INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team VISTA

Vision Spatio-Temporelle et Active

Rennes

THEME 3B

Activity Report

2003

Table of contents

1. Team

Head of project-team
Patrick Bouthemy [DR Inria]

Administrative assistant
Stéphanie Lemaile [TR Inria, shared with Temics team]

Inria research scientists
Frédéric Cao [CR]
François Chaumette [DR]
Pierre Hellier [CR]
Éric Marchand [CR]
Sylvain Prima [CR, from 01/10/2003]
Charles Kervrann [Inra research scientist, on Inria secondment (mise à disposition), from 01/07/2003]

CNRS research scientists
Christian Barillot [DR]
Jean-Pierre Le Cadre [DR]

Faculty members
Étienne Mémin [Associate professor, Ifsic, University of Rennes 1, on CNRS secondment (délégation) until 31/08/2003]
Jian-Feng Yao [Full professor, Irmar, UFR Maths, University of Rennes 1, on Inria secondment (délégation), from 01/09/2003]

Inria technical staff
Fabien Spindler [IR]

Inria project technical staff
François Coldefy [Domus Videum project, from 01/04/2003]
Brigitte Fauvet [Feria project, from 15/09/2003]
Chafik Kermad [Mediaworks project, until 28/02/2003]
Cédric Riou [Sora project, from 01/04/2003]

Inria junior technical staff
Nicolas Ducoin [until 30/09/2003]
Anne-Sophie Tranchant [from 01/09/2003]

Inria Post-doctoral fellows
Peihua Li [from 01/09/2003]
Nathalie Peyrard [Domus Videum project, until 28/02/2003]
Vincent Samson [Lava project]

Teaching assistants
Thomas Corpetti [Ifsic, University of Rennes 1, until 30/09/2003]
Carine Hue [Ifsic, University of Rennes 1, until 31/08/2003]

Ph-D students
Laure Ait-Ali [Inria - Brittany Region grant, from 01/10/2003]
Elise Arnaud [Research Ministry grant]
Vincent Auvray [Cifre grant, GEMS, from 01/11/2003]
Jérôme Boulanger [Inra grant, from 01/12/2003]
Thomas Bréhard [DGA grant]
Cybèle Ciofolo [CNRS - Brittany Region grant]
Anne Cuzol [Research Ministry grant, from 01/10/2003]
Andrew Comport [Inria grant, Sora project]

Isabelle Corouge [Inria grant, until 15/03/2003]
Nicolas Mansard [ENS Cachan-Bretagne member, from 01/09/2003]
Arnaud Ogier [Inria grant, Brittany Region contract]
Gwénaëlle Piriou [Inria grant, Brittany Region contract]
Muriel Pressigout [Research Ministry grant, from 01/10/2003]
Anthony Remazeilles [Research Ministry grant]
François Rousseau [Research Ministry grant]
Omar Tahri [Inria - Brittany Region grant]
Thomas Veit [Inria - Brittany Region grant]

2. Overall Objectives

Key words: *image sequence*, *dynamic scene analysis*, *active vision*, *statistical modeling*, *Markov models*, *Bayesian estimation*, *robust estimation*, *particle filtering*, *statistical learning*, *motion detection*, *motion segmentation*, *optic flow*, *motion recognition*, *fluid motion analysis*, *registration*, *tracking*, *trajectography*, *visual servoing*, *medical image analysis*, *neuro-imaging*, *3D ultrasound images*, *video processing*, *video indexing*, *robot vision*, *augmented reality*, *meteorological imagery*, *experimental fluid mechanics*.

Research activities of the Vista team are concerned with image sequence analysis and active vision. More precisely, we address two broad objectives : - analysis of dynamic scenes and physical phenomena for providing image-based or scene-based motion-related measurements, and for investigating recognition and interpretation of temporal events; - perception and control of automated or robot systems, for handling tracking, grasping, navigation, surveillance or exploration tasks. In that context, we are interested in several types of spatio-temporal images, from optical sensors (video, infra-red image sequences) as well as from acoustics sensors (sonar, 3D echography). We mainly investigate statistical approaches (involving Markovian models and Bayesian methods, robust estimation,...) to properly handle motion and deformation analysis. The aspects of long-term spatio-temporal analysis (e.g., tracking) are particularly studied. We consider situations of active observer, i.e., we study active perception, when the sensor motion or attitude can be controlled. This involves the design and the experimental development of visual servoing methods and of higher level exploration strategies.

The contributions of the Vista team involve optic flow computation, motion detection and segmentation, fluid motion analysis, multiple object tracking, modelling and learning of dynamic video content, motion recognition and classification, theoretical design and experimental validation of a complete visual servoing framework relying on geometrical and kinematical image features, 3D reconstruction by active vision, camera pose computation, object tracking.

Four large domains of applications motivate and validate our work : motion and deformation "metrology" (meteorological satellite imagery, experimental visualization in fluid mechanics), content-based video indexing, medical imagery, robot vision and surveillance systems (underwater robotics, sonar, transportation).

3. Scientific Foundations

3.1. Hierarchical Markov models and Bayesian estimation

Key words: *Markov random fields*, *partial observation*, *Bayesian estimation*, *parameter estimation*, *hierarchical models*, *multigrid algorithm*, *tree*, *non-iterative inference*.

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 [75]. Dans ce cadre, nous nous attachons à définir des modèles ou des algorithmes de type hiérarchique qui permettent de mener de façon efficace les tâches d'estimation (des cartes d'attributs cachés ou des paramètres du modèle), pour des problèmes à espace d'états aussi bien continu que discret.

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

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

où $\mathcal{C}$ est un ensemble de "petites" parties de S ($|c| = 1$ ou 2 en général). Une structure de graphe est associée à cette factorisation (celle pour laquelle tous les éléments de $\mathcal{C}$ sont des cliques), fournissant un outil de codage des indépendances conditionnelles entre variables [74] : si une partie C de S sépare deux autres parties A et B pour ce graphe, alors z_A et z_B sont indépendants sachant z_C ($P(z_A, z_B|z_C, \theta) = P(z_A|z_C, \theta)P(z_B|z_C, \theta)$). La chaîne, la grille bidimensionnelle et le quad-arbre sont les graphes les plus fréquemment rencontrés en analyse d'image ; ils sont illustrés sur la figure jointe.

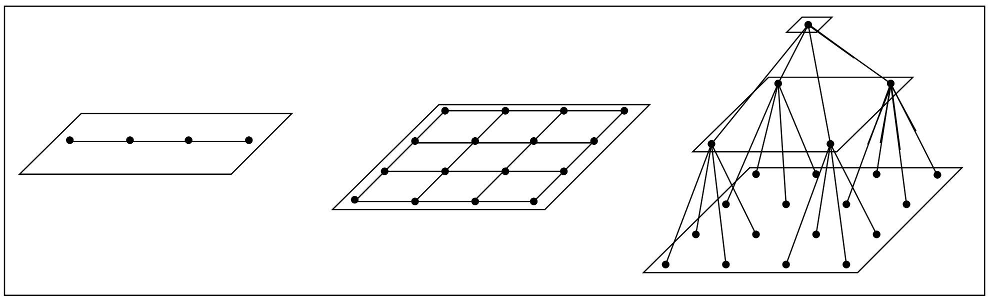

Figure 1.

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 ($\forall s, \quad \arg\max_{x_s} P(x_s, y|\theta)$), 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) [70].
- **estimation des paramètres** : l'estimation du meilleur jeu de paramètres est plus complexe encore, du fait de l'observation partielle du modèle. Les méthodes standard de maximum de vraisemblance ($\hat{\theta} = \arg\max P(x, y|\theta)$) sont prises en défaut par la méconnaissance de x. En conséquence, les outils mis en œuvre sont essentiellement itératifs de type EM ("expectation-maximization") avec pour but la maximisation de $P(y|\theta)$. Ils produisent généralement une estimation de nouveaux paramètres $\theta^{(n+1)}$ sur la base d'échantillons tirés selon la loi *a posteriori* $P(x|y, \theta^{(n)})$. La taille et la structure des problèmes inverses spatiaux qui nous intéressent rendent cette tâche particulièrement délicate (tirage des échantillons, estimation des nouveaux paramètres par minimisation, etc.)

Dans le cas de problèmes de très grande dimension (telle que l'extraction de cartes denses de primitives à partir de plusieurs images), les différents algorithmes génériques précédemment évoqués, même pour des paramètres supposés connus, sont coûteux : en raison de la localité des couplages entre variables, l'information se propage de façon lente et peu efficace.
Dans le but de traiter ces problèmes, une démarche générique souvent fructueuse consiste à *hiérarchiser* les primitives et/ou les données. Cette idée peut être déclinée de diverses manières.
Dans un premier groupe d'approches, des *versions "réduites"* (en terme du nombre des variables inconnues) $P^i(x^i, y|\theta)$ ou $P^i(x^i, y^i|\theta)$, $i = L...0$, du modèle spatial initial sont d'abord construites, puis exploitées au sein d'une *algorithmique descendante*. Les attributs $\hat{x}^i$ inférés à un niveau i servent de point de départ à la procédure d'inférence au niveau $i-1$ de "taille" immédiatement supérieure. Il s'agit ici principalement des méthodes multirésolutions ou multigrilles (selon que des versions réduites de y sont également calculées ou non). Ces méthodes permettent une accélération substantielle des schémas itératifs déterministes standard, et fournissent souvent une amélioration de la qualité des résultats [4].
Une deuxième classe d'approches a pour but la définition de modèles hiérarchiques *globaux* $P(x^L, ..., x^0, y|\theta)$. 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. Visual servoing

Key words: *computer vision*, *robotics*, *visual servoing*.

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 [71] utilisent généralement des informations visuelles 2D extraites de l'image. Les lois de commande consistent alors à contrôler le mouvement de la caméra afin que les mesures dans l'image $s(t)$ atteignent une valeur désirée s^* ou suivent une trajectoire spécifiée $s^*(t)$.
Afin d'élaborer une loi de commande en boucle fermée sur des mesures $s(t)$, il est nécessaire d'estimer ou d'approximer la relation qui lie la variation de s aux variables de contrôle. Dans le cas d'informations visuelles géométriques (telles par exemple les coordonnées 2D d'un point) contrôlées à l'aide des six degrés de liberté d'une caméra, cette relation est définie par [2] :

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

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

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

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

- $\hat{L} = L_{(s(t),\hat{z}(t))}$. On calcule à chaque itération la valeur courante de la matrice d'interaction. Une estimation des paramètres z doit alors être réalisée en ligne, par exemple à l'aide de la connaissance d'un modèle 3D de l'objet [68].
- $\hat{L} = L_{(s^*,z^*)}$. 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 [67] : 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 [73] permet de résoudre de manière efficace et élégante les problèmes de redondance rencontrés lorsqu'une tâche visuelle ne contraint pas l'ensemble des degrés de liberté de la caméra. Il est alors possible d'exécuter, parallèlement à la tâche visuelle, des tâches secondaires telles des opérations de suivi de trajectoires pour des applications d'inspection, ou d'évitement des butées et singularités du robot.
Lorsque la fonction de tâche est correctement modélisée, il est ensuite assez aisé d'élaborer une loi de commande générique permettant une décroissance exponentielle de la fonction de tâche. On obtient :

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

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. Trajectography

Key words: *trajectography*, *tracking*, *path extraction*, *multiple hypotheses*, *combinatorics*.

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, [72]:

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

où Φ désigne la matrice de transition du système, le vecteur $\mathbf{U}$ représente l'accélération (instantanée) de l'observateur et où $\mathbf{W}(k)$ représente un bruit markovien. Par ailleurs, l'équation d'observation prend la forme suivante: $\hat{\theta}_k^i = \theta_k^i + w_k$, où $\theta_k^i = \arctan(r_x^i(k)/r_y^i(k))$ et i est un indice aléatoire correspondant au fait que l'on ne sait pas de quelle source provient la détection (ou s'il s'agit d'une fausse alarme).

Les caractéristiques de ce type de problème sont multiples : *i)* non-linéarité du système, *ii)* possibilité de manœuvre (action sur $\mathbf{U}$); *iii)* caractère aléatoire de i. L'étape d'extraction consiste à attribuer des détections aux sources (estimation de i), et lorsque i est connu, ou estimé, le problème se découple alors sur les différentes sources. L'objectif de l'*extraction* consiste à élaborer des pistes à partir des données provenant des étapes de traitement direct et massif des sorties de capteurs. Les données sont constituées de plots (de nature binaire dans le cas le plus simple). Le problème de l'*extraction* est le suivant: quels sont les sous-ensembles de plots 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 :

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

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

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

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

Une approche séduisante consiste alors à considérer que les probabilités d'assignation des mesures aux pistes sont indépendantes d'un scan à l'autre. Ainsi, la méthode PMHT (Probabilistic Multiple 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), [69].

4. Application Domains

Key words: *video processing*, *video indexing*, *multimedia*, *meteorological imagery*, *experimental fluid mechanics*, *environment*, *medical image analysis*, *neuro-imaging*, *3D ultrasound images*, *image-guided neurosurgery*, *biological imagery*, *health*, *robot vision*, *augmented reality*, *underwater robotics*, *vehicle navigation*, *transportation*, *defense*, *sonar*.

We are dealing with the following application domains (mainly in collaboration with the listed partners) :

- Video processing and indexing (INA, Thomson, Xerox);
- Medical imagery (CHU Rennes, GEMS) and biological imagery (INRA);
- Experimental fluid mechanics (Cemagref) and meterological imagery (Météo-France, LMD);
- Robotics (agronomy, underwater robotics) and augmented reality (Ifremer, Cemagref, Total Immersion);
- Surveillance (Onera, Thales) and vehicle navigation.

5. Software

5.1. Motion2d software - parametric motion model estimation

Participants: Fabien Spindler, Patrick Bouthemy.

Motion2D is a multi-platform object-oriented library to estimate 2D parametric motion models in an image sequence. It can handle several types of motion models, namely, constant (translation), affine, and quadratic models. Moreover, it includes the possibility of accounting for a global variation of illumination. The use of such motion models has been proven adequate and efficient for solving problems such as optic flow computation, motion segmentation, detection of independent moving objects, object tracking, or camera motion estimation, and in numerous application domains, such as dynamic scene analysis, video surveillance, visual servoing for robots, video coding, or video indexing. Motion2D is an extended and optimized implementation of the robust, multi-resolution and incremental estimation method (exploiting only the spatio-temporal derivatives of the image intensity function) we defined several years ago [10]. Real-time processing is achievable for motion models involving up to 6 parameters (for 256x256 images). Motion2D can be applied to the entire image or to any pre-defined window or region in the image. Motion2D is released in two versions :

- Motion2D Free Edition is the version of Motion2D available for development of Free and Open Source software only (no commercial use). It is provided free of charge under the terms of the Q Public License. It includes the source code and makefiles for Linux, Solaris, SunOS, and Irix. The latest version is available for download.
- Motion2D Professional Edition provided for commercial software development. This version also supports Windows 95/98 and NT.

More information on Motion2D can be found at http://www.irisa.fr/vista/Motion2D and the software can be donwloaded at the same Web address.

5.2. d-Change software - motion detection

Participants: Fabien Spindler, Patrick Bouthemy.

D-change is a multi-platform object-oriented software to detect mobile objects in an image sequence acquired by a static camera. It includes two versions : the first one relies on Markov models and supplies a pixel-based binary labeling, the other one introduces rectangular models enclosing the mobile regions to be detected. It simultaneously exploits temporal differences between two successive images of the sequence and differences between the current image and a reference image of the scene without any mobile objects (this reference image is updated on line). The algorithm provides the masks of the mobile objects (mobile object areas or enclosing rectangles according to the considered version) as well as region labels enabling to follow each region over the sequence.

5.3. Dense-Motion software - optical flow computation

Participant: Etienne Mémin.

The Dense-Motion software written in C enables to compute a dense velocity field between two consecutive frames of a sequence. It is based on an incremental robust method encapsulated within an energy modeling framework. The associated minimization is based on a multi-resolution and multigrid scheme. The energy is composed of a data term and a regularization term. The user can choose among two different data models : a robust optical flow constraint or a data model based on an integration of the continuity equation. Two models of regularization can be selected as well : a robust first-order regularization or a second-order Div-Curl regularization. The association of the latter with the data model based on the continuity equation constitutes a dense motion estimator dedicated to image sequences involving fluid flows. It was proven to supply very accurate motion fields on various kinds of sequences in the meteorological domain or in the field of experimental fluid mechanics.

5.4. vistal software : processing and analysis of 3d spatio-temporal images

Participants: Pierre Hellier, Christian Barillot.

VISTAL is a software platform for 3D and 3D+t image analysis in the field of medical imaging allowing the development of generic algorithms used in various contexts (linear and deformable registration, segmentation, statistical modelling or geometric calibration of medical images and systems). This software platform consists of generic classes (Image3D, Image4D and Lattice C++ Templates) and of a set of 3D and 3D+t image processing libraries. VISTAL is an environment independent of the software platform (it runs on Windows, SGI, Linux and Solaris).

5.5. romeo software : robust multigrid elastic registration based on optical flow

Participants: Pierre Hellier, Christian Barillot.

ROMEO (RObust Multigrid Elastic registration based on Optical flow) is a non-rigid registration algorithm based on optical-flow. It is developed using VISTAL (C++ template classes described above). It estimates a regularized deformation field between two volumes in a robust way : two robust estimators are used for both the data term (intensity constancy) and the regularization term (smoothness of the field). An efficient multiresolution and multigrid minimization scheme is implemented so as to estimate large deformations, to increase the accuracy and to speed up the algorithm. ROMEO has been registered at APP.

5.6. juliet software : joint use of landmarks and intensity for elastic registration

Participants: Pierre Hellier, Christian Barillot.

JULIET (Joint Use of Landmarks and Intensity for Elastic regisTration) is a non-rigid registration algorithm that builds on the ROMEO software. It makes it possible to incorporate sparse constraints deduced from the matching of anatomical structures such as cortical sulci for instance. A sparse deformation field is introduced as a soft constraint in the minimization to drive the registration process. A robust estimator is used so as to limit segmentation errors and false matching. JULIET has been registered at APP.

5.7. visp : a visual servoing platform

Participant: Éric Marchand.

Visual servoing is a very active research area in vision-based robotics. A software environment that allows fast prototyping of visual servoing tasks is then of prime interest. The main reason is certainly that it usually requires specific hardware (the robot and, most of the time, dedicated image processing boards). The consequence is that the resulting applications are often not portable and cannot be easily adapted to other environments. Today's software design allows one to propose elementary components that can be combined to build portable high-level applications. Furthermore, the increasing speed of micro-processors allows the development of real-time image processing algorithms on a usual workstation. We have developed a library of canonical vision-based tasks for eye-in-hand visual servoing that contains the most classical linkages that are used in practice. The VISP software environment features all the following capabilities : independence with respect to the hardware, simplicity, extendibility, portability. Moreover, VISP involves a large library of elementary positioning tasks w.r.t. various basic control features (points, lines, circles, spheres, cylinders,...) that can be combined together, and an image processing library that allows the tracking of visual cues (dot, segment, ellipse, spline,...). This modular platform has been primarily developed in C++ on Unix workstations and is now available for Linux and Windows XP.

5.8. Marker software : Marker-based augmented reality kernel

Participant: Éric Marchand.

The Marker software implements an algorithm supplying the computation of camera attitude, the calibration of the cameras using fiducial markers. Pose computation is handled using the virtual visual servoing approach. The principle consists in considering the pose and the calibration as a dual problem of visual control. This method presents many avantages : similar accuracy as for the usual non-linear minimization methods, simplicity, effectiveness. A licence of this software was yielded to Total Immersion company.

5.9. MarkerLess : MarkerLess-based augmented reality kernel

Markerless is an upgrade of the Marker software with additional features. It allows the computation of camera pose with no fiducial marker. It also relies on the virtual visual servoing approach. see paragraph 6.2.5 for more details.

5.10. Development work : Robot vision platforms

Participant: Fabien Spindler.

We exploit several experimental platforms to validate our research work in robot vision, in visual servoing and in active vision. More precisely, we have two robotic vision cells built by Afma Robots : the first one is a cartesian robot with six degrees of freedom, the other one is a cylindrical robot with four degrees of freedom. Each robot is equipped with a CCD camera mounted on the effector and an Imaging Technology framegraber. A PC under Linux communicates with the robot using a SBS Technologies bus adapter. These equiments require specific hardware, but also software maintenance actions and new developments in order to make them evolve. Training and assistance of the users, presentation of demonstrations also form part of the daily activities. Fall 2003, we have acquired a new robotic system with an ultrasound probe dedicated to medical applications, specifically to 3D remote echography.

5.11. Development work : Cycab vehicle

Participants: Fabien Spindler, Nicolas Ducoin.

The Cycab vehicle is a small autonomous non-holonomic mobile robot equipped with a pan-tilt camera and exploited as a research platform. We have continued our work of modelling the platform in order to settle vision-based applications with a view to driver assistance or automatic parking. We have designed a first demonstration which consists in positioning the mobile robot by visual servoing relatively to a target without accounting for non-holomomic constraints. We control the Cycab translational and rotational speed as well as the embedded camera pan and tilt rotational speed. To this end, visual features which are considered to control the vehicle are the distances between the optical center and two vertical segments on the target, deduced from the length and position of the segments in the image. As these distances are independent of a rotation around the optical center, the error between the desired and the current distances is decoupled from the movement of the pan-tilt camera. The camera can then be used to maintain the target within the field of view without affecting the Cycab main task. A second demonstration consisting in detecting and tracking a moving obstacle from the pan-tilt camera mounted on the vehicle was carried out. The aim is to detect and track mobile objects located at less than 20 meters from the car, without any *a priori* knowledge on the image content or obstacle shape. In our case, we consider a pedestrian crossing the Cycab trajectory, but the same approach could be applied to a vehicle, a motor-bike, a bicycle, for instance. After the detection stage, the goal is to achieve an image-based tracking of the pedestrian by controlling the camera pan and tilt so that the detected mobile object remains at the center of the image. Problems such as occlusions or multiple moving objects are not in the scope of this study. Finally, we have mounted an outdoor pan-tilt camera on top of a building in order to have a global view of the area where the Cycab experiments are performed. We have installed the experiment described above in this camera system too, and have designed a demonstration consisting in detecting and tracking the Cycab so that the vehicle remains in the middle of the image when driving in this area.

6. New Results

6.1. Statistical modeling, motion and deformation analysis, video processing

6.1.1. Fluid motion analysis

Participants: Étienne Mémin, Thomas Corpetti.

Fluid flow phenomena abound both in nature and technology. In environmental sciences such as meteorology, climatology, or oceanography, complex flows under concern are caused by the transport of energy and matter within the atmosphere, in oceans, and ashore. In life sciences, complex flows at various scales are also observed in organisms, ranging from the macroscopic level (e.g., deformation of brain tissues) to dynamic structures at the cellular level (e.g., flow through membranes). In industrial production, the engineering of flows is of utmost importance for various processes related to aero- and fluid dynamics, combustion or chemistry. A common feature of these flow phenomena is the lack of a complete physical understanding of the underlying dynamics which prevents their accurate prediction and control. We have proposed a new motion estimation method dedicated to image sequences depicting fluid phenomena. It involves a data-driven term, deduced from the integration of the continuity equation between two successive images, and a regularization term based on the Helmholtz decomposition of vector fields. Such a regularization favors the emergence of homogeneous divergence and vorticity blobs. Our method outperforms on one hand robust dense motion estimation methods defined for computer vision applications and on the other hand PIV (Particle Image Velocimetry) techniques currently used in the field of experimental flow visualisation. The evaluation of our method in the context of experimental fluid mechanics has been conducted whithin a cooperation with the AEROBIO group of the Cemagref Rennes. It was tested on a sequence corresponding to a free turbulent shear layer flow, whose main characteristics are known. This kind of flow appears when the velocities of two fluids that slide along each other have the same direction but different magnitude. Many studies have been carried out on such a kind of flows and well-known representative characteristics can be extracted from the computed motion fields. The two most important ones are the spatial distribution of the longitudinal velocity and the vorticity map. We have also considered other specific statistical parameters to compare our method with an up-to-date PIV technique. These parameters are the blooming parameter which is the derivative of the vorticity magnitude and reveals the evolution of turbulent structures, and the spatial derivative of the absolute motion magnitude which accounts for the speed of the flow.

Besides, we have extended the described fluid motion estimation method relying on the Helmholtz decomposition to compute the associated potential functions (namely, the stream function and the velocity potential) directly from the image sequence. These functions are of primary interest for fluid flow pattern analysis in numerous application domains. The designed solution has been formulated within the variational framework. The corresponding energy function includes regularizers with derivatives up to the third order to obtain unbiased high-quality solutions. The approach has been made computationally tractable by means of auxiliary variables. This work has been carried out within a collaboration with the CVGPR group of Mannheim University (Prof. C. Schnörr).

6.1.2. Motion detection

Participants: Frédéric Cao, Patrick Bouthemy, Thomas Veit.

Motion detection is the first elementary step in image sequence analysis. Therefore, a good algorithm should satisfy several requirements : it should be simple and fast, it should work in a large panel of operational conditions, it should not have extrinsic parameters to be tuned by an expert user, and the number of false alarms should be very limited. We are considering a static camera or a mobile one. For the second case, the global dominant motion, assumed due to the camera motion, is first computed using the Motion-2D software, paragraph 5.1, and then cancelled. The first step is to compute an elementary measure of local motion. A good candidate is the residual normal velocity V_{res}, which has the nice property not to depend too much on contrast. Now, this is unsufficient, since this data is noisy and subject to the well-known aperture problem. The idea is to group high values of this measure of motion into geometrical structures that are quite conform to perceptual objects. Level lines (for a grey level image, level lines are the topological boundaries of the connected components of the sets of the type $\{x,\, u(x) \geq \lambda\}$) are good candidates : indeed, they are closed curves that form a tree-structure. To each level line L, we associate a number of false alarms, defined by

$$NFA(L) = N_{ll} N_v \min_{1 \leq i \leq N_v} B(n, k(i), p_i),$$

where

- n is the number of pixels contained in the interior of L (it makes sense, since level lines are closed).
- $k(i)$ is the number of pixels such that $V_{res} \geq v_i$, where the v_i are a bench of tested values.
- p_i is the empirical probability that $V_{res} \geq v_i$.
- $B(n, k, p) = \sum_{j=k}^{n} \binom{n}{k} p^j (1-p)^j$ is the tail of the binomial distribution.
- N_{ll} is the number of level lines taken into consideration.
- N_v is the number of values v_i to which we compare the observation, typically about 100. This number is sufficient to eliminate quantization effects and choosing a larger value does not make any change in the detection.

A very small number of false alarms gives a quantitative measure of the fact that residual normal velocity can be explained only by chance. We call ε-meaningful motion region a region such that this number of false alarms is smaller than ε. It is quite easy to prove that, if we assume that the values of the residual velocities are i.i.d. in the image, with a distribution equal to the empirical one, then the expected number of ε meaningful region is smaller than ε. It can be proved that ε may be chosen equal to 1 in any type of sequences. The method thus provides an efficient instantaneous motion detector, and all parameters are either automatically fixed, or have nearly no influence.

6.1.3. Probabilistic motion modeling and dynamic content learning

Participants: Patrick Bouthemy, Jian-Feng Yao, Gwénaëlle Piriou, Nathalie Peyrard.

Exploiting the tremendous amount of multimedia data, and specifically video data, requires to develop methods able to extract information at a somewhat semantic level. We tackle the challenging problem of inferring "concepts" from low-level video features. We have designed a statistical approach, involving modeling and (supervised) learning stages, to deal with concepts related to events in videos, and more precisely, to dynamic video content. Therefore, we have focused on motion information. Because of the diversity of dynamic video content (even for a given class of event), no analytic models are available; we have to design such motion models and learn them from videos. To this end, we have introduced new probabilistic motion models which allow us to derive a parsimonious motion representation while coping with errors in the motion measurements and with variability in a given kind of motion content. We handle in a distinct way the camera motion (i.e., the dominant image motion) and the scene motion (i.e., the residual image motion), since these two sources of motion bear important and complementary information on the dynamic video content. As for motion measurements, we consider parametric motion models to capture the camera motion, and low-level local motion features (i.e., normal residual flow magnitude) to account for the scene motion. We suppose that the video has been divided into homogeneous temporal segments. For each video segment, we investigate a "frequency characterization" of their dynamic content, namely, 2D histograms of the velocity vectors inferred from the estimated dominant motion models and 1D histograms of the normal residual flow magnitudes, both computed over all the pixels of the video segment. The latter (along with the histograms of the temporal contrasts of the residual normal flow magnitudes) are modeled by a specific mixture model involving a Gaussian distribution and a Dirac distribution. The former are modeled by a Gaussian mixture. These models are learnt off-line from a training set of video segments using maximum likelihood (ML) criteria along with the ICL criterion to determine the number of components of the Gaussian mixture. To validate this scheme, we have considered sports videos which involve complex contents while being reasonably specifiable. Moreover, events or highlights can be naturally related to motion information in that context. Satisfactory results on athletics and tennis TV programs have been obtained.

6.1.4. Motion classification, event detection and video summarization

Participants: Patrick Bouthemy, Vincent Samson, François Coldefy, Gwénaëlle Piriou.

We are concerned with (supervised) motion classification in video sequences to solve for event detection in videos. We have designed a two-step detection scheme. The first one is a sorting step into two groups ("interesting" versus "not interesting", typically "play" versus "no play" when dealing with sports videos). The second one is a classification step into N classes and amounts to recognize the specified events (in terms of dynamic content) among the pre-selected segments. The proposed framework is general enough. Nevertheless, to validate it, we have focused on sports videos, since they involve complex dynamic contents and events while naturally related to motion information. We have designed and compared three classification schemes. The first one is a Bayesian classification technique exploiting the probabilistic motion models described in paragraph 6.1.3 (in practice, we only consider maximum likelihood (ML) criteria). The two other ones, a SVM technique and an original robust classification method, consider the estimated parameters of these models only. In the SVM classification framework, the goal is to directly determine a decision function from the training data, which may be non-linear thanks to the use of a kernel function, a Gaussian kernel for example. The learning step can be formulated as a quadratic optimization problem that requires appropriate routines. One key issue is how to properly choose the underlying SVM parameters. We have adopted the "leave-one-out" cross-validation adjustment. An alternative approach to SVM classifiers is to model each class by a set of clusters defined by their centers (as in VQ techniques). Furthermore, contrary to classical K-means algorithms, we suggest here to use a non-Euclidean distance in order to be robust to the presence of outliers in the training data. We have chosen a hard-redescending robust function that depends on a single scale parameter. The algorithm alternates allocation of data to clusters and re-estimation of the centers. Actually, we start with only one cluster and we progressively create new clusters from the remaining outliers. The number of clusters can thus be automatically determined along with their centers, and depends on the scale parameter of the robust function. This robust clustering technique is attractive and efficient to deal with a N-class problem involving classes whose content can exhibit a large diversity (which is the case for dynamic video content). Moreover, it easily allows incremental learning (adding new classes, or taking into account new training samples) since the different classes are independently learned and structured.

We have also investigated the issue of video abstraction. We have defined a three-step procedure to create a video summary. First, we perform a temporal segmentation of the video, second, after an off-line learning stage (based on K-means) and exploiting visual and temporal descriptors, we extract the video segments of potential interest, third, we finally select on visual and audio criteria the best segments to be retained in the summary. The video segmentation relies both on the detection of changes in the camera motion and the detection of shot changes. The temporal segmentation is independent of the video genre and can be used for sports videos, documentaries or fiction films also. We use a statistical mean test to decide whether a segment is of potential interest or not. Finally, the construction of the video summary results from the on-line clustering of the preselected segments according to audio (e.g., ball hits sounds, applause, in tennis matches) and visual criteria (see also paragraph 7.5.1).

6.1.5. 4d video-microscopy in biological imagery

Participants: Charles Kervrann, Patrick Bouthemy, Jérôme Boulanger.

One of the most powerful approaches to study cell biological processes has evolved in the form of dynamic analysis of subcellular structures such as organelles and vesicles by localizing and tracking labeled proteins using light microscopy. The development of techniques in which target proteins are fused to Green Fluorescent Protein (GFP) makes it straightforward to label specific macro-molecular structures in vivo. Optical microscopes and digital imaging have advanced to the point where the recording of time-lapse series of 3D images (referred to as 4D imaging) with multiple wavelenghts (referred to as 5D imaging) from living cells, has become routine in many laboratories in biology. However the methods for analyzing these images and for the extraction of accurate quantitative information of the localization and dynamics of biological structures are much less developed. This is particularly true in the case of 3D small objects that are moving in different directions at different speeds and that can aggregate or disapear temporarily. Manual analysis is usually inappropriate in this situation. Besides, most of previous works address this problem only in two dimensions and neglect the case of partially overlapping objects. Our long-term goal is to characterize dynamics and interactions of cell components in 4-5D imaging in sufficient details. We aim at extracting major motion components by statistical learning and spatial statistics from the computed partial or complete trajectories and compare trajectories among normal-type and wild-type cells. The challenge is to track GFP tags with high precision in movies representing several gigabytes of image data and collected and processed automatically to generate information on complex trajectories. For instance, we plan to demonstrate the potential of statistical learning for a typical biological application: the tracking of chromosomes during mitosis. Mitotic chromosome segregation is the process by which replicated sister chromatids are divided equally into two physically separated sets around which daughter cells form. In that case, a large degree of automation is imperative and the object tracking software must perform reliably without the supervision of an operator. Ultimately, the learned dynamics will be combined with molecular genetics to determine the roles of various proteins during segregation processes.

6.1.6. Statistical shape models for medical images

Participants: Christian Barillot, Cybèle Ciofolo, Isabelle Corouge.

We have proposed a framework for building statistical shape models which has been firstly applied to cortical sulci. The model is built from a training population of sulci extracted from MRI volumes with a parametric representation. A coordinate system intrinsic to a sulcus shape is defined in order to align the training population, on which a principal components analysis is then performed. This statistical modeling has been extended to a sulci graph in order to describe not only the morphological features of one sulcus, but also the relationships in terms of relative position and orientation between major sulci. The analysis is concerned with a reduced graph defined by a pair of sulci. This framework has been applied to other applications. On one hand, we used it in an evaluation project of inter-subjects brain registration methods. When performed on local landmarks, the statistical analysis provides a similarity measure between registered shapes, and thus provides a comparison criterion between methods. On the other hand, we have exploited the statistical knowledge acquired by the sulci modeling in the context of anatomical and functional atlases building. More precisely, we have proposed a fusion scheme, local and non-linear, to register inter-subjects functional data (MEG dipoles) toward a single coordinate system linked to the anatomical model of cortical sulci. Experimented on a database of 18 subjects, this method has been shown to reduce the observed inter-individual functional variability. Finally, the methodology proposed to model cortical sulci shape has been applied to determine the functional border shape delimitating low-order visual areas.

6.1.7. 3d ultrasound image analysis

Participants: Christian Barillot, Pierre Hellier, Arnaud Ogier.

The non invasive nature, the moderate cost, and the real-time capability of ultrasound images make them an essential tool to assist medical diagnosis. We intend to better exploit the three-dimensional echographic images in order to improve the diagnosis and the therapeutic processes. Unfortunately, echographic images are corrupted by additive and multiplicative noise, encapsulated in a log-compression filtering. The nature of the resulting acquired signals makes difficult not only the interpretation of these images but also their automatic processing based on image segmentation procedures. We are studying various methods to reduce noise in these images, which is an ill-posed inverse problem. We have first worked on the analysis and the modeling of the noise included in ultrasound images. Then, we have developed adapted filters according to different statistical models of these images (Speckle, K-distribution, Rice distribution or K Homodyne distribution). We are currently evaluating these approaches with 2D ultrasound images, and we are planning to extend these adapted filters to 3D ultrasound images and 3D ultrasound image sequences.

6.1.8. Statistical analysis of MRI times series

Participants: Sylvain Prima, Christian Barillot, Pierre Hellier, Laure Ait-Ali.

This work is just starting and it consists in defining and computing measures of evolution for pathologies in times series of magnetic resonance images (MRI). These measures allow the quantitative follow-up of the disease for subjects with multiple sclerosis. It continues work done during S. Prima's postdoc in the McConnell Brain Imaging Centre of the Montreal Neurological Institute (McGill University), where it was proposed to use univariate and multi-variate change-point statistics for the detection of evolving voxels. It will be carried out in collaboration with Dr. L. Collins (McGill University) and Pr. G. Édan (CHU Rennes).

6.2. Tracking, registration, reconstruction

6.2.1. Particle filtering for tracking multiple objects in image sequences

Participants: Jean-Pierre Le Cadre, Carine Hue.

This study corresponds to C. Hue's Ph-D thesis and to related extensions. It deals with non linear filtering for multiple objects in a cluttered environment. For a single object and linear Gaussian models, the Kalman filter analytically solves the filtering equations. When this linear assumption no longer holds, extended Kalman filters have also been proposed with local linearization of the models. However, these filters diverge when the models are highly non linear, when the posterior density is multimodal or simply if initialization is not convenient. The sequential Monte Carlo methods give an approximation of the posterior law based on a finite sum of Dirac densities centered on points named particles and weighted proportionally to the measurement likelihood. They are based on the Monte Carlo principle. When dealing with multiple objects and clutter, the association between measurements and objects is unknown and must be solved to estimate the states. To that aim, an original algorithm (named MOPF) has been developed, based on sequential Monte Carlo methods and assuming independence of association variables. It has also been extended to a variable number of objects and to multiple receivers. These methods have been applied to the tracking of multiple objects in the image analysis context. Furthermore, the Posterior Cramér-Rao Bound has been extended to this multi-target framework. Various assignment models have been considered, ranging from the exact one to the MOPF and JPDAF ones. Simulation and experimental results tend to prove the interest of the MOPF framework, which is nevertheless not optimal. This work was conducted in cooperation with P. Pérez currently with Microsoft Research in Cambridge.

6.2.2. Point and structure tracker

Participants: Étienne Mémin, Elise Arnaud, Anne Cuzol.

Object tracking is a challenging task in computer vision with a large number of applications. Let us mention for instance the monitoring of passive entities in environmental aerial or satellite images, such as icebergs, polutants,..., or the tracking of 2D or 3D objects in video sequences. One has to face difficult and ambiguous situations generated by cluttered background, occlusion, large geometric deformations, illumination changes or noisy data. To design trackers robust to outliers and occlusions, a classical way consists in resorting to stochastic filtering techniques such as Kalman-Bucy filter or on sequential Monte Carlo approximation methods (also called particle filters). Point tracking in an image sequence forms a basic but essential problem in computer vision. This issue is inherently difficult since one may only rely on very local photometric or geometric time-invariant characteristics of the point. In addition, designing a proper evolution model for a feature point is not straightforward without any *a priori* knowledge on the dynamics of the enclosing object. These difficulties led us to consider systems composed by a state equation and a measure equation which both depend on the image sequence data. A linear filter and a non-linear filter have been designed. The linear filter is derived from a conditional linear minimum variance estimator. It is particularly well suited to image sequences exhibiting global dominant motions. The non-linear filter is implemented with a conditional particle filter. The latter allows us to track points whose motion can be only locally described. In this case, we consider an expression of the measurement likelihood which enables to analytically derive the optimal importance function used in the diffusion process of the particle filter. Such a choice makes the tracker robust with respect to occlusion. The proposed point trackers are very efficient and can track points in various complex situations involving abrupt changes of trajectories, clutter noise, occlusions and large geometric deformations. We will also investigate, on the same basis, the tracking of vector fields described as linear combination of basis shape functions. This last issue will be addressed in the context of image sequences depicting fluid flows. In contrast with the previous case where no prior dynamics were considered, we will rely here on the vorticity expression of Navier-Stokes equation to predict over time the vorticity part of the motion field.

6.2.3. Target motion analysis

Participants: Jean-Pierre Le Cadre, Thomas Bréhard.

Target motion analysis (TMA) has been investigated for a long time. The nature of the problem is also relatively simple since it is assumed that there is no uncertainty about the measurement-to-target assignment. However, some basic difficulties still remain. In a first time, our studies are oriented toward TMA in the Bearings-Only Tracking (BOT) context. Practically, this corresponds to the tracking of a maneuvering target from the outputs (bearings) of infra-red sensors. The outputs (angles) depend non-linearly on the target state, whose variations are modeled by a dynamic system controlled by the observer maneuver and affected by a process noise. In this setting, the system is non-linear and partially observed. Particle filtering then appears as the perfect workhorse and performs quite satisfactorily, at least when correctly initialized. Initialization is thus of crucial importance. Since there is generally no prior about the effective target location, a first step is to render uncoupled the unobservable and observable components. This is achieved by means of the modified polar coordinates. The reduced state vector is made of observable components ($\{\beta,\ \dot{\beta},\ \dot{r}/r\ \}$). While the unobservable component is discretized along a fixed grid, observable components are estimated through a (non-linear) regression. The particle filter is initialized accordingly. As long as no (strong) observer occurs, the range component remains unobservable; however, particle filter performs satisfactorily for tracking observable components. When a strong maneuver occurs, then the range particles converge toward the exact range values, without requiring any additional initialization. The whole process is reliable and performs quite satisfactorily. The initialization problem which was a major difficulty for any particle filter has been solved. This type of approach is also sufficiently general to be extended to various contexts (e.g. 3D-TMA, maneuvering targets, multiple receivers,...). When the process noise is non-zero, the classical (deterministic) Cramér-Rao Bound must be reformulated in a Bayesian framework. This is the aim of the Posterior Cramér-Rao Bound (PCRB). Deriving a reliable formulation of this bound is a prerequisite of multiple issues in sensor management. Recursive calculations of the PCRB have been developed recently. However, it is widely recognized that it is quite over-optimistic in the BOT context. This is due to an unbiaseness assumption which is no longer valid for BOT. The recursive formulation must be abandoned or widely modified, and the calculation of the correct bound is computationally demanding. The new bound we have determined is quite accurate even for not or weakly obervable scenarios.

6.2.4. Spatio-temporal shape matching

Participant: Frédéric Cao.

The purpose of shape matching is to find methods that are able to find a given shape in images (for recognition), in successive images of a sequence (for tracking), or to retrieve a query shape in a shape database. These shapes first need to be detected and encoded. Since most of the time, recognition needs some invariance properties with respect to some groups of transformations, detection and encoding must also be invariant. This problem has been well automated with the works by Lisani (Palma de Mallorca University). Without entering into details, let us just say that the input of the recognition phase is a set of normalized pieces of shapes, that we call codes. Now, in order to assert that two codes are "the same", we need to measure a distance and a threshold under which recognition should occur. The automatic determination of these recognition thresholds is of course a crucial step. The method we propose is a joint work with P. Musé, F. Sur (ENS Cachan) and Y. Gousseau (ENST Paris). It consists in evaluating the probability that two codes match only for casual reasons. Realistic applications with database containing millions of shapes require to compute probabilities of the order of 10^{-15}, that are of course empirically out of reach. The idea is to build a background model, which finely describes the casual distribution of codes in the vicinity of the query code. We empirically show, that near a query, we can locally describe codes by a set of independent features. This independence allows us to build a probabilistic model from a small number of marginal probability distributions that can be empirically evaluated. By independence, we can then evaluate probabilities of the required magnitude. Hence, we can control with a very high confidence, the number of false matchings between a query and a database. By grouping the detected matches, we could still make the detection more robust. Such a method will hopefully bring us to a robust shape tracking in image sequences. The work on motion detection (see paragraph 6.1.2) will be used to decrease the amount of coded information.

6.2.5. Model-based tracking by virtual visual servoing

Participants: Éric Marchand, François Chaumette, Andrew Comport.

A real-time, robust and efficient 3D model-based tracking algorithm for a monocular vision system has been developed. Tracking objects in the scene amounts to computing the pose between the camera and the objects. Non-linear pose computation is formulated by means of a virtual visual servoing approach. In this context, the derivation of point-to-curves interaction matrices are given for different features including lines, circles, cylinders and spheres. A local moving-edge tracker is used in order to provide a real-time estimation of the displacements normal to the object contours. A method is proposed for combining local position uncertainty and global pose uncertainty in an efficient and accurate way by propagating uncertainty. Robustness is obtained by integrating an M-estimator into the visual control law via an iteratively re-weighted least squares implementation. More recently, we also considered the case of non-rigid objects. The proposed method has been validated on several complex image sequences including outdoor environments. Application for this tracker is robotics and visual servoing as well as augmented reality applications.

6.2.6. Camera displacement estimation and real-time tracking for robotics applications

Participants: Éric Marchand, Muriel Pressigout.

Corresponding points in two successive images are linked by a geometric relation described by the essential or fundamental matrix, or, in some cases, an homography. Our goal is to solve for this geometric relation and to estimate the camera displacement using the visual servoing approach. The method more closely exploits the underlying geometric constraints than the classical solutions. The considered application is model-free real-time augmented reality. Indeed, the knowledge of the model of the scene is an important limitation to the development of real-time augmented reality techniques. This approach is therefore an interesting alternative to model-based tracking.

On another side, we aim at developing robust and real-time algorithms to track unknown objects in an image sequence for robotics applications. Our approach is to consider both motion and edge-based information. Contours of the object, along with its motion, can be represented by parametric models whose parameters have to be estimated over time (they are changing due to the relative movement between the camera and the scene, and to apparent deformations related to the image projection,...). It is necessary to take into account visual information of different types and to develop algorithms robust to erroneous measurements and to potential occlusions. The developed algorithm has to run in real-time (video-rate) in order to be integrated in robotics applications such as visual servoing.

6.2.7. Brain image registration

Participants: Pierre Hellier, Christian Barillot.

Our work in brain image resistration has been directed towards two directions : first, the development of an international evaluation project for non-rigid brain image registration, and secondly, the development of a hybrid non-registration framework combining local landmarks with photometric constraints. We have proposed an evaluation framework, based on global and local measures of the relevance of the registration. We have chosen to focus more particularly on the matching of cortical areas, since intersubject registration methods are dedicated to anatomical and functional normalization. Experiments have been conducted on a database of 18 subjects for 7 registration methods, involving the world leading contributors in this field (SPM, Animal, Demons, Romeo,...). The global objective measures used show that the quality of the registration is directly related to the number of d.o.f. of the considered transformation. More surprisingly, local measures based on the matching of cortical sulci did not show significant differences between rigid and non rigid methods. Functional activations have been added to this framework in order to evaluate the capability of non-rigid registration method to cope with inter-individual functional variations. We have shown that our hybrid framework performed statistically significantly better than all other global non rigid registration method tested. The same framework has been used to evaluate the non-rigid local statistical shape modelling using local statistics of brain shapes and a band-limited radial basis function for the functional areas. This new framework has the lower variance among all methods tested in this framework.

6.2.8. Calibration and reconstruction of 3d ultrasound images

Participants: Christian Barillot, Pierre Hellier, François Rousseau.

We have developed a new robust and fully automatic method for calibrating three-dimensional (3D) free-hand ultrasound. 3D free-hand ultrasound consists in mounting a position sensor on a standard probe. The echographic B-scans can be localized in 3D, and can be compounded into a volume. However, especially for quantitative use, this process dramatically requires a calibration procedure that determines its accuracy and usefulness. Calibration aims at determining the 3D transformation (translations, rotations, scaling) between the coordinate system of the echographic images and the coordinate system of the localization system. To calibrate, we acquire images of a phantom whose 3D geometrical properties are known. We have proposed a robust and fully automatic calibration method based on the Hough transform and robust estimators. Experiments have been carried out with synthetic and real sequences. Our calibration method was shown to be easy to perform, accurate, automatic and fast enough for clinical use.

6.3. Visual servoing and active vision

6.3.1. Visual servoing using image moments

Participants: François Chaumette, Omar Tahri.

This year, we have ended our study on the use of image moments in visual servoing. We first recall that we had determined the analytical form of the interaction matrix related to any image moment, either using the classical perspective projection model, either using a spherical projection model. Objects defined by closed edges or by a set of discrete points have been considered. Thanks to the moment invariants theory, we have also determined optimal combinations of moments so that the corresponding interaction matrix is as much decoupled and linear as possible. This result was valid only when the object is parallel to the image plane. This year we have extended this previous result to the more general case where the desired object position may have any orientation with respect to the camera. We have also developed a new method to estimate the pose of an object from its image moments. This method is composed of two successive steps: the first one provides an initialization of the pose by computing a cost function in a two-dimensional space; the second step consists in improving the estimation of the pose using a virtual visual servoing scheme. Satisfactory experimental results have been obtained on the six degree-of-freedom robotics system available in the group.

6.3.2. Vision-based tasks sequencing

Participants: François Chaumette, Nicolas Mansard.

The aim of this new study is to realize global and high level robotics tasks by an adequate sequencing of local vision-based tasks. Constraints (such as obstacles avoidance, visibility preserving, occlusion avoidance, etc.) will have also to be considered. The idea is to develop a reactive scheme less constraining than classical planning techniques. We plan to use the redundancy formalism of the task function approach for such a purpose.

6.3.3. Control law robust to image outliers

Participants: Éric Marchand, François Chaumette, Andrew Comport, Muriel Pressigout.

A fundamental step towards broadening the use of real world visual servoing involves to deal with the important issues of reliability and robustness. To address this issue, we have designed a closed-loop control law that simultaneously performs a visual servoing task and is robust to a general class of external errors. This property allows concurrent consideration of a wide range of errors including noise from image feature extraction, small scale errors in the tracking and even large scale errors in the matching between current and desired features. This is achieved with the exploitation of a robust M-estimator. M-estimation is classically solved by an iteratively re-weighted least-squares technique. The Median Absolute Deviation is used as an estimate of the standard deviation of the inlier data. Such a combination is advantageous because of its high efficiency, its high breakdown point and the resulting influence function. The robustness and stability of the control law was shown to be dependent on a measure of position uncertainty. Furthermore, the convergence criteria of the control law were investigated. Experimental results have been obtained which demonstrate that visual servoing tasks can tolerate severe outlier contamination using the proposed scheme.

6.3.4. Camera control based on salient features

Participant: Éric Marchand.

We proposed (in collaboration with Nicolas Courty from Siames project) an original model to simulate visual perception based on the detection of salient features. Salient features correspond to the maximum of conspicuousness in a static image or in a sequence of images. We intend to use this information to provide visually interesting targets that can be used in multiple contexts. The extraction process of such an information is performed through multiple steps that correspond to different features locally encoding the conspicuousness of a spatial location. We mainly use for those features spatial orientation and motion information, but other types of features could be used too. Two kinds of applications have been considered : video surveillance application and simulation of the visual perception of a synthetic actor.

6.3.5. Navigation from a visual memory

Participants: François Chaumette, Anthony Remazeilles.

This research study is carried out in collaboration with P. Gros (Texmex team). We have developed a new approach for robot motion control using images acquired by an on-board camera. The main particularity of this method is that it is based on a data-basis composed of images acquired off-line and describing the navigation area. It is thus possible to avoid reconstructing the entire scene without limiting the possible displacements. To reach a desired position, a sequence of overlapping images which define the zone that the robot must traverse is extracted from the data-basis. Robot motions are then computed on-line from the tracking of the points of interest extracted from these images. A method based on the potential field approach has been developed to ensure a sufficient visibility of these features during the robot motion. Satisfactory experimental results have been obtained on the six d.o.f. robotics system available in the group.

7. Contracts and Grants with Industry

7.1. Ifremer contract : automatic analysis of otolith images

Participants: Frédéric Cao, Patrick Bouthemy.

no. Inria 1 03 C 1625 00, duration : 6 months

Otoliths are small chunk pieces that can be found in fishes. Their structure is related to the age of the fish. Hence, statistical otolith analysis may help to evaluate the age distribution of fishes. This could have a huge economical impact since fishing quotas may be modified if we had reliable information on sea resources. The contact we had with Ifremer enabled us to propose geometrical tools for the analysis of otolith. The internship work of Emmanuel d'Angelo was related to this through the automatic detection of (non local) parallel structures in images. Additionally, we proposed an automatic detection of otolith nucleus.

7.2. Cifre grant with Thales Systèmes Aéroportés

Participant: Jean-Pierre Le Cadre.

CNRS **contract, duration 36 months.**

This work corresponds to F. Bavencoff's Ph-D thesis and is carried out in collaboration with J.-M. Vanpeperstraete from Thales. The embarked system of maritime patrol airplane must be able to cope with huge data flows, coming from various types of sensors (radar, ESM, IR, buoys). However, observer maneuvers are relatively rare which renders target motion analysis (TMA) difficult. To overcome this problem, a proper way is to take into account constraints relative to the target which may be of various types : e.g., target location, velocity, maneuvers. If we make the assumption that the target motion is rectilinear and uniform, the constrained TMA problem does not seem particularly difficult since it corresponds to a non-linear constrained optimization problem for which many algorithms are available. However, it is well-known that classical optimization methods perform quite poorly in this context. The likelihood functional is almost flat so that the optimization algorithm provides only an estimation of the bounds. One main difficulty is to obtain a convenient representation of the posterior density itself. In this context, it is worth considering separately "easily" and "hardly" estimable parameters. Estimating posterior density is usually restricted to "hardly" estimable parameters. The problem dimensionality is thus considerably reduced so that Monte Carlo methods (here Metropolis-Hastings) can be efficiently used for estimating the highest probability density. This type of algorithm permits to include a wide variety of constraints ranging from kinematic to topological ones and gives considerable improvements for practical situations. For multi-target tracking, a major concern is the evaluation of the probabilities of correct and false association. In the multi-target tracking domain, they play the same basic role than the well-known probabilities of detection and false-alarm for detection systems. Using a series of difficult approximations, a notion of track purity has thus been derived. Another approach is to consider this calculation directly in terms of regression. Using this framework, calculations become rather simple, thus providing essential insights.

7.3. Cemagref-Ofival contract : Quality evaluation of pork meat

Participants: Eric Marchand, François Chaumette.

no. Inria 1 00 C 0494, duration : 36 months.

This study deals with the evaluation of the quality of pork meat by active vision and IRM measurements. It results from a joint proposal made by Cemagref Rennes, Olympig company and Vista team to Ofival which agreed to financially support this project. Our task is to develop a vision-based estimation of the volume of ham pieces. We have designed algorithms for 3D reconstruction and exploration of a scene using a mobile and controlled camera. We use a space carving algorithm to obtain a precise and robust reconstruction of the 3D structure of (unknown) objects. To ensure the complete reconstruction of all the objects of the scene, we have defined a gaze planning strategy that mainly uses a representation of known and unknown areas as a basis for selecting new viewpoints. The trajectory that realizes this exploration is performed using a visual servoing approach.

7.4. European initiatives

7.4.1. fp5-ist project Lava

Participants: Patrick Bouthemy, Frédéric Cao, Vincent Samson.

no. Inria 1 02 G 0424, duration 36 months

The IST project LAVA ("Learning for Adaptable Visual Assistants") started in May 2002 and involves the following partners : XRCE (prime), IDIAP (Switzerland), Lear team from Inria Rhône-Alpes and Vista team, RHUL (UK), University of Graz (Austria), University of Lund (Sweden), et ANU (Australia). It gathers groups from computer vision, machine learning and cognitive sciences and focuses on two key problems : categorizing objects in static images and interpreting events in video. Two formal meetings have been held this year (in Martigny in January and in Lund in June) and a successful mid-term examination review has taken place in Luxemburg in October with other European Cognitive Vision projects. Vista contribution is related to the workpackages "Image models, descriptors and kernels" and "Recognition of objects and events". We have developed and evaluated several learning techniques for event detection based on the analysis of video dynamic content. Our approach can be divided in two main parts : first, extracting motion descriptors, and secondly, recognizing dynamic events in video sequences. The latter is achieved with a supervised classification based on a training step. Besides probabilistic classifiers (maximum likelihood or Bayesian criterion), we have investigated two other classification techniques : support vector machine (SVM), robust clustering (see paragraph 6.1.4). We have validated these methods on sports video on one hand and on traffic scene video on the other hand. The descriptors used as input data are related to dominant and residual motion information in the first case and to spatial position, object size and velocity in the second case.

7.4.2. fp6-ist noe Muscle

Participants: Patrick Bouthemy, Frédéric Cao.

The Vista team is involved in the Network of Excellence MUSCLE ("Multimedia Understanding through Semantics, Computation and Learning") accepted in September 2003.

7.5. French national programme rntl

7.5.1. Domus Videum project

Participants: Patrick Bouthemy, François Coldefy, Nathalie Peyrard.

no. Inria 2 02 C 0100, duration 30 months.

Domus Videum is a RNTL project led by Thomson-Multimedia and in which the academic labs Irccyn, Irin and Irisa (Metiss, Temics and Vista teams) are involved along with INA and SFRS. The main objective is to design a prototype software of video archiving and browsing dedicated to digital Personal Video Recorders (PVR). This software should store videos on the PVR according to the user profile evaluated dynamically. Then, it provides three ways of stored video browsing : a short insight of the video for previewing, which contains the important sequences of the video (highlights), a long summary which presents the video more extensively, and a thematic video browsing, the shots being merged in homogeneous audio and visual clusters. These functionalities would help the user to refine his choice before watching a video. The application focuses on sports videos for which short and long summary are particularly adapted and on documentary films. We are concerned with the short summary module of the project. We have developed a temporal video segmentation based on camera motion. Sequence of interest are extracted from visual motion descriptors after an off-line learning stage. Video abstraction is finally is obtained by considering audio and descriptors. Convincing results on tennis videos have been obtained. Soccer videos are currently handled. The audio analysis is performed by an algorithm developed by Metiss team which segments the sound track according to particular events as applause, speech, music. We have a large test video base (supplied by INA and SFRS : 46 hours of sports videos (soccer, tennis and ice-skating programs), and 4 hours of documentary videos. At the present time, we are working with 8 hours of video.

7.6. French national programme Priamm

7.6.1. Mediaworks Project

Participants: Patrick Bouthemy, Chafik Kermad, Fabien Spindler.

no. Inria 2 00 A 0408, duration 36 months.

The Mediaworks project was labeled within the Priamm programme and was financially supported by the Ministry of Industry. It ended in July 2003. The partners were LIMSI, TF1, Aegis company and Inria (Imedia team in Inria Rocquencourt and Vista team). The project was concerned with the development of an indexing and retrieval system to assist documentalists. It mainly focused on the cooperation between media (image and text) and on a semantic search engine. We have delivered several modules related to video segmentation and to dynamic-content indexing : shot change detection from MPEG streams, camera motion computation from MPEG vectors and labeling, selection of key-frames to represent the shots, face tracking within a shot. They were validated on news programs. This work was carried out in cooperation with P. Gros (Texmex team).

7.7. French national programme Riam

7.7.1. Feria project

Participants: Patrick Bouthemy, Brigitte Fauvet.

no. Inria 2 03 C 1460, duration : 24 months.

The Feria project was labeled within the RIAM programme and is financially supported by the Ministry of Industry. It started in September 2003. The partners are INA, C&S company, Vecsys company, IRIT, Inria (Texmex and Vista teams), Arte, Canal+ Technologies. The goal of the project is to build a general and open framework allowing the easy development of applications for editing interactive video documents. We are concerned with the design of video processing and representation tools.

7.7.2. Sora project

Participants: Éric Marchand, Cédric Riou, Andrew Comport.

no. Inria 2 03 C 1428, duration : 24 months.

The goal of the Sora project (Objects tracking for augmented reality) is to develop tracking algorithm for augmented reality applications. Our partners are Total-Immersion and VideoMage companies. Augmented reality has now progressed to the point where real-time applications are being considered and needed. On the other hand, it is important that synthetic elements are rendered and aligned in the scene in an accurate and visually acceptable way. In order to address these issues in real-time, robust and efficient 3D model-based tracking algorithm is proposed for a “video see through” monocular vision system. Virtual objects can then be projected into the scene using the camera pose computed using a virtual visual appraoch. The tracking rate is 50Hz.

8. Other Grants and Activities

8.1. Regional initiatives

8.1.1. Brittany Region contract : video summarization

Participants: Gwénaëlle Piriou, Patrick Bouthemy.

no. Inria 4 01 C 0576, duration 36 months.

This contract supplies the financial support of the Inria thesis grant allocated to Gwénaëlle Piriou. The thesis started in November 2001 and is described in paragraphs 6.1.3 and 6.1.4.

8.1.2. Brittany Region contract : reconstruction of 3d ultrasound images

Participants: Arnaud Ogier, Christian Barillot.

no. Inria 4 02 C 0527, duration 36 months.

This contract supplies the financial support of the Inria thesis grant allocated to Arnaud Ogier. The thesis started in November 2002 and is described in paragraph 6.1.7.

8.1.3. cper initiative : aerobio

Participants: Etienne Mémin, Thomas Corpetti.

The AEROBIO project associates the Vista team and the Cemagref-Rennes for the study of methods to prevent contamination in food transformation industry. The purpose of this project was to initiate a research program on the analysis of fluid flow from image sequences. Within this framework, we have in particular worked on the physical evaluation of a dense motion estimator dedicated to fluid flows. This statistical evaluation was carried out on a free turbulent shear layer flow.

8.1.4. cper initiative : 3d echography capture by a robotics system

Participants: Fabien Spindler, Christian Barillot, François Chaumette, Anne-Sophie Tranchant.

In November 2003, we have acquired a new robotics system marketed by Sinters company and dedicated to medical applications, specifically to the remote control of medical ultrasound probes. This system is based on a six degrees-of-freedom intrinsically safe robot. It is equipped with a force/torque sensor mounted at the tip of the robot arm which can move an ultrasonic or Doppler probe on the patient's skin while applying a programmable and constant force. To prove the interest of robotics approaches for medical applications, we will pursue two objectives. First, using ultrasound (US) images recorded over the probe trajectory and the corresponding robot positions, we will work on the 3D ultrasound image volume reconstruction. Second, using an external camera, we will develop visual servoing tasks to achieve an automatically probe positioning, or to control that the probe follows a planned trajectory in order to increase the 3D US reconstruction accuracy.

8.2. National initiatives

8.2.1. aci supported by Ministry of Research : Assimage project

Participants: Etienne Mémin, Anne Cuzol.

no. MD 69, duration 36 months.

The ASSIMAGE project involves three Inria teams (Clime in Rocquencourt, Idopt in Grenoble, and Vista), three Cemagref groups (located in Rennes, Montpellier and Grenoble), the LEGI laboratory and the LGGE laboratory both located in Grenoble. It has started in September 2003. The aim of the ASSIMAGE project is to develop methods for the assimilation of images in mathematical models governed by partial differential equations. The targetted applications concern predictions of geophysical flows.

8.2.2. aci supported by Ministry of Research : NeuroBase project

Participants: Christian Barillot, Isabelle Corouge.

no. CNRS 2P 2008, duration 24 months.

The Vista team is the prime contractor of this project which associates the following other groups : Texmex team from Irisa, IFR 49 "Neuroimagerie Fonctionnelle" (CEA SHFJ, Inserm group U494, CHR Pitié Salpétrière, Paris), Caravel team from Inria Rocquencourt and Epidaure team from Inria Sophia-Antipolis, Inserm group U438 (Grenoble), IDM team from Medical Faculty of University of Rennes 1, TIMC lab (Grenoble). This project aims at establishing the conditions allowing the federation through Internet of distributed information data bases in neuroimaging, information being located in various centers of experimentation, clinical departments in neurology or research centers in cognitive neurosciences. This project consists in specifying how to connect and access distributed information data bases in neuroimaging by the definition of a data-processing architecture allowing the access and the sharing of experimentation results or even data processing methods within a same site or between different sites. This would for example enable the search for similar results within these information data bases, the search for images containing singularities, or transverse searches to highlight possible regularities (similarly to a "data mining" approach).

8.2.3. Initiatives supported by cnrs

- DSTIC **specific action 67 : particle filtering methods**
  Participants: Jean-Pierre Le Cadre, Thomas Bréhard, Carine Hue.
  Our contribution to that action led by F. Le Gland (Sigma-2 team) is concerned with the tracking of moving entities in sonar/radar or in video (see paragraphs 6.2.1, 6.2.3).
- DSTIC **specific action : real-time augmented reality**
  Participants: Eric Marchand, Andrew Comport, Christian Barillot, Pierre Hellier.
  This action, led by M.-O. Berger (ISA team in Inria Lorraine), is concerned with augmented reality. A. Comport and E. Marchand have presented their research work in this area in a meeting in Paris in September 2003. Eric Marchand's participation in the IEEE International Symposium on Mixed and Augmented Reality was supported by this action.
- DSTIC **specific action : non-rigid image registration**
  Participants: Christian Barillot, Pierre Hellier.
  The objective of this specific action coordinated by N. Rougon (INT Evry), which is part of the RTP 25 and 42, is to gather people working on non-rigid image registration problems. It encompasses the review of the world community work, especially for the concern of multimodal non-rigid registration, the identification of the methodological and application bottlenecks, the organization of the French community involved in this domain, the dissemination and the promotion of the results in the French image community.

8.2.4. robea Robotics research programme

- **Omnibot project**
  Participant: François Chaumette.
  This two-year project started in October 2002. The aim of the project is to develop robot navigation schemes using a panoramic vision sensor. It involves a collaboration between several academic labs, LASMEA in Clermont-Ferrand, LIRMM in Montpellier, CREA in Amiens and our team. This year, we have determined the analytical form of the interaction matrix related to image moments using the specific geometrical models of panoramic sensors. A visual servoing control law has then been derived from this modeling step. A project meeting has been organized at Irisa in November 2003.
- **Complex dynamic scene interpretation and reactive motion planning.**
  Participants: Eric Marchand, Fabien Spindler.
  This three-year project, which started in October 2002, is a collaboration between Inria Rhône-Alpes (Sharp, Movi and Prima teams), the LAAS (RIA group) and Inria Rennes (Vista team). The objective of this project is the automatic control of a vehicle in road environment. For this year, we have focused on the detection and tracking of mobile obstacles by a moving camera. To this end, we have used motion analysis algorithms (motion detection and tracking) previously developed in the Vista team. One of the originality of our approach is that we do not extract geometrical information from images (no spatial image segmentation) but we directly exploit the intensity image sequence. In addition, we consider a visual servoing module in order to maintain the tracked object of interest in the camera field of view.

- **Tasks sequencing project**
  Participants: François Chaumette, Nicolas Mansard.
  This three-year project started in September 2003. The aim of this project is to develop task sequencing schemes to realize high level robotics tasks from local sensor-based control techniques. Comparison between vision-based control and human beings for gaze control is also considered. It involves a collaboration between LAAS, CERCO and ENIT, all located in Toulouse, and our team.
- **Bodega project**
  Participants: François Chaumette, Anthony Remazeilles.
  This two-year project has started in November 2003. Its goal is to design vision-based and sensor-based methods for the autonomous navigation of mobile vehicles moving around an urban environment. It involves a collaboration between ENSIL in Limoges, UTC UTC in Compiègne, LASMEA in Clermont-Ferrand, the Icare team of Inria Sophia-Antipolis, and our team. A project meeting has been organized at Irisa in November 2003.
- **Robvolint project**
  Participant: François Chaumette.
  This two-year project has started in November 2003. Its aims is to develop vision-based localization techniques and visual servoing schemes for small helicopters moving around an indoor environment. It involves a collaboration between I3S in Nice, CEA in Fontenay-aux-Roses, and our team.

8.3. Bilateral international co-operation

8.3.1. Van Gogh programme, France-Netherlands

Participants: Pierre Hellier, Christian Barillot, François Rousseau.

This scientific exchange action has been set up with the ISI group at U. Hospital Utrecht (M. Viergever, W. Niessen) in order to share our respective experience in the domain of usage of 3D ultrasound in neurosurgery. One of the practical work done in this project was to evaluate spatial and temporal calibration procedures for acquisition performed using free-hand mode. This aspect is particularly acute when these data have to be used in the surgical context. Three techniques have been studied : an industrial one from Medtronic, the one proposed with StradX and the one developed by ourselves (see paragraph 6.2.8).

8.3.2. cnrs-cihr programme, France-Canada

Participant: Christian Barillot.

This action has been granted by CNRS and CIHR (the Canadian National Institute of Health Research). With this grant, Christian Barillot visited the Robarts Research Institute of London at University of Western Ontario for four months (from April to August). The local partners are T. Peters and A. Fenster. The purpose of this visit was first the organisation of the MICCAI conference - http://www.miccai.org - (organized this year in Canada by T. Peters, and in 2004 in France by C. Barillot). The second topic was the development of new applications in the 3D ultrasound domain.

8.3.3. cmcu programme, France-Tunisie

Participants: Christian Barillot, Pierre Hellier, Patrick Bouthemy.

This three-year scientific exchange project involves several teams from France and Tunisia and is concerned with medical image analysis. It was labeled in May 2003.

8.3.4. Inria-grices project, France-Portugal

Participants: François Chaumette, Éric Marchand, Anthony Remazeilles.

This collaboration with IST Lisbon, Portugal (Prof. J. Santos-Victor and Prof. R. Caldas Pinto) is concerned with visual servoing and image sequence analysis for robotics applications. A. Remazeilles and E. Marchand have spent a three days visit at IST in December 2003 and two PhD students from IST have come for a short visit in our team in November 2003.

8.3.5. Visiting scientists

- Jordi Pagès, Ph-D student at the University of Girona, did a three-month visit from January to March 2003. He worked on the modeling of visual features for visual servoing using structured light. His supervisor, Prof. J. Salvi, visited us at the end of his stay.
- Short visits by L. Alvarez (Univ. Las Palmas), A. Doucet (Univ. Cambridge), W. Niessen (Univ. Utrecht), R. Prager (Univ. Cambridge).

8.3.6. Project reviewing

- C. Barillot has been appointed by the European Commission to review project submissions within the IST program of the 6th FP, and more especially FET projects.
- C. Barillot is consultant/collaborator and member of the Advisory Board of the Human Brain Mapping NIH project “A probabilistic referential system for the human brain” (P.I. J.C. Mazziotta). This grant is associating UCLA (A.W. Toga, J.C. Mazziotta), the University of Texas in Austin (P. Fox) and the Montreal Neurological Institute at McGill University (A. Evans). He is associated to this project as consultant/collaborator and member of the Advisory Board, notably on the aspects of segmentation and modeling of distortion applied to the human brain.
- P. Bouthemy expertised two company projects in the domain of video indexing for ANVAR.
- F. Chaumette did a three-day visit at IST Lisbon, Portugal, in July 2003 as external expert of the Portuguese project POCTI/EME/39946/2001 entitled “Visual Servoing Control Applied to Robot Manipulators”.
- C. Kervrann is member of the Scientific Council of the Biometry and Artificial Intelligence Department of Inra.

9. Dissemination

9.1. Leadership within scientific community

- The Vista team is involved in the French network GDR ISIS, “Information, Signal and Images” (http://www-isis.enst.fr/).
- P. Bouthemy participates in the Board Committee of Technovision programme supported by the Ministry of Research and by DGA and aiming at developing evaluation projects of image processing and computer vision techniques (call for proposals to be published early 2004).
- F. Chaumette participates in the Scientific Committee of the CNRS Autonomous and Communicating Robotics program (RTP 17).
- J.-P. Le Cadre participates in the working group “Ultimate performances in trajectography” settled by Thales, and to an evaluation group of the DGA.
- **Editorial boards of journals**
  - C. Barillot is member of the Editorial Board of *Technique et Science Informatiques* (TSI).
  - F. Chaumette is Associate Editor of IEEE *Trans. on Robotics and Automation*.
  - J.-P. Le Cadre is Associate Editor of *Advances in Information Fusion* (ISIF).

- **Conference organization**
  - C. Barillot is General Chairman of MICCAI'2004 which we organize and will be held in September 2004 in Rennes and Saint-Malo. P. Hellier, S. Prima, P. Bouthemy and E. Marchand are members of the Local Organizing Committee.
  - P. Bouthemy was General Chairman of the workshop CBMI'2003 which we organized in Rennes in September 2003 (http://www.irisa.fr/CBMI03). F. Spindler and S. Lemaile were members of the local Organizing Committee.
  - C. Kervrann is member of the Organizing Committee of the 5th French-Danish workshop SSIAB'2004 (Spatial Statistics and Image Analysis in Biology) to be held in Grenoble in May 2004.
  - J.-P. Le Cadre has co-organized a one-day meeting on particle filtering in December 2003 in Paris.
- **Technical program committees of conferences**
  - C. Barillot : member of the board and area chair of MICCAI'2003, TPC member of IPMI'2003, ICPR'2004, IEEE-ISBI'2004.
  - P. Bouthemy : TPC member of CVPR'2003, ICCV'2003, ICIP'2003, TAIMA'2003, ACCV'2004, CVPR'2004, ICPR'2004, ICME'2004, CIVR'2004, Medianet'2004, RFIA'2004.
  - F. Cao : TPC member of Scale-Space'2003, SSIAI'2004, organization and chair of a special session on Image Sequence Analysis at ISSPA'2003.
  - F. Chaumette : TPC member of CVPR'2003, ECCV'2004, JNRR'2003, member of the Editorial Committee of RFIA'2004, member of the ICRA'2003 Best Student Paper Award committee.
  - C. Kervrann : TPC member of ICIP'2003.
  - J.-P. Le Cadre : TPC member of Fusion'2003, NATO "Target Tracking and Sensor Data Fusion for Military Observation Systems" 2003.
  - P. Hellier : TPC member of MICCAI'2003, program co-chair of MICCAI'2004.
  - E. Mémin : TPC member of ICIP'2003, Orasis'2003, CVPR'2004, ECCV'2004.

9.2. Teaching

- Master STIR "Signal, Telecommunications, Image and Radar", University of Rennes 1, (E. Mémin : statistical image analysis, P. Bouthemy : motion analysis, J.-P. Le Cadre : array processing).
- Master of Computer Science, Ifsic, University of Rennes 1 (E. Marchand : 3D models for computer vision and computer graphics; E. Mémin : motion analysis; P. Bouthemy : video indexing).
- DIIC INC, Ifsic, University of Rennes 1 (E. Marchand, E. Mémin, F. Chaumette, E. Arnaud : 3D vision, motion analysis, visual servoing; E. Mémin : Markov models for image analysis; C. Barillot : medical imagery; E. Marchand, F. Spindler : programming tools for image processing), E. Mémin is in charge of the INC (Digital image analysis and communication) channel.
- Insa Rennes, Electrical Engineering Dpt (F. Chaumette, N. Mansard, G. Piriou, T. Veit : computer vision).
- Master GBM, Universities of Tours, Angers, Nantes and Rennes 1 (C.Barillot : 3D medical imagery).

- Master TIS, University of Cergy and ENSEA (P. Bouthemy : video indexing).
- Master PIC and ENSPS Strasbourg, (P. Bouthemy : image sequence analysis).
- Master MVA, ENS Cachan (F. Cao : image filtering and PDE).
- ENSAI Rennes, 3rd year (C. Kervrann : statistical models and image analysis).
- Graduate student interns : T. Boucher (Ecole Polytechnique), G. Boulanger (ESEO, Angers), A. Cuzol (ENSAI and Master Applied Statistics, Rennes 1), E. d'Angelo (Sup-Aero), E. Fouret (Insa Rennes and Master Computer Science, Rennes 1), D. Giot (Master STIR, Rennes 1), Rahul Gupta (IIT, India, Inria internship), J.-F. Marcos Pablo (Univ. Valladolid, Spain, Erasmus programme), M. Pressigout (Insa Rennes and Master Computer Science Rennes 1), Abishek Sharma (Banaras Hindu Univ., India).
- I. Corouge is doing a Post-doc at University of North Carolina (Prof. G. Gerig), Chapel Hill, USA, since March 2003.
- External thesis supervision :
  - A. Alhaj (Cemagref, Rennes) supervised by F. Chaumette;
  - F. Bavencoff (Thales Airborne Systems, Cifre grant) supervised by J.-P. Le Cadre (see 7.2);
  - J. Pages (University of Girona, Spain and Cemagref Rennes) supervised by F. Chaumette.

9.3. Participation in seminars, invitations, awards

- F. Chaumette received with E. Malis the 2002 King Sun Fu Memorial Best IEEE Trans. on Robotics and Automation Paper Award.
- P. Bouthemy gave an invited talk entitled "Modélisation et reconnaissance du mouvement dans des séquences d'images : entre modèles physiques et apprentissage statistique" at the "Journées thématiques" of GDR ISIS which was held in Dourdan in April 2003.
- P. Bouthemy was an invited speaker at the International Workshop on Computer Vision and Image Analysis organized by University of Las Palmas in December 2003 (talk entitled "Probabilistic models of image motion for recognition of dynamic content in video").
- F. Chaumette was invited for a three-day visit at the University of Alicante, Spain in June 2003. He was also invited in September 2003 at the University of Girona, Spain, as external reviewer of the Xavier Quintana's Ph-D thesis defense, and in December 2003 at the University of Waterloo, Canada, as external reviewer of the Lingfeng Deng's Ph-D thesis defense.
- F. Chaumette gave an invited talk at JNRR'2003 in October 2003 on recent French research advances in visual servoing.
- J.-P. Le Cadre gave an invited talk at the SEE COGIS meeting in Paris in May 2003 entitled "Gestion de capteurs pour la détection de cibles mobiles", and at the SEE Radar meeting in Paris in November 2003 entitled "Pistage avant détection".
- J.-P. Le Cadre was an invited speaker at the DSTO workshop in Adelaide in July 2003 (talk entitled "An overview of sensor management for target tracking").
- J.-F. Yao gave an invited talk entitled "A motion-based event detection method for video sequences" at SCRA'2003 (Forum for Interdisciplinary Mathematics on Statistics Combinatorics and Related Areas), Portland, Maine, USA.
- J.-F. Yao was invited for a visit of 3 weeks at Beijing University, Institute of Mathematics (Prof. Qian Mingping). He also visited the National Laboratory of Pattern Recognition (Prof. H. Lu, Image & Video Analysis group) of the Chinese Academy of Sciences.

10. Bibliography

Major publications by the team in recent years

[1] P. BOUTHEMY, E. FRANÇOIS. *Motion segmentation and qualitative dynamic scene analysis from an image sequence.* in « Int. Journal of Computer Vision », number 2, volume 10, avril, 1993, pages 157-182.

[2] B. ESPIAU, F. CHAUMETTE, P. RIVES. *A new approach to visual servoing in robotics.* in « IEEE Transactions on Robotics and Automation », number 6, volume 8, juin, 1992, pages 313-326.

[3] H. GAUVRIT, C. JAUFFRET, J.-P. LE CADRE. *A formulation of multitarget tracking as an incomplete data problem.* in « IEEE Trans. on Aerospace & Electr. Systems », number 4, volume 33, octobre, 1997, pages 1242-1257.

[4] F. HEITZ, P. PÉREZ, P. BOUTHEMY. *Multiscale minimization of global energy functions in some visual recovery problems.* in « CVGIP : Image Understanding », number 1, volume 59, 1994, pages 125-134.

[5] P. HELLIER, C. BARILLOT, E. MÉMIN, P. PÉREZ. *Hierarchical estimation of a dense deformation field for 3D robust registration.* in « IEEE Transactions on Medical Imaging », number 5, volume 20, 2001, pages 388-402.

[6] J.-P. LE CADRE, C. JAUFFRET. *Discrete-time observability and estimability for bearings-only target motion analysis.* in « IEEE Trans. on Aerospace & Electr. Systems », number 1, volume 33, janvier, 1997, pages 178-201.

[7] E. MALIS, F. CHAUMETTE, S. BOUDET. *2 1/2 D visual servoing.* in « IEEE Transactions on Robotics and Automation », number 2, volume 15, avril, 1999, pages 238-250.

[8] E. MARCHAND, F. CHAUMETTE. *An autonomous active vision system for complete and accurate 3D scene reconstruction.* in « Int. Journal of Computer Vision », number 3, volume 32, août, 1999, pages 171-194.

[9] E. MÉMIN, P. PÉREZ. *Dense estimation and object-based segmentation of the optical flow with robust techniques.* in « IEEE Trans. on Image Processing », number 5, volume 7, mai, 1998, pages 703-719.

[10] J.-M. ODOBEZ, P. BOUTHEMY. *Robust multiresolution estimation of parametric motion models.* in « Journal of Visual Communication and Image Representation », number 4, volume 6, décembre, 1995, pages 348-365.

Books and Monographs

[11] F. CAO. *Geometric Curve Evolution and Image Processing.* series Lecture Notes in Mathematics, number 1805, Springer Verlag, 2003.

Doctoral dissertations and “Habilitation” theses

[12] I. COROUGE. *Modélisation statistique de formes en imagerie cérébrale.* Ph. D. Thesis, Université de Rennes 1, mention Informatique, April, 2003.

[13] C. HUE. *Méthodes séquentielles de Monte-Carlo pour le filtrage non linéaire multi-objets dans un environne-*

ment bruité. Applications au pistage multi-cibles et à la trajectographie d'entités dans des séquences d'images 2D. Ph. D. Thesis, Université de Rennes 1, mention Informatique, January, 2003.

[14] E. MÉMIN. *Estimation du flot optique : contributions et panorama de différentes approches.* Habilitation à diriger des recherches, Université de Rennes 1, July, 2003.

[15] P. PÉREZ. *Modèles et algorithmes pour l'analyse probabiliste des images.* Habilitation à diriger des recherches, Université de Rennes 1, December, 2003.

[16] F. ROUSSEAU. *Méthodes d'analyse d'images et de calibration pour l'échographie 3D en mode main-libre.* Ph. D. Thesis, Université de Rennes 1, mention Traitement du Signal et Télécommunications, December, 2003.

Articles in referred journals and book chapters

[17] I. BLOCH, J.-P. LE CADRE, H. MAÎTRE. *Approches probabilistes et statistiques.* I. BLOCH, editor, in « Fusion d'informations en traitement du signal et des images », series Traité IC2, Hermes, 2003, chapter 6, pages 87-118.

[18] F. CAO. *Application of the Gestalt principles to the detection of good continuations and corners in image level lines.* in « Computing and Visualisation in Science », to appear.

[19] F. CHAUMETTE. *Image moments: a general and useful set of features for visual servoing.* in « IEEE Transactions on Robotics », to appear.

[20] I. COROUGE, P. HELLIER, B. GIBAUD, C. BARILLOT. *Inter-individual functional mapping: a non linear local approach.* in « Neuroimage », number 4, volume 19, August, 2003, pages 1337-1348.

[21] T. CORPETTI, E. MÉMIN, P. PÉREZ. *Extraction of singular points from dense motion fields : an analytic approach.* in « Journal of Mathematical Imaging and Vision », number 3, volume 19, 2003, pages 175-198.

[22] F. DAMBREVILLE, J.-P. LE CADRE. *Spatio-temporal multi-mode management for moving target detection.* in « Journal of Information Fusion », to appear.

[23] R. FABLET, P. BOUTHEMY. *Motion recognition using non parametric image motion models estimated from temporal and multiscale cooccurrence statistics.* in « IEEE Trans. on Pattern Analysis and Machine Intelligence », number 12, volume 25, December, 2003, pages 1619-1624.

[24] P. HELLIER, C. BARILLOT, I. COROUGE, B. GIBAUD, G. LE GOUALHER, D. COLLINS, A. EVANS, G. MALANDAIN, N. AYACHE, G. CHRISTENSEN, H. JOHNSON. *Retrospective evaluation of inter-subject brain registration.* in « IEEE Transactions on Medical Imaging », number 9, volume 22, 2003, pages 1120-1130.

[25] P. HELLIER, C. BARILLOT. *A hierarchical parametric algorithm for deformable multimodal image registration.* in « Computer Methods and Programs in Biomedicine », to appear.

[26] P. HELLIER, C. BARILLOT. *Coupling dense and landmark-based approaches for non rigid registration.* in « IEEE Transactions on Medical Imaging », number 2, volume 22, February, 2003, pages 217-227.

[27] E. MARCHAND, F. CHAUMETTE. *Reconstruction 3D par vision dynamique active.* M. DHOME, editor, in « Perception visuelle par imagerie vidéo », series Traité IC2, Hermes, 2003, chapter 6, pages 213-249.

[28] Y. MEZOUAR, F. CHAUMETTE. *Optimal camera trajectory with image-based control.* in « International Journal of Robotics Research », number 10/11, volume 22, October, 2003, pages 781-804.

[29] L. OISEL, E. MÉMIN, L. MORIN, F. GALPIN. *1D dense disparity estimation for 3D reconstruction.* in « IEEE Trans. on Image Processing », number 9, volume 12, September, 2003, pages 1107-1119.

[30] N. PEYRARD, P. BOUTHEMY. *Motion-based selection of relevant video segments for video summarization.* in « Multimedia Tools and Applications Journal », to appear.

[31] I. PRATIKAKIS, C. BARILLOT, P. HELLIER, E. MÉMIN. *Robust multiscale deformable registration of 3D Ultrasound images.* in « International Journal of Image and Graphics », number 4, volume 3, October, 2003, pages 547-565.

[32] J.-N. PROVOST, C. COLLET, P. ROSTAING, P. PÉREZ, P. BOUTHEMY. *Hierarchical Markovian segmentation of multispectral images for the reconstruction of water depth maps.* in « Computer Vision and Image Understanding », to appear.

[33] J.-P. LE CADRE, V. NIMIER, R. REYNAUD. *Fusion en traitement du signal.* I. BLOCH, editor, in « Fusion d'informations en traitement du signal et des images », series Traité IC2, Hermes, 2003, chapter 2, pages 29-52.

Publications in Conferences and Workshops

[34] A. ALHAJ, C. COLLEWET, F. CHAUMETTE. *Couplage entre asservissement visuel et reconstruction 3D par vision dynamique.* in « 19e colloque GRETSI sur le traitement du signal et des images », volume 1, pages 181-184, Paris, September, 2003.

[35] A. ALHAJ, C. COLLEWET, F. CHAUMETTE. *Visual servoing based on dynamic vision.* in « IEEE Int. Conf. on Robotics and Automation, ICRA'03 », volume 3, pages 3055-3060, Taipeh, Taiwan, May, 2003, http://www.irisa.fr/vista/Papers/2003_icra_alhaj.pdf.

[36] E. ARNAUD, E. MÉMIN, B. CERNUSHI FRIAS. *Filtrage conditionnel pour le suivi de points dans des séquences d'images..* in « Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur, ORASIS'2003 », pages 211–220, Gérardmer, France, May, 2003.

[37] T. BRÉHARD, J.-P. LE CADRE, M. PRENAT, Y. COJAN. *Initialization of Particle filter and Posterior Cramer-Rao Bound for Bearings-Only Tracking in Modified Polar Coordinate System.* in « Symposium on Target Tracking and Sensor Data Fusion for Military Observation », Sensors & Electronic Technology Panel, Budapest, Hungary, October, 2003.

[38] F. CAO. *Good continuations in digital image level lines.* in « IEEE Int. Conf. on Computer Vision, ICCV'03 », pages 440-447, Nice, France, October, 2003.

[39] F. CHAUMETTE. *Avancées récentes en asservissement visuel.* in « Journées nationales de la recherche en robotique, JNRR'03 », pages 103-108, Clermont-Ferrand, October, 2003.

[40] A. COMPORT, E. MARCHAND, F. CHAUMETTE. *A real-time tracker for markerless augmented reality.* in « ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR'03 », pages 36-45, Tokyo, Japan, October, 2003.

[41] A. COMPORT, E. MARCHAND, F. CHAUMETTE. *Suivi générique et temps-reel d'objets 3D par asservissement visuel virtuel.* in « Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur, ORASIS'2003 », pages 139–146, Gérardmer, France, May, 2003.

[42] A. COMPORT, M. PRESSIGOUT, E. MARCHAND, F. CHAUMETTE. *A visual servoing control law that is robust to image outliers.* in « IEEE Int. Conf. on Intelligent Robots and Systems, IROS'03 », Las Vegas, Nevada, October, 2003.

[43] I. COROUGE, M. DOJAT, C. BARILLOT. *Statistical shape modeling of unfolded retinotopic maps for a visual areas probabilistic atlas.* in « Medical Image Computing and Computer Assisted Intervention, MICCAI'03 », Montreal, Canada, November, 2003.

[44] T. CORPETTI, E. MÉMIN, A. SANTA-CRUZ, D. HEITZ, G. ARROYO. *Optical flow estimation in experimental fluid mechanics.* in « Seventh Int. Symp. on Signal Processing and its Applications, ISSPA'03 », Paris, France, July, 2003.

[45] N. COURTY, F. LAMARCHE, S. DONIKIAN, E. MARCHAND. *A Cinematography System for Virtual Storytelling.* in « Int. Conf. on Virtual Storytelling, ICVS'03, », series Lecture Notes in Computer Science, O. BALET, G. SUBSOL, P. TORGUET, editors, Toulouse, France, November, 2003.

[46] N. COURTY, E. MARCHAND, B. ARNALDI. *A new application for salliency maps: Synthetic vision of autonomous actors..* in « IEEE Int. Conf. on Image Processing, ICIP'03 », Barcelona, Spain, September, 2003.

[47] N. COURTY, E. MARCHAND. *Visual perception based on salient features.* in « IEEE Int. Conf. on Intelligent Robots and Systems, IROS'03 », Las Vegas, Nevada, October, 2003.

[48] P. HELLIER. *Consistent intensity correction of MR images.* in « IEEE conf. on Image Processing, ICIP'03 », Barcelona, Spain, September, 2003.

[49] T. KOHLBERGER, E. MÉMIN, C. SCHNOERR. *Variational dense motion estimation using a div-curl higher order regularization.* in « Int. Symp. on Signal Processing and Its Applications, ISSPA'03 », Paris, France, July, 2003.

[50] T. KOHLBERGER, E. MÉMIN, C. SCHNOERR. *Variational dense motion estimation using the Helmholtz decomposition.* in « Scale Space'03 », series Lecture Notes in Computer Science, Springer Verlag, volume 2695, L. GRIFFIN, M. LILLHOLM, editors, pages 432-448, Isle of Skye, UK, June, 2003.

[51] P. MUSÉ, F. SUR, F. CAO, Y. GOUSSEAU. *Unsupervised thresholds for shape matching.* in « IEEE Int. Conf. on Image Processing, ICIP'2003 », Barcelona, Spain, September, 2003,

http://www.irisa.fr/vista/Papers/2003_icip_muse.pdf.

[52] N. PEYRARD, P. BOUTHEMY. *Detection of meaningful events in videos based on a supervised classification approach.* in « IEEE Int. Conf. on Image Processing, ICIP 2003 », Barcelona, Spain, September, 2003.

[53] N. PEYRARD, P. BOUTHEMY. *Motion-based selection of relevant video segments for video summarisation.* in « IEEE Int. Conf. on Multimedia & Expo, ICME'2003 », Baltimore, July, 2003.

[54] G. PIRIOU, N. PEYRARD, P. BOUTHEMY, J.-F. YAO. *Détection d'événements dans les vidéos à l'aide de modèles probabilistes de mouvement.* in « Colloque franco-tunisien TAIMA'03 », Hammamet, Tunisia, September, 2003, http://www.irisa.fr/vista/Papers/2003_taima_piriou.pdf.

[55] S. PRIMA, D. ARNOLD, D. COLLINS. *Multivariate statistics for detection of MS activity in serial multimodal MR images.* in « Sixth International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI'2003 », series Lecture Notes in Computer Science, Springer Verlag, Montreal, Canada, November, 2003.

[56] F. ROUSSEAU, R. FABLET, C. BARILLOT. *Robust Statistical Registration of 3D Ultrasound Images Using Texture Information.* in « IEEE Int. Conf. on Image Processing, ICIP'03 », Barcelona, Spain, September, 2003.

[57] F. ROUSSEAU, P. HELLIER, C. BARILLOT. *Calibration method for 3D freehand ultrasound.* in « Medical Image Computing and Computer Assisted Intervention, MICCAI'03 », Montreal, Canada, November, 2003, http://www.irisa.fr/vista/Papers/2003_miccai_rousseau.pdf.

[58] F. ROUSSEAU, P. HELLIER, C. BARILLOT. *Méthode de calibration pour système échographique 3D main-libre.* in « Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur, ORASIS'2003 », pages 7–14, Gérardmer, France, May, 2003, http://www.irisa.fr/vista/Papers/2003_orasis_rousseau.pdf.

[59] O. TAHRI, F. CHAUMETTE. *Application of moment invariants to visual servoing.* in « IEEE Int. Conf. on Robotics and Automation, ICRA'03 », volume 3, pages 4276-4281, Taipeh, Taiwan, May, 2003, http://www.irisa.fr/vista/Papers/2003_icra_tahri.pdf.

[60] O. TAHRI, F. CHAUMETTE. *Asservissement visuel à partir de moments invariants.* in « Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur, ORASIS'2003 », pages 165–174, Gérardmer, France, May, 2003, http://www.irisa.fr/vista/Papers/2003_orasis_tahri.pdf.

[61] J.-F. YAO, G. PIRIOU, P. BOUTHEMY. *A motion-based event detection method for video sequences.* in « Int. Conf. on Statistics, Combinatorics and Related Areas », Portland, ME, October, 2003 (abstract only).

[62] J.-P. LE CADRE. *Detection of a target moving in a network.* in « Fusion'2003 », Cairns, Australia, July, 2003.

Internal Reports

[63] F. CAO, P. MUSÉ, F. SUR. *Extracting meaningful curves from images.* Technical report, number RR-5067, INRIA, Dec., 2003, http://www.inria.fr/rrrt/rr-5067.html.

[64] A. COMPORT, E. MARCHAND, F. CHAUMETTE. *Robust and real-time image-based tracking for markerless augmented reality.* Technical report, number 4847, INRIA, April, 2003, http://www.inria.fr/rrrt/rr-4847.html.

[65] O. TAHRI, F. CHAUMETTE. *Determination of moment invariants and their application to visual servoing.* Technical report, number 4845, INRIA, June, 2003, http://www.inria.fr/rrrt/rr-4845.html.

Miscellaneous

[66] P. HELLIER, C. BARILLOT. *JuLIEt : Joint use of Landmarks and Intensities for Elastic registration.* Inria, Cnrs, April, 2003, Software Copyright.

Bibliography in notes

[67] 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.

[68] D. DEMENTHON, L. DAVIS. *Model-based object pose in 25 lines of code.* in « Int. Journal of Computer Vision », number 1/2, volume 15, juin, 1995, pages 123-141.

[69] H. GAUVRIT. *Extraction multipistes: approches probabilistes et combinatoires.* Ph. D. Thesis, université de Rennes 1, novembre, 1997.

[70] S. GEMAN, D. GEMAN. *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images.* in « IEEE Transactions on Pattern Analysis and Machine Intelligence », number 6, volume 6, 1984, pages 721-741.

[71] S. HUTCHINSON, G. HAGER, P. CORKE. *A tutorial on visual servo control.* in « IEEE Transactions on Robotics and Automation », number 5, volume 12, octobre, 1996, pages 651-670.

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

[73] C. SAMSON, M. L. BORGNE, B. ESPIAU. *Robot Control: the Task Function Approach.* Clarendon Press, Oxford, 1991.

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

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