



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Projet VISTA

Vision Spatio-Temporelle et Active

Rennes

THÈME 3B

*R*apport
d'Activité

2002

Table des matières

1	Composition de l'équipe	4
2	Présentation et objectifs généraux	5
3	Fondements scientifiques	7
3.1	Modèles markoviens hiérarchiques et estimation bayésienne	7
3.2	Commande référencée vision	9
3.3	Extraction de pistes et trajectographie	11
4	Domaines d'applications	13
5	Logiciels	14
5.1	Logiciel RMR d'estimation du mouvement dominant	14
5.2	Logiciel D-Change de détection de mouvement	14
5.3	Logiciel Dense-Motion de mesure de champs de vitesses 2D	15
5.4	Logiciel MD-Shots de découpage d'une vidéo en plans	15
5.5	Logiciel ROMEO de recalage non linéaire en imagerie cérébrale	16
5.6	Vistal : Librairies pour le traitement et l'analyse d'images volumiques spatio-temporelles	16
5.7	Plate-forme VISP d'asservissement visuel	17
5.8	Marker : Logiciel de calcul de pose pour la réalité augmentée	17
5.9	Actions de développement: Plates-formes de vision robotique	18
5.10	Actions de développement: Véhicule Cycab	18
6	Résultats nouveaux	19
6.1	Modèles statistiques, mesure du mouvement et des déformations, traitement vidéo	19
6.1.1	Modèles pour l'analyse du mouvement fluide	19
6.1.2	Modélisation générale du mouvement pour le traitement vidéo	20
6.1.3	Modèles probabilistes de formes en imagerie médicale	21
6.1.4	Détection sûre du mouvement	23
6.1.5	Résumé vidéo par segmentation et classification du mouvement	23
6.1.6	Formation et filtrage des images échographiques 3D	24
6.2	Suivi temporel, recalage, reconstruction	25
6.2.1	Suivi multi-objets en imagerie par filtrage particulière	25
6.2.2	Suivi de points à partir de séquences d'images	26
6.2.3	Trajectographie pour cibles manœuvrantes	27
6.2.4	Détection de structures spatio-temporelles	28
6.2.5	Recalage d'images cérébrales	28
6.2.6	Images ultrasonores 3D, calibration et reconstruction	29
6.3	Asservissement visuel, perception active et exploration	30
6.3.1	Utilisation des moments en asservissement visuel	30
6.3.2	Asservissement visuel virtuel pour la réalité augmentée	31

6.3.3	Asservissement visuel et réalité virtuelle	32
6.3.4	Couplage perception-action par indexation d'images	32
6.3.5	Planification de trajectoire pour l'optimisation de la navigation	33
6.3.6	Théorie de l'exploration	34
7	Contrats industriels (nationaux, européens et internationaux)	35
7.1	Convention Cemagref : Évaluation de la qualité de pièces de viande par vision active	35
7.2	Convention LRBA-DGA : projet Posit	35
7.3	Convention Cifre Thales Systèmes Aéroportés : Traitement d'évaluation de la situation tactique	36
7.4	Convention DGA : Suivi d'amers temps-réel	36
7.5	Convention DGA - ENS Cachan : Détection de trajectoires	36
7.6	Convention Cifre Ina : Reconnaissance de formes spatio-temporelles pour l'indexation vidéo	37
7.7	Actions financées par la Commission Européenne	37
7.7.1	Projet IST Carsense	37
7.7.2	Projet IST Lava	38
7.8	Actions financées par le RNTL	38
7.8.1	Projet Domus Videum	38
7.9	Actions soutenues par le programme Priamm	39
7.9.1	Projet Mediaworks	39
8	Actions régionales, nationales et internationales	40
8.1	Actions régionales	40
8.1.1	Contrat Région Bretagne : Création automatique de résumés de vidéos pour le multimédia domestique et professionnel	40
8.1.2	Contrat Région Bretagne : Méthodes de traitement et de reconstruction d'images échographiques tri-dimensionnelles	40
8.1.3	Contrat CPER Cemagref-Inria AEROBIO	40
8.1.4	Télé-acquisition robotisée d'échographies 3D	40
8.2	Actions nationales	41
8.2.1	ACI Ministère de la Recherche : projet NeuroBase	41
8.2.2	Actions de recherche coopératives de l'Inria	42
8.2.3	Actions MathStic du CNRS	42
8.2.4	AS du département STIC du CNRS	43
8.2.5	Programme Robea du CNRS	43
8.3	Réseaux et groupes de travail internationaux	44
8.3.1	Groupe de travail Ercim	44
8.3.2	Groupe de travail Delos-NSF	44
8.4	Relations bilatérales internationales	44
8.4.1	Programme Procope France-Allemagne	44
8.4.2	Programme Van Gogh France-Pays-Bas	45
8.4.3	Accueil de chercheurs étrangers	45

8.4.4	Expertise de projets	45
9	Diffusion de résultats	46
9.1	Animation de la communauté scientifique	46
9.2	Enseignement universitaire	47
9.3	Participation à des colloques, séminaires, invitations	47
10	Bibliographie	48

1 Composition de l'équipe

Responsable scientifique

Patrick Bouthemy [DR Inria]

Assistante de projet

Édith Blin [TR Inria, avec équipe/projet Imadoc et Temics jusqu'au 30/09/2002]

Stéphanie Lemaile [TR Inria, avec projet Temics, à compter du 01/10/2002]

Personnel Inria

Frédéric Cao [CR]

François Chaumette [DR]

Pierre Hellier [CR]

Éric Marchand [CR]

Fabien Spindler [IR (Atelier)]

Personnel CNRS

Christian Barillot [DR]

Jean-Pierre Le Cadre [DR]

Personnel Université

Étienne Mémin [maître de conférences Ifsic, Université de Rennes 1, en délégation CNRS]

Personnel non permanent

Bruno Cernuschi-Frias [professeur invité Inria, Université de Buenos-Aires, jusqu'au 31/01/2002]

Nicolas Ducoin [ingénieur associé Inria]

Chafik Kermad [ingénieur expert Inria]

Sébastien Paris [ingénieur expert Inria, jusqu'au 15/07/2002]

Chercheur post-doctorant et ATER

Nathalie Peyrard [post-doc Inria]

Thomas Corpetti [ATER Ifsic, à compter du 01/09/2002]

Chercheurs doctorants

Elise Arnaud [allocataire MENRT]

Thomas Bréhard [boursier DGA, à compter du 01/10/2002]

Cybèle Ciofolo [boursière BDI CNRS-Région, à compter du 01/10/2002]

Andrew Comport [boursier Inria]

Isabelle Corouge [boursière Inria]

Thomas Corpetti [allocataire MENRT, jusqu'au 31/08/2002]

Carine Hue [allocataire MENRT]

Arnaud Ogier [boursier Inria, à compter du 01/11/2002]

Gwénaëlle Piriou [boursière Inria]

Anthony Remazeilles [allocataire MENRT]

François Rousseau [allocataire MENRT]

Omar Tahri [boursier Inria-Région]

Emmanuel Veneau [boursier Cifre, Ina, jusqu'au 28/02/2002]

Thomas Veit [boursier Inria-Région, à compter du 01/10/2002]

2 Présentation et objectifs généraux

Mots clés : imagerie spatio-temporelle, vision dynamique, vision active, modélisation statistique, modèles markoviens, estimation bayésienne, estimation robuste, filtrage particulière, segmentation, mesure du mouvement, apprentissage, reconnaissance, suivi temporel, trajectographie, asservissement visuel, stratégie de perception, déformation, mouvement fluide, sonar, imagerie médicale, vision robotique, indexation vidéo, imagerie météorologique, réalité augmentée, traitement vidéo, imagerie physique, imagerie cérébrale, échographie 3D.

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 scènes et de phénomènes physiques de nature dynamique, avec des objectifs de détection, de mesure, d'apprentissage et de reconnaissance d'informations spatio-temporelles ; ii) le couplage perception-commande dans des systèmes automatisés (principalement robotiques), pour des tâches de positionnement, de préhension, de poursuite, de décision ou d'exploration. Nous privilégions une approche statistique (markovienne, bayésienne, estimation robuste, filtrage particulière) de la première catégorie de problèmes, et une approche par asservissement visuel (lois de commande à partir d'informations visuelles) ou par fusion pour la seconde. Les questions d'analyse du mouvement, de déformation, de recalage, de suivi temporel et de trajectographie sont particulièrement étudiées. Une large gamme de modélisations et de lois d'interaction ont été définies pour l'asservissement visuel, et des aspects de vision active sont également abordés. Différents domaines et secteurs d'applications irriguent nos études : imagerie de la physique (phénomènes environnementaux naturels en imagerie météorologique; phénomènes physiques en mécanique des fluides expérimentale liés à la conception industrielle en aérodynamique, aéraulique ou hydrodynamique), imagerie médicale (imagerie cérébrale, chirurgie guidée par l'image, échographie 3D), traitement et indexation vidéo, vision robotique (robotique de manipulation, robotique sous-marine) et réalité augmentée (post-production), systèmes de détection et de navigation (transports, défense).*

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, de mesure, d'apprentissage et de reconnaissance d'informations spatio-temporelles ;
- le couplage perception-commande dans des systèmes automatisés (principalement robotiques), pour des tâches de positionnement, de préhension, de poursuite, de décision ou 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, estimation robuste, filtrage particulaire) de la première catégorie de problèmes, et une approche par asservissement visuel (lois de commande à partir d'informations visuelles) ou par fusion de la seconde. Les questions d'analyse du mouvement, de déformation, de recalage, de suivi temporel et de trajectographie sont particulièrement étudiées. Une large gamme de modélisations et de lois d'interaction (en boucle fermée) ont été définies pour l'asservissement visuel, et des aspects de vision active (systèmes de perception et de détection) sont également abordés pouvant intégrer plusieurs types de capteur.

Plusieurs secteurs et domaines d'applications motivent et irriguent nos études : - *i*) imagerie de la physique (phénomènes environnementaux naturels en imagerie météorologique; phénomènes physiques en mécanique des fluides expérimentale liés à la conception industrielle en aérodynamique, aéraulique ou hydrodynamique), - *ii*) imagerie médicale (imagerie cérébrale, chirurgie guidée par l'image, échographie 3D), - *iii*) traitement et indexation vidéo (filtrage, détection, reconnaissance, structuration), - *iv*) vision robotique (robotique de manipulation, robotique sous-marine), réalité augmentée (post-production), systèmes de détection et de navigation (transports, défense).

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

- **Modèles statistiques pour la détection et la mesure du mouvement et des déformations.** 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 d'analyse d'image. 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 notamment dans ce contexte à des modèles markoviens hiérarchiques, aux estimateurs robustes, à l'introduction de modèles physiques appropriés ainsi qu'à des questions d'apprentissage et de reconnaissance de structures spatio-temporelles ou d'événements temporels.

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. 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. Nous nous intéressons particulièrement à l'analyse du mouvement déformable, notamment le mouvement fluide.

- **Recalage, suivi et trajectographie.** Nous étendons ces approches au problème du recalage non linéaire entre images volumiques en imagerie cérébrale (IRM, IRMf, ainsi qu'en échographie 3D, ce dernier domaine incluant également des questions de calibrage et de reconstruction. Sur les aspects de suivi temporel et de trajectographie, nous considérons des problèmes liés à l'estimation de systèmes dynamiques partiellement observés, ainsi qu'à la formalisation du suivi d'objets multiples par des méthodes de type PMHT (Pro-

babilistic Multiple Hypotheses Tracking), pour résoudre le problème d'association entre observations et modèles. Nous nous intéressons également au filtrage particulière pour appréhender des non-linéarités de l'équation de mesure et/ou du modèle d'état.

– Asservissement visuel et perception active.

Nous avons défini et validé un corpus de méthodes d'asservissement visuel utilisant des informations visuelles de nature aussi bien géométrique que cinématique, et suivant des schémas purement 2D ou combinant des aspects 2D et 3D. Des questions de planification de trajectoires associée à un asservissement visuel sont également explorées. En perception active, nous nous intéressons au contrôle de la position, de la vitesse, ou de la trajectoire de capteurs (ou de leurs porteurs) pour améliorer les performances des algorithmes de vision ou de détection. Deux types de stratégies sont considérés, voire 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. Ces dernières peuvent éventuellement combiner plusieurs capteurs et poser des problèmes de fusion. Nous abordons également 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 secteurs d'applications mentionnés, nous avons eu récemment ou nous avons actuellement des conventions de recherche ou de transfert avec Météo-France, Thales-Optronique, l'Ina, Thomson-Multimedia, l'Ifremer, le Cemagref, l'Onera, des projets multi-partenaires nationaux (programmes RNRT, RNTL, RIAM) et européens (projet IST Carsense, contrat Eumetsat, projet IST Lava).

3 Fondements scientifiques

3.1 Modèles markoviens hiérarchiques et estimation bayésienne

Mots clés : champ de Markov, observation partielle, estimation bayésienne, estimation de paramètre, 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.*

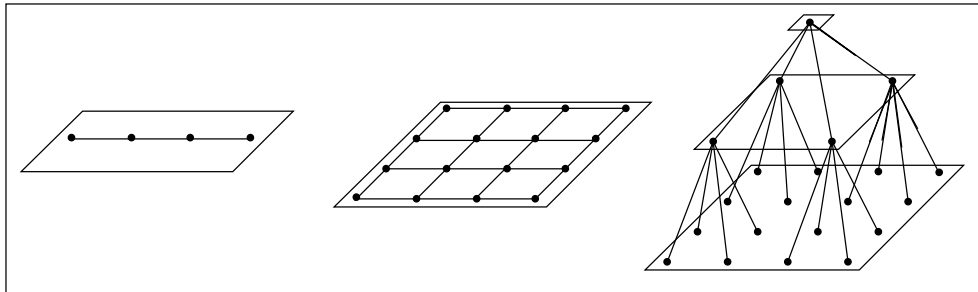
Le type d'approche statistique de l'analyse d'image que nous privilégions nous conduit à mettre en place des *modèles markoviens partiellement observés*, associés à des distributions $P(z|\theta)$, où θ est un jeu de paramètres. L'ensemble $z = \{z_s, s \in S\}$ se décompose en variables

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

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 :

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

où \mathcal{C} est un ensemble de "petites" parties de S ($|c| = 1$ 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 ^[Whi90] : 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** : θ 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) ^[GG84].
- **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

[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.

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.)

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 [4].

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 vi-

[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.

suelles 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 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 [2] :

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

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 :

- $\hat{L} = L_{(s(t), \hat{z}(t))}$. 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

-
- [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.

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 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, hypothèses multiples, 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]:

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

[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.

où Φ 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: $\hat{\theta}_k^i = \theta_k^i + w_k$, où $\theta_k^i = \arctan(r_x^i(k)/r_y^i(k))$ et i est un 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), et lorsque i 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 :

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

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

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

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, il doit être pris en compte des situations où plusieurs sources sont présentes. 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 Hypothesis 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, environnement, imagerie de la physique, mécanique des fluides expérimentale, imagerie médicale, imagerie cérébrale, chirurgie guidée par l'image, échographie 3D, santé, sonar, indexation vidéo, multimédia, robotique manufacturière, robotique sous-marine, réalité augmentée, transports.

Panorama : *Nous nous intéressons à plusieurs grands secteurs d'applications. Le premier concerne ce que l'on peut appeler la "métrologie" du mouvement et des déformations sur des phénomènes physiques naturels ou expérimentaux. Nous explorons dans ce cadre les domaines de la visualisation expérimentale en mécanique de fluides (écoulements turbulents en aéronautique, avec le Cemagref), de l'imagerie satellitale météorologique (images Meteosat, collaborations récentes avec Météo-France, le LMD et Eumetsat), de l'imagerie médicale en collaboration avec le CHU de Rennes (imagerie cérébrale, IRM, IRMf, échographie 3D, chirurgie guidée par l'image). Le second secteur a trait à la robotique, aussi bien robotique manufacturière que robotique d'intervention, et aux systèmes de détection. Les domaines d'applications concernés ont été dans le passé l'énergie nucléaire (avec EDF), et récemment ou sont actuellement le sous-marin (avec*

[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.

l'Ifremer), l'agro-alimentaire (avec le Cemagref), les transports (projet européen IST Carsense, avec notamment le Livic). Nous abordons également le domaine de la réalité augmentée (calcul de pose, post-production) avec la société Total Immersion. Enfin, le troisième secteur d'applications relève du traitement et de l'indexation vidéo. Nous nous intéressons aux aspects de structuration temporelle de vidéos et d'indexation par leur contenu spatio-temporel, avec différents objectifs comme la consultation d'archives audio-visuelles, la création de résumés de documents vidéos (avec des collaborations avec l'Ina, Thomson-Multimedia), ou la reconnaissance d'événements temporels. Nous avons participé à un projet européen Esprit sur le sujet (projet Divan), à un projet Agir dans le cadre du RNRT; actuellement, nous sommes partenaires des projets Mediaworks (avec le Limsi et TF1 notamment) dans le programme Priamm, au projet Domus Videum dans le programme RNTL (avec TMM notamment), au projet Feria récemment labellisé dans le programme RIAM, ainsi qu'au projet européen IST Lava sur des questions d'apprentissage et de vision par ordinateur (avec notamment XRCE et le projet Movi de l'Inria Rhône-Alpes.

5 Logiciels

5.1 Logiciel RMR d'estimation du mouvement dominant

Correspondant: Fabien Spindler

Un logiciel, s'intitulant RMR pour "Robust Multi-Resolution", 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, [10]. Ce logiciel a été optimisé, et déposé à l'APP. Le logiciel RMR 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 comme module d'une application plus conséquente. Une interface réalisée sous TCL-TK permet de le piloter de façon conviviale. De plus, une version optimisée, déposée à l'APP, en a également été développée et est dénommée Motion-2D.

5.2 Logiciel D-Change de détection de mouvement

Correspondant: Fabien Spindler

Le logiciel D-change est un logiciel de détection des zones mobiles dans une séquence d'images acquise par une caméra fixe. Il comprend deux versions : l'une reposant sur une modélisation

markovienne des zones à détecter, l'autre faisant directement l'hypothèse d'un modèle rectangulaire englobant chaque zone à détecter. Il utilise à la fois les informations de différences temporelles entre deux images successives de la séquence et celles entre l'image courante traitée et une image de référence de la scène observée vide d'éléments mobiles (image de référence construite et mise à jour en ligne). L'algorithme fournit le masque des mobiles détectés (régions ou rectangles) ainsi qu'un numéro de suivi qui leur est attribué pour l'intégralité de leur passage dans le champ de la caméra. Les traitements effectués ont été optimisés afin d'être peu coûteux en temps de calcul. Le logiciel D-Change a été déposé à l'APP et correspond à la méthode développée dans la thèse de Y. Ricquebourg (soutenue en 1997) encadrée par P. Bouthemy et F. Heitz, et a été remanié par F. Spindler.

5.3 Logiciel Dense-Motion de mesure de champs de vitesses 2D

Correspondant : Étienne Mémin

Le logiciel Dense-Motion, déposé à l'APP, permet d'estimer un champ dense de vitesses 2D entre deux images consécutives d'une séquence. Il s'appuie sur une formulation énergétique incrémentale robuste. La fonction d'énergie associée est composée d'un terme d'adéquation aux données et d'un terme de régularisation préservant les discontinuités du champ des vitesses à estimer. Ce logiciel permet à l'utilisateur de choisir deux modèles de données différents : un terme classique s'appuyant sur l'équation de contrainte du mouvement apparent ou un terme découlant d'une intégration de l'équation de continuité de la mécanique des fluides. Deux modèles de régularisation robustes peuvent également être choisis: un modèle du premier ordre ou un modèle de régularisation du type "div-curl" du second ordre. Toutes les combinaisons entre ces différentes possibilités sont autorisées. Ce logiciel correspond aux développements méthodologiques réalisés par E. Mémin, P. Pérez et T. Corpetti [9, 18, 32].

5.4 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^[BGG99] 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 (cf paragraphe 5.1). On s'intéresse en fait à l'évolution temporelle de la taille normalisée de ce

[BGG99] 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.

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 de mise à disposition auprès de différents partenaires.

5.5 Logiciel ROMEO de recalage non linéaire en imagerie cérébrale

Correspondant : Pierre Hellier

Le logiciel ROMEO, pour "ROBust Multilevel Elastic registration based on Optical flow", a été écrit en langage C++, et utilise la librairie `vistaL` (décrite au paragraphe 5.6). Il a été déposé à l'APP. Le but du logiciel ROMEO est le recalage non-rigide, ou la mise en correspondance, d'images tridimensionnelles médicales de même modalité. Ce type de recalage peut être utilisé dans les contextes suivants : *i*) les images proviennent du même sujet et ont été acquises à des instants différents. Des variations morphologiques peuvent apparaître, comme le bougé de structures internes, l'absence de certains organes après chirurgie, etc... Un recalage non-rigide sert alors à détecter et à quantifier les déformations rencontrées; *ii*) les images proviennent de sujets différents. La comparaison inter-individuelle nécessite un recalage non-rigide pour tenir compte des variations morphologiques. L'estimation d'une telle transformation permet de reporter, entre chaque sujet, des informations acquises et extraites, numériques et symboliques, anatomiques et fonctionnelles, afin par exemple de construire un atlas probabiliste morphologique. Une description de la méthode de recalage peut être trouvée dans [5].

5.6 Vistal : Bibliothèques pour le traitement et l'analyse d'images volumiques spatio-temporelles

Correspondants : Pierre Hellier, Christian Barillot

Dans le contexte de l'analyse d'images 3D et 3D+t dans le domaine de l'imagerie médicale, nous travaillons à la réalisation de bibliothèques de traitements d'images volumiques spatio-temporelles utilisables dans des problèmes 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 plateforme 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 opérations floues sur des images 3D, manipulation de voisinage dans l'image, filtres linéaires (RIF et RII) et non-linéaires, filtres gaussiens 1D et leurs 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. Ce logiciel a été déposé à l'APP (co-auteurs : C. Barillot, F. Boisse et P. Hellier).

5.7 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 du 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 ensuite :

- 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.

5.8 Marker : Logiciel de calcul de pose pour la réalité augmentée

Correspondant: Éric Marchand

Le logiciel Marker implémente un algorithme de calcul de pose et de calibration des caméras à partir de marqueurs de type point. Il s'appuie sur une approche d'asservissement visuel virtuel. La méthode consiste à considérer la pose et la calibration comme le problème dual de l'asservissement visuel. Elle prend l'image de l'objet comme le motif à atteindre, le motif courant étant

obtenu par la reprojexion du modèle 3D de l'objet sur le plan image d'une caméra virtuelle. Cette méthode présente de nombreux avantages : précision identique aux méthodes de minimisation non-linéaire classiques, plus grande simplicité, efficacité, et possibilité de considérer différentes informations visuelles.

Marker a fait l'objet d'un dépôt à l'APP en septembre 2002. Une licence de ce logiciel a été cédée à la société Total Immersion.

5.9 Actions de développement: Plates-formes de vision robotique

Participant : Fabien Spindler.

Le projet Vista dispose de plusieurs plates-formes expérimentales de vision robotique sur lesquelles sont validées nos recherches en vision robotique (détection et suivi d'objets en mouvement, asservissement visuel, perception active et exploration). Plus précisément, il s'agit d'un robot manufacturier Afma Robots à six degrés de liberté et d'un robot cylindrique à quatre degrés de liberté. Ces deux robots sont chacun équipés d'une caméra CCD embarquée sur l'effecteur et de cartes de numérisation d'images. Chacun des robots est piloté depuis un PC qui communique avec les robots au travers d'un coupleur de bus. Ces matériels nécessitent des opérations de maintenance matérielle, mais également logicielle afin de les faire évoluer. Des activités de formation et d'assistance des utilisateurs, de montage et de présentation de nombreuses démonstrations font également partie du travail quotidien.

5.10 Actions de développement: Véhicule Cycab

Participants : Fabien Spindler, Nicolas Ducoin.

Nous avons tout d'abord œuvré à la mise en route du véhicule. Nous avons intégré toutes les fonctionnalités nécessaires à la commande par logiciel du Cycab. Dans un premier temps, nous avons modifié et affiné la commande en vitesse fournie par Robosoft. Puis, à partir du logiciel Syndex développé par le projet Sosso de l'Inria Rocquencourt, nous avons mis en place une première couche logicielle de contrôle du Cycab à partir d'API écrites en langage C/C++. Ces travaux ont été validés par la mise en place d'une démonstration permettant au pilote de suivre en temps réel, sur l'écran embarqué, la trajectoire du Cycab projetée sur la carte du site. En parallèle, nous avons également mis en place la couche logicielle permettant une téléopération du Cycab depuis un poste fixe grâce à la liaison éthernet sans fil dont il est équipé. D'autre part, nous avons travaillé à la modélisation de l'ensemble "Cycab et sa tourelle site-azimut embarquée", afin de mettre en place une commande référencée vision du véhicule. Nous allons maintenant procéder à sa validation et à la finalisation d'une démonstration d'asservissement visuel. Plus précisément, il s'agira de positionner le Cycab par rapport à un amer visualisé par la caméra embarquée. Enfin, nous allons également installer une tourelle site-azimut déportée qui permettra d'obtenir une vision globale de la zone d'évolution du véhicule.

6 Résultats nouveaux

6.1 Modèles statistiques, mesure du mouvement et des déformations, traitement vidéo

Contexte: *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 de vitesses 2D dans le cas général impose une gestion simultanée des possibles discontinuités ou des structures sous-jacentes (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. L'analyse de séquences d'images d'écoulements fluides est un problème central dans de très nombreux domaines tels que les sciences environnementales (météorologie, océanographie, climatologie), la visualisation expérimentale en mécanique des fluides (aéro- ou hydro-dynamique) ou encore l'imagerie médicale (étude d'écoulements de bio-fluides). Il peut s'agir d'estimer à partir de ces images des champs de vitesses instantanées de l'écoulement ou de détecter (et suivre dans le temps) des structures cohérentes de l'écoulement (structures tourbillonnaires, vortex, puits ou sources). Un autre champ d'investigation porte sur les questions de modélisation non paramétrique et de reconnaissance du mouvement avec des applications notamment en indexation vidéo par le contenu et en création de résumés de vidéos.*

6.1.1 Modèles pour l'analyse du mouvement fluide

Participants : Étienne Mémin, Thomas Corpetti.

Nous avons défini un estimateur de mouvement spécifiquement dédié à des séquences d'images d'écoulements de fluides. Cet estimateur inclut un modèle d'attache aux données original bâti à partir de l'équation de continuité de la mécanique des fluides. Ce terme permet de prendre en compte l'effet des variations de l'intensité lumineuse occasionnées par la présence de mouvements divergents (présence de sources ou puits, variation de la densité de matière ou mouvements transverses au plan de visualisation). Cet estimateur comprend également une fonction régularisante de type "div-curl" du second ordre qui favorise l'émergence de zones compactes à rotationnel et à divergence non nuls. Cette approche a été évaluée et comparée, en premier lieu, sur des images météorologiques à un estimateur "générique" de champs denses de vitesses précédemment élaboré dans le projet [32]. Les trajectoires reconstruites des points de la grille pour des séquences comportant une trentaine d'images ont démontré la nette supériorité de l'approche dédiée, [12, 18, 38], vis à vis de l'estimateur de mouvement générique. Cette étude a été menée en coopération avec Patrick Pérez, actuellement à Microsoft Research à Cambridge, Angleterre.

Cette méthode a d'autre part été évaluée sur des écoulements expérimentaux spécifiques dans le cadre d'une collaboration scientifique avec le Cemagref de Rennes (équipe Aérobio de G. Arroyo). Ce dernier poursuit des recherches dans le domaine de l'aérodynamique pour les industries de transformation de l'agro-alimentaire. Les mesures de vitesses obtenues ont été

validées au moyen de critères spécifiques à la mécanique des fluides expérimentale. Le premier écoulement considéré a été un écoulement de type “couche de mélanges”. Ce type d’écoulement prend naissance à la rencontre de deux courants évoluant dans le même sens avec des amplitudes différentes. Les résultats de l’étude, décrits en détail dans [12], indiquent une cohérence entre les champs mesurés et la théorie attendue. La méthode possède par ailleurs, l’avantage de fournir une densité d’informations bien supérieure par rapport à celles fournies par les techniques usuellement employées en mécanique des fluides expérimentale (techniques dites de PIV, “Particle Image Velocimetry”^[MM02]). Des travaux en cours visent à valider la méthode sur d’autres types d’écoulements expérimentaux, en l’occurrence le cas de sillage proche en aval d’un cylindre et l’écoulement d’une couche de mélange en contre-courant. Les tout premiers résultats obtenus sont très encourageants et démontrent en particulier la possibilité pour cet estimateur de mesurer avec précision des phénomènes existant à des échelles différentes.

Nous avons également développé, en collaboration avec Bernard Delyon de l’Irmar, une technique permettant la détermination des points singuliers (vortex, puits, sources) d’un champ de vitesses de fluides, [12, 17]. Elle s’appuie sur une décomposition du champ des vitesses en une composante à divergence nulle (composante solénoïdale) et en une composante à rotationnel nul (composante irrotationnelle). Cette décomposition, dite de Helmholtz, est effectuée dans le domaine de Fourier, et permet d’estimer les deux fonctions de potentiel associées (le potentiel de vitesse et le potentiel de courant). Il est alors aisé de démontrer que les différents extrema locaux de ces fonctions de potentiel correspondent aux singularités du champ des vitesses. On peut ainsi aboutir à une représentation paramétrique du champ des vitesses, s’exprimant comme la somme de fonctions complexes décrivant l’influence sur le flot de chacune des singularités. L’estimation des différents paramètres du champ est obtenue par maximum de vraisemblance. Cette représentation permet finalement d’extraire les principales structures de l’écoulement. La méthode proposée s’avère d’autre part beaucoup plus robuste du point de vue de la localisation des points singuliers que les méthodes usuelles, basées sur les indices de Poincaré et permet en sus d’extraire directement les lignes de courant du champ.

6.1.2 Modélisation générale du mouvement pour le traitement vidéo

Participants : Patrick Bouthemy, Gwénaëlle Piriou, Nathalie Peyrard.

Dans le contexte de la reconnaissance et de la classification du mouvement en vidéo, nous avons tout d’abord développé une nouvelle modélisation probabiliste du contenu dynamique s’appuyant sur la même information de mouvement que celle utilisée pour les modèles gibbsiens causaux que nous avons précédemment définis, [21, 42], mais qui est plus parcimonieuse en terme de nombre de paramètres. Cette information de mouvement est donnée par une matrice des co-occurrences temporelles de mesures locales partielles (quantifiées) de mouvement. Les modèles gibbsiens causaux sont des modèles performants en terme de reconnaissance. De par leur grand nombre de degrés de liberté, égal à la taille de la matrice de co-occurrences, ces modèles sont capables de s’adapter à et d’apprendre tout type de contenu dynamique. Nous avons toutefois constaté que cela peut conduire à des problèmes de sous-segmentation en seg-

[MM02] S. MCKENNA, W. MCGILLIS, « Performance of digital image velocimetry processing techniques », *Experiments in Fluids* 32, 2002, p. 106–115.

mentation temporelle non supervisée d'une vidéo, et surtout à une possible dégénérescence du nombre de classes en classification non supervisée avec un algorithme du type EM, applications qui nous intéressent particulièrement. Nous modélisons cette fois la matrice des co-occurrences temporelles par un modèle de mélange gaussien en dimension 2, de manière à réduire la complexité du modèle. Un autre avantage de ces modèles est que ce sont des modèles bien connus pour lesquels on dispose d'algorithmes d'estimation efficaces.

Nous explorons par ailleurs une autre voie, différente à plusieurs égards. Cette étude est menée en collaboration avec Jian-Feng Yao de l'Irmar. Nous considérons désormais les variations temporelles (contrastes temporels) des mesures locales (partielles) de mouvement et les substituons à l'évaluation (coûteuse) des co-occurrences (spatio-)temporelles. Nous considérons ainsi un découpage de l'image en blocs spatiaux réguliers, et nous évaluons ces grandeurs pour chaque bloc spatial de l'image et pour tout l'intervalle de temps (bloc temporel de vidéo) considéré. Chaque histogramme est ensuite représenté par un mélange d'une loi normale et d'une mesure de Dirac en 0, les paramètres de ces lois (variance de la loi normale et poids du mélange) étant déterminés par le critère du maximum de vraisemblance. Cependant, en ne considérant que le contraste temporel des mesures locales de mouvement, on ne dispose que d'une information sur l'évolution de l'information de mouvement au cours du temps et non de sa nature en tant que telle. Nous ajoutons, pour chaque bloc spatial considéré, les mesures locales de mouvement calculées pour l'instant initial. Nous avons estimé la loi de ces mesures par un mélange de trois gaussiennes. Avec cette approche, il n'est plus nécessaire de considérer des mesures de mouvement discrétisées. On peut ainsi faire l'économie de l'étape de quantification des mesures locales partielles de mouvement, ce qui apporte un gain supplémentaire en temps de calcul et un gain en précision. Nous obtenons ainsi un modèle complet de mouvement évalué à partir de mesures élémentaires et de calculs immédiats et représentés par des lois simples à très faible nombre de paramètres. Pour l'instant, nous nous sommes intéressés au cas de séquences vidéos où la spatialisation joue un rôle, c'est-à-dire que des mouvements à des endroits différents dans le plan image traduisent des activités différentes. Il peut s'agir, par exemple, de reconnaître différents types d'activités dans un bureau ou d'analyser un trafic urbain à un carrefour important. Pour l'instant, nous avons considéré le modèle joint le plus simple, c'est-à-dire un modèle où les lois par bloc sont postulées indépendantes et le modèle global en est alors le produit. Après avoir déterminé les classes de mouvement considérées et en avoir estimé les paramètres sur un ensemble d'apprentissage, la reconnaissance de ces classes dans une vidéo se fait au sens du maximum de vraisemblance.

6.1.3 Modèles probabilistes de formes en imagerie médicale

Participants : Christian Barillot, Isabelle Corouge, Cybèle Ciofalo.

Nous étudions des techniques d'analyse statistique multidimensionnelle afin de modéliser des formes en imagerie médicale. Nous nous attachons à construire un modèle déformable générique afin de représenter une forme donnée et les déformations qu'elle peut subir. Ce modèle est défini à partir d'une population d'apprentissage, sur laquelle est effectuée une analyse en composantes principales (ACP). Nous avons d'abord considéré les sillons du cortex cérébral extraits d'images IRM. Disposant d'une représentation paramétrique de ces structures, nous

calculons pour chacune d'entre elles un repère local propre à partir de ses axes d'inertie ; cela nous permet d'aligner, de façon rigide, tous les exemplaires de l'ensemble d'apprentissage dans un référentiel local commun. Dans cet "espace local", les déformations d'une structure sont caractérisées par des vecteurs de déplacement par rapport à une structure de référence (structure donnée ou structure moyenne). Les composantes principales de la matrice de covariance de ces vecteurs de déplacement décrivent les modes principaux de déformation. Les tests réalisés montrent la pertinence des premiers modes. Ainsi, nous avons obtenu un modèle probabiliste de sillons corticaux, décrit par une forme moyenne et les variations statistiques, anatomiques dans le cas des sillons, autour de la forme moyenne, [37].

Deux applications de cette modélisation ont été explorées. D'une part, elle a servi aux travaux sur l'évaluation de méthodes de recalage global inter-sujets (méthodes rigides et non rigides, voir paragraphe 6.2.5). En effet, appliquée sur des amers locaux, cette analyse fournit un indicateur de la similarité des formes au sein des populations recalées, et produit ainsi un critère objectif de comparaison entre les méthodes de recalage testées. D'autre part, nous avons exploité la connaissance statistique apportée par le modèle des sillons dans la construction d'atlas anatomique et fonctionnel. En particulier, nous avons considéré le recalage inter-sujets de données fonctionnelles exprimées sous forme de dipôles MEG (localisations d'activations fonctionnelles), à partir des amers anatomiques que sont les sillons corticaux, [61]. L'approche adoptée est locale et non-linéaire, et permet d'étendre par interpolation le champ de déformation, obtenu via l'analyse statistique, entre un sillon donné et le sillon de référence, à un voisinage local du sillon considéré. Il est alors possible de recalculer les dipôles MEG vers un référentiel relatif au sillon de référence. Cette mise en correspondance d'activations fonctionnelles permet d'observer la variabilité fonctionnelle entre individus inhérente au groupe de sujets étudiés. Dans un but de validation, nous avons comparé cette méthode à d'autres approches globales classiques. Notre approche offre précisément l'avantage d'être locale et indépendante d'un cerveau de référence, et de modéliser parfaitement la déformation anatomique entre deux sillons. Nous avons étendu le modèle statistique à un couple de sillons et travaillons sur sa généralisation à un graphe probabiliste de sillons afin de décrire non plus seulement les caractéristiques morphologiques d'un sillon, mais également les relations de voisinage, de position et d'orientation entre sillons principaux. Nous nous intéressons enfin à l'application de ces concepts au recalage de données fonctionnelles inter-individuelles dans le cas d'amers venant d'informations fonctionnelles (délimitation d'aires visuelles), ce dernier travail se fait en collaboration avec l'Inserm U438 de Grenoble (M. Dojat, C. Segebarth).

Nous commençons par ailleurs à étendre cette problématique à la constitution de modèles probabilistes de formes dans un espace $3D+t$, dans le domaine de l'imagerie biomédicale et plus particulièrement pour des structures anatomiques et fonctionnelles en imagerie cérébrale (délimitations fonctionnelles, structures cérébrales corticales ou profondes, . . .). Un des objectifs est de faire émerger de nouveaux outils de quantification en imagerie biomédicale. Nous comptons combiner une modélisation statistique d'un ensemble de formes provenant de différentes observations et des méthodes variationnelles de segmentation de ces formes à l'aide de modèles déformables à topologie et co-dimension variables.

6.1.4 Détection sûre du mouvement

Participants : Frédéric Cao, Patrick Bouthemy, Thomas Veit.

Nous avons commencé des travaux sur la détection “sûre” du mouvement (selon un critère probabiliste de détection “sans paramètre”). L’objectif est de comprendre quels sont les éléments géométriques les plus simples, mais qui sont également les plus robustes et les plus génériques, qui font qu’un changement temporel entre deux images est jugé significatif et devient un mouvement perceptuel. Cette approche repose sur un principe perceptuel, appelé principe de Helmholtz, selon lequel une conjonction d’une même qualité sur plusieurs objets locaux devient significative lorsqu’on suppose que, toute chose égale par ailleurs, la qualité est uniformément et indépendamment distribuée sur les objets en question. Un premier exemple d’application a fait l’objet d’une étude pour la DGA, [60](voir paragraphe 7.5). Le but est de trouver des seuils de détection à partir desquels on peut être sûr qu’un objet “ponctuel” (traduit par une trajectoire) a traversé la scène. Pour ce faire, on calcule la probabilité de certains événements dans ce qui a été par ailleurs appelé MHI (Motion History Image^[BD01]) et utilisé pour de l’apprentissage du mouvement.

De manière générale, un objectif est de rechercher des structures semi-locales caractéristiques d’un mouvement ; on les considérera significatives lorsque leur nombre d’occurrences fortuites est trop faible pour qu’elles soient dues au hasard. On montre dans de nombreux cas, que cela permet de régler des seuils de manière automatique et que leur valeur n’influence la détection que de manière logarithmique. La thèse de Thomas Veit (commencée début octobre 2002) s’inscrit dans cette démarche. Le travail va en fait démarrer par la définition d’une méthode de calcul rapide et compact d’un flot optique approximatif sur certaines bases de fonctions. Dans un deuxième temps, on cherchera à mettre en application le principe de Helmholtz sur des séquences d’images de fluides en mouvement et notamment pour une détection sûre de structures tourbillonnaires. Si les résultats sont concluants, on essaiera de généraliser la démarche à d’autres événements spatio-temporels, qui pourront s’insérer dans le projet IST Lava (voir paragraphe 7.7.2).

6.1.5 Résumé vidéo par segmentation et classification du mouvement

Participants : Patrick Bouthemy, Nathalie Peyrard.

Construire un résumé d’une vidéo (voire également le paragraphe 7.8.1) peut se réaliser en deux étapes principales : segmentation temporelle de la vidéo en plages homogènes, puis sélection (par classification supervisée) des plages pertinentes. Nous avons proposé une approche originale pour la segmentation séquentielle de vidéo, basée sur l’information de mouvement contenue dans la vidéo, [55]. Cette méthode ne nécessite pas de connaissance a priori sur la vidéo, ni de phase d’apprentissage. Elle repose sur l’analyse de l’évolution temporelle du contenu dynamique dans la vidéo. Ce contenu dynamique est caractérisé par la modélisation probabiliste de la distribution empirique des co-occurrences temporelles de mesures locales liées au mouvement et quantifiées, selon un modèle gibbsien causal que nous avons

[BD01] A. BOBICK, J. DAVIS, « The recognition of human movement using temporal templates », *IEEE Trans. on Pattern Analysis and Machine Intelligence* 23, 3, mars 2001, p. 257–267.

précédemment proposé [21]. Ce cadre probabiliste nous permet de construire une mesure de similarité en terme de contenu dynamique à partir de la divergence de Kullback-Leibler, que nous utilisons dans la formulation d'un critère de décision pour la segmentation. La méthode de segmentation consiste alors à appliquer séquentiellement ce critère entre la plage homogène en cours de constitution et le bloc temporel suivant de la vidéo (bloc élémentaire formé de quelques images), de manière à former progressivement des segments homogènes en terme de contenu dynamique. Des expérimentations sur plusieurs vidéos (fiction, reportages sportifs) ont été menées. Nous avons également exploré une autre voie pour la segmentation temporelle d'une vidéo selon son contenu. Les changements de mouvement de caméra (entre plan fixe, panoramique, zoom, travelling, . . .), généralement assimilables aux changements du mouvement dominant dans l'image, représentent aussi une source d'information importante. Pour cela nous étudions, à travers des tests de Hinkley, les variations temporelles de certains coefficients du modèle de mouvement affine 2D global, estimé de manière robuste entre deux images successives de la vidéo. Les segments ainsi obtenus fournissent des unités temporelles raisonnables sur lesquelles on peut appliquer des algorithmes de reconnaissance ou de classification supervisée du contenu dynamique pour sélectionner les plages pertinentes devant former le résumé vidéo.

Nous nous intéressons par ailleurs à la classification non supervisée du contenu dynamique de vidéos, et nous voulons également aborder des problèmes plus complexes de catégorisation du mouvement. Par catégorisation, il faut entendre classification accompagnée d'information sur l'organisation des classes entre elles. Comme déjà décrit précédemment, les méthodes étudiées se basent sur une information de mouvement donnée par une matrice des co-occurrences temporelles de mesures locales partielles de mouvement. Cette information est ensuite caractérisée par un modèle probabiliste (modèle gibbsien ou modèle de mélange gaussien). Dans le cadre du projet IST Lava (voir paragraphe 7.7.2), une première approche étudiée est celle de l'apprentissage par méthodes de noyaux, notamment comment combiner modèle probabiliste et noyaux et quelles performances peut-on en espérer ? Parallèlement, nous nous intéressons aux méthodes de type "Self-Organizing Maps" (SOM), algorithmes qui permettent non seulement d'obtenir une classification des données, mais aussi un ordonnement des classes entre elles, fournissant ainsi une information d'un niveau supérieur à une simple classification. Pour cette deuxième approche, les modèles de mélanges gaussiens paraissent bien adaptés du fait de leur petit nombre de paramètres.

6.1.6 Formation et filtrage des images échographiques 3D

Participants : Christian Barillot, Arnaud Ogier.

Le but de ce travail est de pouvoir mieux exploiter les images échographiques tridimensionnelles afin d'améliorer le diagnostic et le suivi thérapeutique. Nous nous plaçons plus particulièrement dans le contexte d'acquisition de type "main libre". L'imagerie échographique présente de nombreux avantages, comme sa relative innocuité, son faible coût (en comparaison d'autres systèmes en cours comme le scanner X ou l'IRM), sa portabilité et la grande diffusion des imageurs 2D, son caractère spatio-temporel intrinsèque. En contre-partie, cette imagerie souffre encore de nombreux inconvénients qui limitent son impact clinique, tels que la faible qualité des images acquises (notamment en terme de rapport signal-à-bruit), les contraintes opéra-

tionnelles pour l'acquisition et la nécessité d'une grande expertise pour l'interprétation des données (le praticien doit effectuer lui-même l'examen pour l'interpréter), la difficulté d'accès à des informations géométriques quantitatives. Nous nous intéressons dans un premier temps, dans le cadre de la thèse d'Arnaud Ogier qui vient de débiter, à l'étude de la nature des images acquises pour aider à la définition d'outils de reconstruction et de traitement mieux adaptés. Il s'agit notamment de mieux rendre compte des différentes composantes du bruit de formation des images échographiques (mélange de bruit additif venant de la chaîne d'acquisition électronique, et de bruit multiplicatif venant de la physique de formation des images), ainsi que des irrégularités photométriques et géométriques des images reconstruites (inhomogénéités, effet de volume partiel, distorsions, . . .).

6.2 Suivi temporel, recalage, reconstruction

Contexte : Le problème générique du suivi dans des séquences d'images 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. Cela peut être étendu au problème du recalage entre images volumiques, en particulier dans le domaine de l'imagerie médicale. 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" ou de "leurres" autour des données utiles. Nous traitons ce problème à deux niveaux : d'un point de vue général et théorique, ce qui comprend le choix et la représentation des entités suivies et du modèle d'évolution temporelle, la reconstruction de trajectoires complètes, le suivi d'objets multiples, la gestion des occultations et des croisements, la prise en compte des changements de topologie des formes suivies; d'un point de vue plus spécialisé dans des contextes d'applications donnés.

6.2.1 Suivi multi-objets en imagerie par filtrage particulière

Participants : Jean-Pierre Le Cadre, Carine Hue.

Pour traiter les problèmes de suivi temporel, des méthodes probabilistes à base de filtrage de Kalman sont généralement employées. De telles techniques présentent cependant des limitations notables pour gérer les non-linéarités induites par les équations d'observation. Elles se heurtent en particulier au problème d'entités multiples, qui nécessite la manipulation de lois multimodales. Pour pallier ces problèmes, des techniques d'association de données, basées sur des mélanges de lois normales, ont été développées dans le domaine de l'extraction multipistes. D'autre part, les méthodes de filtrage particulière^[BDM01], basées sur la manipulation de jeux d'échantillons assortis de poids, permettent de traiter directement les problèmes d'estimation de paramètres variables dans le temps, dans un cadre non linéaire.

Nos travaux se sont focalisés sur le suivi d'objets, mobiles et partiellement observés, par des méthodes de filtrage particulière. Dans un premier temps, les efforts ont consisté à analyser les techniques classiques de filtrage particulière dans le cadre du suivi d'un objet unique

[BDM01] N. BARTOLI, P. DEL MORAL, *Simulation et algorithmes stochastiques*, Cépaduès-Éditions, 2001.

et manœuvrant, partiellement observé (par exemple, dans le cas de mesures d'angles seules en trajectographie sonar). Cette analyse a permis de mettre en évidence l'intérêt de telles méthodes. On peut ainsi souligner leur robustesse, que ce soit par rapport au modèle de trajectoire de l'objet, au bruit d'estimation, ou, bien sûr, aux non-linéarités du modèle. Il est cependant certain que de nombreuses applications requièrent le suivi d'objets multiples. Les difficultés sont alors nombreuses. Les trajectoires des différents objets doivent être estimées à partir d'informations relativement pauvres et en présence de fausses alarmes. La difficulté essentielle réside cependant dans la méconnaissance de l'association des mesures aux modèles des objets. Il s'agit alors de résoudre simultanément un problème d'association et d'estimation. Le formalisme que nous avons adopté est celui des lois de mélanges de densités, où il s'agit d'estimer à la fois les variables d'affectation des mesures aux pistes et les états des objets, [25]. On utilise des hypothèses d'association semblables à celles des méthodes de type PMHT. Dans notre contexte, ceci a été réalisé en combinant d'une part le filtrage particulaire et, d'autre part, un échantillonneur de Gibbs. Ainsi, les probabilités d'assignation des mesures aux modèles sont-elles aussi estimées par simulation. L'algorithme correspondant permet d'estimer les trajectoires d'objets obéissant à des trajectoires markoviennes. Celui-ci hérite des propriétés de robustesse du filtrage particulaire mono-objet, tout en requérant une charge de calcul raisonnable.

Du fait de la généralité de l'approche et de la flexibilité des hypothèses, il est possible d'étendre cette approche à de nombreux autres problèmes : extraction multi-pistes distribuée, association actif/passif, optimisation des mesures. La mise en œuvre de ces méthodes pour le suivi d'objets multiples dans des séquences d'images a été l'objet de développements (suivi de piétons). L'analyse des performances a constitué un autre volet de notre activité; en particulier la comparaison entre les performances empiriques et théoriques (bornes de Cramér-Rao), [45]. Plus précisément, le calcul des bornes a été développé pour diverses logiques d'association de données et dans le cadre de cibles diffusantes. Ce travail continue à se faire en coopération avec Patrick Pérez, actuellement à Microsoft Research à Cambridge, Angleterre.

6.2.2 Suivi de points à partir de séquences d'images

Participants : Étienne Mémin, Elise Arnaud, Bruno Cernuschi-Frias.

Le suivi de points à partir d'une séquence d'images constitue une tâche fondamentale pour un bon nombre d'applications en vision par ordinateur. Il s'agit de reconstruire la trajectoire de points donnés à partir d'informations partiellement cachées. Contrairement au suivi de formes structurées, on ne peut dans ce cas s'appuyer que sur des informations locales pour lesquelles il est difficile d'élaborer des modèles dynamiques *a priori* adéquats. Nous avons spécifié une méthode de suivi qui couple des aspects de corrélation photométrique avec des contraintes différentielles de conservation photométrique du type "flot optique". Cela conduit à faire coopérer des modèles de mouvement apparents calculés dans la séquence d'images avec des mesures de corrélation invariantes aux déformations affines. Ces dernières mesures sont établies entre les positions initiales des points considérés et leurs positions courantes, tandis que les modèles de mouvement sont évalués entre deux instants successifs.

La coopération des modèles de mouvement et des mesures de corrélation est formulée de

manière probabiliste dans un cadre de filtrage stochastique. L'association de contraintes différentielles et de mesures de corrélation a permis d'obtenir d'excellents résultats de suivi dans deux types de situations générales. La première concerne les séquences caractérisées par un mouvement d'ensemble linéaire défini sur l'ensemble de l'image. Dans ce cas, nous obtenons un filtre linéaire conditionnellement à la séquence d'images. Ce filtre, de type Kalman, est dérivé au moyen d'un estimateur de variance minimum conditionnelle. La seconde situation correspond au cas où chaque point exhibe un modèle de mouvement qui lui est propre. Le filtre obtenu appréhende alors des mouvements d'ensemble non linéaires mais pouvant être localement décrits au moyen de modèles linéaires. Il correspond à un filtre particulière conditionnel pour lequel la distribution d'importance optimale est connue. Ces deux filtres s'avèrent bien adaptés à la reconstruction de trajectoires caractérisées par des changements de direction très marqués, comme cela peut être notamment le cas pour des écoulements fluides.

6.2.3 Trajectographie pour cibles manœuvrantes

Participants : Jean-Pierre Le Cadre, Thomas Bréhard.

Le domaine de la trajectographie passive se démarque nettement de l'extraction multi-pistes puisqu'il n'y a pas fondamentalement d'incertitude sur l'origine des mesures. Cette relative simplicité ne doit pas masquer les difficultés auxquelles on doit faire face : initialisation des filtres, manœuvres de la cible, pertes de détection, robustesse vis-à-vis du modèle de cible, pureté de la piste extraite. Dans le but de remédier à certains de ces problèmes, nos efforts ont tout d'abord porté sur le filtrage n -conditionnel. Cet algorithme a été développé par Noyer^[Noy96], et consiste à conditionner le tirage aléatoire des particules aux n dernières mesures. En d'autres termes, l'algorithme génère à chaque pas de temps des parties de trajectoires de particules puis retient la meilleure. Plus précisément, l'idée de base est de diffuser N fois chaque particule et de ne retenir que la meilleure en termes de vraisemblance. Cette méthode donne des résultats sensiblement meilleurs que les algorithmes classiques de filtrage particulière, lorsque le mouvement de la cible présente une forte variabilité.

Pour modéliser des trajectoires complexes, nous avons ensuite proposé une méthode originale, dénommée "Batch Sequential Importance Sampling". Nous avons considéré des modèles hybrides constituées d'un état à valeurs continues, mais dont l'équation d'évolution est déterminée par un mode à valeurs discrètes, pouvant par exemple correspondre à un mouvement rectiligne uniforme, à un virage à taux de giration constant, . . . Ces modes peuvent être liés à des données d'identification (type de la cible). Le passage d'une loi à une autre peut aussi être modélisé par une chaîne de Markov. La méthode la plus utilisée est sans doute l'IMM (Interacting Multiple Models). En effet, elle permet de réduire l'arbre des hypothèses à une simple loi de mélange. Cependant, les algorithmes IMM sont basés sur une hypothèse d'observation linéaire ce qui en réduit très fortement l'intérêt pour cette étude. Le filtrage particulière présente par contre un intérêt certain. Nous considérons que chaque particule représente un bout de trajectoire. Notre méthode ne se contente pas d'estimer l'état (continu) et le mode courant, elle met de plus à jour le passé proche de la trajectoire. Le problème est alors de

[Noy96] J. NOYER, *Traitement non-linéaire du signal radar par filtrage particulière*, thèse de doctorat, Université de Toulouse, novembre 1996.

développer une formule de mise à jour des poids normalisés. Les résultats obtenus en suivi d'une cible effectuant des manœuvres très serrées sont des plus encourageants. Par ailleurs, les mesures (gisements estimés) sont fréquemment entachées de biais, généralement variables dans le temps. Il apparaît nécessaire, dans un premier temps, de quantifier l'influence de ceux-ci au moyen de la borne de Cramér-Rao a posteriori, puis, de considérer l'estimation-poursuite conjointe de la trajectoire de la cible et des biais.

6.2.4 Détection de structures spatio-temporelles

Participant : Frédéric Cao.

Il s'agit de détecter, avec des seuils fixés automatiquement, des structures géométriques simples et génériques qui sont à la base de la vision de bas niveau, autant dans la perception des images fixes que dans la perception du mouvement. Dans les deux cas, un principe géométrique important, et annoncé comme tel par les fondateurs de la théorie de la Gestalt, est celui de bonne continuation. Grâce à ce principe, nous sommes capables de reconstruire des objets ou des trajectoires en partie occultés. Plus fondamentalement, une entité est perçue comme un tout, lorsqu'on peut grouper les éléments qui la composent en vertu d'un principe de bonne continuation. La définition d'un tel principe peut se faire dans une formulation probabiliste utilisant le principe de Helmholtz, formulée il y a quelques années par des chercheurs de l'ENS Cachan^[DMM00]. Appliqué aux lignes de niveau d'une image, ce principe permet de détecter, de manière surprenante car invariante par changement de contraste, une très grande partie des contours des objets [59]. Plus précisément, on se donne un entier $N_\theta \geq 1$ et on appelle $\theta_i = \frac{i\pi}{2N_\theta}$. On se donne un ensemble de courbes échantillonnées régulièrement. Si Γ est une sous-courbe telle qu'il existe $1 \leq i \leq N_\theta$ tel que la variation de l'angle de la tangente est en tout point inférieure à θ_i , on appelle nombre de fausses alarmes le nombre

$$NFA(\Gamma) = N_c N_\theta \left(\frac{\theta_i}{\pi} \right)^l,$$

où θ_i est choisi minimal, N_c est le nombre total de sous-courbes (qu'on peut donc calculer une fois donné l'ensemble de courbes), et l est la longueur de la sous-courbe considérée. Si $NFA(\Gamma) \leq \varepsilon$, on dit que Γ est une ε -bonne continuation. Cette définition s'explique par le résultat suivant : si on suppose que, toute chose égale par ailleurs, les courbes sont des réalisations de marche aléatoire avec des accroissements indépendants et uniformément distribués dans $(-\pi, \pi)$, alors l'espérance du nombre de ε -bonnes continuation est inférieur à ε . La suite naturelle de ces travaux est de mettre en œuvre ce principe de bonne continuation dans des séquences d'images afin de grouper des structures spatiales correspondant à un objet mobile pour en effectuer le suivi (obtention de trajectoires).

6.2.5 Recalage d'images cérébrales

Participants : Pierre Hellier, Christian Barillot, Isabelle Corouge.

[DMM00] A. DESOLNEUX, L. MOISAN, J.-M. MOREL, « Maximal meaningful events and application to image analysis », *rapport de recherche n°2000-22*, CMLA, ENS-Cachan, 2000.

Nous avons poursuivi notre travail portant sur l'évaluation de différentes méthodes de recalage non-rigide d'images cérébrales, [35]. Nous avons dans ce domaine initié et mené un projet international d'évaluation de différentes méthodes de recalage non-rigide inter-sujets d'images cérébrales. Sur une base commune de dix-huit sujets, six méthodes importantes avaient été évaluées. La méthode SPM (Statistical Parametric Mapping, <http://www.fil.ion.ucl.ac.uk/spm/>), très répandue dans la communauté de recherche en neurosciences, est venue enrichir la liste des méthodes évaluées dans ce projet. Ce travail a été réalisé en collaboration avec le laboratoire Wellcome Department of Cognitive Neurology situé à Londres. Pour les mesures globales, les résultats ont confirmé l'importance du nombre de degrés de liberté de la transformation considérée dans le recalage. Par contre, la méthode SPM semble donner des résultats statistiquement plus satisfaisants pour les mesures locales, basées en grande partie sur le recalage de sillons corticaux, [44].

Nos avons également étendu ce cadre méthodologique d'évaluation à des données fonctionnelles, [61]. Les méthodes de recalage non-rigide inter-sujets visent en particulier à construire des atlas probabilistes anatomiques et fonctionnels du cerveau. L'hypothèse sous-jacente est que la variabilité fonctionnelle inter-individuelle peut se décomposer grossièrement en une variabilité anatomique, corrigée par les approches de recalage, et une variabilité fonctionnelle résiduelle, qui est l'objet d'étude. Sur une partie de la base des 18 sujets, nous disposons d'acquisitions fonctionnelles correspondant à une stimulation sensorielle de trois doigts. Ce type de paradigme, connu et maîtrisé, est particulièrement adapté dans ce cadre. Nous avons donc étudié l'impact des approches de recalage sur la dispersion des données fonctionnelles. Alors que la dispersion des données fonctionnelles recalées varie selon l'approche de recalage, il est toutefois possible de retrouver la somatotopie des aires sensorielles.

6.2.6 Images ultrasonores 3D, calibration et reconstruction

Participants : Christian Barillot, Pierre Hellier, François Rousseau.

L'échographie tridimensionnelle en mode "main libre" offre de nombreux avantages comme une grande souplesse d'utilisation, un domaine d'utilisation plus large, mais nécessite une phase cruciale de calibration pour repérer précisément les plans de chaque acquisition d'images échographiques 2D dans l'espace tridimensionnel afin de reconstruire un volume d'images échographiques 3D. Dans ce domaine, nous avons travaillé sur deux points : d'une part, le développement d'une approche automatique et robuste de calibration; d'autre part, la validation de différentes méthodes de calibration, [65].

Afin que l'utilisation clinique de l'échographie 3D en mode "main libre" soit possible, il est nécessaire que la phase de calibration soit aisée, automatique, robuste et rapide. Ceci est particulièrement vrai lorsque la liaison entre la sonde et le capteur de position ne peut être fixée et est variable dans le temps. Cet objectif nous a conduit à développer une méthode de calibration basée sur un fantôme simple (une bassine d'eau). Les paramètres de calibration peuvent être estimés à partir de la structure géométrique connue du fantôme. L'approche développée, [58], est automatique, contrairement à d'autres méthodes où l'interaction manuelle et le réglage des paramètres peuvent être fastidieux. La robustesse de l'approche a été améliorée d'une part en utilisant la transformée de Hough pour détecter la ligne dans les images échographiques

correspondant au fond de la bassine, et d'autre part par un test de vraisemblance inter-image sur la détection des caractéristiques géométriques de la bassine. Ce travail a fait l'objet d'une collaboration avec le professeur Pierre Darnault du CHU Pontchaillou, Rennes.

Par ailleurs, nous avons élaboré une méthodologie de comparaison objective de plusieurs méthodes de calibration. Un des standards du domaine est le logiciel StradX développé à l'Université de Cambridge. La société SofamorDanek a également développé une méthode de calibration qui emploie une mire géométrique différente. L'évaluation des différentes méthodes de calibration se base sur la comparaison de segmentations réalisées sur des volumes échographiques reconstruits. Pour cela, nous travaillons en collaboration avec l'Université d'Utrecht dans le cadre d'une action intégrée du programme VanGogh (voir paragraphe 8.4.2). Les premiers résultats de cette étude tendent à montrer que la calibration a une influence sensible sur les volumes mesurés, cette influence se faisant plus particulièrement ressentir sur des mouvements particuliers de la sonde lors de l'acquisition.

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 différents types. 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 Utilisation des moments en asservissement visuel

Participants : François Chaumette, Omar Tahri.

Le but de cette étude est de déterminer une représentation optimale et générique des informations visuelles pour l'asservissement visuel 2D. Il s'agit de sélectionner dans l'ensemble des informations visuelles qu'il est possible de suivre dans une séquence d'images le nombre minimal de celles-ci, tout en offrant de bonnes propriétés de découplage des degrés de liberté du système, de robustesse de la commande aux erreurs de mesure, d'absence de singularités et de minima locaux, d'adéquation des mouvements 3D générés. L'ensemble de ces propriétés peut être étudié à partir de la matrice d'interaction associée aux informations visuelles. Dans ce contexte, nous avons considéré les moments de la projection d'un objet, facilement mesurables à partir d'une segmentation spatiale de l'image, [36]. Nous avons tout d'abord élaboré une nouvelle méthode pour l'obtention de moments invariants aux translations, aux rotations, et aux changements d'échelle valable dans un espace de dimension quelconque. En utilisant un modèle de projection sphérique plutôt que le classique modèle de projection perspective, les invariants obtenus par cette méthode ont été exploités pour réaliser un découplage des rotations vis à vis des translations de la caméra. Un choix judicieux des invariants à introduire dans la boucle de commande a permis d'exhiber un domaine de convergence très large. Plus précisément, nous avons considéré les coordonnées dans l'image du centre de gravité de l'objet,

la valeur de sa surface calculée par projection sphérique (qui est entièrement invariante à tout mouvement de rotation), deux combinaisons des invariants aux translations, aux rotations et aux changements d'échelle, et enfin, l'orientation de la projection de l'objet dans l'image. Lorsque l'objet est parallèle au plan image, la matrice d'interaction (ou jacobien) associée à ces six informations visuelles est alors triangulaire, ce qui répond à l'objectif poursuivi. Des résultats expérimentaux sur la cellule robotique du projet ont validé les résultats théoriques obtenus pour des objets continus de forme complexe (segmentés dans l'image par une simple binarisation).

Ces résultats ont aussi été étendus au cas d'un objet constitué par une nuage de points d'intérêt. Les moments ne sont alors plus calculés par intégration sur la surface de la projection de l'objet, mais par sommation discrète des points considérés. Pour effectuer cette généralisation, il a fallu déterminer la forme analytique de la matrice d'interaction pour les moments d'ordre quelconque dans ce cas de figure. Il a fallu alors introduire certaines modifications dans le choix des six informations visuelles pour contrôler les six degrés de liberté du système. À titre d'exemple, la surface n'est plus définie dans le cas d'un objet constitué d'un ensemble discret de points, et il a fallu lui substituer un nouvel invariant aux rotations.

6.3.2 Asservissement visuel virtuel pour la réalité augmentée

Participants : Éric Marchand, François Chaumette, Andrew Comport.

L'asservissement visuel virtuel consiste à considérer le calcul de pose (c'est à dire le calcul de la position de la caméra par rapport à un objet) comme le problème dual de l'asservissement visuel. Cette méthode prend l'image de l'objet comme le motif à atteindre, le motif courant étant obtenu par la reprojection du modèle 3D de l'objet sur le plan image d'une caméra virtuelle. La position initiale de cette caméra virtuelle peut être *a priori* quelconque et sa position finale (obtenue lorsque l'erreur entre le motif courant et le motif désiré est nulle) correspond à la pose recherchée. Cette méthode est en fait très proche de celle proposée par D. Lowe^[Low92] qui consiste à minimiser une fonction non-linéaire par l'algorithme de Levenberg-Marquardt. L'avantage de cette méthode par rapport à celle de Lowe tient dans le fait que la matrice Jacobienne reliant le mouvement de la caméra (la modification de la pose) au déplacement de l'objet dans le plan image est parfaitement connue, que la nature des informations visuelles utilisables peut être très variée et que ces dernières peuvent même être aisément combinées, [51, 52]. Le gain se situe donc à la fois sur la simplicité de la modélisation et donc de la mise en œuvre, mais aussi sur la richesse de l'information utilisable.

Afin de prendre en compte d'éventuelles mesures aberrantes, nous avons introduit dans les lois de commande, sur lesquelles repose l'asservissement visuel virtuel, des estimateurs robustes (des M-estimateurs). Nous avons en effet abordé l'application de ces estimateurs robustes au problème d'asservissement visuel. L'objectif est de pouvoir prendre en compte, directement au niveau de la loi de commande, la présence éventuelle de mesures aberrantes résultant d'erreurs dans l'extraction, la mise en correspondance ou le suivi des informations visuelles. Cette approche statistique permet en l'occurrence d'obtenir une estimation beaucoup plus précise

[Low92] D. LOWE, « Robust model-based motion tracking through the integration of search and estimation », *Int. J. of Computer Vision IJCV-8*, 2, 1992, p. 113–122.

de la pose. Un travail important a aussi été réalisé sur les primitives visuelles utilisées dans la loi de commande. Dans un premier temps, des primitives géométriques simples avaient été considérées. Afin d'englober dans le calcul de pose un maximum d'information visuelles, nous considérons désormais comme primitives des distances entre des points et des primitives géométriques particulières (droites, ellipses, ...). Les matrices d'interaction reliant la variation de ces distances au mouvement de la caméra ont été explicitement calculées. Les algorithmes résultants permettent un suivi d'objet complexe à la cadence vidéo. Les résultats ont été validés sur de multiples séquences d'images.

6.3.3 Asservissement visuel et réalité virtuelle

Participant : Éric Marchand.

Ces travaux ont été menés en collaboration avec Nicolas Courty et Bruno Arnaldi (projet Siames). Notre objectif est de proposer un cadre général permettant le contrôle d'une caméra dans un environnement virtuel. La caméra doit non seulement pouvoir se positionner par rapport à son environnement, mais elle doit de plus être à même de réagir à des modifications de celui-ci. La méthode proposée, [29], repose sur l'asservissement visuel qui permet de générer automatiquement des mouvements 3D de la caméra à partir d'une tâche spécifiée dans l'image. Cette approche de contrôle des humanoïdes a aussi été intégrée dans des systèmes plus vastes de gestion du comportement. En particulier, des modules de gestion des mouvements de la caméra spécialisés dans la cinématographie virtuelle ont été définis.

6.3.4 Couplage perception-action par indexation d'images

Participants : François Chaumette, Anthony Remazeilles.

Cette étude est menée en collaboration avec Patrick Gros (projet TexMex). Elle porte sur la réalisation de très grands déplacements par un système robotique muni d'une caméra embarquée, [53]. Ce déplacement est défini par une image initiale, fournie par la caméra avant que le système robotique ne se déplace, et une image finale que la caméra doit retrouver une fois le déplacement réalisé. Dans ce contexte, ces deux images sont a priori totalement différentes, et ne présentent le plus souvent aucune primitive commune. Cette absence de recouvrement rend impossible l'utilisation des techniques actuelles d'asservissement visuel. Nous avons décomposé ce problème en deux phases : *i*), localisation du système robotique dans son environnement; il est aussi nécessaire de localiser la position à atteindre à la fin du déplacement et de mettre en relation ces deux positions; *ii*), réalisation du déplacement, en contrôlant celui-ci par asservissement visuel. Dans la première étape, nous utilisons une base d'images acquises lors d'une phase préalable d'apprentissage.

Nous avons réalisé un module qui permet tout d'abord de déterminer les images de la base les plus proches des images initiale et finale correspondant au déplacement visé, et de sélectionner dans cette même base une succession d'images intermédiaires permettant de décrire le chemin à parcourir. Ce module exploite une approche d'indexation d'images et utilise des descripteurs locaux invariants de points d'intérêt fournis par un détecteur de coins. L'intérêt de l'utilisation de techniques d'indexation d'images est notamment le gain substantiel en temps

de calcul. Chaque couple d'images successives du chemin permet ensuite de définir un asservissement visuel. L'ensemble des associations d'images possibles est représenté par un graphe, dont les poids des arcs quantifient la "difficulté" pour le système robotique à effectuer le déplacement défini par ces deux images. L'algorithme du plus court chemin de Dijkstra, entre les noeuds du graphe correspondant aux images initiale et finale, permet d'obtenir la succession d'images décrivant le chemin à parcourir.

Nous nous intéressons actuellement à la réalisation du déplacement défini par cette collection d'images. Les primitives mises en correspondance entre ces images successives doivent permettre de planifier une trajectoire entre l'image initiale et l'image finale, mais sans forcément passer par les positions liées aux images intermédiaires. Cependant, nous devons nous assurer que le système robotique reste dans un environnement connu, c'est-à-dire que des primitives mises en correspondance entre images du chemin sont toujours visibles sur l'image fournie par la caméra lors du déplacement du système robotique. Nous avons implanté un algorithme de reconstruction pour disposer dans un même espace de représentation de l'ensemble des primitives détectées dans les images du chemin. Nous effectuons plus précisément une reconstruction projective, afin d'être indépendant des paramètres de la caméra ayant acquis les images de la base, et de celle embarquée sur le système robotique. Cette reconstruction projective nous permet, par reprojection dans le plan image, de décider si la position considérée de la caméra est valide ou non. Enfin, nous avons implanté un algorithme de suivi robuste de primitives, à la cadence vidéo, pour la réalisation de l'asservissement visuel. Il s'appuie sur les travaux de Shi et Tomasi^[ST94].

6.3.5 Planification de trajectoire pour l'optimisation de la navigation

Participants : Jean-Pierre Le Cadre, Sébastien Paris.

L'objectif général est d'étudier et de développer un outil d'aide à l'élaboration d'une trajectoire optimisée, afin d'améliorer la précision de navigation, [54]. Cette tâche est particulièrement importante lorsque l'on dispose d'observations extérieures au système de navigation lui-même. Ces informations extérieures sont fusionnées avec les observations internes via l'algorithme de suivi de trajectoire. Il faut noter que cette trajectoire optimisée doit se trouver à l'intérieur d'un domaine de contraintes : contraintes de manœuvres, contraintes opérationnelles, positions initiales et finales, contraintes de temps, . . . Il s'agit d'un problème de planification dont le but est de déterminer la suite des décisions qui optimise une fonctionnelle de coût globale. Pour quantifier la qualité de l'information, une fonctionnelle basée sur le calcul des matrices de Fisher associées à l'estimation de la suite des états est parfaitement adaptée. Les difficultés sont cependant nombreuses : non existence d'une relation d'ordre total sur les matrices de Fisher associées aux trajectoires; absence de monotonie des fonctionnelles adaptées pour l'évaluation des gains d'information; impossibilité de calculer une borne supérieure (autrement que par énumération directe des trajectoires possibles) pour le gain d'information associé à un début de chemin, et conséquence pratique, l'utilisation de méthodes de type "Branch and Bound" n'apporterait aucun gain combinatoire; difficultés liées à la définition de l'information

[ST94] J. SHI, C. TOMASI, « Good features to track », *in: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'94*, p. 593–600, Seattle, juin 1994.

globale apportée par une trajectoire; définition des objectifs de l'optimisation.

D'une part, le calcul de cette borne a été développé et validé pour les modèles complets; cette dernière semble être un bon outil pour quantifier les performances du filtrage particulière. D'autre part, les fonctionnelles d'information locales ne représentent que localement le gain d'information. Alors que pour celles-ci, on dispose de méthodes d'optimisation rapides et incluant toutes les contraintes opérationnelles, il n'en est pas de même pour les fonctionnelles d'information basées sur le calcul de la borne. L'étude s'est donc orientée sur l'utilisation de méthodes d'optimisation statistique. Il s'agit tout d'abord de générer des trajectoires candidates, suivant un modèle de diffusion convenable et incluant les hypothèses opérationnelles relatives aux contraintes de manœuvre ainsi que leurs fréquences. Cette étape est déjà disponible au travers des algorithmes de programmation dynamique du type MDP (Markov Decision Process). Il est alors possible d'obtenir une approximation de l'histogramme des valeurs d'une fonctionnelle de la borne associée à une trajectoire donnée. Cette fonctionnelle peut être la valeur du déterminant de la matrice de Fisher à l'instant final. Ayant obtenu cet histogramme par tirage des coûts élémentaires, on est alors parfaitement en mesure d'encadrer la probabilité de trouver une trajectoire, respectant les contraintes opérationnelles et associée à une valeur de fonctionnelle supérieure à un seuil donné. Ces travaux ont été menés dans le cadre d'un contrat avec le LRBA (voir paragraphe 7.2).

6.3.6 Théorie de l'exploration

Participant : Jean-Pierre Le Cadre.

Ce travail s'est fait en collaboration avec Frédéric Dambreville, ancien doctorant du projet et actuellement en post-doctorat à NPS, Monterey aux USA. Les jeux de recherche sont le cadre naturel pour l'optimisation de la stratégie de l'observateur en présence d'incertitudes sur celle de la cible. À ce titre, le modèle de Koopman^[Koo99] présente un certain nombre d'inconvénients, contrepartie de sa capacité d'optimisation massive. Ainsi est-il basé sur un a priori probabiliste markovien du mouvement de la cible qui est plutôt restrictif. L'aspect probabiliste ne prend pas en compte les capacités réactives de la cible et l'aspect markovien réduit l'ordre de complexité de la représentation. Aussi, une autre approche consiste à modéliser l'ensemble des trajectoires possibles de la cible (*a priori* ensembliste. Ces considérations conduisent à des jeux de recherche *a priori* sur des ensembles de trajectoires, [40]. De nombreux problèmes de ce type peuvent se ramener à des problèmes de programmation linéaire. La simplicité de ce type d'approches a cependant deux inconvénients majeurs : l'usage de fonctions de détection locales de type exponentiel est souvent nécessaire, ce qui est une approximation, et l'ensemble des stratégies doit être de taille limitée (description exhaustive). Dans un premier temps, il s'est agi de développer des méthodes originales pour des jeux de recherche ; d'abord pour une cible statique, puis pour une cible mobile, en présence de contraintes sur la distribution de celle-ci (optimisation minimax des stratégies de recherche). Afin de rendre les problèmes faisables en terme de complexité, une décomposition semi-markovienne de l'*a priori* ensembliste a été défini, garantissant malgré tout une complexité suffisante de la modélisation. Une résolution

[Koo99] B. KOOPMAN, *Search and screening. General principles with historical applications*, MORS, Heritage Series, 1999.

directe du jeu s'étant avérée inadaptée, cette difficulté a été contournée par approximation du jeu par une séquence d'optimisation unilatérales de type Brown.

Dans un cadre opérationnel réel, il apparaît raisonnable de faire l'hypothèse (pessimiste) que l'objet à détecter pourra se comporter au cours de la phase de recherche de manière à gêner le plus possible sa détection. Le fait d'envisager une certaine réactivité de la cible nous a alors amené à étudier le problème d'optimisation de l'effort de recherche d'un point de vue dynamique ; c'est à dire en prenant en compte l'évolution des informations contextuelles tout au long du processus de recherche. La grande difficulté de ce type de problème tient à l'objectif paradoxal, qui consiste à établir un plan de recherche *a priori* optimal pour des situations de recherche fondamentalement dynamiques. Ceci peut se faire par la prise en compte d'informations contextuelles (à modéliser) à l'intérieur même du problème d'optimisation. Ces informations de contexte peuvent être par exemple la connaissance possible d'informations sur la trajectoire courante de la cible, ou d'informations sur la stratégie effective de l'observateur.

7 Contrats industriels (nationaux, européens et internationaux)

7.1 Convention Cemagref : Évaluation de la qualité de pièces de viande par vision active

Participants : Eric Marchand, François Chaumette.

no. Inria 1 00 C 0494, durée : 36 mois.

Cette étude porte sur l'évaluation de la qualité de pièces de viande de porc par vision active et mesures IRM. Elle fait suite à une soumission conjointe du Cemagref Rennes, de la société Olympig et du projet Vista auprès de l'Ofival qui apporte un soutien financier. Les travaux ont porté sur l'élaboration de stratégies de perception et d'action afin de déterminer le volume englobant des pièces de porc considérées et de détecter d'éventuels défauts visibles sur leur surface. Pour cela, nous avons utilisé une caméra embarquée sur l'effecteur d'un robot manipulateur. Ces travaux ont été validés sur la cellule de vision robotique du projet.

7.2 Convention LRBA-DGA : projet Posit

Participants : Jean-Pierre Le Cadre, Sébastien Paris.

no. Inria 1 00 C 0607, durée 15 mois, avenant 9 mois

Les partenaires de ce projet sont la société CRIL-Technology et l'Onera. L'objectif global est d'étudier et de développer un outil d'aide à l'élaboration d'une trajectoire optimisée, afin d'améliorer la précision de navigation. Cette tâche est particulièrement importante lorsque l'on dispose d'observations extérieures au système de navigation lui-même. Il faut noter que cette trajectoire optimisée doit se trouver à l'intérieur d'un domaine de contraintes : contraintes de manœuvres, contraintes opérationnelles, positions initiales et finales, contraintes de temps. L'étude réalisée est décrite au paragraphe 6.3.5.

7.3 Convention Cifre Thales Systèmes Aéroportés : Traitement d'évaluation de la situation tactique

Participant : Jean-Pierre Le Cadre.

durée 36 mois.

Cette convention correspond à l'encadrement de la thèse de F. Bavencoff à TSA et une collaboration avec J.-M. Vanpeperstraete de TSA. Le système embarqué d'un avion de patrouille maritime doit fusionner et intégrer une quantité importante de données fournies par des capteurs de différentes natures (radar, ESM, optronique,...) pour proposer une évaluation de la situation la plus précise possible. Il doit être en mesure de traiter des scènes contenant un très grand nombre de cibles (plusieurs centaines). Dans ce but, l'orientation générale des travaux se trouve dans le cadre des problèmes d'association piste-à-piste, piste-à-mesure, ainsi que sur l'apport des données d'identification. Pour ce faire, l'approche retenue consiste à utiliser conjointement les outils classiques des statistiques (lois des estimateurs, tests d'hypothèses,...) d'une part, et l'optimisation combinatoire d'autre part (algorithmes de Munkres, des enchères). Pour cette année, les efforts ont porté sur l'évaluation des performances en association multiscan. Plus précisément, le calcul de la probabilité d'association correcte a été réexaminé dans le cadre de l'approche de Mori et al.^[MCC92]. Les hypothèses statistiques de ce calcul ont été aussi étendues ou justifiées. Un autre volet de l'activité a été constitué par l'étude de l'utilisation de techniques d'estimation statistique du type tirage MCMC (Monte Carlo Markov Chain) dans ce même contexte.

7.4 Convention DGA : Suivi d'amers temps-réel

Participant : Frédéric Cao.

no. Inria 1 02 C 0559, durée 9 mois.

Il s'agit d'une activité de conseil pour l'implantation d'un algorithme de suivi d'amers en temps réel, utilisant une méthode de recalage des ensembles de niveau d'une image.

7.5 Convention DGA - ENS Cachan : Détection de trajectoires

Participant : Frédéric Cao.

durée 6 mois.

Le but de cette étude est de montrer qu'on peut décider de la présence d'une trajectoire dans un environnement très bruité avec des seuils de détection fixés de manière entièrement automatique (voir paragraphe 6.1.4).

[MCC92] S. MORI, K. CHANG, C. CHONG, « Performance analysis of optimal data association with applications to multiple target tracking », *in: Multitarget-Multisensor Tracking: Applications and Advances, vol. II*, Artech House, 1992, ch. 7, p. 183-234.

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

Participants : Patrick Bouthemy, Emmanuel Veneau.

no. Inria 1 99 C 222, durée 36 mois.

Cette convention concernait la thèse de Emmanuel Veneau, encadrée à l'Ina par Laurent Vinet, thèse débutée en mars 1999 et soutenue en février 2002, [13]. Cette étude s'est inscrite dans le domaine de l'indexation de vidéos par le contenu. Elle concernait plus précisément 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 avait porté sur une structuration temporelle, de niveau intermédiaire, de la vidéo, le regroupement des plans élémentaires en séquences représentant des entités plus pertinentes et plus exploitables pour la phase de reconnaissance d'événements spatio-temporels et plus généralement pour l'annotation effectuée par les documentalistes. Une méthode combinant plusieurs critères, couleur, durée, mouvement, a été élaborée. Nous nous sommes ensuite intéressés à la caractérisation de l'activité présente dans les plans ou les séquences extraites. Pour cela, nous avons considéré plusieurs approches s'appuyant sur la construction de cartes d'"historiques" issues de mesures spatio-temporelles locales (différences temporelles d'images après compensation du mouvement de la caméra d'une part, sorties de filtres spatio-temporels de Gabor d'autre part), et sur un schéma de classification hiérarchique par SVM ("Support Vector Machines"). Des évaluations expérimentales ont été menées sur des documents vidéos de retransmission d'événements sportifs, qui forment une étude de cas représentative et complexe.

7.7 Actions financées par la Commission Européenne

7.7.1 Projet IST Carsense

Participants : Eric Marchand, Patrick Bouthemy.

no. Inria 1 00 B 0152, durée 36 mois

Les partenaires du projet européen IST Carsense sont les sociétés Autocruise Ltd (GB), BMW AG (All.), CRF Societa Consortile per Azioni (It.), Thomson-CSF Detexis (Fr.), Jena-Optronik (All.), Renault Recherche Innovation (Fr.), IBEO Lasertechnik Hipp KG (All.), TRW Automotive (GB), l'ENSMP, le Livic (LCPC), l'Inrets-Leost et l'Inria (M. Parent, projets Sharp et Vista). L'objectif de ce projet est de concevoir un système multi-capteur d'assistance à la conduite automobile à basse vitesse. Ce système doit utiliser des capteurs de différents types (caméras vidéo, laser et radar). Les informations extraites des signaux et images issus de ces capteurs sont ensuite fusionnées pour créer une carte de l'environnement du véhicule. Nous sommes intervenus dans le workpackage "Computer Vision" de ce projet où nous avons développé un module de détection d'obstacles par analyse spatio-temporelle d'une séquence d'images monoculaires. Le principe est de calculer le mouvement apparent induit dans l'image par le déplacement de la caméra pour une route dégagée de tout obstacle et assimilée localement à un plan, de le compenser et d'analyser la séquence d'images résultantes. La présence d'un élément sur la route, mobile ou non, se traduit par des erreurs de compensation. Outre la

localisation des obstacles potentiels, nous avons développé un module de suivi temporel de ces obstacles et nous calculons le temps-à-collision associé à ces obstacles, [47].

7.7.2 Projet IST Lava

Participants : Patrick Bouthemy, Frédéric Cao, Nathalie Peyrard.

no. Inria 1 02 G 0424, durée 36 mois

Le projet IST Lava (“Learning for Adaptable Visual Assistants”) a débuté en mai 2002 et réunit les partenaires suivants : XRCE (“prime” du projet), l’Idiap (Suisse), l’Inria (projets Movi de l’Inria Rhône-Alpes et projet Vista), le RHUL (G.-B.), l’Université de Graz (Autriche), l’Université de Lund (Suède), et l’ANU (Australie). Il réunit des équipes en vision par ordinateur, en apprentissage statistique (“Machine Learning”) et en sciences cognitives. L’objectif est d’étudier de manière couplée des questions d’apprentissage automatique et de catégorisation (en particulier avec des approches de type “méthodes par noyaux”) ainsi que de reconnaissance de scènes à partir d’images, sur des aspects spatiaux (formes, objets) et dynamiques (mouvements, événements). Des applications liées à l’informatique diffuse avec capteurs image ou vidéo (“ambient intelligence scenarios”) sont visées. Nous intervenons sur les aspects de détection de mouvement, de catégorisation et de reconnaissance de mouvements et d’événements particuliers. Des premiers éléments sont décrits aux paragraphes 6.1.4 et 6.1.5.

7.8 Actions financées par le RNTL

7.8.1 Projet Domus Videum

Participants : Patrick Bouthemy, Nathalie Peyrard.

no. Inria 2 02 C 0100, durée 30 mois

Le projet Domus Videum est un projet relevant du programme RNTL. Ce projet regroupe Thomson-Multimedia, l’Ina, l’Ircsyn, l’Irin, l’Irisa (projets Metiss, Temics et Vista) et le SFRS, et concerne la conception et à la réalisation d’une plate-forme d’indexation (vidéo et son) et d’exploitation de documents audio-visuels dans un contexte domestique. Notre contribution porte sur la création de résumés visuels courts. Dans un avenir proche, les décodeurs numériques associés aux téléviseurs seront munis de capacités très importantes de stockage (disques). La fonctionnalité de résumé vidéo permettra d’offrir à l’utilisateur un aperçu rapide du contenu des enregistrements réalisés, exploitable par exemple dans une phase de pré-sélection des documents susceptibles de l’intéresser. Un descriptif des travaux réalisés cette année sur la segmentation temporelle de la vidéo selon son contenu dynamique est fourni au paragraphe 6.1.5. Nous avons également travaillé en collaboration avec Guillaume Gravier (projet Metiss) à une coopération multi-modale entre image et son pour la sélection des plages pertinentes de la vidéo. L’objectif est d’extraire les plages pertinentes du document vidéo à partir des informations issues non seulement de la bande vidéo du document, mais aussi de la bande audio. Nous avons réalisé un travail préliminaire afin d’étudier différentes possibilités de fusion de l’information audio correspondant à une segmentation temporelle (étiquetage) du signal audio en classes pré-définies (parole, musique, applaudissements, bruit), et de l’information sur le

contenu dynamique de la vidéo exprimée par des modèles probabilistes gibbsiens de mouvement par plage vidéo préalablement segmentée. Nous avons constaté que les segmentations correspondant aux deux vérités-terrain, respectivement audio et contenu dynamique, sont en fait peu corrélées. En conséquence, nous comptons intégrer l'information audio non pas au niveau de la création de la segmentation temporelle de la vidéo (nous utiliserons pour cela uniquement une information de mouvement), mais au niveau de la décision de la sélection d'un segment pour constituer le résumé vidéo. Dans cette approche, un segment est alors associé à une classe (ce qui implique une phase préalable d'apprentissage) par maximisation de la loi des observations de mouvement conditionnellement au label audio affiché sur ce segment.

7.9 Actions soutenues par le programme Priamm

7.9.1 Projet Mediaworks

Participants : Patrick Bouthemy, Chafik Kermad, Fabien Spindler.

no. Inria 2 00 A 0408, durée 36 mois

Le projet Mediaworks est un projet labellisé par le programme Priamm et financé par le Ministère de l'industrie. Ce projet a débuté en septembre 2000. Il regroupe TF1, le Limsi, la société Aegis, l'Inria (projets Imedia de l'Inria Rocquencourt et Vista), et traite de systèmes d'aide à l'indexation pour des documentalistes. Les aspects principaux de ce projet en sont la coopération entre les media texte et image, et la mise au point d'un moteur de recherche sémantique. Pour ce qui nous concerne, le travail consiste à fournir des outils de structuration automatique en plans, de typage et de représentation iconique de ces plans. Nous collaborons avec Patrick Gros (projet TexMex) sur cette étude. Nous avons développé une technique temps-réel de découpage en plans pour des changements de plans abrupts ("cuts") exploitant directement les informations contenues dans le flux compressé MPEG (type des macro-blocs et coefficients de DCT). Une extension à la prise en compte des transitions progressives est en cours. Une représentation iconique des plans a été réalisée qui consiste à sélectionner des images du plan en fonction du mouvement de la caméra, calculé directement à partir des vecteurs MPEG. Le typage du plan selon le mouvement de la caméra est simultanément effectué en temps réel. Nous allons d'autre part développé une technique de suivi de visages dans un plan (visages initialement détectés dans la première image du plan par une méthode définie par le projet Imedia), combinant l'estimation d'un modèle paramétrique du mouvement et la mise en correspondance de primitives particulières. Cela permettra de propager aisément la détection des visages sur tout le plan et d'en faciliter la reconnaissance.

8 Actions régionales, nationales et internationales

8.1 Actions régionales

8.1.1 Contrat Région Bretagne : Création automatique de résumés de vidéos pour le multimédia domestique et professionnel

Participants : Gwénaëlle Piriou, Patrick Bouthemy.

no. Inria 4 01 C 0576, durée 36 mois.

Ce contrat permet de financer la bourse Inria de Gwénaëlle Piriou pour sa thèse qui a débuté en novembre 2001 et dont le descriptif est donné au paragraphe 6.1.2.

8.1.2 Contrat Région Bretagne : Méthodes de traitement et de reconstruction d'images échographiques tri-dimensionnelles

Participants : Arnaud Ogier, Christian Barillot.

no. Inria 4 02 C 0527, durée 36 mois.

Ce contrat permet de financer la bourse Inria de Arnaud Ogier pour sa thèse qui a débuté en novembre 2002 et dont le descriptif est donné au paragraphe 6.1.6.

8.1.3 Contrat CPER Cemagref-Inria AEROBIO

Participants : Etienne Mémin, Patrick Bouthemy, Thomas Corpetti.

Dans le cadre du Contrat de Plan État-Région, un financement pluri-annuel a été attribué sur la période 2000-2003 au projet Cemagref-Inria Rennes dénommé AEROBIO. Ce projet vise à renforcer les recherches menées en matière de lutte contre la contamination des aliments dans les ateliers de transformation alimentaire. Il présente deux volets complémentaires : *i*) renforcer les moyens expérimentaux du Cemagref Rennes par la création d'un hall aéraulique basse vitesse, équipé de matériels de mesure performants pour l'étude de la turbulence et des instabilités, parmi lesquels des équipements de visualisation d'écoulements sous plans laser; *ii*) permettre un programme de recherche commun entre le Cemagref et l'Irisa sur l'étude des écoulements de fluides par l'analyse d'images numériques (voir paragraphe 6.1.1). Le financement du CPER ne porte que sur l'équipement.

8.1.4 Télé-acquisition robotisée d'échographies 3D

Participants : Christian Barillot, François Chaumette, Fabien Spindler.

Une subvention de la Région a été obtenue pour contribuer à l'achat d'un équipement d'acquisition robotisée d'échographies 3D. Ce projet a pour but de développer un système d'acquisition d'échographies 3D piloté à distance à l'aide d'images vidéos et du contenu même des images échographiques. Les applications visées par ce système sont de deux types : d'une part, l'étude de pathologies vasculaires périphériques (plaques d'athéromes carotidiennes, thromboses veineuses), et d'autre part, le diagnostic et le suivi thérapeutique de pathologies en obstétrique

et chez le nourrisson. L'intérêt de ce système sera de permettre à un médecin, expert d'une pathologie, de réaliser lui-même, et à distance, l'examen échographique, et donc de pouvoir l'interpréter dans les meilleures conditions sans pour autant avoir à se déplacer, ni faire se déplacer le patient. L'apport de la troisième dimension dans l'examen échographique viendra renforcer à la fois la qualité du diagnostic fait et surtout les possibilités d'accès plus précis à des informations quantitatives. Les originalités du projet résident : *i*) dans sa capacité à fournir des échographies 3D sans pour autant recourir à des sondes spécifiques tout en optimisant la procédure, aujourd'hui classique, d'acquisition de type "main libre" par la programmation et la réalisation robotisée de la trajectoire de la sonde; *ii*) dans l'absence de recours à un robot maître (système haptique) par la mise au point d'un système de commande par asservissement visuel utilisant à la fois l'image vidéo de la salle et du champ opératoire et les images ultrasonores acquises. Ce projet se fera en collaboration avec deux centres d'expérimentations cliniques, les CHR de Tours (Drs Pourcelot et Tranquart) et de Rennes (Drs Darnault et Tréguier).

8.2 Actions nationales

8.2.1 ACI Ministère de la Recherche : projet NeuroBase

Participants : Christian Barillot, Isabelle Corouge.

no. CNRS 2P 2008, durée 24 mois.

Deux des préoccupations majeures des chercheurs et des cliniciens impliqués dans des expérimentations en neuro-imagerie sont la gestion de la quantité immense de données produites et la confrontation de leurs expériences (entre personnes de différents centres, ou vis-à-vis de publications). De plus, et particulièrement dans des centres de taille moyenne voire petite, les chercheurs ou les cliniciens n'ont pas la possibilité de réaliser des expérimentations d'envergure suffisante faute à la fois de moyens humains et de capacités de recrutement de sujets. En outre, la validité statistique des résultats est parfois insuffisante. Enfin, s'ajoute le souci d'assurer au mieux la rentabilité des équipements existants, coûteux à l'achat comme à l'utilisation, en facilitant leur accès à une plus large communauté d'utilisateurs. Pour toutes ces raisons, la mise en commun, entre centres, des résultats d'expérimentations, à travers un réseau, permettra d'étendre la portée scientifique des études menées et permettra également d'impliquer un public plus large dans des études de neuro-imagerie. Cela nécessite que les utilisateurs puissent diffuser, échanger ou accéder à des informations de neuro-imagerie avec des moyens d'accès appropriés, à la fois précis dans la recherche de l'information et souples dans la manière de constituer ou d'interroger ces bases. Dans ce contexte, l'action concertée incitative (ACI) "Neurobase" (<http://www.irisa.fr/vista/neurobase>) a pour objectif d'établir les conditions permettant de fédérer au travers de l'Internet des bases d'informations en neuro-imagerie, situées dans différents centres d'expérimentation, cliniques neurologiques ou centres de recherche en neurosciences cognitives. Il s'agit de spécifier une architecture informatique permettant l'accès et le partage de résultats d'expérimentations ou bien encore de méthodes de traitement des données au sein d'un même site ou entre sites différents. Cela permettra par exemple la recherche de résultats similaires, la recherche d'images contenant des singularités, ou encore des recherches transversales de type "fouille de données" pour mettre en évidence d'éventuelles régularités.

Ce projet, référencé sous le numéro TS 2001/23, a pour partenaires les projets Vista (responsable de l'ACI) et Texmex de l'Irisa, l'IFR 49 "Neuroimagerie Fonctionnelle" (CEA SHFJ, Inserm U494, CHR Pitié Salpêtrière, Paris), le projet Caravel de l'Inria Rocquencourt et le projet Epidaure de l'Inria Sophia-Antipolis, l'Unité Inserm U438 (Grenoble), l'équipe IDM de la Faculté de Médecine de Rennes 1, le laboratoire TIMC (Grenoble).

8.2.2 Actions de recherche coopératives de l'Inria

– Imagerie cérébrale

Participants : Christian Barillot, Pierre Hellier, Isabelle Corouge.

Nous avons participé à deux actions de recherche coopérative entre l'Inria (projets Epidaure, Robotvis et Vista) et l'IFR de neuro-imagerie de Paris-Orsay (CEA-SHJF Orsay, CHU la Pitié Salpêtrière). Ces actions, l'une pilotée par le projet Robotvis, ARC MC2 ("Mieux Connaître le Cerveau"), et l'autre par le projet Epidaure, ont eu pour objectif de développer de nouvelles méthodes de traitement d'image et de recalage en imagerie fonctionnelle cérébrale. L'une a été principalement axée sur la MEG (Magnéto-Encéphalographie), l'autre sur l'IRM fonctionnelle. Notre contribution dans ces deux actions a porté sur le développement de méthodes de recalage non rigide en imagerie cérébrale, appliquées d'une part à la correction des distorsions de phase en IRM f et d'autre part au recalage inter-sujets avec contraintes anatomiques (par les sillons corticaux). Nous avons également contribué à la définition d'un cadre méthodologique de validation des techniques de recalage en imagerie cérébrale, pour la correction du bougé du sujet lors d'une acquisition EPI en IRM f (recalage rigide) et pour la validation du recalage non rigide entre images IRM anatomiques de sujets différents (cf. paragraphe 6.2.5). Ces actions se sont conclues en mars 2002.

8.2.3 Actions MathStic du CNRS

– Action : Mesure du mouvement

Participants : Patrick Bouthemy, Frédéric Cao, Étienne Mémin, Gwénaëlle Piriou, Nathalie Peyrard, Thomas Corpetti.

durée 12 mois

Cette action, dont nous sommes responsables, est menée en collaboration avec l'Irmar (Bernard Delyon, Jian-Feng Yao), et s'intitule précisément "Modélisations mathématiques pour l'analyse de séquences d'images : mesure et reconnaissance du mouvement". Elle correspond à des éléments décrits dans les paragraphes 6.1.1 et 6.1.2.

– Action : Chaînes de Markov cachées et filtrage particulière

Participants : Jean-Pierre Le Cadre, Etienne Mémin, Elise Arnaud, Thomas Bréhard, Carine Hue.

Les méthodes particulières forment une classe de techniques de simulation permettant de traiter les questions d'inférence récursive (filtrage, prédiction, détection de ruptures)

pour les modèles d'état non linéaires et non gaussiens. Les enjeux de cette action spécifique dont les coordinateurs sont O. Cappé (ENST Paris) et F. Le Gland (projet Sigma-2) et qui comprend une dizaine d'équipes, sont à la fois méthodologiques et autour de développements applicatifs (télécommunications, vision, finance, . . .). Notre contribution dans cette action porte essentiellement sur des thèmes liés à la poursuite d'objets ou d'entités mobiles en sonar/radar et image (voir paragraphes 6.2.1, 6.2.2, 6.2.3). Deux journées consacrées aux divers aspects du filtrage particulière ont été organisées en décembre 2002. Nous participons également à l'action spécifique du département STIC intitulée "Méthodes particulières" (AS-67).

8.2.4 AS du département STIC du CNRS

- Action spécifique : Réalité augmentée temps-réel

Participants : Christian Barillot, François Chaumette, Pierre Hellier, Eric Marchand.

Cette action spécifique du CNRS est pilotée par Marie-Odile Berger du Loria et a débuté en novembre 2002. La réalité augmentée (RA) est un concept relativement récent dont l'objectif est d'augmenter la perception visuelle par des éléments virtuels ajoutés de manière réaliste au monde réel. L'objectif est d'obtenir un environnement interactif faisant coopérer objets virtuels et objets réels dans le même espace. Les applications potentielles de ce concept sont très nombreuses et ambitieuses : commerce électronique, héritage culturel (visualiser un site tel qu'il était au temps de sa splendeur), imagerie médicale (visualisation sur le patient des informations extraites de différentes imageries), systèmes réalistes d'apprentissage, . . . En dehors des problèmes d'ordre technologique (ordinateurs portables, systèmes de visualisation adaptés), ces nouvelles applications posent des problèmes méthodologiques difficiles et intéressants, car les environnements concernés sont souvent vastes, non structurés, et que les opérations de recalage des mondes réels et virtuels doivent être réalisées en temps réel. Le rôle de la vision par ordinateur dans les applications de RA est important, et concerne plus particulièrement la détection, le suivi et le calcul de pose en temps réel.

8.2.5 Programme Robea du CNRS

- Projet Omnibot

Participant : François Chaumette.

Ce projet de deux ans, acté en octobre 2002 et financé dans le cadre du programme ROBEA sur la ROBotique et les Entités Artificielles, correspond à une collaboration entre le LASMEA à Clermont-Ferrand, le LIRMM à Montpellier, le CREA à Amiens et le projet Vista. Il va s'agir d'étudier l'apport de la vision omnidirectionnelle pour la commande de robots mobiles.

- Projet Interprétation de scènes dynamiques complexes et planification réactive de mouvement
Participant : Eric Marchand.

Ce projet de trois ans, acté en octobre 2002, porte sur une collaboration entre l'In-

ria Rhône-Alpes (projets Movi et Sharp), le laboratoire Gravir (équipe Prima), le LAAS (groupe RIA) et le projet Vista. L'objectif de ce projet est l'automatisation de la conduite d'un véhicule en environnement de type routier. Deux problématiques scientifiques difficiles sont abordées : l'interprétation de scènes dynamiques complexes à l'aide de caméras et la planification réactive de mouvement en environnement dynamique. Le véhicule Cy-cab dont nous disposons (voir paragraphe 5.10) servira de base expérimentale.

8.3 Réseaux et groupes de travail internationaux

8.3.1 Groupe de travail Ercim

Participants : Patrick Bouthemy, Frédéric Cao.

Nous participons au groupe de travail provisoirement intitulé "Image Understanding" d'Ercim qui s'est créé cette année sous la responsabilité de Eric Pauwels du CWI (participation et présentation de nos travaux aux réunions qui se sont tenues à Amsterdam et Vienne).

8.3.2 Groupe de travail Delos-NSF

Participant : Patrick Bouthemy.

P. Bouthemy a été invité à participer au groupe de travail "Digital Imagery for Significant Cultural and Historical Materials" qui s'est constitué sous la responsabilité de Alberto Del Bimbo (Univ. Florence) et qui relève du réseau DELOS et de la NSF. Cela s'est notamment traduit par une présentation intitulée "Towards extraction of meaningful temporal video segments" à la session spéciale "Multimedia Content-based Retrieval for Cultural Heritage Applications" organisée au sein du workshop MIR'2002 (4th Int. Workshop on "Multimedia Information Retrieval") en décembre 2002 à Juan-les-Pins, [56].

8.4 Relations bilatérales internationales

8.4.1 Programme Procope France-Allemagne

Participants : Etienne Mémin, Patrick Bouthemy, Frédéric Cao, Elise Arnaud.

Nous menons, dans le cadre du programme d'actions intégrées franco-allemand Procope, une coopération avec l'équipe de C. Schnörr de l'université de Mannheim. Cette coopération porte sur l'estimation du mouvement de structures déformables, et plus particulièrement sur la comparaison et la combinaison d'approches variationnelles et statistiques, correspondant aux expertises respectives des deux équipes. Elle s'est traduite cette année par une étude visant à estimer directement à partir de séquences d'images les fonctions de potentiel associées au champ des vitesses d'un écoulement fluide (potentiel de vitesse et fonction de courant). E. Arnaud a effectué un séjour de deux semaines à Mannheim, F. Cao et E. Mémin une visite de quelques jours.

8.4.2 Programme Van Gogh France-Pays-Bas

Participants : Pierre Hellier, Christian Barillot, François Rousseau.

Nous avons monté cette année, dans la continuité du séjour postdoctoral de Pierre Hellier, une collaboration avec l'Université d'Utrecht portant sur l'échographie tridimensionnelle en neurochirurgie. L'équipe "Image Sciences Institute" possède une renommée internationale dans le domaine de l'imagerie médicale. Cette collaboration a donné lieu à une visite de quelques jours de C. Barillot, P. Hellier et F. Rousseau à Utrecht (voir paragraphe 6.2.6).

8.4.3 Accueil de chercheurs étrangers

- Prof. Bruno Cernuschi-Frias, Université de Buenos-Aires, a effectué un séjour de cinq mois dans le projet de septembre 2001 à janvier 2002, sur un poste d'accueil Inria de spécialiste académique étranger (cf paragraphe 6.2.2).
- Jordi Pages, doctorant de l'Université de Girona (Catalogne, Espagne) effectue un séjour de trois mois au sein du projet qui a débuté début novembre 2002. Ses travaux concernent la modélisation d'informations visuelles obtenues à partir de lumière structurée pour l'asservissement visuel.
- Prof. Christoph Schnörr (Université de Mannheim, Allemagne) a effectué une visite de quelques jours en octobre, et ses deux étudiants en thèse, Timo Kolhberger et Alexander Dejon, un séjour respectivement de trois et deux semaines en septembre-octobre (cf paragraphe 8.4.1).
- Prof. Wiro Niessen du ISI de l'Université d'Utrecht et son étudiante en thèse Marloes Letteboer ont effectué une visite de quelques jours dans le cadre de notre collaboration relevant du PAI Van Gogh.
- Luis Alvarez (Univ. Las Palmas), Ferran Marques (UPC Barcelone), Richard Prager (Univ. Cambridge) ont effectué des visites de courte durée au sein du projet.

8.4.4 Expertise de projets

- C. Barillot est expert de la Commission Européenne pour l'évaluation des projets de recherche du 5ème programme cadre du thème "Quality of life and management of living resources : research and technological development activities of a generic nature" ("Mid-term reviewer").
- C. Barillot est membre de l'"Advisory Board" du projet nord-américain "Human Brain Mapping" intitulé « A probabilistic referential system for the human brain » (resp. J.C. Mazziotta), renouvelé pour la période 2002-2005, et associant l'université UCLA (A.W. Toga, J.C. Mazziotta), l'université du Texas à Austin (P. Fox) et l'Institut Neurologique de Montréal à l'université McGill (A. Evans), et est consultant/collaborateur sur les aspects de segmentation et de modèles de déformation.
- P. Bouthemy a été membre du Comité d'évaluation et d'orientation du département DTIM de l'Onera qui a évalué les activités en traitement d'images du DTIM en mars 2002.
- F. Spindler est membre de la commission d'évaluation des ingénieurs de l'Inra sur la période 2001-2002.

9 Diffusion de résultats

9.1 Animation de la communauté scientifique

- C. Barillot anime un groupe de réflexion sur la neuro-imagerie qui a pour but de fédérer, sur Rennes et sa région, les personnes et les groupes de recherche qui s'intéressent au domaine de la neuro-imagerie clinique et cognitive. Ce groupe comprend à la fois des chercheurs en STIC et en sciences de la vie (CNRS, Inria, Inserm, Université de Rennes), des médecins du CHU de Rennes (neurologues, neurochirurgiens, radiologues, neurophysiologistes, pharmacologistes, ...) et des chercheurs en sciences de l'homme et de la société (linguistes et cognitivistes). Des séminaires réguliers sont organisés.
- Participation au GDR ISIS "Information, signal et images" : le projet est membre du GDR ISIS.
- J.-P. Le Cadre est membre du groupe de travail "Performances ultimes en trajectographie" de Thales, et d'un groupe d'évaluation horizontal de la DGA. J.-P. Le Cadre a co-organisé avec François Le Gland du projet Sigma-2 des journées thématiques du GDR-ISIS sur le filtrage particulaire qui se sont tenues à Paris en décembre 2002.
- **Comités éditoriaux de revues**
 - C. Barillot est membre du comité de rédaction de la revue *Technique et Science Informatiques* (TSI).
 - P. Bouthemy est éditeur associé de la revue *IEEE Trans. on Image Processing*.
 - F. Chaumette est éditeur associé de la revue *IEEE Trans. on Robotics and Automation*.
 - J.-P. Le Cadre est éditeur associé de la revue *Advances in Information Fusion* (ISIF).
- **Organisation de congrès internationaux**
 - C. Barillot est "General Chairman" du congrès MICCAI'2004 qui se tiendra à Saint-Malo en septembre 2004 et est organisé par le projet.
 - P. Bouthemy est président du comité d'organisation du workshop CBMI'2003 qui se tiendra à Rennes, à l'Irisa, en septembre 2003 (<http://www.irisa.fr/CBMI03>). F. Spindler est membre du comité d'organisation, chargé des aspects logiciels.
 - P. Bouthemy a été membre du comité d'organisation du workshop "Statistical Methods in Video Processing" (SMVP'2002), qui s'est tenu les 1er et 2 juin 2002 à Copenhague en conjonction avec ECCV'2002.
 - F. Chaumette a co-organisé un workshop sur l'asservissement visuel couplé au congrès international IROS'2002 (avec P. Martinet et N. Andreff du Lasmea, ainsi que E. Cervera de l'Université de Castille, Espagne).
- **Participation à des comités de programme**
 - C. Barillot : conférences internationales IEEE ISBI'2002, IPMI'2003, membre du comité directeur ("board") des congrès MICCAI et IPMI.
 - P. Bouthemy : conférences internationales ICPR'2002, SSIAT'2002, CVPR'2003, CBMI'2003, TAIMA'2003, congrès national RFIA'2002.
 - F. Cao : conférence internationale Scale-Space'2003.
 - F. Chaumette : conférences internationales ECCV'2002, IROS'2002, CVPR'2003. F. Chaumette a été l'un des quatre membres du jury de sélection du prix du meilleur papier

- étudiant à la conférence ICRA'2002.
- J.-P. Le Cadre : conférence internationale Fusion'2002, symposium OTAN "Target Tracking and Sensor Data Fusion for Military Observation Systems" 2003.
- P. Hellier : conférence internationale MICCAI'2002.
- E. Mémin : conférence nationale Orasis'2002.

9.2 Enseignement universitaire

- DEA Stir "Signal, Télécommunications, Image et Radar", université de Rennes 1, option Image (E. Mémin : approches markoviennes en imagerie, P. Bouthemy : analyse multi-images), option Signal (J.-P. Le Cadre : traitement d'antenne).
- DEA informatique, Ifsic, Rennes 1 (E. Marchand : modèles 3D pour la synthèse et la vision par ordinateur; E. Mémin : analyse du mouvement).
- DHC INC, Ifsic, Rennes 1 (E. Marchand, E. Mémin, F. Chaumette, T. Corpetti : vision 3D, analyse du mouvement, asservissement visuel; E. Mémin : modèles markoviens en analyse d'images; C. Barillot : imagerie médicale; E. Marchand, F. Spindler : méthode de programmation pour le traitement d'image).
- Insa Rennes, Génie électrique 5ème année, option Image (F. Chaumette, G. Piriou, F. Rousseau, O. Tahri : vision par ordinateur).
- DEA GBM universités Tours-Angers-Nantes-Rennes (C.Barillot : imagerie 3D appliquée à la médecine).
- DEA Traitement informatique pour l'anatomie et la morphométrie, Univ. Paris 5 et 7 (C.Barillot).
- DEA TIS, Université de Cergy, Ensea (P. Bouthemy : indexation vidéo).
- DEA PIC et ENSPS Strasbourg, 3ème année (P. Bouthemy : analyse et évaluation de séquences d'images).
- DEA MVA, ENS Cachan (F. Cao : filtrage d'images et équations aux dérivées partielles).

- Stages effectués au sein du projet : N. Bonnel (Insa Rennes et DEA Info Ifsic), T. Bréhard (ENSAI), A. Fardoun (ENSTBr), A. Kissi (DEA Génie Biomédical, Paris 12), J.-F. Marcos Pablo (Univ. Valladolid, Espagne, programme Erasmus), M. Pressigout (Insa Rennes, Génie Info, 4ème année).

- Direction de thèse hors projet:
 - A. Alhaj (Cemagref, Rennes) par F. Chaumette;
 - F. Bavencoff (Thales Airborne Systems, bourse Cifre) par J.-P. Le Cadre (voir paragraphe 7.3).

9.3 Participation à des colloques, séminaires, invitations

- P. Bouthemy a donné deux cours invités en avril 2002 à Karlsruhe, intitulés "Modélisation générique du mouvement à partir d'une séquence d'images : une approche non paramétrique et statistique" et "Applications de ce schéma à la reconnaissance, la classification et la segmentation du mouvement, à l'indexation et la recherche de vidéos sur des infor-

- mations de mouvement”, dans le cadre de l’Ecole doctorale virtuelle “Visuelle Kognition” co-organisée par l’Université de Karlsruhe (H.-H. Nagel) et l’ENSPS (E. Hirsch).
- P. Bouthemy a présenté une conférence invitée intitulée “Analyse et caractérisation de la vidéo” aux journées “Analyse et indexation multimédia” organisées au Labri à Bordeaux en juin 2002.
 - F. Cao a été invité à la conférence *Algorithm’2002* qui s’est tenue à Podbanske, Slovaquie, et a présenté un exposé intitulé “Application of Gestalt principles to the detection of good continuations, corners and terminators in image level lines”.
 - F. Cao a présenté un exposé intitulé “Détection invariante par changement de contraste et sans paramètre de bonnes continuations” au séminaire “Méthodes mathématiques du traitement d’images” du laboratoire Jacques-Louis Lions de Paris 6 (organiseurs A. Cohen et P. Combettes).
 - F. Cao a participé au *Cemracs’2002* (Centre d’été mathématique de recherche avancée en calcul scientifique) à travers un projet de recherche sur la mise en correspondance automatique de formes en collaboration avec Yann Gousseau (ENST Paris), Pablo Musé et Frédéric Sur (ENS Cachan). Cela s’est notamment traduit par un séjour de deux semaines au Centre International de Rencontres Mathématiques à Luminy en juillet-août 2002.
 - F. Chaumette a participé à l’Ecole d’été franco-mexicaine *SSIR’2002*, organisée à Toulouse en juillet 2002 et portant sur la vision et la robotique, intervention sur l’asservissement visuel.
 - F. Chaumette a participé en tant qu’intervenant à l’Ecole d’été sur l’asservissement visuel organisée à Benicassim, Espagne, en septembre 2002.
 - F. Chaumette a été invité pour un séjour de deux semaines en Australie en février 2002 à l’occasion d’un workshop et pour la visite des universités de New South Wales (Sydney), de Newcastle, et ANU (Canberra).
 - F. Chaumette a été invité à l’IST de Lisbonne (Portugal) en juillet 2002 en tant qu’expert sur le projet portugais *POCTI/EME/39946/2001* intitulé “Visual servoing control applied to robot manipulators”.
 - R. Fablet et P. Bouthemy ont obtenu le prix *AFRIF’2002* de la meilleure communication en Reconnaissance et Interprétation des Formes présentée au congrès *RFIA’2002*, [42].

10 Bibliographie

Ouvrages et articles de référence de l’équipe

- [1] P. BOUTHEMY, E. FRANÇOIS, « Motion segmentation and qualitative dynamic scene analysis from an image sequence », *Int. Journal of Computer Vision* 10, 2, avril 1993, p. 157–182.
- [2] 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.
- [3] 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, octobre 1997, p. 1242–1257.
- [4] 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.

-
- [5] P. HELLIER, C. BARILLOT, E. MÉMIN, P. PÉREZ, « Hierarchical estimation of a dense deformation field for 3D robust registration », *IEEE Transactions on Medical Imaging* 20, 5, 2001, p. 388–402.
 - [6] 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, janvier 1997, p. 178–201.
 - [7] 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.
 - [8] E. MARCHAND, F. CHAUMETTE, « An autonomous active vision system for complete and accurate 3D scene reconstruction », *Int. Journal of Computer Vision* 32, 3, août 1999, p. 171–194.
 - [9] 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.
 - [10] 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.

Livres et monographies

- [11] F. CAO, *Geometric Curve Evolution and Image Processing*, Lecture Notes in Mathematics, Springer Verlag, à paraître.

Thèses et habilitations à diriger des recherches

- [12] T. CORPETTI, *Estimation et analyse de champs denses de vitesses d'écoulements fluides*, thèse de doctorat, Université de Rennes 1, Mention Traitement du Signal, juillet 2002.
- [13] E. VENEAU, *Macro-segmentation multi-critère et classification de séquences par le contenu dynamique pour l'indexation vidéo*, thèse de doctorat, Université de Rennes 1, Mention Traitement du Signal, février 2002.

Articles et chapitres de livre

- [14] A. ALMANSA, F. CAO, Y. GOUSSEAU, B. ROUGÉ, « Interpolation of digital elevation models using AMLE and related methods », *IEEE Transactions on Geoscience and Remote Sensing* 40, 2, février 2002, p. 314–325.
- [15] F. CHAUMETTE, « Asservissement visuel », in : *La commande des robots manipulateurs*, W. Khalil (éditeur), *Traité IC2*, Hermès, 2002, ch. 3, p. 105–150.
- [16] C. COLLEWET, F. CHAUMETTE, « Positioning a camera with respect to planar objects of unknown shape by coupling 2D visual servoing and 3D estimations », *IEEE Trans. on Robotics and Automation* 18, 3, juin 2002, p. 322–333.
- [17] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Extraction of singular points from dense motion fields: an analytic approach », *Journal of Mathematical Imaging and Vision*, à paraître.
- [18] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Dense estimation of fluid flows », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 3, mars 2002, p. 365–380.
- [19] F. DAMBREVILLE, J.-P. LE CADRE, « Detection of a markovian target with optimization of the search efforts under generalized linear constraints », *Naval Research Logistics* 49, 2, février 2002, p. 117–142.

- [20] R. DONATI, J.-P. LE CADRE, « Detection of oceanic electric fields based on the GLRT », *IEE Proc. Radar, Sonar and Navigation* 149, 4, avril 2002, p. 221–230.
- [21] R. FABLET, P. BOUTHEMY, P. PÉREZ, « Non-parametric motion characterization using causal probabilistic models for video indexing and retrieval », *IEEE Trans. on Image Processing* 11, 4, avril 2002, p. 393–407.
- [22] G. FLANDIN, F. CHAUMETTE, « Fusion d'informations visuelles pour la localisation d'objets complexes », *Traitement du Signal* 19, 1, juin 2002, p. 49–57.
- [23] P. HELLIER, C. BARILLOT, « Coupling dense and landmark-based approaches for non rigid registration », *IEEE Transactions on Medical Imaging*, à paraître.
- [24] C. HUE, J.-P. LE CADRE, P. PÉREZ, « Sequential Monte Carlo methods for multiple target tracking and data fusion », *IEEE Trans. on Signal Processing* 50, 2, février 2002, p. 309–325.
- [25] C. HUE, J.-P. LE CADRE, P. PÉREZ, « Tracking multiple objects with particle filtering », *IEEE Trans. on Aerospace and Electronic Systems* 38, 3, juillet 2002, p. 791–812.
- [26] J.-P. LE CADRE, « Data association and multitarget tracking », *in: Multisensor Fusion*, A. Hyder, E. Shakhbazian, et E. Waltz (éditeurs), *NATO Sciences Series, II. Mathematics, Physics, Kluwer, 70*, NATO ASI on Multisensor Data Fusion, 2002, p. 331–351.
- [27] E. MALIS, F. CHAUMETTE, « Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods », *IEEE Trans. on Robotics and Automation* 18, 2, avril 2002, p. 176–186.
- [28] E. MARCHAND, F. CHAUMETTE, F. SPINDLER, M. PERRIER, « Controlling an uninstrumented manipulator by visual servoing », *The International Journal of Robotics Research, IJRR* 21, 7, juillet 2002, p. 635–648.
- [29] E. MARCHAND, N. COURTY, « Controlling a camera in a virtual environment. », *The Visual Computer Journal* 18, 1, février 2002, p. 1–19.
- [30] Y. MEZOUAR, F. CHAUMETTE, « Avoiding self-occlusions and preserving visibility by path planning in the image », *Robotics and Autonomous Systems* 41, 2-3, novembre 2002, p. 77–87.
- [31] Y. MEZOUAR, F. CHAUMETTE, « Path planning for robust image-based control », *IEEE Trans. on Robotics and Automation* 18, 4, août 2002, p. 534–549.
- [32] E. MÉMIN, P. PÉREZ, « Hierarchical estimation and segmentation of dense motion fields », *Int. Journal of Computer Vision* 46, 2, février 2002, p. 129–155.
- [33] L. OISEL, E. MÉMIN, L. MORIN, F. GALPIN, « 1D dense disparity estimation for 3D reconstruction », *IEEE Trans. on Image Processing*, à paraître.
- [34] C. PAPIN, P. BOUTHEMY, G. ROCHARD, « Unsupervised segmentation of low clouds from infrared METEOSAT images based on a contextual spatio-temporal labeling approach », *IEEE Trans. on Geoscience and Remote Sensing* 40, 1, janvier 2002, p. 104–114.

Communications à des congrès, colloques, etc.

- [35] C. BARILLOT, P. HELLIER, I. COROUGE, « Non rigid registration in neuroimaging: from a retrospective evaluation of global approaches to new advances in local and hybrid methods », *in: IEEE Int. Symp. on Biomedical Imaging*, Washington, D.C., sept. 2002.
- [36] F. CHAUMETTE, « A first step toward visual servoing using image moments », *in: IEEE/RSJ Int Conf on Intelligent Robots and Systems, IROS'2002*, 1, p. 378–383, Lausanne, octobre 2002.
- [37] I. COROUGE, C. BARILLOT, « Statistical modeling of pairs of sulci in the context of neuroimaging probabilistic atlas », *in: Conf. Medical Image Computing and Computer Assisted Intervention, MICCAI'2002, Vol. LNCS 2489*, p. 655–662, Tokyo, septembre 2002.

-
- [38] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Dense motion analysis in fluid imagery », *in: European Conf. on Computer Vision, ECCV'2002, LNCS 2350*, p. 676–691, Copenhague, mai 2002.
- [39] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Régularisation div-curl et équation de continuité pour l'estimation du mouvement fluide », *in: 13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, RFIA '2002, 3*, p. 887–898, Angers, janvier 2002.
- [40] F. DAMBREVILLE, J.-P. LE CADRE, « Search game for a moving target with dynamically generated informations », *in: Conf. Fusion'2002*, p. 243–250, Annapolis, juillet 2002.
- [41] R. FABLET, P. BOUTHEMY, « Extraction of regions of interest based on motion activity for video retrieval with partial queries », *in: 9th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU'2002*, Annecy, juillet 2002.
- [42] R. FABLET, P. BOUTHEMY, « Modélisation statistique non paramétrique et reconnaissance du mouvement », *in: 13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, RFIA '2002, 2*, p. 549–558, Angers, janvier 2002.
- [43] G. FLANDIN, F. CHAUMETTE, « Visual data fusion for objects localization by active vision », *in: European Conf. on Computer Vision, ECCV'2002, LNCS 2353*, p. 312–326, Copenhague, mai 2002.
- [44] P. HELLIER, J. ASHBURNER, I. COROUGE, C. BARILLOT, K. FRISTON, « Inter subject registration of functional and anatomical data using SPM », *in: Conf. Medical Image Computing and Computer Assisted Intervention, MICCAI'2002, LNCS 2489*, p. 590–587, Tokyo, septembre 2002.
- [45] C. HUE, J.-P. LE CADRE, P. PÉREZ, « Performance analysis of two sequential Monte Carlo methods and posterior Cramer-Rao bounds for multi-target tracking. », *in: 5th International Conference on Information Fusion, FUSION'2002*, p. 464–473, Annapolis, juillet 2002.
- [46] C. LACOSTE, R. FABLET, P. BOUTHEMY, Y.-F. YAO, « Création de résumés de vidéos par une approche statistique », *in: 13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, RFIA '2002, 1*, p. 153–162, Angers, janvier 2002.
- [47] G. LEFAIX, E. MARCHAND, P. BOUTHEMY, « Motion-based obstacle detection and tracking for car driving assistance », *in: IAPR Int. Conf. on Pattern Recognition, ICPR '2002, 4*, p. 74–77, Québec, août 2002.
- [48] M. LETTEBOER, P. WILLEMS, P. HELLIER, W. NIESSEN, « Acquisition of 3D ultrasound images during neuronavigation », *in: Proc. of Computer Assisted Radiology and Surgery*, H.-U. Lemke, M.-W. Vannier (éditeurs), Springer, Paris, juin 2002.
- [49] R. MAHONY, P. CORKE, F. CHAUMETTE, « Choice of image features for depth-axis control in image based visual servo control », *in: IEEE/R SJ Int Conf on Intelligent Robots and Systems, IROS'2002, 1*, p. 390–395, Lausanne, octobre 2002.
- [50] R. MAHONY, T. HAMEL, F. CHAUMETTE, « A decoupled image space approach to visual servo control of a robotic manipulator », *in: IEEE Int. Conf. on Robotics and Automation, ICRA '2002, 3*, p. 3781–3786, Washington DC, mai 2002.
- [51] E. MARCHAND, F. CHAUMETTE, « Virtual visual servoing: a framework for real-time augmented reality », *in: EUROGRAPHICS'2002 Conference Proc.*, G. Drettakis, H.-P. Seidel (éditeurs), *Computer Graphics Forum*, 21(3), p. 289–298, Sarrebruck, septembre 2002.
- [52] E. MARCHAND, J. ROYAN, F. CHAUMETTE, « Calcul de pose et calibration par asservissement visuel virtuel : application à la réalité augmentée », *in: 13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, RFIA '2002, 3*, p. 839–848, Angers, janvier 2002.

- [53] Y. MEZOUAR, A. REMAZEILLES, P. GROS, F. CHAUMETTE, « Image interpolation for image-based control under large displacement », *in: IEEE Int. Conf. on Robotics and Automation, ICRA'2002*, 3, p. 3787–3794, Washington DC, mai 2002.
- [54] S. PARIS, J.-P. LE CADRE, « Planification for terrain-aided navigation », *in: Conf. Fusion'2002*, p. 1007–1014, Annapolis, juillet 2002.
- [55] N. PEYRARD, P. BOUTHEMY, « Content-based video segmentation using statistical motion models », *in: British Machine Vision Conference, BMVC'2002*, 2, p. 527–536, Cardiff, septembre 2002.
- [56] N. PEYRARD, P. BOUTHEMY, « Towards extraction of meaningful temporal video segments », *in: 4th Int. Workshop on Multimedia Information Retrieval, MIR'2002*, Juan-les-Pins, décembre 2002.
- [57] P. PÉREZ, C. HUE, J. VERMAAK, M. GANGNET, « Color-based probabilistic tracking », *in: European Conf. on Computer Vision, ECCV'2002, LNCS 2350*, p. 661–675, Copenhagen, juin 2002.
- [58] F. ROUSSEAU, P. HELLIER, C. BARILLOT, « A fully automatic calibration procedure for freehand 3D ultrasound », *in: IEEE Int. Symp. on Biomedical Imaging*, Washington D.C., juillet 2002.

Rapports de recherche et publications internes

- [59] F. CAO, « Contrast invariant, parameterless detection of good continuations, corners and terminators », *rapport de recherche n°1487*, Irisa, octobre 2002, <ftp://www.irisa.fr/techreports/2002/PI-1487.ps.gz>.
- [60] F. CAO, « Sûreté de la détection et application à la détection de trajectoires », *rapport de recherche*, IRISA, étude pour la DGA, 2002.
- [61] I. COROUGE, P. HELLIER, B. GIBAUD, C. BARILLOT, « A local approach for inter-individual functional registration », *rapport de recherche n°4415*, Inria, mars 2002.
- [62] C. HUE, J.-P. LE CADRE, P. PÉREZ, « Performance analysis of two sequential Monte Carlo methods and posterior Cramer-Rao bounds for multi-target tracking », *rapport de recherche n°1457*, Irisa, avril 2002.
- [63] N. JAIN, E. MÉMIN, C. PÉREZ, « Parallelization of dense fluid motion estimation application using OpenMP », *rapport de recherche n°1483*, Irisa, septembre 2002, <ftp://ftp.irisa.fr/techreports/2002/PI-1483.ps.gz>.
- [64] I. PRATIKAKIS, C. BARILLOT, P. DARNAULT, « Towards free-hand 3-D ultrasound », *rapport de recherche n°4399*, Inria, mars 2002.
- [65] F. ROUSSEAU, C. BARILLOT, « Quality assessment of electromagnetic localizers in the context of 3D ultrasound », *rapport de recherche n°4408*, Inria, mars 2002, <http://www.inria.fr/rrrt/rr-4408.html>.