INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Projet VISTA

Vision Spatio-Temporelle et Active

Rennes

THÈME 3B

Rapport d'Activité

2000

Table des matières

1 Composition de l'équipe **3**

2 Présentation et objectifs généraux **4**

3 Fondements scientifiques **6**
3.1 Modèles markoviens hiérarchiques et estimation bayésienne 6
3.2 Commande référencée vision 8
3.3 Extraction de pistes et trajectographie 10

4 Domaines d'applications **12**

5 Logiciels **13**
5.1 Logiciel RMRm d'estimation du mouvement dominant 13
5.2 Logiciel MD-Shots de découpage d'une vidéo en plans 13
5.3 Plate-forme VISP d'asservissement visuel 14
5.4 Vistal : Librairies pour le traitement et l'analyse d'images volumiques spatio-temporelles 15

6 Résultats nouveaux **15**
6.1 Modèles statistiques, métrologie du mouvement et des déformations 15
6.1.1 Modèles pour la mesure du mouvement fluide 16
6.1.2 Amélioration de séquences d'images et super-résolution 17
6.1.3 Modélisation non paramétrique de l'information spatio-temporelle et reconnaissance du mouvement 18
6.1.4 Modèles probabilistes de formes en imagerie médicale 19
6.1.5 Analyse d'images ultrasonores 20
6.1.6 Structuration et classification de vidéos, indexation de données numériques 21
6.2 Recalage, suivi, trajectographie 22
6.2.1 Suivi en imagerie par filtrage particulaire 23
6.2.2 Suivi de structures déformables en imagerie météorologique 24
6.2.3 Segmentation et recalage 3D non linéaire en imagerie cérébrale 24
6.2.4 Appariements d'objets et de scènes 25
6.2.5 Extraction de pistes et poursuite 26
6.3 Asservissement visuel, perception active et exploration 26
6.3.1 Planification dans l'image 26
6.3.2 Évitement des butées articulaires 27
6.3.3 Utilisation des moments en asservissement visuel 28
6.3.4 Couplage vision locale et vision globale 28
6.3.5 Théorie de l'exploration 29
6.3.6 Asservissement visuel et animation 30
6.3.7 Plate-forme véhicule CyCab 31

7 Contrats industriels (nationaux, européens et internationaux) 31
7.1 Convention DGA - Thomson-CSF-Optrosys : Conception d'une chaîne de traitements pour la restauration de séquences d'images . . . 31
7.2 Convention Adfac-Eumetsat : Estimation de champs de vents sur nuages bas . . 32
7.3 Convention Ifremer : Commande d'un bras manipulateur non instrumenté . . . 32
7.4 Convention Cifre Thomson-RCM : Méthodes et algorithmes de traitements multi-capteurs . . . 33
7.5 Convention Cifre Ina : Reconnaissance de formes spatio-temporelles pour l'indexation vidéo . . . 33
7.6 Actions financées par la Commission Européenne . . . 34
7.6.1 Projet Esprit (R&D MS) Divan . . . 34
7.6.2 Projet IST Carsense . . . 34
7.7 Actions financées par le RNRT . . . 35
7.7.1 Projet Agir . . . 35
7.8 Actions soutenues par le programme Priamm . . . 36
7.8.1 Projet Mediaworks . . . 36

8 Actions régionales, nationales et internationales 36
8.1 Actions régionales . . . 36
8.1.1 Contrat CPER Cemagref-Inria AEROBIO . . . 36
8.1.2 Télé-acquisition robotisée d'échographies 3D . . . 36
8.2 Actions nationales . . . 37
8.2.1 Pôle CNRS Micro-robotique : Asservissement visuel et micro-manipulation 37
8.2.2 Convention GIS Sciences de la Cognition : Imagerie cérébrale . . . 37
8.2.3 Actions de recherche coopératives de l'Inria . . . 37
8.3 Relations bilatérales internationales . . . 38
8.3.1 Programme Alliance France-Angleterre . . . 38
8.3.2 Programme France-Tunisie . . . 39
8.3.3 Accueil de chercheurs étrangers . . . 39
8.3.4 Expertise de projets . . . 39

9 Diffusion de résultats 40
9.1 Animation de la communauté scientifique . . . 40
9.2 Enseignement universitaire . . . 41
9.3 Participation à des colloques, séminaires, invitations . . . 42

10 Bibliographie 42

1 Composition de l'équipe

Responsable scientifique

Patrick Bouthemy [DR Inria]

Assistante de projet

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

Personnel Inria

François Chaumette [CR, puis DR à compter du 01/10/2000]
Isabelle Lacoffrette [TR (Atelier), à mi-temps du 15/01/2000 au 30/06/2000]
Éric Marchand [CR]
Patrick Pérez [CR, jusqu'au 31/03/2000]
Fabien Spindler [IR (Atelier)]

Personnel CNRS

Christian Barillot [CR]
Patrick Gros [CR]
Jean-Pierre Le Cadre [DR]

Personnel Université

Etienne Mémin [maître de conférences Ifsic, Université de Rennes 1]

Personnel non permanent

Abdeljabar Benzougar [ingénieur expert Inria]
Gildas Lefaix [ingénieur expert Inria, à compter du 01/09/2000]
Fabien Boissé [service national, jusqu'au 15/04/2000]

Chercheurs post-doctorants et ATER

Caroline Baillard [post-doc Inria, jusqu'au 31/07/2000]
Christophe Papin [IE-postdoc Inria, jusqu'au 30/04/2000]
Ioannis Pratikakis [post-doc Inria, jusqu'au 15/03/2000]
Stelios Zimeras [post-doc Inria, jusqu'au 30/09/2000]

Chercheurs doctorants

Damien Albert [allocataire MENRT, jusqu'au 30/09/2000]
Sid Ahmed Berrani [boursier TMM, à compter du 01/12/2000]
Annabelle Chardin [boursière Inria-Région, jusqu'au 31/01/2000]
Isabelle Corouge [boursière Inria]
Thomas Corpetti [allocataire MENRT]
Frédéric Dambreville [boursier DGA]
Fabien Dekeyser [boursier DGA]
Ronan Fablet [boursier BDI CNRS-Région]
Grégory Flandin [boursier Inria-Région]
Pierre Hellier [boursier Inria-Région, jusqu'au 31/10/2000]
Carine Hue [allocataire MENRT]
Isabelle Leibowicz [Boursière Cifre, Thomson RCM]
Youcef Mezouar [allocataire MENRT]
François Rousseau [allocataire MENRT, à compter du 01/10/2000]
Omar Tahri [boursier Inria-Région, à compter du 01/10/2000]
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 "métrologie" 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 surveillance (sonar, transports).*

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 de suivi temporel. 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 : *"métrologie" 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 surveillance* (sonar, transports).

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. 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'indexa-

tion de vidéo par le contenu. Enfin, sur les aspects de trajectographie, nous considérons des problèmes liés à l'estimation de systèmes dynamiques partiellement observés dans un contexte passif, notamment dans le domaine sonar, ainsi qu'à la formalisation du suivi d'objets mutiples par des méthodes de type PMHT ou 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 de type vidéo ou sonar. Deux types de stratégies sont considérés et combinés : d'une part, des stratégies locales, liées aux techniques d'asservissement visuel et fournissant une perception de bas niveau, et d'autre part, des stratégies globales, activées sur des événements particuliers et remontant à une perception de plus haut niveau. Nous abordons le couplage d'une vision locale embarquée et d'une vision globale de la scène. Par ailleurs, nous étudions des questions liées à la théorie de l'exploration ("search theory").

Dans les trois secteurs d'applications mentionnés, nous avons eu récemment ou nous avons actuellement des conventions de recherche ou de transfert avec Météo-France, Thomson-CSF-Optronique, l'Ina, Thomson-Multimedia, EDF, l'Ifremer, le Cemagref, des projets multi-partenaires nationaux (programme RNRT, programme Priamm) et européens (projet Esprit R&D Divan, projet IST Carsense, contrat Eumetsat), et des collaborations avec des partenaires académiques comme notamment le LMD et l'École Navale.

3 Fondements scientifiques

3.1 Modèles markoviens hiérarchiques et estimation bayésienne

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

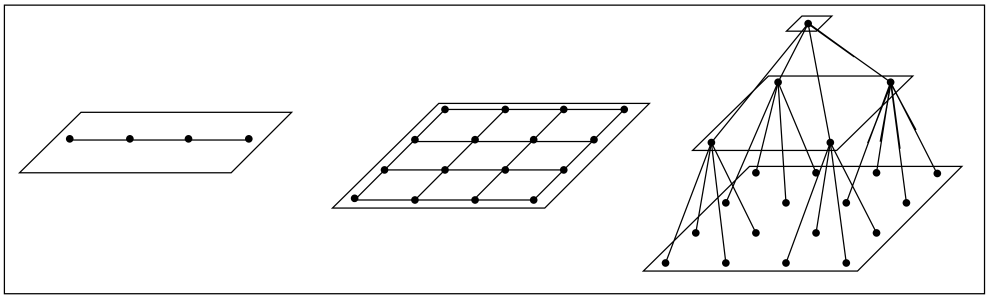
Résumé : *Dans le but de traiter les problèmes inverses multi-images liés aux axes de recherche du projet, nous nous intéressons au formalisme markovien associé à la théorie bayésienne de la décision* [Win95]. *Dans ce cadre, nous nous attachons à définir des modèles ou des algorithmes de type hiérarchique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

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

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

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

où $\mathcal{C}$ est un ensemble de "petites" parties de S ($|c| = 1$ 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.



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.)

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

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

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

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

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

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

3.2 Commande référencée vision

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[DD95] D. Dementhon, L. Davis, « Model-based object pose in 25 lines of code », *Int. Journal of Computer Vision 15*, 1/2, juin 1995, p. 123–141.

[Cha97] F. Chaumette, « Potential problems of stability and convergence in image-based and position-based visual servoing », *in : Workshop on Vision and Control*, Block Island, Rhode Island, juin 1997.

[SBE91] C. Samson, M. L. Borgne, B. Espiau, *Robot Control: the Task Function Approach*, Clarendon Press, Oxford, 1991.

tâche. On obtient :

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

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

3.3 Extraction de pistes et trajectographie

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

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

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

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

Le problème d'extraction-poursuite de cibles mobiles se pose dans le contexte du traitement de l'information. On dispose d'observations constituées de fonctions non-linéaires de l'état de la source (azimuts, dopplers, etc.). Les observations sont décrites comme les sorties d'un système variable dans le temps. Le modèle d'évolution temporelle de l'état est lié aux hypothèses faites sur la trajectoire des sources. Il est alors possible de décrire les observations par un système non-linéaire variable dans le temps pour lequel les paramètres à déterminer sont les composantes de l'état initial. Plus précisément, on appelle $\mathbf{X}^i$ le vecteur de l'état relatif (i.e. dans un repère lié à l'observateur) de la ième 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 $\mathbf{X} = (\mathbf{X}^1,\mathbf{X}^2,..,\mathbf{X}^n)$ 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: $\hat{\theta}^i_k = \theta^i_k + w_k$, où $\theta^i_k = \arctan(r^i_x(k)/r^i_y(k))$ et i est un

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

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

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

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

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

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

$$\begin{aligned} p(Z_t|\mathbf{X}_0) = u^{-m_t}(1-P_d)\mu_u(m_t) + u^{1-m_t}\frac{P_d.\mu_u(m_t-1)}{m_t} \\ \sum_{j=1}^{m_t} \det(\Sigma_t)^{-1/2} exp\left(-1/2||z_{t,j} - h_t(\mathbf{X}_0)||^2_{\Sigma_t^{-1}}\right) \end{aligned}$$

où m_t désigne le nombre de détections dans la fenêtre de validation, u le volume de la fenêtre de validation, Z_t l'observation au temps t, $\mu_u(m_t)$ la probabilité de m_t et P_d la probabilité de détection. Le problème de l'*extraction* revient alors à chercher le vecteur $\mathbf{X}_0$ maximisant la vraisemblance $P(\mathcal{Z}|\mathbf{X}_0)$ définie ci-dessus.
Cependant, la réalité de l'extraction est faite de problèmes multi-sources. L'assignation des mesures aux pistes peut être réalisée au moyen d'une énumération exhaustive des hypothèses d'assignation. Ceci constitue le cadre classique des algorithmes MHT (Multiple Hypotheses Tracking). Pour éviter l'explosion combinatoire de tels algorithmes, diverses méthodes ont été développées. Il est cependant assez évident que la faiblesse des algorithmes de tri et, dans une moindre mesure, de fusion réside dans le risque d'éliminer certaines bonnes séquences.

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

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

4 Domaines d'applications

Mots clés : imagerie météorologique, environnement, mécanique des fluides expérimentale, métrologie dynamique, imagerie médicale, santé, sonar, indexation vidéo, multimedia, 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, avec Meteo-France, le* LMD *et Eumetsat), de l'imagerie médicale (imagerie cérébrale,* IRM*,* IRM*f, échographie 3*D*). Le second secteur a trait à la robotique, aussi bien robotique manufacturière que robotique d'intervention, et aux systèmes de surveillance et de navigation. 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*

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

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

actuellement au projet Agir dans le cadre du RNRT *et au projet Mediaworks (le Limsi et* TF*1 notamment) dans le cadre du programme Priamm.*

5 Logiciels

5.1 Logiciel RMRm d'estimation du mouvement dominant

Correspondant: Fabien Spindler

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

[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.

entre ces deux images, et on constate une chute soudaine de cette valeur. Lors de transitions progressives, cette chute est moins nette mais plus prolongée. Cette détection de sauts est effectuée par un test statistique de Hinkley, dont l'implantation est simple et efficace. Un atout original de la méthode est d'appréhender ainsi la détection des différentes transitions par un même test, ne comprenant qu'un seul paramètre à fixer par l'utilisateur. Une même valeur de ce paramètre est de plus utilisée pour les deux types de transitions traitées.

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

5.3 Plate-forme VISP d'asservissement visuel

Correspondant: Éric Marchand

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

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

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

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

Correspondant : Christian Barillot

Dans le contexte du développement de logiciels d'analyse d'images 3D et 3D+t dans le domaine de l'imagerie médicale, nous travaillons à la réalisation de librairies de traitements d'images volumiques spatio-temporelles utilisables dans des contextes variés comme le recalage linéaire et le recalage non-linéaire d'images, la segmentation d'images et la visualisation d'images et de résultats. Cette plate-forme logicielle, dénommée VISTAL pour "Volumetric ImageS Treatments and Analysis Libraries", est constituée d'une classe générique (template) C++ ainsi que d'une série de librairies de traitements, incluant calcul algébrique et d'opérations floues sur des images 3D, manipulation de voisinage dans l'image, filtres linéaires (RIF et RII) et non-linéaires, filtres gaussiens 1D et 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).

6 Résultats nouveaux

6.1 Modèles statistiques, métrologie du mouvement et des déformations

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

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

6.1.1 Modèles pour la mesure du mouvement fluide

Participants : Étienne Mémin, Patrick Pérez, Thomas Corpetti, Stelios Zimeras.

Nous nous intéressons au problème de la mesure du mouvement dans des séquences d'images représentant des phénomènes fluides. L'analyse de séquences d'images de fluides est d'une importance capitale dans des domaines comme l'imagerie médicale lorsque l'on s'intéresse à la mécanique des bio-fluides, les sciences environnementales (océanographie, météorologie et climatologie) si l'on étudie la façon dont évolue l'atmosphère à travers les formations nuageuses, l'océan en terme de courants ou si l'on cherche à suivre un polluant, ou encore la mécanique des fluides expérimentale si l'on cherche à mesurer les vitesses d'écoulement du fluide observé et en analyser la structure dans le cas d'écoulements turbulents par exemple. Dans tous ces domaines, l'analyse du mouvement de ces phénomènes à partir des images numériques acquises requiert l'extraction du champ des vitesses apparentes à partir de la fonction intensité (plus précisément de ses variations spatio-temporelles). Notons également que ce type de mesures peut fournir des données utiles aux modèles numériques d'écoulements (comme les champs de vent pour les modèles de prévision météorologique).

Dans le contexte de l'imagerie fluide, 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 (mouvement occasionnant une variation de l'intensité lumineuse et caractérisés par un mouvement apparent de divergence non nulle), nous avons proposé une contrainte de mouvement apparent spécifique issue de *l'équation de continuité* de la mécanique des fluides. Cette nouvelle contrainte rend compte de la divergence du mouvement apparent et de la variation d'intensité qui lui est associée. Cette contrainte 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 conduit à l'introduction d'un terme de lissage favorisant l'émergence de zones compactes à rotationel et à divergence non nuls, ce que ne peut autoriser un lissage classique au premier ordre. À cette fin, nous avons élaboré une fonction de régularisation originale inspirée des techniques de régularisation "div-curl" du second ordre.

Cette approche développée dans le cadre de la thèse de Thomas Corpetti a été évaluée et comparée, sur des images météorologiques [50, 51], à l'approche générique de mesure de champ de vitesses précédemment élaborée dans le projet [MP98]. 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 été également adaptée à la mesure de champs de vent des basses couches nuageuses et validée expérimentalement dans le cadre du post-doctorat de Stellios Zimeras et de la collaboration avec le LMD (Laboratoire de Météorologie Dynamique) (voir sous-section 7.2).

6.1.2 Amélioration de séquences d'images et super-résolution

Participants : Patrick Bouthemy, Patrick Pérez, Fabien Dekeyser.

Notre objectif est 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 l'amélioration du rapport signal-à-bruit de la séquence d'images et 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, 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'inconvénient majeur est que l'utilisation d'un seul modèle paramétrique de mouvement ne permet que l'estimation du mouvement dominant et non des mouvements secondaires, ce dont nous devrons tenir compte dans les traitements de filtrage et de super-résolution.

Après avoir traité la question du filtrage spatio-temporel adaptatif [56], nous avons cette année abordé le problème de la super-résolution. Nous avons 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. Ce procédé est ensuite itéré pour chaque image de la séquence originelle. 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, ce qui permet de ne pas créer d'artéfacts [55]. Nous étudions à présent des approches conjointes réalisant la restauration des images et l'estimation du mouvement apparent simultanément, chaque étape bénéficiant ainsi des performances améliorées de l'autre. Pour cela, nous avons d'une part défini un schéma original

[MP98] E. Mémin, P. Pérez, « A multigrid approach to hierarchical motion estimation », *in : Proc. Int. Conf. on Computer Vision, ICCV'98*, p. 933–938, Bombay, Inde, janvier 1998.

[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.

et performant de multi-résolution temporelle qui a montré une amélioration significative de la qualité de la suite d'images super-résolues reconstruites. Il s'appuie sur 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, et ainsi de suite jusqu'à atteindre la cadence temporelle de la suite d'images super-résolues visée. Nous travaillons d'autre part à une autre approche du problème par la définition et la minimisation alternée d'une fonction d'énergie régularisante (modélisation markovienne) spécifiant de manière conjointe l'estimation des paramètres du mouvement et la reconstruction d'une image super-résolue. Cette étude est menée en collaboration avec Thomson-CSF-Optronique (voir sous-section 7.1).

6.1.3 Modélisation non paramétrique de l'information spatio-temporelle et reconnaissance du mouvement

Participants : Patrick Bouthemy, Ronan Fablet.

Nous avons proposé une approche originale de modélisation statistique et de caractérisation non paramétrique du mouvement, et considéré son application à des problématiques de reconnaissance et de classification, ainsi que de segmentation des images, basées sur une notion générale de mouvement que l'on appelera "activité de scène". L'analyse du mouvement dans des séquences d'images s'appuie généralement sur des modèles de mouvement 2D paramétriques ou 3D. Toutefois, certaines classes de mouvement (mouvement articulé comme celui d'une marche ou d'une course, mouvement non structuré comme celui d'une foule ou de feuillages au vent, mouvement fluide comme celui d'une rivière ou de flammes) ne peuvent être aisément appréhendées par ces techniques classiques. Notre approche repose sur l'analyse statistique de distributions spatio-temporelles de mesures locales de mouvement et leur représentation par des modèles de Gibbs temporels et causaux. Cette modélisation nous permet une évaluation complète de la vraisemblance conditionnelle des séquences d'observations locales de mouvement relativement à un modèle de mouvement. Cette caractéristique nous permet d'établir un schéma direct pour l'estimation de ces modèles au sens du maximum de vraisemblance. Une phase de réduction de l'ordre des modèles ainsi identifiés 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 de classification et de reconnaissance des contenus dynamiques des vidéos. En premier lieu, nous avions exploité cette mesure de similarité de mouvement pour construire une classification hiérarchique non supervisée d'un ensemble de vidéos. Cette dernière peut être exploitée pour des applications de reconnaissance au sens large, et en particulier en indexation vidéo pour la visualisation rapide du contenu d'une base de vidéos ou pour accélérer le processus de recherche d'exemples similaires à une vidéo proposée comme requête [58].

Cette année, nous avons abordé le problème de la segmentation des images au sens de l'activité de scène. Cette tâche est formulée comme un problème d'étiquetage régularisé (markovien) d'un graphe d'adjacence résultant d'un découpage initial de l'image en blocs, un modèle d'activité de scène étant initialement identifié dans chaque bloc. La mesure de similarité introduite précédemment est à nouveau exploitée entre blocs cette fois. L'intérêt de cette approche est de pouvoir segmenter des zones d'intérêt dans l'image sans modèle préétabli de mouvement; elle est donc d'une très grande applicabilité puisque des contenus très divers peuvent être ainsi appréhendés. En quelque sorte, le modèle est appris en même temps que la classification s'opère, et la segmentation n'est plus un préalable (souvent très délicat) à la reconnaissance, mais en résulte. Notre approche fournit donc un cadre unifié pour l'extraction et la caractérisation d'entités d'intérêt relativement à des propriétés générales de mouvement. Elle peut englober dans son principe tout type de mouvement (rigide, articulé, ou déformable) et tout type de situation (entité unique -un véhicule ou une personne-, ou groupe d'éléments -vol d'oiseaux ou groupe de joueurs).

6.1.4 Modèles probabilistes de formes en imagerie médicale

Participants : Christian Barillot, Isabelle Corouge.

Nous étudions des techniques d'analyse statistique multidimensionnelle pour modéliser des 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. Ce dernier nous permet d'aligner, de façon rigide, tous les exemplaires de l'ensemble d'apprentissage dans un référentiel local commun. Dans cet "espace local", les déformations d'une structure sont caractérisées par 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 calculés et leur caractère discriminant, intéressant notamment pour la classification d'éléments inconnus ou la détection de sous-groupes au sein d'une population.
D'autre part, le champ de déformation obtenu, entre un sillon donné et le sillon de référence, est étendu à un voisinage local du sillon considéré en utilisant l'interpolation basée sur les splines de type "thin plates". Il peut alors être appliqué à tout objet qui aura été associé à ce sillon. C'est le cas par exemple de dipôles MEG (localisation d'activation fonctionnelle), que nous recalons ainsi de façon locale et non-linéaire. Il est alors possible de caractériser de manière statistique la population des dipôles ainsi recalés, et de comparer le recalage inter-sujets des dipôles dans un cadre local (par rapport aux sillons corticaux) et dans un cadre global (par rapport aux IRM). Cette technique a de plus été utilisée dans le contexte de l'évaluation de méthodes de recalage global inter-sujets (méthodes rigides et non rigides). Pratiquée sur des amers locaux, l'analyse statistique fournit un indicateur de la similarité des formes au sein des populations recalées. Nous nous efforçons de généraliser cette approche à d'autres entités et à d'autres supports comme l'imagerie échographique 3D. Une des difficultés réside dans l'extraction et la représentation des structures d'intérêt. Un modèle déformable bien construit est utilisable dans de nombreux contextes que nous n'avons pas encore explorés : reconnaissance de structures complexes, création d'atlas, aide à la formalisation des liens entre représentation

numérique et représentation symbolique de structures dans un contexte de fouille de données.

6.1.5 Analyse d'images ultrasonores

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

Notre travail dans ce domaine s'est orienté dans trois directions :
- la réalisation d'un système d'acquisition d'échographies 3D et son évaluation en site hospitalier sur des cas cliniques;
- la définition de méthodes de recalage et d'analyse des déformations en imagerie échographique 3D neurochirurgicale per-opératoire;
- le développement de méthodes de segmentation d'images ultrasonores 3D par l'utilisation d'ensembles de niveaux et l'introduction d'une modélisation statistique des régions à segmenter.

La première étude a consisté en la mise en place d'un système d'acquisition d'échographies 3D par la technique dite "main libre" avec comme objectif à moyen terme de réaliser le couplage entre une imagerie ultrasonore 3D et un système de neuro-navigation dans un contexte d'imagerie per-opératoire en neurochirurgie. Nous avons conçu un dispositif associant le logiciel StradX de l'Université de Cambridge (voir sous-section 8.3.1) et un équipement de neuro-navigation fourni par le CHU de Rennes (système de capteurs optiques de type Flash-Point 5000) afin de référencer les acquisitions ultrasonores dans la géométrie du système de neuro-navigation. Ce système a été évalué, en collaboration avec Pierre Darnault (responsable de l'unité d'échographie radiologique du CHU de Rennes), dans trois domaines cliniques : imagerie orthopédique et cérébrale chez le nourrisson, imagerie obstétrique et imagerie vasculaire périphérique. Ces expérimentations ont permis de dégager un certains nombre d'enseignements pour l'utilisation future de tels systèmes d'acquisition qui nous conduisent à nous intéresser à trois aspects cruciaux de la phase d'acquisition: le développement de méthodes robustes de calibration, le développement de méthodes de reconstruction 3D exploitant la corrélation entre le mouvement mesuré de la sonde et l'analyse du mouvement apparent dans les images (thèse de F. Rousseau), le développement de système d'aide à l'élaboration de trajectoires optimales de la sonde échographique par l'utilisation d'un système robotique asservi (contrat CPER).
La deuxième étude nous a conduit à développer une méthode de recalage non rigide pour la correction de mouvement dans une séquence échographique 3D spatio-temporelle per-opératoire en neurochirurgie. Nous avons extrait des structures similaires entre données IRM et ultrasons 3D (ventricules par exemple), et généré des déformations synthétiques pour simuler l'influence de l'action opératoire. Ces données nous ont été fournies par Xavier Pennec du projet Epidaure. La compensation de cette déformation se fait par la méthode de calcul de champs denses multigrilles 3D développée par ailleurs dans le projet [66].
Le dernier aspect de notre travail porte sur la segmentation de structures présentes dans des images échographiques 3D (artères carotidiennes par exemple) par l'utilisation des ensembles de niveaux. Une surface 3D fermée se propage d'une position initiale vers les contours désirés par l'évolution itérative d'une fonction implicite 4D. La contribution principale de ce travail est la conception d'un modèle robuste auto-adaptatif de la fonction d'évolution. Le pas d'itération et le coefficient de la force de propagation externe, tous deux habituellement fixés empirique-

ment, sont ainsi automatiquement calculés à chaque itération en fonction des données et de l'état courant du processus. En outre, une information de type région, liée à la distribution statistique estimée des intensités de l'image par un algorithme SEM, est utilisée à la place des gradients d'intensité. Cette méthode a été appliquée à des images échographiques 3D de la carotide, mais a également pu être adaptée pour traiter des images IRM du cerveau (pour la segmentation entre matière grise et matière blanche).

6.1.6 Structuration et classification de vidéos, indexation de données numériques

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

Nous continuons à développer notre schéma de structuration et d'indexation de vidéos lié au contenu dynamique de ces dernières. Il comprend trois grandes phases, la première réalisant la segmentation temporelle de la vidéo en plans et séquences, la seconde concernant la caractérisation des plans extraits et l'extraction d'entités mobiles pertinentes pour l'indexation, la troisième portant sur l'extraction de descripteurs de mouvement ou d'activité du contenu de la scène. Des travaux sont par ailleurs entrepris concernant des algorithmes d'indexation proprement dits, et leur adaptation aux problèmes spécifiques posés par les données multimédia : imprécision, grande dimension, évolution possible des bases et des descripteurs utiles. Concernant le premier point, nous avons cette année développé des techniques temps-réel de caractérisation du mouvement de la caméra et de découpage associé s'attaquant directement au document vidéo compressé et utilisant les vecteurs de mouvement transmis dans le flux MPEG. Un estimateur robuste du modèle de mouvement dominant dans l'image courante exploite en effet un champ de vitesses 2D entre les deux images successives reformé à partir des vecteurs MPEG. Nous avons également défini un schéma régularisé de détection des entités mobiles dans chaque plan utilisant comme données d'entrée les vecteurs MPEG et les coefficients DCT d'ordre 0. Nous nous intéressons par ailleurs à des techniques de classification de type SVM ("Support Vector Machines") pour caractériser les plans en termes de type d'activité à partir de mesures spatio-temporelles locales. Enfin, nous exploitons le schéma statistique non paramétrique de reconnaissance et de segmentation d'entités dynamiques pertinentes, exposé à la sous-section 6.1.3, pour formaliser et satisfaire des requêtes partielles (c.a.d. portant sur une partie du contenu du plan vidéo) [60].

Nous avons attaqué, en collaboration avec Laurent Amsaleg du projet Temics, un nouveau problème : le stockage des descripteurs extraits des vidéos. Cela peut également être étendu à d'autres descripteurs concernant les images fixes ou la bande sonore, auxquels s'intéressent d'autres projets de l'Inria (projets Imedia, Movi, Sigma-2) avec lesquels nous collaborons au demeurant dans différents contextes. Une fois calculés de nombreux descripteurs, 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 qui sont 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 [80]. 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. Diverses solutions, basées sur des schémas d'approximation des données, sont envisagées et ont commencé à être testées. Il s'agit de regrouper les données dans des cases prédéfinies, puis à pouvoir déterminer le plus rapidement possible lors d'une requête quelles sont les cases dont le contenu peut être intéressant. Beaucoup de techniques existent, mais leur degré de complexité est très élevé, aussi un travail spécifique est nécessaire dans le cas d'espaces de grande dimension et de données de grande taille. Dans un premier temps, nous travaillons à partir de l'algorithme des VA-files développé à ETHZ de Zürich par Weber[WSB98] pour l'adapter à une taille de données plus importante. Nous avons plus particulièrement regardé le cas où les images sont décrites dans des descripteurs locaux (par exemple ceux développés au sein du projet Movi) et où la requête est donc constituée de nombreuses interrogations de la base. Dans ce cas, il est possible de regrouper et d'organiser ces interrogations de manière à minimiser le temps de réponse. 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 des plus grands volumes de données. L'objectif est de passer de tests sur 1000 à 5000 images à des tests sur un million d'images. Par ailleurs, l'intérêt d'utiliser des descripteurs liés à la couleur par rapport à des descripteurs de niveaux de gris a aussi été montré [63].

6.2 Recalage, suivi, trajectographie

Contexte : *Le problème générique de* suivi *dans des séquences d'images, qu'elles soient 1D, 2D, ou 3D, occupe une place centrale dans de très nombreuses applications. Il consiste à extraire et reconnaître les entités d'intérêt, puis à établir un lien temporel entre les instances successives de ces entités. Les principales difficultés résident dans la variabilité au cours du temps des attributs caractérisant une même entité, allant jusqu'à la disparition temporaire, totale ou partielle de l'entité, dans la présence simultanée éventuelle de plusieurs entités, pouvant être très proches, au sein des données, et dans la présence de "fouillis" 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 pouvant inclure le suivi de structures rigides (véhicules) et articulées (piétons), le suivi d'entités déformables (formes nuageuses, vortex), le suivi d'entités d'intérêt pour l'indexation vidéo, le suivi d'objets pour l'asservissement visuel, l'extraction de pistes en sonar ou radar.*

[WSB98] R. Weber, H. Schek, S. Blott, « A quantitative analysis of performance study for similarity-search methods in high-dimensional spaces », *in : Proc. of the 24th Int. Conf. on Very Large Data Bases, New-York, USA*, p. 194–205, août 1998.

6.2.1 Suivi en imagerie par filtrage particulaire

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

Pour traiter les problèmes de suivi, des méthodes probabilistes à base de filtrage de Kalman sont généralement employées. De telles techniques présentent des limitations notables pour gérer les non-linéarités induites par les équations d'observation. Elles se heurtent en particulier au problème d'entités multiples qui nécessite la manipulation de lois multimodales. Pour pallier ces problèmes, les techniques d'association de données, basées sur des mélanges de lois normales, ont été développées dans le domaine de l'extraction multi-pistes. Plus récemment, des techniques dites de "condensation" (pour "conditional density propagation") ont été introduites en vision par ordinateur. Apparentées aux techniques de filtrage particulaire, ces méthodes reposent sur la manipulation de jeux d'échantillons assortis de poids. Dans le cadre des applications de suivi intéressant le projet (extraction de pistes en sonar et radar, extraction et suivi de structures telles que les vortex en imagerie de phénomènes fluides, suivi de véhicules en trafic routier,...), les travaux se sont focalisés sur le suivi d'objets multiples et partiellement observés (thèse de Carine Hue) par des méthodes de type filtrage particulaire. Dans un premier temps, les efforts ont consisté à analyser les techniques de rééchantillonnage (systématique et adaptatif) dans le cadre du suivi d'un objet unique et manœuvrant, partiellement observé (par exemple, dans le cas de mesures d'angles seules comme en trajectographie sonar). Cette analyse a permis de mettre en évidence l'interêt de ces méthodes pour ce problème. On peut 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 ou encore à l'initialisation. Enfin, le coût de calcul n'est pas prohibitif.

Il est cependant certain que de nombreuses applications pratiques 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 majeure réside cependant dans la méconnaissance de l'association des mesures aux modèles des objets. Il s'agit alors de résoudre simultanément un problème d'association et d'estimation. Le formalisme que nous avons adopté est celui des lois de mélange de densités, où il s'agit d'estimer à la fois les variables d'affectation des mesures aux pistes et les états des objets. Celui-ci est à la base des méthodes de type PMHT. Ceci a été réalisé en combinant d'une part le filtre particulaire mono-objet et, d'autre part, un échantillonneur de Gibbs. Celui-ci est aussi connu, dans ce contexte, sous le nom de "data augmentation method". Les probabilités d'assignation des mesures aux modèles sont elles-aussi estimées par filtrage particulaire. 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. Les performances s'avèrent, elles-aussi, tout à fait satisfaisantes [94]. De plus, la méthode originale de filtrage particulaire multi-objets que nous avons ainsi définie a été étendue au cas de récepteurs multiples, ce qui inclut alors une diversité des types de mesures et de leur fréquence. Une application au suivi de cibles multiples, dans un contexte de radar ou de sonar, a été développée et met en évidence l'apport d'informations de nature différente et l'aptitude du filtre particulaire à gérer celles-ci. D'autre part, l'application de notre méthode au suivi d'objets multiples dans des séquences d'images est en cours de réalisation.

6.2.2 Suivi de structures déformables en imagerie météorologique

Participants : Patrick Bouthemy, Etienne Mémin, Christophe Papin.

Dans le cadre de la thèse de Christophe Papin, soutenue en décembre 1999, nous avons développé une méthode permettant de détecter et de suivre des cellules nuageuses convectives, ainsi que d'en caractériser l'état [74]. Cette méthode s'appuie sur la méthode des ensembles de niveaux ("level sets") proposée par Osher et Sethian. Cette dernière consiste à décrire l'évolution d'une courbe fermée par l'évolution d'une hyper-surface d'un degré supérieur. Notre contribution dans ce cadre a consisté à définir une équation d'évolution de cette surface à deux composantes agissant successivement. La première est une composante dynamique originale s'appuyant sur un champ de vecteurs de vitesses estimé au préalable sur des zones d'intérêt potentiel. Cette composante permet notamment de gérer proprement une propagation temporelle (prédiction) de l'hyper-surface estimée à l'instant précédant en intégrant de façon naturelle les changements de topologie de cette surface. La seconde est une étape photométrique permettant de déformer les contours des cellules convectives en prenant en compte les températures locales des différentes cellules et leurs températures moyennes prédites à partir de l'équation de conservation de la masse de la mécanique des fluides. Cette deuxième étape imprime au contour une dilatation ou une contraction suivant le contexte photométrique local. Une consolidation de ce travail a été effectuée par Christophe Papin, à la suite de la thèse, par des évaluations expérimentales extensives sur de nombreuses situations météorologiques pertinentes et difficiles, en relation avec Météo-France. De plus, un effort logiciel particulier a été accompli pour optimiser le code et produire un logiciel documenté et modulaire (séparation en particulier des aspects génériques de l'implantation des ensembles de niveaux et de la spécification des aspects relatifs à l'application considérée).

6.2.3 Segmentation et recalage 3D non linéaire en imagerie cérébrale

Participants : Christian Barillot, Etienne Mémin, Pierre Hellier, Caroline Baillard.

Nous avons défini des méthodes de recalage non-rigide en imagerie cérébrale, mono-modalité et multi-modalité et avons abordé leur coopération avec des méthodes de segmentation 3D. Dans un premier temps, nous avons développé une approche originale pour estimer une transformation 3D dense entre deux volumes IRM de deux sujets différents [65]. Cette méthode est basée sur une estimation du flot optique, exprimée dans un cadre statistique robuste, afin d'une part de limiter l'influence du bruit d'acquisition de l'IRM, et d'autre part d'autoriser localement des changements de topologie. Nous avons également proposé un schéma de minimisation efficace, à la fois multirésolution et multigrille de la fonction d'énergie ainsi construite. Une segmentation anatomique du cerveau est introduite dans le partitionnement adaptatif sur lequel repose la minimisation multigrille, dans le but de limiter l'estimation des champs de déformation à un support cohérent, d'accélérer l'algorithme et de raffiner l'estimation dans les zones importantes. Nous avons validé cette méthode sur des données synthétiques, puis sur une base d'images IRM de 20 sujets. Nous nous sommes également intéressés au problème du recalage non-rigide multi-modalité [66]. Les images fonctionnelles IRM*f*, 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 gênent l'interprétation des données fonctionnelles en raison du déphasage entre les cartes d'activations et les régions anatomiques d'intérêt. Nous avons proposé un algorithme de recalage multimodalité reposant sur une mesure de similarité adaptée (l'information mutuelle). Il exploite une minimisation multigrille en estimant un modèle paramétrique affine à chaque niveau de grille. Nous avons validé cette méthode sur des données simulées et sur des données cliniques.

Par ailleurs, une autre étude a permis de montrer qu'il est fructueux de faire coopérer des procédés de segmentation et de recalage. Une méthode de type ensemble de niveaux ("level sets") de segmentation de structures anatomiques a pu ainsi bénéficier d'une initialisation appropriée issue d'un recalage non rigide même relativement grossier (mais rapide) [43]. Le champ de déformation calculé permet en effet de prédire correctement la localisation des modèles déformables, entraînant une segmentation automatique plus performante et plus précise. Nous avons validé cette approche sur des données simulées et réelles. Nous avons également étudié le mode réciproque de coopération entre segmentation et recalage. Une segmentation de structures anatomiques d'intérêt comme les sillons corticaux a permis de contraindre le procédé de recalage non-rigide présenté plus haut. En particulier, la modélisation statistique retenue permet d'intégrer cette contrainte locale et éparse de manière naturelle. Sur la base de vingt sujets, dont nous disposons pour chacun d'une image IRM et de six sillons majeurs segmentés par hémisphère, nous avons montré l'apport de cette contrainte anatomique dans le recalage non rigide.

6.2.4 Appariements d'objets et de scènes

Participants : Patrick Gros, Damien Albert.

Un problème récurrent en robotique et particulièrement en asservissement visuel est la nécessité de disposer d'un appariement entre l'image initiale vue par la caméra et une image connue, en l'occurrence correspondant à la réalisation de la tâche considérée, pour lancer la réalisation de cette tâche (tâche de positionnement par exemple). Cela est également nécessaire dans le domaine de la robotique mobile pour connaître la position initiale du robot. Les techniques d'indexation peuvent être un moyen de résoudre cette difficulté, car elles permettent d'une part de stocker de très nombreuses images correspondant aux divers positions possibles de la caméra et de l'effecteur, mais aussi de comparer ces images à une nouvelle image de manière efficace. Un tel couplage n'est pas sans difficulté. D'une part, les algorithmes d'asservissement supposent, jusqu'à présent, que les appariements entre le motif courant et le motif cible qui leur sont fournis sont exacts, hypothèse qui ne peut tenir lorsque ces appariements sont effectués de manière entièrement automatique. Il sera donc nécessaire de développer des algorithmes d'asservissement robustes, qui acceptent qu'une partie des appariements fournis soient erronés. D'autre part, les environnements de travail des robots présentent des particularités fortes, en termes des objets perçus ou de conditions d'illumination, qui peuvent permettre ou qui obligent de développer des descripteurs particuliers. L'objectif est ainsi de pouvoir utiliser l'asservissement visuel de manière moins contrainte, par la suppression de toute intervention manuelle en phase initiale. Dans un premier temps, notre effort s'est porté sur les problèmes posés par l'interfaçage entre des techniques d'asservissement et des techniques de reconnaissance, et sur la constitution de base contenant non seulement des images mais aussi des relations géométriques entre ces images.

6.2.5 Extraction de pistes et poursuite

Participants : Jean-Pierre Le Cadre, Isabelle Leibowicz.

Quelques questions spécifiques de trajectographie passive ont été abordées cette année en complément des travaux précédents que nous avions réalisés sur ce sujet, en particulier une méthode de trajectographie pour des mesures d'angles coniques avec la prise en compte de l'attitude du récepteur. Un travail important a aussi été réalisé dans le cadre des problèmes d'association pistes-à-pistes, pistes-à-mesures et sur l'apport des données d'identification (association et fusion multi-capteurs). Pour cela, l'approche retenue consiste à utiliser conjointement les outils classiques de statistiques (lois des estimateurs, tests d'hypothèses, etc...) d'une part, et d'optimisation combinatoire d'autre part (algorithmes de type "Auction"). Pour tous ces problèmes, les capacités d'"auto-contrôle" du système sont un volet essentiel. Ainsi, il est rare de disposer de la variance des données observées, des biais (erreur de positionnement,...), ou encore des probabilités de détection et de fausses alarmes. Comme nous sommes amenés à considérer des systèmes complexes, une étape essentielle consiste à estimer la qualité et la fiabilité des observations élémentaires. Nous avons montré que ces problèmes pouvaient être résolus par des méthodes statistiques classiques, avec des résultats satisfaisants. Par contre, l'utilisation de méthodes plus sophistiquées, telles que la détection répartie, s'est heurtée à des problèmes fondamentaux comme la méconnaissance des rapports signaux-à-bruits élémentaires.

6.3 Asservissement visuel, perception active et exploration

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

6.3.1 Planification dans l'image

Participants : François Chaumette, Youcef Mezouar.

Les techniques actuelles d'asservissement visuel 2D, même si elles sont efficaces pour une grande classe d'application, se heurtent à des difficultés lorsque la position initiale de la caméra est fortement éloignée de sa position désirée. Nous avons résolu ce problème en développant une technique de planification de trajectoires dans l'image de plusieurs points d'un objet [72]. Le suivi de ces trajectoires entraîne alors une faible erreur entre la consigne et la mesure courante, cadre où les techniques d'asservissement visuel 2D sont connues pour être stables et robustes aux diverses erreurs de modélisation. Pour générer de telles trajectoires dans l'image, nous avons adopté la méthode des fonctions de potentiel, classique dans le domaine de la planification. La fonction de potentiel que nous avons modélisée est définie dans l'espace opérationnel du robot et admet un pôle attractif (la configuration désirée). Elle contraint la trajectoire du robot pour atteindre ce pôle à suivre une ligne droite dans l'espace opérationnel, ce qui est d'un grand intérêt pratique. Il est également possible de rajouter des contraintes dans la fonction de potentiel (sous forme de pôles répulsifs), afin de faire dévier la trajectoire du robot de sa trajectoire nominale quand cela s'avère nécessaire. Les contraintes que nous avons considérées sont de maintenir l'objet en permanence dans le champ de vision de la caméra et de permettre au robot d'éviter ses butées articulaires. De plus, en tirant bénéfice des résultats acquis dans le domaine de l'asservissement visuel 2D1/2 [31], nous avons développé une méthode qui ne nécessite pas la connaissance a priori des modèles 3D des objets considérés.

Une fois la trajectoire 3D du robot planifiée, il suffit ensuite de calculer, par simple projection perspective, les trajectoires dans l'image des points d'intérêt de l'objet servant à la spécification de la tâche. Une étude sur la sensibilité aux erreurs de calibration de la caméra a été effectuée. Elle aboutit à une grande robustesse de notre technique de planification. Par ailleurs, pour améliorer le comportement de la loi de commande utilisée pour suivre les trajectoires planifiées, celles-ci sont interpolées par des fonctions B-splines. On peut alors supprimer autant que possible les erreurs de traînage au cours de l'asservissement. Les résultats expérimentaux obtenus sur la cellule de vision robotique du projet se sont avérés très satisfaisants, aussi bien en ce qui concerne la forme des trajectoires réalisées, la réalisation des contraintes spécifiées, et la robustesse de notre méthode aux diverses erreurs de modélisation du système.

6.3.2 Évitement des butées articulaires

Participants : François Chaumette, Eric Marchand.

Pour toutes les tâches robotiques faisant intervenir un bras manipulateur, un problème important consiste à éviter les butées articulaires de ce bras au cours de la réalisation de la tâche. Des travaux ont déjà été menés sur ce thème dans le projet par le passé. Ceux-ci étaient basés sur le formalisme de la redondance de l'approche par fonction de tâche. La méthode développée consistait à exprimer l'évitement des butées comme la minimisation d'un coût secondaire, sous contrainte que la tâche principale soit réalisée. Elle est valide lorsque la tâche principale ne contraint pas l'ensemble des degrés de liberté du système, et est donc complémentaire des travaux sur la planification dans l'image (décrits à la sous-section 6.3.1). Toutefois, un défaut important de l'approche par redondance réside dans le choix très délicat du paramètre réglant l'importance de la tâche secondaire par rapport à la tâche principale. C'est pourquoi nous avons développé cette année une nouvelle méthode ne nécessitant le réglage d'aucun paramètre [45]. Elle consiste à bloquer les axes s'approchant trop dangeureusement de leur butée sans pour autant perturber la tâche principale. Elle repose sur la résolution d'une succession de simples systèmes linéaires permettant de trouver une solution au problème tant que celle-ci existe. Cette méthode générale a été validée pour une tâche d'asservissement visuel de centrage et de suivi d'un objet dans l'image par un bras manipulateur à six degrés de liberté, ainsi que par un buste d'un humanoïde de synthèse (cette dernière validation ayant été réalisée par N. Courty du projet Siames, voir sous-section 6.3.6).

6.3.3 Utilisation des moments en asservissement visuel

Participants : François Chaumette, Omar Tahri.

Le but de cette étude est de déterminer une représentation optimale et générique des informations visuelles pour l'asservissement visuel 2D. Il s'agit de sélectionner, dans l'ensemble des informations visuelles qu'il est possible d'extraire et de suivre dans une séquence d'images, le nombre minimal de celles-ci, tout en offrant de bonnes propriétés générales, à savoir, découplage des degrés de liberté du système, robustesse de la commande aux erreurs de mesure, absence de singularités et de minima locaux, adéquation des mouvements 3D générés. L'ensemble de ces propriétés peut être étudié à partir de la matrice d'interaction associée aux informations visuelles. L'an dernier, nous avions réussi à déterminer la forme analytique de la matrice d'interaction associée aux moments d'ordre quelconque d'un objet. La forme obtenue est générale et permet notamment de retrouver les résultats des cas particuliers des primitives géométriques, points, segments et ellipses. L'objectif de généricité est donc atteint, même si des occultations partielles de l'objet peuvent s'avérer problématiques dans l'utilisation des moments de sa projection dans l'image.
Nous avons ensuite considéré un objet particulier (en l'occurrence un rectangle), et avons pu sélectionner six informations visuelles basées sur les moments telles que, pour une configuration donnée (centrée et parallèle au plan image), la matrice d'interaction associée soit triangulaire. Les premières expérimentations que nous avons réalisées cette année ont permis de mettre en évidence des comportements très satisfaisants dans certaines configurations, améliorant très nettement les résultats obtenus précédemment. Malheureusement, dans certaines autres configurations, le conditionnement de la matrice d'interaction ne permet pas d'obtenir des résultats acceptables. Cela va nous entraîner à considérer des informations visuelles supplémentaires pour la suite de cette étude qui concernera la thèse de Omar Tahri qui vient de débuter.

6.3.4 Couplage vision locale et vision globale

Participants : François Chaumette, Grégory Flandin.

Les études menées précédemment au sein du projet dans le domaine de l'asservissement visuel et de la vision active concernaient l'utilisation d'une seule caméra. Dans un contexte de scènes plus complexes, encombrées ou comportant des objets mobiles, nous souhaitons à présent généraliser cette approche à l'emploi de deux capteurs, l'un fournissant une vision locale de la scène, l'autre en ayant une vue globale [62]. Dans ce cadre, nous nous sommes intéressés au problème de l'exploration de scènes pour la localisation d'objets complexes. Les modèles utilisés usuellement considèrent des objets constitués de primitives géométriques (points, segments, cylindres,…) ou bien une description exhaustive de l'environnement en terme de voxels. Pour notre part, nous nous sommes intéressés à l'estimation du volume englobant des objets et de leur localisation, cette connaissance étant suffisante pour de nombreuses applications telles que la planification de trajectoires ou l'évitement d'obstacles. De plus, il est ensuite possible de se focaliser sur un objet particulier si une information plus précise est nécessaire.

Le modèle de connaissance que nous avons développé résulte d'un mélange de modèle statistique et de modèle à erreur bornée sous une hypothèse ellipsoïdale des supports de l'erreur. A travers ce modèle, nous pouvons calculer, en chaque point de la scène, la probabilité que ce point appartienne à un objet donné. Plus précisément, l'algorithme d'estimation intègre successivement l'ensemble des informations acquises au cours de la trajectoire de la caméra locale et les fusionne avec celles acquises par la caméra globale. En se basant sur ce modèle de connaissance, nous avons ensuite déterminé les mouvements de la caméra permettant de raffiner de façon optimale le modèle de l'objet. La stratégie que nous avons développée consiste à réduire de façon isotropique l'incertitude liée à ce modèle. Les premiers résultats expérimentaux, obtenus sur des scènes simples permettant d'utiliser des algorithmes de traitements d'image élémentaires (mais temps-réel), se sont avérés très satisfaisants, validant ainsi la méthode et la stratégie développées. Nous comptons à présent recourir à des algorithmes de segmentation fondés sur une analyse du mouvement dans la séquence d'images, afin de valider nos travaux sur des objets de texture complexe.

6.3.5 Théorie de l'exploration

Participants : Jean-Pierre Le Cadre, Frédéric Dambreville.

Nous nous préoccupons de l'optimisation de l'effort de recherche d'un objet, généralement appelé "cible", qui peut être fixe, mobile (avec par exemple une trajectoire supposée markovienne) ou même réactif. On dispose de ressources, nécessairement limitées, et on cherche à optimiser des fonctions telles que la probabilité de détection, le temps de détection ou encore une fonction de risque. Notre objectif a été, dans un premier temps, de mettre en œuvre les méthodes connues de la théorie de l'exploration. On notera parmi les problèmes traités : l'optimisation de la trajectoire de recherche par des méthodes de type "branch and bound", l'optimisation de l'effort de recherche pour un problème à n-étapes (algorithme "forward-backward"). Bien que les résultats obtenus correspondent le plus souvent à l'intuition, l'aspect quantificatif est évidemment important. On voit cependant qu'il n'est pas nécessaire de compliquer beaucoup le cadre de l'analyse, en introduisant par exemple la "logique" d'association des décisions ou la prise en compte de fausses alarmes, pour arriver à des problèmes très difficiles. C'est le cas lorsque l'on considère des cibles réactives ou encore une fonction de risque liée à l'effort de recherche.

La thèse de Frédéric Dambreville s'est orientée dans un premier temps, vers ce type de problèmes (voir le rapport d'activité 1999). Cette analyse a été étendue cette année à l'optimisation de systèmes de détection multi-modes, en présence de contraintes syntaxiques (coopérations des modes et des ressources de détection, conditions de dépendence) [87]. Pour cela, nous avons développé un formalisme original, nommé "lateral splitting", qui, joint aux opérateurs de gestion des ressources, permet de traiter des problèmes très complexes. La répartition spatio-temporelle des efforts de recherche, [54], est alors un résultat de l'algorithme d'optimisation, qui serait d'ailleurs bien difficile à deviner. Par ailleurs, les travaux se sont aussi largement orientés vers les problèmes de type "Search games". Un des buts est de s'affranchir de connaissances trop restrictives relatives à la distribution, donnée a priori, de la cible. Il s'agit alors d'optimiser la répartition des efforts de recherche pour la "pire" distribution de la cible; il s'agit d'un problème de type "min-max". Nous avons obtenu des conditions générales d'optimalité. Nous avons ensuite étudié les implications pratiques de celles-ci, avec prise en compte des contraintes, et développé l'algorithme associé. Il faut noter que celui-ci permet de prendre en compte des cibles de trajectoire markovienne. Parmi les exemples d'application de ces méthodes, nous citerons plus particulièrement : l'inclusion des coûts de déplacement de l'observateur (algorithme "branch-and-bound"), l'optimisation d'un système de veille multi-modes (petit-champ, grand-champ) en liaison avec Thomson-CSF-Optrosys, l'évaluation des performances d'un réseau de détecteurs (avec le Gesma).

6.3.6 Asservissement visuel et animation

Participant : Éric Marchand.

Notre objectif dans le domaine de l'animation est de proposer un schéma général permettant le contrôle d'une caméra dans un environnement virtuel. La caméra doit non seulement pouvoir se positionner par rapport à son environnement, mais elle doit de plus être à même de réagir à des modifications de celui-ci. L'approche originale que nous proposons repose sur le formalisme de l'asservissement visuel, puisque ce dernier permet de générer automatiquement des mouvements 3D de la caméra à partir d'une tâche spécifiée dans l'image. 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 comme la réalité virtuelle ou les jeux vidéo.

Dans un premier temps, nous nous sommes intéressés à la commande d'une caméra dans des environnements de dynamique a priori inconnue. La trajectoire de la caméra est donc contrainte à la fois par la tâche de positionnement à réaliser et par la scène elle-même (risque de collision, d'occultation,...). Nous avons exploité à cette fin la notion de redondance de l'approche par fonction de tâche; les contraintes du système sont ainsi exprimées par la tâche secondaire, la tâche "visuelle" restant prioritairement satisfaite. Nous avons considéré un grand nombre de problèmes différents allant de tâches simples comme la poursuite d'un objet à des tâches plus complexes comme l'évitement d'occultations ou d'obstacles [70]. On peut voir un intérêt plus grand à utiliser ces techniques dans le cadre d'applications temps-réel, où l'utilisateur se déplace dans un environnement où l'on ne connaît pas *a priori* les mouvements des objets. Une autre application possible est celle de la commande référencée vision des mouvements d'un humanoïde de synthèse. L'asservissement visuel est en effet, là encore, une approche intéressante pour générer les commandes de bas niveau de tels "avatars". Il faut noter dans ce cas que les mouvements de la caméra virtuelle (en l'occurrence, les yeux de l'humanoïde) ne sont plus libres. Il faut considérer la caméra comme étant montée sur l'effecteur d'un robot hyper-redondant (typiquement 9 degrés de liberté si l'on considère la seule animation du buste). Notons que les recherches menées cette année sur l'évitement des butées articulaires (voir la sous-section 6.3.2) ont trouvé une autre application dans ce cadre.
Dans le cadre de notre collaboration avec le projet Siames sur ce sujet (N. Courty et B. Arnaldi), un couplage entre la bibliothèque ViSP et la plate-forme d'animation GASP a été réalisé. Ce

couplage permet de considérer une caméra comme un acteur à part entière de GASP ce qui autorise, outre de pouvoir se déplacer dans l'environnement, de communiquer avec les autres acteurs du monde virtuel dans lequel elle évolue. Ce couplage devrait d'autre part autoriser, à court terme, le portage des algorithmes de navigation interactive sur la machine de réalité virtuelle Imersia de l'Irisa.

6.3.7 Plate-forme véhicule CyCab

Participants : Isabelle Lacoffrette, Fabien Spindler.

Dans le cadre des travaux de recherche et développement menés dans le domaine des transports, l'Inria a développé un prototype expérimental de petit véhicule électrique équipé de capteurs: le CyCab. Ces véhicules ont été fabriqués par la société Robosoft et équipent les cinq unités de recherche de l'Inria, dont celle de Rennes et notre projet depuis juillet 2000. Notre objectif est de constituer une plate-forme expérimentale permettant de développer des travaux dans les domaines de la robotique mobile et des transports. Sont visées des stratégies de navigation automatisées à l'aide d'une caméra embarquée et d'une autre déportée, du traitement de séquences d'images et de l'analyse du mouvement, et enfin du couplage entre action et perception. Nous avons principalement travaillé pour l'instant à l'instrumentation des capteurs associés à cette opération. Cela concerne d'une part la mise en route matérielle et logicielle d'une tourelle site et azimut VIC-2400 destinée à observer l'évolution du CyCab (installation fixe d'observation de la scène), et d'autre part le développement d'une démonstration d'asservissement visuel utilisant la tourelle site et azimut PTU-46 permettant de valider ses fonctionnalités et destinée à être embarquée sur le véhicule.

7 Contrats industriels (nationaux, européens et internationaux)

7.1 Convention DGA - Thomson-CSF-Optrosys : Conception d'une chaîne de traitements pour la restauration de séquences d'images

Participants : Patrick Bouthemy, Patrick Pérez, Fabien Dekeyser.

no. 1 98 C 500, durée 24 mois

Cette étude est menée en collaboration avec Thomson-CSF-Optronique avec un financement DGA, dans le cadre du programme Civil-Défense. Deux applications sont prises en compte, l'une relative au traitement de séquences d'images infra-rouges aériennes, l'autre au traitement d'images médicales endoscopiques vidéo. Notre rôle consiste d'une part en l'étude et le développement d'algorithmes originaux de 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 est non seulement de débruiter ces images ou de supprimer des défauts localisés, mais également de construire des images de plus grande résolution à partir de la séquence originelle. Il est évident qu'une restauration de nature spatio-temporelle efficace doit tenir compte du mouvement présent dans la séquence d'images. Le travail effectué cette année est décrit à la sous-section

6.1.2) et dans l'article [55]. Des validations expérimentales des algorithmes développés de filtrage et de super-résolution ont donné des résultats très prometteurs sur des séquences de 100 à 200 images infra-rouges aériennes fournies par TCO. Dans le cadre de cette application, où l'on rencontre des mouvements du capteur aéroporté de translation, de rotation et de zoom de faible amplitude, pour une scène quasi-planaire étant donné la distance d'observation, un modèle paramétrique de mouvement apparent 2D à 6 (modèle affine) ou à 8 paramètres (modèle quadratique) s'avère très bien adapté. Le faible coût de calcul de ces algorithmes permet d'envisager leur utilisation à terme dans des applications temps-réel.

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

Participants : Étienne Mémin, Patrick Pérez, Stelios Zimeras.

no. 1 99 C 351, durée 15 mois.

Dans le cadre des études lancées par le consortium météorologique européen Eumetsat autour de l'exploitation des images du futur satellite Météosat Seconde Génération, nous avons mené un travail en collaboration avec le LMD (Laboratoire de Météorologie Dynamique) de comparaison de méthodes d'estimation de champs de vents couplées avec des méthodes de détection de nuages bas. Pour notre part, nous avons dans ce cadre comparé le comportement, sur des images Météosat, d'un estimateur de mouvement générique [MP98] et d'un estimateur de mouvement adapté aux séquences d'images de fluides [50]. L'estimateur de mouvement dédié s'est avéré plus performant que l'estimateur générique en termes de précision des mesures et de respect de la structure du champ des vitesses [88, 76].

7.3 Convention Ifremer : Commande d'un bras manipulateur non instrumenté

Participants : François Chaumette, Eric Marchand, Fabien Spindler.

no. 1 00 C 001, durée 12 mois

L'objectif de cette étude est de commander les mouvements du bras manipulateur embarqué sur le robot sous-marin Victor 6000 à l'aide d'une caméra observant l'effecteur du manipulateur. Sur le plan pratique, l'intérêt d'utiliser les informations visuelles fournies par la caméra vient du fait que ce bras n'est pas instrumenté de capteurs proprioceptifs. On ne dispose alors plus de moyen de mesure des mouvements de ce bras, ce qui rend toute commande imprécise, à moins de fermer la boucle de commande à l'aide d'un capteur extéroceptif (dans le cas présent, la caméra). Un deuxième intérêt réside dans l'indépendance de la position atteinte par le bras par rapport aux diverses erreurs de modélisation et de calibration, aussi bien du capteur que du bras, s'il est possible d'exprimer la tâche assignée au manipulateur directement dans l'espace de mesure du capteur (en l'occurrence le plan image de la caméra). Avant d'effectuer un déplacement, une phase préalable de calcul de la consigne dans l'image est tout d'abord réalisée. Celle-ci consiste à localiser la position de l'effecteur dans le repère de la caméra

[MP98] E. Mémin, P. Pérez, « A multigrid approach to hierarchical motion estimation », *in : Proc. Int. Conf. on Computer Vision, ICCV'98*, p. 933–938, Bombay, Inde, janvier 1998.

par un calcul de pose, à calculer sa position désirée à partir du déplacement spécifié, puis à projeter cette position désirée dans le plan image de la caméra, ce qui fournit la consigne attendue. L'orientation de la caméra étant également contrôlée par asservissement visuel 2D (afin que l'effecteur soit toujours dans le champ de vision de celle-ci), il convient de tenir compte de son déplacement (mesuré par odométrie) pour calculer la nouvelle consigne à atteindre à chaque itération de la loi de commande. Ensuite, nous avons utilisé une technique classique d'asservissement visuel 3D pour amener la position courante de l'organe terminal à sa position désirée. Cette phase d'asservissement requiert l'utilisation du jacobien inverse du manipulateur. Comme ce jacobien dépend de la valeur des coordonnées articulaires du manipulateur, et que ces valeurs sont inconnues, la position de l'organe terminal est estimée par un calcul de pose, et on en déduit les valeurs recherchées en résolvant les équations du modèle géométrique inverse. L'ensemble du procédé a été mis en œuvre et validé sur la cellule expérimentale du projet [96].

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

Participants : Jean-Pierre Le Cadre, Isabelle Leibowicz.

no. KTNR, durée 36 mois.

L'objet de la thèse d'Isabelle Leibowicz (bourse Cifre) est l'optimisation des fonctions de surveillance et d'identification d'un avion de patrouille maritime. Un tel système constitue un concentré de traitements et de problèmes multi-récepteurs et multi-cibles. Ainsi, si l'avion dispose de nombreux récepteurs, radars, ESM passif, infrarouge, sonar, il doit aussi traiter, extraire, associer, identifier un grand nombre de pistes. Les problèmes relatifs à l'association piste-à-piste et des mesures aux pistes, à l'apport d'informations d'identification, à la combinaison des diverses sources d'information radar, ESM, infrarouge, ont constitué une partie importante de l'activité. Les résultats obtenus mettent en évidence l'importance de l'information d'identification, essentiellement radar. L'accroissement des performances, en termes de probabilité d'association correcte, a pu être quantifié. Un autre volet important est relatif à l'utilisation des données ESM capteur passif, pour les fonctions de pistage. Une phase d'analyse est également nécessaire afin de déterminer la nature de l'environnement et de décider de l'utilisation des ressources disponibles. Il s'agit en particulier de retracer l'historique des cibles et de reconnaître une situation opérationnelle spécifique. La modélisation de la situation opérationnelle permet d'intégrer les aspects probabilistes et causaux du déroulement d'une mission et des informations perçues par les capteurs et les opérateurs.

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

Participants : Patrick Bouthemy, Emmanuel Veneau.

no. 1 99 C 222, durée 36 mois.

Cette convention concerne la thèse de Emmanuel Veneau, encadrée à l'Ina par Rémi Ronfard et plus récemment 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 [78]. 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 simplement de mesures spatio-temporelles locales et par la définition de schémas de classification hiérarchique par SVM ("Support Vector Machines"). Cette approche est appliquée à un document vidéo de retransmission d'un événement sportif, en l'occurrence des championnats d'athlétisme, qui forme une étude de cas représentative et complexe.

7.6 Actions financées par la Commission Européenne

7.6.1 Projet Esprit (R&D MS) Divan

Participants : Patrick Bouthemy, Fabien Spindler.

no. 1 97 C 696, durée 30 mois

Nous avons participé, avec le projet Sigma-2, au projet européen Divan ("Distributed audio-Visual Archives Network"), relevant du programme Esprit R&D Multimedia Systems, qui s'est clos à la mi-2000. Les partenaires en étaient les sociétés Intracom (Gr.) et Tecmath (All.), la RAI (It.), l'ERT (Gr.), Forth-ICS (Gr.), THD-ITO (All.), l'Ina et l'Inria. Ce projet a porté sur des problèmes de structuration, d'annotation, de consultation et d'exploitation d'archives audiovisuelles multiples. Ceci comprend notamment la prise en compte de requêtes liées aux contenus des vidéos, et donc implique en amont la réalisation d'une indexation par le contenu. Il visait à inclure de plus la consultation des archives audio-visuelles dans un mode distribué sur le réseau. Nous sommes intervenus sur des questions de segmentation et d'indexation par le contenu spatio-temporel de documents vidéos. Nous avons porté dans le prototype développé nos logiciels de découpage de la vidéo en plans, de caractérisation du mouvement de la caméra, et de détection d'éléments mobiles. Nous les avons adaptés pour une insertion adéquate dans la phase de décodage du flux MPEG et dans les schémas de description retenus. Nous avons de plus développé et validé une extension de l'algorithme de découpage en plans pour inclure la détection des transitions progressives de type volets. Nous avons aussi conçu une technique robuste et temps réel de détection de logos dans les images d'un plan (logos de chaînes de télévision par exemple). Nous avons enfin construit un module permettant de manière simple de caractériser le "niveau d'activité" dans la scène filmée.

7.6.2 Projet IST Carsense

Participants : Patrick Bouthemy, Eric Marchand, Gildas Lefaix.

no. 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 intervenons dans le workpackage “Computer Vision” de ce projet où nous devons développer un module de détection d'obstacles par analyse spatio-temporelle d'une séquence monoculaire. 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, doit se traduire par des erreurs de compensation. Outre la localisation des obstacles potentiels, nous cherchons à calculer des informations de temps à collision et les trajectoires des obstacles détectés.

7.7 Actions financées par le RNRT

7.7.1 Projet Agir

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

no. 2 99 C 006, durée 18 mois

Le projet AGIR est un projet RNRT précompétitif qui a officiellement débuté au 1er septembre 1999. Il regroupe l'Ina (maître d'œuvre), les sociétés CS-Systèmes d'Information, Arts Vidéo Interactive et Mémodata, pour l'Irisa les projets Sigma-2 et Vista, le projet Movi de l'Inria Rhône-Alpes, et des équipes de recherche de l'Irit, du Clips-Imag, de l'INT, et du LIP6. Ce projet est financé pour 18 mois, avec un renouvellement possible de 12 mois.
Le but du projet est de développer un système complet permettant d'indexer des vidéos, éventuellement accompagnées de textes, puis d'utiliser cette indexation pour rechercher des documents par leur contenu, ceci en mélangeant les divers médias. Ce projet prend donc naturellement la suite du travail fait dans le cadre du projet européen Divan, avec un accent plus fort mis sur l'aspect multimédia des données, des index et des requêtes, sur l'architecture globale du projet et sur les liens avec les efforts de normalisation actuels (MPEG-7). Les travaux développés par Vista concernent principalement la structuration temporelle automatique des séquences d'images par l'utilisation directe des informations issues de la compression MPEG-2, comme les vecteurs de mouvement, l'extraction et le suivi d'entités mobiles pertinentes et la fourniture de descripteurs d'activité et de mouvement (voir sous-section 6.1.6 et rapport [85]). Le projet intervient aussi, mais de manière moins forte, dans la mise au point d'algorithmes de macro-segmentation et de regroupement de plans en entités de plus haut niveau sémantiquement plus significatives [78]. Par ailleurs, les sorties des modules développés seront exploitées dans d'autres tâches du projet.

7.8 Actions soutenues par le programme Priamm

7.8.1 Projet Mediaworks

Participants : Patrick Bouthemy, Patrick Gros.

no. 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 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 va consister 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 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.

Une subvention de la Région a été obtenue pour contribuer à l'achat d'un équipement d'acquisition d'échographies 3D robotisé. 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 des images échographiques. Les originalités du projet résideront d'une part 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. Ce projet se fera en collaboration avec deux centres d'expérimentation cliniques: les CHU de Tours (Docteurs Pourcelot et Tranquart) et de Rennes (Docteurs Darnault et Tréguier).

8.2 Actions nationales

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

Participants : Éric Marchand, François Chaumette.

no. 2 L 008, durée 24 mois.

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

8.2.2 Convention GIS Sciences de la Cognition : Imagerie cérébrale

Participants : Christian Barillot, Pierre Hellier.

no. 1 97 C 731, durée 24 mois

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

La dernière phase de cette étude a consisté à comparer cinq méthodes de recalage non-rigide de volumes cérébraux, le quadrillage proportionnel de Talairach, la méthode ANIMAL développée par Louis Collins à Montréal, la méthode fluide de Christensen (Univ. Iowa), la méthode des démons développée dans le projet Epidaure et notre méthode de recalage [66]. Ces méthodes ont été comparées selon des critères globaux (recouvrement des volumes Lvv) et locaux (appariement des sillons corticaux) sur la base de 18 sujets. Dans ce même contexte, nous avons étudié également l'influence de différents recalages sur la localisation de données fonctionnelles (activités motrice et somesthésique enregistrées en MEG). Cette évaluation de méthodes de recalage inter-indivus s'est de plus étendue à une action coordonnée par C. Barillot avec l'Université d'Iowa (G. Christensen, M. Vannier, J. Haller).

8.2.3 Actions de recherche coopératives de l'Inria

- Imagerie cérébrale

  Participants : Christian Barillot, Pierre Hellier.

  Nous participons 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 Salpétrière). 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-MagnétoEncéphaloGraphie, l'autre sur l'IRM fonctionnelle. Notre contribution dans ces deux actions porte sur le développement de méthodes de recalage non rigides 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). L'objectif est également de définir un cadre méthodologique de validation des techniques de recalage en imagerie cérébrale, à la fois pour la correction du bougé de 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.

– AVEC, Asservissement Visuel en Environnement Complexe
  Participants : François Chaumette, Éric Marchand, Gregory Flandin, Youcef Mezouar.

  Cette action de recherche coopérative, intitulée Asservissement Visuel en Environnement Complexe (AVEC), regroupe l'ensemble des projets de l'Inria travaillant dans le domaine de l'asservissement visuel, à savoir, outre Vista, les projets Icare (Sophia-Antipolis), Movi et Bip (Rhône-Alpes). Elle est coordonnée par F. Chaumette. Le premier thème scientifique majeur traité dans cette ARC porte sur l'élaboration d'algorithmes de suivi de primitives géométriques dans des séquences d'images complexes. L'objectif est de pouvoir ouvrir le champ applicatif des techniques d'asservissement visuel à des environnements réels (et donc complexes) sur lesquels peu de connaissance a priori est disponible (pas de modèle CAO des objets considérés par exemple). Le second thème porte sur le choix des informations visuelles à utiliser dans la commande en fonction de la tâche à réaliser. La difficulté provient de la très grande masse d'informations contenues dans une séquence d'images qui rend extrêmement délicate la sélection automatique et optimale des seules informations pertinentes. Différents critères de sélection sont étudiés, notamment la robustesse du suivi par traitement d'images et les éventuelles propriétés de découplage des lois de commande associées. Deux réunions de deux jours se sont déroulées cette année, l'une à Rennes et l'autre à Sophia-Antipolis. Comme l'an dernier, des séjours ponctuels de membres d'un projet dans un autre projet ont également eu lieu. Ainsi, P. Rives, du projet Icare, est venu à Rennes pour porter et valider sur notre cellule expérimentale de vision robotique un nouveau schéma d'asservissement visuel exploitant des contraintes épipolaires .

8.3 Relations bilatérales internationales

8.3.1 Programme Alliance France-Angleterre

Participants : Christian Barillot, Patrick Bouthemy, Ioannis Pratikakis, François Rousseau.

Dans le cadre de cette action, nous nous proposons d'étudier la faisabilité de l'utilisation des ultrasons 3D de type “main libre” dans différents domaines applicatifs, comme par exemple la radiologie chez le nourrisson et le contrôle par l'image per-opératoire ultrasonore de l'asservissement et du guidage de gestes chirurgicaux dans un environnement de neuronavigation (en lien avec des partenaires cliniques à Rennes et à Cambridge). Cette collaboration a permis la mise en place sur le site de Rennes d'une plate-forme d'expérimentation d'acquisition d'échographies 3D utilisant un localisateur 3D de type optique. Le système testé offre une plus grande précision que le système classique, précision nécessaire pour le développement futur d'applications en neurochirurgie dans un contexte de neuro-navigation. Deux réunions ont eu lieu cette année (une à Rennes en septembre avec la venue de A. Gee et R. Prager, et une à Cambridge en décembre, séjour de C. Barillot et F. Rousseau). Une campagne de mesure a pu avoir lieu au CHU de Rennes, en relation avec les Docteurs P. Darnault et C. Treguier, associant le logiciel StradX de l'Université de Cambridge permettant la reconstruction du volume échographique 3D à partir d'une suite d'acquisitions 2D, et le système de localisation optique développé à Rennes.

8.3.2 Programme France-Tunisie

Participants : Patrick Bouthemy, Patrick Gros.

Nous participons à la collaboration franco-tunisienne mise en place par l'Inria, le Cert et l'Ensi de Tunis. Le projet retenu concerne l'indexation multimédia et regroupe les projets Vista, Movi de l'U.R. Rhône-Alpes, et Imedia de l'U.R. Rocquencourt. Dans le cadre de ce projet, P. Gros et P. Bouthemy ont effectué un séjour d'une semaine à Tunis en avril 2000. Ils ont donné un cours de DEA à l'Ensi de Tunis sur l'indexation des images et des vidéos. Ils ont également participé à un colloque sur l'indexation multimédia à Tunis réunissant des chercheurs tunisiens et français (exposé de P. Bouthemy intitulé "Description du mouvement pour l'indexation et la consultation de vidéos" et de P. Gros intitulé "Algorithmes d'indexation pour données multidimensionnelles imprécises") et à une émission scientifique sur Radio Tunis International. En décembre, P. Gros a donné un nouveau cours en DEA à l'Ensi. Rim Mezhoud, doctorante de l'Ensi de Tunis sous la direction du Prof. Faouzi Ghorbel, a effectué dans ce cadre un séjour de 3 mois dans le projet, de mi-novembre 1999 à mi-février 2000, sous la direction de P. Gros et de L. Amsaleg (du projet Temics), et qui a donné lieu à un travail sur l'indexation de descripteurs numériques de grande taille [80].

8.3.3 Accueil de chercheurs étrangers

- Mario Ramalho, Maître de conférence à l'IST de Lisbonne, effectue un séjour de six mois au sein du projet pour travailler dans le domaine de l'asservissement visuel et du suivi.
- Prof. Christoph Schnörr et Daniel Cremers (Université de Mannheim, Allemagne) ont effectué une visite d'une semaine dans le projet en septembre.
- Roy Streit (NUWC, Newport, USA), M. Malyutov (Northeastern Univ, Boston, USA), D. Vandermeulen (Katholieke Universiteit Leuven, Belgique), E. Wersink (Thomson-Signaal, Pays-Bas), et J.-F. Lots (Heriot Watt Univ., Edimbourg) ont effectué des visites de courte durée au sein du projet.

8.3.4 Expertise de projets

- C. Barillot est expert de la Commission Européenne pour l'évaluation des projets de recherche du 5ème programme cadre du thème "Quality of life and management of living resources : research and technological development activities of a generic nature"; 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.

9 Diffusion de résultats

9.1 Animation de la communauté scientifique

- C. Barillot anime un groupe de réflexion sur la neuroimagerie 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 neuroimagerie clinique et cognitive. Ce groupe comprend à la fois des chercheurs en sciences pour l'ingénieur et en sciences de la vie (CNRS, Inria, Inserm, Université de Rennes), des médecins du CHU de Rennes (neurologues, neurochirurgiens, radiologues, neurphysiologistes, pharmacologistes, ...) et des chercheurs en sciences de l'homme et de la société (linguistes et cogniticiens). 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 posé par l'indexation des documents multimédia 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.
  - P. Bouthemy a fait un exposé sur l'indexation vidéo lors de la réunion commune aux GT-4 (Modèles déformables dynamiques) et GT-10 (Indexation multimedia) en avril 2000. G. Flandin a présenté un poster sur ses travaux en vision active lors des journées thématiques 2000 du GDR ISIS.
  - F. Chaumette co-anime avec M. Devy du LAAS le thème "vision dynamique et active" du GT-5 Vision. Une réunion sur le thème de l'asservissement visuel et de la vision active a été organisée à l'Irisa les 14 et 15 décembre 2000.
- J.-P. Le Cadre est membre du groupe de travail "Performances ultimes en trajectographie" de Thomson-CSF.
- J.-P. Le Cadre a fait partie du comité d'évaluation du département DTIM de l'Onera.
- P. Bouthemy est éditeur associé de la revue IEEE *Trans. on Image Processing* et est membre du comité de rédaction de la revue TSI.

- P. Gros est membre du comité de rédaction de la revue *Traitement du Signal*.
- P. Pérez est éditeur associé de la revue IEEE *Trans. on Image Processing*.
- F. Chaumette est co-éditeur avec R. Horaud d'un numéro spécial de la revue *Int. Journal of Computer Vision* dédié à l'asservissement visuel et paru en juin 2000 [24].
- P. Bouthemy est co-éditeur avec C. Schnörr (Université de Mannheim, Allemagne) d'un numéro spécial de la revue *Journal of Mathematical Imaging and Vision* sur l'analyse du mouvement fluide dans des images, prévu pour fin 2001.
- Participation aux Comités de Programme:
  - C. Barillot : conférences internationales ICIP'2001, IPMI'2001.
  - P. Bouthemy : conférences internationales ECCV'2000, Visual'2000, RIAO'2000, congrès national RFIA'2000 (également membre du Comité éditorial).
  - F. Chaumette : conférences internationales ICRA'2000, CVPR'2000, ICCV'2001.
  - J.-P. Le Cadre : conférence internationale Fusion'2000 (également, membre du Comité d'organisation).
  - P. Gros : journées nationales Orasis'2001.
  - E. Mémin : conférence internationale ICCV'2001.
  - P. Pérez : conférence internationale ICIP'2000.
- D. Albert, G. Flandin et Y. Mezouar (avec N. Courty du projet Siames) ont organisé les 13èmes Journées Jeunes Chercheurs en Robotique (JJCR) qui se sont tenues à l'Irisa les 28 et 29 septembre.

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, trajectographie et extraction-poursuite).
- DEA informatique, Ifsic, Rennes 1 (E. Marchand : modèles 3D pour la synthèse et la vision par ordinateur; E. Mémin : analyse du mouvement).
- Diic-INC, Ifsic, Rennes 1 : (E. Mémin, F. Chaumette, G. Flandin : vision 3D, analyse du mouvement, asservissement visuel; E. Mémin : modèles markoviens en analyse d'images; C. Barillot : imagerie médicale).
- Insa Rennes, Génie électrique 5ème année, option image (F. Chaumette, T. Corpetti, R. Fablet, Y. Mezouar : 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 Electronique, option Image, Ireste, université de Nantes (P. Bouthemy : analyse du mouvement).

- DEA GBM universités Tours-Angers-Nantes-Rennes (C. Barillot : imagerie 3D appliquée à la médecine).
- Stages effectués au sein de l'équipe : R. Chasle (DEA et Insa Info, Rennes), D. Duval (DEA Info, Ifsic, Rennes), M. Lalesnel (INSA Rennes, génie électrique, 5ème année), R. Peteri (DEA Photonique et Image, ENSPS, Strasbourg), G. Teniou (DIIC INC, Ifsic, Rennes), Y. Tréglos (INSA Rennes, 4ème année).
- Direction de thèse hors projet:
  - J.-N. Provost (Ecole Navale, Brest) par P. Bouthemy et P. Pérez;
  - A. Alhaj (Cemagref, Rennes) par F. Chaumette;
  - E. Kijak (Thomson Multimedia, bourse Cifre) par P. Gros;
  - R. Donati (Gesma Brest) par J.P. Le Cadre, thèse soutenue en juin 2000 à Rennes 1, titre: "Détection spatio-temporelle de signaux électriques océaniques - Application à la surveillance de zones";
  - direction de DRT de S. Marrec (Thomson Multimédia) par P. Gros, soutenance en novembre 2000 à l'Ifsic, titre: "Sélection d'images-clés par analyse de champs de vecteurs MPEG".

9.3 Participation à des colloques, séminaires, invitations

- F. Chaumette a été invité pour une visite de quelques jours à l'IST de Lisbonne et y a donné une conférence invitée lors du 5th Ibero-American Symposium on Pattern Recognition en septembre 2000, [46].
- F. Chaumette a été invité en novembre pour une semaine au sein du Human Computer Interaction group de Rockwell Science Center à Thousand Oaks, Californie, USA (conseil en asservissement visuel et calcul de pose).
- P. Gros a été invité au colloque "Sciences des réseaux" organisé conjointement par les deux académies des sciences française et chinoise et qui s'est tenu à Beijing en octobre 2000. L'exposé avait pour titre "The non-networks aspects of image and video archives".
- J.-P. Le Cadre a fait une présentation relative à l'extraction multipistes au séminaire de statistique de Paris-VII et au NATO ASI "Multisensor Data Fusion" (Pitlochry, juillet 2000).
- R. Fablet a été invité à présenter ses travaux lors du workshop "Mathematics and Image Analysis" (MIA'2000) qui s'est tenu à Paris en septembre 2000 (titre de l'exposé: "Non-parametric scene activity analysis for statistical retrieval with partial query").
- F. Dambreville et F. Dekeyser ont participé aux journées doctorales de la DGA qui se sont tenues à Paris en octobre 2000. Ils ont présenté à cette occasion un poster sur leurs travaux de thèse.

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] A. Chardin, *Modèles énergétiques hiérarchiques pour la résolution des problèmes inverses en analyse d'images, Application à la télédétection*, thèse de doctorat, Université de Rennes 1, mention Traitement du signal et Télécommunications, janvier 2000.

[12] P. Hellier, *Recalage non rigide en imagerie cérébrale : méthode et validation*, thèse de doctorat, Université de Rennes 1, mention Traitement du signal et Télécommunications, décembre 2000.

Articles et chapitres de livre

[13] L. Amsaleg, P. Gros, « Content-based retrieval using local descriptors : problems and issues from a database perspective », *Pattern Analysis and Applications*, à paraître.

[14] C. Baillard, P. Hellier, C. Barillot, « Segmentation of brain 3D MR images using level sets and dense registration », *Medical Image Analysis*, à paraître.

[15] F. Chaumette, « Asservissement visuel. », *in : Les bases de la robotique*, E. Dombre (éditeur), *Traité IC2*, Hermes, 2000, ch. 7.

[16] N. Courty, E. Marchand, « Navigation et contrôle d'une caméra dans un environnement virtuel: une approche référencée image. », *Technique et Science Informatiques, TSI*, à paraître.

[17] A. Crétual, F. Chaumette, « Dynamic stabilization of a pan and tilt camera for sub-marine image visualization », *Computer Vision and Image Understanding 79*, 1, juillet 2000, p. 47–65.

[18] F. Dekeyser, P. Pérez, P. Bouthemy, « Restoration of noisy, blurred, undersampled image sequences using a parametric motion model », *Réseaux et systèmes répartis - calculateurs parallèles 12*, 3-4, numéro spécial Image et vidéo, décembre 2000, p. 387–405.

[19] R. Fablet, P. Bouthemy, « Non-parametric scene activity analysis for statistical retrieval with partial query », *Journal of Mathematical Imaging and Vision*, à paraître.

[20] M. Gelgon, P. Bouthemy, « A region-level motion-based graph representation and labeling for tracking a spatial image partition », *Pattern Recognition 33*, 4, avril 2000, p. 725–745.

[21] N. Giordana, P. Bouthemy, F. Chaumette, F. Spindler, J.-C. Bordas, V. Just, « Two dimensional model-based tracking of complex shapes for visual servoing tasks », *in : Robust vision for vision-based control of motion*, M. Vincze et G. Hager (éditeurs), IEEE Press, 2000, ch. 6, p. 67–75.

[22] P. Hellier, C. Barillot, E. Mémin, P. Pérez, « Hierarchical estimation of a dense deformation field for 3D robust registration », *IEEE Transactions on Medical Imaging*, à paraître.

[23] P. Hellier, C. Barillot, E. Mémin, P. Pérez, « Estimation robuste 3D d'un champ de déformation pour le recalage non-linéaire inter-sujet d'images cérébrales », *Traitement du Signal 17*, 2, 2000.

[24] R. Horaud, F. Chaumette, « Guest editorial of special issue on image-based robot servoing », *Int. Journal of Computer Vision 37*, 1, juin 2000, p. 5–6.

[25] L. Hubert-Moy, A. Cotonnec, L. Le Du, A. Chardin, P. Pérez, « A comparison of parametric classification procedures of remotly sensed data applied on different landscape units », *Remote Sensing of Environment*, à paraître.

[26] 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.

[27] J.-P. Le Cadre, H. Gauvrit, F. Trarieux, « Approximations of the Cramér-Rao bound for multiple-target motion analysis », *IEE Proceedings Radar, Sonar and Navigation 147*, 3, juin 2000, p. 105–113.

[28] J.-P. Le Cadre, G. Souris, « Searching tracks », *IEEE Transactions on Aerospace & Electr. Systems 36*, 4, octobre 2000.

[29] J.-P. Le Cadre, « Data association and multitarget tracking », *in : Multisensor data fusion*, A. Hyder (éditeur), Kluwer, 2000.

[30] G. Le Goualher, A.-M. Argenti, M. Duyme, W. Baare, H. Hulshoff Pol, C. Barillot, A. Evans, « Statistical sulcal shape comparisons: application to the detection of genetic encoding of the central sulcus shape. », *NeuroImage 11*, 5, mai 2000, p. 564–574.

[31] E. Malis, F. Chaumette, « 2 1/2 D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement », *Int. Journal of Computer Vision 37*, 1, juin 2000, p. 79–97.

[32] E. Marchand, P. Bouthemy, F. Chaumette, « A 2D-3D model-based approach to real-time visual tracking », *Image and Vision Computing*, à paraître.

[33] E. Marchand, F. Chaumette, « Reconstruction 3D par vision dynamique active », *in : Inférence d'informations volumiques par imagerie vidéo*, M. Dhome (éditeur), *Traité IC2*, Hermes, 2000.

[34] M. Mignotte, C. Collet, P. Pérez, P. Bouthemy, « Hybrid genetic optimization and statistical model-based approach for the classification of shadow shapes in sonar imagery », *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 2, février 2000, p. 129–141.

[35] M. Mignotte, C. Collet, P. Pérez, P. Bouthemy, « Markov random field and fuzzy logic modeling in sonar imagery: application to the classification of underwater floor », *Computer Vision and Image Understanding 79*, 1, juillet 2000, p. 4–24.

[36] M. Mignotte, C. Collet, P. Pérez, P. Bouthemy, « Sonar image segmentation using a hierarchical MRF model », *IEEE Transactions on Image Processing IP-9*, 7, juillet 2000, p. 1216–1231.

[37] E. Mémin, T. Risset, « VLSI design methodoloy for edge-preserving image reconstruction », *Real Time Imaging*, à paraître.

[38] E. Mémin, T. Risset, « On the study of VLSI derivation for optical flow estimation », *Int. Journal of Pattern Recogition and Artificial Intelligence 14*, 4, juin 2000, p. 441–462.

[39] P. Pérez, A. Chardin, J.-M. Laferté, « Noniterative manipulation of discrete energy-based models for image analysis », *Pattern Recognition 33*, 4, avril 2000, p. 573–586.

[40] Y. Ricquebourg, P. Bouthemy, « Real-time tracking of moving persons by exploiting spatio-temporal image slices », *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 8, août 2000, p. 797–808.

Communications à des congrès, colloques, etc.

[41] C. Baillard, C. Barillot, « Robust 3D segmentation of anatomical structures with level sets », *in: Proc. of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2000, LNCS 1935*, p. 236–245, Pittsburgh, Pennsylvania, octobre 2000.

[42] C. Baillard, P. Hellier, C. Barillot, « Cooperation between level set techniques and dense 3D registration for the segmentation of brain structures », *in: IAPR Int. Conf. on Pattern Recognition, ICPR 2000*, *1*, p. 991–994, Barcelone, Espagne, septembre 2000.

[43] C. Baillard, P. Hellier, C. Barillot, « Segmentation of brain structures using level sets and dense registration », *in: IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, p. 94–103, Hilton Head Island, Caroline du Sud, juin 2000.

[44] F. Chaumette, E. Malis, « 2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoings », *in: IEEE Int. Conf. on Robotics and Automation*, *1*, p. 630–635, San Francisco, avril 2000.

[45] F. Chaumette, E. Marchand, « A new redundancy-based iterative scheme for avoiding joint limits: application to visual servoing », *in: IEEE Int. Conf. on Robotics and Automation*, *2*, p. 1720–1725, San Francisco, avril 2000.

[46] F. Chaumette, « Modeling issues in visual servoing », *in: 5th Ibero-american Symp. on Pattern Recognition, SIARP 2000*, p. 3–18, Lisbonne, Portugal, septembre 2000.

[47] C. Collet, J.-N. Provost, P. Rostaing, P. Pérez, P. Bouthemy, « Spot satellite data analysis for bathymetric mapping », *in: IEEE Int. Conf. on Image Processing, ICIP'2000*, Vancouver, Canada, septembre 2000.

[48] C. Collewet, F. Chaumette, L. Wallian, « 2D visual servoing on complex shapes based on a polar description », *in: 6th IFAC Int. Symp. on Robot Control, SYROCO'2000*, *2*, p. 711–716, Vienne, Autriche, septembre 2000.

[49] C. Collewet, F. Chaumette, « A contour approach for image-based control of objects with complex shape », *in: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'2000*, Takamatsu, Japon, novembre 2000.

[50] T. Corpetti, E. Mémin, P. Pérez, « Adaptation of standard optic methods to fluid motion », *in: Int. Symposium on Flow Visualization*, *paper 62*, p. 1–10, Edimbourg, août 2000.

[51] T. Corpetti, E. Mémin, P. Pérez, « Estimating fluid optical flow », *in: IAPR Int. Conf. on Pattern Recognition, ICPR'2000*, *3*, p. 1045–1049, Barcelone, Espagne, septembre 2000.

[52] A. Crétual, F. Chaumette, « Image-based positioning with respect to a non-structured scene using 2D image motion », *in: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'2000*, Takamatsu, Japon, novembre 2000.

[53] F. Dambreville, J.-P. Le Cadre, « Detection with spatial and temporal optimization of search efforts involving multiple modes and multiple resources management », *in: Int. Conf. on Information Fusion, FUSION'2000, II*, p. 9–16, Paris, France, juillet 2000.

[54] F. Dambreville, J.-P. Le Cadre, « Spatial and temporal optimization of search efforts for the detection of a markovian target », *in: Imacs World Congress*, Lausanne, Suisse, août 2000.

[55] F. Dekeyser, P. Bouthemy, P. Pérez, E. Payot, « Super-resolution from noisy image sequences exploiting a 2D parametric motion model », *in : IAPR Int. Conf. on Pattern Recognition, ICPR'2000*, *3*, p. 354–357, Barcelone, Espagne, septembre 2000.

[56] F. Dekeyser, P. Bouthemy, P. Pérez, « Spatio-temporal Wiener filtering of image sequences using a parametric motion model », *in : IEEE Int. Conf. on Image Processing, ICIP'2000*, Vancouver, Canada, septembre 2000.

[57] F. Dekeyser, P. Pérez, P. Bouthemy, « Restoration of noisy, blurred, undersampled image sequences using a parametric motion model », *in : Inter. Symposium on Image and Video Communications, ISIVC'2000*, Rabat, Maroc, avril 2000.

[58] R. Fablet, B. Bouthemy, P. Pérez, « Statistical motion-based video indexing and retrieval », *in : Conf. on Content-Based Multimedia Information Access, RIAO'2000*, Paris, France, avril 2000.

[59] R. Fablet, P. Bouthemy, « Statistical motion-based object indexing using optic flow field », *in : IAPR Int. Conf. on Pattern Recognition, ICPR 2000*, *4*, p. 287–290, Barcelone, Espagne, septembre 2000.

[60] R. Fablet, P. Bouthemy, « Statistical motion-based retrieval with partial query », *in : 4th Int. Conf. on Visual Information Systems, Visual'2000, Vol. LNCS 1929, Springer-Verlag*, p. 96–107, Lyon, novembre 2000.

[61] G. Flandin, F. Chaumette, E. Marchand, « Coopération caméra embarquée / caméra déportée pour l'asservissement visuel », *in : 12ème Journées des Jeunes Chercheurs en Robotique*, p. 58–63, Bourges, février 2000.

[62] G. Flandin, F. Chaumette, E. Marchand, « Eye-in-hand / eye-to-hand cooperation for visual servoing », *in : IEEE Int. Conf. on Robotics and Automation*, *3*, p. 2741–2746, San Francisco, avril 2000.

[63] P. Gros, « Color illumination models for image matching and indexing », *in : IAPR Int. Conf. on Pattern Recognition, ICPR'2000*, *3*, p. 590–883, Barcelone, Espagne, septembre 2000.

[64] P. Gros, « Experimental evaluation of color illumination models for image matching and indexing », *in : Conf. on Content-Based Multimedia Information Access, RIAO'2000*, p. 567–574, Paris, France, avril 2000.

[65] P. Hellier, C. Barillot, E. Mémin, P. Pérez, « An energy-based framework for dense 3D registration of volumetric brain image », *in : IEEE Conf. Computer Vision and Pattern Recognition, CVPR'2000*, *II*, p. 270–275, Hilton Head Island, South Carolina, USA, juin 2000.

[66] P. Hellier, C. Barillot, « Multimodal non-rigid warping for correction of distortions in functional MRI », *in : Proc. of Medical Image Computing and Computer-Assisted Intervention, MICCAI'00, LNCS 1935*, p. 512–520, Pittsburgh, Pennsylvania, octobre 2000.

[67] B. Lamiroy, P. Gros, S. Picard, « Combining local recognition methods for better image recognition », *in : British Machine Vision Conference, BMVC'2000*, *2*, p. 735–744, Bristol, UK, septembre 2000.

[68] J.-P. Le Cadre, « Scheduling active and passive measurements », *in : Int. Conf. on Information Fusion, FUSION'2000*, Paris, juillet 2000.

[69] E. Malis, F. Chaumette, S. Boudet, « Multi-cameras visual servoing », *in : IEEE Int. Conf. on Robotics and Automation*, *4*, p. 3183–3188, San Francisco, avril 2000.

[70] E. Marchand, N. Courty, « Image-based virtual camera motion strategies », *in : Graphics Interface Conference, GI'2000*, S. Fels, P. Poulin (éditeurs), Morgan Kaufmann, p. 69–76, Montréal, mai 2000.

[71] Y. Mezouar, F. Chaumette, « Génération de trajectoire dans l'image : vers l'asservissement visuel robuste », *in : 12ème Journées des Jeunes Chercheurs en Robotique*, p. 131–137, Bourges, février 2000.

[72] Y. MEZOUAR, F. CHAUMETTE, « Path planning in image space for robust visual servoing », *in : IEEE Int. Conf. on Robotics and Automation*, *3*, p. 2759–2764, San Francisco, avril 2000.

[73] L. OISEL, E. MÉMIN, L. MORIN, « Geometric-driven optical flow estimation and segmentation for 3D reconstruction », *in : European Conference on Computer Vision, ECCV'2000, LNCS 1843*, *2*, p. 849–863, Dublin, Irlande, juin 2000.

[74] C. PAPIN, P. BOUTHEMY, E. MÉMIN, G. ROCHARD, « Tracking and characterization of highly deformable cloud structures », *in : European Conference on Computer Vision, ECCV'2000, LNCS 1843*, *2*, p. 428–442, Dublin, Irlande, juin 2000.

[75] C. PAPIN, P. BOUTHEMY, E. MÉMIN, « Suivi de l'activité convective de cellules nuageuses dans des séquences d'images satellitaires », *in : 12ème Congrès Reconnaissance des Formes et Intelligence Artificielle, RFIA'2000*, *1*, p. 369–378, Paris, février 2000.

[76] A. SZANTAI, F. DESALMAND, M. DESBOIS, P. LECOMTE, S. MÉMIN, E. ZIMERAS, « Tracking low-level clouds over Central Africa on Meteosat images », *in : 2000 Eumetsat Meteorological Satellite Data Users' Conference*, p. 813–820, Bologne, Italie, mai 2000.

[77] A. SZANTAI, F. DESALMAND, M. DESBOIS, P. LECOMTE, P. PÉREZ, S. ZIMERAS, P. BOUTHEMY, « Tracking low-level clouds over land on Meteosat images », *in : 5th International Winds Workshop, NOAA, JMA, WMO, and Eumetsat*, p. 1–8, Lorne, Australie, février 2000.

[78] E. VENEAU, R. RONFARD, P. BOUTHEMY, « From video shot clustering to sequence segmentation », *in : IAPR Int. Conf. on Pattern Recognition, ICPR'2000*, *4*, p. 254–257, Barcelone, Espagne, septembre 2000.

[79] S. ZIMERAS, T. CORPETTI, M. MÉMIN, P. PÉREZ, « Motion estimation using satellite data », *in : Workshop on The Statistics of Directions, Shapes and Images*, 19th Leeds Annual Statistical Research (LASR), Leeds, U.K., 2000.

Rapports de recherche et publications internes

[80] L. AMSALEG, P. GROS, R. MEZHOUD, « Mise en base d'images indexées par des descripteurs locaux: problèmes et perspectives », *rapport de recherche n° 1316*, Irisa, mars 2000, `http://www.irisa.fr/bibli/publi/pi/2000/1316/1316.html`.

[81] L. AMSALEG, P. GROS, « A robust technique to recognize objects in images, and the DB problems it raises », *rapport de recherche n° 1371*, Irisa, novembre 2000, `http://www.irisa.fr/bibli/publi/pi/2000/1371/1371.html`.

[82] C. BAILLARD, C. BARILLOT, P. BOUTHEMY., « Robust adaptive segmentation of 3D medical images with level sets », *rapport de recherche n° 1369*, Irisa, décembre 2000, `ftp://ftp.irisa.fr/techreports/2000/PI-1369.ps.gz`.

[83] C. BAILLARD, P. HELLIER, C. BARILLOT, « Segmentation of 3D brain structures using level sets », *rapport de recherche n° 1291*, Irisa, janvier 2000, `http://www.irisa.fr/bibli/publi/pi/2000/1291/1291.html`.

[84] C. BARILLOT, « Modèle probabiliste de sillons corticaux : Comparaison avec des méthodes de fusion globales en IRM et MEG », *Rapport de convention*, Convention GIS, Région Bretagne 1 99 C 184, mai 2000.

[85] A. BENZOUGAR, P. BOUTHEMY, F. SPINDLER, « Spécification de la tâche de structuration de la vidéo », *Rapport de convention*, Rapport intermédiaire, Convention RNRT Agir 2 99 C 006, 2000.

[86] T. CORPETTI, E. MÉMIN, P. PÉREZ, « Dense fluid flow estimation », *rapport de recherche n° 1352*, Irisa, septembre 2000, `ftp://ftp.irisa.fr/techreports/2000/PI-1352.ps.gz`.

[87] F. DAMBREVILLE, J.-P. LE CADRE, « Continuous management of modes and resources in detection of a moving target », *rapport de recherche n° 1334*, Irisa, juin 2000.

[88] F. Dealmand, M. Desbois, P. Leconte, A. Szantai, E. Mémin, P. Pérez, S. Zimeras, « Calculation of low level winds over land », *Rapport de convention*, Rapport intermédiaire, Convention Inria Rennes-LMD-Eumetsat 1 99 C 351, février 2000.

[89] F. Dekeyser, P. Pérez, P. Bouthemy, « Restauration de séquences d'images numériques : rapport bibliographique », *Rapport de convention*, Rapport intermédiaire, convention TIRESIAS Inria Rennes-TCO-DGA 1 98 C 500, 2000.

[90] F. Dekeyser, P. Pérez, P. Bouthemy, « Restauration de séquences d'images numériques : étude d'algorithmes de traitement spatio-temporel », *Rapport de convention*, Rapport intermédiaire, convention TIRESIAS Inria Rennes-TCO-DGA 1 98 C 500, 2000.

[91] R. Fablet, P. Bouthemy, P. Pérez, « Non parametric statistical analysis of scene activity for motion-based video indexing and retrieval », *rapport de recherche n° 1351*, Irisa, septembre 2000.

[92] P. Gros, « Introduction à l'algèbre tensorielle », *rapport de recherche n° 1298*, Irisa, février 2000, `ftp://ftp.irisa.fr/techreports/2000/PI-1298.ps.gz`.

[93] P. Hellier, C. Barillot, « Coupling dense and landmark-based approaches for non-rigid registration », *rapport de recherche n° RR-4076*, Inria, décembre 2000, `http://www.inria.fr/RRRT/RR-4076.html`.

[94] C. Hue, J.-P. Le Cadre, P. Pérez, « Tracking multiple objects with particle filtering », *rapport de recherche n° 1361*, Irisa, octobre 2000.

[95] E. Marchand, P. Bouthemy, F. Chaumette, « A 2D-3D model-based approach to real-time visual tracking. », *rapport de recherche n° 1309*, Irisa, mars 2000, `ftp://ftp.irisa.fr/techreports/2000/PI-1309.ps.gz`.

[96] E. Marchand, F. Chaumette, F. Spindler, « Étude sur la commande des mouvements d'un manipulateur de ROV à l'aide d'une caméra pan-tilt », *Rapport de convention*, Contrat Ifremer-Inria Rennes 100 C 001, juin 2000.

[97] E. Marchand, N. Courty, « Visual servoing in computer animation: What you want to see is really what you get ! », *rapport de recherche n° 1310*, Irisa, mars 2000, `http://www.irisa.fr/bibli/publi/pi/2000/1310/1310.html`.

Divers

[98] F. Spindler, F. Chaumette, T. Fauconnier, C. Blonz, « Control of a camera by the glance », vidéo Inria, IEEE Int. Conf. on Robotics and Automation, Video Session, San Francisco, USA, avril 2000.