

Projet VISTA

Vision Spatio-Temporelle et Active

Rennes

THÈME 3B



*R*apport
d'Activité

2001

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	13
5.1	Logiciel RMR d'estimation du mouvement dominant	13
5.2	Logiciel MD-Shots de découpage d'une vidéo en plans	14
5.3	Logiciel D-Change de détection de mouvement	15
5.4	Logiciel Dense-Motion de mesure de champs de vitesses 2D	15
5.5	Logiciel ROMEO de recalage non linéaire en imagerie cérébrale	15
5.6	Plate-forme VISP d'asservissement visuel	16
5.7	Vistal : Bibliothèques pour le traitement et l'analyse d'images volumiques spatio-temporelles	16
5.8	Actions de développement : Plates-formes de vision robotique	17
5.9	Actions de développement : Véhicule Cycab	17
6	Résultats nouveaux	18
6.1	Modèles statistiques, métrologie du mouvement et des déformations, indexation	18
6.1.1	Modèles pour l'analyse du mouvement fluide	18
6.1.2	Super-résolution de séquences d'images	19
6.1.3	Modélisation non paramétrique et reconnaissance du mouvement	20
6.1.4	Modèles probabilistes de formes en imagerie médicale	21
6.1.5	Résumé et indexation vidéo	22
6.1.6	Indexation de données numériques	23
6.2	Suivi, trajectographie, recalage 3D	24
6.2.1	Suivi multi-objets en imagerie par filtrage particulaire et trajectographie	24
6.2.2	Suivi de structures caractéristiques dans des images de fluides	25
6.2.3	Méthodologie d'évaluation de méthodes de recalage non-rigide	26
6.2.4	Analyse d'images ultrasonores 3D	27
6.3	Asservissement visuel, perception active et exploration	28
6.3.1	Utilisation des moments en asservissement visuel	28
6.3.2	Asservissement visuel virtuel pour la réalité augmentée	29
6.3.3	Navigation réactive dans des espaces virtuels	29
6.3.4	Reconstruction 3D par vision dynamique active	30
6.3.5	Planification dans l'image	31
6.3.6	Couplage perception-action par indexation d'images	32

6.3.7	Théorie de l'exploration	32
7	Contrats industriels (nationaux, européens et internationaux)	33
7.1	Convention DGA - Thales Optronique : Conception d'une chaîne de traitements pour la restauration de séquences d'images	33
7.2	Convention Cemagref : Évaluation de la qualité de pièces de viande de porc par vision active	34
7.3	Convention Ifremer : Commande en vitesse d'un bras manipulateur non instrumenté par asservissement visuel	34
7.4	Convention LRBA, DGA : projet Posit	35
7.5	Convention Cifre Thales Airborne Systems : Méthodes et algorithmes de traitements multi-capteurs	35
7.6	Convention Onera :	36
7.7	Convention Cifre Ina : Reconnaissance de formes spatio-temporelles pour l'indexation vidéo	36
7.8	Actions financées par la Commission Européenne	37
7.8.1	Projet IST Carsense	37
7.9	Actions financées par le RNRT	37
7.9.1	Projet Agir	37
7.10	Actions financées par le RNTL	38
7.10.1	Projet Domus Videum	38
7.11	Actions soutenues par le programme Priamm	38
7.11.1	Projet Mediaworks	38
8	Actions régionales, nationales et internationales	39
8.1	Actions régionales	39
8.1.1	Contrat CPER Cemagref-Inria AEROBIO	39
8.1.2	Télé-acquisition robotisée d'échographies 3D	39
8.1.3	Contrat CRITT Santé : Environnement d'acquisition et de traitement d'échographies 3D	40
8.2	Actions nationales	40
8.2.1	ACI Ministère de la Recherche : projet NeuroBase	40
8.2.2	ACI jeunes chercheurs : Techniques d'exploitation des données multimédias	40
8.2.3	Actions de recherche coopératives de l'Inria	41
8.3	Relations bilatérales internationales	41
8.3.1	Programme Alliance France-Angleterre	41
8.3.2	Programme Procope France-Allemagne	41
8.3.3	Programme France-Tunisie	42
8.3.4	Accueil de chercheurs étrangers	42
8.3.5	Expertise de projets	42

9	Diffusion de résultats	43
9.1	Animation de la communauté scientifique	43
9.2	Enseignement universitaire	44
9.3	Participation à des colloques, séminaires, invitations	45
10	Bibliographie	46

1 Composition de l'équipe

Responsable scientifique

Patrick Bouthemy [DR Inria]

Assistante de projet

Édith Blin [TR Inria, avec équipe/projet Imadoc et Temics]

Personnel Inria

Frédéric Cao [CR à compter du 01/09/2001]

François Chaumette [DR]

Pierre Hellier [CR à compter du 01/09/2001]

Éric Marchand [CR]

Fabien Spindler [IR (Atelier)]

Personnel CNRS

Christian Barillot [CR, puis DR à compter du 01/10/2001]

Patrick Gros [CR]

Jean-Pierre Le Cadre [DR]

Personnel Université

Étienne Mémin [maître de conférences Ifsic, Université de Rennes 1, en délégation CNRS à compter du 01/09/2001]

Personnel non permanent

Abdeljabar Benzougar [ingénieur expert Inria, jusqu'au 28/02/2001]

Bruno Cernuschi-Frias [professeur invité Inria, Université de Buenos-Aires, à compter du 15/09/2001]

Nicolas Ducoin [ingénieur associé Inria, à compter du 01/10/2001]

Chafik Kermad [ingénieur expert Inria, à compter du 01/12/2001]

Gildas Lefaix [ingénieur expert Inria, jusqu'au 30/09/2001]

Sébastien Paris [ingénieur expert Inria, à compter du 15/05/2001]

Chercheur post-doctorant et ATER

Nathalie Peyrard [post-doc Inria, à compter du 01/12/2001]

Chercheurs doctorants

Elise Arnaud [allocataire MENRT, à compter du 01/10/2001]

Sid Ahmed Berrani [boursier Cifre TMM]

Andrew Comport [boursier Inria, à compter du 01/10/2001]

Isabelle Corouge [boursière Inria]

Thomas Corpetti [allocataire MENRT]

Frédéric Dambreville [boursier DGA, jusqu'au 30/11/2001]

Fabien Dekeyser [boursier DGA, jusqu'au 30/11/2001]

Ronan Fablet [boursier BDI CNRS-Région, jusqu'au 31/08/2001]

Grégory Flandin [boursier Inria-Région, jusqu'au 30/09/2001]

Carine Hue [allocataire MENRT]

Isabelle Leibowicz [Boursière Cifre, Thales-AS, jusqu'au 28/02/2001]

Youcef Mezouar [allocataire MENRT, jusqu'au 31/08/2001]

Gwénaëlle Piriou [boursière Inria, à compter du 01/11/2001]

Anthony Remazeilles [allocataire MENRT, à compter du 01/10/2001]

François Rousseau [allocataire MENRT]
Omar Tahri [boursier Inria-Région]
Emmanuel Veneau [boursier Cifre, Ina]

2 Présentation et objectifs généraux

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

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

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 mesures, référencées image ou scène, des besoins de reconnaissance et de décision sur des événements temporels ;
- la perception de systèmes automatisés ou robotiques, pour des tâches de surveillance et de détection, de guidage et de manipulation, de navigation et d'exploration.

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

Nous privilégions une approche statistique des questions d'analyse du mouvement et des déformations. Les aspects d'analyse spatio-temporelle long-terme sont particulièrement étudiés, notamment les questions de suivi temporel et de trajectographie. Nous abordons également les situations d'observateur actif, lorsque le capteur est commandable. Nous nous plaçons dans un cadre de contrôle en boucle fermée sur des informations visuelles (asservissement visuel), et à un plus haut niveau nous nous intéressons à des schémas de perception active. Trois

secteurs d'applications motivent principalement nos études : *"métérologie" du mouvement et des déformations* (imagerie météorologique, imagerie médicale, visualisation expérimentale en mécanique des fluides), *indexation de vidéos* par le contenu, *vision robotique et systèmes de détection*.

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

- **Modèles statistiques et analyse d'images spatio-temporelles.** Dans le cadre de l'analyse d'images spatio-temporelles, et du suivi dans le temps des primitives extraites, les modélisations statistiques à base de dépendances locales offrent un cadre méthodologique souple et riche, permettant d'aborder à l'aide d'outils variés de nombreux problèmes. Associés à l'estimation bayésienne, de tels modèles permettent de spécifier le lien entre les données et les primitives inconnues, et d'y adjoindre une connaissance a priori sur ces dernières prenant en compte le contexte local (spatial ou temporel). La spécificité des différents modèles d'analyse spatio-temporelle que nous développons, dépend de la nature des variables, continue, discrète, ou symbolique, et de leurs interactions, linéaires ou non, mais aussi de la structuration de ces interactions locales. Nous nous intéressons notamment dans ce contexte à des modèles markoviens hiérarchiques ainsi qu'à des estimateurs robustes.
- **Analyse du mouvement et trajectographie.** L'analyse du mouvement dans une séquence d'images doit être posée comme un problème joint d'estimation et de segmentation, puisqu'il s'agit d'appréhender des informations partiellement observables et discontinues. C'est un problème particulièrement difficile, mais dont la résolution est cruciale pour la plupart des tâches en analyse de scène dynamique. Ce problème peut se présenter en fait sous plusieurs variantes, suivant que l'objectif prioritaire se trouve être l'obtention d'une mesure dense ou paramétrique du mouvement, d'une partition de l'image en régions, ou l'extraction d'entités pertinentes. Un thème central et encore très ouvert sur ce sujet est l'analyse du mouvement déformable, qu'il soit articulé, élastique ou fluide. Nous étendons également ces approches au problème du recalage non linéaire entre images volumiques (dans le domaine médical). Un autre champ d'investigation motivant ces études sur la segmentation et la caractérisation du mouvement, ainsi que sur le suivi temporel, est représenté par la structuration et l'indexation de vidéo par le contenu. Enfin, sur les aspects de trajectographie, nous considérons des problèmes liés à l'estimation de systèmes dynamiques partiellement observés, ainsi qu'à la formalisation du suivi d'objets multiples par des méthodes de type PMHT et filtrage particulaire.
- **Asservissement visuel et perception active.** La perception active consiste à élaborer des stratégies de perception et d'action dans le but d'améliorer, par le contrôle des paramètres du capteur, les performances des algorithmes de vision ou de détection. Nous nous intéressons au contrôle de la position, de la vitesse, ou de la trajectoire de capteurs. Deux types de stratégies sont considérés et combinés : d'une part, des stratégies locales, liées aux techniques d'asservissement visuel et fournissant une perception de bas niveau, et d'autre part, des stratégies globales, activées sur des événements particuliers et remontant à une perception de plus haut niveau. Nous abordons le couplage d'une vision locale embarquée et d'une vision globale de la scène. Par ailleurs, nous étudions des questions liées à la théorie de l'exploration ("search theory").

Dans les trois secteurs d'applications mentionnés, nous avons eu récemment ou nous avons

actuellement des conventions de recherche ou de transfert avec Météo-France, Thales-Optronique, l'Ina, Thomson-Multimedia, EDF, l'Ifremer, le Cemagref, l'Onera, des projets multi-partenaires nationaux (programmes RNRT, RNTL, Priamm) et européens (projet Esprit R&D Divan, projet IST Carsense, contrat Eumetsat).

3 Fondements scientifiques

bayésienne

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.*

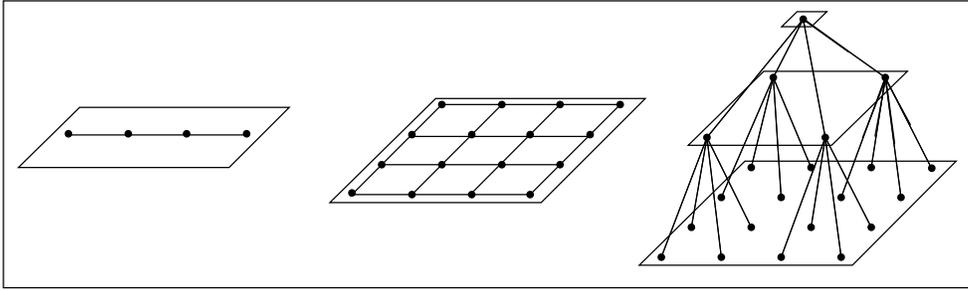
Le type d'approche statistique de l'analyse d'image que nous privilégions nous conduit à mettre en place des *modèles markoviens partiellement observés*, associés à des distributions $P(z|\theta)$, où θ est un jeu de paramètres. L'ensemble $z = \{z_s, s \in S\}$ se décompose en variables observées y 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 bi-dimensionnelle et le quad-arbre sont les graphes les plus fréquemment rencontrés en analyse d'image ; ils sont illustrés sur la figure jointe.

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

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



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 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

[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.

construites, puis exploitées au sein d'une *algorithmique descendante*. Les attributs \hat{x}^i inférés à un niveau i servent de point de départ à la procédure d'inférence au niveau $i - 1$ de "taille" immédiatement supérieure. Il s'agit ici principalement des méthodes multirésolutions ou multigrilles (selon que des versions réduites de y sont également calculées ou non). Ces méthodes permettent une accélération substantielle des schémas itératifs déterministes standard, et fournissent souvent une amélioration de la qualité des résultats [3].

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

3.2 Commande référencée vision

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

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

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

Afin d'élaborer une loi de commande en boucle fermée sur des mesures $s(t)$, il est nécessaire d'estimer ou d'approximer la relation qui lie la variation de s aux variables de contrôle. Dans le cas d'informations visuelles géométriques (telles par exemple les coordonnées 2D d'un point) contrôlées à l'aide des six degrés de liberté d'une caméra, cette relation est définie par [1] :

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

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

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

[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.

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

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

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

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

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

-
- [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.

3.3 Extraction de pistes et trajectographie

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

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

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

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

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

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

où 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

[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.

S_i ayant une origine commune ? On entend ici par origine commune le fait que les différents points de S_i sont associés à une même source en mouvement. Un sous-ensemble S_i s'appelle alors une piste extraite. Il importe de noter que cette suite est d'une nature statistique plus compliquée que précédemment puisque, dans le cas général, elle inclut des fausses alarmes, des fausses associations ainsi que des non-détections.

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

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

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

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

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

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

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

Une approche séduisante consiste alors à considérer que les probabilités d'assignation des mesures aux pistes sont indépendantes d'un scan à l'autre. Ainsi, la méthode PMHT (Probabilistic Multiple 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].

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

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, mécanique des fluides expérimentale, imagerie médicale, santé, sonar, indexation vidéo, multimédia, robotique manufacturière, robotique sous-marine, transports.

Panorama : *Nous nous intéressons à trois grands types d'applications. Le premier concerne ce que l'on peut appeler la "métrologie" du mouvement et des déformations. Nous explorons dans ce cadre les domaines de la visualisation expérimentale en mécanique de fluides (écoulements turbulents, avec le Cemagref), de l'imagerie satellitale météorologique (images Meteosat, collaborations récentes avec Météo-France, le LMD et Eumetsat), de l'imagerie médicale (imagerie cérébrale, IRM, IRMf, échographie 3D). Le second secteur a trait à la robotique, aussi bien robotique manufacturière que robotique d'intervention, et aux systèmes de détection. Les domaines d'applications concernés ont été récemment ou sont actuellement l'énergie (avec EDF), le sous-marin (avec l'Ifremer), l'agro-alimentaire (avec le Cemagref), les transports (projet européen IST Carsense, avec notamment le Livic). Enfin, nous menons un investissement significatif dans le domaine de l'indexation vidéo. Nous sommes concernés par les aspects de structuration temporelle de vidéo et d'indexation par leur contenu spatio-temporel, avec différents objectifs comme la consultation d'archives audio-visuelles ou la création de résumés de documents vidéos (avec des collaborations avec l'Ina, Thomson-Multimedia). Nous avons participé récemment à un projet européen Esprit sur le sujet (projet Divan, avec le projet Sigma-2), et nous participons au projet Agir dans le cadre du RNRT, au projet Mediaworks (avec le Limsi et TF1 notamment) dans le cadre du programme Priamm, et au projet Domus Videum récemment labellisé dans le cadre du programme RNTL (avec TMM notamment).*

5 Logiciels

5.1 Logiciel RMR d'estimation du mouvement dominant

Correspondant : Fabien Spindler

[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.

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, [9]. 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.

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

Correspondant : Fabien Spindler

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

La méthode proposée pour réaliser le découpage de la vidéo en plans élémentaires^[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 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

[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.

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.

5.3 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.

5.4 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 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 [8, 22, 39].

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 5.7). Il a été déposé à l'APP. Le but du logiciel ROMEO est le recalage non-rigide, ou 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 [66].

5.6 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.

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

Correspondant : 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 plate-forme logicielle, dénommée VISTAL pour "Volumetric ImageS Treatments and Analysis Libra-

ries”, 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 en 2000 (co-auteurs : C. Barillot, F. Boisse et P. Hellier).

5.8 Actions de développement : Plates-formes de vision robotique

Participant : Fabien Spindler.

Nous avons procédé cette année au basculement des applications liées au robot cylindrique et au robot cartésien, de l’environnement Sun Solaris vers l’environnement PC Linux. Ce travail a tout d’abord nécessité l’installation et la mise en route sous Linux de nouveaux équipements périphériques comme des cartes de numérisation d’images (IC-comp de la société Imaging Technology) et des coupleurs de bus (PCI-VME bus de la société Bit3). Ensuite, tous les outils et démonstrations existants ont été portés sous Linux. Nous avons également créé de nouvelles interfaces de manipulation des robots et des matériels associés.

5.9 Actions de développement : Véhicule Cycab

Participants : Fabien Spindler, Nicolas Ducoin.

Le début de cette année 2001 a été consacré à l’installation de différents périphériques sur le véhicule. Dans un premier temps, nous avons réalisé une interface mécanique permettant d’installer la tourelle site-azimut PTU de la société Directed Perception équipée d’une caméra CCD Sony. À l’arrière du véhicule, nous avons installé et configuré un PC Linux au format “shoe box” destiné à accueillir des traitements demandant d’importantes ressources de calcul. Une carte de numérisation IC-comp de la société Imaging Technology permet de numériser les images provenant de la PTU. À l’aide de ces matériels, une démonstration de détection d’obstacles via la caméra embarquée est en cours de développement.

Le dernier trimestre de cette année est consacré à la mise en route du véhicule. Nous allons intégrer toutes les fonctionnalités nécessaires à la commande par logiciel du Cycab. Pour cela, il s’agira d’intégrer les développements réalisés par Robosoft et le projet Sosso de l’Inria Rocquencourt pour commander le véhicule à l’aide de l’environnement Syndex à partir du PC Linux embarqué.

6 Résultats nouveaux

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

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. Un thème central et encore très ouvert sur ce sujet concerne l'analyse du mouvement déformable et en particulier fluide ; nous visons dans ce contexte une certaine métrologie du mouvement et des déformations. Un autre champ d'investigation porte sur les questions de reconnaissance du mouvement avec des applications notamment en indexation vidéo par le contenu.*

6.1.1 Modèles pour l'analyse du mouvement fluide

Participants : Étienne Mémin, Thomas Corpetti.

L'analyse de séquences d'images montrant des écoulements fluides est un problème central dans des domaines tels que les sciences environnementales (météorologie, océanographie, climatologie), la visualisation en mécanique des fluides expérimentale (aéro- ou hydro-dynamique) ou encore l'imagerie médicale (étude d'écoulements de bio-fluides). Dans tous ces domaines, il peut s'agir d'estimer à partir de ces images des champs de vitesses instantanées de l'écoulement visualisé ou de suivre dans le temps des structures cohérentes de l'écoulement (structures tourbillonnaires, vortex, puits/sources).

En particulier, dans un tel contexte, l'intensité lumineuse peut subir de fortes distorsions spatiales et temporelles rendant difficile l'utilisation de techniques génériques d'estimation de mouvement issue de la vision par ordinateur et originellement conçues pour des mouvements rigides et une hypothèse d'invariance de la fonction d'intensité. Afin de pallier ce type de défaut et notamment de rendre compte des mouvements du fluide transverses au plan de visualisation (mouvements occasionnant une variation de l'intensité lumineuse et caractérisés par un mouvement apparent de divergence non nulle), nous avons proposé une contrainte spécifique pour l'estimation du mouvement apparent. Cette nouvelle contrainte, issue de l'équation de continuité de la mécanique des fluides, rend compte de la divergence du mouvement apparent et de la variation d'intensité qui lui est associée. Elle a été en particulier élaborée pour être intégrée de façon naturelle dans notre schéma d'estimation hiérarchique et pour pouvoir, contrairement à une utilisation directe de l'équation de continuité, faire face à des déplacements de grande amplitude. L'association de ce modèle d'attache aux données et d'une fonction régularisante de type "div-curl" du second ordre favorise l'émergence de zones à rotationnel et à divergence non nuls, [22]. Cette approche a été évaluée et comparée, en premier lieu sur des images météorolo-

giques, à l'approche générique de mesure de champs de vitesses précédemment élaborée dans le projet [8]. 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 vis-à-vis de l'estimateur de mouvement générique, qui malgré l'utilisation d'estimateurs robustes, conduit à des champs de vitesses trop lisses. Cette méthode a d'autre part été évaluée dans le cas d'images montrant l'évolution de l'interface de couches de mélange. Des comparaisons avec une approche classique de type PIV ("Particle Image Velocimetry") en mécanique des fluides expérimentale ont montré également la supériorité de notre estimateur.

Nous avons également développé une technique permettant la détermination de tous les points singuliers d'intérêt (vortex, puits, sources) d'un champ de vitesses 2D continu et différentiable. Cette méthode s'appuie sur une décomposition du champ en une composante à divergence nulle (composante solénoïdale) et en une composante à rotationnel nul (composante irrotationnelle). Cette décomposition, dite de Helmholtz, est réalisée en fait dans le domaine de Fourier, et permet d'estimer les deux fonctions de potentiel dont sont issues ces deux composantes, le potentiel de vitesse et le potentiel de courant. Il est alors aisé de démontrer que les différents extréma locaux de ces fonctions de potentiel correspondent à des singularités du champ des vitesses de type vortex, puits ou sources. La détermination des points singuliers nous a permis d'autre part d'estimer une représentation paramétrique du champ des vitesses. Cette représentation s'exprime comme la somme de fonctions complexes décrivant l'influence sur le flot de chacune des singularités. Ainsi, à titre d'exemple, l'influence d'un vortex donné peut être modélisée par une fonction à rotationnel constant sur un disque centré sur la singularité et par une fonction analytique (à rotationnel et à divergence nuls) en dehors de ce disque. L'estimation des différents paramètres du champ est obtenue par maximum de vraisemblance. Cette représentation permet au final d'extraire les structures principales de l'écoulement. La méthode proposée s'avère 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 Super-résolution de séquences d'images

Participants : Patrick Bouthemy, Fabien Dekeyser.

Cette étude s'est conclue cette année, [12]. Notre objectif était de développer de nouvelles méthodes pour la restauration de séquences d'images utilisant explicitement la redondance d'information existant entre les images successives. Nous nous intéressons plus particulièrement à la création d'une image, ou d'une suite d'images, de haute résolution à partir de la séquence d'images fournies par le capteur. Un tel objectif implique bien évidemment la connaissance du mouvement apparent dans la séquence d'images. Afin de limiter le coût de calcul, critère important dans le contexte applicatif étudié, nous avons cherché à estimer un modèle paramétrique du mouvement apparent plutôt qu'un champ dense de vitesses 2D, ce qui est apparu comme une approche originale dans ce contexte. Cependant, l'utilisation d'un seul modèle paramétrique de mouvement ne permet d'estimer que le mouvement dominant et non des mouvements secondaires, ce qui conduit à la mise en œuvre de traitements adaptatifs de filtrage et de super-résolution. Cette étude a été menée en collaboration avec Thales-Optronique (voir

sous-section 7.1). Nous avons tout d'abord conçu un algorithme itératif qui remet à jour à chaque nouvelle image une estimation de l'image super-résolue. Pour cela, l'image observée est simulée à partir de l'image super-résolue estimée (sous-échantillonnage et transformation géométrique correspondant au mouvement dominant paramétrique estimé entre les deux images). L'erreur entre l'image observée et l'image simulée est utilisée pour remettre à jour l'estimée. Une alternative par correction multiplicative a été explorée cette année [57]. Rappelons qu'une carte de détection des points ne relevant pas du mouvement dominant, carte obtenue par une méthode préalablement développée dans le projet^[OB97], permet d'inhiber en ces points la remise à jour de l'image super-résolue.

Nous avons ensuite étudié des approches réalisant conjointement la restauration des images et l'estimation du mouvement apparent. Nous avons premièrement défini une version "pyramidale" du schéma original, par multi-résolution temporelle, qui a montré une amélioration significative de la qualité de la suite d'images super-résolues reconstruites. Cette version pyramidale comprend la génération successive d'images super-résolues intermédiaires (dans l'échelle temporelle) entre lesquelles est ré-estimé le mouvement dominant qui est ensuite exploité pour créer les images super-résolues à l'échelle temporelle suivante. Nous avons ensuite développé une approche plus explicite du problème par la définition, et la minimisation alternée, d'une fonction d'énergie globale régularisante (modélisation markovienne) spécifiant, de manière conjointe, l'estimation des paramètres du mouvement et la reconstruction d'une suite d'images super-résolues. Le terme de régularisation spatiale sur les images reconstruites doit notamment être défini avec attention, puisque le lissage introduit ne doit pas pénaliser la reconstruction super-résolue impliquant des informations liées aux hautes fréquences spatiales. Nous avons ainsi opté pour une régularisation du second ordre (modèle de Schultz-Stevenson) associée à un estimateur robuste (fonction de Huber). Ce travail a continué à connaître des interactions avec Patrick Pérez, actuellement à Microsoft Research à Cambridge, Angleterre.

6.1.3 Modélisation non paramétrique et reconnaissance du mouvement

Participants : Patrick Bouthemy, Ronan Fablet, Gwénaëlle Piriou.

Nous avons proposé une approche originale de modélisation statistique et de caractérisation non paramétrique du mouvement dans des séquences d'images vidéos, et considéré son application à des problèmes de reconnaissance et de classification, ainsi que de segmentation des images, basés sur une notion générale de mouvement que l'on a appelée "activité de scène", [13]. L'objectif était d'appréhender une très large gamme de situations ou classes de mouvements englobant des mouvements rigides (comme ceux de véhicules), des mouvements articulés (comme celui d'une marche ou d'une course), des mouvements non structurés (comme celui d'un groupe de personnes ou de feuillages au vent), des mouvements fluides (comme celui d'une rivière ou de flammes). Notre approche repose sur l'analyse de statistiques de co-occurrence de mesures locales de mouvement et leur représentation par des modèles de Gibbs causaux, [28]. Cette modélisation permet une évaluation complète de la vraisemblance conditionnelle des

[OB97] J.-M. ODOBEZ, P. BOUTHEMY, « Separation of moving regions from background in an image sequence acquired with a mobile camera », *in: Video Data Compression for Multimedia Computing*, H. H. Li, S. Sun, and H. Derin (eds.), Kluwer Academic Publisher, 1997, ch. 8, p. 283–311.

séquences d'observations locales de mouvement relativement à un modèle de mouvement, ce qui conduit à un schéma simple et direct à la fois pour l'apprentissage de ces modèles et pour la reconnaissance ou la classification d'une séquence donnée. Une phase de réduction (par tests de vraisemblance) de l'ordre des modèles estimés a de plus été introduite. La mesure de similarité entre modèles de mouvement, quant à elle, peut s'exprimer naturellement à partir de la divergence de Kullback-Leibler (une version symétrisée en fait). Nous avons ainsi un cadre statistique général pour la segmentation, la classification et la reconnaissance des contenus dynamiques de vidéos, avec en particulier des applications en indexation vidéo par le contenu, aussi bien pour la satisfaction de requêtes globales que de requêtes partielles.

Le travail le plus récent a principalement porté sur l'investigation de plusieurs schémas d'évaluation des co-occurrences des quantités locales de mouvement considérées. Le but était d'enrichir les modèles d'activité (jusqu'à présent estimés à partir des co-occurrences temporelles des quantités de mouvement) par la prise en compte d'éléments reflétant non seulement les propriétés temporelles des mouvements étudiés (évolution temporelle eulérienne), mais également des propriétés spatiales de ces derniers (agencement spatial du mouvement instantané dans l'image). Néanmoins, nous voulions rester dans le cadre des modèles de Gibbs causaux (ou chaînes de Markov) en raison de leurs bonnes propriétés (capacité d'évaluer complètement et simplement la vraisemblance conditionnelle, ce qui est primordial pour l'estimation des modèles ainsi que pour les aspects de reconnaissance et de classification). Ainsi, deux schémas originaux ont été définis. Le premier repose sur la génération de marches aléatoires spatio-temporelles au sein du volume que représente une séquence d'images, [59]. Le second introduit des cartes multi-résolutions des quantités locales de mouvement et l'évaluation de co-occurrences en échelle, conduisant à un schéma séquentialisant la prise en compte des co-occurrences temporelles et celles en échelle dans la suite des cartes multi-résolutions de quantités de mouvement, [60]. Les expérimentations ont montré une amélioration substantielle des performances des algorithmes associés de classification et de reconnaissance du mouvement. Les perspectives de ces travaux sont nombreuses, qu'elles concernent la considération d'autres types de mesures locales de mouvement, la combinaison de modèles paramétriques et non paramétriques de mouvement, ou la sélection (par apprentissage, supervisé ou non), la catégorisation et la reconnaissance de modèles d'activité pertinents en liaison avec différentes applications (détection d'événements particuliers, création de résumés vidéos, ...).

6.1.4 Modèles probabilistes de formes en imagerie médicale

Participants : Christian Barillot, Isabelle Corouge.

Nous étudions des techniques d'analyse statistique multidimensionnelle pour la modélisation de formes en imagerie médicale. Nous nous attachons à construire un modèle déformable générique représentant une forme donnée et les déformations qu'elle peut subir. Ce modèle est bâti à partir d'une population d'apprentissage, sur laquelle nous pratiquons ensuite 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, [50]. Dans cet espace local, les déformations d'une structure sont caractérisées par un vecteur de déplacement par rapport à une structure de référence (structure X ou structure moyenne). Les composantes principales de la matrice de covariance des vecteurs de déplacement décrivent les modes principaux de déformation. Les tests réalisés montrent la pertinence des premiers modes.

Deux applications de cette modélisation ont été abordées. D'une part, elle a servi aux travaux sur l'évaluation de méthodes de recalage global inter-sujets (méthodes rigides et non rigides). En effet, pratiquée sur des amers locaux, l'analyse statistique fournit un indicateur de la similarité des formes au sein des populations recalées, et produit ainsi un critère de comparaison entre les méthodes. D'autre part, nous avons exploité la connaissance statistique apportée par le modèle sur les sillons dans le contexte de la construction d'atlas anatomique et fonctionnel. En particulier, nous avons travaillé sur 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. L'approche adoptée est locale et non-linéaire, et vise à étendre le champ des déformations, obtenu par l'analyse statistique, entre un sillon donné et le sillon de référence, à un voisinage local du sillon considéré, le rendant ainsi applicable à tout point associé à ce sillon. Ceci est réalisé par une interpolation de type "thin-plate splines", et permet 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 inter-indivuelle inhérente au groupe de sujets étudiés. Dans un but de validation, nous avons comparé cette méthode à d'autres approches globales classiques. Notre méthode offre précisément l'avantage d'être locale et indépendante d'un cerveau de référence, et de modéliser parfaitement la déformation anatomique entre deux sillons. Les travaux en cours cherchent à étendre le modèle statistique à un graphe 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.

6.1.5 Résumé et indexation vidéo

Participants : Patrick Bouthemy, Ronan Fablet, Emmanuel Veneau, Nathalie Peyrard.

Nous nous sommes cette année particulièrement attachés aux deux problèmes suivants relatifs à l'indexation de vidéos par leur contenu : *i*) la création de résumés de vidéos ; *ii*) la classification de scènes d'après leur contenu dynamique. La création de résumés de vidéos est un sujet encore très peu exploré. L'objectif est de fournir une version très courte d'un document vidéo (typiquement dans un rapport 1 à 100), "raisonnablement" représentative de son contenu. De tels résumés de vidéos pourront être exploités dans des contextes professionnels, ou domestiques, d'archivage et de consultation de bases de documents audiovisuels pour faciliter des opérations de visualisation, de navigation ou de présélection. Nous avons commencé, à travers le stage de DEA de Caroline Lacoste et une collaboration avec J.-F. Yao de l'Irmar, à développer une approche originale pour la création automatique de résumés de vidéos. Elle s'appuie sur l'analyse du contenu dynamique de micro-segments de la vidéo, contenu décrit par les distributions de co-occurrences temporelles de quantités locales de mouvement, distributions elles-mêmes représentées par des modèles de Gibbs temporels (cf paragraphe 6.1.3).

Nous avons tout d'abord réalisé une segmentation temporelle de la vidéo en plages homogènes au sens de cette notion d'activité de mouvement, résultant d'une classification hiérarchique ascendante et conduisant à une sélection d'instants représentant au mieux la vidéo. Une seconde méthode repose sur la sélection directe de plages pertinentes et se base sur une modélisation par chaîne de Markov cachée à deux états. Les modèles d'activité associés aux états pertinent et non-pertinent sont obtenus sur critère entropique. Les premiers résultats expérimentaux ont permis de vérifier l'intérêt de l'approche statistique adoptée. Cette étude sera poursuivie dans le cadre du projet RNTL Domus Videum (cf paragraphe 7.10.1). Elle comprendra notamment l'apprentissage et la sélection de modèles pertinents sur d'autres critères que le critère entropique, et pourra s'étendre à la prise en compte d'éléments extraits de la bande sonore.

Nous continuons à développer, en collaboration avec l'Ina (cf paragraphe 7.7), un schéma de structuration et d'indexation de vidéos lié au contenu dynamique de ces dernières. Nous nous intéressons à la caractérisation de l'activité des plans ou des séquences extraits, selon un schéma de classification hiérarchique par SVM ("Support Vector Machines"). Nous avons exploré différentes variantes de mesures locales spatio-temporelles, simples différences temporelles d'images, différences temporelles après compensation du mouvement de la caméra, sorties de filtres de Gabor spatio-temporels, ainsi que de descripteurs, cartes d'"historiques" (cartes MHI introduites par Bobick), cartes de co-occurrences temporelles. La méthode développée a été validée sur des exemples de documents audio-visuels de sport, qui forment une étude de cas représentative et complexe.

6.1.6 Indexation de données numériques

Participants : Patrick Gros, Sid Ahmed Berrani.

Nous avons poursuivi, en collaboration avec Laurent Amsaleg du projet TEMICS, notre étude des problèmes posés par l'indexation et la recherche dans de grandes bases de documents multimédias, [17]. Cela concerne la vidéo, les images fixes mais aussi la bande sonore. Une fois calculés les descripteurs associés à ces médias, il apparaît en effet nécessaire de pouvoir les stocker, puis les retrouver de manière rapide et efficace, lorsque l'on doit répondre à une requête. Soit ce stockage se fait en mémoire principale, ce qui est le cas dans la quasi-totalité des systèmes actuels et ne peut gérer de trop grandes quantités de données, soit il peut se faire en mémoire secondaire, mais les SGBD actuels ne sont pas conçus pour le type de données extraites des médias, et de nouvelles techniques d'indexation doivent être mises au point. Notre travail a débuté par un état de l'art et la comparaison de quelques méthodes publiées, ce qui nous a permis de faire un bilan des problèmes posés. Le défi est d'arriver à battre une recherche séquentielle et exhaustive. Les résultats publiés sont toutefois fortement biaisés par le fait que la taille des données utilisées permet de charger ces données simultanément en mémoire centrale où la recherche est bien plus rapide. Nous nous proposons donc de réexaminer ces algorithmes dans le cas où les données sont très volumineuses et où leur lecture en une seule fois n'est pas possible, [43].

Diverses solutions, basées sur des algorithmes de clustering et des schémas d'approximation des données, sont envisagées et ont commencé à être testées. Il s'agit d'organiser les données de manière à ne pas avoir à toutes les considérer lors d'une requête, mais aussi d'éviter d'avoir

à gérer des structures arborescentes dont la complexité est le plus souvent exponentielle avec la dimension de l'espace des descripteurs. Le but de notre travail est de pouvoir faire une recherche séquentielle, mais sur une petite partie des données seulement. D'autres pistes sont aussi explorées. L'une d'entre elles consiste à utiliser des accélérateurs matériels de type FPGA afin de filtrer les données au plus près de leur lieu de stockage, pour n'avoir à traiter au niveau du processeur central que le petit ensemble des données les plus pertinentes. Cet axe de recherche est mené en collaboration avec Dominique Lavenier du projet AIDA. Ce travail, outre son intérêt pour ce qui est des algorithmes d'indexation, permettra aussi de valider les capacités de reconnaissance des descripteurs utilisés sur de plus grands volumes de données. Pour cela, nous avons commencé des tests sur une collection de 690 000 images. L'objectif est de passer de tests sur 1000 à 5000 images à des tests sur un million d'images.

6.2 Suivi, trajectographie, recalage 3D

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 et trajectographie

Participants : Jean-Pierre Le Cadre, Carine Hue.

Pour traiter les problèmes de suivi, 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 présentes dans les équations d'observation. Elles se heurtent également 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, 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. Dans le contexte des applications de suivi intéressant le projet (extraction de pistes en sonar ou radar, suivi de véhicules en imagerie vidéo, analyse d'images de fluides en mouvement), les travaux se sont focalisés sur le suivi, par des méthodes de filtrage particulière, d'objets mobiles et partiellement observés. Dans un premier temps, les efforts ont consisté à analyser les techniques classiques de filtrage particulière pour le suivi

d'un objet unique et manœuvrant, partiellement observé (par exemple, dans le cas de mesures d'angle seules en trajectographie sonar). Cette analyse a permis de mettre en évidence l'intérêt de telles méthodes pour ce problème. On peut ainsi souligner la robustesse de celles-ci, que ce soit par rapport au modèle de trajectoire de l'objet, au bruit d'estimation, aux non-linéarités du modèle des observations et encore à l'initialisation. Enfin, le coût de calcul n'apparaît pas prohibitif.

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 les problèmes 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. 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 (aussi connu sous le vocable de "data augmentation method"). 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 [33]. Du fait de la généralité de l'approche et de la flexibilité des hypothèses prises, il est possible d'étendre cette approche à de nombreux 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 d'importants développements (suivi de piétons) [68, 70]. 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). Ce travail continue à se faire en coopération avec Patrick Pérez, actuellement à Microsoft Research à Cambridge, Angleterre.

6.2.2 Suivi de structures caractéristiques dans des images de fluides

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

Le champ des vitesses d'un écoulement fluide peut être décrit par un certain nombre de structures cinématiques cohérentes tels que les vortex, puits ou sources. L'extraction et le suivi de ces singularités du flot représentent des questions ardues dans l'analyse du mouvement fluide à partir de séquences d'images. En effet, la connaissance de ces points singuliers est d'une importance capitale pour la caractérisation du flot, et, réciproquement, leur localisation nécessite une connaissance au moins partielle du flot. Localisation des points singuliers et estimation du flot sont par conséquent deux problèmes étroitement couplés qui ne peuvent être appréhendés de façon indépendante [MBD98]. Fort heureusement, le champ des vitesses est assimilable à un

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

champ linéaire dans le voisinage de ces singularités; cette propriété offre la possibilité d'une description partielle du flot.

Le problème dans son ensemble se pose comme l'estimation de paramètres pour un système non linéaire (par rapport aux coordonnées des points singuliers à déterminer et aux domaines de linéarité associés) qui peut parfois être partiellement observé (en météorologie par exemple certaines couches nuageuses peuvent être en partie occultées). Le suivi temporel de ces paramètres s'entend naturellement dans un cadre de filtrage (ou de poursuite). Néanmoins, la complexité du modèle nécessitera sans doute l'emploi de méthodes d'estimation sophistiquées faisant intervenir des méthodes de type Monte-Carlo ou quasi-Monte-Carlo (MCMC, filtrage particulière), ainsi que des versions approchées mais plus réalistes en temps de calcul. D'autre part, afin d'appréhender au mieux la variabilité et l'évolution de ces phénomènes, nous tâcherons d'introduire dans les différentes méthodes étudiées des connaissances relatives à la nature physique du phénomène. Ainsi, l'introduction des lois de la mécanique des fluides pourrait conduire à des méthodes de filtrage originales guidées par des équations aux dérivées partielles traduisant l'évolution des phénomènes observés. Il s'agit d'un travail qui vient de débiter avec l'arrivée des personnes concernées entre septembre et octobre 2001.

6.2.3 Méthodologie d'évaluation de méthodes de recalage non-rigide

Participants : Christian Barillot, Pierre Hellier, Etienne Mémin, Isabelle Corouge.

Nous avons 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 18 sujets, 6 méthodes ont été évaluées : la méthode "Animal" développée au Montreal Neurological Institute (McGill University), la méthode des démons développée par le projet Epidaure de l'Inria Sophia-Antipolis, la méthode de l'université d'Iowa, un recalage rigide par maximisation de l'information mutuelle, le quadrillage proportionnel de Talairach, ainsi que la méthode Romeo développée dans notre projet (voir 5). Nous avons pour cela spécifié des critères globaux et locaux afin de quantifier la précision des recalages obtenus, [65]. Ces critères, basés sur des structures anatomiques d'intérêt extraites des images IRM des sujets, sont indépendants de la manière dont les transformations ont été estimées, ce qui garantit l'objectivité de cette évaluation. Les critères globaux s'appuient sur le calcul d'un volume moyen, sur le recouvrement après recalage des matières grises et blanches des sujets, sur la corrélation après recalage des courbures locales, et sur une mesure des singularités éventuelles du champ des déformations calculé. Les mesures locales sont basées sur la mise en correspondance des principaux sillons corticaux, qui sont des structures corticales particulièrement intéressantes d'un point de vue anatomique et fonctionnel. Pour les mesures globales, nous avons montré que les méthodes de recalage non rigide obtiennent de meilleurs résultats que les méthodes de recalage rigide. La qualité du recalage est de plus directement reliée au nombre de degrés de liberté de la transformation. Nous avons pu également montrer la nécessité d'imposer des contraintes sur la transformation afin d'éviter la présence de singularités. Les mesures locales n'ont par contre pas permis de distinguer les méthodes de recalage rigide et non rigide, ce qui constitue un résultat surprenant et nouveau pour la communauté.

Forts de cette évaluation, nous avons proposé une méthode mixte pour recalcr les images des

cerveaux de différents individus. Cette approche s’appuie sur une estimation régularisée du flot optique (logiciel Romeo), et comprend l’introduction explicite de contraintes géométriques éparées, [66]. Le cadre statistique d’estimation permet d’incorporer naturellement la mise en correspondance d’amers anatomiques, en l’occurrence des sillons corticaux. L’apport objectif de cette contrainte a pu être montré sur une base de 18 sujets, sur des critères anatomiques mais aussi sur le recalage de données fonctionnelles. Nous avons également continué à nous intéresser au problème du recalage non rigide multimodal. Les images fonctionnelles *IRMf*, qui proviennent de séquences rapides d’acquisition (séquences EPI - Echo Planar imaging), présentent des déformations géométriques dues à une non linéarité du champ magnétique, ces déformations gênant l’interprétation des données fonctionnelles en déplaçant les cartes d’activations par rapport aux régions anatomiques d’intérêt. Nous avons proposé une amélioration de l’algorithme de recalage multimodal comprenant une mesure de similarité adaptée (l’information mutuelle) et un cadre de minimisation multigrille incorporant un modèle paramétrique affine à chaque niveau de grille. Nous avons introduit des contraintes nouvelles à la fois sur la régularité du champ des déformations (régularisation d’ordre 1 de type “Thin Plate Spline”) et sur sa topologie (introduction d’une contrainte sur le Jacobien du champ).

6.2.4 Analyse d’images ultrasonores 3D

Participants : Christian Barillot, Pierre Hellier, François Rousseau.

Dans le cadre de la réalisation d’un système d’acquisition d’échographies 3D par la technique dite “main libre”, nous nous sommes intéressés à la phase d’acquisition des images, [80]. Cela nous a tout d’abord conduits à évaluer la robustesse et la précision des localiseurs 3D de type électromagnétique. Ce type de localiseur 3D permet de connaître la position et l’orientation d’un capteur posé sur la sonde échographique, ce qui donne la possibilité de construire un volume à partir des images 2D (après une phase de calibrage). Il est important de connaître le comportement de ce type de localiseur dans un milieu où les interférences métalliques et électromagnétiques sont nombreuses, pour pouvoir ensuite l’utiliser en milieu hospitalier. Nous avons étudié les systèmes les plus utilisés du commerce : “FastTrack” de la société Polhemus, “Flock of Birds” de la société Ascension Technology, ainsi que le système de capteur optique portable “laserBIRD”. Nous avons mis au point une procédure d’évaluation sur notre plate-forme robotique afin de calibrer ces différents systèmes, les paramètres de calibrage retenus étant les paramètres d’angularité, de translation, ainsi que d’écartement à la sonde et à un environnement métallique. Les résultats de cette étude permettent de conclure qu’il est ainsi possible d’utiliser un système de repérage électromagnétique dans un environnement hospitalier sous la condition de contraindre l’acquisition des images échographiques à un rayon d’action de l’ordre de 70 cm du capteur.

Nous nous sommes également intéressés au problème du calibrage d’un système échographique 3D. La phase de calibrage se doit d’être simple, automatique et rapide en vue d’une utilisation en routine clinique. On trouve à l’heure actuelle des méthodes semi-automatiques qui nécessitent l’intervention du médecin (extraction des points d’intérêt ou localisation de la mire), et qui se révèlent fastidieuses dans leur utilisation et sans réelle possibilité de contrôle avant l’acquisition d’une séquence 3D sur le patient. Une nouvelle méthode robuste de calibrage

totalement automatique a été développée, et est actuellement en cours de validation sur un ensemble d'images synthétiques et réelles. Notre méthode repose sur le principe qu'un calibrage correct des paramètres donne un recalage optimal des images ultrasonores de la mire, cette mire étant formée d'une bassine remplie d'eau et ayant un fond plat. Le schéma de modélisation développé permet d'inclure les paramètres de localisation de la mire et ceux de calibrage de la sonde. L'optimisation du critère ainsi construit se fait dans un cadre multirésolution sur les multi-plans échographiques 2D et inclut des estimateurs robustes. Nous utilisons de plus la séquence échographique dans sa totalité pour effectuer un recalage 3D entre la mire plane et un ensemble de points extraits de la séquence d'images échographiques. Cette procédure pourra être complétée pour prendre en compte plusieurs paramètres de focalisation de la sonde échographique, afin notamment de réaliser un calibrage indépendant du facteur de profondeur (facteur pouvant varier au cours d'une acquisition de volume 3D).

active et exploration

6.3 Asservissement visuel, perception active et exploration

Contexte : La perception active consiste à élaborer des stratégies de perception et d'action dans le but d'améliorer les performances des algorithmes de vision ou de détection par le contrôle des paramètres du capteur, ou de réaliser des tâches robotiques (positionnement, saisie, suivi, . . .). Nous nous intéressons au contrôle de la position, de la vitesse, ou de la trajectoire de capteurs de 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 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 d'informations, 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.

Nous avons ainsi considéré les moments de la projection d'un objet dans l'image, facilement mesurables à partir d'une segmentation spatiale de l'image (la segmentation en tant que telle ne faisant pas partie de nos travaux). Nous avons tout d'abord élaboré une méthode de détermination de l'ensemble des moments invariants aux divers mouvements rigides (en translation et en rotation), ainsi qu'aux changements d'échelle. À partir de ces invariants, nous sommes à la recherche d'une sélection adéquate de moments. Il semble que deux cas doivent être considérés : le cas des objets symétriques et celui des objets non symétriques, un grand nombre de moments s'annulant dans le cas des objets symétriques.

6.3.2 Asservissement visuel virtuel pour la réalité augmentée

Participants : François Chaumette, Éric Marchand, Andrew Comport.

Nous avons récemment proposé l'utilisation de techniques d'asservissement visuel pour la réalité augmentée. Le problème fondamental posé dans ce contexte est en effet celui du calcul de pose (c.a.d., le calcul de la position de la caméra par rapport à un objet). L'asservissement visuel virtuel consiste à considérer le calcul de pose comme le problème inverse de l'asservissement visuel. Cette méthode prend l'image de l'objet comme le motif désiré 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 (un résidu) par l'algorithme de Levenberg-Marquardt. L'avantage de notre 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 et que les informations visuelles utilisables peuvent être très variées et même être combinées entre elles. Nous avons instancié cette approche pour des informations visuelles de type point, droite, cercle, cylindre, . . . L'apport se situe à la fois sur la simplicité de la modélisation et donc de la mise en œuvre, et sur la richesse de l'information utilisable. Nous avons développé une extension de cette méthode afin de réaliser l'estimation des paramètres intrinsèques de la caméra. Il est alors nécessaire de calculer la variation de la position d'un point de l'image en fonction de la variation des différents paramètres de la caméra (paramètres intrinsèques et pose). La qualité des résultats obtenus est à comparer à celle résultant des meilleures méthodes de calibration actuellement disponibles. Il reste que, là encore, la modélisation du problème et les algorithmes sous-jacents sont beaucoup plus simples sous cette forme.

6.3.3 Navigation réactive dans des espaces virtuels

Participant : Éric Marchand.

Notre objectif est de proposer un cadre général permettant le contrôle d'une caméra dans un environnement virtuel. Ce travail est mené en collaboration avec le projet Siames. 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 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, [54]. La possibilité de prendre en compte, de manière dynamique et en temps-réel des modifications de l'environnement, découle de l'utilisation de contraintes intégrées dans les lois de commande considérées. Cette approche est ainsi adaptée à des contextes hautement réactifs (réalité virtuelle, jeux vidéos). De façon générale, une des originalités de cette approche (dans ce contexte d'applications) est le contrôle défini dans l'image. On est en effet ramené à spécifier des tâches à accomplir dans un espace

[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.

2D, et à obtenir des trajectoires de caméra dans un espace 3D. C'est en fait une approche de l'animation très intuitive, tout en étant formalisée, que nous proposons, puisque celle-ci est réalisée en fonction de ce que l'on souhaite observer dans la séquence d'images produites. Nous avons comparé cette année cette approche purement 2D à des approches 3D et 2D 1/2. Nous avons d'autre part élaboré la commande d'humanoïdes de synthèse. L'asservissement visuel est en effet une approche intéressante pour générer les commandes de bas niveau de tels avatars. Il faut noter que les mouvements de la caméra virtuelle (ici les yeux de l'humanoïde) ne sont plus libres. Il faut considérer la caméra virtuelle comme étant montée sur l'effecteur d'un robot hyper-redondant, typiquement à neuf degrés de liberté si l'on considère la seule animation du buste, et plus si l'on considère la marche. Cette approche du contrôle des humanoïdes a aussi été intégrée dans des systèmes plus larges de gestion du comportement (cf. rapport d'activité du projet Siames).

6.3.4 Reconstruction 3D par vision dynamique active

Participants : François Chaumette, Grégory Flandin.

Nous avons achevé cette année l'étude portant sur la localisation et l'estimation du volume englobant d'objets complexes par vision dynamique active, [14, 61, 63]. Nous rappelons que le modèle d'objet que nous avons développé résulte d'un mélange de modèle probabiliste et de modèle à erreur bornée sous une hypothèse ellipsoïdale du support de l'erreur. De plus, l'algorithme d'estimation intègre progressivement l'ensemble des informations acquises au cours de la trajectoire de la caméra locale embarquée sur le robot, et les fusionne avec celles acquises par la caméra observant globalement la scène. Nous avons déterminé les mouvements de la caméra permettant de raffiner de façon optimale le modèle de l'objet. La stratégie développée consiste à réduire de façon isotropique l'incertitude liée au modèle estimé. Cette technique a été validée sur des objets de forme relativement complexe et sur des images fortement texturées (aussi bien pour le fond de la scène que pour les objets considérés) grâce à l'emploi d'un algorithme de segmentation fondé sur une analyse du mouvement apparent dans la séquence d'images.

Nous avons également développé une méthode d'exploration d'une scène dont le but est de détecter la présence (ou l'absence) d'objets en son sein. Le modèle que nous utilisons s'insère dans un cadre probabiliste et prend en compte, en plus des observations successives, d'éventuels *a priori* structurels ou relationnels sur la distribution des objets dans la scène. Nous avons choisi d'exprimer la connaissance sur la scène sous forme de probabilités d'occupation, celles-ci étant définies sur des zones (ou cellules) issues d'un échantillonnage préalable de la scène. L'échantillonnage que nous proposons est construit sur la base des observations issues de la vision globale : les zones de recherche sont les troncs de cône engendrés par la projection des objets dans l'image globale. À l'instant initial, l'état de la connaissance sur la scène se résume à un certain *a priori* sur le type de scène (encombrée ou non, dense ou éparse, ...) représenté par la loi jointe sur toutes les cellules. Aux instants suivants et tout au long de la recherche, le modèle intègre les observations issues du traitement des images fournies par la caméra locale, sous la forme de données homogènes et compatibles avec le modèle de la scène, en l'occurrence des contraintes sur les probabilités d'occupation des cellules observées. En effet, les *a priori*

sur la scène ainsi que les observations (une probabilité de détection le long de l'axe optique) se traduisent par des contraintes linéaires sur la loi jointe. L'intégration de contraintes linéaires s'exprime comme la résolution d'un système sous-contraint, résolution que nous construisons de manière itérative pour nous affranchir des problèmes d'inversion et surtout pour pouvoir intégrer les observations à mesure qu'elles se présentent. Enfin, nous avons défini une stratégie de recherche optimale déterminant la position à atteindre de la caméra locale pour optimiser la recherche. Il s'agit de prédire le gain d'information apporté par de nouvelles observations. Nous avons proposé quatre critères possibles dont trois sont basés sur la norme de la variation des marginales (ou probabilités d'occupation) et diffèrent par le choix de la norme. Le dernier critère proposé porte sur la variation d'entropie de ces marginales.

6.3.5 Planification dans l'image

Participants : François Chaumette, Youcef Mezouar.

Nous avons achevé cette année notre étude sur la planification de trajectoires dans l'image, [16]. Rappelons que cette approche permet de prendre en compte de grands déplacements de la caméra entre sa position initiale et sa position désirée. Elle permet aussi d'accroître la robustesse du système vis-à-vis des erreurs de modélisation et de calibration. La première méthode que nous avons développée était basée sur l'approche par fonctions de potentiels. Nous pouvions alors considérer aussi bien des contraintes sur la trajectoire 3D de la caméra (comme par exemple une trajectoire la plus proche possible d'une ligne droite) que des contraintes supplémentaires telles l'évitement des butées articulaires du robot, l'évitement d'auto-occultations de parties de l'objet, l'assurance que l'objet reste dans le champ de vision du capteur, [77]. Cette méthode fournit un ensemble de positions à atteindre dans l'image, ensemble à partir duquel il est possible de calculer des trajectoires temporelles par une approximation par des B-splines. Ces trajectoires sont ensuite suivies par les méthodes classiques d'asservissement visuel 2D. Nous nous sommes attachés cette année au développement d'une méthode différente, reposant sur un calcul explicite des trajectoires à suivre (ce qui évite la phase d'approximation par des B-splines). La méthode consiste à calculer les trajectoires de points dans l'image issues d'une trajectoire de la caméra, [78]. À partir de la matrice de colinéation relative à un plan de référence et aux images initiale et finale (matrice de colinéation qui peut s'estimer à partir d'un certain nombre de points appariés dans ces deux images), nous avons obtenu des trajectoires qui sont indépendantes des paramètres de calibration de la caméra (sous l'hypothèse que la matrice d'homographie à l'infini, directement liée à la matrice de colinéation et à la rotation effectuée par la caméra, peut être estimée). Ce résultat très intéressant assure une très grande robustesse aux erreurs de modèle des trajectoires générées puis suivies par asservissement visuel. Les trajectoires auxquelles nous nous sommes particulièrement intéressés sont des géodésiques, assurant une distance de parcours minimale. Un comportement temporel sur ces trajectoires peut aussi être obtenu par la minimisation d'un critère comme le minimum d'énergie ou le minimum d'accélération. Ces résultats s'obtiennent à travers le formalisme de la commande optimale. Nous avons en outre tenté d'employer ce formalisme pour ajouter des contraintes (telles que la préservation de la visibilité de l'objet). Il est alors malheureusement impossible d'aboutir à une forme analytique des trajectoires dans l'image. En lien avec les

aspects de couplage entre indexation d'images et asservissement visuel, nous avons abordé la planification de trajectoires entre N images (avec $N > 2$). Ces méthodes ont été validées sur notre cellule de vision robotique. Nous avons notamment utilisé un suivi de points d'intérêt pour prendre en compte des images naturelles dans nos expérimentations.

6.3.6 Couplage perception-action par indexation d'images

Participants : François Chaumette, Patrick Gros, Anthony Remazeilles.

Dans les applications pratiques, l'asservissement visuel est souvent limité par la nécessité de disposer d'un appariement entre l'image de départ et l'image à atteindre. Bien que nos travaux en planification de trajectoire (voir paragraphe 6.3.5) permettent de traiter le cas où la distance entre les images est grande, cela nécessite quand même que les deux images aient des éléments en commun. Nous avons donc proposé d'utiliser une base d'images de l'environnement du robot et des techniques d'indexation et d'appariement pour faire sauter ce verrou. Le principe est de constituer préalablement une base d'images de l'environnement du robot, puis, lorsqu'une tâche est définie, de rechercher dans cette base les images les plus proches de l'image de départ et de l'image cible. Un chemin est alors recherché dans la base, sous la forme d'une suite d'images intermédiaires, permettant de passer progressivement de l'image initiale à l'image cible, de telle manière que deux images consécutives de ce chemin aient toujours suffisamment d'éléments en commun pour permettre la mise en œuvre d'un asservissement visuel. On passe alors de l'image initiale à l'image cible par une suite d'asservissements. Pour obtenir une vitesse constante d'exécution du mouvement et éviter de devoir converger vers chaque image intermédiaire, plusieurs améliorations ont été proposées. L'utilisation d'homographie permet de prédire la position dans l'image courante de l'image cible suivante sur le chemin, et d'abandonner au plus vite l'image cible courante dès qu'on peut passer à l'asservissement suivant. La technique de planification de trajectoire permet alors d'assurer la convergence de ces asservissements et d'obtenir une vitesse constante sur l'ensemble de la trajectoire. Une première expérimentation avec une scène plane a permis de valider ce schéma. Il s'agit maintenant de passer à des scènes tridimensionnelles et à des chemins plus complexes. Pour cela, il faut remplacer l'usage des homographies par celui des tenseurs trifocaux, et affiner les critères de choix du chemin. En particulier, on peut vouloir non seulement limiter le nombre d'asservissements intermédiaires, mais aussi imposer des contraintes sur le mouvement effectué par la caméra, comme la limitation des rotations ou la prise en compte de la non holonomie du robot.

6.3.7 Théorie de l'exploration

Participants : Jean-Pierre Le Cadre, Frédéric Dambreville.

Nous étudions l'optimisation de l'effort de recherche d'un objet, généralement appelé cible, pouvant être fixe, mobile (avec par exemple une trajectoire supposée markovienne) ou même réactive. On dispose de ressources nécessairement limitées, et on cherche à optimiser des fonctions telles que la probabilité de détection, le temps de détection ou encore une fonction de risque. L'analyse classique a été étendue à l'optimisation de systèmes de détection multimodaux, en présence de contraintes syntaxiques (coopérations des modes et des ressources de

détection, conditions de dépendance). Il est ainsi possible de traiter l'optimisation de systèmes de détection complexes, incorporant plusieurs types de récepteurs avec des relations de conditionnement des ressources, [11]. Pour cela, un langage de description composé d'opérateurs laissant stable l'ensemble des problèmes à contraintes linéaires et permettant de prendre en compte leur comportement temporel (modèle quotient) a été développé [26].

Les jeux de recherche ("search games") sont le cadre naturel pour l'optimisation de la stratégie de l'observateur en présence d'incertitudes sur celle de la cible. Un *a priori* probabiliste markovien du mouvement de la cible est en effet plutôt restrictif : *i*) l'aspect probabiliste ne prend pas en compte les capacités réactives de la cible ; *ii*) la markovianité réduit l'ordre de complexité de la représentation. Dans les cas où la cible envisage et planifie son déplacement, dans une intention de furtivité par exemple, ce type de modélisation du comportement de la cible est tout-à-fait insuffisant. Aussi, une autre approche est de modéliser l'ensemble des trajectoires possibles de la cible (*a priori* ensembliste). Ces remarques amènent à considérer des jeux de recherche *a priori* sur des ensembles de trajectoires. On peut constater que 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 : *i*) l'usage de fonctions de détection locales de type exponentiel est une approximation ; *ii*) 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). Une redéfinition de l'*a priori*, rendant compte la réactivité de la cible, nous a amené à considérer une formulation de type minimax des problèmes d'optimisation de Brown [Bro80]. Une résolution directe du jeu s'étant avérée inadaptée, cette difficulté a été contournée par approximation du jeu par une séquence d'optimisations unilatérales de type Brown. Nos efforts se sont ensuite orientés vers une meilleure modélisation du comportement de la cible. Dans un cadre opérationnel réel, il apparaît en effet raisonnable de faire l'hypothèse que l'objet à détecter pourra se comporter de manière à gêner le plus possible sa détection. Le fait d'envisager une certaine réactivité de la cible nous a alors amenés 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. Les conditions d'optimalité de la stratégie du chercheur ont été développées [81].

7 Contrats industriels (nationaux, européens et internationaux)

7.1 Convention DGA - Thales Optronique : Conception d'une chaîne de traitements pour la restauration de séquences d'images

Participants : Patrick Bouthemy, Fabien Dekeyser.

no. Inria 1 98 C 500, durée 24 mois

[Bro80] S. BROWN, « Optimal search for a moving target in discrete time and space », *Operations Research* 28, 6, nov.-déc. 1980, p. 1275–1289.

Cette étude a été menée en collaboration avec Thalès-Optronique avec un financement DGA, dans le cadre du programme Civil-Défense, et s'est conclue en mars 2001. Deux applications ont été prises en compte, l'une relative au traitement de séquences d'images infra-rouges aériennes, l'autre au traitement d'images médicales endoscopiques vidéos. Notre rôle a consisté d'une part en l'étude et le développement d'algorithmes de restauration de séquences d'images incluant filtrage spatio-temporel et obtention d'images super-résolues, et d'autre part en une assistance pour l'intégration et l'évaluation des divers éléments de la chaîne de traitement. L'objectif a été non seulement de débruiter ces images ou de supprimer des défauts localisés, mais également de construire des images de plus grande résolution à partir de la séquence originelle. Une telle restauration de nature spatio-temporelle doit tenir compte du mouvement présent dans la séquence d'images et nous avons exploité des modèles paramétriques de mouvement 2D (cf. paragraphe 6.1.2). Des validations expérimentales des algorithmes de filtrage et de super-résolution développés ont donné des résultats très probants sur des séquences d'images infra-rouges aériennes fournies par Thales-Optronique. Le faible coût de calcul de ces algorithmes permet d'envisager leur utilisation à terme dans des applications temps-réel.

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

Participants : François Chaumette, Eric Marchand.

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 a accepté d'y apporter un soutien financier. Les travaux doivent porter 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 utiliserons une caméra embarquée sur l'effecteur d'un robot manipulateur.

7.3 Convention Ifremer : Commande en vitesse d'un bras manipulateur non instrumenté par asservissement visuel

Participants : François Chaumette, Eric Marchand, Fabien Spindler.

no. Inria 1 00 C 0606, durée : 9 mois.

Cette étude s'est déroulée dans la prolongation d'une convention passée l'an dernier intitulée "Étude sur la commande des mouvements d'un manipulateur de ROV à l'aide d'une caméra pan-tilt". Nous avons alors réalisé une commande par asservissement visuel, permettant au robot manipulateur d'effectuer un déplacement spécifié par l'opérateur, [74]. Dans cette nouvelle étude, nous avons réalisé une commande par asservissement visuel afin que le manipulateur suive une vitesse spécifiée par l'opérateur. On passe alors d'une commande exprimée sous la forme d'une position à atteindre à une commande exprimée sous la forme d'une vitesse à suivre. À partir de la consigne, il est possible de déterminer les trajectoires dans le plan image que doivent suivre les points d'intérêt de l'organe terminal observés par la caméra, puis de réaliser

une loi de commande permettant de suivre ces trajectoires en en minimisant les erreurs de traînage. La nouvelle loi de commande a été validée sur notre système robotique à l'Irisa. La validation sur le site de l'Ifremer Toulon est en cours de réalisation.

7.4 Convention LRBA, DGA : projet Posit

Participants : Jean-Pierre Le Cadre, Sébastien Paris.

no. CNRS 509707, durée 12 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 manoeuvres, contraintes opérationnelles, positions initiales et finales, contraintes de temps.

7.5 Convention Cifre Thales Airborne Systems : Méthodes et algorithmes de traitements multi-capteurs

Participants : Jean-Pierre Le Cadre, Isabelle Leibowicz.

no. Université KTNR, durée 38 mois.

L'objet de cette thèse [15] était l'optimisation des fonctions de surveillance et d'identification d'un avion de patrouille maritime. Un tel système constitue un concentré de traitements et de problèmes multi-récepteurs et multi-cibles. Ainsi, si l'avion dispose de nombreux récepteurs (radars, ESM passif, infrarouge, sonar), il doit aussi traiter (extraire, associer, identifier) un très grand nombre de pistes. Les problèmes relatifs à l'association piste-à-piste, mesures à pistes, à l'apport d'informations d'identification, à la combinaison des diverses sources d'information (radar, ESM, infrarouge) ont constitué une partie importante de l'activité. Pour ces problèmes, les capacités d'auto-contrôle sont un volet essentiel. Pratiquement, il est rare de disposer de la variance des mesures et de leurs biais, ou encore des probabilités de détection et de fausses alarmes. Une étape essentielle consiste à estimer la qualité de la fiabilité des observations élémentaires (en entrée du système). L'optimisation de l'observation est au cœur de nombreux problèmes d'estimation de systèmes dynamiques partiellement observés. La difficulté du problème tient au fait qu'il s'agit d'optimiser les mesures pour un système dont l'état (ou la suite des états pour un système markovien) est inconnu. Ceci concerne aussi bien l'optimisation de la trajectoire de l'observateur, que la disposition des capteurs, l'utilisation optimale des ressources et la gestion de facteurs d'indiscrétions. Une partie de ces problèmes a été analysée de façon satisfaisante. Les résultats obtenus mettent en évidence l'importance de l'information d'identification (radar, ESM). L'accroissement des performances (en termes de probabilité d'association correcte) a pu être quantifié. Un autre volet important est relatif à l'utilisation des données ESM (capteur passif) pour les fonctions de pistage. L'étude des performances des méthodes de pistage a été poursuivie. On doit mentionner aussi les tests d'acceptation de pistes, la résolution de problèmes d'association par des méthodes combinatoires (mise en œuvre de la méthode de Munkres, des "enchères"). Des modélisations de la situation opérationnelle per-

mettent d'intégrer les aspects probabilistes et causaux du développement d'une mission et d'y associer les informations perçues par les capteurs et les opérateurs.

7.6 Convention Onera :

Participant : Jean-Pierre Le Cadre.

no. CNRS 550227, durée 6 mois.

Dans un cadre compétitif, l'optimisation des moyens d'exploration paraît d'autant plus importante que la valeur des informations recueillies est fondamentale pour l'évaluation de la situation. Tandis que les méthodes classiques d'optimisation de l'effort de recherche ("search theory") visent, avant tout, à optimiser une fonction de détection, il s'agit plus en l'occurrence de déterminer la séquence de répartition des efforts de reconnaissance qui assure la meilleure description de l'environnement. L'optimisation de telles fonctionnelles pose cependant des problèmes spécifiques, car les fonctionnelles considérées sont des fonctions non séparables des variables élémentaires. En fait, il paraît impossible de déterminer directement la fonction duale alors que ceci est le point central des méthodes d'optimisation de l'effort de recherche. Nous avons développé une méthode de type "plans sécants". Une extension naturelle en est l'optimisation de la reconnaissance multipériodes, pour laquelle nous avons étendu les algorithmes de type forward-backward (Brown). Bien que ce soit un problème d'optimisation difficile (non séparable, non convexe, très grand nombre de variables), l'algorithme permet d'atteindre la solution en un temps raisonnable. La prise en compte des capacités de réaction de la cible a constitué l'autre volet. Le comportement des cibles intéressantes n'est pas neutre ; il est tout à fait vraisemblable de les supposer réactives aux efforts de reconnaissance. Cette situation se rencontre aussi dans les "dynamic search games" où la cible a une parfaite connaissance de la séquence des cellules cherchées mais peut miser sur l'épuisement des ressources du chercheur. Certains de ces problèmes ne relèvent pas du formalisme des jeux, c'est le cas si la stratégie de la cible est déterministe, alors que d'autres se situent exactement dans un cadre classique de jeux. Nous avons considéré plus particulièrement ces problèmes dans un contexte de réseaux de topographies connues.

spatio-temporelles pour l'indexation vidéo

7.7 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 concerne la thèse de Emmanuel Veneau, encadrée à l'Ina par Laurent Vinet, thèse qui a débuté en mars 1999. Cette étude s'inscrit dans le domaine de l'indexation de vidéos par le contenu. Elle concerne 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 a porté sur une structuration temporelle, de niveau intermédiaire, de la vidéo, le regroupement des plans élémentaires détectés en

séquences devant représenter des entités plus pertinentes et plus exploitables pour la phase de reconnaissance d'événements spatio-temporels. Une méthode combinant plusieurs critères, couleur, durée, mouvement, a été élaborée. Actuellement, nous nous intéressons à la caractérisation de l'activité, dans les plans ou les séquences extraites, par la construction de cartes d'"historiques" issues de mesures spatio-temporelles locales et par la définition de schémas de classification hiérarchique par SVM ("Support Vector Machines"). Cette approche est appliquée à des documents vidéos de retransmission d'événements sportifs, en l'occurrence des championnats d'athlétisme, qui forment une étude de cas représentative et complexe.

7.8 Actions financées par la Commission Européenne

7.8.1 Projet IST Carsense

Participants : Patrick Bouthemy, Eric Marchand, Gildas Lefaix.

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 utilisera des capteurs de différents types (caméras vidéo, laser et radar). Les informations extraites des signaux et images issus de ces capteurs seront 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. L'idée de base 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.

7.9 Actions financées par le RNRT

7.9.1 Projet Agir

Participants : Patrick Bouthemy, Patrick Gros, Fabien Spindler, Abdeljabar Benzougar, Emmanuel Veneau.

no. Inria 2 99 C 006, durée 18 mois

Le projet AGIR est un projet RNRT précompétitif qui a officiellement débuté au 1er septembre 1999. Il regroupe l'Ina (maître d'œuvre), les sociétés CS-Systèmes d'Information, Arts Vidéo Interactive et Mémodata, pour l'Irisa les projets Sigma-2 (Metiss) et Vista, le projet Movi de l'Inria Rhône-Alpes, et des équipes de recherche de l'Irit, du Clips-Imag, de l'INT, et du LIP6. Ce projet a été financé pour 18 mois, la réunion de recette du projet par le ministère

a eu lieu en septembre 2001. La mise au point de la deuxième phase du projet est en cours et donnera en fait lieu à la soumission d'un nouveau projet de 24 mois dans le cadre du programme RIAM. Le but du projet est de développer un système complet permettant d'indexer des vidéos, éventuellement accompagnées de textes, puis d'utiliser cette indexation pour rechercher des documents par leur contenu, ceci en combinant les divers médias. La première phase du projet a consisté à développer un "framework" pour le développement d'applications d'indexation de recherche et de manipulation de documents multimédias. Les travaux développés par Vista concernaient principalement la structuration temporelle automatique des séquences d'images par l'utilisation directe des informations issues de la compression MPEG-2, [44], comme les vecteurs de mouvement, l'extraction et le suivi d'entités mobiles pertinentes, et la fourniture de descripteurs d'activité et de mouvement. Le projet est aussi intervenu en collaboration avec l'Ina dans la mise au point d'algorithmes de macro-segmentation et de regroupement de plans en entités de plus haut niveau plus significatives. Par ailleurs, les sorties des modules développés sont exploitées dans d'autres modules du projet.

7.10 Actions financées par le RNTL

7.10.1 Projet Domus Videum

Participants : Patrick Bouthemy, Nathalie Peyrard.

labélisé, durée 30 mois

Le projet Domus Videum est un projet qui vient d'être labellisé par le 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 portera 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.

7.11 Actions soutenues par le programme Priamm

7.11.1 Projet Mediaworks

Participants : Patrick Bouthemy, Patrick Gros, Chafik Kermad.

no. Inria 2 00 A 0408, durée 36 mois

Le projet Mediaworks est un projet labellisé par le programme Priamm et le programme Société de l'information, financé par le Ministère de l'industrie, et qui a débuté au 1er septembre 2000. Ce projet 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. Ses éléments principaux en sont la coopération entre les media texte et image, et la mise au point d'un moteur de recherche sémantique. Pour Vista, le travail consiste à fournir des outils

de structuration automatique en plans et de représentation iconique de ces plans, ainsi que des descripteurs de niveaux d'activités au sein des plans.

8 Actions régionales, nationales et internationales

8.1 Actions régionales

8.1.1 Contrat CPER Cemagref-Inria AEROBIO

Participants : Patrick Bouthemy, Etienne Mémin, 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.

8.1.2 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 vidéo 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 sur 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 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).

8.1.3 Contrat CRITT Santé : Environnement d'acquisition et de traitement d'échographies 3D

Participants : Christian Barillot, François Rousseau.

no. Inria 4 00 C 0468 00, durée 12 mois.

L'objectif de ce projet est de mettre en place, notamment en site hospitalier (en lien avec le Prof. Pierre Darnault, CHR Pontchaillou, Rennes), un environnement d'acquisition et de traitement d'échographies 3D et de développer des outils appropriés de calibrage et segmentation d'images échographiques 3D et 3D+t. Nous comptons utiliser cet environnement dans le contexte d'applications périnatales (orthopédie du nourrisson, imagerie trans-fontanelle) et d'applications vasculaires périphériques (carotide,...).

8.2 Actions nationales

8.2.1 ACI Ministère de la Recherche : projet NeuroBase

Participants : Christian Barillot, Patrick Gros, Isabelle Corouge.

durée 24 mois.

Nous sommes responsables d'une action concertée incitative (ACI) du Ministère de la recherche qui vient d'être labellisée et est intitulée "NeuroBase : Système de gestion de données et de connaissances réparties en neuro-imagerie". L'objectif est de faire coopérer, au travers de l'Internet, des bases d'informations en neuro-imagerie situées dans différents centres d'expérimentation, cliniques neurologiques ou établissements de recherche en neurosciences cognitives. Ce projet consistera à spécifier la façon de relier de telles bases et d'y accéder efficacement par la définition d'une architecture informatique permettant 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 Temics 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 Epidaur 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 ACI jeunes chercheurs : Techniques d'exploitation des données multimédias

Participants : Patrick Gros, Sid Ahmed Berrani.

durée 36 mois.

Ce programme du ministère de la recherche a pour but d'aider et d'accompagner la création de nouvelles équipes par de jeunes chercheurs. Le projet qui a été accepté vise à monter une nouvelle équipe sur le thème de l'indexation, de la recherche et plus largement de l'exploitation de grandes bases de documents multimédias. Cette équipe sera effectivement créée à l'Irisa sous forme de l'action TEXMEX au 1er janvier 2002.

8.2.3 Actions de recherche coopératives de l’Inria

- Imagerie cérébrale

Participants : Christian Barillot, Pierre Hellier.

Nous avons participé à deux actions de recherche coopérative entre l’Inria (projets Epidaure, Robotvis et Vista) et l’IFR de neuroimagerie 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, visent à développer de nouvelles méthodes de traitement d’image et de recalage en imagerie fonctionnelle cérébrale. L’une est 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 6.2.3).

8.3 Relations bilatérales internationales

8.3.1 Programme Alliance France-Angleterre

Participants : Christian Barillot, Patrick Bouthemy, François Rousseau.

Cette action est menée avec A. Gee et R. Prager de l’Université de Cambridge. Nous nous intéressons à l’étude d’images ultrasonores 3D acquises dans un mode de type “main libre” pour différents domaines d’application, comme la radiologie chez le nourrisson ou le contrôle par l’image per-opératoire ultrasonore de l’asservissement et du guidage de gestes chirurgicaux dans un environnement de neuro-navigation (en liaison avec des partenaires cliniques à Rennes et à Cambridge). Grâce à cette collaboration, nous mettons en place sur le site de Rennes une plate-forme expérimentale d’acquisition d’échographies 3D. Ce dispositif repose sur la localisation de la sonde échographique par un capteur magnétique, le tout étant relié à une station de travail PC Linux. La technique de reconstruction du volume échographique utilisée est le logiciel StradX développé par Cambridge. Notre contribution a d’une part été de mettre au point un ensemble de tests calibrés pour évaluer différentes solutions de localisation électromagnétique. D’autre part, nous travaillons à améliorer la procédure de calibrage d’un examen échographique 3D afin de la rendre la plus possible automatique et transparente à l’utilisateur.

8.3.2 Programme Procope France-Allemagne

Participants : Patrick Bouthemy, Etienne Mémin, Thomas Corpetti.

Nous avons mis en place, dans le cadre du programme d’actions intégrées franco-allemand Procope, une coopération avec le groupe du Professeur 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. T. Corpetti a effectué un séjour de deux semaines à Mannheim, E. Mémin et P. Bouthemy une visite de quelques jours.

8.3.3 Programme France-Tunisie

Participants : Patrick Bouthemy, Patrick Gros, Ronan Fablet.

Nous participons à la collaboration franco-tunisienne mise en place par l'Inria et l'École supérieure de Communications de Tunis. Le projet retenu concerne l'indexation multimédia et regroupe les projets Imedia de l'Inria Rocquencourt, Movi de l'Inria Rhône-Alpes et Vista. Dans le cadre de ce projet, P. Gros et R. Fablet ont participé au module de cours « Indexation multimédia » du DEA de cette école lors d'un séjour en Tunisie en juin 2001. P. Gros a aussi été invité au colloque TAIMA'2001, qui s'est tenu en octobre 2001 à Hammamet, pour présenter une communication sur les problèmes posés par l'indexation d'images vus du côté des bases de données.

8.3.4 Accueil de chercheurs étrangers

- Prof. Bruno Cernuschi-Frias, Université de Buenos-Aires, effectue un séjour de douze mois dans le projet à compter du 15 septembre 2001, sur un poste d'accueil Inria de spécialiste académique étranger (cf paragraphe 6.2.2).
- Mario Ramalho, maître de conférence à l'IST de Lisbonne, a effectué un séjour de six mois au sein du projet jusqu'au 1er mars 2001. Il a développé une méthode de suivi de points d'intérêt dans une séquence d'images basée sur les techniques de corrélation.
- Juanma Sanchez, doctorant à l'UAB, Barcelone, sous la direction du Prof. Xavier Binefa, a effectué un séjour d'un mois, de mi-mai à mi-juin, sur le sujet de l'indexation vidéo par le contenu.
- Prof. Christoph Schnörr et Daniel Cremers (Université de Mannheim, Allemagne) ont effectué respectivement une visite de quelques jours, en octobre, et un séjour de deux semaines et demie en septembre-octobre (cf paragraphe 8.3.2).
- Alberto Del Bimbo (Univ. Florence), Philippe Mulhem (IPAL, Singapour), Lyn Thomas (Univ. Southampton), Zheng Wentao (NHK) ont effectué des visites de courte durée au sein du projet.

8.3.5 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" ; évaluation de soumissions de projets en janvier 2000 à Bruxelles.
- 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 1998-2002, 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 et P. Gros ont chacun effectué une expertise pour l'Anvar Ile-de-France dans le domaine de l'indexation vidéo.
- F. Spindler est membre de la commission d'évaluation des ingénieurs de l'Inra sur la période 2001-2002, a participé au jury du concours 2001 d'ingénieurs de recherche de l'Inra.

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.
- P. Gros est animateur du groupe de travail CIDRE ("Content-based Indexing and REtrieval"). Ce groupe, d'abord interne à l'Irisa, est maintenant ouvert à la communauté rennaise et rassemble, outre des chercheurs de différents projets de l'Irisa intéressés par l'indexation multimedia, des personnes de Thomson Multimedia, de l'université de Rennes 2 et de France Télécom R&D. Le groupe de travail est concerné par les problèmes posés par l'indexation des documents multimédias et par le stockage de descripteurs numériques en grande quantité. Il se réunit tous les mois et comprend des spécialistes de médias (images fixes et animées, son et parole, documents, texte) ainsi que d'autres issus des communautés des bases de données et de la recherche d'information.
- Participation au GDR ISIS "Information, signal et images" : le projet est membre du GDR ISIS et P. Gros en est le correspondant.
- 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.
- P. Gros a organisé deux des quatre ateliers de la journée Inria-Industrie sur la "Recherche d'information dans les documents électroniques" qui s'est tenue le 6 décembre 2001 à Rocquencourt.
- 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*.
- C. Barillot est membre du comité de rédaction de la revue *Technique et Science Informatiques* (TSI).
- P. Gros est membre du comité de rédaction de la revue *Traitement du Signal*.
- Participation aux Comités de Programme :
 - C. Barillot : conférences internationales ICIP'2001, IPMI'2001, MICCAI'2001.
 - P. Bouthemy : conférences internationales CVPR'2001, TAIMA'2001, ICPR'2002, SSIAT'2002,

- congrès national RFIA'2002.
- F. Chaumette : conférences internationales ICCV'2001, SIRS'2001, ECCV'2002.
- J.-P. Le Cadre : conférence internationale Fusion'2001.
- P. Gros : journées nationales Orasis'2001.
- P. Hellier : conférence internationale MICCAI'2001.
- E. Mémin : conférence internationale ICCV'2001.
- L'article "Retrospective evaluation of inter-subject brain registration", [65], présenté à la conférence MICCAI'2001 par P. Hellier, a été sélectionné pour l'obtention du prix du "meilleur article étudiant" de la conférence (5 articles sélectionnés sur 136).
- P. Hellier a obtenu une Mention spéciale au Prix Jeune Chercheur 2001 du Conseil régional de Bretagne dans la catégorie "Structure et propriétés de la matière".

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).
- DIIC INC, Ifsic, Rennes 1 : (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).
- Insa Rennes, Génie électrique 5ème année, option Image (F. Chaumette, I. Corouge, T. Corpetti, O. Tahri : vision par ordinateur).
- ENSAI, Rennes, 3ème année : (F. Dambreville, E. Mémin, F. Dekeyser : modèles statistiques pour l'image).
- ENPC Marne-La-Vallée, 2ème année (F. Chaumette : couplage vision-commande en robotique).
- École Polytechnique, 3ème année (P. Gros : synthèse et analyse d'images).
- ENSTA Paris, 3ème année (P. Gros : géométrie projective et vision 3D).
- 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 de l'École Supérieure de Communications de Tunis, option Images (P. Gros, R. Fablet : indexation d'images fixes, indexation de vidéos).
- Stages effectués au sein de l'équipe : N. David (CNAM, Marne-la-Vallée), G. Gouzien (Insa Toulouse), S. Hammer (ENST Paris), F. Hubert (DEA Stir, Rennes), C. Lacoste (Insa et DEA Toulouse), C. Larboulette (DEA Info, Rennes), M.-C. Pitis (ENSAI, Rennes), A. Remazeilles (DEA Info, Rennes), J. Royan (Insa et DEA Info, Rennes), M. Saunier (Mastère ENSTBr, Brest).
- Direction de thèse hors projet :
 - J.-N. Provost (Ecole Navale, Brest) par P. Bouthemy, thèse soutenue en juin 2001 à l'UBO, Brest, titre : "Classification bathymétrique en imagerie multispectrale SPOT" ;

- A. Alhaj (Cemagref, Rennes) par F. Chaumette ;
- F. Bavencov (Thales Airborne Systems, bourse Cifre) par J.-P. Le Cadre ;
- M. Chakchouk (co-tutelle Tunis, Paris 5) par P. Bouthemy ;
- E. Kijak (Thomson Multimedia, bourse Cifre) par P. Gros.

9.3 Participation à des colloques, séminaires, invitations

- C. Barillot a donné un tutoriel invité intitulé “Linear versus non-linear registration of brainscans - what can we expect? Some evaluation issues”, dans le cadre d’une école à Dagstuhl, Allemagne, en mars 2001.
- C. Barillot a donné un tutoriel invité intitulé “Deformable Registration”, dans le cadre de la conférence internationale IPMI qui s’est tenue à Davis, USA en juin 2001.
- P. Bouthemy a participé à la rencontre HKU-Inria qui s’est tenue à Hong-Kong en avril 2001 et a donné un exposé intitulé “Parametric and non-parametric approaches for image motion modeling, segmentation, recognition, and indexing”.
- P. Bouthemy et P. Gros ont été invités à participer au workshop MMBIR (“Multimedia Content-based Indexing and Retrieval”) co-organisé par NSF, IBM, UC Berkely et l’Inria à Rocquencourt en septembre 2001, et ont donné un exposé respectivement sur “Statistical modeling for motion-based video classification and retrieval” et “Content-based retrieval based on local descriptors : problems and issues from a database perspective”.
- F. Chaumette a participé à l’Ecole d’été franco-mexicaine SSIR’2001 organisée à Cuernavaca (Mexique) en juin 2001 portant sur la vision et la robotique, en tant qu’intervenant sur le sujet de l’asservissement visuel.
- F. Chaumette a été invité pendant une semaine à l’Université de Girona (Espagne) en mai 2001, notamment pour y donner des cours en asservissement visuel.
- F. Dekeyser a participé au workshop Inria-NSC Taiwan qui s’est tenu à Taiwan en février 2001, et a donné un exposé intitulé “Spatio-temporal filtering of image sequences and super-resolution, application to key-frame construction for video indexing”.
- P. Gros a donné une conférence invitée intitulée “Les invariants : introduction et application à la vision et à la robotique” aux troisièmes journées nationales de la recherche en robotique (JNRR) qui se sont tenues à Giens en octobre 2001.
- C. Hue a effectué un séjour de 3 mois (avril, mai, juin 2001) à Microsoft Research à Cambridge, Angleterre, sous la direction de Patrick Pérez.
- P. Hellier a effectué un post-doctorat (bourse Inria) de 10 mois à l’Université d’Utrecht sous la direction de Max Viergever, R. Fablet a entrepris un post-doctorat (bourse Inria) à Brown University sous la direction de Michael Black, Y. Mezouar (financement local) à Columbia University sous la direction de Peter Allen et F. Dambreville (bourse DGA) au Naval Postgraduate School à Monterey sous la direction de Alan Washburn.

10 Bibliographie

Ouvrages et articles de référence de l'équipe

- [1] B. ESPIAU, F. CHAUMETTE, P. RIVES, « A new approach to visual servoing in robotics », *IEEE Transactions on Robotics and Automation* 8, 6, juin 1992, p. 313–326.
- [2] H. GAUVRIT, 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.
- [3] F. HEITZ, P. PÉREZ, P. BOUTHEMY, « Multiscale minimization of global energy functions in some visual recovery problems », *CVGIP : Image Understanding* 59, 1, 1994, p. 125–134.
- [4] J.-M. LAFERTÉ, P. PÉREZ, F. HEITZ, « Discrete Markov modeling and inference on the quad-tree », *IEEE Transactions on Image Processing* 9, 3, mars 2000, p. 390–404.
- [5] J.-P. LE CADRE, C. JAUFFRET, « Discrete-time observability and estimability for bearings-only target motion analysis », *IEEE Trans. on Aerospace & Electr. Systems* 33, 1, janvier 1997, p. 178–201.
- [6] G. LE GOUALHER, C. BARILLOT, Y. BIZAIS, « Modeling cortical sulci with active ribbons », *Int. Journal of Pattern Recognition and Artificial Intelligence* 8, 11, 1997, p. 1295–1315.
- [7] E. MARCHAND, F. CHAUMETTE, « An autonomous active vision system for complete and accurate 3D scene reconstruction », *Int. Journal of Computer Vision* 32, 3, août 1999, p. 171–194.
- [8] E. MÉMIN, P. PÉREZ, « Dense estimation and object-based segmentation of the optical flow with robust techniques », *IEEE Trans. on Image Processing* 7, 5, mai 1998, p. 703–719.
- [9] J.-M. ODOBEZ, P. BOUTHEMY, « Robust multiresolution estimation of parametric motion models », *Journal of Visual Communication and Image Representation* 6, 4, décembre 1995, p. 348–365.
- [10] V. SUNDARESWARAN, P. BOUTHEMY, F. CHAUMETTE, « Exploiting image motion for active vision in a visual servoing framework », *Int. Journal of Robotics Research* 15, 6, décembre 1996, p. 629–645.

Thèses et habilitations à diriger des recherches

- [11] F. DAMBREVILLE, *Optimisation de la gestion des capteurs et des informations pour un système de détection*, thèse de doctorat, Université de Rennes 1, mention Traitement du signal et télécommunications, novembre 2001.
- [12] F. DEKEYSER, *Restauration de séquences d'images par des approches spatio-temporelles : filtrage et super-résolution par le mouvement*, thèse de doctorat, Université de Rennes 1, mention Traitement du signal et télécommunications, novembre 2001.
- [13] R. FABLET, *Modélisation statistique non paramétrique et reconnaissance du mouvement dans des séquences d'images ; application à l'indexation vidéo*, thèse de doctorat, Université de Rennes 1, mention Traitement du signal et télécommunications, juillet 2001.
- [14] G. FLANDIN, *Modélisation probabiliste et exploration visuelle autonome pour la reconstruction de scènes inconnues*, thèse de doctorat, Université de Rennes 1, mention Informatique, novembre 2001.
- [15] I. LEIBOWICZ, *Traitements multicapteurs pour les systèmes aéroportés de patrouille maritime*, thèse de doctorat, Université de Rennes 1, mention Traitement du signal et télécommunications, février 2001.

- [16] Y. MEZOUAR, *Planification de trajectoires pour l'asservissement visuel*, thèse de doctorat, Université de Rennes 1, mention Informatique, novembre 2001.

Articles et chapitres de livre

- [17] L. AMSALEG, P. GROS, « Content-based retrieval using local descriptors : problems and issues from a database perspective », *Pattern Analysis and Applications* 4, 2/3, 2001, p. 108–124.
- [18] C. BAILLARD, P. HELLIER, C. BARILLOT, « Segmentation of brain 3D MR images using level sets and dense registration », *Medical Image Analysis* 5, 3, septembre 2001, p. 185–194.
- [19] F. CHAUMETTE, « Asservissement visuel », in : *Les bases de la robotique*, E. Dombre (éditeur), *Traité IC2*, Hermes, ch. 7, à paraître.
- [20] F. CHAUMETTE, E. MARCHAND, « A redundancy-based iterative approach for avoiding joint limits : application to visual servoing », *IEEE Trans. on Robotics and Automation* 17, 5, octobre 2001.
- [21] C. COLLET, J.-N. PROVOST, P. ROSTAING, P. PÉREZ, P. BOUTHEMY, « Segmentation bathymétrique d'images multispectrales SPOT », *Traitement du Signal* 18, 1, janvier 2001, p. 1–16.
- [22] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Dense estimation of fluid flows », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, à paraître.
- [23] N. COURTY, E. MARCHAND, « Navigation et contrôle d'une caméra dans un environnement virtuel : une approche référencée image. », *Technique et Science Informatiques, TSI* 20, 6, juin 2001, p. 779–803.
- [24] A. CRÉTUAL, F. CHAUMETTE, « Application of motion-based visual servoing to target tracking », *Int. Journal of Robotics Research* 20, 11, novembre 2001.
- [25] A. CRÉTUAL, F. CHAUMETTE, « Visual servoing based on image motion », *Int. Journal of Robotics Research* 20, 11, novembre 2001.
- [26] F. DAMBREVILLE, J.-P. LE CADRE, « Detection of a Markovian target with optimization of the search efforts under generalized linear constraints », *Naval Research Logistics*, à paraître.
- [27] R. DONATI, J.-P. LE CADRE, « Detection of oceanic electric fields based on the GLRT », *IEE Proc. Radar, Sonar and Navigation*, à paraître.
- [28] R. FABLET, P. BOUTHEMY, P. PÉREZ, « Non parametric motion characterization using causal probabilistic models for video indexing and retrieval », *IEEE Trans. on Image Processing*, à paraître.
- [29] R. FABLET, P. BOUTHEMY, « Non-parametric scene activity analysis for statistical retrieval with partial query », *Journal of Mathematical Imaging and Vision* 14, 3, mai 2001, p. 257–270.
- [30] P. GROS, R. FABLET, P. BOUTHEMY, « New descriptors for image and video indexing », in : *State-of-the-Art in Content-Based Image and Video Retrieval*, R. Veltkamp, H. Burkhardt, et H.-P. Kriegel (éditeurs), *Computational Imaging and Vision, Vol.22*, Kluwer Academic Publ., ch. 10, p. 213–234, 2001.
- [31] 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.
- [32] L. HUBERT-MOY, A. COTONNEC, L. LE DU, A. CHARDIN, P. PÉREZ, « A comparison of parametric classification procedures of remotely sensed data applied on different landscape units », *Remote Sensing of Environment* 75, 2, février 2001, p. 174–187.

- [33] C. HUE, J.-P. LE CADRE, P. PÉREZ, « The (MR)MTPF : particle filters to track multiple targets using multiple receivers », *IEEE Trans. on Signal Processing*, à paraître.
- [34] B. LAMIROY, P. GROS, S. PICARD, « Combining local recognition methods for better image recognition », *Vision 17*, 2, Second Quarter 2001, p. 1–6.
- [35] 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*, à paraître.
- [36] E. MALIS, G. MOREL, F. CHAUMETTE, « Robot control from disparate multiple sensors », *Int. Journal of Robotics Research* 20, 5, mai 2001, p. 364–378.
- [37] E. MARCHAND, P. BOUTHEMY, F. CHAUMETTE, « A 2D-3D model-based approach to real-time visual tracking », *Image and Vision Computing* 19, 13, novembre 2001, p. 941–955.
- [38] E. MARCHAND, N. COURTY, « Controlling a camera in a virtual environment. », *The Visual Computer Journal*, à paraître.
- [39] E. MÉMIN, P. PÉREZ, « Hierarchical estimation and segmentation of dense motion fields », *Int. Journal of Computer Vision*, à paraître.
- [40] E. MÉMIN, T. RISSET, « VLSI design methodology for edge-preserving image reconstruction », *Real Time Imaging* 7, 1, février 2001, p. 109–126.
- [41] 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*, à paraître.

Communications à des congrès, colloques, etc.

- [42] L. AMSALEG, P. GROS, S. BERRANI, « A robust technique to recognize objects in images, and the DB problems it raises », in : *Proc. of the Workshop on Multimedia Information Systems*, Capri, novembre 2001.
- [43] L. AMSALEG, P. GROS, S. BERRANI, « A robust technique to recognize objects in images, and the DB problems it raises », in : *Actes des ateliers Traitement et analyse d'images : méthodes et applications, TAIMA'2001*, Hammamet, octobre 2001.
- [44] A. BENZOUGAR, P. BOUTHEMY, R. FABLET, « MRF-based moving object detection from MPEG coded video », in : *IEEE Int. Conf. on Image Processing, ICIP'2001*, p. 402–405, Thessalonique, octobre 2001.
- [45] F. CHAUMETTE, A. CRÉTUAL, « Dynamic stabilization of a pan and tilt camera for sub-marine image visualization », in : *IARP Workshop on Underwater Robotics for Sea Exploitation and Environmental Monitoring*, p. 146–157, Rio de Janeiro, octobre 2001.
- [46] C. COLLEWET, F. CHAUMETTE, P. LOISEL, « Image-based visual servoing on planar objects of unknown shape », in : *IEEE Int. Conf. on Robotics and Automation*, 1, p. 247–252, Séoul, mai 2001.
- [47] I. COROUGE, C. BARILLOT, P. HELLIER, P. TOULOUSE, B. GIBAUD, « Non-linear local registration of functional data », in : *Medical Image Computing and Computer Assisted Intervention 2001 (MICCAI'01)*, M.-A. Viergever, T. Dohi, M. Vannier (éditeurs), *Lecture Notes in Computer Sciences 2208*, Springer-Verlag, p. 948–956, Utrecht, octobre 2001.
- [48] I. COROUGE, C. BARILLOT, « Use of a probabilistic shape model for non-linear registration of 3D scattered data », in : *IEEE Int. Conf. on Image Processing, ICIP'2001*, p. 149–152, Thessalonique, octobre 2001.

- [49] I. COROUGE, C. BARILLOT, « Utilisation d'un modèle probabiliste de formes pour le recalage non-linéaire de données 3D éparses », *in : Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle, ORASIS'2001*, p. 145–154, Cahors, juin 2001.
- [50] I. COROUGE, P. HELLIER, P. TOULOUSE, B. GIBAUD, C. BARILLOT, « A new inter-individual functional mapping scheme based on a local non linear registration method », *in : Proc. of Human Brain Mapping, HBM'2001, Neuroimage, 16, 3*, p. 100–100, Brighton, juin 2001.
- [51] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Estimation de mouvement fluide basée sur l'équation de continuité associée à une régularisation Div-Curl », *in : Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle, ORASIS'2001*, p. 481–490, Cahors, juin 2001.
- [52] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Estimation dense du mouvement en imagerie fluide », *in : 9ème Colloque Francophone de Visualisation et de Traitement d'Images en Mécanique des Fluides, FLUVISU'2001*, Rouen, juin 2001.
- [53] N. COURTY, E. MARCHAND, B. ARNALDI, « Through-the-eyes control of a virtual humanoïd », *in : IEEE Int. Conf. on Computer Animation 2001*, H.-S. Ko, N. Magnenat-Thalmann (éditeurs), Séoul, novembre 2001.
- [54] N. COURTY, E. MARCHAND, « Computer animation : a new application for image-based visual servoing », *in : IEEE Int. Conf. on Robotics and Automation, 1*, p. 223–228, Séoul, mai 2001.
- [55] F. DAMBREVILLE, J.-P. LE CADRE, « Distribution of continuous search effort for the detection of a target with optimal moving strategy », *in : SPIE Conf. on Data and Signal Processing of Small Targets*, San Diego, juillet 2001.
- [56] F. DAMBREVILLE, J.-P. LE CADRE, « Optimal distribution of continuous search effort for detection of a target in a min-max game context », *in : 4th International Conference on Information Fusion, FUSION'01*, Montréal, août 2001.
- [57] F. DEKEYSER, P. BOUTHEMY, P. PÉREZ, « A new algorithm for super-resolution from image sequences », *in : 9th International Conference on Computer Analysis of Images and Patterns, CAIP'2001, LNCS Vol.2124, Springer*, W. Sharbek (éditeur), p. 473–481, Varsovie, septembre 2001.
- [58] F. DEKEYSER, P. BOUTHEMY, P. PÉREZ, « Une approche multi-échelle de super-résolution à partir de séquences d'images », *in : Colloque GRETSI'2001*, Toulouse, septembre 2001.
- [59] R. FABLET, P. BOUTHEMY, « Motion recognition using spatio-temporal random walks in sequence of 2D motion-related measurements », *in : IEEE Int. Conf. on Image Processing, ICIP'2001*, p. 652–655, Thessalonique, octobre 2001.
- [60] R. FABLET, P. BOUTHEMY, « Non parametric motion recognition using temporal multiscale Gibbs models. », *in : IEEE Int. Conference on Computer Vision and Pattern Recognition, CVPR '01*, Kauai Marriott, Hawaii, décembre 2001.
- [61] G. FLANDIN, F. CHAUMETTE, « Autonomous visual exploration of complex objects », *in : IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'01*, Maui, Hawaii, octobre 2001.
- [62] G. FLANDIN, F. CHAUMETTE, « Fusion d'informations visuelles pour la localisation d'objets complexes », *in : Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle, ORASIS'2001*, p. 233–242, Cahors, juin 2001.
- [63] G. FLANDIN, F. CHAUMETTE, « Vision-based control using probabilistic geometry for objects representation », *in : IEEE Int. conf. on Decision and Control, CDC'01*, Orlando, décembre 2001.

- [64] P. GROS, A. REMAZEILLES, F. CHAUMETTE, « Les invariants : introduction et application à la vision et à la robotique », *in : Actes des troisièmes journées nationales de la recherche en robotique*, Giens, octobre 2001.
- [65] P. HELLIER, C. BARILLOT, I. COROUGE, B. GIBAUD, LE GOUALHER, G., L. COLLINS, A. EVANS, G. MALANDAIN, N. AYACHE, « Retrospective evaluation of inter-subject brain registration », *in : Medical Image Computing and Computer Assisted Intervention 2001, MICCAI'01*, M.-A. Viergever, T. Dohi, M. Vannier (éditeurs), *Lecture Notes in Computer Sciences 2208*, Springer-Verlag, p. 258–265, Utrecht, octobre 2001.
- [66] P. HELLIER, C. BARILLOT, « Cooperation between local and global approaches to register brain images », *in : Int. Conf. on Information Processing in Medical Imaging, IPMI'01*, M.-F. Insana, R.-M. Leahy (éditeurs), *Lecture Notes in Computer Science 2082*, p. 315–328, Davis, juin 2001.
- [67] C. HUE, J.-P. LE CADRE, P. PÉREZ, « The (MR)MTPF :particle filters to track multiple targets using multiple receivers », *in : 4th Int. Conf. on Information Fusion, FUSION'01*, Montréal, août 2001.
- [68] C. HUE, J.-P. LE CADRE, P. PÉREZ, « A particle filter to track multiple objects », *in : IEEE Workshop on Multi-Object Tracking*, p. 61–68, Vancouver, juillet 2001.
- [69] C. HUE, J.-P. LE CADRE, P. PÉREZ, « Tracking multiple objects with particle filtering using multiple receivers », *in : International Seminar on Target Tracking, Algorithms & Applications*, p. 61–64, Twente, octobre 2001.
- [70] C. HUE, J.-P. LE CADRE, P. PÉREZ, « Trajectographie multi-objets par filtrage particulaire », *in : Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle, ORASIS'2001*, p. 511–520, Cahors, juin 2001.
- [71] J.-P. LE CADRE, « Allocation of search effort to optimize information », *in : 4th International Conference on Information Fusion, FUSION'01*, Montréal, août 2001.
- [72] J.-P. LE CADRE, « Approximations de la probabilité de détection d'une cible mobile », *in : Colloque GRETSI sur le Traitement du Signal et des Images*, Toulouse, septembre 2001.
- [73] J.-F. LOTS, D. LANE, E. TRUCCO, F. CHAUMETTE, « A 2-D visual servoing for underwater vehicle station keeping », *in : IEEE Int. Conf. on Robotics and Automation, 3*, p. 2767–2772, Séoul, mai 2001.
- [74] E. MARCHAND, F. CHAUMETTE, F. SPINDLER, M. PERRIER, « Controlling an uninstrumented ROV manipulator by visual servoing. », *in : MTS/IEEE OCEANS 2001 Conference*, p. 1047–1053, Honolulu, novembre 2001.
- [75] E. MARCHAND, F. CHAUMETTE, F. SPINDLER, M. PERRIER, « Controlling the manipulator of an underwater ROV using a coarse calibrated pan tilt camera », *in : IEEE Int. Conf. on Robotics and Automation, 3*, p. 2773–2778, Séoul, 2001.
- [76] Y. MEZOUAR, F. CHAUMETTE, « Avoiding self-occlusions and preserving visibility by path planning in the image », *in : 9th Int. Symp. on Intelligent Robotic systems, SIRS'2001*, Toulouse, juillet 2001.
- [77] Y. MEZOUAR, F. CHAUMETTE, « Design and tracking of desirable trajectories in the image space by integrating mechanical and visibility constraints », *in : IEEE Int. Conf. on Robotics and Automation, 1*, p. 731–736, Séoul, mai 2001.
- [78] Y. MEZOUAR, F. CHAUMETTE, « Model-free optimal trajectories in the image space : application to robot vision control », *in : IEEE Int. Conference on Computer Vision and Pattern Recognition, CVPR '01*, Kauai Marriott, Hawaï, décembre 2001.
- [79] Y. MEZOUAR, F. CHAUMETTE, « Visual servoing by path planning », *in : European Conf. on Control, ECC'01*, p. 2904–2909, Porto, septembre 2001.

- [80] I. PRATIKAKIS, C. BARILLOT, P. HELLIER, « Robust multi-scale non-rigid registration of 3D ultrasound images », in : *3rd Int. Conf. on Scale-Space and Morphology in Computer Vision, Scale-Space'01, LNCS 2106*, M. Kerckhove (éditeur), p. 389–397, Vancouver, juillet 2001.

Rapports de recherche et publications internes

- [81] F. DAMBREVILLE, J.-P. LE CADRE, « Minimax optimization of continuous search efforts for the detection of a target », *rapport de recherche n°1403*, Irisa, juin 2001, <http://www.irisa.fr/bibli/publi/pi/2001/1403/1403.html>.
- [82] G. FLANDIN, F. CHAUMETTE, « Visual data fusion : application to objects localization and exploration », *rapport de recherche n°1394*, Irisa, avril 2001, <http://www.irisa.fr/bibli/publi/pi/2001/1394/1394.html>.
- [83] E. MARCHAND, F. CHAUMETTE, « A new formulation for non-linear camera calibration using virtual visual servoing », *rapport de recherche n°1366*, Irisa, janvier 2001, <http://www.irisa.fr/bibli/publi/pi/2001/1366/1366.html>.
- [84] Y. MEZOUAR, F. CHAUMETTE, « Path planning for robust image-based visual servoing », *rapport de recherche n°1377*, Irisa, janvier 2001, <http://www.irisa.fr/bibli/publi/pi/2001/1377/1377.html>.

Divers

- [85] C. BARILLOT, « Linear versus non-linear registration of brainscans - What can we expect ? Some evaluation issue », Dagstuhl Seminar, mars 2001.
- [86] P. GROS, « Comment retrouver une image dans un paquet d'un million d'images... en moins d'un million de secondes ? », Actes des premières rencontres des sciences et technologies de l'information ASTI'2001, Paris, avril 2001.