

# *Projet VISTA*

*Vision Spatio-Temporelle et Active*

*Rennes*

THÈME 3B

*R* *apport*  
*d'Activité*

1999



## Table des matières

<b>1</b>	<b>Composition de l'équipe</b>	<b>3</b>
<b>2</b>	<b>Présentation et objectifs généraux</b>	<b>4</b>
<b>3</b>	<b>Fondements scientifiques</b>	<b>6</b>
3.1	Modèles markoviens hiérarchiques et estimation bayésienne . . . . .	6
3.2	Commande référencée vision . . . . .	8
3.3	Extraction de pistes et trajectographie . . . . .	10
<b>4</b>	<b>Domaines d'applications</b>	<b>12</b>
<b>5</b>	<b>Logiciels</b>	<b>13</b>
5.1	Logiciel RMRM d'estimation du mouvement dominant . . . . .	13
5.2	Logiciel MD-Shots de découpage d'une vidéo en plans . . . . .	13
5.3	Plate-forme VISP d'asservissement visuel . . . . .	14
5.4	Vistal: Librairies pour le traitement et l'analyse d'images volumiques spatio-temporelles . . . . .	14
<b>6</b>	<b>Résultats nouveaux</b>	<b>15</b>
6.1	Modèles statistiques, métrologie du mouvement et des déformations . . . . .	15
6.1.1	Conception de modèles hiérarchiques et d'algorithmiques efficaces pour l'analyse multi-images . . . . .	16
6.1.2	Modèles pour la mesure du mouvement fluide . . . . .	17
6.1.3	Amélioration de séquences d'images . . . . .	18
6.1.4	Modélisation statistique et classes de mouvement non paramétrées . . . . .	19
6.1.5	Structuration spatio-temporelle de vidéo pour l'indexation . . . . .	19
6.1.6	Modèles probabilistes de formes en imagerie médicale . . . . .	20
6.2	Recalage, suivi, trajectographie . . . . .	21
6.2.1	Suivi de structures déformables en imagerie météorologique . . . . .	21
6.2.2	Suivi en imagerie par filtrage particulière . . . . .	22
6.2.3	Suivi hybride 2D-3D . . . . .	22
6.2.4	Recalage 3D non linéaire en imagerie cérébrale IRM et IRMf . . . . .	23
6.2.5	Recalage et suivi de structures en imagerie ultrasonore . . . . .	24
6.2.6	Extraction de pistes et poursuite . . . . .	24
6.3	Asservissement visuel, perception active et exploration . . . . .	25
6.3.1	Nouvelle représentation pour l'asservissement visuel . . . . .	25
6.3.2	Planification dans l'image . . . . .	26
6.3.3	Asservissement visuel et animation . . . . .	26
6.3.4	Appariements d'objets et de scènes . . . . .	27
6.3.5	Couplage vision locale et vision globale . . . . .	27
6.3.6	Théorie de l'exploration . . . . .	28

<b>7 Contrats industriels (nationaux, européens et internationaux)</b>	<b>29</b>
7.1 Convention EDF : Suivi 2D de formes complexes avec exploitation de modèles CAO	29
7.2 Convention Edixia : Industrialisation de techniques d'asservissement visuel . . . . .	29
7.3 Convention Météo-France: Analyse spatio-temporelle d'images satellitaires appliquée à l'observation météorologique . . . . .	30
7.4 Convention Alcatel-CRC : Indexation vidéo . . . . .	30
7.5 Convention DGA - Thomson-CSF-Optronique: Conception d'une chaîne de traitements pour la restauration de séquences d'images . . . . .	31
7.6 Convention Adfac-Eumetsat : Estimation de champs de vents sur nuages bas . .	31
7.7 Convention Ifremer: Contribution au portage du logiciel Motion-2D sur une plate-forme temps-réel . . . . .	32
7.8 Convention Cifre Thomson-RCM: Méthodes et algorithmes de traitements multi-capteurs . . . . .	32
7.9 Convention Cifre Ina : Reconnaissance de formes spatio-temporelles pour l'indexation vidéo . . . . .	32
7.10 Actions financées par la Commission Européenne . . . . .	33
7.10.1 Projet Esprit (R&D MS) Divan . . . . .	33
7.11 Actions financées par le RNRT . . . . .	33
7.11.1 Projet Agir . . . . .	33
<b>8 Actions régionales, nationales et internationales</b>	<b>34</b>
8.1 Actions régionales . . . . .	34
8.1.1 GIS GSTB "Groupement de Télédétection en Bretagne" . . . . .	34
8.2 Actions nationales . . . . .	35
8.2.1 Pôle CNRS Micro-robotique: Asservissement visuel et micro-manipulation	35
8.2.2 Convention GIS Sciences de la Cognition: Imagerie cérébrale . . . . .	35
8.2.3 Actions de recherche coopératives de l'Inria . . . . .	35
8.2.4 Action incitative du Ministère de la Recherche, projet Visir . . . . .	37
8.3 Relations bilatérales internationales . . . . .	37
8.3.1 Programme France-Israël . . . . .	37
8.3.2 Programme France-Tunisie . . . . .	37
8.3.3 Accueil de chercheurs étrangers . . . . .	37
<b>9 Diffusion de résultats</b>	<b>38</b>
9.1 Animation de la communauté scientifique . . . . .	38
9.2 Enseignement universitaire . . . . .	38
9.3 Participation à des colloques, séminaires, invitations . . . . .	39
<b>10 Bibliographie</b>	<b>40</b>

# 1 Composition de l'équipe

## Responsable scientifique

Patrick Bouthemy [DR Inria]

## Assistante de projet

Edith Blin [TR Inria, avec projets Imadoc et Temics]

## Personnel Inria

François Chaumette [CR]

Éric Marchand [CR]

Patrick Pérez [CR]

Fabien Spindler [IR (Atelier)]

## Personnel CNRS

Christian Barillot [CR]

Patrick Gros [CR, à compter du 01/07/99]

Jean-Pierre Le Cadre [DR]

## Personnel Université

Etienne Mémin [maître de conférences Ifsic, Université de Rennes 1, à compter du 01/09/99]

## Personnel non permanent

Abdeljabar Benzougar [ingénieur expert Inria, à compter du 15/11/99]

Fabien Boissé [service national]

Didier Zugaj [ingénieur expert Inria, jusqu'au 31/08/99]

## Chercheurs post-doctorants et ATER

Caroline Baillard [post-doc Inria, à compter du 01/10/99]

Ioannis Pratikakis [post-doc Inria, à compter du 15/03/99]

Jean-Marc Laferté [Ater Ifsic, jusqu'au 31/08/99]

Stelios Zimeras [post-doc Inria, à compter du 01/09/99]

## Collaborateurs extérieurs

Etienne Mémin [maître de conférences UBS, IUT Vannes, jusqu'au 31/08/99]

### Chercheurs doctorants

Damien Albert [allocataire MENRT, à compter du 01/10/99]  
Annabelle Chardin [boursière Inria-Région]  
Isabelle Corouge [boursière Inria, à compter du 01/11/99]  
Thomas Corpetti [allocataire MENRT, à compter du 01/10/99]  
Frédéric Dambreville [boursier DGA]  
Fabien Dekeyser [boursier DGA]  
Ronan Fablet [boursier BDI CNRS-Région]  
Grégory Flandin [boursier Inria-Région]  
Pierre Hellier [boursier Inria-Région]  
Carine Hue [allocataire MENRT, à compter du 01/10/99]  
Isabelle Leibowicz [Boursière Cifre, Thomson RCM]  
Youcef Mezouar [allocataire MENRT]  
Christophe Papin [boursier Inria]  
Emmanuel Veneau [boursier Cifre, Ina, à compter du 01/03/99]

## 2 Présentation et objectifs généraux

**Mots clés :** imagerie spatio-temporelle, vision dynamique, vision active, modélisation statistique, modèle markovien, estimation bayésienne, segmentation, mouvement, suivi, trajectographie, asservissement visuel, stratégie de perception, déformation, mouvement fluide, fusion, sonar, imagerie médicale, vision robotique, indexation vidéo, imagerie météorologique.

**Résumé :** *Le projet Vista s'intéresse à plusieurs types d'imageries spatio-temporelles, pouvant relever de l'imagerie optique (vidéo, infra-rouge) comme acoustique (sonar, échographie). Nos travaux en vision spatio-temporelle et active se structurent en deux grandes catégories : i) l'analyse de phénomènes physiques, pour des besoins de mesures, référencées image ou scène, des besoins d'interprétation et de décision sur des événements temporels ; ii) la perception de systèmes automatisés ou robotiques, pour des problèmes de surveillance et de détection, de guidage et de manipulation, de navigation ou d'exploration. Nous privilégions une approche statistique (markovienne-bayésienne) de ces problèmes. Les aspects d'analyse spatio-temporelle long-terme et de suivi sont particulièrement étudiés, comme les questions d'analyse du mouvement et des déformations. Nous abordons les situations d'observateur actif, c.a.d. d'asservissement visuel ou de perception active, lorsque le capteur est commandable. Trois secteurs d'applications irriguent nos études: la métrologie du mouvement et des déformations (imagerie météorologique, imagerie médicale, visualisation expérimentale en mécanique des fluides), la vision robotique et systèmes de surveillance (sonar, transports, endoscopie), l'indexation vidéo.*

Les travaux du projet Vista portent sur deux grandes catégories de problèmes pouvant interagir :

- l'analyse de scènes ou de phénomènes physiques dynamiques, pour des objectifs de détection, d'interprétation et de décision sur des événements temporels, ainsi que pour des besoins de mesures ;
- la perception de systèmes automatisés ou robotiques, pour des tâches de surveillance, de guidage et de manipulation, de navigation et d'exploration.

Dans ces contextes, nous nous intéressons à plusieurs types d'imageries spatio-temporelles, 1D, 2D ou 3D +  $t$ , relevant principalement de l'imagerie optique (vidéo, infra-rouge), mais aussi acoustique (sonar, échographie).

Nous privilégions une approche statistique (markovienne-bayésienne) des questions d'analyse du mouvement et des déformations. Les aspects d'analyse spatio-temporelle long-terme sont particulièrement étudiés, notamment de suivi temporel. Nous abordons également les situations d'observateur actif, c.a.d. de perception active, lorsque le capteur est commandable. Trois secteurs d'applications motivent principalement nos études : *métrologie du mouvement et des déformations* (imagerie météorologique, imagerie médicale, visualisation expérimentale en mécanique des fluides), *vision robotique et systèmes de surveillance* (sonar, transports, endoscopie), *indexation de vidéos* par le contenu.

Plus précisément, nous structurons nos activités de recherche en trois axes principaux :

- **Modèles probabilistes et analyse d'images spatio-temporelles.** Dans le cadre de l'analyse d'images spatio-temporelles, et du suivi dans le temps des primitives extraites, les modélisations statistiques à base de dépendances locales offrent un cadre méthodologique souple et riche, permettant d'aborder à l'aide d'outils variés de nombreux problèmes. Associés à l'estimation bayésienne, de tels modèles permettent de spécifier le lien entre les données et les primitives inconnues, et d'y adjoindre une connaissance a priori sur ces dernières prenant en compte le contexte local (spatial ou temporel). La spécificité des différents modèles d'analyse spatio-temporelle que nous développons, dépend de la nature des variables, continue, discrète, ou symbolique, et de leurs interactions, linéaires ou non, mais aussi de la structuration de ces interactions locales. Nous nous intéressons dans ce contexte à des modèles markoviens hiérarchiques, pour des aspects de mesure, de segmentation ou de fusion.
- **Analyse du mouvement et trajectographie.** L'analyse du mouvement dans une séquence d'images doit être posée comme un problème joint d'estimation et de segmentation, puisqu'il s'agit d'appréhender des informations partiellement observables et discontinues. C'est un problème particulièrement difficile, mais dont la résolution est cruciale pour la plupart des tâches en analyse de scène dynamique. Ce problème peut se présenter en fait sous plusieurs variantes, suivant que l'objectif prioritaire se trouve être l'obtention d'une mesure dense ou paramétrique du mouvement, d'une partition de l'image en régions, ou l'extraction d'entités pertinentes. Un thème central et encore très ouvert sur ce sujet est l'analyse du mouvement déformable, qu'il soit articulé, élastique ou fluide. Un autre champ d'investigation motivant ces études sur la segmentation et la caractérisation du mouvement, ainsi que sur le suivi temporel, est représenté par la structuration et l'indexation de vidéo par le contenu. Enfin, sur les aspects de trajectographie, nous considérons

des problèmes liés à l'estimation de systèmes dynamiques partiellement observés dans un contexte passif, notamment dans le domaine sonar.

- **Asservissement visuel et perception active.** La perception active consiste à élaborer des stratégies de perception et d'action dans le but d'améliorer, par le contrôle des paramètres du capteur, les performances des algorithmes de vision ou de détection. Nous nous intéressons au contrôle de la position, de la vitesse, ou de la trajectoire de capteurs de type vidéo ou sonar. Deux types de stratégies sont considérés et combinés: d'une part, des stratégies locales, liées aux techniques d'asservissement visuel et fournissant une perception de bas niveau, et d'autre part, des stratégies globales, activées sur des événements particuliers et remontant à une perception de plus haut niveau. Nous abordons le couplage d'une vision locale embarquée et d'une vision globale de la scène. Par ailleurs, nous étudions des questions liées à la théorie de l'exploration ("search theory").

Dans les trois secteurs d'applications mentionnés, nous avons des conventions de recherche ou de transfert avec Météo-France, EDF, Ifremer, DCN, Thomson-CSF-Optronique, Alcatel-CRC, Ina, Arts Vidéo, des projets multi-partenaires nationaux (programme RNRT) et européens (projet Esprit R&D Divan), et des collaborations académiques avec notamment l'École Navale et le Cemagref.

### 3 Fondements scientifiques

bayésienne

#### 3.1 Modèles markoviens hiérarchiques et estimation bayésienne

**Mots clés :** champ markovien, observation partielle, estimation bayésienne, estimation de paramètres, modèle hiérarchique, algorithme multigrille, arbre, inférence non-itérative.

**Résumé :** *Dans le but de traiter les problèmes inverses multi-images liés aux axes de recherche du projet, nous nous intéressons au formalisme markovien associé à la théorie bayésienne de la décision [Win95]. Dans ce cadre, nous nous attachons à définir des modèles ou des algorithmes de type hiérarchique qui permettent de mener de façon efficace les tâches d'estimation (des cartes d'attributs cachés ou des paramètres du modèle), pour des problèmes à espace d'états aussi bien continu que discret.*

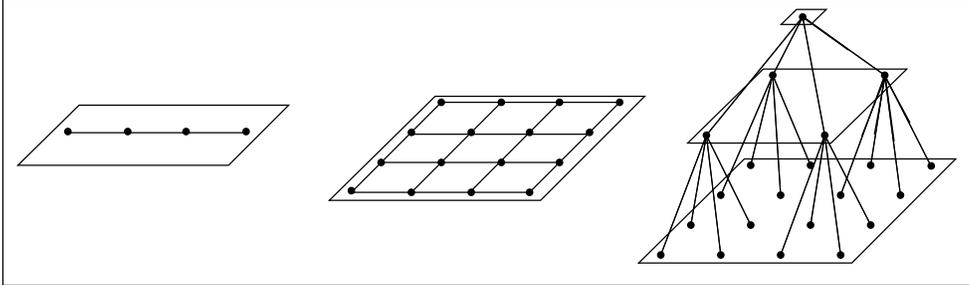
Le type d'approche statistique de l'analyse d'image que nous privilégions nous conduit à mettre en place des *modèles markoviens partiellement observés*, associés à des distributions  $P(z|\theta)$ , où  $\theta$  est un jeu de paramètres. L'ensemble  $z = \{z_s, s \in S\}$  se décompose en variables observées  $y$  et variables cachées  $x$ :  $z = (x, y)$ . Le caractère markovien se manifeste par une factorisation de cette loi jointe en un produit d'interactions locales :

$$P(z|\theta) \propto \prod_{c \in C} f_c(z_c|\theta)$$

---

[Win95] G. WINKLER, *Image analysis, random fields and dynamic Monte Carlo methods*, Springer, 1995.

où  $\mathcal{C}$  est un ensemble de “petites” parties de  $S$  ( $|\mathcal{c}| = 1$  ou 2 en général). Une structure de graphe est associée à cette factorisation (celle pour laquelle tous les éléments de  $\mathcal{C}$  sont des cliques), fournissant un outil de codage des indépendances conditionnelles entre variables <sup>[Whi90]</sup> : si une partie  $C$  de  $S$  sépare deux autres parties  $A$  et  $B$  pour ce graphe, alors  $z_A$  et  $z_B$  sont indépendants sachant  $z_C$  ( $P(z_A, z_B | z_C, \theta) = P(z_A | z_C, \theta)P(z_B | z_C, \theta)$ ). La chaîne, la grille bidimensionnelle et le quad-arbre sont les graphes les plus fréquemment rencontrés en analyse d’image ; ils sont illustrés sur la figure jointe.



Sur la base de la distribution  $P(z|\theta)$ , se posent les deux problèmes suivants (outre celui préalable du choix de la famille paramétrique) :

- **inférence statistique** :  $\theta$  étant fixé, nous recherchons le “meilleur”  $x$  connaissant  $y$  à l’aide d’estimateurs bayésiens tels que le MAP ( $\arg \max_x P(x, y | \theta)$ ), le MPM ( $\forall s, \arg \max_{x_s} P(x_s, y | \theta)$ ), et l’espérance  $E(x | y, \theta)$ .

Quels que soient l’estimateur et l’algorithmique d’inférence associée, une propagation de l’ensemble de l’information doit être opérée au travers du modèle par le biais des interactions locales. Les estimées vont résulter de la combinaison de décisions locales basées sur la fusion de l’information fournie par l’ensemble ou une partie du voisinage spatial ou temporel (donnant lieu à une prédiction) avec celle apportée localement par les données. C’est ce même principe qui, par exemple, sous-tend aussi bien les différents filtrages de Kalman temporels que nous sommes amenés à mettre en œuvre en matière de suivi, que l’estimation lissée (en espace) de champs de vitesses instantanées. En revanche cette inférence devra être itérative pour la plupart des structures de graphe (en particulier sur la grille bidimensionnelle usuelle) <sup>[GG84]</sup>.

- **estimation des paramètres** : l’estimation du meilleur jeu de paramètres est plus complexe encore, du fait de l’observation partielle du modèle. Les méthodes standard de maximum de vraisemblance ( $\hat{\theta} = \arg \max P(x, y | \theta)$ ) sont prises en défaut par la méconnaissance de  $x$ . En conséquence, les outils mis en œuvre sont essentiellement itératifs de type EM (“expectation-maximization”) avec pour but la maximisation de  $P(y | \theta)$ . Ils produisent généralement une estimation de nouveaux paramètres  $\theta^{(n+1)}$  sur la base d’échantillons tirés selon la loi *a posteriori*  $P(x | y, \theta^{(n)})$ . La taille et la structure des problèmes inverses spatiaux qui nous intéressent rendent cette tâche particulièrement délicate (tirage des échantillons, estimation des nouveaux paramètres par minimisation, etc.)

[Whi90] J. WHITTAKER, *Graphical models in applied multivariate statistics*, Wiley, 1990.

[GG84] S. GEMAN, D. GEMAN, « Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 6, 1984, p. 721–741.

Dans le cas de problèmes de très grande dimension (telle que l'extraction de cartes denses de primitives à partir de plusieurs images), les différents algorithmes génériques précédemment évoqués, même pour des paramètres supposés connus, sont coûteux : en raison de la localité des couplages entre variables, l'information se propage de façon lente et peu efficace.

Dans le but de traiter ces problèmes, une démarche générique souvent fructueuse consiste à *hiérarchiser* les primitives et/ou les données. Cette idée peut être déclinée de diverses manières.

Dans un premier groupe d'approches, des *versions "réduites"* (en terme du nombre des variables inconnues)  $P^i(x^i, y|\theta)$  ou  $P^i(x^i, y^i|\theta)$ ,  $i = L \dots 0$ , du modèle spatial initial sont d'abord construites, puis exploitées au sein d'une *algorithmique descendante*. Les attributs  $\hat{x}^i$  inférés à un niveau  $i$  servent de point de départ à la procédure d'inférence au niveau  $i - 1$  de "taille" immédiatement supérieure. Il s'agit ici principalement des méthodes multirésolutions ou multigrilles (selon que des versions réduites de  $y$  sont également calculées ou non). Ces méthodes permettent une accélération substantielle des schémas itératifs déterministes standard, et fournissent souvent une amélioration de la qualité des résultats [3].

Une deuxième classe d'approches a pour but la définition de modèles hiérarchiques *globaux*  $P(x^L, \dots, x^0, y|\theta)$ . Il s'agit de champs markoviens sous-tendus par un graphe *hiérarchique causal*, un arbre le plus souvent. Dans ce cas, les différents niveaux sont liés au sein du même modèle. L'inférence est donc menée de façon globale, et est rendue efficace par l'exploitation de la causalité. Plus précisément, des estimateurs exacts peuvent être obtenus en deux "passes" rappelant celles connues dans le cadre des chaînes de Markov. En revanche, il s'agit d'une classe restrictive de modèles dont la capacité à rendre compte des interactions spatiales entre primitives images est encore à étudier de façon complète.

### 3.2 Commande référencée vision

**Mots clés :** vision, robotique, commande référencée capteurs, asservissement visuel.

**Résumé :** *L'asservissement visuel consiste à utiliser les informations fournies par une caméra afin d'en contrôler le mouvement. L'approche que nous avons retenue repose sur la modélisation de fonctions de tâches appropriées et consiste à spécifier le problème en termes de régulation dans l'image. Elle permet de compenser les imprécisions des modèles (erreurs de calibration), aussi bien du capteur que du porteur de la caméra, par des lois de commande robustes en boucle fermée sur les informations visuelles extraites de l'image.*

Les techniques d'asservissement visuel [HHC96] utilisent généralement des informations visuelles 2D extraites de l'image. Les lois de commande consistent alors à contrôler le mouvement de la caméra afin que les mesures dans l'image  $s(t)$  atteignent une valeur désirée  $s^*$  ou suivent une trajectoire spécifiée  $s^*(t)$ .

Afin d'élaborer une loi de commande en boucle fermée sur des mesures  $s(t)$ , il est nécessaire d'estimer ou d'approximer la relation qui lie la variation de  $s$  aux variables de contrôle. Dans

---

[HHC96] S. HUTCHINSON, G. HAGER, P. CORKE, « A tutorial on visual servo control », *IEEE Transactions on Robotics and Automation* 12, 5, octobre 1996, p. 651–670.

le cas d'informations visuelles géométriques (telles par exemple les coordonnées 2D d'un point) contrôlées à l'aide des six degrés de liberté d'une caméra, cette relation est définie par [1] :

$$\dot{s} = L_{(s,z)} T$$

où  $T$  est le torseur cinématique de la caméra et où  $L$  est la matrice d'interaction associée à  $s$ . Cette matrice dépend de la valeur courante de  $s$ , mais aussi de la profondeur de l'objet considéré, représentée par les paramètres notés  $z$ . L'asservissement visuel 2D consiste schématiquement à réguler (c'est-à-dire amener et conserver à 0) la fonction de tâche suivante :

$$e = \hat{L}^+ (s - s^*)$$

où  $\hat{L}^+$  est la pseudo-inverse d'un modèle ou d'une approximation de  $L$ . Pour assurer la stabilité et la convergence de  $e$ , il faut avoir  $\hat{L}^+ L > 0$ ,  $\forall t$ , c'est-à-dire que  $\hat{L}$  doit être suffisamment correct et proche de  $L$  pour ne pas trop perturber le système. Deux choix sont couramment utilisés :

- $\hat{L} = L_{(s(t), z(t))}$ . On calcule à chaque itération la valeur courante de la matrice d'interaction. Une estimation des paramètres  $z$  doit alors être réalisée en ligne, par exemple à l'aide de la connaissance d'un modèle 3D de l'objet [DD95].
- $\hat{L} = L_{(s^*, z^*)}$ . Dans ce cas, la matrice choisie est constante et correspond à la configuration désirée. Une valeur de profondeur à la position finale, même très approximative, est nécessaire.

Ces deux possibilités ont chacune leurs avantages et leurs inconvénients [Cha97] : dans le premier cas, mouvements de la caméra inadéquats, voire impossibles à réaliser, rencontre éventuelle de minima locaux ; dans le second cas, possible passage de l'objet hors du champ de vue de la caméra. Finalement, il est possible de rencontrer une singularité de la matrice d'interaction, entraînant soit une instabilité de la commande, soit un échec dans la convergence du système. Nous travaillons à lever ces limitations, notamment par le développement des techniques dites d'asservissement visuel 2D 1/2.

Ces différents problèmes ne doivent pas faire oublier que l'asservissement visuel 2D apporte toute satisfaction dans un grand nombre de cas. De plus, l'intégration de l'asservissement visuel dans l'approche générale de la fonction de tâche [SBE91] permet de résoudre de manière efficace et élégante les problèmes de redondance rencontrés lorsqu'une tâche visuelle ne contraint pas l'ensemble des degrés de liberté de la caméra. Il est alors possible d'exécuter, parallèlement à la tâche visuelle, des tâches secondaires telles des opérations de suivi de trajectoires pour des applications d'inspection, ou d'évitement des butées et singularités du robot.

Lorsque la fonction de tâche est correctement modélisée, il est ensuite assez aisé d'élaborer une loi de commande générique permettant une décroissance exponentielle de la fonction de

- 
- [DD95] D. DEMENTHON, L. DAVIS, « Model-based object pose in 25 lines of code », *Int. Journal of Computer Vision* 15, 1/2, juin 1995, p. 123-141.
- [Cha97] F. CHAUMETTE, « Potential problems of stability and convergence in image-based and position-based visual servoing », *in: Workshop on Vision and Control*, Block Island, Rhode Island, juin 1997.
- [SBE91] C. SAMSON, M. L. BORGNE, B. ESPIAU, *Robot Control: the Task Function Approach*, Clarendon Press, Oxford, 1991.

tâche. On obtient :

$$T = -\hat{\lambda} e - \frac{\widehat{\partial e}}{\partial t}$$

où  $\hat{\lambda}$  est un gain qui permet de régler la vitesse de convergence, et où  $\frac{\widehat{\partial e}}{\partial t}$  représente une compensation de la vitesse éventuelle de l'objet considéré. Ce terme, s'il est correctement estimé, permet de supprimer les erreurs de traînage inhérentes à tout problème de poursuite de cibles.

### 3.3 Extraction de pistes et trajectographie

**Mots clés :** trajectographie, extraction de pistes, poursuite, hypothèse multiple, combinatoire.

**Résumé :** *Nous décrivons les problèmes liés à l'extraction de pistes et à l'estimation des paramètres cinématiques. L'accent est mis sur la nature des observations et des problèmes d'estimation associés.*

Nous considérons les problèmes liés à l'estimation de systèmes dynamiques partiellement observés. Pratiquement, ces problèmes sont généralement posés dans un cadre passif, l'observateur est uniquement un récepteur.

Un corollaire de ceci est que le système est, en général, uniquement partiellement observé, ce qui signifie que l'on n'observe pas directement l'état du système, mais seulement une fonction (non-linéaire) bruitée de cet état. Ainsi, en sonar passif, l'état du système (i.e. les paramètres définissant la trajectoire de la source) n'est observé qu'au travers des gisements (angles) estimés par l'antenne.

Le problème d'extraction-poursuite de cibles mobiles se pose dans le contexte du traitement de l'information. On dispose d'observations constituées de fonctions non-linéaires de l'état de la source (azimuts, dopplers, etc.). Les observations sont décrites comme les sorties d'un système variable dans le temps. Le modèle d'évolution temporelle de l'état est lié aux hypothèses faites sur la trajectoire des sources. Il est alors possible de décrire les observations par un système non-linéaire variable dans le temps pour lequel les paramètres à déterminer sont les composantes de l'état initial. Plus précisément, on appelle  $\mathbf{X}^i$  le vecteur de l'état relatif (i.e. dans un repère lié à l'observateur) de la  $i$ ème source dont les composantes sont les coordonnées relatives  $(r_x, r_y)$  et les vitesses relatives  $(v_x, v_y)$  de la source. L'état du système est  $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^n)$  et son équation à temps discret prend alors la forme ci-dessous, [NLG84]:

$$\mathbf{X}(k) = \Phi(k, k-1)\mathbf{X}(k-1) + \mathbf{U}(k) + \mathbf{W}(k)$$

où  $\Phi$  désigne la matrice de transition du système, le vecteur  $\mathbf{U}$  représente l'accélération (instantanée) de l'observateur et où  $\mathbf{W}_k$  représente un bruit markovien. Par ailleurs, l'équation d'observation prend la forme suivante:  $\theta_k^i = \theta_k^i + w_k$ , où  $\theta_k^i = \arctan(r_x^i(k)/r_y^i(k))$  et  $i$  est un

---

[NLG84] S. NARDONE, A. LINDGREN, K. GONG, « Fundamental properties and performance of conventional bearings-only target motion analysis », *IEEE Trans. on Automatic Control* 29, 9, sept. 1984, p. 775-787.

indice aléatoire correspondant au fait que l'on ne sait pas de quelle source provient la détection (ou s'il s'agit d'une fausse alarme).

Les caractéristiques de ce type de problème sont multiples : *i*) non-linéarité du système, *ii*) possibilité de manœuvre (action sur  $\mathbf{U}$ ); *iii*) caractère aléatoire de  $i$ . L'étape d'extraction consiste à attribuer des détections aux sources (estimation de  $i(n, \omega)$ ), et lorsque  $i(n, \omega)$  est connu, ou estimé, le problème se découple alors sur les différentes sources. L'objectif de l'*extraction* consiste à élaborer des pistes à partir des données provenant des étapes de traitement direct et massif des sorties de capteurs. Les données sont constituées de plots (de nature binaire dans le cas le plus simple). Le problème de l'*extraction* est le suivant: quels sont les sous-ensembles de plots  $S_i$  ayant une origine commune? On entend ici par origine commune le fait que les différents points de  $S_i$  sont associés à une même source en mouvement. Un sous-ensemble  $S_i$  s'appelle alors une piste extraite. Il importe de noter que cette suite est d'une nature statistique plus compliquée que précédemment puisque, dans le cas général, elle inclut des fausses alarmes, des fausses associations ainsi que des non-détections.

Une approche naturelle consiste donc à considérer un modèle probabiliste des observations incluant à la fois des hypothèses sur la nature statistique des observations (probabilités de détection, de fausse-alarme, distribution et nombre de fausses alarmes) et sur la nature de la trajectoire de la (ou des) source(s). Utilisant cette structure, il s'agit de séparer l'ensemble des plots en des sous-ensembles de plots ayant une origine cinématique commune et un ensemble de fausses alarmes.

Ainsi, dans le cas mono-source, la densité des observations est un mélange de lois normale (la densité d'une vraie détection), de Poisson (le nombre de fausses alarmes par unité de volume) et uniforme (répartition des fausses alarmes). On montre ainsi que la densité des observations prend la forme ci-dessous :

$$P(\mathcal{Z}|\mathbf{X}_0) = \prod_{t=1}^T p(Z_t|\mathbf{X}_0) \quad \text{avec}$$

$$p(Z_t|\mathbf{X}_0) = u^{-m_t}(1 - P_d)\mu_u(m_t) + u^{1-m_t} \frac{P_d \cdot \mu_u(m_t - 1)}{m_t}$$

$$\sum_{j=1}^{m_t} \det(\Sigma_t)^{-1/2} \exp\left(-1/2 \|z_{t,j} - h_t(\mathbf{X}_0)\|_{\Sigma_t^{-1}}^2\right)$$

où  $m_t$  désigne le nombre de détections dans la fenêtre de validation,  $u$  le volume de la fenêtre de validation,  $Z_t$  l'observation au temps  $t$ ,  $\mu_u(m_t)$  la probabilité de  $m_t$  et  $P_d$  la probabilité de détection. Le problème de l'*extraction* revient alors à chercher le vecteur  $\mathbf{X}_0$  maximisant la vraisemblance  $P(\mathcal{Z}|\mathbf{X}_0)$  définie ci-dessus.

Cependant, la réalité de l'extraction est faite de problèmes multi-sources. L'assignation des mesures aux pistes peut être réalisée au moyen d'une énumération exhaustive des hypothèses d'assignation. Ceci constitue le cadre classique des algorithmes MHT (Multiple Hypotheses Tracking). Pour éviter l'explosion combinatoire de tels algorithmes, diverses méthodes ont été développées. Il est cependant assez évident que la faiblesse des algorithmes de tri et, dans une moindre mesure, de fusion réside dans le risque d'éliminer certaines bonnes séquences.

Une approche séduisante consiste alors à considérer que les probabilités d'assignation des mesures aux pistes sont indépendantes d'un scan à l'autre. Ainsi, la méthode PMHT (Probabilistic Multiple Hypothese Tracking) ne nécessite aucune énumération des hypothèses, mais un nombre (maximal) de pistes est fixé. L'algorithme PMHT repose sur deux étapes, l'une d'optimisation et l'autre de calcul d'espérance conditionnelle. Ces deux étapes ont été reconsidérées dans le contexte plus général de l'estimation des paramètres de mélanges (algorithme EM), [Gau97].

Le problème de l'initialisation des pistes, lui-même, consiste en l'association de diverses détections élémentaires prises au cours du temps. C'est avant tout un problème de combinatoire. Il s'agit, en général, d'un problème d'assignation  $n$ -dimensionnel de variables binaires. Notre approche a été celle de la relaxation Lagrangienne où  $n - 2$  variables d'assignation sont relaxées et où l'on se ramène à une série de problèmes d'assignation 2D, [Gau97]. L'estimation des variables duales est elle-même accomplie au moyen d'algorithmes du type "sous-gradient". Les étapes d'obtention de solutions réalisables (calcul du saut de dualité), d'optimisation (sous-gradients, calcul du pas) ainsi que l'inclusion de fausses mesures ont été étudiées avec attention. Ce type d'approche a été appliqué à des problèmes d'association de données multi-capteurs, [PDBSW92].

## 4 Domaines d'applications

**Mots clés :** imagerie météorologique, imagerie médicale, sonar, indexation vidéo, mécanique des fluides, robotique nucléaire, robotique sous-marine, transports.

**Panorama :** *Nous nous intéressons à trois grands types d'applications. Le premier concerne ce que l'on peut appeler la métrologie du mouvement et des déformations. Nous explorons dans ce cadre les domaines de la visualisation expérimentale en mécanique de fluides (écoulements turbulents), de l'imagerie satellitale météorologique (images Meteosat, avec Meteo-France), de l'imagerie médicale (imagerie cérébrale, IRM, IRMf, échographie 3D). Le second secteur a trait à la robotique, aussi bien robotique manufacturière que robotique d'intervention, et aux systèmes de surveillance. Les domaines d'applications concernés sont d'une part l'énergie (avec EDF), le sous-marin (avec l'Ifremer), l'agro-alimentaire (avec le Cemagref), les transports. Enfin, nous menons un investissement significatif dans le domaine de l'indexation vidéo. Nous sommes concernés par les aspects de structuration temporelle de vidéo et d'indexation par leur contenu spatio-temporel, avec différents objectifs comme la consultation d'archives audio-visuelles ou l'édition de documents vidéo interactifs (avec des collaborations avec l'Ina et Alcatel-CRC notamment). Nous participons à un projet européen sur le sujet (projet Divan, avec le projet Sigma-2) et au projet Agir dans le cadre du RNRT.*

---

[Gau97] H. GAUVRIT, *Extraction multipistes: approches probabilistes et combinatoires*, thèse de doctorat, université de Rennes 1, novembre 1997.

[PDBSW92] K. PATTIPATI, S. DEB, Y. BAR-SHALOM, R. WASHBURN, « A new relaxation algorithm and passive sensor data association », *IEEE Trans. on Automatic Control* 37, 2, fév. 1992, p. 198–213.

## 5 Logiciels

### 5.1 Logiciel RMRm d'estimation du mouvement dominant

*Correspondant:* Fabien Spindler

Un logiciel, s'intitulant RMRm pour "Robust Multi-Resolution modified", d'estimation du mouvement dominant dans une séquence d'images a été développé en langage C ANSI. Il correspond à une méthode d'estimation robuste, multi-résolution, et incrémentale de modèles 2D paramétrés de mouvement, n'exploitant que les gradients spatio-temporels de l'intensité, [9]. Ce logiciel a été optimisé, et déposé à l'APP. Le logiciel RMRm permet d'estimer plusieurs types de modèles de mouvement, respectivement, modèles constants (translations), affines, et quadratiques. Il intègre de plus la possibilité de prendre en compte les variations globales d'illumination. Pour agir directement sur la cadence de traitement, il est également possible de spécifier la taille du support d'estimation, ainsi que le premier et le dernier niveau considérés dans les pyramides d'images multi-résolutions. Il atteint actuellement selon les options retenues, des cadences de traitement comprises entre 1 Hz et 3 Hz sur station Ultra-Sparc 1 à 170 MHz. Pour évaluer les performances de l'algorithme et la validité du modèle de mouvement estimé, le logiciel permet aussi de générer une séquence d'images compensées, par rapport à un instant de référence sélectionné, par le mouvement estimé. Ce logiciel a été mis à disposition chez plusieurs partenaires industriels. Il peut être exploité tel quel, ou être un module d'une application plus conséquente. Une interface réalisée sous Tcl-Tk permet de le piloter de façon conviviale.

### 5.2 Logiciel MD-Shots de découpage d'une vidéo en plans

*Correspondant:* Fabien Spindler

Le logiciel MD-**shots**, pour "motion-based detection of shots", de découpage d'une vidéo en plans élémentaires a été développé en langages C-ANSI et C++. Préalablement à une indexation des informations contenues dans une vidéo, il est nécessaire de bâtir une représentation temporelle structurée de cette vidéo, correspondant au découpage en plans élémentaires. Il s'agit de détecter les "cuts" et les transitions progressives marquant les changements de plans. Par transition progressive, on entend des effets spéciaux comme le fondu, le fondu enchaîné, les volets,...

La méthode proposée pour réaliser le découpage de la vidéo en plans élémentaires [13] s'appuie sur la cohérence temporelle du support associé au mouvement global dominant entre deux images successives au sein d'un même plan. Ce dernier est estimé à l'aide du logiciel RMR-m. On s'intéresse en fait à l'évolution temporelle de la taille normalisée de ce support. Lors d'un "cut", aucun modèle de mouvement ne peut correctement décrire la transformation entre ces deux images, et on constate une chute soudaine de cette valeur. Lors de transitions progressives, cette chute est moins nette mais plus prolongée. Cette détection de sauts est effectuée par un test statistique de Hinkley, dont l'implantation est simple et efficace. Un atout original de la méthode est d'appréhender ainsi la détection des différentes transitions par un même test, ne comprenant qu'un seul paramètre à fixer par l'utilisateur. Une même valeur de ce paramètre est de plus utilisée pour les deux types de transitions traitées.

Pour agir directement sur la cadence de traitement du logiciel, il est possible de spécifier le premier et le dernier niveau considérés dans les pyramides d'images multi-résolution. Il est également possible de retenir un sous-échantillonnage temporel de la séquence à traiter. Selon les options retenues, le logiciel atteint actuellement des cadences de traitement comprises entre 0,6 Hz et 2,6 Hz sur station Ultra-Creator à 300 MHz. Ce logiciel a été déposé à l'APP. Il a fait l'objet d'une mise à disposition auprès de Alcatel-CRC et de l'Ina. D'autre part, un contrat de licence d'exploitation a été conclu avec la société Arts Vidéo Interactive (insertion dans le produit commercialisé MOVIDEO-2 STUDIO).

### 5.3 Plate-forme ViSP d'asservissement visuel

*Correspondant:* Éric Marchand

Face à l'absence d'un environnement logiciel permettant le prototypage rapide de tâches d'asservissement visuel, absence principalement due au recours à des matériels spécifiques (robot, cartes d'acquisition et de traitement d'image...), nous avons développé la plate-forme ViSP (pour "Visual Servoing Platform"). L'objectif est de pouvoir concevoir des applications portables et facilement adaptables à d'autres contextes de travail. L'environnement réalisé permet la mise à disposition à un programmeur d'un ensemble de briques élémentaires pouvant être combinées pour la construction d'applications plus complexes.

ViSP présente les fonctionnalités requises pour ce type d'utilisation : indépendance vis-à-vis du matériel, simplicité, évolutivité, portabilité. ViSP est écrit en C++ et dispose :

- d'une bibliothèque de tâches élémentaires de positionnement par rapport à des primitives visuelles variées (points, lignes, cercles, sphères, cylindres, etc.) ; elles peuvent être combinées pour la prise en compte d'objets plus complexes.
- d'une bibliothèque de traitements d'image qui permet le suivi de primitives visuelles (point, segment, ellipse, spline, etc.) à la cadence vidéo.

Nous avons ajouté cette année :

- une bibliothèque de calcul de pose (méthodes de Lagrange, Dementhon, Lowe).
- une bibliothèque de simulation de processus d'asservissement visuel avec visualisation sous OpenGL.

ViSP a été utilisé dans la réalisation de diverses expérimentations, telles que le suivi d'un tube par asservissement visuel, l'évitement d'occultations ou des butées articulaires, ainsi que diverses tâches de positionnement (point, ligne, cercle, sphère, cylindre, carré, quadrilatère). Trois des modules de ce logiciel *ViSP\_pose*, *ViSP\_tracking* et *ViSP\_control* ont fait l'objet d'un dépôt à l'APP en décembre 1999.

### 5.4 Vistal: Librairies pour le traitement et l'analyse d'images volumiques spatio-temporelles

*Correspondants:* Fabien Boisse, Christian Barillot

Dans le contexte du développement de logiciels d'analyse d'images 3D et 3D+t dans le domaine de l'imagerie médicale, nous travaillons à la réalisation de librairies de traitements

d'images volumiques spatio-temporelles utilisables dans des contextes variés comme le recalage linéaire et le recalage non-linéaire d'images, la segmentation d'images et la visualisation d'images et de résultats. Cette plate-forme logicielle, dénommée VISTAL pour "Volumetric ImageS Treatments and Analysis Libraries", est constituée d'une classe générique (template) C++ ainsi que d'une série de bibliothèques de traitements, incluant calcul algébrique et d'opérations floues sur des images 3D, manipulation de voisinage dans l'image, filtres linéaires (RIF et RII) et non-linéaires, filtres gaussiens 1D et ses dérivées en RIF, calcul de gradient, de laplacien et de leurs dérivées, gestion de traitements sur des histogrammes, calcul d'interpolations sur des images 3D, opérations élémentaires de morphologie mathématique, manipulation de contour actif 3D de forme cylindrique, calcul de textures d'ordre 1 et d'ordre 2 sur les images.

VISTAL est un environnement indépendant de la plate-forme logicielle Unix utilisée (Solaris, Linux) ainsi que du matériel. Son évolutivité est assurée par l'utilisation, comme noyau, d'une classe "template" C++. Cet environnement sert actuellement de base au développement de logiciels correspondant aux études menées en recalage d'images linéaire et non-linéaire (dense) ou encore en segmentation d'images à partir d'ensembles de niveaux (voir sous-sections 6.1.6, 6.2.4 et 6.2.5).

## 6 Résultats nouveaux

### 6.1 Modèles statistiques, métrologie du mouvement et des déformations

*Contexte: Dans le cadre de l'analyse d'images spatio-temporelles, les modélisations statistiques à base de dépendances locales offrent un cadre méthodologique souple et riche, permettant d'aborder à l'aide d'outils variés de nombreux problèmes. Associés à la théorie bayésienne de la décision, de tels modèles permettent de spécifier le lien entre les données et les primitives inconnues, en y adjoignant une connaissance a priori sur ces dernières prenant en compte le contexte local (spatial ou temporel). La spécificité des différents modèles d'analyse spatio-temporelle que nous avons déjà définis ou que nous comptons développer, dépend de la nature des variables, continue, discrète, ou symbolique, et de leurs interactions, linéaires ou non, mais aussi de la structuration de ces interactions locales. Nous nous intéressons dans ce contexte à des modèles markoviens hiérarchiques ainsi qu'aux estimateurs robustes.*

*L'analyse du mouvement dans une séquence d'images doit être posée comme un problème joint d'estimation et de segmentation, puisqu'il s'agit d'appréhender des informations partiellement observables et discontinues. Une partition de l'image en régions cohérentes au sens du mouvement nécessite, sous une forme ou une autre, une mesure du mouvement. Inversement, le calcul d'un champ des vitesses 2D dans le cas général impose une détection et une gestion simultanée des possibles discontinuités (inconnues a priori) du mouvement. C'est donc un problème particulièrement difficile, mais dont la résolution est cruciale pour la plupart des tâches en analyse de scène dynamique et plus généralement de phénomènes physiques dynamiques. Ce problème peut se présenter en fait sous plusieurs variantes, suivant que l'objectif prioritaire se trouve être l'obtention d'une mesure dense ou paramétrique du mouvement, d'une partition de l'image en régions, ou l'extraction d'entités pertinentes. Un thème central et encore très ouvert sur ce sujet concerne l'analyse du mouvement déformable et en particulier fluide ; nous visons dans ce contexte une véritable métrologie du mouvement et des déformations. Un autre champ*

*d'investigation motivant ces études sur la segmentation et la caractérisation du mouvement est représenté par l'indexation vidéo par le contenu.*

### 6.1.1 Conception de modèles hiérarchiques et d'algorithmiques efficaces pour l'analyse multi-images

**Participants :** Patrick Pérez, Etienne Mémin, Annabelle Chardin, Carine Hue, Jean-Marc Laferté.

Pour analyser efficacement des volumes d'images très importants (séquences d'images vidéo, images satellitales de très grande taille obtenues à différentes résolutions, dans différentes longueurs d'onde, et à différentes dates, données volumiques multimodales et/ou multipatients en imagerie médicale), nous étudions, dans le cadre de la modélisation probabiliste, d'une part de *nouveaux modèles hiérarchiques*, et d'autre part de *nouvelles algorithmiques* s'appliquant aux modèles existants.

Le premier axe d'étude concerne la définition de modèles hiérarchiques autorisant l'emploi d'une algorithmique légère, tout en conservant la souplesse de définition et la richesse de modélisation des champs markoviens. Dans le cas particulier où le graphe d'interaction est triangulé (cas des arbres en particulier), des méthodes non itératives d'inférence (i.e., complexité par pixel constante) à paramètres fixés, et des procédures EM ("Expectation Maximization") associées pour l'estimation des paramètres, ont ainsi été étudiées [4, 23].

Dans les problèmes d'analyse d'images, les limitations inhérentes à la condition de triangulation s'avèrent cependant assez contraignantes. Ceci nous a conduits à étudier des modèles hiérarchiques originaux reposant sur une classe plus vaste de structures d'interactions obtenues par la troncation des arbres et par le placement au niveau le plus "grossier" d'une structure spatiale. De cette façon, une propagation de l'information sur l'ensemble du domaine est assurée au niveau du réseau non-causal de petite taille par un schéma itératif léger, et sur le reste de la structure hiérarchique par les techniques non itératives en deux passes. Après avoir étudié ce type d'approche dans le cadre de la classification supervisée au sens du MAP, nous avons poursuivi par la mise en place sur ces structures de l'inférence au sens du MPM [35, 36] et de l'estimation des divers paramètres par des techniques EM [37] et ICE ("Iterated Conditional Expectation").

Le travail théorique sur les modèles à interactions arborescentes se poursuit par ailleurs. Après avoir obtenu une vision unifiée des méthodes d'inférence non itératives sur l'arbre pour les modèles aussi bien discrets que gaussiens, nous avons commencé à examiner de façon critique et comparative, pour ces deux types de modèles sur arbre, les diverses paramétrisations proposées, souvent sans justification, dans l'état-de-l'art. Il a entre autres été montré lors du stage de Carine Hue que la paramétrisation gaussienne avec décroissance géométrique de l'innovation le long de la structure arborescente est à déconseiller. L'identification de ses paramètres s'avère difficile voire impossible si seuls les niveaux les plus fins sont observés.

Le second axe d'étude vise à améliorer l'algorithmique itérative associée aux modèles markoviens spatiaux pour lesquels le graphe d'interaction est typiquement une grille. L'introduction d'estimateurs robustes dans ce type de modèles conduit, comme c'est le cas dans nos outils d'estimation de champs denses de mouvement, à des problèmes d'optimisation non convexe

de très grande dimension. Pour mener à bien de telles minimisations, nous avons dans un premier temps défini des schémas de type multigrille reposant d'une part sur des contraintes paramétriques variables, et d'autre part sur une adaptativité des grilles emboîtées exploitant de façon originale les variables auxiliaires que la réécriture semi-quadratique des fonctions fait apparaître. Nous nous sommes ensuite intéressés à d'autres techniques d'optimisation qui pourraient être utilisées de façon concurrente ou complémentaire. Dans cette optique, la comparaison de diverses techniques de résolution des très grands systèmes linéaires creux rencontrés à chaque étape des techniques de moindres carrés pondérés itérés est poursuivie en collaboration avec le projet Aladin (J. Erhel et F. Guyomarch). Par ailleurs, la collaboration avec le projet Api (T. Risset) autour de la synthèse de composants VLSI dédiés nous a conduits à étudier d'un point de vue théorique différentes techniques de moindres carrés pondérés itérés. Un schéma de Jacobi non linéaire original, mieux adapté aux fonctionnelles manipulées, a en particulier été introduit. L'amélioration des performances a été établie expérimentalement, et sa convergence a été théoriquement établie sous certaines conditions [26].

### 6.1.2 Modèles pour la mesure du mouvement fluide

**Participants :** Étienne Mémin, Patrick Pérez, Thomas Corpetti.

Dans la continuité des études que nous menons sur l'estimation et la segmentation du mouvement 2D dans des séquences d'images <sup>[MP98a]</sup>, nous nous intéressons à la définition de modèles spécifiquement adaptés aux mouvements fluides. En effet, les modèles génériques d'estimation du mouvement basés sur l'équation de contrainte du mouvement apparent et un lissage au premier ordre, ne se comportent pas de façon satisfaisante lorsque les champs à mesurer exhibent localement de fortes vorticités et/ou divergences qui sont des structures capitales pour la description et l'étude de tels flots. Nous avons tout d'abord étudié le remplacement, dans le modèle de mesure, de l'équation de conservation de la luminance, par l'équation de continuité de la mécanique des fluides. Cette équation cinématique doit néanmoins être maniée avec précaution lorsqu'elle est introduite dans un contexte multirésolution incrémental tel que celui usuellement employé pour la mesure des grands déplacements. Nous avons donc adapté l'équation de continuité au cadre multirésolution, et abouti à une nouvelle équation de contrainte du mouvement apparent mieux adaptée au mouvement fluide et valide pour les grands déplacements. Parallèlement, la fonction de lissage permettant de régulariser le problème est définie en termes de divergence et de rotationnel, selon le principe de la régularisation "div-curl" préservant les concentrations significatives de l'une ou l'autre de ces deux grandeurs différentielles. L'estimateur basé sur ces deux ingrédients fournit des résultats prometteurs pour un certain nombre de situations. Initié dans le cadre du stage de DEA de Thomas Corpetti, ce travail est actuellement poursuivi en thèse.

Le second travail se situe dans le prolongement plus directe des méthodes d'estimation et de segmentation conjointes que nous avons proposées récemment [8], et de structuration de mouvements fluides par la localisation des points singuliers étudiée précédemment dans le

---

[MP98a] E. MEMIN, P. PEREZ, « Dense estimation and object-based segmentation of the optical flow with robust techniques », *IEEE Transactions on Image Processing* 7, 5, 1998, p. 703–719.

projet [MBD98]. Dans cette étude [62], l'estimation dense du champ des vitesses est couplée à un modèle paramétrique non linéaire. Ce modèle permet de décrire la forme du champ des vitesses apparentes de l'écoulement à l'aide d'entités particulières telles que les vortex, les puits ou les sources. Il permet de rendre compte des composantes solénoïdale et irrotationnel du champ des vitesses. La composante laminaire du flot, responsable d'un mouvement d'ensemble, forme un champ très lisse qu'il est possible d'extraire au moyen de l'estimateur de mouvement générique que nous avons développé [8]. Les vortex, puits et sources sont extraits grâce à la technique des indices de Poincaré appliquée au champ résiduel résultant de la différence entre l'estimé courant du champ de mouvement et la composante laminaire du flot. L'estimation du champ dense, des points singuliers et des modèles paramétriques non-linéaires associés, est effectuée de façon conjointe et coopérative. Ceci fournit, à la convergence, une bonne mesure dense de l'écoulement fluide, mais aussi une description paramétrique compacte permettant de caractériser la structure de l'écoulement, et, éventuellement de suivre cette structure dans le temps.

### 6.1.3 Amélioration de séquences d'images

**Participants :** Patrick Bouthemy, Patrick Pérez, Fabien Dekeyser.

Nous étudions des problèmes d'amélioration de vidéos incluant des aspects de débruitage spatio-temporel et de construction d'images de plus grande résolution à partir de la séquence originelle. Il est évident qu'un filtrage spatio-temporel approprié doit tenir compte du mouvement présent dans la séquence d'images. Nous cherchons plus précisément à combiner estimation du mouvement et filtrage spatio-temporel des images. Le premier mode de combinaison étudié cette année est à caractère séquentiel, phase initiale d'estimation du mouvement suivie du filtrage spatio-temporel appliqué à la séquence d'images compensées par le mouvement estimé. Deux types de mesures de mouvement ont été pris en compte et réalisés par les méthodes déjà développées dans le projet : une estimation paramétrique globale robuste du mouvement dominant d'une part [9], ce qui est apparu original et efficace dans ce contexte, une estimation hiérarchique robuste d'un champ dense de vitesses avec prise en compte des discontinuités d'autre part [8]. Nous avons mené une étude comparative de différentes techniques de filtrage spatio-temporel pour le débruitage des séquences d'images avec un accent plus particulier mis sur le filtrage de Wiener. Parmi les différentes variantes possibles, de très bons résultats expérimentaux ont été obtenus avec un filtre de Wiener spatio-temporel 3D calculé dans le domaine fréquentiel. En effet, ce filtre a permis d'obtenir une bonne réduction du bruit tout en préservant les détails de l'image et en offrant un faible coût de calcul. La prise en compte des erreurs de compensation du mouvement conduit de plus à opter pour une version adaptative de ce filtre. Ce type de débruitage permet ensuite de recourir efficacement à des techniques simples de régularisation, telles que la troncation de valeurs singulières, pour la déconvolution du flou. Nous avons commencé à utiliser ce type de filtrage comme pré-traitement dans un schéma de super-résolution. Nous obtenons ainsi une image non-uniformément échantillonnée dont on

---

[MBD98] M. MAURIZOT, P. BOUTHEMY, B. DELYON, « 2D fluid motion analysis from a single image », in : *Proc. Conf. on Computer Vision and Pattern Recognition, CVPR'98*, p. 184–189, Santa Barbara, juin 1998.

pourra interpoler les valeurs sur une grille fine. Par ailleurs, nous abordons le deuxième mode d'amélioration de séquences d'images, où estimation du mouvement et filtrage spatio-temporel sont considérés conjointement et interagissent. Cette étude a été menée en collaboration avec Thomson-CSF-Optronique (voir sous-section 7.5).

#### 6.1.4 Modélisation statistique et classes de mouvement non paramétrées

**Participants :** Patrick Bouthemy, Patrick Pérez, Ronan Fablet.

Dans le cadre de l'indexation d'une vidéo par le contenu, nous avons proposé une mesure de similarité sur les informations de mouvement reposant sur une approche bayésienne. Le choix d'une telle mesure se pose dans la phase d'indexation d'une vidéo (prise comme un problème de classification), comme dans la phase de recherche (recherche des éléments les plus semblables à une requête donnée). L'utilisation de descripteurs sous la forme directe d'attributs numériques calculés s'avère délicate en raison de la grande dimension de l'espace d'attributs ainsi générés, d'une formalisation difficile de la notion d'incertitude de mesure, et d'un choix non évident d'une distance adéquate. Aussi, il nous paraît préférable d'identifier des modèles de distribution sur ces attributs. Les attributs considérés pour caractériser l'information de mouvement contenue dans une vidéo, résultent de l'évaluation de co-occurrences temporelles de mesures locales de mouvement appropriées [45]. Notre approche exploite ainsi une modélisation par champs de Gibbs causaux de la distribution spatio-temporelle, associée à la séquence vidéo traitée, des mesures locales de mouvement. L'utilisation de modèles causaux rend possible le calcul exact de la vraisemblance d'une séquence de mesures locales de mouvement pour un modèle donné. L'estimation du modèle au sens du maximum de vraisemblance peut être réalisée par des techniques d'optimisation déterministes classiques. Par ailleurs, nous pouvons définir de manière immédiate et fondée une mesure de similarité entre séquences à partir de la divergence de Kullback appliquée aux modèles estimés. Cette mesure permet d'extraire une représentation hiérarchique d'une base de vidéos exprimant des similarités graduelles en termes de contenu de mouvement. Cet arbre peut ensuite être exploité dans la recherche des exemples similaires, au sens du mouvement, à une vidéo proposée comme requête par un utilisateur.

#### 6.1.5 Structuration spatio-temporelle de vidéo pour l'indexation

**Participants :** Patrick Bouthemy, Patrick Gros, Ronan Fablet, Didier Zugaj.

Nous continuons à développer notre schéma de structuration et d'indexation de vidéos lié au contenu dynamique de ces dernières. Il comprend trois grandes phases, la première réalisant la segmentation temporelle de la vidéo en plans et séquences, la seconde concernant l'extraction d'entités mobiles pertinentes pour l'indexation, la troisième portant sur l'indexation proprement dite de l'information extraite. Nous avons complété la méthode de détection des changements de plans [13], par un module de détection des transitions progressives particulières que sont les effets de volets [73, 72]. L'originalité de la solution proposée réside dans l'exploitation de la distribution spatiale des pixels non conformes au mouvement dominant estimé entre images successives. Les résultats expérimentaux obtenus sur de nombreux exemples réels montrent la capacité de la méthode à détecter une grande variété de volets (directions de

déplacement variées, effets générés différents - frontière franche, graduelle, ou enroulement), et à être robuste aux fausses alarmes dans le cas de scènes complexes.

Pour ce qui concerne le deuxième objectif, nous nous sommes intéressés cette année à la détection d'objets mobiles dans des séquences d'images couleur acquises avec une caméra mobile [42]. Nous formulons ce problème comme la recherche des zones de l'image dont le mouvement n'est pas conforme au mouvement dominant dans l'image supposé dû au mouvement de la caméra. Pour ce faire, nous procédons en trois étapes. Un graphe d'adjacence de régions spatiales est déduit d'une partition spatiale de l'image au sens de la couleur. Nous estimons par ailleurs un modèle affine 2D global représentant le mouvement dominant entre deux images successives. La détection des objets mobiles se ramène alors à l'étiquetage des noeuds du graphe considéré en deux classes correspondant aux régions spatiales conformes ou non au mouvement dominant. Cette formulation du problème par un étiquetage bayésien permet de conserver la topologie du graphe et ainsi d'autoriser un étiquetage réversible (contrairement aux approches par fusion de régions), et de ne pas introduire de dépendance par rapport à l'ordre de considération des régions. Cette étape exploite une modélisation markovienne originale sur le graphe d'adjacence des régions, exploitant d'une part des mesures locales de mouvement appropriées et incorporant d'autre part une information contextuelle au niveau des régions. Du fait de l'utilisation d'une partition spatiale au sens de la couleur, cette approche permet d'obtenir une meilleure localisation des frontières de mouvement.

Des travaux sont par ailleurs entrepris concernant des algorithmes d'indexation proprement dits, et leur adaptation aux problèmes spécifiques posés par les données multimédia : imprécision, grande dimension, évolution possible des bases et des descripteurs utiles.

### 6.1.6 Modèles probabilistes de formes en imagerie médicale

**Participants :** Christian Barillot, Isabelle Corouge.

Notre objectif est d'établir un cadre générique permettant la construction de modèles numériques statistiques de formes en imagerie IRM et échographique à partir d'un apprentissage sur une base d'entités préalablement segmentées, ces formes pouvant être surfaciques ou volumiques. Nous avons tout d'abord considéré les sillons du cortex cérébral extraits d'images IRM. Nous exploitons un ensemble de modèles reconstruits de ces sillons pour constituer une base de déformations à partir d'une forme moyenne obtenue après recalage rigide, de chaque exemplaire traité, dans un référentiel local commun (axes d'inertie, minimisation des distances entre sillons, etc.). Les déformations d'une entité donnée sont caractérisées par des vecteurs de déplacement par rapport au modèle moyen, après recalage rigide global. Les composantes principales de la matrice de covariance des vecteurs de déplacement décrivent les modes principaux de déformation. Ces modes permettent de déterminer une probabilité d'occurrence (présence, localisation, forme, ...) d'un sillon autour de son modèle moyen. Le champ de déformation peut ensuite être appliqué à tout objet (structure, localisation d'activation MEG ou IRMf, etc...) qui aura été associé à ce sillon [29].

Ce travail relève désormais de la thèse d'Isabelle Corouge. Il a pour but de généraliser l'approche à d'autres entités et mesures de déformation, ainsi qu'à des sources de données différentes en l'occurrence les ultrasons 3D et l'imagerie IRM. L'objectif est de mettre en

oeuvre les outils nécessaires pour segmenter et modéliser des structures d'intérêt dans des séquences spatio-temporelles d'images 3D. La modélisation statistique proposée permettra de mieux aborder le problème de la constitution de modèles génériques de formes utilisables dans des contextes variés comme la classification de pathologies, la reconnaissance de structures complexes, la création d'atlas, voire encore pour aider à formaliser les liens entre représentation numérique et représentation symbolique de structures dans un contexte de fouille de données.

## 6.2 Recalage, suivi, trajectographie

*Contexte: Le problème générique de suivi dans des séquences 1D ou 2D de données occupe une place centrale dans de très nombreuses applications. Il consiste à extraire et reconnaître les entités d'intérêt, puis à établir un lien temporel entre les instances successives de ces entités. Les principales difficultés résident dans la variabilité au cours du temps des attributs caractérisant une même entité (allant jusqu'à la disparition temporaire, totale ou partielle de l'entité), dans la présence simultanée éventuelle de plusieurs entités (pouvant être très proches) au sein des données, et dans la présence de "fouillis" autour des données utiles. Nous traitons ce problème à deux niveaux: d'un point de vue général et théorique (choix et représentation des entités suivies, modèle d'évolution temporelle, reconstruction de trajectoires complètes, gestion des occultations et des croisements, prise en compte des changements de topologie des formes suivies); d'un point de vue plus spécialisé dans des contextes d'applications donnés (suivi de structures articulées, suivi de formes nuageuses, suivi pour l'indexation vidéo, suivi pour l'asservissement visuel).*

### 6.2.1 Suivi de structures déformables en imagerie météorologique

**Participants :** Patrick Bouthemy, Etienne Mémin, Christophe Papin.

Nous avons développé une méthode de détection, de suivi et de caractérisation de l'activité de structures nuageuses dites convectives, [67, 12]. Ces structures, responsables de phénomènes météorologiques violents et soudains subissent couramment des fusions ou des scissions au cours de leur évolution. Notre approche est organisée en deux phases distinctes. La phase de suivi exploite un formalisme particulièrement adapté au suivi d'objets déformables avec changement de topologie : celui des ensembles de niveaux introduit par Sethian et Osher [Set96]. Elle comprend deux étapes successives. La première s'appuie sur un champ de mouvement 2D préalablement estimé et propage dans le sens du flot dominant les contours obtenus à l'instant précédent. Elle utilise dans un second temps des informations de nature photométrique afin d'ajuster au mieux les contours des cellules nuageuses. Ces contours sont déformés localement conditionnellement à une température caractéristique prédite relative à la cellule nuageuse considérée. La fonction d'évolution définie a la propriété originale de pouvoir dilater ou comprimer le contour en fonction du contexte photométrique local, ce qui est primordial dans la mesure où les formes suivies se chevauchent d'un instant à l'autre. Dans une deuxième phase, nous cherchons à caractériser l'état d'activité de chacune des structures nuageuses traitées. Nous avons défini deux descripteurs de la convection atmosphérique. Le premier correspond

---

[Set96] J. SETHIAN, *Level Set Methods*, Cambridge University Press, 1996.

à la divergence du mouvement apparent et est calculé à partir du champ de mouvement 2D estimé. Le second représente la variation temporelle de température. L'étude du comportement couplé de ces deux descripteurs permet de qualifier la tendance du mouvement vertical associé aux cellules d'origine convective. Cette tâche est formulée comme un problème de segmentation markovienne à trois classes : deux dites d'activité, nommées "ascendance" et "subsidence" et traduisant des mouvements verticaux dans l'atmosphère, et une appelée "non-active" ou encore "indéterminée". Les résultats obtenus sont présentés selon une codification simple, ce qui permet aux prévisionnistes d'appréhender rapidement la situation atmosphérique. Cette approche a été testée et validée sur un grand nombre de séquences d'images infrarouges thermiques ou de vapeur d'eau acquises par les satellites géostationnaires Meteosat et Goes. Cette étude a été menée en collaboration avec Météo-France (voir sous-section 7.3).

### 6.2.2 Suivi en imagerie par filtrage particulière

**Participants** : Jean-Pierre Le Cadre, Patrick Pérez, Carine Hue.

Pour traiter les problèmes de suivi, des méthodes probabilistes à base de filtrage de Kalman sont généralement employées. De telles techniques présentent des limitations notables pour gérer les non-linéarités des modèles introduits (de mesure en particulier). Elles se heurtent en outre au problème d'entités multiples qui nécessite la manipulation de lois multimodales. Pour pallier ces problèmes, les techniques d'*association de données*, basées sur des mélanges de lois normales, ont été développées dans le domaine de l'extraction multipistes. Plus récemment, des techniques dites de "*condensation*" (conditional density propagation) ont été introduites en vision par ordinateur. Apparentées aux techniques de *filtrage particulière*, ces méthodes reposent sur la manipulation de jeux d'échantillons assortis de poids. Dans le cadre des applications de suivi intéressant le projet (extraction de pistes en sonar et radar, extraction d'entités cinématiques telles que les vortex en imagerie de phénomènes fluides, suivi de véhicules en trafic routier,...), nous nous proposons d'étudier, de comparer et d'étendre ces diverses techniques de filtrage non linéaire. Ce travail fait l'objet de la thèse de Carine Hue.

### 6.2.3 Suivi hybride 2D-3D

**Participants** : Patrick Bouthemy, François Chaumette, Éric Marchand.

Nous avons développé une méthode originale de suivi d'objets complexes approximativement modélisés par un polyèdre [58, 56]. L'approche repose sur l'estimation du mouvement de l'objet dans l'image, ainsi que sur un calcul de pose. Nous avons développé une méthode exploitant un modèle CAO "frustré" de la pièce manufacturée suivie, et un calcul de pose à partir d'un calibrage approximatif de la caméra. Elle enchaîne deux étapes de transformation globale, la première à caractère 2D, la seconde à caractère 3D. Un premier recalage de la silhouette suivie, pouvant appréhender des grands déplacements, est réalisé via l'estimation d'un modèle affine 2D de mouvement à l'aide d'une méthode robuste. Cette dernière exploite en entrée les composantes perpendiculaires au contour du déplacement des points <sup>[Bou89]</sup> de la projection

---

[Bou89] P. BOUTHEMY, « A Maximum-Likelihood Framework for Determining Moving Edges », *IEEE Trans. on Pattern Analysis and Machine Intelligence* 11, 5, mai 1989, p. 499-511.

des arêtes du modèle CAO. Nous pouvons ensuite évaluer les paramètres d'une transformation 3D par la minimisation itérative d'une fonction de coût non linéaire. Elle consiste à positionner au mieux la projection des arêtes du modèle CAO de l'objet sur les contrastes d'intensité dans l'image avec prise en compte de l'orientation de ces derniers. On utilise alors explicitement dans le suivi la propriété de rigidité des objets. Le résultat de l'étape précédente fournit une initialisation correcte du calcul de pose. Cette étape inclut de plus la gestion des faces apparaissant et disparaissant au cours du temps. Les deux étapes (estimation du mouvement 2D et de la pose) ne requièrent pas de phases explicites de segmentation, notamment d'extraction de contours dans les images successives. Elles sont robustes aux occultations partielles de l'objet. Finalement, l'algorithme peut être exécuté à une cadence très rapide (10 Hz sur un PC à 450 MHz). Des expérimentations nombreuses sur la cellule robotique du projet ont confirmé les performances de cet algorithme de suivi et son utilisabilité en asservissement visuel. Cette étude a été menée en relation avec EDF (voir sous-section 7.1).

#### 6.2.4 Recalage 3D non linéaire en imagerie cérébrale IRM et IRMf

**Participants :** Christian Barillot, Etienne Mémin, Patrick Pérez, Pierre Hellier.

Afin de calculer un champ de déplacements entre les images des cerveaux de deux patients (recalage non linéaire mono-modalité), nous faisons coopérer deux approches, l'une globale et l'autre locale. Dans la méthode globale, [51, 49], nous estimons un champ dense de déformations 3D par des techniques de type "flot optique" [MP98b]. L'hypothèse de conservation de la luminance étant perturbée par les artefacts d'acquisition des images IRM (bruit, inhomogénéité du champ magnétique), nous introduisons des estimateurs robustes. D'autre part, les estimateurs robustes sont également utiles pour préserver les discontinuités du champ de déformation, car les cerveaux de deux patients différents ne sont pas topologiquement équivalents. Enfin le schéma d'optimisation est à la fois multirésolution et multigrille afin d'accélérer le traitement et d'améliorer la qualité finale du champ obtenu. Nous avons réalisé une validation numérique des performances de cet algorithme de recalage sur le "fantôme" de l'université Mc Gill de Montréal. L'algorithme de recalage défini a été utilisé pour compenser un champ de déformation synthétique et nous avons mis en œuvre une évaluation basée sur des mesures de recouvrement des matières grises et blanches. Nous avons ainsi pu montrer que même en présence de bruit fort et de déformations violentes, nous pouvons compenser jusqu'à 95% de la déformation. D'autre part la même méthode, sans la relaxation multigrille, nous fournit un algorithme de recalage rigide monomodal. Deux champs d'applications sont alors possibles: la correction de bougé pendant une acquisition dynamique (IRMf par exemple), et le recalage de volumes SPECT acquis au repos et pendant une crise d'épilepsie (SPECT Ictal et Inter-Ictal). L'approche locale [LGBB97] consiste à soumettre le champ recherché à des contraintes anatomiques de type sillons [LGBB97] afin d'appréhender au mieux la grande variabilité inter-individuelle, et en

---

[MP98b] E. MÉMIN, P. PÉREZ, « A multigrid approach to hierarchical motion estimation », *in: Proc. Int. Conf. on Computer Vision, ICCV'98*, p. 933–938, Bombay, Inde, Janvier 1998, [http://www.irisa.fr/vista/Papers/1998/memin\\_ICCV.ps.gz](http://www.irisa.fr/vista/Papers/1998/memin_ICCV.ps.gz).

[LGBB97] G. LE GOUALHER, C. BARILLOT, Y. BIZAIS, « Three-dimensional segmentation and representation of cortical sulci using active ribbons », *Int. J. of Pattern Recognition and Artificial Intelligence* 11, 8, 1997, p. 1295–1315.

particulier les changements de topologie. La formulation bayésienne permet d'intégrer dans un même contexte ces deux approches et de préserver le caractère local de la contrainte anatomique.

D'autre part nous développons une version multimodale de cet algorithme afin d'effectuer en particulier un recalage non-rigide  $IRM-IRMf$ . L'acquisition en  $IRMf$  peut en effet subir de fortes distorsions géométriques qui faussent l'interprétation de ces données si un recalage rigide seul est effectué. Pour effectuer le recalage, nous utilisons une mesure de similarité adaptée (information mutuelle, rapport de corrélation [RMAP98]).

### 6.2.5 Recalage et suivi de structures en imagerie ultrasonore

**Participants :** Christian Barillot, Patrick Bouthemy, Ioannis Pratikakis, Caroline Baillard.

L'étude concerne le couplage entre imagerie ultra-sonore 3D et neuro-navigation dans un contexte d'imagerie per-opératoire en neurochirurgie. Le principe consiste à segmenter les données  $IRM$  pré-opératoires, afin d'initialiser la localisation des structures anatomiques de référence (sillons, ventricule) visibles dans l'image échographique 3D per-opératoire. Le recalage multi-modal s'appuie sur l'information fournie par les capteurs optiques du système de neuro-navigation. Ces structures sont ensuite suivies dans la séquence ultrasons 3D au cours de l'opération pour mettre à jour le planning chirurgical. Deux aspects ont été abordés. Le premier concerne l'acquisition des données. Une interface entre le logiciel StradX fourni par l'université de Cambridge et l'équipement de neuro-navigation du CHU de Rennes<sup>1</sup> a été développée afin de référencer les acquisitions ultrasonores dans la géométrie du système de neuro-navigation. Des capteurs optiques liés à la sonde échographique sont utilisés. Nous avons d'autre part traité une séquence de données volumiques  $IRM$  et ultrasons 3D, fournie et recalée par le projet Epidaure. Nous avons extrait des structures similaires entre données  $IRM$  et ultrasons 3D (ventricules par ex.), et généré des déformations synthétiques pour simuler l'influence de l'action opératoire. La compensation de cette déformation se fait par la méthode de calcul de champs denses multigrilles développée dans le projet.

Un autre objectif que nous venons d'aborder est la détection directe d'entités de forme complexe dans des données volumiques ultrasonores et le suivi de ces structures dans des séquences 3D+t.

### 6.2.6 Extraction de pistes et poursuite

**Participants :** Jean-Pierre Le Cadre, Isabelle Leibowicz.

Un premier volet est relatif à l'analyse de l'estimabilité, une mesure de la précision d'estimation, de la trajectoire d'une source mobile. Nous avons développé un formalisme original qui permet d'analyser les performances de systèmes relativement complexes (e.g. multiplateformes,

---

1. Neuronavigateur de type FlashPoint 5000.

---

[RMAP98] A. ROCHE, G. MALANDAIN, N. AYACHE, X. PENNEC, « Multimodal Image Registration by Maximization of the Correlation Ratio », *rapport de recherche n° RR-3378*, INRIA, 1998.

par temps de retards, etc.). Nous l'avons ensuite appliqué au problème de la fusion de mesures d'origines différentes, telles que radar et infrarouge, actif et passif,... Nous obtenons une bonne corrélation de l'analyse théorique et des résultats de simulation. Cette analyse a ensuite été étendue à l'estimation de trajectoires complexes et à l'analyse de systèmes de détection et de poursuite multistatique.

Le deuxième volet a trait à l'approximation des performances en extraction de multipistes. Le formalisme utilisé est celui des lois de mélange de densités, où il s'agit d'estimer à la fois les variables d'affectation des mesures aux pistes et les vecteurs d'états des sources. Celui-ci est à la base des méthodes de type PMHT. L'analyse repose sur les propriétés algébriques de la matrice de Fisher. Le problème réside dans la dimension du problème et la présence des termes d'interaction des pistes, variables dans le temps. Il a été résolu par la mise en œuvre d'un formalisme basé, d'une part, sur l'utilisation des coordonnées polaires modifiées, et d'autre part, sur les propriétés des produits de Kronecker. Les résultats ainsi obtenus sont simples et relativement généraux. La simplicité de cette analyse est cependant due à l'"oubli" des problèmes combinatoires. C'est pourquoi notre effort s'est aussi porté sur l'analyse de critères de performances tels que la probabilité d'association correcte en extraction multipistes, la pureté de piste, etc. . .

active et exploration

### 6.3 Asservissement visuel, perception active et exploration

*Contexte: La perception active consiste à élaborer des stratégies de perception et d'action dans le but d'améliorer les performances des algorithmes de vision ou de détection par le contrôle des paramètres du capteur, ou de réaliser des tâches robotiques (positionnement, saisie, suivi, . . .). Nous nous intéressons au contrôle de la position, de la vitesse, ou de la trajectoire de capteurs de type vidéo ou sonar. Deux types de stratégies sont considérés et combinés : d'une part, des stratégies locales, liées aux techniques d'asservissement visuel et fournissant une perception de bas niveau, et d'autre part, des stratégies globales, activées sur des événements particuliers et remontant à une perception de haut niveau.*

#### 6.3.1 Nouvelle représentation pour l'asservissement visuel

**Participant :** François Chaumette.

Cette étude a pour objectif de déterminer une représentation optimale et générique en asservissement visuel 2D. Il s'agit de sélectionner dans l'ensemble des informations visuelles qu'il est possible d'extraire et de suivre dans une séquence d'images, un nombre minimal d'entre elles, offrant de bonnes propriétés : découplage des degrés de liberté du système, robustesse de la commande aux erreurs de mesure, absence de singularités et de minima locaux. L'ensemble de ces propriétés peut être étudié à partir de la matrice d'interaction associée aux informations visuelles. Nous avons ainsi déterminé la forme analytique de la matrice d'interaction associée aux moments d'ordre quelconque d'un objet. Cette forme dépend de la valeur des moments dans l'image et des coefficients du polynôme décrivant la surface des limbes de l'objet. La forme obtenue est générale et permet notamment de retrouver les résultats des cas particuliers

des segments et des ellipses. L'objectif de généralité est donc atteint, même si des occultations partielles de l'objet peuvent s'avérer problématiques.

Nous avons ensuite considéré un objet particulier (en l'occurrence un rectangle) et avons pu sélectionner six informations visuelles basées sur les moments telles que, pour une configuration donnée (centrée et parallèle au plan image), la matrice d'interaction associée soit triangulaire, ce qui permet d'assurer le découplage visé dans la commande des degrés de liberté du système. Les expérimentations validant cette représentation sont en cours.

### 6.3.2 Planification dans l'image

**Participants** : François Chaumette, Youcef Mezouar.

Les techniques actuelles d'asservissement visuel 2D, même si elles sont efficaces pour une grande classe d'applications, se heurtent à des difficultés lorsque la position initiale de la caméra est fortement éloignée de sa position désirée. Nous nous proposons de résoudre ce problème par la planification de trajectoires dans l'image de plusieurs points d'un objet. Le suivi de ces trajectoires entraîne alors une faible erreur entre la consigne et la mesure courante, cadre où les techniques d'asservissement visuel 2D sont connues pour être stables et robustes aux diverses erreurs de modélisation. Pour générer de telles trajectoires dans l'image, nous avons adopté la méthode des fonctions de potentiel, classique dans le domaine de la planification. La fonction de potentiel que nous avons modélisée est définie dans l'espace opérationnel du robot et admet un pôle attractif (la configuration désirée) et un pôle répulsif (les configurations indésirables telles celles pour lesquelles la cible sort du champ de vision de la caméra). Il suffit alors de projeter l'objet d'intérêt dans l'image pour l'ensemble des positions planifiées de la caméra pour obtenir les trajectoires 2D souhaitées. Une étude sur la sensibilité aux erreurs de calibration de la caméra a été effectuée. De plus, nous avons généralisé cette méthode au cas où la forme et les dimensions de l'objet sont inconnues. Les résultats expérimentaux obtenus sur la cellule de vision robotique du projet se sont avérés, comme attendu, très satisfaisants.

### 6.3.3 Asservissement visuel et animation

**Participant** : Éric Marchand.

Les problèmes associés à la gestion d'une caméra dans un environnement 3D virtuel sont nombreux. Il faut non seulement pouvoir réaliser une tâche visuelle (souvent une tâche de focalisation ou de manière générale une tâche de positionnement) de manière efficace, mais il faut en plus être en mesure de réagir de manière appropriée à des modifications de l'environnement. Nous avons choisi d'utiliser le savoir-faire acquis dans le domaine de l'asservissement visuel pour proposer une solution à ce problème. De plus, contrairement aux travaux antérieurs sur ce thème [GW92], nous ne nous sommes pas limités à des tâches de positionnement dans des environnements statiques. Nous avons considéré l'introduction de contraintes dans la commande de la caméra grâce au formalisme de redondance. Nous avons abordé plusieurs aspects incluant des tâches simples, comme la poursuite d'un objet, et des tâches plus complexes

---

[GW92] M. GLEICHER, A. WITKIN, « Through-the-lens camera control », in : *ACM Computer Graphics, SIGGRAPH'92*, p. 331-340, Chicago, July 1992.

comme l'évitement d'occultations ou d'obstacles, ou encore le positionnement par rapport à la face éclairée d'un objet, [39]. Néanmoins, le contrôle se faisant dans l'image, l'animateur ne maîtrise pas la trajectoire 3D de la caméra. De telles techniques sont plutôt pertinentes dans le cadre d'applications temps-réel.

### 6.3.4 Appariements d'objets et de scènes

**Participants :** Patrick Gros, Damien Albert.

Cette étude vise à appliquer les techniques d'appariement d'images et d'indexation à l'asservissement visuel et au guidage de véhicule. Les techniques d'asservissement visuelles pour fonctionner ont besoin d'établir un premier lien entre une image cible et une image courante, ces deux images pouvant, dans les situations réelles, avoir peu d'éléments en commun. L'idée est donc d'utiliser une base d'images établie préalablement, pour trouver un lien entre l'image courante et une des images de cette base, et de se servir de ce premier lien et de connaissances géométriques pour amener la caméra dans une position où elle puisse voir la cible et où l'asservissement visuel puisse effectivement débiter.

### 6.3.5 Couplage vision locale et vision globale

**Participants :** Éric Marchand, François Chaumette, Grégory Flandin.

Les études menées précédemment au sein du projet dans le domaine de l'asservissement visuel et de la vision active concernaient l'utilisation d'une seule caméra. Dans un contexte de scènes plus complexes, encombrées ou comportant des objets mobiles, nous souhaitons à présent généraliser cette approche à l'emploi de deux capteurs, l'un fournissant une vision locale embarquée et l'autre ayant une vue globale de la scène. Dans un premier temps, nous nous sommes intéressés à l'aspect "bas niveau" de l'utilisation de plusieurs caméras dans la boucle de commande d'un robot. Nous avons réalisé une tâche de positionnement où la caméra globale contrôle les degrés de translation du robot et la caméra locale ceux de rotation par focalisation sur une cible.

Nous avons tout d'abord utilisé le formalisme de redondance de l'approche fonction de tâche, qui, après quelques réglages, s'est avéré satisfaisant. Cependant, pour conserver le maximum d'indépendance entre les deux systèmes de vision, nous nous sommes ensuite intéressés à la manière de découpler les deux sous-tâches considérées. En effet, si le mouvement de rotation de la caméra embarquée, lié à la focalisation sur la cible, n'induit pas de changement dans l'image déportée, en revanche, son mouvement de translation induit un mouvement dans l'image embarquée. La non prise en compte de ce couplage dans la modélisation conduit alors à une erreur de traînage dans la tâche de suivi. Pour pallier ce problème et tout en conservant un maximum d'indépendance entre les tâches, nous avons choisi de modéliser l'action du mouvement de translation sur l'image embarquée sous forme d'une perturbation à estimer et à compenser. Les techniques d'estimation et de filtrage que nous avons développées dans le cadre de la poursuite d'un objet mobile peuvent alors s'appliquer.

### 6.3.6 Théorie de l'exploration

**Participants :** Jean-Pierre Le Cadre, Frédéric Dambreville.

Nous nous préoccupons de l'optimisation de l'effort de recherche d'un objet, généralement appelé "cible", qui peut être fixe, mobile (e.g. trajectoire markovienne) ou même réactif. On dispose de ressources, nécessairement limitées, et on cherche à optimiser des fonctions telles que la probabilité de détection, le temps de détection ou encore une fonction de risque. Notre objectif a été, dans un premier temps, de mettre en oeuvre les méthodes (connues) de la théorie de l'exploration. On notera parmi les problèmes traités : l'optimisation de la trajectoire de recherche par des méthodes de type "branch and bound", l'optimisation de l'effort de recherche pour un problème à  $n$ -étapes (algorithme "forward-backward"). Bien que les résultats obtenus correspondent le plus souvent à l'intuition, l'aspect quantitatif est évidemment important. On notera aussi que ces méthodes ne nécessitent généralement pas une charge de calcul déraisonnable. Dans un registre plus innovant, notre effort s'est porté sur l'utilisation conjointe de ce formalisme et de celui de l'association de données pour étendre l'analyse à la détection de pistes (et non d'objets ponctuels). Le formalisme utilisé est celui de l'optimisation convexe discrète (variables discrètes et dualité). Des résultats originaux, relatifs à l'optimisation de l'effort de recherche d'une piste, ont ainsi été obtenus. La fonction de détection fait intervenir une logique d'association de détections élémentaires (règles "AND ou OR", ou encore " $p$  sur  $n$ "), conditionnellement à une hypothèse cinématique. Bien que la fonctionnelle de détection ne soit pas concave, on peut se ramener assez aisément à un problème dual qui possède de bonnes propriétés de convexité. Sur un plan pratique, la complexité des algorithmes s'accroît fortement, mais le coût de calcul demeure raisonnable. On voit cependant qu'il n'est pas nécessaire de compliquer beaucoup le cadre de l'analyse (e.g. "logique" d'association des décisions) pour arriver à des problèmes très difficiles. C'est le cas lorsque l'on considère des cibles réactives ou encore une fonction de risque liée à l'effort de recherche.

La thèse de Frédéric Dambreville, [41, 74], s'est orientée dans un premier temps vers ce type de problèmes. Alors que dans les travaux actuels l'effort de recherche est fixé à chaque étape, on se pose la question de l'optimisation des allocations des quantités d'effort de recherche aux différentes étapes temporelles. Diverses contraintes de renouvellement de l'effort de recherche sont considérées dans un cadre unique. Après avoir obtenu les conditions générales d'optimalité, nous avons développé des approches algorithmiques originales. La répartition des efforts de recherche aux différentes étapes temporelles dépend non seulement du type de trajectoire de la cible (e.g. diffusion) mais aussi de la répartition spatiale, à chaque étape, des efforts de recherche élémentaires. Cette étude a été illustrée par plusieurs types d'applications. À ce travail s'ajoutent des études plus théoriques relatives aux équations différentielles associées à la répartition de l'effort de recherche et aux aspects dynamiques en théorie de la recherche (e.g. prise en compte de détections multiples).

## 7 Contrats industriels (nationaux, européens et internationaux)

### 7.1 Convention EDF : Suivi 2D de formes complexes avec exploitation de modèles CAO

**Participants :** Patrick Bouthemy, François Chaumette, Éric Marchand.

no. 1 97 C 234, durée 30 mois.

Il apparaît primordial pour introduire les techniques d'asservissement visuel dans un large éventail d'applications, de pouvoir appréhender des scènes naturelles, à savoir, sans marqueurs, objets non polyédriques, conditions d'illumination variables, . . . Nous nous intéressons ainsi au suivi 2D d'objets dans un environnement non maîtrisé dans un contexte de commande référencée vision. Nous souhaitons satisfaire des critères à la fois de généralité suffisante et de complexité calculatoire faible. La structure 2D à suivre correspond à la projection dans l'image d'objets relativement complexes, comme par exemple un objet métallique. Les situations devant être traitées présentent des caractéristiques difficiles comme zone brillante sur zone brillante, ombres portées, arêtes arrondies, occultations partielles, éclairage variable.

Nous avons achevé cette étude cette année. La version finale de l'algorithme de suivi enchaîne deux étapes de transformation globale, la première à caractère 2D, la seconde à caractère 3D (voir [56] ou le paragraphe 6.2.3. pour plus de détails). Nous avons considéré à la demande d'EDF plusieurs objets manufacturés, et l'algorithme a été testé lors d'expérimentations nombreuses sur la cellule robotique de l'Irisa en considérant de fortes variations des conditions expérimentales (lumière, vitesse du robot, etc.). Ces expériences ont confirmé les performances de cet algorithme de suivi tant 2D que 3D. La fréquence du suivi (entre 2 et 10 Hz suivant l'objet considéré) permet d'utiliser cet algorithme dans des schémas d'asservissement visuel. Comme l'ont démontré des tests de répétabilité, une précision de l'ordre de 0.7 mm en translation et 0.2 deg en rotation a pu être atteinte.

d'asservissement  
visuel

### 7.2 Convention Edixia : Industrialisation de techniques d'asservissement visuel

**Participants :** Fabien Spindler, François Chaumette.

no: 1 99 C 140, durée: 6 mois.

Cette convention avec la société Edixia, menée pour la DER-EDF Chatou, a porté sur l'amélioration des logiciels d'asservissement visuel transférés les années précédentes par l'apport de nouvelles fonctionnalités: calibration grossière automatique, calculs de pose 3D par les méthodes de Lagrange, Dementhon et Lowe, test de convergence dans l'espace cartésien (qui s'ajoute au test de convergence dans l'image déjà présent), ajout d'une loi de commande supplémentaire 2D 1/2 (pour que la caméra effectue une ligne droite tout en s'orientant de manière à conserver l'objet d'intérêt dans l'image). Edixia a pour sa part reconçu l'IHM de l'application, réalisé une meilleure gestion de l'éclairage à leds automatique, et apporté une plus grande robustesse aux phases de reconnaissance et de suivi des objets. Ce système a été transféré à la

DER-EDF et intégré avec succès à une expérimentation de pose et de dépose d'un obturateur de bol de générateur de vapeur d'une centrale nucléaire.

satellitaires appliquée à l'observation météorologique

### 7.3 Convention Météo-France : Analyse spatio-temporelle d'images satellitaires appliquée à l'observation météorologique

**Participants** : Christophe Papin, Patrick Bouthemey, Étienne Mémin.

no. 1 97 C 818, durée 36 mois.

Cette collaboration avec Météo-France, et plus précisément avec le CMS (Centre de Météorologie Spatiale) de Lannion et la DIRO de Rennes (centre de prévisions météorologiques) sur trois ans correspond au doctorat de C. Papin [12], et se termine fin 1999. Elle a permis de développer de nouvelles méthodes d'analyse d'images satellitaires météorologiques, afin d'accroître qualitativement et numériquement l'ensemble des informations fournies aux prévisionnistes concernant certaines configurations météorologiques critiques. Nous avons concentré nos efforts sur l'étude de deux situations importantes portant, d'une part, sur la détection précoce des nuages bas (et des brouillards) de nuit dans les images infrarouges thermiques acquises par le satellite Meteosat, et d'autre part, sur la détection, le suivi et la caractérisation des cellules nuageuses convectives. Les résultats obtenus sur le premier point ont été décrits dans le rapport d'activité de l'année 1998 ainsi que dans différents articles dont [PBR98]. Sur le second point, on peut se référer au paragraphe 6.2.1, ainsi qu'à l'article [67]. Tous ces développements originaux ont été validés sur de très nombreux exemples réels représentatifs d'une grande variété de situations météorologiques, et ont fourni des résultats très satisfaisants. Ils font de plus l'objet d'une évaluation équivalente à un mode opérationnel sur une semaine, courant novembre, par les prévisionnistes du centre de Rennes. L'ensemble des logiciels seront transférés à Météo-France.

### 7.4 Convention Alcatel-CRC : Indexation vidéo

**Participants** : Fabien Spindler, Patrick Bouthemey.

no. 1 98 G 292, durée 10 mois

Cette convention avec Alcatel - CRC, qui concernait principalement le projet Movi de l'Inria Rhône-Alpes, s'est clos au printemps 99. Pour ce qui nous concerne, elle a consisté dans la mise à disposition de logiciels, et dans leur intégration à un environnement d'indexation vidéo (dénommé VideoPrep) pour la création de documents vidéo structurés et interactifs. De tels documents contiennent ainsi la possibilité de génération d'actions suite à une désignation dans l'image d'une zone reconnaissable préalablement indexée (environnement VideoClic). Dans un contexte de préparation de la vidéo dans un mode semi-automatique, ce qui est le bon point de vue dans ce type d'application, nous avons aussi développé une technique de suivi, avant ou arrière au sein d'un plan, d'une zone quelconque préalablement détournée par l'opérateur

---

[PBR98] C. PAPAN, P. BOUTHEMY, G. ROCHARD, « Segmentation of low clouds in Meteosat IR night-time images based on a contextual spatio-temporal labeling approach », *in: Conf. AMS-Eumetsat'98*, 2, p. 370-373, Paris, mai 1998.

[47]. Le suivi est automatique et temps-réel. Un mécanisme détectant les décrochages dans le suivi a été de plus défini et intégré, permettant d'alerter l'opérateur pour une réinitialisation éventuelle de la procédure.

## 7.5 Convention DGA - Thomson-CSF-Optronique : Conception d'une chaîne de traitements pour la restauration de séquences d'images

**Participants :** Patrick Bouthemy, Patrick Pérez, Fabien Dekeyser.

no. 1 98 C 500, durée 24 mois

Cette étude est menée en collaboration avec Thomson-CSF-Optronique avec un financement DGA, dans le cadre du programme Civil-Défense. Deux applications sont prises en compte, l'une relative au traitement de séquences d'images infra-rouges aériennes, l'autre au traitement d'images médicales endoscopiques vidéo. Notre rôle consiste d'une part en l'étude et le développement d'algorithmes originaux de filtrage spatio-temporel, et d'autre part en une assistance pour l'intégration et l'évaluation des divers éléments de la chaîne de traitement. L'objectif est non seulement de débruiter ces images ou de supprimer des défauts localisés, mais également de construire des images de plus grande résolution à partir de la séquence originale. Il est évident qu'un filtrage spatio-temporel efficace doit tenir compte du mouvement présent dans la séquence d'images. Nous avons cette année construit un schéma séquentiel produisant, dans une première étape, une estimation de mouvement et réalisant, dans une seconde étape, un filtrage spatio-temporel après compensation du mouvement estimé. Deux types de mesures de mouvement ont été pris en compte : une estimation paramétrique globale du mouvement dominant d'une part, une estimation d'un champ dense de vitesses d'autre part. Plusieurs techniques de filtrage ont été explorées et comparées sur des exemples réels, la plus intéressante apparaissant être une formulation par filtrage de Wiener dans le domaine fréquentiel (voir paragraphe 6.1.3).

## 7.6 Convention Adfac-Eumetsat : Estimation de champs de vents sur nuages bas

**Participants :** Patrick Pérez, Étienne Mémin, Stelio Zimeras.

no. 1 99 C 351, durée 15 mois.

Dans le cadre des études lancées par le consortium météorologique européen Eumetsat autour de l'exploitation des images du futur satellite Meteosat Seconde Génération (MSG), le présent travail mené en collaboration avec le LMD (Laboratoire de Météorologie Dynamique, Palaiseau) vise à développer de nouveaux outils d'estimation de champs de vents à basse altitude. Il s'agit plus particulièrement de coupler des méthodes de détection de nuages bas sur images multispectrales, de suivi d'entités photométriques et/ou cinématiques, et d'estimation de champs denses de déplacements. Les données traitées sont des séquences GOES (satellite météorologique américain), dont les caractéristiques sont proches de celles des futures séquences d'images qui devront être fournies par le satellite MSG.

du logiciel Motion-2D sur une plate-forme temps-réel

### 7.7 Convention Ifremer : Contribution au portage du logiciel Motion-2D sur une plate-forme temps-réel

**Participants** : Patrick Bouthemy, Fabien Spindler.

no. 1 99 C 015, durée 12 mois

Cette convention a donné lieu au transfert à l'Ifremer du logiciel déposé à l'A.P.P. MOTION-2D, version optimisée du logiciel RMRM d'estimation de modèles paramétriques 2D de mouvement [9]. L'objectif était d'une part de le porter sur une plate-forme temps-réel, et d'autre part de l'exploiter, conjointement à des données inertielles, dans la création de mosaïques vidéo du fond sous-marin pendant l'exploration d'un véhicule sous-marin. Notre tâche a consisté à effectuer un suivi-conseil de ces développements menés à l'Ifremer Toulon. Une version opérationnelle, aux résultats probants, de ce logiciel de création de mosaïques est désormais disponible et va être utilisée prochainement dans des campagnes scientifiques en mer.

traitements multi-capteurs

### 7.8 Convention Cifre Thomson-RCM : Méthodes et algorithmes de traitements multi-capteurs

**Participants** : Jean-Pierre Le Cadre, Isabelle Leibowicz.

no. KTNR, durée 36 mois.

L'objet de la thèse d'Isabelle Leibowicz est l'optimisation des fonctions de surveillance et d'identification d'un avion de patrouille maritime. Un tel système constitue un concentré de traitements et de problèmes multi-récepteurs et multi-cibles. Ainsi, si l'avion dispose de nombreux récepteurs, radars, ESM passif, infrarouge, sonar, il doit aussi traiter, extraire, associer, identifier un grand nombre de pistes. Les problèmes relatifs à l'association piste-à-piste et des mesures aux pistes, à l'apport d'informations d'identification, à la combinaison des diverses sources d'information radar, ESM, infrarouge, ont constitué une partie importante de l'activité. Les résultats obtenus mettent en évidence l'importance de l'information d'identification, essentiellement radar. L'accroissement des performances, en termes de probabilité d'association correcte, a pu être quantifié. Un autre volet important est relatif à l'utilisation des données ESM capteur passif, pour les fonctions de pistage. L'étude des performances des méthodes de pistage a été poursuivie. On doit mentionner aussi les tests d'acceptation de pistes par des critères de pureté de pistes, la résolution des problèmes d'association, les méthodes de réduction de l'information et, plus généralement, l'analyse du système.

spatio-temporelles pour l'indexation vidéo

### 7.9 Convention Cifre Ina : Reconnaissance de formes spatio-temporelles pour l'indexation vidéo

**Participants** : Patrick Bouthemy, Emmanuel Veneau.

no. 1 99 C 222, durée 36 mois.

Cette convention concerne la thèse de Emmanuel Veneau, encadrée à l'Ina par Rémi Ronfard et qui a débuté en mars 1999. Cette étude s'inscrit dans le domaine de l'indexation de

vidéo par le contenu. Elle s'intéresse plus précisément à la reconnaissance de formes spatio-temporelles dans le cadre d'une certaine sémantique d'événements audio-visuels, liée en particulier aux mouvements et actions de personnages, et plus généralement d'éléments de la scène d'intérêt. Dans un premier temps, le travail a porté sur une structuration temporelle, de niveau intermédiaire, de la vidéo : le regroupement des plans élémentaires détectés en séquences devant représenter des entités plus pertinentes et plus exploitables pour la phase de reconnaissance d'événements spatio-temporels.

## 7.10 Actions financées par la Commission Européenne

### 7.10.1 Projet Esprit (R&D MS) Divan

**Participants :** Didier Zugaj, Patrick Bouthemy, Fabien Spindler.

no. 1 97 C 696, durée 30 mois

Nous participons, avec le projet Sigma-2, au projet européen Divan ("Distributed audio-Visual Archives Network"), relevant du programme Esprit R&D Multimedia Systems, qui a débuté au 1er septembre 97. Le maître d'œuvre en est la société Intracom (Gr.), les autres partenaires sont Tecmath (All.), la RAI (It.), l'ERT (Gr.), Forth-ICS (Gr.), THD-ITO (All.), l'Ina et l'Inria. Nous intervenons sur des questions d'indexation de vidéo par le contenu spatio-temporel des images. Une présentation du projet Divan peut être consultée à l'adresse Web suivante <http://divan.intranet.gr/info>. Ce projet s'intéresse à des problèmes de consultation et d'exploitation d'archives audiovisuelles [31]. Ceci nécessite de pouvoir répondre à des requêtes liées aux contenus des vidéos, et donc implique en amont la réalisation d'une indexation par le contenu. Il est envisagé de plus la consultation des archives audio-visuelles dans un mode distribué sur le réseau.

Nous avons contribué aux spécifications des modules d'extraction d'informations visuelles et d'indexation des vidéos. Nous avons porté dans le système développé nos logiciels de découpage de la vidéo en plans, de caractérisation du mouvement de la caméra, et de détection des éléments mobiles. Nous les avons adaptés pour une insertion adéquate dans la phase de décodage du flux MPEG et dans les schémas de description retenus. Nous avons de plus développé et validé une extension de l'algorithme de découpage en plans pour inclure la détection des transitions progressives de type volets. Elle consiste à considérer d'autres informations relatives au support associé au mouvement dominant estimé que sa taille. En l'occurrence, nous exploitons des caractéristiques particulières de la géométrie des zones non conformes au mouvement dominant, lors de l'utilisation de volets dans la génération de transitions progressives au sein de la vidéo. Nous avons aussi conçu une technique robuste et temps réel de détection de logos dans les images d'un plan (logos de chaînes de télévision par exemple). Nous développons enfin un module permettant de manière simple de caractériser le "niveau d'activité" dans la scène filmée.

## 7.11 Actions financées par le RNRT

### 7.11.1 Projet Agir

**Participants :** Patrick Bouthemy, Patrick Gros, Fabien Spindler, Abdeljabar Benzougar.

no. 2 99 C 006, durée 18 mois

Le projet AGIR est un projet RNRT précompétitif qui a officiellement débuté au 1er septembre 1999. Il regroupe Alcatel-CRC, Arts Vidéo Interactive, l'Ina, pour l'Irisa les projets Sigma-2 et Vista, le projet Movi de l'Inria Rhône-Alpes, et des équipes de recherche de l'Irit, du Clips-Imag, de l'INT, et du LIP6. Ce projet est financé pour 18 mois, avec un renouvellement possible de 12 mois.

Le but du projet est de développer un système complet permettant d'indexer des vidéos, éventuellement accompagnées de textes, puis d'utiliser cette indexation pour rechercher des documents par leur contenu, ceci en mélangeant les divers médias. Ce projet prend donc naturellement la suite du travail fait dans le cadre du projet européen Divan, avec un accent plus fort mis sur l'aspect multimédia des données, des index et des requêtes, sur l'architecture globale du projet et sur les liens avec les efforts de normalisation actuels (MPEG-7).

Les travaux développés par Vista concerneront principalement la structuration temporelle automatique des séquences d'images par l'utilisation directe des informations issues de la compression MPEG-2, comme les vecteurs de mouvement, l'extraction et le suivi d'entités mobiles pertinentes et la fourniture de descripteurs d'activité et de mouvement. Le projet intervient aussi, mais de manière moins forte, dans la mise au point d'algorithmes de macro-segmentation et de regroupement de plans en entités de plus haut niveau sémantiquement plus significatives. Par ailleurs, les sorties des modules développés seront exploités dans d'autres tâches du projet.

## 8 Actions régionales, nationales et internationales

### 8.1 Actions régionales

#### 8.1.1 GIS GSTB "Groupement de Télédétection en Bretagne"

**Participants :** Patrick Pérez, Annabelle Chardin.

no. 2 98 C 306, durée 24 mois.

Le GSTB regroupe le Ceva, l'Ensar, l'ENSTBr, l'Ifremer, l'Inria, l'UBO, l'UHB (Costel) et l'université Rennes 1 (Enssat). Sur la base de financements européens Feder et régionaux, ce groupement d'intérêt scientifique, mis en place en 1997, a pour vocation de concentrer les moyens et fédérer les compétences régionales en matière de télédétection. Outre l'achat et l'exploitation d'un capteur aéroporté CASI, ce groupement développe une synergie entre "thématiciens" et "traiteurs d'images" autour de trois opérations pluri-annuelles de recherche d'intérêt régional. Le projet Vista s'est impliqué plus particulièrement aux côtés du Costel (université UHB Rennes 2) dans une étude concernant l'analyse des structures paysagères et de l'occupation des sols dans la baie de Lannion. Cette étude a pris fin en 1999.

Les nouveaux outils d'inférence discrète hiérarchique que nous avons définis (voir paragraphe 6.1.1.) ont été exploités dans la dernière phase de l'étude que nous avons menée avec le laboratoire Costel (université Rennes 2) dans le cadre du GSTB, sur la classification d'images Landsat multispectrales et multitemporelles. Grâce à la nomenclature détaillée et au jeu d'échantillons de classification associé (validés par test sur le terrain) constitués par le Costel, nous avons pu valider le module d'estimation des paramètres du modèle Markovien hiérarchique de classification.

## 8.2 Actions nationales

### 8.2.1 Pôle CNRS Micro-robotique : Asservissement visuel et micro-manipulation

**Participants** : Éric Marchand, François Chaumette.

no. 2 L 008, durée 24 mois.

L'objectif de cette étude est de participer à la réalisation d'un démonstrateur de micro-manipulation automatique. Cette étude est menée en collaboration avec le Laboratoire de Robotique de Paris (LRP) et avec l'antenne de Bruz de l'École Normale Supérieure de Cachan. En complément des travaux des autres laboratoires sur la définition et le développement de préhenseurs et de manipulateurs dédiés aux objets particuliers manipulés, notre rôle est de définir les techniques de commande des mouvements de haute précision, à la fois des capteurs et des actionneurs. La solution proposée repose sur l'utilisation de l'asservissement visuel. Nous avons principalement abordé cette année la gestion de la faible profondeur de champ disponible, en développant une loi de commande effectuant automatiquement la recherche du plan focal à partir des mesures fournies par un critère de netteté.

### 8.2.2 Convention GIS Sciences de la Cognition : Imagerie cérébrale

**Participants** : Christian Barillot, Pierre Hellier.

no. 1 97 C 731, durée 24 mois

*Partenaires* : projet Epidaure, laboratoire SIM (université Rennes 1), Institut Neurologique de Montréal (Univ. Mc Gill).

Notre tâche cette année s'est surtout concentrée sur la génération de champs denses de déformation à partir de la modélisation statistique obtenue des sillons corticaux (voir paragraphe 6.1.6). Nous avons utilisé une interpolation par splines de plaques minces. Nous avons également collaboré avec l'Université Mc Gill et l'Unité Inserm 155 (M. Duyme, A-M Argenti) pour analyser le codage génétique des sillons centraux pris dans une population de jumeaux monozygotes. Le travail restant consiste à recalcr dans un même espace les données anatomiques IRM (sillons et volumes) et les données fonctionnelles MEG sur la base des deux techniques de recalage développées : le recalage par modélisation statistique des formes des sillons et le recalage global par la mesure de champs de déformation.

### 8.2.3 Actions de recherche coopératives de l'Inria

– Ultrasons 3D

**Participants** : Christian Barillot, Ioannis Pratikakis.

Cette ARC regroupe les projets Epidaure (UR Sophia-Antipolis), Isa (UR Lorraine) et Vista. Elle est coordonnée par C. Barillot. Un site web présente les différents travaux menés à ce jour (<http://www.irisa.fr/vista/Ultrason3D>). Dans ce cadre, des études exploratoires ont pu être entreprises. Pour ce qui nous concerne, nous avons travaillé,

dans le cadre du stage post-doctoral de Ioannis Pratikakis, à la réalisation d'un environnement de vision per-opératoire en neurochirurgie permettant de visualiser en temps réel la déformation de structures anatomiques cérébrales (comme les sillons corticaux par exemple) à l'aide d'une imagerie échographique 3D et de corriger la planification faite préalablement sur l'imagerie anatomique (IRM) voire fonctionnelle (IRMf, MEG). Ce projet se fait également en collaboration avec le laboratoire SIM de l'université de Rennes 1, et le laboratoire de Bio-engineering de l'université de Cambridge (R. Prager et A. Gee). Dans ce contexte, nous interagissons de plus avec différents centres hospitalo-universitaires de l'Ouest (Tours - Drs Pourcelot et Tranquart, Brest - Dr. Bizais, et Rennes - Dr. Darnaux) pour réaliser un environnement logiciel visant à reconstruire et quantifier en trois dimensions un certain nombre de pathologies à partir d'une imagerie échographique (comme l'imagerie vasculaire périphérique pour les plaques d'athérome carotidiennes par exemple).

– Avec, Asservissement visuel en environnement complexe

**Participants** : François Chaumette, Éric Marchand, Youcef Mezouar, Gregory Flandin.

Cette action de recherche coopérative, intitulée "Asservissement Visuel en Environnement Complexe", regroupe l'ensemble des projets de l'Inria travaillant dans le domaine de l'asservissement visuel: Vista, Icare (UR Sophia-Antipolis), Movi et Bip (UR Rhône-Alpes). Elle est coordonnée par F. Chaumette.

Le premier thème scientifique majeur traité dans cette ARC porte sur l'élaboration d'algorithmes de suivi de primitives géométriques dans des séquences d'images complexes. L'objectif est de pouvoir ouvrir le champ applicatif des techniques d'asservissement visuel à des environnements réels (et donc complexes) sur lesquels peu de connaissance a priori est disponible (pas de modèle CAO des objets considérés par exemple). Nous avons ainsi développé un algorithme de suivi d'un trièdre composé de trois plans orthogonaux. Il correspond à une simplification de la méthode présentée au paragraphe 6.2.3. Le suivi s'effectue à la cadence d'environ 5Hz (sur une station Sun ultra 5).

Le second thème porte sur le choix des informations visuelles à utiliser dans la commande en fonction de la tâche à réaliser. La difficulté provient de la très grande richesse des informations fournies dans une séquence d'images qui rend extrêmement délicate la sélection automatique et optimale des seules informations pertinentes. Différents critères de sélection sont étudiés, notamment la robustesse du suivi par traitement d'image et les éventuelles propriétés de découplage des lois de commande associées (voir paragraphe 6.3.1.).

Trois réunions de deux jours se sont déroulées dans les trois sites impliqués et des séjours ponctuels de membres d'un projet dans un autre projet ont eu lieu (en ce qui concerne Vista, séjour d'Éric Marchand à Grenoble en juillet pour le transfert d'un algorithme de suivi).

### 8.2.4 Action incitative du Ministère de la Recherche, projet Visir

**Participants :** Patrick Bouthemy, Patrick Gros, Ronan Fablet.

Le projet VISIR est une action incitative soutenue par le Ministère de la Recherche qui regroupe cinq équipes de recherche (Vista et Sigma-2 de l'Irisa, Movi de Gravir et de l'Inria Rhône-Alpes, IHMPT de l'Irit, et MRIM de Clips-Imag) et l'Ina autour du thème de l'indexation multimédia. Elle est dirigée par P. Gros. Cette action, qui ne regroupe donc que des partenaires impliqués dans le projet AGIR, sera consacrée à développer des techniques plus en amont de celles qui seront implantées dans le projet AGIR qui a une visée plus applicative. Seront particulièrement concernées les techniques permettant de faire collaborer descripteurs et techniques issus de média différents pour améliorer les résultats de recherche par le contenu.

## 8.3 Relations bilatérales internationales

### 8.3.1 Programme France-Israël

**Participants :** Patrick Bouthemy, Ronan Fablet.

Le projet de coopération France-Israël avec le groupe du Prof. S. Peleg au Hebrew University of Jérusalem et le projet Movi (Inria Rhône-Alpes) sur le thème de l'indexation de vidéo intitulé "Video indexing and browsing" s'est terminé début 1999 par la présentation des résultats obtenus lors du séminaire d'évaluation finale du programme France-Israël "Autoroutes de l'Information" à Juan-les-Pins. Ces travaux ont conduit à la réalisation et à la validation de plusieurs fonctionnalités pour la navigation dans des vidéos ou des bases de vidéos à travers des liens entre les contenus (même objet en mouvement ou non, ou même classe d'objets) d'images clés ou d'images mosaïques représentant les plans extraits. Ce projet a donné lieu à la rédaction d'un article en commun [30].

### 8.3.2 Programme France-Tunisie

**Participants :** Patrick Bouthemy, Patrick Gros.

Nous participons à la collaboration franco-tunisienne qui vient d'être mise en place par l'Inria, le Cert et l'Ensi à Tunis. Le projet retenu concerne l'indexation multimédia et regroupe les projets Vista, Movi de l'U.R. Rhône-Alpes, et Imedia de l'U.R. Rocquencourt. Dans le cadre de ce projet pour 1999, nous recevons une doctorante tunisienne, Rim Mezhoud, pour un stage de trois mois (à partir de novembre 99) pour un travail sur les algorithmes d'indexation.

### 8.3.3 Accueil de chercheurs étrangers

T.-M. Peters (Univ. of Western Ontario, Canada), M. Viergever (University Hospital Utrecht), R. Streit et T. Luginbuhl (NUWC, USA), Z. Zhang (Microsoft Research, Seattle), V. Sundareswaran (Rockwell Science Center), F. Miyazaki (Univ. Osaka) ont effectué des visites courtes au sein du projet.

## 9 Diffusion de résultats

### 9.1 Animation de la communauté scientifique

- Participation au GDR-PRC ISIS "Information, Signal et ImageS" :
  - P. Bouthemy est membre du Comité de direction du GDR-PRC ISIS, et co-anime, avec R. André-Obrecht de l'Irit, le GT-10 "Indexation multimedia et vidéo" (organisation par le GT du workshop européen CBMI'99, "Content-based Multimedia Indexing", qui s'est tenu à Toulouse en octobre 99).
  - F. Chaumette co-anime avec M. Devy du LAAS l'opération thématique sur la vision dynamique et active du GT-5 "Vision", (co-organisation d'une réunion de ce pôle sur l'asservissement visuel en mars 1999 à Toulouse).
- Participation au pôle CNRS sur la robotique sous-marine (F. Chaumette).
- P. Bouthemy a animé la 10ème rencontre Irisatech portant sur l'"Indexation de video par le contenu: enjeux et problematiques (de l'exploitation des archives audio-visuelles a la creation de documents interactifs hypervideo)" (autres intervenants: J.-C. Stella, Arts Vidéo Interactive, R. Ronfard, Ina, avril 99).
- P. Bouthemy est éditeur associé de la revue IEEE *Trans. on Image Processing* et est membre du comité de rédaction de la revue TSI.
- P. Gros est membre du comité de rédaction de la revue *Traitement du Signal*.
- F. Chaumette est co-éditeur avec R. Horaud d'un numéro spécial de la revue *Int. Journal of Computer Vision* dédié à l'asservissement visuel, à paraître en 2000.
- Participation aux Comités de Programme:
  - P. Bouthemy : conférences internationales ICCV'99, Vision Interface'99, EMMCVPR'99, CBMI'99, RIAO'2000, ECCV'2000, SSIAT'2000, congrès national RFIA'2000.
  - F. Chaumette: conférences internationales ICRA'2000, CVPR'2000, journées nationales de la Robotique'99.
  - J.-P. Le Cadre: conférences internationales IDC'99, Fusion'2000.
  - P. Bouthemy est membre du Comité éditorial de RFIA'2000.
  - P. Gros a été responsable des comités d'organisation et de programme des journées Orasis'99 réunissant les jeunes chercheurs en vision par ordinateur et perception visuelle, membre du comité d'organisation du workshop CBMI'99.
  - J.-P. Le Cadre a été président du Comité d'organisation du Grets'99 et É. Mémin membre.

### 9.2 Enseignement universitaire

- DEA Stir "Signal, Télécommunications, Image et Radar", université de Rennes 1, option Image (P. Pérez: approches markoviennes en imagerie et analyse multi-images), option Signal (J.-P. Le Cadre: traitement d'antenne et trajectographie-pistage).
- DEA informatique, Ifsic, Rennes 1 (E. Marchand: modèles 3D pour la synthèse et la vision par ordinateur; E. Mémin: analyse du mouvement).

- Diic-INC, Ifsic, Rennes 1 : (E. Mémin, F. Chaumette, G. Flandin : vision 3D, analyse du mouvement, asservissement visuel; E. Mémin : modèles markoviens en analyse d'images; C. Barillot : imagerie médicale).
- Insa Rennes, Génie électrique 5ème année, option image (F. Chaumette, P. Hellier, G. Flandin, Y. Mezouar : vision par ordinateur).
- ENPC Marne-La-Vallée, 2ème année (F. Chaumette : couplage vision-commande en robotique).
- École Polytechnique, 2ème et 3ème années (P. Gros : algorithmique, imagerie tridimensionnelle).
- ENSTA Paris, 3ème année (P. Gros : géométrie projective et vision 3D).
- Supelec Paris (J.-P. Le Cadre : trajectographie et pistage).
- DEA Electronique, option Image, Ireste, université de Nantes (P. Bouthemy : analyse du mouvement).
- DEA GBM universités Tours-Angers-Nantes-Rennes (C. Barillot : imagerie 3D appliquée à la médecine).
- Stages effectués au sein de l'équipe : C. Bonsangue (DESS CCI, Ifsic), N. Courty (DEA Informatique-Insa), T. Corpetti (DEA Stir - Insa GE), O. Dournon (Supelec, en collaboration avec le projet Aladin), Y. Francois (Enssat Lannion), J.-N. Guiavarc'h (DEA Informatique), C. Hue (Dea Mathématiques, filière Statistiques, Rennes 1), S. Mbang (Institut Africain d'Informatique, Libreville, Gabon).
- Direction de thèse de J.-N. Provost (Ecole Navale Brest) par P. Bouthemy et P. Pérez; direction de thèse de R. Donati (Gesma Brest) par J.P. Le Cadre; direction de DRT de S. Marrec (société Thomson Multimédia) par P. Gros.

### 9.3 Participation à des colloques, séminaires, invitations

- C. Barillot a donné une conférence invitée "Active contour modelling for the segmentation of medical images" au Workshop on Segmentation of Medical Images, SPIE Medical Imaging à San Diego en février 1999.
- C. Barillot a donné une conférence invitée "Anatomical and functional image fusion in brain imaging" au 1st Annual Symposium on Virtual Reality In Vascular Medicine au London Health Sciences Centre, University Campus London, Canada en octobre 1999.
- P. Bouthemy a donné une conférence invitée "Spatio-temporal video segmentation and representation for video indexing: new issues in motion analysis", au Colloque MRF'99, Mathematics and Shape Recognition, Luminy, avril 1999.
- P. Bouthemy a donné une conférence invitée "Segmentation spatio-temporelle, suivi, représentation et caractérisation du mouvement pour l'indexation vidéo", au Colloque franco-tunisien TAIMA'99, Tunis, mars 1999.
- F. Chaumette a été invité pour une visite de quelques jours dans le laboratoire de robotique sous-marine du Prof. D. Lane (Heriot-Watt University, Edimbourg).
- P. Gros a été invité au séminaire "Content-Based Image and Video Indexing" organisé par le centre de conférence de Dagstuhl en Allemagne, décembre 1999.

## 10 Bibliographie

### Ouvrages et articles de référence de l'équipe

- [1] B. ESPIAU, F. CHAUMETTE, P. RIVES, « A new approach to visual servoing in robotics », *IEEE Transactions on Robotics and Automation* 8, 6, juin 1992, p. 313–326.
- [2] H. GAUVRET, C. JAUFFRET, J.-P. LE CADRE, « A formulation of multitarget tracking as an incomplete data problem », *IEEE Trans. on Aerospace & Electr. Systems* 33, 4, oct. 1997, p. 1242–1257.
- [3] F. HEITZ, P. PÉREZ, P. BOUTHEMY, « Multiscale minimization of global energy functions in some visual recovery problems », *CVGIP: Image Understanding* 59, 1, 1994, p. 125–134.
- [4] J.-M. LAFERTÉ, P. PÉREZ, F. HEITZ, « Discrete Markov modeling and inference on the quadtree. », *IEEE Transactions on Image Processing*, à paraître, [http://www.irisa.fr/vista/Papers/1999/laferte\\_ip.ps.gz](http://www.irisa.fr/vista/Papers/1999/laferte_ip.ps.gz).
- [5] J.-P. LE CADRE, C. JAUFFRET, « Discrete-time observability and estimability for bearings-only target motion analysis », *IEEE Trans. on Aerospace & Electr. Systems* 33, 1, jan. 1997, p. 178–201.
- [6] G. LE GOUALHER, C. BARILLOT, Y. BIZAIS, « Modeling cortical sulci with active ribbons », *Int. Journal of Pattern Recognition and Artificial Intelligence* 8, 11, 1997, p. 1295–1315.
- [7] E. MARCHAND, F. CHAUMETTE, « An autonomous active vision system for complete and accurate 3D scene reconstruction », *Int. Journal of Computer Vision* 32, 3, August 1999, p. 171–194.
- [8] E. MÉMIN, P. PÉREZ, « Dense estimation and object-based segmentation of the optical flow with robust techniques », *IEEE Trans. on Image Processing* 7, 5, mai 1998, p. 703–719.
- [9] J.-M. ODOBEZ, P. BOUTHEMY, « Robust multiresolution estimation of parametric motion models », *Journal of Visual Communication and Image Representation* 6, 4, décembre 1995, p. 348–365.
- [10] V. SUNDARESWARAN, P. BOUTHEMY, F. CHAUMETTE, « Exploiting image motion for active vision in a visual servoing framework », *Int. Journal of Robotics Research* 15, 6, décembre 1996, p. 629–645.

### Thèses et habilitations à diriger des recherches

- [11] C. BARILLOT, *Fusion de données et imagerie 3D en médecine*, Habilitation à diriger des recherches, Université de Rennes 1, septembre 1999.
- [12] C. PAPIN, *Analyse spatio-temporelle d'images météorologiques satellitaires: détection et suivi de structures nuageuses critiques*, thèse de doctorat, Université de Rennes 1, mention Traitement du signal et télécommunications, décembre 1999.

### Articles et chapitres de livre

- [13] P. BOUTHEMY, M. GELGON, F. GANANSIA, « A unified approach to shot change detection and camera motion characterization », *IEEE Trans. on Circuits and Systems for Video Technology* 9, 7, octobre 1999, p. 1030–1044.
- [14] N. GIORDANA, P. BOUTHEMY, F. CHAUMETTE, F. SPINDLER, J.-C. BORDAS, V. JUST, « 2D model-based tracking of complex shapes for visual servoing tasks », *in: Robust Vision for Vision-based Control of Motion*, M. Vincze et G. Hager (éditeurs), IEEE Press, à paraître.
- [15] J.-M. LAFERTÉ, P. PÉREZ, F. HEITZ, « Discrete Markov modeling and inference on the quadtree. », *IEEE Transactions on Image Processing*, à paraître, [http://www.irisa.fr/vista/Papers/1999/laferte\\_ip.ps.gz](http://www.irisa.fr/vista/Papers/1999/laferte_ip.ps.gz).

- [16] J.-P. LE CADRE, H. GAUVRIT, F. TRARIEUX, « Approximations of the Cramér-Rao bounds for multi-target tracking », *IEE Proc. Radar, Sonar Navig.*, à paraître.
- [17] J.-P. LE CADRE, C. JAUFFRET, « On the convergence of iterative methods for bearings - only tracking », *IEEE Trans. on Aerospace and Electr. Systems (AES)* 35, 3, juillet 1999, p. 801–818.
- [18] J.-P. LE CADRE, S. LAURENT-MICHEL, « Optimizing the receiver maneuvers for bearings-only tracking », *Automatica* 35, 4, avril 1999, p. 591–606.
- [19] G. LE GOUALHER, E. PROCYK, L. COLLINS, R. VENEGOPAL, C. BARILLOT, A. EVANS, « Automated extraction and variability analysis of sulcal neuroanatomy », *IEEE Transactions on Medical Imaging, TMI* 18, 3, mars 1999, p. 206–217.
- [20] E. MALIS, F. CHAUMETTE, S. BOUDET, « 2 1/2 D visual servoing », *IEEE Transactions on Robotics and Automation* 15, 2, avril 1999, p. 238–250, [http://www.irisa.fr/vista/Papers/1999/malis\\_RA.ps.gz](http://www.irisa.fr/vista/Papers/1999/malis_RA.ps.gz).
- [21] E. MARCHAND, F. CHAUMETTE, « Active vision for complete scene reconstruction and exploration », *IEEE Trans. on Pattern Analysis and Machine Intelligence* 21, 1, janvier 1999, p. 65–72, [http://www.irisa.fr/vista/Papers/1999/marchand\\_PAMI.pdf](http://www.irisa.fr/vista/Papers/1999/marchand_PAMI.pdf).
- [22] E. MARCHAND, F. CHAUMETTE, « An autonomous active vision system for complete and accurate 3D scene reconstruction », *Int. Journal of Computer Vision* 32, 3, August 1999, p. 171–194.
- [23] M. MIGNOTTE, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Hybrid genetic optimization and statistical model-based approach for the classification of shadow shapes in sonar imagery », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, à paraître.
- [24] M. MIGNOTTE, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Sonar image segmentation using a hierarchical MRF model », *IEEE Transactions on Image Processing*, à paraître.
- [25] M. MIGNOTTE, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Three-classes Markovian segmentation of high resolution sonar images », *Computer Vision and Image Understanding* 76, 3, décembre 1999.
- [26] E. MÉMIN, T. RISSET, « VLSI design methodology for edge-preserving image reconstruction », *Real Time Imaging*, à paraître.
- [27] G. SOURIS, J.-P. LE CADRE, « Un panorama des méthodes d’optimisation de l’effort de recherche en détection », *Traitement du Signal*, à paraître.
- [28] O. TRÉMOIS, J.-P. LE CADRE, « Optimal observer trajectory in bearings-only tracking for maneuvering sources », *IEE Proceedings Radar, Sonar and Navigation* 146, 2, avril 1999, p. 31–39.

### Communications à des congrès, colloques, etc.

- [29] C. BARILLOT, G. LE GOUALHER, P. HELLIER, B. GIBAUD, « Statistical analysis of brain sulci based on active ribbon modeling », *in: Medical Imaging: Image Processing, 3661*, p. 312–321, San Diego, CA, février 1999.
- [30] P. BOUTHEMY, Y. DUFOURNAUD, R. FABLET, R. MOHR, S. PELEG, A. ZOMET, « Video hyper-link creation for content-based browsing and navigation », *in: Workshop on Content-Based Multimedia Indexing, CBMI’99*, Toulouse, octobre 1999, [http://www.irisa.fr/vista/Papers/1999/bouthemy\\_CBMI99.ps.gz](http://www.irisa.fr/vista/Papers/1999/bouthemy_CBMI99.ps.gz).
- [31] P. BOUTHEMY, C. GARCIA, R. RONFARD, G. TZIRITAS, E. VENEAU, D. ZUGAJ, « Scene segmentation and image feature extraction for video indexing and retrieval », *in: Proc. of the 3rd Int. Conf. on Visual Information Systems, Visual’99, Lecture Notes in Computer Science, 1614*, Springer Verlag, p. 245–252, Amsterdam, juin 1999.

- [32] P. BOUTHEMY, « Segmentation spatio-temporelle, suivi, représentation et caractérisation du mouvement pour l'indexation vidéo », *in: Colloque franco-tunisien TAIMA'99*, Tunis, mars 1999.
- [33] P. BOUTHEMY, « Spatio-temporal video segmentation and representation for video indexing: new issues in motion analysis », *in: Colloque MRF'99: Mathematics and Shape Recognition, Mathématique et reconnaissance de formes*, Luminy, avril 1999.
- [34] A. CHARDIN, P. PÉREZ, « Estimation hiérarchique semi-itérative du mode des marginales a posteriori en imagerie », *in: 17ème Colloque Gretsi sur le Traitement du signal et des images, 1*, p. 207–210, Vannes, septembre 1999, [http://www.irisa.fr/vista/Papers/1999/chardin\\_GRETSI.ps.gz](http://www.irisa.fr/vista/Papers/1999/chardin_GRETSI.ps.gz).
- [35] A. CHARDIN, P. PÉREZ, « Modes of posterior marginals with hierarchical models », *in: IEEE Int. Conf. on Image Processing, ICIP'99*, Kobe, Japon, octobre 1999, [http://www.irisa.fr/vista/Papers/1999/chardin\\_ICIP.ps.gz](http://www.irisa.fr/vista/Papers/1999/chardin_ICIP.ps.gz).
- [36] A. CHARDIN, P. PÉREZ, « Semi-iterative inferences with hierarchical energy-based models for image analysis », *in: Int. Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Computer Science, 1654*, Springer, p. 83–98, York, juillet 1999, [http://www.irisa.fr/vista/Papers/1999/chardin\\_EMMCVPR.ps.gz](http://www.irisa.fr/vista/Papers/1999/chardin_EMMCVPR.ps.gz).
- [37] A. CHARDIN, P. PÉREZ, « Unsupervised image classification with a hierarchical EM algorithm », *in: IEEE Int. Conf. on Computer Vision, ICCV'99, 2*, p. 969–974, Corfou, Grèce, septembre 1999, [http://www.irisa.fr/vista/Papers/1999/chardin\\_ICCV.ps.gz](http://www.irisa.fr/vista/Papers/1999/chardin_ICCV.ps.gz).
- [38] F. CHAUMETTE, « Potential problems of unstability and divergence in image-based and position-based visual servoing », *in: Proc. of the European Control Conference, ECC'99*, Karlsruhe, Allemagne, septembre 1999.
- [39] N. COURTY, E. MARCHAND, « Navigation et contrôle d'une caméra dans un environnement virtuel : une approche référencée image. », *in: AFIG'99, 7èmes journées de l'Association Française d'Informatique Graphique*, Reims, novembre 1999.
- [40] G. CSURKA, P. BOUTHEMY, « Direct identification of moving objects and background from 2D motion models », *in: IEEE Int. Conf. on Computer Vision, ICCV'99, 1*, p. 566–571, Corfou, Grèce, septembre 1999.
- [41] F. DAMBREVILLE, J.-P. LE CADRE, « Distribution optimale d'un effort de recherche multi-étapes pour des contraintes linéaires généralisées », *in: 17ème Colloque Gretsi sur le Traitement du signal et des images, 4*, p. 1005–1008, Vannes, septembre 1999.
- [42] R. FABLET, P. BOUTHEMY, M. GELGON, « Moving object detection in color image sequences using region-level graph labeling », *in: IEEE Int. Conf. on Image Processing, ICIP'99*, Kobe, Japon, octobre 1999, [http://www.irisa.fr/vista/Papers/1999/fablet\\_ICIP.ps.gz](http://www.irisa.fr/vista/Papers/1999/fablet_ICIP.ps.gz).
- [43] R. FABLET, P. BOUTHEMY, M. GELGON, « Étiquetage statistique d'un graphe de régions pour la détection d'objets mobiles dans des séquences d'images couleur », *in: 17ème Colloque Gretsi sur le Traitement du signal et des images, 2*, p. 499–502, Vannes, septembre 1999, [http://www.irisa.fr/vista/Papers/1999/fablet\\_GRETSI.ps.gz](http://www.irisa.fr/vista/Papers/1999/fablet_GRETSI.ps.gz).
- [44] R. FABLET, P. BOUTHEMY, « Extraction d'attributs de mouvement et classification hiérarchique pour l'indexation vidéo par le contenu », *in: ORASIS'99, Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle*, p. 15–24, Aussois, avril 1999, [http://www.irisa.fr/vista/Papers/1999/fablet\\_ORASIS.ps.gz](http://www.irisa.fr/vista/Papers/1999/fablet_ORASIS.ps.gz).
- [45] R. FABLET, P. BOUTHEMY, « Motion-based feature extraction and ascendant hierarchical classification for video indexing and retrieval », *in: Proc. of the 3rd Int. Conf. on Visual Information Systems, VISual99, Lecture Notes in Computer Science, 1614*, Springer Verlag, p. 221–228, Amsterdam, juin 1999, [http://www.irisa.fr/vista/Papers/1999/fablet\\_VISual.ps.gz](http://www.irisa.fr/vista/Papers/1999/fablet_VISual.ps.gz).

- [46] R. FABLET, P. BOUTHEMY, « Spatio-temporal segmentation and general motion characterization for video indexing and retrieval », in : *10th DELOS Workshop on Audio-Visual Digital Libraries*, Santorin, Grèce, juin 1999, [http://www.irisa.fr/vista/Papers/1999/fablet\\_DELOS.ps.gz](http://www.irisa.fr/vista/Papers/1999/fablet_DELOS.ps.gz).
- [47] M. GELGON, P. BOUTHEMY, T. DUBOIS, « A region tracking method with failure detection for an interactive video indexing environment », in : *Proc. of the 3rd Int. Conf. on Visual Information Systems, Visual'99, Lecture Notes in Computer Science, 1614*, Springer Verlag, p. 261–268, Amsterdam, juin 1999.
- [48] M. GELGON, P. BOUTHEMY, J.-P. LE CADRE, « Associating and estimating trajectories of multiple moving regions with a probabilistic multi-hypothesis tracking approach », in : *Physics in Signal and Image Processing Symposium, PSIP'99*, Paris, janvier 1999.
- [49] P. HELLIER, C. BARILLOT, E. MÉMIN, P. PÉREZ, « Medical image registration with robust multigrid techniques », in : *Proc. of the 2nd Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, MICCAI*, C. Taylor, A. Colchester (éditeurs), *Lecture Notes in Computer Science, 1679*, Springer Verlag, p. 680–687, septembre 1999.
- [50] P. HELLIER, C. BARILLOT, E. MÉMIN, P. PÉREZ, « Recalage d'images cérébrales par estimation robuste contextuelle du flot optique », in : *ORASIS'99, Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle*, p. 85–91, Aussois, avril 1999.
- [51] P. HELLIER, C. BARILLOT, E. MÉMIN, P. PÉREZ, « Recalage d'images cérébrales par estimation robuste contextuelle du flot optique », in : *17ème Colloque Gretsi sur le Traitement du signal et des images, 4*, p. 1077–1080, Vannes, septembre 1999.
- [52] P. JANNIN, D. SCHWARTZ, C. GROVA, C. BARILLOT, B. GIBAUD, « Fusion of multimodal functional neuro-imaging (MEG, fMRI, SPECT): definition of different matching paradigms », in : *Computer Assisted Radiology and Surgery*, H. Lemke, M. Vannier, K. Inamura, A. Farman (éditeurs), Elsevier, p. 238–243, Paris, juin 1999.
- [53] C. JAUFFRET, J.-P. LE CADRE, « Observabilité et trajectographie passive d'un objet en mouvement uniformément accéléré en dimension 3 », in : *17ème Colloque Gretsi sur le Traitement du signal et des images, 2*, p. 673–676, Vannes, septembre 1999.
- [54] J.-P. LE CADRE, « Searching Tracks », in : *Proc. of Fusion'99*, p. 478–486, Palo Alto CA, USA, juillet 1999.
- [55] I. LEIBOWICZ, P. NICOLAS, « Mixing kinematics and identification data for track-to-track association », in : *SPIE Conf. on Sensor Fusion: Architectures, Algorithms and Applications III*, Orlando, avril 1999.
- [56] E. MARCHAND, P. BOUTHEMY, F. CHAUMETTE, V. MOREAU, « Robust real-time visual tracking using a 2D-3D model-based approach », in : *IEEE Int. Conf. on Computer Vision, ICCV'99, 1*, p. 262–268, Corfou, Grèce, septembre 1999, [http://www.irisa.fr/vista/Papers/1999/marchand\\_ICCV.ps.gz](http://www.irisa.fr/vista/Papers/1999/marchand_ICCV.ps.gz).
- [57] E. MARCHAND, P. BOUTHEMY, F. CHAUMETTE, V. MOREAU, « Suivi robuste d'objets en temps-réel : une approche hybride 2D-3D », in : *17ème Colloque Gretsi sur le Traitement du signal et des images, 1*, p. 115–118, Vannes, septembre 1999, [http://www.irisa.fr/vista/Papers/1999/marchand\\_Gretsi.ps.gz](http://www.irisa.fr/vista/Papers/1999/marchand_Gretsi.ps.gz).
- [58] E. MARCHAND, P. BOUTHEMY, F. CHAUMETTE, V. MOREAU, « Visual tracking by coupling 2D motion and 3D pose estimation. », in : *IEEE Int Conf. on Image Processing, ICIP'99*, Kobe, Japon, octobre 1999, [http://www.irisa.fr/vista/Papers/1999/marchand\\_ICIP.ps.gz](http://www.irisa.fr/vista/Papers/1999/marchand_ICIP.ps.gz).
- [59] E. MARCHAND, « ViSP: A Software Environment for Eye-in-Hand Visual Servoing », in : *IEEE Int. Conf. on Robotics and Automation, ICRA'99, 4*, p. 3224–3229, Detroit, Michigan, mai 1999, [http://www.irisa.fr/vista/Papers/1999/marchand\\_ICRA.ps.gz](http://www.irisa.fr/vista/Papers/1999/marchand_ICRA.ps.gz).
- [60] M. MIGNOTTE, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Bayesian inference and optimization

- strategies for some detection and classification problems in sonar imagery », *in* : *SPIE Conf. Non Linear Image Processing*, 3646, p. 14–27, San Jose, Californie, janvier 1999.
- [61] M. MIGNOTTE, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Markov random field models and fuzzy formalism-based data modeling for the sea-bed classification in sonar imagery », *in* : *SPIE Conf. Mathematical Modeling, Bayesian Estimation and Inverse Problems*, 3816, p. 229–240, Denver, Colorado, juillet 1999.
- [62] E. MÉMIN, P. PÉREZ, « Dense/parametric estimation of fluid flows », *in* : *IEEE Int. Conf. on Image Processing, IICIP'99*, Kobe, Japon, octobre 1999.
- [63] E. MÉMIN, P. PÉREZ, « Estimation conjointe de champs de vitesse denses et paramétriques pour des écoulements fluides », *in* : *17ème Colloque Gretsi sur le Traitement du signal et des images*, 2, p. 247–250, Vannes, septembre 1999.
- [64] E. MÉMIN, P. PÉREZ, « Fluid motion recovery by coupling dense and parametric fields », *in* : *IEEE Int. Conf. on Computer Vision, ICCV'99*, 1, p. 620–625, Corfou, Grèce, septembre 1999, [http://www.irisa.fr/vista/Papers/1999/memin\\_ICCV.ps.gz](http://www.irisa.fr/vista/Papers/1999/memin_ICCV.ps.gz).
- [65] L. OISEL, L. MORIN, E. MÉMIN, C. LABIT, « Reconstruction 3D de scènes complexes par maillage de cartes de disparité », *in* : *17ème Colloque Gretsi sur le Traitement du signal et des images*, 2, p. 391–394, Vannes, septembre 1999.
- [66] C. PAPIN, P. BOUTHEMY, E. MÉMIN, G. ROCHARD, « Détection, suivi et caractérisation des nuages convectifs à partir d'images satellitaires », *in* : *ORASIS'99, Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle*, p. 51–60, Aussois, avril 1999.
- [67] C. PAPIN, P. BOUTHEMY, E. MÉMIN, G. ROCHARD, « Tracking and characterization of convective clouds from satellite images », *in* : *1999 Eumetsat Meteorological Satellite Data User's Conference*, Copenhague, Danemark, septembre 1999.
- [68] J.-N. PROVOST, C. COLLET, P. PÉREZ, P. BOUTHEMY, « A hierarchical unsupervised multispectral model to segment SPOT images for ocean cartography », *in* : *IEEE Int. Conf. on Image Processing, IICIP'99*, Kobe, Japon, octobre 1999.
- [69] J.-N. PROVOST, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Segmentation d'images radiométriques multispectrales pour l'obtention de cartes bathymétriques », *in* : *ORASIS'99, Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle*, p. 15–24, Aussois, avril 1999.
- [70] J.-N. PROVOST, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Segmentation non-supervisée d'images multispectrales SPOT. Application à la cartographie maritime », *in* : *17ème Colloque Gretsi sur le Traitement du signal et des images*, 1, p. 27–30, Vannes, septembre 1999.
- [71] J.-N. PROVOST, C. COLLET, P. PÉREZ, P. BOUTHEMY, « Unsupervised multispectral segmentation of Spot images applied to nautical cartography », *in* : *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, NSIP'99*, Antalya, Turquie, juin 1999.
- [72] D. ZUGAJ, P. BOUTHEMY, « Détection de volets pour l'indexation de vidéo par le contenu », *in* : *17ème Colloque Gretsi sur le Traitement du signal et des images*, 4, p. 1033–1036, Vannes, septembre 1999.
- [73] D. ZUGAJ, P. BOUTHEMY, « Wipe detection in the temporal segmentation of video », *in* : *European Workshop on Content-Based Multimedia Indexing, CBMI'99*, Toulouse, octobre 1999.

## Rapports de recherche et publications internes

- [74] F. DAMBREVILLE, J.-P. LE CADRE, « Detection of a Markovian target with optimization and distribution of the search efforts under generalized linear constraints », *rapport de recherche n° 1278*, Irisa, novembre 1999.

- 
- [75] P. HELLIER, C. BARILLOT, E. MÉMIN, P. PÉREZ, « Estimation robuste 3D d'un champ de déformation pour le recalage non-linéaire inter-sujet d'images cérébrales », *Publication interne n° 1281*, Irisa, <http://www.irisa.fr/EXTERNE/bibli/pi/1281/1281.html>, 1999.
- [76] E. MARCHAND, P. BOUTHEMY, F. CHAUMETTE, « Étude sur le suivi 2D de formes complexes en vue d'une utilisation en commande référencée vision », *rapport de recherche*, contrat EdF - Inria Rennes, Irisa 1.97.C234, septembre 1999.
- [77] E. MÉMIN, T. RISSET, « Hardware driven considerations for energy based applications. », *rapport de recherche n° 1220*, Irisa, avril 1999, <ftp://ftp.irisa.fr/techreports/1999/PI-1220.ps.gz>.