking of motion discontinuities. – *International Conference on Computer Vision*, vol. 2, pp. 551–558, 1999.
- [Blake 98] A. Blake, M. Isard. – *Active contours*. – Springer-Verlag, 1998.
- [Blake 99] A. Blake, B. North, M. Isard. – Learning multi-class dynamics. *Advances in Neural Information Processing Systems*, 1999.
- [Boucher 03] T. Boucher. – Suivi de visages dans des vidéos par couplage entre un détecteur de visages et un algorithme de suivi de points. – Rapport de stage, École Polytechnique, 2003.
- [Breidt 00] F.J. Breidt, A.L. Carriquiry. – Highest density gates for target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 2000.
- [Bréhard 04] T. Bréhard, J.-P. Le Cadre. – A new approach for the bearings-only problem : estimation of the variance-to-range ratio. – *International Conference on Information Fusion*, 2004.
- [CahnvonSeelen 97] U. Cahn von Seelen, R. Bajcsy. – Adaptive correlation tracking of targets with changing scale. *Reconnaissance, Surveillance, and Target Acquisition for the Unmanned Ground Vehicle*, éd. par Thomas M. Strat Oscar Firschein, pp. 313–322. – Morgan Kaufmann, 1997.
- [Cai 95] Z. Cai, F. Le Gland, H. Zang. – *An Adaptive Local Grid Refinement Method for Nonlinear Filtering*. – Rapport de Recherche n RR-2679, INRIA Institut National de Recherche en Informatique et Automatique, 1995.
- [Carpenter 99] J. Carpenter, P. Clifford, P. Fearnhead. – An improved particle filter for non-linear problems. *IEE Proceedings - F : Radar, Sonar and Navigation*, 146 :2–7, 1999.
- [Casella 96] G. Casella, C. Robert. – Rao-Blackwellisation of sampling schemes. *Biometrika*, 83(1) :81–94, 1996.
- [Celeux 04] G. Celeux, J. Nascimento, J. Marques. – Learning switching dynamic models for objects tracking. *Pattern Recognition*, 37 :1841–1853, 2004.
- [Chambon 02] S. Chambon, A. Crouzil. – *Évaluation et comparaison de mesures de corrélation robustes*. – Rapport de Recherche n2002-34-R, IRIT Institut de Recherche en Informatique de Toulouse, 2002.
- [Chateau 04] T. Chateau, F. Jurie, R. Marc, M. Dhome. – Suivi et reconnaissance de gestes par vision monoculaire en temps réel : application à la formation des chargés de manoeuvres pour la conduite des ponts

- polaires. – *Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, 2004.
- [Chen 00a] R. Chen, J.S. Liu. – Mixture kalman filters. *Journal of the Royal Statistical Society, Serie B*, 62(3) :493–508, 2000.
- [Chen 00b] R. Chen, X. Wang, J.S. Liu. – Adaptative joint detection and decoding in flat-fading channels via mixture Kalman filtering. *IEEE Transactions on Information Theory*, 46(6) :2079–2094, 2000.
- [Chen 03] Y. Chen, J.S. Liu. – *Sequential importance sampling with resampling in molecular population genetics*. – Discussion paper 03-09, Institute of Statistics and Decision Sciences, Duke University., 2003.
- [Chen 04] Y. Chen, Y. Rui. – Speaker tracking using particle filter sensor fusion. – *Asian Conference on Computer Vision*, vol. 2, pp. 782–787, 2004.
- [Choo 01] K. Choo, D. Fleet. – People tracking using hybrid Monte Carlo filtering. – *International Conference on Computer Vision*, vol. 2, pp. 321–328, 2001.
- [Clapp 00] T. Clapp. – *Statistical methods for the processing of communication data*. – PhD. Thesis, University of Cambridge, Signal Processing Group, 2000.
- [Comaniciu 03] D. Comaniciu, V. Ramesh, P. Meer. – Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5) :564–574, 2003.
- [Crisan 99] D. Crisan, M. Grunwald. – *Large deviation comparison of branching algorithms versus resampling algorithms : application to discrete time atochastic filtering*. – Rapport de recherche, Cambridge University, Statistical laboratory, 1999.
- [Crisan 01] D. Crisan. – Particle filters - a theoretical perspective. [*Doucet 01a*], chap. 2. – Springer-Verlag, 2001.
- [Crisan 02] D. Crisan, A. Doucet. – Survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50(3) :736–746, 2002.
- [Del Moral 96] P. Del Moral. – Non linear filtering : Interacting particle solution. *Markov Processes and Related Fields*, 2(4) :555–580, 1996.
- [Del Moral 00] P. Del Moral, L. Miclo. – Branching and interacting particle systems approximations of feynman-kac formulae with applications to non-linear filtering. *Séminaire de probabilités XXXIV*, éd. par J. Azéma, M. Emery, M. Ledoux, M. Yor. – Springer-Verlag Berlin, 2000.
- [Delakis 03] M. Delakis, C. Garcia. – Training convolutional filters for robust face detection. – *International Workshop of Neural Networks for Signal Processing*, pp. 739–748, 2003.
- [Deutscher 00] J. Deutscher, A. Blake, I. Reid. – Articulated body motion capture by annealed particle filtering. – *Conference on Computer Vision and Pattern Recognition*, 2000.

- [Dezert 01] J. Dezert. – Introduction au pistage multi-cibles multi-senseurs. – notes de cours ENSTA, 1998-2001.
- [Doucet 98] A. Doucet. – *On sequential simulation-based methods for Bayesian filtering*. – Rapport technique CUED/F-INFENG/TR.310, University of Cambridge, Engineering department, 1998.
- [Doucet 00a] A. Doucet, N. de Freitas, K. Murphy, S. Russell. – Rao-Blackwellised particle filtering for dynamic bayesian networks. – *Conference on Uncertainty in Artificial Intelligence*, 2000.
- [Doucet 00b] A. Doucet, S. Godsill, C. Andrieu. – On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3) :197–208, 2000.
- [Doucet 01a] A. Doucet, N. de Freitas, N. Gordon (édité par). – *Sequential Monte Carlo methods in practice*. – New York : Springer-Verlag, Series Statistics for Engineering and Information Science, 2001.
- [Doucet 01b] A. Doucet, N. Gordon, V. Krishnamurthy. – Particle filters for state estimation of jump Markov linear systems. *IEEE Transactions on signal processing*, 49(3) :613–624, 2001.
- [Doucet 04] A. Doucet, S. Sénécal. – Fixed-lag sequential Monte Carlo. – *European Signal Processing Conference*, 2004.
- [Duda 73] R. Duda, P. Hart. – *Pattern recognition and scene analysis*. – John Wiley and Sons Inc, 1973.
- [Faugeras 88] O. Faugeras, F. Lustman. – *Motion and structure from motion in a piecewise planar environment*. – Rapport de recherche, INRIA, Institut National de Recherche en Informatique et Automatique, 1988.
- [Fearnhead 98] P. Fearnhead. – *Sequential Monte Carlo methods in filter theory*. – Phd. thesis, Merton College, University of Oxford, England, 1998.
- [Fei 04] H. Fei, I. Reid. – Joint Bayes filter : a hybrid tracker for non-rigid hand motion recognition. – *European Conference on Computer Vision*, pp. 497–508, 2004.
- [Fox 01] D. Fox, S. Thrun and W. Burgard, F. Dellaert. – Particle filters for mobile robot localization. *[Doucet 01a]*, chap. 19. – Springer-Verlag, 2001.
- [Freitas 01] N. de Freitas, C. Andrieu, P. Højen Sørensen, M. Niranjan, A. Gee. – Sequential Monte Carlo methods for neural networks. *[Doucet 01a]*, chap. 17. – Springer-Verlag, 2001.
- [Freitas 02] N. de Freitas. – Rao-Blackwellised particle filtering for fault diagnosis. – *IEEE Aerospace Conference*, 2002.
- [Freitas 04] N. de Freitas, R. Dearden, F. Hutter, R. Morales-Menéndez, J. Mutch, D. Poole. – Diagnosis by a waiter and a Mars explorer. *[Haykin 04]*. – IEEE, 2004.

- [Garcia 02] C. Garcia, M. Delakis. – A neural architecture for fast and robust face detection. – *International Conference on Pattern Recognition*, vol. 2, pp. 40–43, 2002.
- [Gelb 74] A. Gelb (édité par). – *Applied Optimal Estimation*. – MIT press, Cambridge Mass, 1974.
- [Gelgon 98] M. Gelgon. – *Segmentation spatio-temporelle et suivi dans une séquence d'images : application à la structuration et à l'indexation de vidéo*. – PhD. Thesis, Université de Rennes 1, 1998.
- [Georgescu 04] B. Georgescu, P. Meer. – Point matching under large image deformations and illumination changes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6), 2004.
- [Giachetti 00] A. Giachetti. – Matching techniques to computer image motion. *Image and Vision Computing*, 18(3) :247–260, 2000.
- [Gilks 01] W.R. Gilks, C. Berzuini. – Following a moving target - Monte Carlo inference for dynamic bayesian models. *Journal of the Royal Statistical Society - Series B Statistical Methodology*, 63(1) :127–146, 2001.
- [Godsill 01] S. Godsill, T. Clapp. – Improvement strategies for Monte Carlo particle filters. [*Doucet 01a*], chap. 7. – Springer-Verlag, 2001.
- [Gordon 93] N. Gordon, D. Salmond, A. Smith. – Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, 140(2) :107–113, 1993.
- [Gordon 95] N. Gordon, D. Salmond, C. Ewing. – Bayesian state estimation for tracking and guidance using the Bootstrap filter. *Journal of Guidance Control and Dynamics*, 18(6) :1434–1443, 1995.
- [Gordon 97] N. Gordon. – A hybrid Bootstrap filter for target tracking in clutter. *IEEE Transactions in Aerospace and Electronic Systems*, 33(1) :353–358, 1997.
- [Gouiffès 04] M. Gouiffès, C. Collewet, C. Fernandez-Maloigne, A. Trémeau. – Suivi de points : extension aux objets non-lambertiens. – *Journées des Jeunes Chercheurs en Robotique*, 2004.
- [Gustafsson 02] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, P-J Nordlund. – Particle filters for positioning, navigation and tracking. *IEEE Transactions on Signal Processing*, 50(2) :425–437, 2002.
- [Hager 98] G. Hager, P. Belhumeur. – Efficient region tracking with parametric models of geometry and illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(10) :1025–1039, 1998.
- [Handschin 70] J. Handschin. – Monte Carlo techniques for prediction and filtering of non-linear stochastic processes. *Automatica*, 6 :555–563, 1970.
- [Harris 88] C. Harris, M. Stephens. – A combined corner detection and edge detector. – *Proceedings of the Alvey Vision Conference*, pp. 147–151, 1988.

- [Hastings 70] W.K. Hastings. – Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 1970.
- [Haykin 04] S. Haykin, N. de Freitas (édité par). – *special issue on sequential state estimation : from Kalman filters to particle filters*. – 2004.
- [Hinkley 71] D. Hinkley. – Inference about the change-point from cumulative sum tests. *Biometrika*, 58(509-523), 1971.
- [Horn 81] B. Horn, B. Schunck. – Determining optical flow. *Artificial Intelligence*, 17 :185–203, 1981.
- [Hu 94] X. Hu, N. Ahuja. – Feature extraction and matching as signal detection. *International Journal of Pattern Recognition and Artificial Intelligence*, 8(6) :1343–1379, 1994.
- [Hue 01] C. Hue, J.-P. Le Cadre, P. Pérez. – A particle filter to track multiple objects. – *IEEE Workshop on Multi-Object Tracking*, pp. 61–68, 2001.
- [Hue 02] C. Hue, J.-P. Le Cadre, P. Pérez. – Tracking multiple objects with particle filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 38(3) :791–812, 2002.
- [Hue 03] C. Hue. – *Méthodes séquentielles de Monte Carlo pour le filtrage non linéaire multi-objets dans un environnement bruité. Applications au pistage multi-cibles et à la trajectographie d'entités dans des séquences d'images 2D*. – Thèse de doctorat, Université de Rennes 1, 2003.
- [Hürzeler 98] M. Hürzeler, H.R. Künsch. – Monte carlo approximations for general state space models. *Journal of Computational and Graphical Statistics*, 7(2) :175–193, 1998.
- [Iba 01] Y. Iba. – Population Monte Carlo algorithms. *Transactions for the Japanese Society for Artificial Intelligence*, 16(2) :279–286, 2001.
- [Immerkær 96] J. Immerkær. – Fast noise variance estimation. *Computer Vision and Image Understanding*, 64(2), 1996.
- [Irani 94] M. Irani, B. Rousso, S. Peleg. – Computing occluding and transparent motions. *International Journal of Computer Vision*, 12(1) :5–16, 1994.
- [Isard 96] M. Isard, A. Blake. – Contour tracking by stochastic propagation of conditional density. – *European Conference on Computer Vision*, vol. 1, pp. 343–356, 1996.
- [Isard 98a] M. Isard, A. Blake. – CONDENSATION – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1) :5–28, 1998.
- [Isard 98b] M. Isard, A. Blake. – ICONDENSATION : unifying low-level and high-level tracking in a stochastic framework. – *European Conference on Computer Vision*, 1998.

- [Isard 98c] M. Isard, A. Blake. – A mixed-state CONDENSATION tracker with automatic model-switching. – *International Conference on Computer Vision*, pp. 107–112, 1998.
- [Jazwinski 70] A.M. Jazwinski. – *Stochastic processes and filtering theory*. – Academic Press, 1970.
- [Jepson 01] A. Jepson, D. Fleet, T. El-Maraghi. – Robust online appearance models for visual tracking. – *Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 415–422, 2001.
- [Jin 01] H. Jin, P. Favaro, S. Soatto. – Real-time feature tracking and outlier rejection with changes in illumination. – *International Conference on Computer Vision*, vol. 1, pp. 684–689, 2001.
- [Julier 97] S. Julier, J. Uhlmann. – A new extension of the Kalman filter to nonlinear systems. – *International Symposium Aerospace/Defense Sensing, Simulations and Controls*, 1997.
- [Jurie 02] F. Jurie, M. Dhome. – Hyperplane approximation for template matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7) :996–1000, 2002.
- [Kalman 60] R. Kalman. – A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(D) :35–45, 1960.
- [Kalman 61] R. Kalman, R. Bucy. – New results in linear filtering and prediction theory. *Transactions of the ASME–Journal of Basic Engineering*, 83(D) :93–108, 1961.
- [Kanatani 99] K. Kanatani, N. Ohta. – Accuracy bound and optimal computation of homography for image mosaicing applications. – *International Conference on Image Processing*, 1999.
- [Kanazawa 95] K. Kanazawa, D. Koller, S. Russell. – Stochastic simulation algorithms for dynamic probabilistic networks. – *Eleventh Conference on Uncertainty in Artificial Intelligence*, pp. 346–351, 1995.
- [Kang 03] J. Kang, I. Cohen, G. Medioni. – Continuous tracking within and across camera streams. – *Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 267–272, 2003.
- [Kermad 01] C. Kermad, C. Collewet. – Improving feature tracking by robust points of interest selection. – *Proceedings of the International Fall Workshop on Vision, Modeling, and Visualization*, 2001.
- [Kervrann 95] C. Kervrann. – *Modèles statistiques pour la segmentation et le suivi de structures déformables bidimensionnelles dans une séquence d’images*. – Thèse de doctorat, Université de Rennes 1, 1995.
- [Khan 04] Z. Khan, T. Balch, F. Dellaert. – A Rao-Blackwellized particle filter for eigentracking. – *Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 980–986, 2004.
- [Kirubarajan 04] T. Kirubarajan, Y. Bar-Shalom. – Probabilistic data association techniques for target tracking in clutter. [*Haykin 04*]. – IEEE, 2004.

- [Kitagawa 87] G. Kitagawa. – Non-Gaussian state-space modelling of non-stationary time series (with discussion). *Journal of the American Statistical Association*, 82 :1032–1063, 1987.
- [Kitagawa 96] G. Kitagawa. – Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5 :1–25, 1996.
- [Kong 94] A. Kong, J.S. Liu, W.H. Wong. – Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89 :278–288, 1994.
- [Kramer 88] S.C. Kramer, H.W. Sorenson. – Recursive Bayesian estimation using piece-wise constant approximations. *Automatica*, 24 :789–801, 1988.
- [Kwok 04] C. Kwok, D. Fox, M. Meilă. – Real-time particle filter. *Proceedings of the IEEE (issue on State Estimation)*, 92(2), 2004.
- [Kwolek 04] B. Kwolek. – Stereovision-based head tracking using color and ellipse fitting in a particle filter. – *European Conference on Computer Vision*, pp. 192–204, 2004.
- [Lan 97] Z.-D. Lan, R. Mohr. – *Robust Location based Partial Correlation*. – Rapport de recherche, INRIA Institut National de Recherche en Informatique et en Automatique, 1997.
- [Lefebvre 04] T. Lefebvre, H. Bruyninckx, J. De Schutter. – Kalman filters for nonlinear systems : a comparison of performance. *The International Journal of Control*, à paraître, 2004.
- [LeGland 03] F. Le Gland. – Filtrage particulaire. – *Actes du 19ème colloque GRETSI sur le traitement du signal et des images*, vol. 2, pp. 1–8, 2003.
- [Liu 95] J.S. Liu, R. Chen. – Blind deconvolution via sequential imputation. *Journal of the American Statistical Association*, 90 :567–576, 1995.
- [Liu 98a] J.S. Liu, R. Chen. – Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93 :1032–1044, 1998.
- [Liu 98b] J.S. Liu, R. Chen, W.H. Wong. – Rejection control and sequential importance sampling. *Journal of the American Statistical Association*, 93 :1022–1031, 1998.
- [Liu 04] T.L. Liu, H.T. Chen. – Real-time tracking using trust-region methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3) :397–402, 2004.
- [Lu 04] L. Lu, X.-T. Dai, G. Hager. – A particle filter without dynamics for robust 3d face tracking. – *Conference on Computer Vision and Pattern Recognition*, 2004.
- [Lucas 81] B. Lucas, T. Kanade. – An iterative image registration technique with an application to stereo vision. – *Proceeding of the International Joint Conference on Artificial Intelligence*, pp. 674–679, 1981.

- [Ma 04] B. Ma, R.E. Ellis. – Surface-based registration with a particle filter. – *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, 2004.
- [MacCormick 99] J. MacCormick, A. Blake. – A probabilistic exclusion principle for tracking multiple objects. – *International Conference on Computer Vision*, pp. 572–578, 1999.
- [MacCormick 00] J. MacCormick, M. Isard. – Partitioned sampling, articulated objects, and interface-quality hand tracking. – *European Conference on Computer Vision*, 2000.
- [Marrs 02] A. Marrs, S. Maskell, Y. Bar-Shalom. – Expected likelihood for tracking in clutter with particle filters. – *SPIE Conference on Signal and Data Processing of Small Targets*, vol. 4728, 2002.
- [Masson 04] L. Masson, F. Jurie and M. Dhome. – Suivi de motifs texturés : approche hybride texture/contours. – *Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, 2004.
- [Matthews 04] I. Matthews, T. Ishikawa, S. Baker. – The template update problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6) :810–815, 2004.
- [Matthies 89] L. Matthies, T. Kanade, R. Szeliski. – Kalman filter-based algorithms for estimating depth from image sequences. *International Journal of Computer Vision*, 3 :209–238, 1989.
- [Merwe 00] R. van der Merwe, A. Doucet, N. de Freitas, E. Wan. – *The unscented particle filter*. – Rapport technique CUED/F-INFENG/TR 380, University of Cambridge, Engineering department, 2000.
- [Merwe 01] R. van der Merwe, E. Wan. – The square-root unscented Kalman filter for state and parameter-estimation. – *International Conference on Acoustics, Speech, and Signal Processing*, 2001.
- [Merwe 03] R. van der Merwe, E. Wan. – Sigma-point kalman filters for probabilistic inference in dynamic state-space models. – *Workshop on Advances in Machine Learning*, 2003.
- [Metropolis 49] N. Metropolis, S. Ulam. – The Monte Carlo method. *Journal of the American Statistical Association*, 44 :335–341, 1949.
- [Metropolis 53] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.h. Teller, E. Teller. – Equations of state calculations by fast computing machine. *The Journal of Chemical Physics*, 1953.
- [Meyer 94] F. Meyer, P. Bouthemy. – Region-based tracking using affine motion models in long image sequences. *CVGIP : Image Understanding*, 60(2) :119–140, 1994.
- [Mezouar 01] Y. Mezouar, F. Chaumette. – Model-free optimal trajectories in the image space : Application to robot vision control. – *Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 1155–1162, 2001.

- [Mikolajczyk 02] K. Mikolajczyk, C. Schmid. – An affine invariant interest point detector. – *European Conference on Computer Vision*, vol. 1, pp. 128–142, 2002.
- [Milheiro 94] P. Milheiro, de Oliveira. – Approximate filters for a nonlinear discrete time filtering problem with small observation noise. *Stochastics and Stochastics Reports*, 46(24), 1994.
- [Moghaddam 97] B. Moghaddam, A. Pentland. – Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 1997.
- [Morency 03] L.-P. Morency, A. Rahimi, T. Darrell. – Adaptive view-based appearance models. – *Conference on Computer Vision and Pattern Recognition*, 2003.
- [Musso 01] C. Musso, N. Oudjane, F. Le Gland. – Improving regularised particle filters. [*Doucet 01a*], chap. 12. – Springer-Verlag, 2001.
- [Neal 01] R. Neal. – Annealed importance sampling. *Statistics and Computing*, 11 :125–139, 2001.
- [Nguyen 01] H.T. Nguyen, M. Worring, R. van den Boomgaard. – Occlusion robust adaptive template tracking. – *International Conference on Computer Vision*, vol. 1, pp. 678–683, 2001.
- [Nguyen 04] H.T. Nguyen, A. Smeulders. – Tracking aspect of the foreground against the background. – *European Conference on Computer Vision*, pp. 446–456, 2004.
- [Nickels 98] K. Nickels. – *Model Based Tracking of Articulated Objects*. – PhD. Thesis, University of Illinois, Department of Electrical Engineering, 1998.
- [Nickels 99] K. Nickels, S. Hutchinson. – Measurement error estimation for feature tracking. – *Conference on Robotics and Automation*, pp. 2527–2532, 1999.
- [North 98] B. North, A. Blake. – Learning dynamical models using Expectation-Maximisation. – *International Conference on Computer Vision*, 1998.
- [Nummiaro 03] K. Nummiaro, E. Koller-Meier, L. Van Gool. – An adaptive color-based particle filter. *Image and Vision Computing*, 21(1) :99–110, 2003.
- [Odobez 94] J.-M. Odobez. – *Estimation, détection et segmentation du mouvement : une approche robuste et markovienne*. – PhD. Thesis, Université de Rennes 1 - France, 1994.
- [Odobez 95a] J.-M. Odobez, P. Bouthemy. – Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6(4) :348–365, 1995.
- [Odobez 95b] J.M. Odobez, P. Bouthemy. – Motion-based motion segmentation exploiting a 2d motion model robust estimation. – *International Conference on Image Processing*, vol. 3, pp. 628–632, 1995.

- [Odobez 03] J.-M. Odobez, S. Ba, D. Gatica-Perez. – An implicit motion likelihood for tracking with particle filters. – *British Machine Vision Conference*, 2003.
- [Okuma 04] K. Okuma, A. Taleghani, N. De Freitas, J. Little, D. Lowe. – A boosted particle filter : multitarget detection and tracking. – *European Conference on Computer Vision*, 2004.
- [Oudjane 00] N. Oudjane. – *Stabilité et approximations particulières en filtrage non linéaire - Application au pistage*. – PhD. Thesis, Université de Rennes 1, 2000.
- [Papanikolopoulos 93] N.P. Papanikolopoulos, P.K. Khosla, 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.
- [Pham 98] D.T. Pham. – *Stochastic methods for sequential data assimilation in strongly nonlinear systems*. – Rapport de Recherche n3597, INRIA Institut National de Recherche en Informatique et Automatique, 1998.
- [Picard 91] J. Picard. – Efficiency of the extended kalman filter for nonlinear systems with small noise. *SIAM Journal on Applied Mathematics*, 51(3) :843–885, 1991.
- [Pitt 99] M. Pitt, N. Shepard. – Filtering via simulation : Auxiliary particle filter. *Journal of the American Statistical Association*, 94(446) :590–599, 1999.
- [Pitt 01] M. Pitt, M. Shephard. – Auxiliary variable based particle filters. *[Doucet 01a]*, chap. 13. – Springer-Verlag, 2001.
- [Pressigout 05] M. Pressigout, E. Marchand. – Fast multimodal 2D tracker : Cue fusion for robustness and accuracy. – *sousmis à IEEE International Conference on Robotics and Automation*, 2005.
- [Pérez 02] P. Pérez, C. Hue, M. Gangnet. – Color-based probabilistic tracking. – *European Conference on Computer Vision*, pp. 661–675, 2002.
- [Pérez 03] P. Pérez. – *Modèles et algorithmes pour l'analyse probabiliste des images*. – Habilitation à diriger des recherches, Université de Rennes 1, 2003.
- [Pérez 04] P. Pérez, Jaco Vermaak, Andrew Blake. – Data fusion for visual tracking with particles. *Proceedings of the IEEE (issue on State Estimation)*, 92(3) :495–513, 2004.
- [Rahimi 01] A. Rahimi, L.-P. Morency, T. Darrell. – Reducing drift in parametric motion tracking. – *International Conference on Computer Vision*, 2001.
- [Raja 98] Y. Raja, S. McKenna, S. Gong. – Colour model selection and adaptation in dynamic scenes. – *European Conference on Computer Vision*, pp. 460–474, 1998.

- [Remazeilles 04] A. Remazeilles, F. Chaumette, P. Gros. – Robot motion control from a visual memory. – *IEEE International Conference on Robotics and Automation*, vol. 4, pp. 4695–4700, 2004.
- [Ristic 04a] B. Ristic, S. Arulampalam, N. Gordon. – *Beyond the Kalman filter - Particle filters for tracking applications*. – Artech House, 2004.
- [Ristic 04b] B. Ristic, S. Arulampalam, N. Gordon. – A tutorial on particle filters. *in [Ristic 04a]*, chap. 3. – Springer Verlag, 2004.
- [Rittscher 99] J. Rittscher, A. Blake. – Classification of human body motion. – *International Conference on Computer Vision*, pp. 634–639, 1999.
- [Robert 99] C. Robert, G. Casella. – *Monte Carlo statistical methods*. – Springer Verlag, 1999.
- [Rong Li 96a] X. Rong Li, Y. Bar-Shalom. – Tracking in clutter with nearest neighbor filter : analysis and performance. *IEEE Transactions on Aerospace and Electronic Systems*, 32(3) :995–1010, 1996.
- [Rong Li 96b] X. Rong Li, X. Zhi. – Probabilistic strongest neighbor filter for tracking in clutter. – *SPIE Conference on Signal and Data Processing of Small Targets*, vol. 2759, 1996.
- [Rong Li 97] X. Rong Li. – Data association probability and measurement density function of tracking in clutter with strongest neighbor measurements. – *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 3163, 1997.
- [Rong Li 01] X. Rong Li, V.P. Jilkov. – A survey of maneuvering target tracking - part III : Measurement models. – *SPIE Conference on Signal and Data Processing of Small Targets*, pp. 4473–4571, 2001.
- [Rotella 03] F. Rotella. – Filtrage optimal. – cours de l'École Nationale d'Ingénieurs de Tarbes, 2003.
- [Rui 01] Y. Rui, Y. Chen. – Better proposal distributions : object tracking using unscented particle filter. – *Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 786–793, 2001.
- [Salmond 90] D. Salmond. – Mixture reduction algorithms for target tracking and guidance. – *SPIE Conference on Signal and Data Processing of Small Targets*, vol. 1305, pp. 434–445, 1990.
- [Saporta 90] G. Saporta. – *Probabilités Analyse des données et Statistiques*. – Editions Technip, 1990.
- [Särkkä 04] S. Särkkä, A. Vehtari, J. Lampinen. – Rao-Blackwellized Monte Carlo data association for multiple target tracking. – *International Conference on Information Fusion*, 2004.
- [Schmetz 93] J. Schmetz, K. Holmlund, J. Hoffman, B. Strauss, B. Mason, V. Gaertner, A. Koch, L. Van De Berg. – Operational cloud-motion winds from meteosat infrared images. *Journal of Applied Meteorology*, 32 :1206–1225, 1993.

- [Schmid 96] C. Schmid. – *Appariement d'images par invariants locaux de niveaux de gris - Application à l'indexation d'une base d'objets.* – PhD. Thesis, Institut National Polytechnique de Grenoble, 1996.
- [Schmid 00] C. Schmid, R. Mohr, C. Bauckhage. – Evaluation of interest point detector. *International Journal of Computer Vision*, 37(2) :151–172, 2000.
- [Schön 03a] T. Schön. – *On Computational Methods for Nonlinear Estimation.* – Rapport de DEA, Department of Electrical Engineering, Linköping University, 2003.
- [Schön 03b] T. Schön, F. Gustafsson, P.-J. Nordlund. – *Marginalized Particle Filters for Nonlinear State-space Models.* – Rapport de Recherche n LiTH-ISY-R-2548, Department of Electrical Engineering, Linköping University, 2003.
- [Schulz 01] D. Schulz, W. Burgard, D. Fox, A.B. Cremers. – Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. – *Conference on Robotics and Automation*, 2001.
- [Shi 94] J. Shi, C. Tomasi. – Good features to track. – *Conference on Computer Vision and Pattern Recognition*, pp. 593–600, 1994.
- [Sidenbladh 00] H. Sidenbladh, M. Black, D. Fleet. – Stochastic tracking of 3d human figures using 2d image motion. – *European Conference on Computer Vision*, vol. 2, pp. 702–718, 2000.
- [Sidenbladh 01] H. Sidenbladh. – *Probabilistic tracking and reconstruction of 3d human motion in monocular video sequences.* – PhD. Thesis, Université de Stockholm, 2001.
- [Sidenbladh 02] H. Sidenbladh, M. Black, L. Sigal. – Implicit probabilistic models of human motion for synthesis and tracking. – *European Conference on Computer Vision*, vol. 1, pp. 784–800, 2002.
- [Singer 74] R. Singer, R. Sea, K. Housewright. – Derivation and evaluation of improved tracking filters for use in dense multitarget environment. *IEEE Transactions on Information Theory*, 20(4) :423–432, 1974.
- [Singh 92] A. Singh, P. Allen. – Image-flow computation : An estimation-theoretic framework and a unified perspective. *CVGIP : Image Understanding*, 56(2) :152–177, 1992.
- [Sminchisescu 02] C. Sminchisescu, B. Triggs. – Hyperdynamic importance sampling. – *European Conference on Computer Vision*, 2002.
- [Smith 95] P. Smith, G. Buechler. – A branching algorithm for discriminating and tracking multiple objects. *IEEE Transactions on Automatic Control*, 20 :101–104, 1995.
- [Song 03] T. Song, S. Shin. – A probabilistic nearest neighbor filter for m validated measurements. – *International Conference on Information Fusion*, 2003.

- [Song 04] T. Song, K. Rhee, D. Lee. – A probabilistic strongest neighbor filter algorithm for m validated measurements. – *International Conference on Information Fusion*, 2004.
- [Sorenson 71] H.W. Sorenson, D.L. Alspach. – Recursive Bayesian estimation using Gaussian sums. *Automatica*, 7 :465–479, 1971.
- [Sorenson 88] H.W. Sorenson. – *Bayesian analysis of time series and dynamical models*, chap. 6 Recursive estimation for nonlinear dynamic systems. – J.C. Spall, 1988.
- [Streit 94] R. Streit, T. Luginbuhl. – Maximum likelihood method for probabilistic multi-hypothesis tracking. – *SPIE Conference on Signal and Data Processing of Small Targets*, vol. 2235, 1994.
- [Sturm 02] P. Sturm. – Structure and motion for dynamic scenes - the case of points moving in planes. – *European Conference on Computer Vision*, vol. 2, pp. 867–882, 2002.
- [Sullivan 01] J. Sullivan, J. Rittscher. – Guiding random particles by deterministic search. – *International Conference on Computer Vision*, pp. 323–330, 2001.
- [Tarel 02] J.-P. Tarel, S.-S. Ieng and P. Charbonnier. – Using robust estimation algorithms for tracking explicit curves. – *European Conference on Computer Vision*, pp. 492–507, 2002.
- [Tissainayagam 01] P. Tissainayagam, D. Suter. – Performance prediction analysis of a point feature tracker based on different motion models. *Computer Vision and Image Understanding*, 84 :104–125, 2001.
- [Tissainayagam 04] P. Tissainayagam, D. Suter. – Assessing the performance of corner detectors for point feature tracking applications. *Image and Vision Computing*, 22(8) :663–679, 2004.
- [Tomasi 91] C. Tomasi, T. Kanade. – *Detection and tracking of point features*. – Rapport de Recherche nCMU-CS-91-132, Carnegie Mellon University, 1991.
- [Tommasini 98] T. Tommasini, A. Fusiello, E. Trucco, V. Roberto. – Making good features to track better. – *Conference on Computer Vision and Pattern Recognition*, pp. 178–183, 1998.
- [Torma 04] P. Torma, C. Szepesvári. – Enhancing particle filters using local likelihood sampling. – *European Conference on Computer Vision*, pp. 16–27, 2004.
- [Torr 03] P. Torr, C. Davidson. – IMPSAC : synthesis of importance sampling and random sample consensus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25 :354–364, 2003.
- [Torresani 02] L. Torresani, C. Bregler. – Space-time tracking. – *European Conference on Computer Vision*, 2002.
- [Toyama 02] K. Toyama, A. Blake. – Probabilistic tracking with exemplars in a metric space. *International Journal of Computer Vision*, 48(1) :9–19, 2002.

- [Turk 91] M. Turk, A. Pentland. – Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1) :71–86, 1991.
- [Vastel 04] G. Vastel. – Localisation précise d’éléments faciaux pour la reconnaissance de visages dans des séquence vidéos. – Rapport de projet de fin d’étude, École Nationale Supérieure des Sciences Appliquées et de Technologie de Lannion, 2004.
- [Vermaak 01] J. Vermaak, A. Blake, M. Gangnet, P. Pérez. – Sequential Monte Carlo fusion of sound and vision for speaker tracking. – *International Conference on Computer Vision*, vol. 1, pp. 741–746, 2001.
- [Vermaak 02a] J. Vermaak, C. Andrieu, A. Doucet, S.J. Godsill. – Particle methods for Bayesian modeling and enhancement of speech signals. *IEEE Transactions on Speech and Audio Processing*, 10(3) :173–185, 2002.
- [Vermaak 02b] J. Vermaak, P. Pérez, M. Gangnet, A. Blake. – Towards improved observation models for visual tracking : selective adaptation. – *European Conference on Computer Vision*, vol. 1, pp. 645–660, 2002.
- [Vlassis 02] N. Vlassis, B. Terwijn, B. Kröse. – Auxiliary particle filter robot localization from high-dimensional sensor observations. – *Conference on Robotics and Automation*, pp. 7–12, 2002.
- [Wan 00] E.A. Wan, R. van der Merwe. – The unscented kalman filter for nonlinear estimation. – *IEEE Symposium on Adaptive Systems for Signal Processing, Communication and Control*, 2000.
- [Wang 02] B. Wang, K.K. Sung, T.K. Ng. – The localized consistency principle for image matching under non-uniform illumination variation and affine distortion. – *European Conference on Computer Vision*, vol. 1, pp. 205–219, 2002.
- [Wolf 02] J. Wolf, W. Burgard, H. Burkhardt. – Using an image retrieval system for vision-based mobile robot localization. – *International Conference on Image and Video Retrieval*, 2002.
- [Wu 04] Y. Wu, T. Huang. – Robust vision tracking by integrating multiple cues based on co-inference learning. *International Journal of Computer Vision*, 58(1) :55–71, 2004.
- [Zhang 95] Z. Zhang, R. Deriche, O. Faugeras, Q.-T. Luong. – A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal*, 78 :87–119, 1995.
- [Zhou 04] S. Zhou, R. Chellapa, B. Moghaddam. – Visual tracking and recognition using appearance-based modeling in particle filters. *IEEE Transactions on Image Processing (to appear)*, 2004.

Publications

Revue

E. Arnaud, E. Mémin

Conditional filters for image sequence based tracking - Application to point tracking. – IEEE *Trans. of Image Processing* – accepté pour publication.

Congrès internationaux

E. Arnaud, E. Mémin

Optimal importance sampling for tracking in image sequences : application to point tracking. – Dans *European Conference on Computer Vision, ECCV'04* – Prague, République Tchèque, Mai 2004.

E. Arnaud, E. Mémin, et B. Cernuschi-Frías

A robust stochastic filter for point tracking in image sequences. – Dans *Asian Conference on Computer Vision, ACCV'04* – Ile de Jeju, Corée du sud, Janvier 2004.

Congrès et workshops nationaux

E. Arnaud, E. Mémin, et B. Cernuschi-Frías

Filtrage conditionnel pour la trajectographie dans des séquences d'images - Application au suivi de points. – Dans *14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, RFIA'04* – Toulouse, Janvier 2004.

E. Arnaud, E. Mémin, et B. Cernuschi-Frías

Filtrage conditionnel pour le suivi de points dans des séquences d'images. – Dans *Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle, ORASIS'03* – Gerardmer, Mai 2003.

Résumé

Cette étude traite de l'utilisation de méthodes de filtrage pour du suivi dans des séquences d'images. Ces algorithmes reposent sur une représentation du système dynamique par une chaîne de Markov cachée, décrite par une loi dynamique et une vraisemblance des données. Pour construire une méthode générale, une loi dynamique estimée sur les images est considérée. Ce choix met en évidence les limitations du modèle simple de chaîne de Markov cachée, qui ne décrit pas la dépendance des éléments du système aux images.

Nous proposons d'abord une modélisation originale du problème. Celle-ci rend les images explicites et permet de construire des algorithmes sans information *a priori*. Les filtres associés à cette nouvelle représentation sont dérivés sur la base des filtres classiques, en considérant un conditionnement par rapport à la séquence. Il est également présenté comment ce nouveau schéma permet de considérer des modèles simples, pour lesquels la fonction d'importance optimale est disponible.

Ensuite, nous nous intéressons à la validation pratique de la modélisation proposée sur une application de suivi de points caractéristiques. Les systèmes mis en œuvre sont entièrement estimés sur la séquence. Ils associent des mesures de similarité à une dynamique définie à partir d'un mouvement instantané estimé par une méthode différentielle robuste. Afin de contrôler l'importance des différents éléments du système, les matrices de covariance de bruit des modèles sont estimées. Trois algorithmes de suivi de points sont ainsi construits et validés sur de nombreuses séquences réelles. Enfin, cette approche est étendue au suivi de motifs plans texturés. Le modèle considéré introduit une information géométrique par homographie et amène à un algorithme robuste aux occultations totales.

Mots-clefs : filtre de Kalman, filtre particulière, méthodes séquentielles de Monte Carlo, estimateurs de variance minimale, analyse du mouvement, suivi de points caractéristiques, méthodes de corrélation.

Abstract

This thesis is concerned with the use of filtering methods for tracking in image sequences. For the algorithms introduced here, the system is represented by a Hidden Markov Chain, described by a dynamic law and a likelihood. In order to construct a general method, the dynamic law is estimated from the images. This choice underlines some limitations of the simple model of Hidden Markov Chains. Indeed, such a modelisation does not describe the dependance of the system's components to the image sequence.

We first propose an original modelisation of the problem where the image data are explicitly taken into account. Such a model allows us to consider algorithms that do not rely on a priori information. Different kinds of filters associated with this new modelisation are derived. Then, a validation of this modelisation is presented. Three feature point trackers are proposed on this basis. They combine a dynamical law relying on an estimated motion, and measurements provided by a matching technique. Finally, this approach is extended to planar object tracking.

Keywords : Kalman filter, particle filter, sequential Monte Carlo methods, minimum variance estimate, motion analysis, feature-point tracking, block-matching.