

Modèles déformables statistiques pour l’analyse de structures 2D dans des séquences d’images

Charles Kervrann[†] – Fabrice Heitz[‡]

[†] *IRISA / INRIA*
Campus Universitaire de Beaulieu
35042 Rennes Cedex

[‡] *LSIIT – CNRS / École Nationale Supérieure de Physique de Strasbourg*
Boulevard Sébastien Brant
67400 Illkirch

RÉSUMÉ. L’interprétation de scènes dynamiques complexes requiert, dans un nombre croissant de domaines d’application, l’analyse de structures déformables. Dans cet article nous présentons un cadre statistique de modélisation de structures déformables 2D. La représentation adoptée repose sur la définition d’une forme prototype (“*template*”) à laquelle sont appliquées des déformations globales et locales, décrites par un modèle markovien. Une estimation statistique des déformations alliée à un suivi temporel de la structure assurent une grande robustesse dans l’extraction d’objets déformables variés de séquences d’images longues.

ABSTRACT. In many applications of dynamic scene analysis, the objects or structures to be analysed show deformations. In this paper, we introduce a statistical framework for the representation of 2D deformable structures in image sequences. The model relies on the specification of a template, on which global as well as local deformations (described by a Markov process) are defined. The estimation and tracking of deformations by statistical estimation methods enables a robust extraction of deformable parts from long image sequences.

MOTS-CLÉS : analyse de séquences d’images, modèles déformables, segmentation, suivi, modèles markoviens, filtrage de Kalman.

KEY WORDS : image sequence analysis, deformable models, segmentation, tracking, Markov models, Kalman filtering.

1. Introduction

Jusqu'à la fin des années quatre-vingts, les modèles développés en vision par ordinateur ont essentiellement porté sur la représentation et l'analyse d'objets solides animés de mouvements rigides. La modélisation des déformations et leur exploitation en analyse de séquences d'images ont connu un développement important ces cinq dernières années, développement dicté en particulier par les besoins croissants des applications dans ce domaine (imagerie biomédicale, analyse du mouvement humain, analyse de phénomènes destructifs, etc.). L'introduction de "modèles déformables" [KAS 87, PEN 91a, STA 92] pour la représentation des déformations s'est révélée importante à plus d'un titre :

- d'une part, ces modèles permettent la prise en compte d'informations *a priori* sur la structure des objets et leurs déformations, ce qui introduit une robustesse appréciable dans les techniques d'extraction ou de suivi ;
- d'autre part, la modélisation des déformations ouvre des perspectives prometteuses, mais peu explorées pour l'instant, dans le domaine de la caractérisation et de l'interprétation du comportement dynamique d'objets complexes.

On peut noter que la représentation des déformations se heurte à la grande variété des objets ou phénomènes naturels qui peuvent se présenter : objets polyarticulés, structures présentant des déformations élastiques ou phénomènes déformables fluides. Ceci rend nécessaire l'introduction de modèles mathématiques adaptés aux classes de déformations observées. Les modèles que nous présentons dans cet article s'appliquent essentiellement aux deux premières classes d'objets que nous venons de citer. Ils ont pour but de représenter des structures bidimensionnelles déformables en mouvement dans des séquences d'images.

Notre approche repose sur une description de la forme d'intérêt par une structure prototype ("template déformable") qui incorpore des connaissances de nature statistique sur la forme étudiée et sur ses déformations. La forme prototype, s'interprétant comme un représentant de la classe d'objets d'intérêt ("forme moyenne"), est décrite par une suite discrète de points caractéristiques identifiés sur le contour de l'objet analysé (figure 1). Les déformations de la forme prototype sont représentées de façon hiérarchisée par des paramètres globaux et locaux (figure 1). Les paramètres de déformation globaux incluent les transformations du groupe des similitudes planes (rotations, translations et homothéties) ainsi que des paramètres additionnels qui contrôlent les modes de déformation principaux, issus d'une analyse statistique. Les transformations du groupe des similitudes permettent une mise en correspondance globale et grossière de la forme prototype avec les données-image observées. Les modes de déformation principaux assurent un recalage plus précis prenant en compte les déformations globales de la structure (la figure 3b présente un exemple de recalage d'un modèle de main avec les seules transformations du groupe des similitudes, le recalage de la figure 3c intégrant les modes de déformation). Les modes de déformation globaux sont obtenus grâce à une technique de décomposition modale s'appuyant sur une analyse en composantes principales (ou

“transformée de Karhunen-Loeve) des déformations observées sur une population d’apprentissage représentative [COO 94b]. Cette transformée, classiquement utilisée en analyse des données, permet de réduire considérablement la dimension de l’espace de représentation des déformations en ne retenant que les axes principaux d’inertie qui portent la majeure partie de la variabilité observée sur la population d’apprentissage. Les déformations locales sont par ailleurs représentées par des perturbations aléatoires se superposant à la structure globalement déformée (figure 1). Le processus des déformations locales est supposé suivre un modèle de Gauss-Markov du premier ordre, qui prend en compte la dépendance statistique entre les déformations appliquées à des points voisins. Ce processus de déformations locales peut être vu comme un raffinement des déformations globales appliquées à la forme prototype initiale et prenant en compte les variabilités fines rencontrées.

Ce modèle statistique paramétrique est utilisé dans des tâches de segmentation, dans le cadre classique de l’estimation bayésienne. Une estimée statistique des déformations (au sens du critère du “Maximum A Posteriori” – MAP) peut alors être obtenue à partir de données observables constituées par des champs de gradients spatio-temporels extraits préalablement de la séquence d’images. L’estimateur ainsi défini conduit à la minimisation d’une fonction d’énergie globale, fortement non linéaire, qui décrit les interactions entre déformations et données-image. Le minimum de cette fonction d’énergie correspond à une configuration du modèle déformable réalisant une “bonne” segmentation de l’image initiale (figure 3). Des algorithmes d’optimisation globale peuvent être utilisés pour estimer les paramètres globaux, avec un coût calculatoire acceptable, vu le faible nombre de paramètres en jeu (de l’ordre d’une dizaine en général). Le processus de déformations locales, qui constitue un raffinement de la solution, est estimé grâce à des algorithmes d’optimisation locale (algorithme du gradient ou ICM [BES 86]).

La prise en compte de la cohérence du mouvement de la structure déformable améliore par ailleurs sensiblement la qualité de la segmentation. Dans ce but, un suivi des paramètres de déformation par filtrage récursif [BLA 92, KER 94b] est intégré dans la boucle de segmentation. Le suivi par filtrage de Kalman [BAS 94, BLA 92, BLA 93, DER 90, MEY 94, PEN 91a] permet de prendre en compte des déplacements importants de l’objet d’intérêt, facilitant l’optimisation lors de l’estimation des paramètres de déformation [KER 94b].

L’article est organisé de la façon suivante. Nous rappelons tout d’abord en section 2, les grandes familles d’approches décrites dans la littérature pour la représentation et le traitement des objets déformables. Nous y situons en particulier notre approche par rapport aux modèles existants. Le modèle déformable que nous proposons est décrit de façon détaillée en section 3. La procédure de segmentation et les techniques d’optimisation qui lui sont associées sont présentées en section 4. Le suivi du modèle déformable par filtrage récursif fait l’objet de la section 5. Des exemples d’apprentissage, de segmentation et de suivi de structures déformables variées sur des séquences réelles, sont proposés en section 6. Quatre études de cas illustrent l’aptitude de l’approche proposée à

s'adapter à des applications diverses ainsi que sa robustesse dans le traitement de séquences longues (plus d'une centaine de trames) et bruitées.

2. Modèles déformables en analyse d'image

Les modèles déformables introduits en imagerie intègrent des connaissances *a priori* sur la structure des objets et sur les déformations de ces structures.

Depuis les travaux de Kass, Witkin et Terzopoulos sur les contours actifs, les modèles déformables ont connu des déclinaisons variées en imagerie, [COH 92, GRE 91]. D'abord utilisés en segmentation d'images statiques, ces modèles sont maintenant exploités avec profit en restauration d'images [GRE 91], pour le suivi de structures déformables dans des séquences d'images [BAS 94, COH 93, HER 92, MAR 92] ou pour la caractérisation d'objets par l'analyse des déformations.

Différentes classes de modèles déformables ont été développées ces dernières années. Une typologie de ces modèles peut s'organiser autour des caractéristiques suivantes :

- le cadre de modélisation (déterministe ou statistique),
- le niveau de spécialisation de la représentation (modèles à structure peu contrainte ou approches par forme prototype - "templates déformables"),
- le caractère continu ou discret du modèle considéré.

Les modèles déformables sont généralement associés (de façon directe ou indirecte) à des techniques de minimisation de fonctions d'énergie par des procédures d'optimisation locale ou globale. Les techniques d'optimisation locale nécessitent une initialisation du modèle proche de la configuration optimale [COH 93, COO 94b, KAS 87]. Les techniques d'optimisation globale (souvent stochastiques [GEM 84]) sont théoriquement insensibles à la configuration initiale du modèle déformable [GRE 91, KER 94a]. L'emploi d'une optimisation locale ou globale dépend essentiellement de la définition du problème à résoudre qui parfois nécessite la recherche d'un minimum global [KER 94a] et parfois se contente d'un minimum local, dans une région donnée, qui dépend de l'initialisation [COH 93, KAS 87].

Parmi les modèles déterministes, les modèles de contours actifs 1D (à usage général), dérivés des "snakes" de Kass, Witkin et Terzopoulos, [KAS 87] ont connu un succès important, même s'ils présentent un certain nombre de limitations [BER 91, COH 93, KAS 87, LEI 90]. Ces modèles continus (arcs paramétrés fermés ou non) sont contrôlés par des fonctions d'énergies intégrant des contraintes génériques sur le contour de l'objet à extraire. Dans ces approches, une courbe paramétrée, dont la configuration initiale est généralement spécifiée par l'opérateur, évolue vers les contours les plus proches, sous l'effet de forces créées par l'image [KAS 87]. L'optimisation de la fonction d'énergie conduit dans ce cas à la résolution d'un système d'équations aux dérivées partielles, menée par des techniques de différences finies, [BER 91, KAS 87]) qui nécessitent une initialisation adéquate pour converger vers des solutions satisfaisantes.

Des procédures ont été proposées, pour certains types d’images, pour produire de telles initialisations sans intervention de l’opérateur, [ROU 91]. Les contours actifs se prêtent toutefois mal à la modélisation de formes présentant des discontinuités (de courbure par exemple) quoique des techniques aient été proposées pour les prendre en compte dans la modélisation, [BER 91]. Ces modèles ont été généralisés à la représentation de surfaces déformables par Cohen *et al*, [COH 92, COH 93] pour la segmentation de volumes de données en imagerie médicale. Des méthodes de résolution par éléments finis ont été introduites dans ce contexte.

Les modèles décrits par Yuille, [LIP 90, YUI 92] se rattachent à la classe précédente, mais s’appuient sur une description plus spécifique des formes à représenter. Des exemples de modèles de lèvres et d’œil, sont par exemple présentés dans [YUI 92]. Les formes sont décrites par un ensemble de courbes 2D paramétrées simples (cercles, ellipses, paraboles, etc.), ajustées dans l’image par minimisation de fonctions d’énergie grâce à des techniques de gradient. Ces modèles doivent être construits “manuellement” pour chaque application à partir des connaissances sur la structure (géométrique et photométrique) des formes à représenter. Des descripteurs de formes plus génériques ont également été introduits : Staib et Duncan, [STA 92] proposent d’utiliser des descripteurs de Fourier, associés à des techniques itératives de minimisation d’énergie pour analyser des objets déformables.

Les modèles physiques déformables (2D ou 3D) (inspirés de la mécanique) constituent une autre classe de représentations, d’abord utilisées dans le domaine de la synthèse d’images, de la simulation et de l’animation [TER 88]. Ils ont depuis été exploités avec succès en analyse d’images, pour segmenter et suivre des structures déformables 2D et 3D dans des séquences d’images [NAS 93, PEN 91b, PEN 91a, TER 91]. Dans ces approches, les structures déformables sont représentées par des modèles paramétriques (ellipsoïdes, superquadriques, modèles polynomiaux) ou par des éléments finis. Leur évolution est contrôlée par les lois de la dynamique, [PEN 91a, TER 91]. Ils nécessitent généralement une initialisation très proche de la bonne position du modèle pour produire des segmentations satisfaisantes [PEN 91b, PEN 91a, TER 91]. Les équations de mouvement de la structure déformable peuvent, dans le cas des modèles physiques, être résolues efficacement en réalisant une décomposition des déformations selon les modes propres de vibration du système physique [PEN 91b, PEN 91a, NAS 93]. Cette technique d’analyse modale conduit à une très bonne approximation des déformations dans un espace de dimension réduite.

L’analyse modale a également été exploitée pour adapter les modes de déformations de la structure aux déformations effectivement observées sur une population d’apprentissage [COO 94b, SZE 95] (on parle alors de modes “statistiques”). Cootes *et al*. [COO 94b] utilisent ainsi une décomposition de Karhunen-Loeve des déformations observées sur une population, qui permet de ne retenir qu’un ensemble restreint de modes de déformation significatifs, associés aux axes d’inertie principaux. Typiquement de cinq à dix modes autorisent une

description précise des déformations observées sur la population de départ, [COO 94b]. Dans l’approche de Cootes *et al.* [COO 94b], le modèle est ensuite classiquement ajusté sur les contours préalablement extraits de l’image par des techniques de gradient (nécessitant une initialisation proche de la configuration souhaitée [COO 94b]).

Récemment, ces deux approches modales (modes vibratoires “physiques” et modes “statistiques”) ont été partiellement unifiées et utilisées conjointement pour extraire des structures 3D anatomiques du cerveau [MAR 94] ou pour modéliser des visages [COO 95].

Notons enfin que des résultats encourageants ont été obtenus en restauration et segmentation d’images bruitées par Grenander *et al* [GRE 91] grâce à l’utilisation de modèles déformables stochastiques. Ces modèles, peu explorés pour l’heure [MAR 92, LAI 94], s’appuient sur des formes prototypes (“templates”) dont les déformations locales sont décrites par des modèles statistiques markoviens [BES 86, GEM 84]. Leur mise en œuvre est réalisée dans le cadre de l’estimation statistique bayésienne et conduit à la minimisation de fonctions d’énergie globale (menée par des techniques de Monte Carlo, [GRE 91], peu sensibles aux configurations initiales du modèle). Les propriétés globales de ces modèles peu contraints s’avèrent toutefois difficiles à maîtriser et les techniques de Monte Carlo associées restent extrêmement coûteuses.

L’approche que nous décrivons dans cet article combine, de façon hiérarchisée, dans un cadre cohérent et de nature statistique deux techniques de modélisation d’objets déformables précédemment décrites. Le modèle que nous proposons permet en effet d’allier les avantages d’une description compacte des déformations globales (par décomposition modale) à une modélisation statistique (markovienne) des déformations locales. Ce cadre de modélisation permet la mise en œuvre d’algorithmes d’estimation des hyperparamètres du modèle et des déformations s’appuyant sur le cadre mathématique bien posé de l’estimation bayésienne.

Cette approche présente les avantages suivants par rapport aux approches existantes :

- Le modèle statistique hiérarchisé permet de représenter, dans un *même* modèle mathématique, les variabilités globales *et* locales de structures déformables diverses.
- Notre approche permet une description statistique parcimonieuse des déformations observées. Typiquement une dizaine de paramètres modaux permettent une première représentation approchée des déformations, facilitant l’estimation statistique des déformations locales. Ce n’est pas le cas des modèles stochastiques précédemment proposés par Grenander *et al* [GRE 91] qui décrivent les déformations globales à partir d’une multitude de déformations locales, conduisant à des temps de calcul prohibitifs.
- Par le biais de la phase d’apprentissage, le modèle s’adapte, d’un point de vue statistique, à la classe d’objets et de déformations à représenter. Ce n’est pas le cas de la représentation proposée dans [GRE 91], dans laquelle les déformations ne sont pas contraintes par un apprentissage.

- Notre approche intègre une estimation *en ligne* des hyperparamètres du modèle statistique, qui contrairement à l’approche classique EM est non itérative et donc peu coûteuse.
- Notre approche conduit à des schémas de segmentation et de suivi qui sont robustes, en particulier aux niveaux de bruit élevés, aux bruits non blancs, aux données partielles ou manquantes ainsi qu’aux occlusions entre objets. Ce n’est pas le cas du modèle développé par Cootes [COO 94a, COO 94b], qui s’avère très sensible au bruit.
- L’algorithme de segmentation proposé peut, lorsque le problème s’y prête, s’appuyer sur des techniques d’optimisation globale (relaxation stochastique), qui convergent vers un minimum global. Ceci peut être souhaitable dans certaines applications et permet alors d’éviter toute interaction manuelle pour initialiser la segmentation par une configuration du modèle proche de la solution. Ce point est important lorsqu’il s’agit de développer des algorithmes non supervisés.

Les caractéristiques du modèle que nous proposons sont décrites dans la section suivante.

3. Un modèle déformable statistique hiérarchique

Les techniques de décomposition modale se sont révélées très efficaces pour décrire les déformations par un nombre réduit de paramètres significatifs [TER 91, PEN 91a, NAS 93]. Dans le cadre de notre travail, la représentation des déformations globales est obtenue grâce à une technique récemment décrite par Cootes *et al.* [COO 94b] qui proposent de retenir un nombre réduit de modes significatifs de variation dans une décomposition de Karhunen-Loeve des déformations observées sur une population d’apprentissage représentative. Un processus de déformation locale, décrit par un modèle markovien, est introduit en second lieu pour affiner cette première description. Cette représentation hiérarchique originale (figure 1) se révèle très efficace pour décrire l’ensemble des déformations spécifiques à une classe d’objets.

3.1. Modélisation des déformations globales

La mise en œuvre du modèle déformable passe par une phase préalable d’apprentissage qui peut être réalisée hors-ligne, ou en ligne (c’est-à-dire simultanément avec la segmentation et le suivi). Cette phase d’apprentissage a pour objectif d’identifier la structure du modèle moyen (le “template”) et de caractériser ses déformations à partir d’une population représentative. Nous présentons dans [KER 95] une méthode permettant de réaliser cet apprentissage partiellement en ligne, l’identification des caractéristiques du modèle s’effectuant sur les premières images de la séquence. Nous supposons ici disposer d’une séquence d’apprentissage caractéristique de la structure déformable à représenter et l’apprentissage est mené hors ligne.

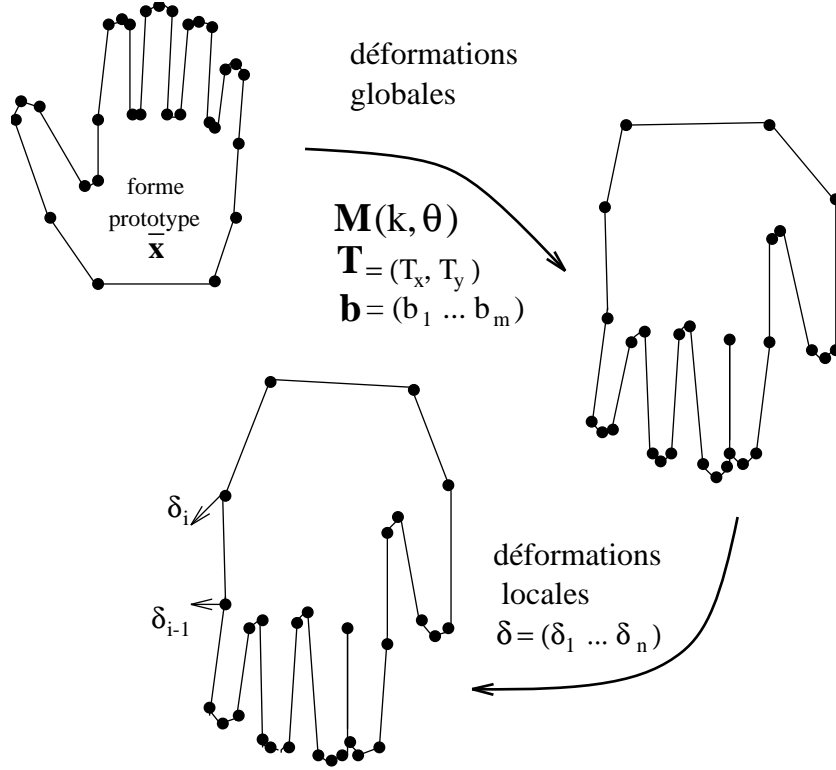


Figure 1. Description hiérarchique des déformations

Une forme particulière \mathbf{x}_i est représentée par une suite donnée de n points caractéristiques de son contour, définissant une approximation polygonale de l'objet à représenter (figure 1) :

$$\mathbf{x}_i = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{in}, y_{in})^T.$$

Si ces n points caractéristiques sont associés aux lieux de forte courbure identifiés sur le contour de l'objet, il est alors possible de les extraire de manière automatique. En pratique, nous les avons désigné manuellement ici. Les déformations d'une forme \mathbf{x}_i , observées sur une population représentative de N_T formes, sont ensuite caractérisées par un vecteur de déplacement $\tilde{\mathbf{x}}_i = \mathbf{x}_i - \bar{\mathbf{x}}$ où $\bar{\mathbf{x}}$ est l'élément représentant de la classe d'objets considérée (forme prototype). La forme prototype est définie par la position moyenne des différentes approximations polygonales des objets, préalablement recalées par similitude plane (selon la technique décrite dans [COO 94b]) (figure 1) :

$$\bar{\mathbf{x}} = \frac{1}{N_T} \sum_{i=1}^{N_T} \mathbf{x}_i. \quad [1]$$

La décomposition de Karhunen-Loeve (KL) des déformations s'obtient par diagonalisation de la matrice d'autocovariance du vecteur de déplacement \mathbf{x}_i , définie par :

$$\mathbf{C} = \frac{1}{N_T} \sum_{i=1}^{N_T} (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})^T \quad [2]$$

Les vecteurs de base associés à la transformée de Karhunen-Loeve sont obtenus en résolvant l'équation aux valeurs propres :

$$\Lambda = \Phi_{2n}^T \mathbf{C} \Phi_{2n} \quad [3]$$

où $\Phi_{2n} = (\phi_1, \phi_2, \dots, \phi_{2n})$ est la matrice contenant les vecteurs propres de \mathbf{C} et Λ est une matrice diagonale dont les éléments non nuls sont les valeurs propres de \mathbf{C} . Le vecteur de déformation $\tilde{\mathbf{x}}$ s'exprime dans la base des vecteurs propres par :

$$\mathbf{b}_{2n} = \Phi_{2n}^T \tilde{\mathbf{x}} = \Phi_{2n}^T (\mathbf{x} - \bar{\mathbf{x}}) \quad [4]$$

où \mathbf{b}_{2n} est un vecteur ($2n \times 1$). La transformation KL conduit à des paramètres de déformation décorrélés et autorise une représentation compacte en ne conservant que les m modes de variation les plus significatifs, associés aux m valeurs propres les plus fortes. Une forme particulière \mathbf{x} est représentée en conséquence par :

$$\mathbf{x} = \bar{\mathbf{x}} + \Phi \mathbf{b} \quad [5]$$

où $\mathbf{b} = (b_1, \dots, b_m)^T$ représente le vecteur ($m \times 1$) associé aux m modes de déformation les plus significatifs et $\Phi = (\phi_1, \phi_2, \dots, \phi_m)$ est une sous-matrice de Φ_{2n} contenant les m vecteurs propres correspondants.

En résumé, la phase d'apprentissage (hors-ligne) réalise donc l'identification préalable de $\bar{\mathbf{x}}$ et Φ , qui sont fixés et définissent la structure du modèle et la façon dont ce modèle peut se déformer. Les déformations particulières du modèle dans une image sont alors contrôlées par les paramètres b_i , $i = 1, \dots, m$ qui varient d'une image à l'autre. Quelques modes sont suffisants pour décrire convenablement la variabilité totale observée sur la population d'apprentissage. En pratique nous avons conservé dans nos expériences un nombre de modes permettant de prendre en compte 90% de la variabilité totale (soit de 5 à 7 modes). À titre d'illustration, le tableau 1 présente les valeurs propres normalisées, calculées sur une population de onze mains représentées par un modèle à 29 points. Le rapport d'une valeur propre sur l'ensemble total des valeurs propres exprime le pourcentage d'erreur introduit si le mode associé n'est pas sélectionné. Sur cet exemple, on remarque qu'un nombre réduit de paramètres (cinq à sept) intègre déjà plus de 90% de la variabilité totale.

Le modèle déformable global est alors décrit par la représentation paramétrique suivante (figure 1) :

$$\mathbf{X} = \mathbf{M}(k, \theta) [\bar{\mathbf{x}} + \Phi \mathbf{b}] + \mathbf{T} \quad [6]$$

Valeur propre	$\frac{\lambda_i}{\sum_{i=1}^{2n} \lambda_i}$	$\frac{\lambda_i}{\sum_{i=1}^{2n} \lambda_i}$ cumulé
λ_1	44%	44%
λ_2	21%	65%
λ_3	14%	79%
λ_4	6%	85%
λ_5	5%	90%
λ_6	4%	94%
λ_7	3%	97%
λ_8	2%	99%

Table 1. *Tableau de valeurs propres normalisées (modèle de main à 29 points, $2n = 54$, $N_T = 11$)*

où le vecteur de translation $\mathbf{T} = (t_x, t_y, t_x, t_y, \dots)$ de dimension $2n$, la rotation d'angle θ et l'homothétie de facteur k (regroupées dans la matrice de transformation $\mathbf{M}(k, \theta)$),

$$\mathbf{M}(k, \theta) = \begin{pmatrix} M(k, \theta) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & M(k, \theta) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & M(k, \theta) \end{pmatrix} \quad \text{avec} \quad M(k, \theta) = \begin{pmatrix} k \cos \theta & -k \sin \theta \\ k \sin \theta & k \cos \theta \end{pmatrix}. \quad [7]$$

ainsi que les modes de déformation \mathbf{b} définissent les transformations géométriques globales, appliquées à la forme prototype $\bar{\mathbf{x}}$ (figure 1). L'équation 6 définit un premier niveau de représentation de la structure déformable. Cette représentation est raffinée par une modélisation des déformations locales, décrite dans la section suivante.

3.2. Modélisation des déformations locales

Un processus de déformation locale $\boldsymbol{\delta}$, appliqué aux n points, est introduit pour affiner la description de la section précédente. Les déformations locales peuvent être vues comme des *perturbations aléatoires* (translations locales aléatoires) se superposant à la structure *déterministe* globalement déformée (figure 1). Une hypothèse raisonnable, dans notre cas, consiste à supposer que $\boldsymbol{\delta}$ suit un modèle markovien-gaussien, modélisant les interactions entre points voisins [KER 94a]. Le modèle déformable complet peut alors s'écrire sous la forme (figure 1) :

$$\mathbf{Y} = \mathbf{X} + \boldsymbol{\delta} = \mathbf{M}(k, \theta) [\bar{\mathbf{x}} + \Phi \mathbf{b}] + \mathbf{T} + \boldsymbol{\delta}. \quad [8]$$

Dans notre cas, $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_n)$ et $\delta_i = (\delta_{x_i}, \delta_{y_i})$, est un processus markovien-gaussien centré du 1^{er} ordre :

$$P(\boldsymbol{\delta}) = \frac{1}{Z_p} \exp -\frac{1}{2} \boldsymbol{\delta}^T \mathbf{R}^{-1} \boldsymbol{\delta} \quad [9]$$

où \mathbf{R} est la matrice de covariance de $\boldsymbol{\delta}$ et Z_p est une constante de normalisation. Sous cette hypothèse, la loi conjointe de $\boldsymbol{\delta}$ s'écrit :

$$P(\boldsymbol{\delta}) = \frac{1}{Z_p} \exp -\frac{1}{2} \sum_{i=1}^n \left[\frac{1}{\varepsilon_i^2} \|\delta_i - \delta_{i-1}\|^2 + \frac{1}{\sigma_i^2} \|\delta_i\|^2 \right]. \quad [10]$$

σ_i^2 et ε_i^2 sont les paramètres de variance du modèle; ε_i^2 permet d'agir sur la régularité des déformations locales (interactions entre points voisins) tandis que des valeurs faibles pour σ_i^2 ont tendance à minimiser l'effet des déformations locales. Ces paramètres sont supposés constants quel que soit le point i considéré sur la forme: $\sigma_i^2 = \sigma^2$ et $\varepsilon_i^2 = \varepsilon^2$, $\forall i$. En pratique, il est souhaitable de favoriser une solution régulière, en particulier en présence de données manquantes (occlusions, bruits, etc.). Les valeurs $\sigma = 4$ et $\varepsilon = 1$ ont été adoptées dans ce but.

4. Segmentation d'une image par modèle déformable

Notre approche du problème de segmentation par modèle déformable repose sur une formulation bayésienne du problème, intégrant le modèle défini ci-dessus pour la structure déformable (équation 8) comme modèle *a priori*, ainsi que des modèles d'observations dépendant de l'application traitée.

4.1. Estimation bayésienne des déformations

Nous nous intéressons au problème de l'extraction d'objets mobiles (déformables) d'une séquence d'images. Soit $\mathbf{d} = (d_s, s \in S)$ l'ensemble des observations définies sur une grille rectangulaire (les observations sont liées aux gradients spatio-temporels extraits de la séquence d'images). Le problème de la segmentation est posé comme étant celui de l'estimation (au sens du Maximum A Posteriori – MAP) de la réalisation (cachée) du processus \mathbf{Y} portant la structure déformable à partir du champ des observations \mathbf{d} :

$$\mathbf{Y}^* = \arg \max_{\mathbf{Y}} P(\mathbf{d} | \mathbf{Y}) P(\mathbf{Y}). \quad [11]$$

D'après les hypothèses statistiques énoncées plus haut (équation 9), \mathbf{Y} est un processus markovien-gaussien du 1^{er} ordre. On en déduit :

$$P(\mathbf{Y}) = \frac{1}{Z_p} \exp -\frac{1}{2} (\mathbf{Y} - \mathbf{X}(\boldsymbol{\Theta}))^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{X}(\boldsymbol{\Theta})). \quad [12]$$

Les paramètres de similitude plane $\mathbf{M}(k, \theta)$, \mathbf{T} et les modes de déformation globaux \mathbf{b} sont ici considérés comme des hyperparamètres (déterministes) $\boldsymbol{\Theta} = (\mathbf{M}(k, \theta), \mathbf{T}, \mathbf{b})$ de ce modèle probabiliste.

La loi $P(\mathbf{d} | \mathbf{Y})$ exprime la vraisemblance des observations conditionnellement au processus de déformation à estimer; elle est spécifiée de façon "ad

hoc” par une distribution de Gibbs [GEM 84] qui intègre des connaissances spécifiques liées au type d’images ou d’objets à traiter :

$$P(\mathbf{d} | \mathbf{Y}) = \frac{1}{Z_d} \exp -E_d(\mathbf{Y}, \mathbf{d}) \quad [13]$$

où $E_d(\mathbf{Y}, \mathbf{d})$ est une fonction d’énergie et Z_d est une constante de normalisation. L’expression de $E_d(\mathbf{Y}, \mathbf{d})$ est donnée dans la section suivante, dans deux cas particuliers (segmentation au sens du mouvement, segmentation au sens des contours photométriques).

Si Θ représente le vecteur des hyperparamètres, l’estimée au sens du MAP des déformations peut s’écrire :

$$\begin{aligned} \mathbf{Y}^* &= \arg \max_{\mathbf{Y}} P(\mathbf{d} | \mathbf{Y}) p(\mathbf{Y}) \\ &= \arg \max_{\mathbf{Y}} P(\mathbf{Y}, \mathbf{d} | \Theta) \end{aligned} \quad [14]$$

La distribution conjointe est alors également une distribution de Gibbs :

$$P(\mathbf{Y}, \mathbf{d} | \Theta) = \frac{1}{Z} \exp -E_{\Theta}(\mathbf{Y}, \mathbf{d}), \quad [15]$$

dont la constante de normalisation Z ne dépend pas de Θ [KER 95]. Estimer le processus de déformation au sens du MAP revient donc à minimiser la fonction d’énergie suivante :

$$\begin{aligned} E_{\Theta}(\mathbf{Y}, \mathbf{d}) &= E_d(\mathbf{Y}, \mathbf{d}) + E_p(\mathbf{Y}) \\ &= E_d(\mathbf{Y}, \mathbf{d}) + \frac{1}{2} (\mathbf{Y} - \mathbf{X}(\Theta))^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{X}(\Theta)). \end{aligned} \quad [16]$$

Soulignons que la modélisation du seul terme de vraisemblance $E_d(\mathbf{Y}, \mathbf{d})$ conduirait à des résultats de segmentation peu robustes au bruit et à la variabilité des objets observés. L’introduction du modèle *a priori* $E_p(\mathbf{Y})$ permet de contourner cette difficulté en *contraignant* le processus de déformation. Le modèle d’observation (terme $E_d(\mathbf{Y}, \mathbf{d})$) est décrit dans la section suivante.

4.2. Modèles d’observations

Pour la classe des problèmes qui nous intéresse ici, une énergie d’interaction entre processus de déformation et observations $E_d(\mathbf{Y}, \mathbf{d})$, fonction des paramètres du modèle déformable, est définie. La fonction d’énergie $E_d(\mathbf{Y}, \mathbf{d})$ que nous spécifions s’appuie sur des observations liées aux gradients spatio-temporels extraits de la séquence d’image. Cette énergie liant les variables à estimer et les observations doit être une mesure pertinente permettant l’extraction de l’objet d’intérêt de l’image. Elle est bâtie différemment selon la nature des observations disponibles et l’objectif poursuivi : extraction de régions déformables homogènes au sens du mouvement ou extraction des contours photométriques d’un objet déformable (nous présentons en section 6 plusieurs exemples venant illustrer ces deux cas).

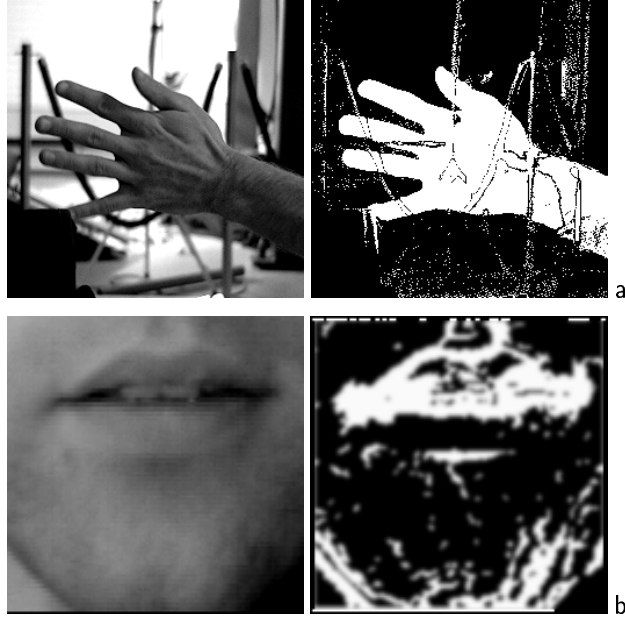


Figure 2. Cartes d'observations utilisées dans la segmentation : a-changements temporels, b-gradients spatiaux

Segmentation au sens du mouvement Soit I_s^t , $s \in S$ l'image de luminance à l'instant t où $s = (x, y)$ désigne les coordonnées spatiales. On suppose que la caméra est fixe. Dans le but d'extraire les objets en mouvement, les changements temporels de la fonction luminance (figure 2a) sont évalués par deux modules complémentaires. Le premier mesure les changements temporels $d_1(s)$ sur trois images successives, le second, dont la sortie $d_2(s)$ sera fusionnée avec le premier, procède en utilisant une image de référence construite dynamiquement "en ligne". Plus précisément :

$$\begin{aligned} d_1(s) &= \min (|I_t(s) - I_{t-\Delta t}(s)|, |I_{t+\Delta t}(s) - I_t(s)|), \\ d_2(s) &= |I_{ref}(s) - I_t(s)|. \end{aligned} \quad [17]$$

où Δt exprime l'intervalle de temps entre deux acquisitions d'images.

La construction de l'image de référence s'inspire d'un estimateur linéaire du fond (fixe) de l'image décrit dans [DON 88] (voir également Irani *et al.* [IRA 92]). Les observations d_1 présentent des valeurs élevées sur les objets mobiles et des valeurs faibles sur le fond (statique). Elle se dégradent toutefois lorsque les objets analysés sont photométriquement très homogènes et présentent un fort recouvrement sur eux-mêmes dans leur déplacement. La seconde source d'observations d_2 ne présente pas cet inconvénient et est utilisée en complément.

Le champ des observations est alors défini par :

$$d_s = \max (s_\alpha(d_1(s)), s_\beta(d_2(s))) \quad [18]$$

$$\begin{aligned} \text{où : } \quad s_\eta(y) &= 1 \quad \text{si } y > \eta \\ s_\eta(y) &= 0 \quad \text{sinon.} \end{aligned} \quad [19]$$

L'énergie $E_d(\mathbf{Y}, \mathbf{d})$ décrit l'interaction entre les changements inter-images seuillés d_s et la configuration du modèle déformable. Pour une configuration donnée du modèle déformable, l'image est partitionnée en deux régions : l'intérieur du modèle $\Gamma_{\mathbf{Y}}^I$ et l'extérieur du modèle $\Gamma_{\mathbf{Y}}^O$ correspondant au fond. L'énergie $E_d(\mathbf{Y}, \mathbf{d})$ vise à regrouper à l'intérieur du modèle déformable les points de l'image présentant un mouvement significatif et à rejeter à l'extérieur du modèle les points appartenant au fond (fixe) de l'image :

$$E_d(\mathbf{Y}, \mathbf{d}) = \sum_{s \in \Gamma_{\mathbf{Y}}^I} |d_s - 1| + \sum_{s \in \Gamma_{\mathbf{Y}}^O} |d_s - 0|. \quad [20]$$

Segmentation au sens des contours photométriques Dans de nombreuses applications, les contours photométriques (gradients spatiaux) de l'objet d'intérêt représentent la seule information exploitable pour réaliser la segmentation (figure 2b). Ces observations permettent cependant une localisation précise de l'objet, sous réserve que le modèle déformable soit convenablement initialisé, manuellement ou de manière automatique, par des procédures adéquates (morphologie mathématique [ROU 91], transformée de Hough [LAI 94]). Dans ce cas, l'énergie liée aux observations est classiquement spécifiée par [YUI 92] :

$$E_d(\mathbf{Y}, \mathbf{d}) = - \sum_{s \in \Gamma_{\mathbf{Y}}} \|\vec{\nabla} I_s\| \quad [21]$$

où $\|\vec{\nabla} I_s\|$ représente la norme du gradient au site s et $\Gamma_{\mathbf{Y}}$ désigne la frontière de la structure déformable. Cette forme de l'énergie $E_d(\mathbf{Y}, \mathbf{d})$, introduite pour la première fois dans [KAS 87] puis adaptée dans [YUI 92] pour une forme "prototype", est définie comme l'intégrale de la norme du gradient observée le long de la frontière définie par le modèle.

4.3. Estimation des hyperparamètres et des déformations

Le problème qui se pose dès lors est celui de l'estimation conjointe du champ \mathbf{Y} et du vecteur (inconnu) des hyperparamètres Θ du modèle [MOH 93]. De manière générale, la loi conjointe (équation 15) est une fonction non-linéaire des paramètres du modèle $\mathbf{M}(k, \theta)$, \mathbf{T} et \mathbf{b} . Par commodité d'écriture, le modèle stochastique est redéfini comme :

$$\mathbf{Y} = \mathbf{X}(\Theta) + \delta. \quad [22]$$

Puisque Θ est inconnu, le problème de segmentation revient à estimer le couple (Θ^*, \mathbf{Y}^*) selon le critère d'optimalité :

$$(\Theta^*, \mathbf{Y}^*) = \arg \max_{\Theta, \mathbf{Y}} P(\mathbf{Y}, \mathbf{d} | \Theta). \quad [23]$$

Ce critère satisfait le maximum global, par rapport à \mathbf{Y} et Θ , de la distribution $P(\mathbf{Y}, \mathbf{d} | \Theta)$ mais est extrêmement difficile à implanter sous cette forme. Ce critère considère les *hyperparamètres* Θ au même niveau que le processus de déformation \mathbf{Y} . La résolution de ce problème d'estimation à deux variables peut être menée de manière itérative par une méthode de type EM ("*Expectation-Maximization*"). Cette méthode s'avère toutefois coûteuse, car itérative et convergeant lentement. L'approche bayésienne nous permet d'envisager plusieurs autres voies pour développer un estimateur des hyperparamètres. L'estimateur du *Maximum de Vraisemblance Marginale* (MVM) défini par :

$$\Theta^* = \arg \max_{\Theta} \int_{\mathbf{Y}} P(\mathbf{Y}, \mathbf{d} | \Theta) d\mathbf{Y}. \quad [24]$$

conduit (moyennant une hypothèse simplificatrice pertinente) à un algorithme d'estimation peu coûteux car non itératif (contrairement à l'approche EM).

L'estimée Θ^* est exploitée dans une seconde étape pour l'estimation au sens du MAP de \mathbf{Y} :

$$\mathbf{Y}^* = \arg \max_{\mathbf{Y}} P(\mathbf{Y} | \mathbf{d}, \Theta^*). \quad [25]$$

Estimation de Θ Nous proposons ici un estimateur de Θ permettant d'approcher l'estimateur MVM (équation 24). D'après la règle de Bayes, nous pouvons écrire :

$$P(\mathbf{Y}, \mathbf{d} | \Theta) = P(\mathbf{d} | \mathbf{Y}, \Theta) P(\mathbf{Y} | \Theta). \quad [26]$$

Une hypothèse raisonnable consiste à supposer que le processus aléatoire δ est de variance faible ($\sigma_i^2 \rightarrow 0, \forall i$) : \mathbf{Y} reste très "concentré" autour de $\mathbf{X}(\Theta)$. Dans ces conditions, la loi gaussienne de \mathbf{Y} peut être approchée par une masse de Dirac :

$$P(\mathbf{Y} | \Theta) \approx \delta(\mathbf{Y} - \mathbf{X}(\Theta)) \quad [27]$$

où $\delta(\cdot)$ désigne la masse de Dirac. Il vient alors :

$$\begin{aligned} \Theta^* &= \arg \max_{\Theta} \int_{\mathbf{Y}} P(\mathbf{d} | \mathbf{Y}, \Theta) P(\mathbf{Y} | \Theta) d\mathbf{Y} \\ &= \arg \max_{\Theta} \int_{\mathbf{Y}} P(\mathbf{d} | \mathbf{Y}, \Theta) \delta(\mathbf{Y} - \mathbf{X}(\Theta)) d\mathbf{Y} \\ &= \arg \max_{\Theta} P(\mathbf{d} | \mathbf{Y} = \mathbf{X}(\Theta)), \end{aligned} \quad [28]$$

ce qui conduit à l'estimateur suivant :

$$\Theta^* = \arg \min_{\Theta} E_d(\mathbf{X}(\Theta), \mathbf{d}). \quad [29]$$

Estimation de \mathbf{Y} La solution au sens du MAP est obtenue selon le critère défini par l'équation 25, qui conduit à :

$$\mathbf{Y}^* = \arg \min_{\mathbf{Y}} E_{\Theta^*}(\mathbf{Y}, \mathbf{d}). \quad [30]$$

La recherche d'une solution (c'est-à-dire d'une "bonne" segmentation de l'image) passe donc par l'estimation des hyperparamètres du modèle au sens du MVM et la minimisation de l'énergie globale, selon l'approche définie précédemment. Ces optimisations (non linéaires) sont menées différemment en fonction des données observables.

Segmentation au sens du mouvement Pour être insensible à la configuration initiale du modèle déformable, des algorithmes de minimisation globale sont appliqués, dans ce cas, sur la *première* image de la séquence, pour déterminer les hyperparamètres Θ^* . Les déformations locales (estimation de \mathbf{Y}^*) sont ensuite obtenues par un algorithme déterministe rapide de type ICM [BES 86]. Ceci conduit à une segmentation robuste et évite de recourir à une initialisation manuelle du modèle (proche de la segmentation souhaitée). L'optimisation globale utilisée [KER 94a] s'appuie sur un algorithme de recuit simulé basé sur l'échantillonneur de Gibbs [GEM 84] et peut être mise en œuvre dans des conditions calculatoires acceptables puisque le nombre de paramètres à optimiser est réduit (les temps de calcul sont de 3 à 4 mn CPU sur une station SUN/SPARC10 pour la première image - de taille 256×256). Notons que les images suivantes de la séquence sont entièrement traitées grâce à l'algorithme ICM, en tirant profit d'une initialisation proche de la solution, fournie par une procédure de suivi par filtrage récursif décrite en section 5. Les temps de calcul tombent alors à 45 s par image. Ces résultats sont encourageants car des temps de calcul de quelques heures avaient été évoqués par Grenander *et al.* pour traiter une image de taille 128 × 70 [GRE 91] par un modèle statistique. Bien entendu, l'approche locale proposées par Cootes *et al.* est plus rapide, mais requiert une initialisation manuelle du modèle [COO 94b]. Récemment, une procédure d'optimisation globale (*algorithme génétique* [HIL 92]) a été proposée pour éviter d'initialiser le modèle déformable dans une image [COO 94a], mais les temps de calcul (pour ces techniques, en général comparables à ceux des méthodes que nous décrivons ici) ne sont pas donnés par les auteurs.

La robustesse de l'optimisation globale aux conditions initiales, au bruit ainsi qu'à des observations partielles ou manquantes est illustrée figures 3 et 4. Dans chacune de ces figures, des champs d'observations représentant des gradients temporels seillés bruités ont été construits à partir de silhouettes réelles de mains. Différents rapports signal-à-bruit, ainsi que des occlusions partielles ont été simulés. Dans tous les cas, la forme prototype est initialisée

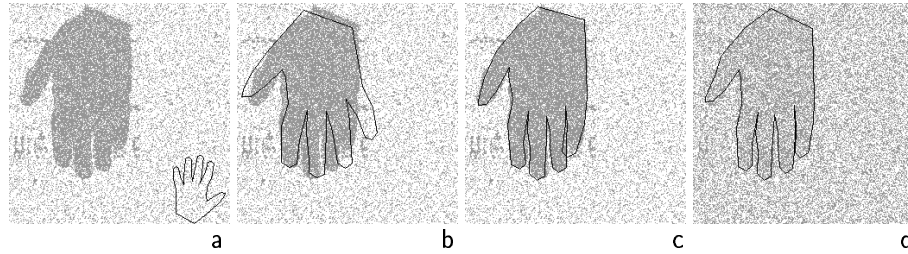


Figure 3. *Initialisation et convergence du modèle déformable (voir texte)*

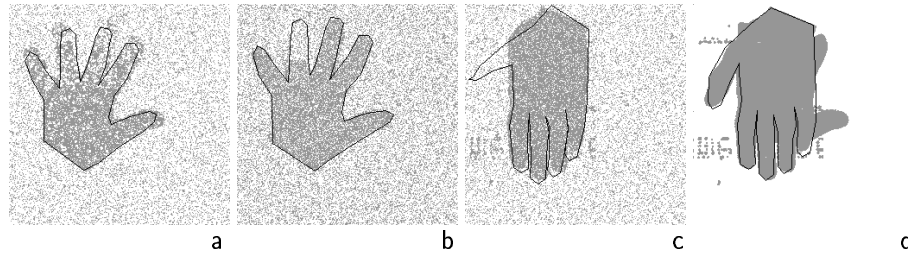


Figure 4. *Segmentations avec données manquantes et occlusions partielles*

de manière quelconque dans l'image (dans la partie inférieure droite de l'image – voir figure 3a).

La figure 3b présente le résultat de l'estimation des hyperparamètres Θ en prenant uniquement en compte les hyperparamètres décrivant les similitudes planes (rotation 2D, translation et homothétie). L'estimation complète de Θ (incluant les modes de déformation issus de la décomposition KL) conduit à la segmentation de la figure 3c (jugée satisfaisante). Un résultat similaire est obtenu figure 3d avec un rapport signal-à-bruit particulièrement bas. La figure 4 présente différents résultats de segmentations (en partant toujours de la même configuration initiale du modèle), dans des situations considérées généralement comme difficiles, d'occlusions partielles (doigts bruités ou supprimés - figure 4-a-b-c) ou de superposition de deux structures (figure 4d). Ces résultats démontrent l'intérêt de l'approche statistique et de l'optimisation globale dans ces situations adverses. Elle fournit en particulier une robustesse appréciable aux occlusions ou disparitions partielles et momentanées de la structure à segmenter (un exemple sur une séquence réelle de mains est présenté en section 6).

Segmentation au sens des contours photométriques On adopte dans ce cas un schéma d'optimisation similaire au précédent. Soulignons toutefois qu'une minimisation globale par des techniques stochastiques n'est plus appropriée ici : le critère de segmentation n'est valide que dans une *région d'intérêt*

dans le voisinage de la solution désirée [COO 94b, KAS 87, NAS 93, STA 92]. Une optimisation par descente de gradient, s'appuyant sur le calcul analytique des dérivées partielles de la fonction d'énergie, est utilisée dans ce contexte [KER 95, STA 92]. Dans les exemples que nous présentons, pour ce modèle d'observation (voir section 6), l'initialisation est réalisée de façon manuelle sur la première image de la séquence, le suivi par filtrage récursif assurant des initialisations satisfaisantes pour les images suivantes. La procédure de suivi est décrite dans la section qui suit.

5. Suivi du modèle déformable sur une séquence longue

La procédure de segmentation décrite dans la section précédente peut être complétée avec profit par un suivi temporel des hyperparamètres du modèle, utilisant un filtrage récursif (filtre de Kalman). Un filtrage récursif des hyperparamètres $\Theta = (\mathbf{M}(k, \theta), \mathbf{T}, \mathbf{b})$ du modèle permet d'assurer un suivi des grands déplacements de la structure déformable. Il garantit, mieux qu'un suivi de primitives discrètes (points, segments [DER 90], arcs paramétrés [CUR 92]) une segmentation robuste et permet de faciliter l'optimisation des hyperparamètres par une meilleure initialisation d'un instant à l'autre. Un recours à des algorithmes d'optimisation locale est alors possible, conduisant à une réduction significative du temps de calcul.

Un suivi de paramètres de mouvements *rigides* associés à des régions homogènes a été proposé dans [MEY 94], et a été décrit récemment pour des régions délimitées par des contours actifs dans [BAS 94]. Cootes [COO 94a] propose également de réaliser une projection d'un modèle déformable dans le sens du mouvement, en utilisant une équation de prédiction heuristique du premier ordre sur les paramètres de déformation. Cette approche [COO 94a] n'intègre toutefois aucune modélisation du bruit, pas de filtrage *a posteriori* sur les paramètres du modèle, ni de détection de rupture du suivi, comme nous le proposons ici, et s'avère donc peu fiable [COO 94a].

5.1. Équations d'évolution temporelle du modèle

On construit un estimateur récursif linéaire des hyperparamètres Θ du modèle. Le processus de déformation locale n'est pas pris en compte dans les équations du filtre car trop sensible au bruit inhérent aux observations. Les mesures considérées dans ce filtre, sont le résultat de l'estimation, au sens du MVM, des paramètres globaux de déformation, comme nous l'avons exposé en section 4. Dans le cadre du traitement d'une séquence longue, la représentation est notée à l'instant t par :

$$\mathbf{X}_t = \mathbf{M}_t [\bar{\mathbf{x}} + \Phi \mathbf{b}_t] + \mathbf{T}_t. \quad [31]$$

Le modèle déformable global et ses dérivées temporelles développées au 1^{er} ordre et au 2^{ème} ordre s'expriment par :

$$\begin{cases} \mathbf{X}_t = \mathbf{M}_t [\bar{\mathbf{x}} + \Phi \mathbf{b}_t] + \mathbf{T}_t, \\ \dot{\mathbf{X}}_t = \dot{\mathbf{M}}_t [\bar{\mathbf{x}} + \Phi \mathbf{b}_t] + \mathbf{M}_t \Phi \dot{\mathbf{b}}_t + \dot{\mathbf{T}}_t, \\ \ddot{\mathbf{X}}_t = \ddot{\mathbf{M}}_t [\bar{\mathbf{x}} + \Phi \mathbf{b}_t] + \mathbf{M}_t \Phi \ddot{\mathbf{b}}_t + 2\dot{\mathbf{M}}_t \Phi \dot{\mathbf{b}}_t + \ddot{\mathbf{T}}_t. \end{cases} \quad [32]$$

L'évolution temporelle du modèle s'exprime au moyen d'un développement au 2^{ème} ordre de \mathbf{X}_t et de $\dot{\mathbf{X}}_t$:

$$\begin{cases} \mathbf{X}_{t+\Delta t} = \mathbf{X}_t + \Delta t \dot{\mathbf{X}}_t + \frac{\Delta t^2}{2} \ddot{\mathbf{X}}_t, \\ \dot{\mathbf{X}}_{t+\Delta t} = \dot{\mathbf{X}}_t + \Delta t \ddot{\mathbf{X}}_t. \end{cases} \quad [33]$$

Par identification, on dérive les équations d'évolution temporelle des hyperparamètres du modèle (les termes croisés $\dot{\mathbf{M}}_t \Phi \mathbf{b}_t$ sont supposés négligeables, voir [KER 95]) :

$$\begin{cases} \mathbf{M}_{t+\Delta t} = \mathbf{M}_t + \Delta t \dot{\mathbf{M}}_t + \frac{\Delta t^2}{2} \ddot{\mathbf{M}}_t, \\ \dot{\mathbf{M}}_{t+\Delta t} = \dot{\mathbf{M}}_t + \Delta t \ddot{\mathbf{M}}_t, \\ \mathbf{T}_{t+\Delta t} = \mathbf{T}_t + \Delta t \dot{\mathbf{T}}_t + \frac{\Delta t^2}{2} \ddot{\mathbf{T}}_t, \\ \dot{\mathbf{T}}_{t+\Delta t} = \dot{\mathbf{T}}_t + \Delta t \ddot{\mathbf{T}}_t, \\ \mathbf{b}_{t+\Delta t} = \mathbf{b}_t + \Delta t \dot{\mathbf{b}}_t + \frac{\Delta t^2}{2} \tilde{\Phi}_{t+\Delta t}^{-1} \mathbf{M}_t \Phi \ddot{\mathbf{b}}_t, \\ \dot{\mathbf{b}}_{t+\Delta t} = [\mathbf{I} + \frac{3\Delta t}{2} \tilde{\Phi}_{t+\Delta t}^{-1} (\dot{\mathbf{M}}_t \Phi - \dot{\mathbf{M}}_{t+\Delta t} \Phi)] \dot{\mathbf{b}}_t \\ + \left[\Delta t \mathbf{I} - \frac{\Delta t^2}{2} \tilde{\Phi}_{t+\Delta t}^{-1} \dot{\mathbf{M}}_{t+\Delta t} \Phi \right] \tilde{\Phi}_{t+\Delta t}^{-1} \mathbf{M}_t \Phi \ddot{\mathbf{b}}_t \end{cases} \quad [34]$$

avec :

$$\tilde{\Phi}_{t+\Delta t}^{-1} = [(\mathbf{M}_{t+\Delta t} \Phi)^T (\mathbf{M}_{t+\Delta t} \Phi)]^{-1} (\mathbf{M}_{t+\Delta t} \Phi)^T. \quad [35]$$

À ces équations d'évolution, on associe un filtre Kalman portant sur \mathbf{M} , \mathbf{T} et \mathbf{b} (les calculs sont détaillés dans [KER 95]).

Équations du filtre de Kalman L'évolution dynamique du système est décrite par :

$$\mathbf{s}_{t+\Delta t} = \mathbf{A}_t \mathbf{s}_t + \mathbf{B}_t \mathbf{u}_t + \boldsymbol{\xi}_t \quad [36]$$

où \mathbf{s}_t est le vecteur d'état, \mathbf{u}_t est une entrée déterministe, \mathbf{A}_t est la matrice de transition et $\boldsymbol{\xi}_t$ est un bruit blanc gaussien centré de matrice covariance $\mathbf{Q}_t = \mathbb{E} [\boldsymbol{\xi}_t \boldsymbol{\xi}_t^T]$. La mesure \mathbf{w}_t est une fonction linéaire du vecteur d'état :

$$\mathbf{w}_t = \mathbf{H}_t \mathbf{s}_t + \boldsymbol{\eta}_t \quad [37]$$

où $\boldsymbol{\eta}_t$ correspond à un bruit blanc gaussien centré de matrice de covariance $\mathbf{V}_t = \mathbb{E} [\boldsymbol{\eta}_t \boldsymbol{\eta}_t^T]$ (dans notre cas, \mathbf{H}_t est la matrice identité \mathbf{I} et $\mathbf{V}_t = \sigma_\eta^2 \mathbf{I}$).

Si $\widehat{\mathbf{s}}_{t_2 | t_1}$ représente l'estimée linéaire optimale de \mathbf{s}_{t_2} connaissant les mesures jusqu'à t_1 ($t_1 \leq t_2$) et si $\mathbf{P}_{t_2 | t_1}$ désigne la matrice de covariance de l'erreur associée, les équations de remise à jour du filtre s'écrivent :

$$\widehat{\mathbf{s}}_{t|t} = \widehat{\mathbf{s}}_{t|t-\Delta t} + \mathbf{K}_t [\mathbf{w}_t - \mathbf{H} \widehat{\mathbf{s}}_{t|t-\Delta t}], \quad [38]$$

$$\mathbf{P}_{t|t} = [\mathbf{I} - \mathbf{K}_t \mathbf{H}] \mathbf{P}_{t|t-\Delta t} \quad [39]$$

où le gain de Kalman est défini par :

$$\mathbf{K}_t = \mathbf{P}_{t|t-\Delta t} \mathbf{H}^T [\mathbf{H} \mathbf{P}_{t|t-\Delta t} \mathbf{H}^T + \mathbf{V}_t]^{-1}. \quad [40]$$

Les équations de prédiction sont définies par :

$$\widehat{\mathbf{s}}_{t+\Delta t|t} = \mathbf{A}_t \widehat{\mathbf{s}}_{t|t} + \mathbf{B}_t \mathbf{u}_t, \quad [41]$$

$$\mathbf{P}_{t+\Delta t|t} = \mathbf{A}_t \mathbf{P}_{t|t} \mathbf{A}_t^T + \mathbf{Q}_t. \quad [42]$$

Nous avons développé deux filtres : un filtre à vitesse constante et un filtre reposant sur une estimée en ligne de la vitesse d'évolution des paramètres (filtre "à vitesse instantanée").

5.2. Filtre à vitesse constante

Soit \mathbf{z}_t représentant \mathbf{T}_t , \mathbf{b}_t ou \mathbf{M}_t . On définit le vecteur d'état $\mathbf{s}_t = (\mathbf{z}_t \ \dot{\mathbf{z}}_t)^T$ et $\mathbf{u}_t = 0$. Les matrices de transition respectives pour les paramètres \mathbf{M} et \mathbf{T} d'une part, et \mathbf{b} d'autre part sont données par :

$$\mathbf{A}_t = \begin{pmatrix} \mathbf{I} & \Delta t \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \text{ et } \mathbf{A}_t = \begin{pmatrix} \mathbf{I} & \Delta t \mathbf{I} \\ \mathbf{0} & \mathbf{I} + \Delta t \tilde{\boldsymbol{\Phi}}_{t+\Delta t}^{-1} [\dot{\mathbf{M}}_t - \dot{\mathbf{M}}_{t+\Delta t}] \boldsymbol{\Phi} \end{pmatrix}. \quad [43]$$

Les vecteurs de bruit d'état sont définis respectivement pour \mathbf{M} et \mathbf{T} d'une part, et \mathbf{b} d'autre part :

$$\boldsymbol{\xi}_t = \begin{pmatrix} \frac{\Delta t^2}{2} \ddot{\mathbf{z}}_t \\ \Delta t \dot{\mathbf{z}}_t \end{pmatrix}, \quad \boldsymbol{\xi}_t = \begin{pmatrix} \frac{\Delta t^2}{2} \tilde{\boldsymbol{\Phi}}_{t+\Delta t}^{-1} \mathbf{M}_t \boldsymbol{\Phi} \ddot{\mathbf{z}}_t \\ \left[\Delta t \mathbf{I} - \frac{\Delta t^2}{2} \tilde{\boldsymbol{\Phi}}_{t+\Delta t}^{-1} \dot{\mathbf{M}}_{t+\Delta t} \boldsymbol{\Phi} \right] \tilde{\boldsymbol{\Phi}}_{t+\Delta t}^{-1} \mathbf{M}_t \boldsymbol{\Phi} \dot{\mathbf{z}}_t \end{pmatrix}. \quad [44]$$

On remarque que la prédiction et la remise à jour de \mathbf{b} dépend de $\dot{\mathbf{M}}$ (couplage entre hyperparamètres). Nous avons adopté une formulation linéaire simple du filtre en découplant les filtres sur les hyperparamètres \mathbf{M} , \mathbf{T} et \mathbf{b} et en approchant $\dot{\mathbf{M}}_t$ par $\dot{\mathbf{M}}_{t|t}$ et $\dot{\mathbf{M}}_{t+\Delta t}$ par $\dot{\mathbf{M}}_{t+\Delta t|t}$. Cette approximation donne des résultats satisfaisants en pratique. L'enchaînement des étapes de

la procédure de suivi est le suivant : partant de l'estimée obtenue à l'instant t , l'étape de prédiction permet de positionner le modèle à l'instant $t + \Delta t$, proche de la solution désirée. Les mesures à l'instant $t + \Delta t$ (estimation au sens du MVM des hyperparamètres du modèle, voir section 4) sont obtenues par une technique d'optimisation locale rapide (ICM). Enfin, l'état est remis à jour selon l'équation 38. L'état ainsi filtré est le résultat final de l'estimation des déformations à l'instant $t + \Delta t$.

Ce filtre (de même que le filtre décrit dans la section suivante) est complété par une procédure de détection de rupture, permettant une réinitialisation du suivi lorsque le mouvement observé ne répond plus au modèle cinématique sous-jacent. Ce test est décrit de façon détaillée dans [KER 95]. En pratique le nombre de ruptures reste faible.

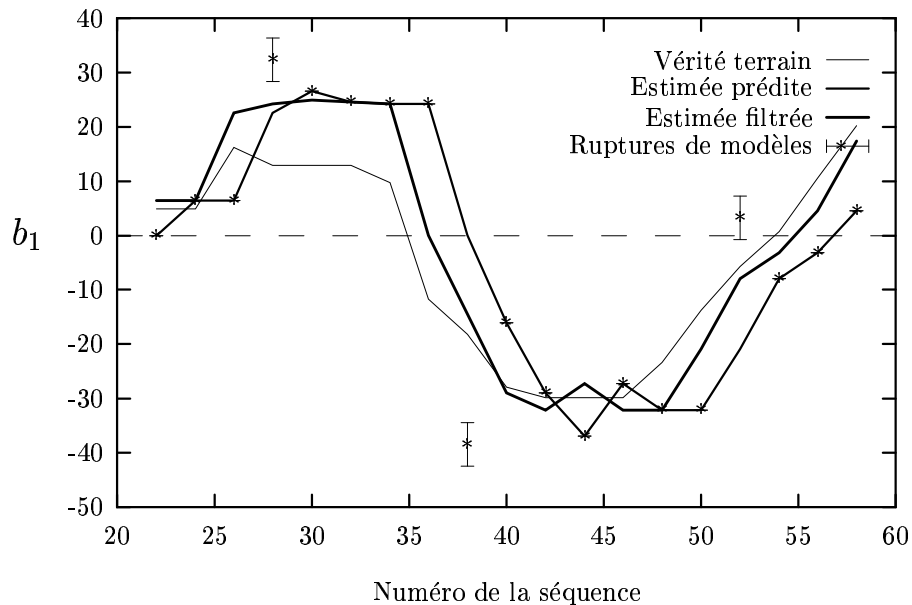


Figure 5. Suivi de l'amplitude modale b_1

À titre d'illustration, la figure 5 présente l'évolution de la prédiction et de la remise à jour de l'amplitude modale b_1 sur une séquence synthétique d'une main en mouvement pour laquelle on connaît la "vérité terrain". Les détections de rupture du modèle sont signalées par des barres verticales: le filtre est alors réinitialisé et une optimisation globale (recuit simulé) est utilisée sur l'image correspondante. Cette technique, contrairement à l'approche heuristique de prédiction proposée dans [COO 94a], assure un suivi robuste en toutes circonstances dans la séquence. D'autres résultats de suivi sur séquences longues sont consignés dans [KER 95] et en section 6.

5.3. Filtre “à vitesse instantanée”

Le vecteur d’état est défini ici par $\mathbf{s}_t = \mathbf{z}_t$ et on introduit une entrée déterministe (vitesse instantanée) $\mathbf{u}_t = \dot{\mathbf{z}}_t$ calculée par *différences finies* ou à l’aide du flot optique calculé le long de la frontière du modèle au cours du temps [KER 95]. Ce choix conduit à des estimations plus bruitées mais s’avère plus robuste lorsque la structure présente un comportement cinématique complexe (évolution du modèle à vitesse non constante).

Les matrices \mathbf{A}_t et \mathbf{B}_t sont identiques pour tous les paramètres :

$$\mathbf{A}_t = \mathbf{I} \quad \text{et} \quad \mathbf{B}_t = \Delta t \mathbf{I}. \quad [45]$$

Les vecteurs de bruit d’état correspondent respectivement pour \mathbf{M} et \mathbf{T} d’une part, et \mathbf{b} d’autre part à :

$$\boldsymbol{\xi}_t = \frac{\Delta t^2}{2} \ddot{\mathbf{z}}_t \quad \text{et} \quad \boldsymbol{\xi}_t = \frac{\Delta t^2}{2} \tilde{\boldsymbol{\Phi}}_{t+\Delta t}^{-1} \mathbf{M}_t \boldsymbol{\Phi} \ddot{\mathbf{z}}_t. \quad [46]$$

$\boldsymbol{\xi}_t$ intègre dans sa modélisation toutes les dérivées d’ordre supérieur assimilées à un bruit blanc. La procédure de suivi mise en œuvre est identique à celle présentée dans la section 5.2.

6. Résultats expérimentaux

L’approche proposée est illustrée ici sur des applications variées correspondant à des classes d’images et à des modèles de déformations différents.

Une première étude de cas concerne l’extraction et le suivi d’une main en mouvement dans un environnement texturé comportant des occlusions partielles (figure 6). La séquence se compose de plus d’une centaine de trames. L’apprentissage a été effectué avec un modèle de main de 30 points à partir d’une séquence d’apprentissage de 22 images, différentes de celles qui sont traitées ici. L’initialisation du modèle dans la première image de la séquence est aléatoire. Une procédure d’optimisation globale est utilisée pour cette image, les images suivantes étant traitées par des techniques d’optimisation déterministe rapides, en couplage avec le suivi par filtrage de Kalman des hyperparamètres du modèle. Le suivi s’est révélé fiable sur l’ensemble de la séquence, sans “décrochage” du modèle ou segmentation erronée. La figure 6 présente trois images (non consécutives) extraites de la séquence. On notera en particulier (figure 6c) l’absence de “décrochage” du modèle lorsque la main disparaît partiellement derrière une boîte disposée en avant-plan. La figure 7 présente les observations correspondantes sur lesquelles s’appuie la segmentation (on pourra noter le caractère bruité des gradients temporels utilisés ici ainsi que les observations manquantes figure 7c). Le temps de calcul moyen mesuré sur la séquence complète est de 45s par image (taille : 256×256) sur station Sun Sparc10.

Pour visualiser l’apport du suivi par filtrage de Kalman, la figure 8 présente le résultat de la segmentation sans la procédure de suivi décrite dans la section précédente. L’estimée finale au sens du MVM, à l’instant $t - \Delta t$ est

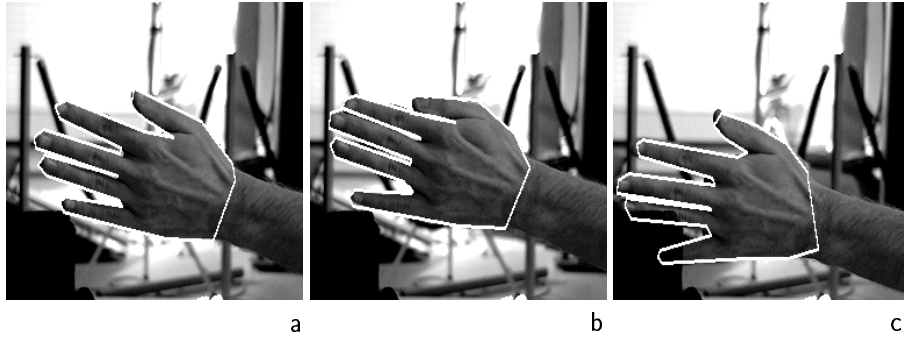


Figure 6. *Segmentation et suivi d'une main en mouvement sur un fond texturé*

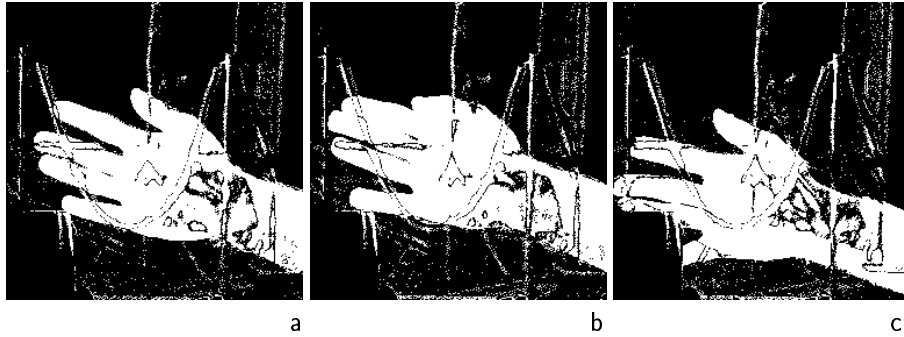


Figure 7. *Champs d'observations (gradients temporels) d_s utilisés dans la segmentation de la main en mouvement*

présentée figure 8a. Sur la figure 8b, l'estimée finale est simplement projetée à l'instant suivant et sert d'estimée initiale à l'instant t . L'optimisation du processus de déformation par des algorithmes déterministes rapides (ICM) ne garantit pas dans ce cas un résultat satisfaisant (problèmes de segmentation sur le petit doigt dû à la présence de minima locaux dans le critère – voir figure 8c). Une optimisation des paramètres du modèle par des techniques stochastiques conduit à une segmentation satisfaisante (mais coûteuse!) de la main (voir figure 8d – 4 mn de temps CPU). La procédure de suivi permet d'obtenir une meilleure initialisation et donc de réduire notablement le temps de calcul par utilisation d'un algorithme d'optimisation déterministe (moins d'une minute de temps CPU) tout en maintenant une qualité de segmentation analogue. Les résultats présentés figure 9 ont été obtenus par un suivi (avec un modèle à vitesse constante) des paramètres de déformation (des résultats similaires sont observés avec le filtre à vitesse instantanée). La figure 9a présente l'estimée obtenue à l'instant $t - \Delta t$. La configuration du modèle prédite vers l'instant t apparaît figure 9b. La figure 9c correspond à la "mesure" exploitée par le

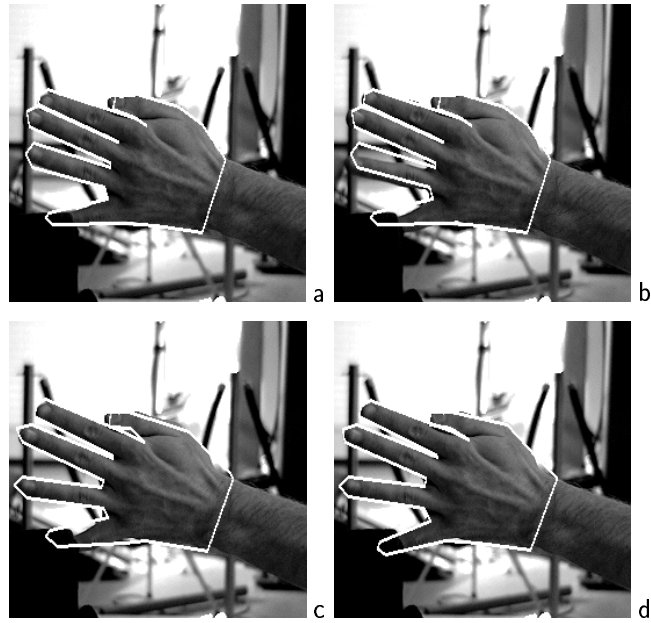


Figure 8. *Extraction d'une main en mouvement - sans procédure de suivi*

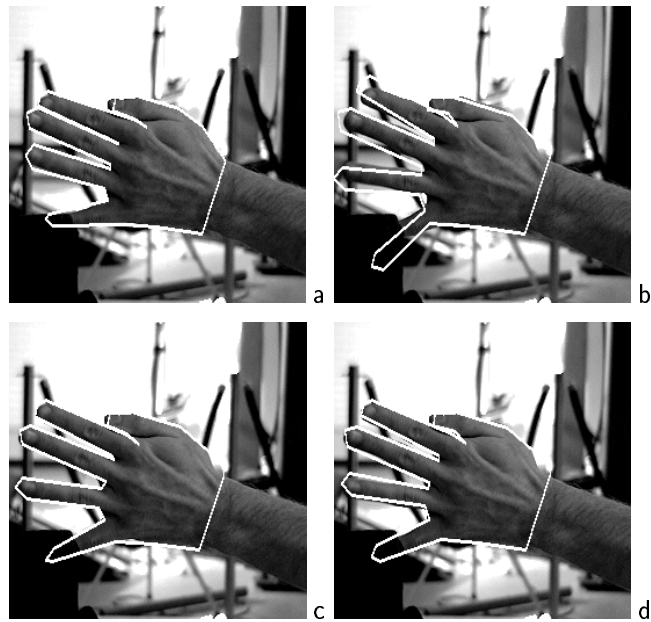


Figure 9. *Extraction d'une main en mouvement - avec procédure de suivi*

filtre à ce même instant (il s'agit d'une optimisation par algorithme ICM des hyperparamètres du modèle). Enfin la figure 9d présente la configuration correspondant à la remise à jour de l'état du filtre (prise en compte de la mesure). On constate qu'on obtient une bonne segmentation finale, sans recours à une coûteuse optimisation globale pour l'estimation des déformations.

Une seconde étude de cas concerne le suivi de lèvres en mouvement dans des séquences de visages (figure 10). Les gradients spatiaux représentent ici la seule information exploitée pour réaliser la segmentation. Dans ce cas, le filtre à "vitesse instantanée" s'est révélé plus robuste, car mieux adapté pour analyser le comportement cinématique complexe de la structure analysée. Le modèle de lèvres utilisé (composé de 29 points interpolés par une B-spline cubique) a été identifié préalablement sur une séquence d'apprentissage de 10 images. L'initialisation de la première image est manuelle. Les résultats de segmentation et de suivi sont satisfaisants eu égard au caractère bruité des observations utilisées (voir figure 2 un exemple de champ d'observations) et très comparables à ceux présentés dans [YUI 92]. L'énergie à minimiser (équation 16) est cependant beaucoup plus simple dans notre cas et l'aspect dynamique a été pris en compte. Les temps de calcul, sont ici de 15s par image (taille des images : 256×256).

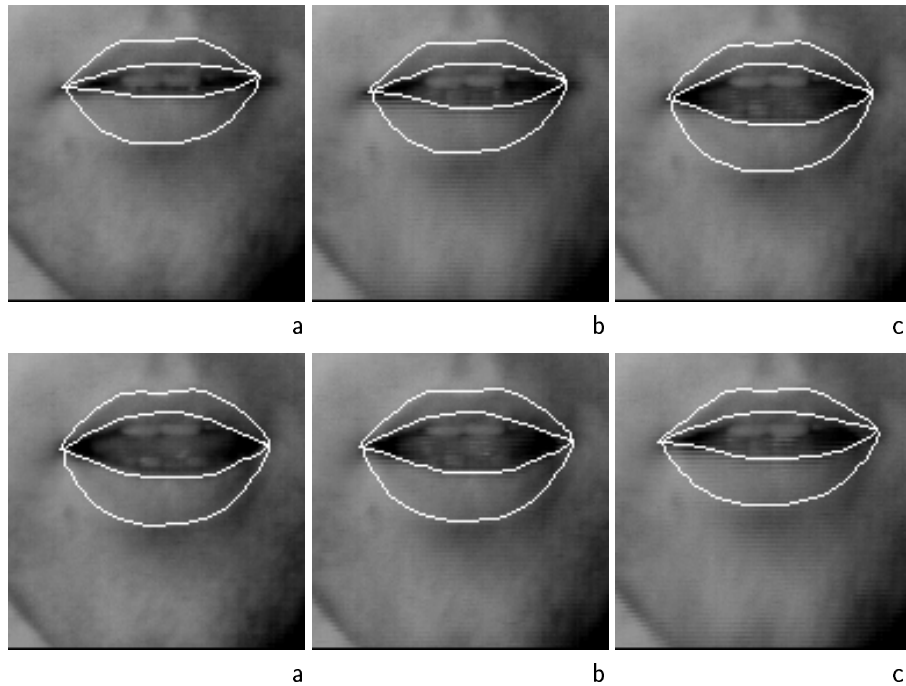


Figure 10. *Suivi du mouvement de lèvres*

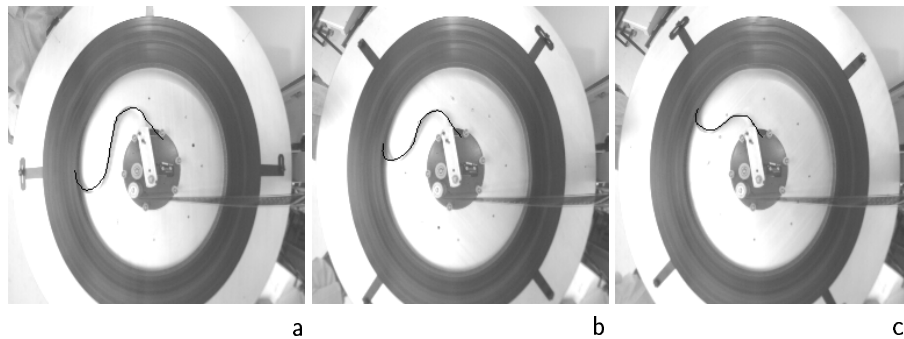


Figure 11. *Suivi de la forme d'une boucle de déroulement (images du Laboratoire d'Automatique des Arts et Métiers)*

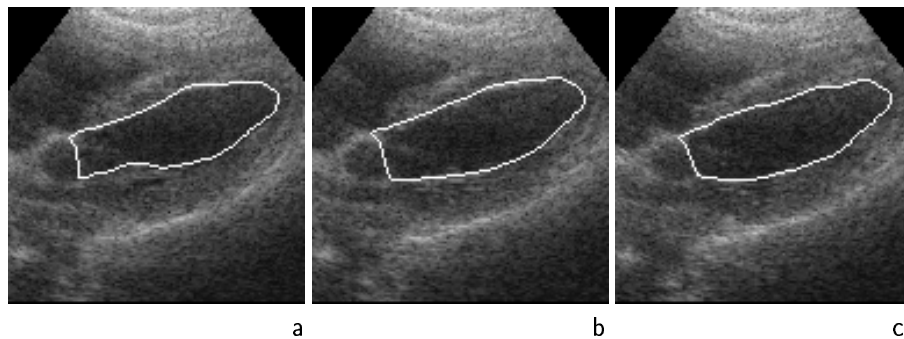


Figure 12. *Suivi du muscle cardiaque (ventricule gauche) en imagerie échocardiographique*

Une troisième étude de cas plus originale, présentée figure 11, illustre la diversité des champs d'applications possibles de cette technique de segmentation : il s'agit ici de suivre la boucle de déroulement d'un film cinématographique afin d'assurer un asservissement du moteur entraînant la bobine [ZUG 95]. Le modèle déformable, particulièrement simple dans ce cas, se compose de 16 points interpolés par une B-spline. L'initialisation est ici également manuelle, seuls les gradients spatiaux étant utilisés pour le suivi. Des résultats très encourageants ont été obtenus (figure 11), avec des temps de calcul relativement courts de l'ordre de 5 secondes par image (en raison de la simplicité du modèle). Le modèle de boucle dont le suivi est assuré ici apparaît en surimpression (en noir), près du centre de la bobine.

Enfin le suivi s'appuyant sur un modèle statistique déformable s'est également révélé performant dans un problème classique de segmentation en imagerie médicale. La figure 12 présente le suivi, au cours d'un cycle, du muscle

cardiaque (ventricule gauche) dans une séquence échocardiographique. Les observations utilisées sont là encore liées aux gradients spatiaux (le modèle de muscle cardiaque se compose de 20 points). Les temps de calcul sont de l'ordre de 25s par image (taille des images: 256×256). Les segmentations obtenues dans ce cas ont été jugées satisfaisantes par des experts du monde médical.

7. Conclusion

Dans cet article nous avons présenté un cadre de modélisation statistique pour la représentation, la segmentation et le suivi d'objets déformables dans une séquence d'images. L'approche repose sur la définition d'une structure prototype ("template déformable") à laquelle sont appliquées des déformations hiérarchisées (similitudes planes, modes de déformations globaux et déformations locales). Une décomposition modale des déformations, obtenue par apprentissage sur une population représentative, permet une adaptation du modèle à la classe particulière des objets d'intérêt et conduit à une représentation parcimonieuse des déformations globales avec un nombre réduit de paramètres (de l'ordre de la dizaine). Associé à des techniques d'optimisation globales et locales, ce modèle permet la segmentation de structures déformables dans des séquences d'images longues. L'utilisation d'un filtrage récursif sur les paramètres du modèle assure par ailleurs un suivi de la structure au cours du temps et une bonne propagation des initialisations. Cette approche a été illustrée sur des séquences d'images réelles, pour quatre classes différentes de structures déformables bidimensionnelles. Elle s'est révélée robuste, même dans des situations jugées habituellement difficiles pour les approches de type "contour actif" (observations bruitées, données manquantes, occlusions partielles entre objets).

Soulignons que le cadre de modélisation et les solutions algorithmiques proposées ici sont suffisamment généraux pour s'appliquer à une large classe de structures déformables. Une extension à des structures tridimensionnelles est ainsi à l'étude. La nécessité de recourir à un apprentissage hors-ligne peut également être largement levée, comme le montre une étude préliminaire sur ce point [KER 95]. On aboutit ainsi à des algorithmes presque entièrement guidés par les données. La disponibilité d'outils de segmentation et de suivi performants pour des structures déformables variées, laisse ainsi entrevoir des perspectives prometteuses dans le domaine de l'interprétation qualitative et quantitative de phénomènes déformables complexes.

8. Bibliographie

- [BAS 94] B. BASCLE, P. BOUTHEMY, N. DERICHE, et F. MEYER. « Tracking complex primitives in an image sequence ». *Proc. of Int. Conf. Pattern Recognition*, pages 426–431, Jerusalem, Israel, octobre 1994.
- [BER 91] M.O BERGER. « *Les contours actifs : modélisation, comportement et convergence* ». Thèse de doctorat, Institut National Polytechnique de Lorraine, INRIA Lorraine, Nancy, février 1991.

- [BES 86] J. BESAG. « On the statistical analysis of dirty pictures ». *J. Royal Statist. Soc. B*, 48(3):259–302, 1986.
- [BLA 92] A. BLAKE et A. YUILLE. *Active Vision*. MIT-Press, 1992.
- [BLA 93] A. BLAKE, R. CURWEN, et A. ZISSERMAN. « A Framework for Spatiotemporal Control in the Tracking of Visual Contours ». *Int. J. Computer Vision*, 11(2):127–145, octobre 1993.
- [COH 92] I. COHEN, L.D. COHEN, et N. AYACHE. « Using deformable surfaces to segment 3D images and infer differential structures ». *CVGIP: Image Understanding*, 56(2):242–263, septembre 1992.
- [COH 93] L.D. COHEN et I. COHEN. « Finite element methods for active contour models and balloons for 2D and 3D images ». *IEEE Trans. Pattern Anal. Machine Intell.*, 5(11):587–600, novembre 1993.
- [COO 94a] T.F. COOTES, A.H. HILL, C.J. TAYLOR, et J. HASLAM. « The use of active models for locating structures in medical images ». *Image and Vision Computing*, 12(6):355–365, juillet 1994.
- [COO 94b] T.F. COOTES, C.J. TAYLOR, D.H. COOPER, et J. GRAHAM. « Active shape models - Their training and application ». *CVGIP: Image Understanding*, 61(1):38–59, janvier 1994.
- [COO 95] T.F. COOTES et C.J. TAYLOR. « Combining point distribution models with shape models based on finite element analysis ». *Image and Vision Computing*, 13(5):403–409, juin 1995.
- [CUR 92] R. CURWEN et A. BLAKE. « Dynamic contours: real-time active splines ». *Active Vision*, pages 39–57, chapitre 3, MIT-Press, édité par A. Blake and A. Yuille, 1992.
- [DER 90] R. DERICHE et O. FAUGERAS. « Tracking Line Segments ». *Proc. of European Conf. Computer Vision*, pages 259–268, Antibes, avril 1990.
- [DON 88] G.W. DONOHOE, D.R. HUSH, et Ahmed. N.. « Change detection for target detection and classification in video sequences ». *Proc. of Int. Conf. Acoust., Speech, Signal Processing*, pages 1084–1087, New-York, USA, avril 1988.
- [GEM 84] S. GEMAN et D. GEMAN. « Stochastic relaxation, Gibbs distributions and the bayesian restoration of images ». *IEEE Trans. Pattern Anal. Machine Intell.*, 6(6):721–741, novembre 1984.
- [GRE 91] U. GRENANDER, Y. CHOW, et D.M. KEENAN. *Hands. A Pattern Theoretic Study of Biological Shapes*. Springer, 1991.
- [HER 92] I. HERLIN et N. AYACHE. « Features extraction and analysis methods for sequences of ultrasound images ». *Proc. of European Conf. Computer Vision*, pages 43–57, Santa Margherita, Italie, mai 1992.
- [HIL 92] A.H. HILL et C.J. TAYLOR. « Model based image interpretation using genetic algorithms ». *Image and Vision Computing*, 10(5):295–300, juin 1992.
- [IRA 92] M. IRANI, B. ROUSO, et S. PELEG. « Detecting and tracking multiple moving objects using temporal integration ». *Proc. of European Conf. Computer Vision*, pages 282–287, Santa Margherita, Italie, mai 1992.
- [KAS 87] M. KASS, A. WITKIN, et D. TERZOPOLOUS. « Snakes: Active Contour Models ». *Proc. of Int. Conf. Computer Vision*, pages 259–268, London, UK, juin 1987.
- [KER 94a] C. KERVRANN et F. HEITZ. « A hierarchical Statistical Framework for the Segmentation of Deformable Objects in Image Sequences ». *Proc. of IEEE Conf. Comp. Vision Pattern Rec.*, pages 724–728, Seattle, USA, juin 1994.

- [KER 94b] C. KERVRANN et F. HEITZ. « Robust tracking of stochastic deformable models in image sequences ». *Proc. of Int. Conf. Image Processing*, pages 88–92, Seattle, USA, novembre 1994.
- [KER 95] C. KERVRANN. « Modèles statistiques pour la segmentation et le suivi de structures déformables bidimensionnelles dans une séquence d'images ». Thèse de doctorat, Université Rennes I, novembre 1995.
- [LAI 94] K.L. LAI et R.T. CHIN. « Deformable contours: modeling and extraction ». *Proc. of IEEE Conf. Comp. Vision Pattern Rec.*, pages 601–608, Seattle, USA, juin 1994.
- [LEI 90] F. LEITNER, I. MARQUE, S. LAVALLEE, et Cinquin P.. « Dynamic segmentation. Finding the edge with snake splines ». *Proc. of Int. Conf. on Curves and Surfaces*, pages 1–4, Chamonix, juin 1990.
- [LIP 90] P. LIPSON, A.L. YUILLE, D. O'KEEFE, J. CAVANAUGH, J. TAAFLE, et D. ROSENTHAL. « Deformable templates for feature extraction from medical images ». *Proc. of European Conf. Computer Vision*, pages 413–417, Antibes, avril 1990.
- [MAR 92] K.V. MARDIA et T.F. HAINSWORTH. « Deformable templates in image sequences ». *Proc. of Int. Conf. Pattern Recognition*, pages 132–135, La Haye, Pays-Bas, septembre 1992.
- [MAR 94] J. MARTIN, A. PENTLAND, et R. KIKINIS. « Shape analysis of brain structures using physical and experimental modes ». *Proc. of IEEE Conf. Comp. Vision Pattern Rec.*, pages 752–755, Seattle, USA, juin 1994.
- [MEY 94] F. MEYER et P. BOUTHEMY. « Region-based tracking using affine motion models in long image sequences ». *CVGIP: Image Understanding*, 60(2):119–140, septembre 1994.
- [MOH 93] A. MOHAMMAD DJAFARI. « Estimation des hyperparamètres dans une approche bayésienne de la résolution des problèmes inverses linéaires ». *Actes du Colloque GRETSI*, pages 201–204, Juan-Les-Pins, septembre 1993.
- [NAS 93] C. NASTAR et N. AYACHE. « Fast segmentation, tracking and analysis of deformable objects ». *Proc. of Int. Conf. Computer Vision*, pages 275–279, Berlin, Allemagne, mai 1993.
- [PEN 91a] A. PENTLAND et B. HOROWITZ. « Recovery of non-rigid motion and structure ». *IEEE Trans. Pattern Anal. Machine Intell.*, 13(7):730–742, juillet 1991.
- [PEN 91b] A. PENTLAND et S. SCLAROFF. « Closed-form solutions for physically based shape modeling and recognition ». *IEEE Trans. Pattern Anal. Machine Intell.*, 13(7):715–729, juillet 1991.
- [ROU 91] N. ROUGON et F. PRETEUX. « Marqueurs déformables: segmentation par contour actif et morphologie mathématique ». *Actes du 8e congrès AF-CET/INRIA Reconnaissance des Formes et Intelligence Artificielle*, pages 955–966, Lyon, 1991.
- [STA 92] L.H. STAIB et J.S. DUNCAN. « Boundary Finding with Parametrically Deformable Models ». *IEEE Trans. Pattern Anal. Machine Intell.*, 14(11):1061–1075, novembre 1992.
- [SZE 95] G. SZEKELY, A. KELEMEN, C. BRECHBUHLER, et G.. GERIG. « Segmentation of 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier surface models ». *Proc. of Comp. Vis., Virtual Reality and Robotic in Medicine*, pages 495–505, Nice, avril 1995.

- [TER 88] D. TERZOPOULOS et A. WITKIN. « Deformable Models. Physically Based Models with Rigid and Deformable Components ». *IEEE Computer Graphics and Applications*, pages 41–51, novembre 1988.
- [TER 91] D. TERZOPOULOS et D. METAXAS. « Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics ». *IEEE Trans. Pattern Anal. Machine Intell.*, 13(7):703–714, juillet 1991.
- [YUI 92] A. YUILLE, P.W. HALLINAN, et D.S. COHEN. « Feature extraction from faces using deformable templates ». *Int. J. Computer Vision*, 8(2):99–111, 1992.
- [ZUG 95] D. ZUGAJ et V. LATTUATI. « Contrôle en temps réel de la forme d'une boucle de déroulement par Vision Artificielle ». *QCAV' 95*, Le Creusot, 1995.

Charles Kervrann est né en 1969. Titulaire du DEA "Signal Télécommunications Image et Radar" de l'université de Rennes 1 en 1992, il a préparé un doctorat en "Traitement du Signal et Télécommunications" à l'IRISA (Rennes, projet TEMIS). Ses travaux de thèse, soutenue en novembre 1995, ont porté sur la segmentation et le suivi de structures déformables 2D dans des séquences d'images par des techniques d'analyse statistiques de l'image et de la forme. Il mène actuellement ses travaux de recherche dans le cadre du projet européen VIDAS visant à coupler les signaux audio et vidéo pour la compression d'images visiophoniques.

Fabrice Heitz est né en 1961. Diplômé de Télécom Bretagne (1984), Docteur en automatique et traitement du signal (Télécom Paris, 1988), il a de 1988 à 1994 été Chargé de Recherches à l'IRISA/INRIA (Rennes, projet TEMIS). Il est, depuis 1994, Professeur à l'ENSPS et mène ses activités de recherche au LSIT/URA CNRS 1871, où il est responsable d'une opération de recherche sur l'analyse et la compression multi-images. Ses domaines d'intérêt actuels couvrent la modélisation de structures déformables 2D et 3D, la modélisation statistique multi-images ainsi que l'analyse et la compression de séquences d'images rapides.