



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team VISTA

Vision Spatio-Temporelle et Apprentissage

Rennes

THEME COG

Activity
R *eport*

2006

Table of contents

1. Team	1
2. Overall Objectives	1
2.1. Overall Objectives	1
3. Scientific Foundations	2
3.1. Motion estimation and motion segmentation with mrf models	2
3.2. Fluid motion analysis	4
3.3. Object tracking with non-linear probabilistic filtering	5
3.4. Perceptual grouping for image and motion analysis	7
4. Application Domains	8
4.1. Application Domains	8
4.1.1. Content-aware video applications	9
4.1.2. Experimental fluid mechanics and meteorological imagery	9
4.1.3. Biological imagery	10
5. Software	10
5.1. Motion2d software - parametric motion model estimation	10
5.2. d-Change software - motion detection	11
5.3. Dense-Motion software - optical flow computation	11
5.4. VTrack software - generic interactive visual tracking platform	11
5.5. safir-nd - image denoising software	12
6. New Results	12
6.1. Image sequence processing and modeling	12
6.1.1. Biological image sequence denoising	12
6.1.2. Motion-compensated temporal denoising of transparent image sequences	13
6.1.3. Object detection and dynamical background estimation	13
6.1.4. Motion detection in adverse conditions	14
6.1.5. Conditional mixed-state Markov models	14
6.1.6. Motion texture modeling and segmentation	15
6.2. Motion estimation and matching	16
6.2.1. Fluid motion estimation	16
6.2.2. Motion estimation and spatial support segmentation of transparent layers	17
6.2.3. Repetitive motion and video alignment	17
6.3. Tracking	17
6.3.1. Stochastic filtering for fluid motion tracking	17
6.3.2. Robust visual tracking without prior	18
6.3.3. Variational data assimilation for visual tracking	19
6.3.4. Image sequence simulation and trafficking analysis	20
6.4. Dynamic event modeling, learning and recognition	20
6.4.1. Target behaviour analysis	20
6.4.2. Event detection in video	21
6.4.3. Video content recognition using trajectories	21
6.4.4. Extracting information from a video sensor network	22
6.4.5. Hierarchical surveillance: a solution learning approach	22
7. Contracts and Grants with Industry	22
7.1. ft-rd contract : Probabilistic visual tracking for annotation of team sport broadcasts	22
7.2. Cifre grant with Thomson-rd	23
7.3. Cifre grant with General Electric Healthcare	23
7.4. cea-Cesta contract: Target tracking in a dense environment	24
7.5. Grant from Dupont De Nemours	24
7.6. European initiatives	24

7.6.1.	fp6 Fluid project	24
7.6.2.	fp6 Pegase project	25
7.6.3.	fp6-ist noe Muscle	25
7.6.4.	fp6-ist tn Visiontrain	26
8.	Other Grants and Activities	26
8.1.	Regional initiatives	26
8.1.1.	prir Brittany council program : Imafluid project	26
8.2.	National initiatives	26
8.2.1.	aci Ministry grant : Assimage project	26
8.2.2.	anr “Jeunes chercheurs” program, stands-msg project	27
8.2.3.	aci Ministry grant : Behaviour project	27
8.2.4.	aci Ministry grant : Modyncell5d project	28
8.2.5.	aci Ministry grant : project related to the development of new fluorescent probes	28
8.2.6.	Other involvements	28
8.3.	Bilateral international co-operation	29
8.3.1.	Collaboration with University of Buenos-Aires, Inria associate team fim	30
8.3.2.	Advanced Research Program (pra) France-China	31
8.3.2.1.	Visiting scientists	31
9.	Dissemination	31
9.1.	Leadership within scientific community	31
9.2.	Teaching	32
9.3.	Participation in seminars, invitations, awards	33
10.	Bibliography	34

1. Team

Head of project-team

Patrick Bouthemy [DR Inria, HdR]

Administrative assistant

Huguette Béchu [TR Inria, shared with Temics project-team]

Inria research scientists

Ivan Laptev [CR]

Patrick Pérez [DR, HdR]

Charles Kervrann [Inria research scientist, on Inria secondment (mise à disposition)]

CNRS research scientists

Jean-Pierre Le Cadre [DR, HdR]

Faculty members

Étienne Mémin [Associate professor, Ifsic, University of Rennes 1, HdR]

Jian-Feng Yao [Full professor, Irmr, UFR Maths, University of Rennes 1, (collaborateur extérieur), HdR]

Bruno Cernuschi-Frias [Visiting professor, UBA, Inria associate team FIM, June 2006]

Inria technical staff

Nicolas Gengembre [FT-RD contract until 31/05/06, Fluid IST project from 01/06/06]

Patrick Heas [Fluid IST project]

Post-doctoral fellows

Gwénaëlle Piriou [Lecturer (ATER) Ifsic, University of Rennes 1]

Ph-D students

Kamel Aouichat [Cifre grant, Video Metrix, from 01/03/06]

Vincent Auvray [Cifre grant, General Electric Healthcare, until 31/10/06]

Vijay Badrinarayanan [Cifre grant, Thomson-RD, from 01/03/06]

Jérôme Boulanger [Inria grant]

Aurélien Bugeau [Research Ministry grant]

Tomas Crivelli [Faculty of Engineering, University of Buenos-Aires, “thèse en co-tutelle”]

Anne Cuzol [Research Ministry grant until 30/09/06, Inria grant from 01/10/06]

Emilie Dexter [DGA grant, from 01/10/06]

Alexandre Hervieu [Inria grant, FP6-NOE Muscle]

Adrien Ickowicz [Inria grant, from 01/11/06, Cesta contract]

Nicolas Papadakis [Inria grant, FP6-IST Fluid project]

Thierry Pécot [Inria grant, from 01/10/06]

Cécile Simonin [DGA grant, with CEP-DGA Arcueil]

2. Overall Objectives

2.1. Overall Objectives

Keywords: *Bayesian estimation, Markov models, a contrario decision, biological imagery, dynamic scene analysis, experimental fluid mechanics, fluid motion analysis, image sequence, meteorological imagery, motion detection, motion recognition, motion segmentation, optic flow, particle filtering, registration, robust estimation, statistical learning, statistical modeling, tracking, trajectography, video indexing, video processing.*

Vista research work is concerned with various types of spatio-temporal images, (mainly video images, but also meteorological satellite images, video-microscopy, X-ray images). We are investigating methods to analyze dynamic scenes, and, more generally, dynamic phenomena, within image sequences. We address the full range of problems raised by the analysis of such dynamic contents with a focus on image motion analysis issues: denoising, motion detection, motion estimation, motion-based segmentation, tracking, motion recognition and interpretation with learning. We usually rely on statistical approaches, resorting to: Markov models, Bayesian inference, robust estimation, *a contrario* decision, particle filtering, learning. Application-wise, we focus our attention on three main domains: content-aware video applications, meteorological imaging and experimental visualization in fluid mechanics, biological imaging. For that, a number of collaborations, academic and industrial, national and international, are set up.

3. Scientific Foundations

3.1. Motion estimation and motion segmentation with mrf models

Keywords: *Markov models, energy function, motion estimation, motion segmentation, optic flow, parametric motion models, robust estimation.*

Assumptions (i.e., data models) must be formulated to relate the observed image intensities to motion, and other constraints (i.e., motion models) must be added to solve problems like motion segmentation, optical flow computation, or motion recognition. The motion models are supposed to capture known, expected or learned properties of the motion field ; this implies to somehow introduce spatial coherence or more generally contextual information. The latter can be formalized in a probabilistic way with local conditional densities as in Markov models. It can also rely on predefined spatial supports (e.g., blocks or pre-segmented regions). The classic mathematical expressions associated with the visual motion information are of two types. Some are continuous variables to represent velocity vectors or parametric motion models. The others are discrete variables or symbolic labels to code motion detection (binary labels), motion segmentation (numbers of the motion regions or layers) or motion recognition output (motion class labels).

In the past years, we have addressed several important issues related to visual motion analysis, in particular with a focus on the type of motion information to be estimated and the way contextual information is expressed and exploited. Assumptions (i.e., data models) must be formulated to relate the observed image intensities to motion, and other constraints (i.e., motion models) must be added to solve problems like motion segmentation, optical flow computation, or motion recognition. The motion models are supposed to capture known, expected or learned properties of the motion field ; this implies to somehow introduce spatial coherence or more generally contextual information. The latter can be formalized in a probabilistic way with local conditional densities as in Markov models. It can also rely on predefined spatial supports (e.g., blocks or pre-segmented regions). The classic mathematical expressions associated with the visual motion information are of two types. Some are continuous variables to represent velocity vectors or parametric motion models. The others are discrete variables or symbolic labels to code motion detection (binary labels), motion segmentation (numbers of the motion regions or layers) or motion recognition output (motion class labels). We have also recently introduced new models, called mixed-state models and mixed-state auto-models, whose variables belong to a domain formed by the union of discrete and continuous values. We briefly describe here how such models can be specified and exploited in two central motion analysis issues: motion segmentation and motion estimation.

The brightness constancy assumption along the trajectory of a moving point $p(t)$ in the image plane, with $p(t) = (x(t), y(t))$, can be expressed as $dI(x(t), y(t), t)/dt = 0$, with I denoting the image intensity function. By applying the chain rule, we get the well-known motion constraint equation:

$$r(p, t) = \mathbf{w}(p, t) \cdot \nabla I(p, t) + I_t(p, t) = 0, \quad (1)$$

where ∇I denotes the spatial gradient of the intensity, with $\nabla I = (I_x, I_y)$, and I_t its partial temporal derivative. The above equation can be straightforwardly extended to the case where a parametric motion model is considered, and we can write:

$$r_\theta(p, t) = \mathbf{w}_\theta(p, t) \cdot \nabla I(p, t) + I_t(p, t) = 0, \quad (2)$$

where θ denotes the vector of motion model parameters.

One important step ahead in solving the motion segmentation problem was to formulate the motion segmentation problem as a statistical contextual labeling problem or in other words as a discrete Bayesian inference problem. Segmenting the moving objects is then equivalent to assigning the proper (symbolic) label (i.e., the region number) to each pixel in the image. The advantages are mainly two-fold. Determining the support of each region is then implicit and easy to handle: it merely results from extracting the connected components of pixels with the same label. Introducing spatial coherence can be straightforwardly (and locally) expressed by exploiting MRF models. Here, by motion segmentation, we mean the competitive partitioning of the image into motion-based homogeneous regions. Formally, we have to determine the hidden discrete motion variables (i.e., region numbers) $l(i)$ where i denotes a site (usually, a pixel of the image grid; it could be also an elementary block). Let $l = \{l(i), i \in S\}$. Each label $l(i)$ takes its value in the set $\Lambda = \{1, \dots, N_{reg}\}$ where N_{reg} is also unknown. Moreover, the motion of each region is represented by a motion model (usually, a 2D affine motion model of parameters θ which have to be conjointly estimated; we have also explored non-parametric motion modeling [81]). Let $\Theta = \{\theta_k, k = 1, \dots, N_{reg}\}$. The data model of relation (2) is used. The *a priori* on the motion label field (i.e., spatial coherence) is expressed by specifying a MRF model (the simplest choice is to favour the configuration of the same two labels on the two-site cliques so as to yield compact regions with regular boundaries). Adopting the Bayesian MAP criterion is then equivalent to minimizing an energy function E whose expression can be written in the general following form:

$$E(l, \Theta, N_{reg}) = \sum_{i \in S} \rho_1[r_{\theta_{l(i)}}(i)] + \sum_{i \sim j} \rho_2[l(i), l(j)], \quad (3)$$

where $i \sim j$ designates a two-site clique. We first considered [73] the quadratic function $\rho_1(x) = x^2$ for the data-driven term in (3). The minimization of the energy function E was carried out on l and Θ in an iterative alternate way, and the number of regions N_{reg} was determined by introducing an extraneous label and using an appropriate statistical test. We later chose a robust estimator for ρ_1 [90], [91]. It allowed us to avoid the alternate minimization procedure and to determine or update the number of regions through an outlier process in every region.

Specifying (simple) MRF models at a pixel level (i.e., sites are pixels and a 4- or 8-neighbour system is considered) is efficient, but remains limited to express more sophisticated properties on region geometry or to handle extended spatial interaction. Multigrid MRF models [84] is a means to address somewhat the second concern (and also to speed up the minimization process while usually supplying better results). An alternative is to first segment the image into spatial regions (based on grey level, colour or texture) and to specify a MRF model on the resulting graph of adjacent regions [82]. The motion region labels are then assigned to the nodes of the graph (which are the sites considered in that case). This allowed us to exploit more elaborated and less local *a priori* information on the geometry of the regions and their motion. However, the spatial segmentation stage is often time consuming, and getting an effective improvement on the final motion segmentation accuracy remains questionable.

By definition, the velocity field formed by continuous vector variables is a complete representation of the motion information. Computing optical flow based on the data model of equation (1) requires to add a motion model enforcing the expected spatial properties of the motion field, that is, to resort to a regularization method. Such properties of spatial coherence (more specifically, piecewise continuity of the motion field) can be expressed on local spatial neighborhoods. First methods to estimate discontinuous optical flow fields were based on MRF models associated with Bayesian inference (i.e., minimization of a discretized energy function).

A general formulation of the global (discretized) energy function to be minimized to estimate the velocity field \mathbf{w} can be given by:

$$E(\mathbf{w}, \zeta) = \sum_{p \in S} \rho_1[r(p)] + \sum_{p \sim q} \rho_2[\|\mathbf{w}(p) - \mathbf{w}(q)\|, \zeta(p'_{p \sim q})] + \sum_{A \in \chi} \rho_3(\zeta_A), \quad (4)$$

where S designates the set of pixel sites, $r(p)$ is defined in (1), $S' = \{p'\}$ the set of discontinuity sites located midway between the pixel sites and χ is the set of cliques associated with the neighborhood system chosen on S' . We first used quadratic functions and the motion discontinuities were handled by introducing a binary line process ζ [83]. Then, robust estimators were popularized leading to the introduction of so-called auxiliary variables ζ now taking their values in $[0, 1]$ [89]. Multigrid MRF are moreover involved, and multiresolution incremental schemes are exploited to compute optical flow in case of large displacements. Dense optical flow and parametric motion models can also be jointly considered and estimated, which enables to supply a segmented velocity field [88]. Depending on the followed approach, the third term of the energy $E(\mathbf{w}, \zeta)$ can be optional.

3.2. Fluid motion analysis

Keywords: *continuity equation, div-curl regularization, experimental fluid mechanics, fluid motion analysis, meteorological image sequences, singular points.*

Analyzing fluid motion is essential in number of domains and can rarely be handled using generic computer vision techniques. In this particular application context, we study several distinct problems. We first focus on the estimation of dense velocity maps from image sequences. Fluid flows velocities cannot be represented by a single parametric model and must generally be described by accurate dense velocity fields in order to recover the important flow structures at different scales. Nevertheless, in contrast to standard motion estimation approach, adapted data model and higher order regularization are required in order to incorporate suitable physical constraints. In a second step, analysing such velocity fields is also a source of concern. When one wants to detect particular events, to segment meaningful areas, or to track characteristic structures, dedicated methods must be devised and studied.

Since several years, the analysis of video sequences showing the evolution of fluid phenomena has attracted a great deal of attention from the computer vision community. The applications concern domains such as experimental visualization in fluid mechanics, environmental sciences (oceanography, meteorology, ...), or medical imagery.

In all these application domains, it is of primary interest to measure the instantaneous velocity of fluid particles. In oceanography, one is interested to track sea streams and to observe the drift of some passive entities. In meteorology, both at operational and research levels, the task under consideration is the reconstruction of wind fields from the displacements of clouds as observed in various satellite images. In medical imaging, the issue can be to visualize and analyze blood flow inside the heart, or inside blood vessels. The images involved in each domain have their own characteristics and are provided by very different sensors. The huge amount of data of different kinds available, the range of applicative domains involved, and the technical difficulties in the processing of all these specific image sequences explain the interest of the image analysis community.

Extracting dense velocity fields from fluid images can rarely be done with the standard computer vision tools. The latter were originally designed for quasi-rigid motions with stable salient features, even if these techniques have proved to be more and more efficient and provide accurate results for natural images [89], [88]. These generic approaches are based on the brightness constancy assumption of the points along their trajectory ($\frac{df}{dt} = 0$), along with the spatial smoothness assumption of the motion field. These estimators are defined as the minimizer of the following energy function:

$$\int_{\Omega} \rho[\nabla f \cdot w + \frac{\partial f}{\partial t}] ds + \alpha \int_{\Omega} \rho(\|\nabla w\|) ds. \quad (5)$$

The penalty function ρ is usually the L_2 norm, but it may be substituted for a robust function attenuating the effect of data that deviate significantly from the brightness constancy assumption [89], and enabling also to implicitly handle the spatial discontinuities of the motion field.

Contrary to usual video image sequence contents, fluid images exhibit high spatial and temporal distortions of the luminance patterns. The design of alternative approaches dedicated to fluid motion thus constitutes a widely-open research problem. It requires to introduce some physically relevant constraints which must be embedded in a higher-order regularization functional [77]. The method we have devised for fluid motion involves the following global energy function:

$$\int_{\Omega} \rho\{f(\mathbf{s} + \mathbf{w}, t + 1) \exp(\operatorname{div} \mathbf{w}) - f(\mathbf{s}, t)\} ds + \alpha \int_{\Omega} (\|\nabla \operatorname{div} \mathbf{w}\|^2 + \|\nabla \operatorname{curl} \mathbf{w}\|^2) ds.$$

The first term comes from an integration of the continuity equation (assuming the velocity of a point is constant between instants t and $t + \Delta t$). Such a data model is a “fluid counterpart” of the usual “Displaced Frame Difference” expression. Instead of expressing brightness constancy, it explains a loss or gain of luminance due to a diverging motion. The second term is a smoothness term designed to preserve divergence and vorticity blobs. This regularization term is nevertheless very difficult to implement. As a matter of fact, the associated Euler-Lagrange equations consist in two fourth-order coupled PDE’s, which are tricky to solve numerically. We proposed to simplify the problem by introducing auxiliary functions, and by defining the following alternate smoothness function:

$$\int_{\Omega} |\operatorname{div} \mathbf{w} - \xi|^2 + \lambda \rho(\|\nabla \xi\|) ds + \alpha \int_{\Omega} |\operatorname{curl} \mathbf{w} - \zeta|^2 + \lambda \rho(\|\nabla \zeta\|) ds. \quad (6)$$

The new auxiliary scalar functions ξ and ζ can be respectively seen as estimates of the divergence and the curl of the unknown motion field, and λ is a positive parameter. The first part of each integral enforces the displacement to comply with the current divergence and vorticity estimates ξ and ζ , through a quadratic goodness-of-fit enforcement. The second part associates the divergence and the vorticity estimates with a robust first-order regularization enforcing piece-wise smooth configurations. From a computational point of view, such a regularizing function only implies the numerical resolution of first-order PDE’s. It may be shown that, at least for the L_2 norm, the regularization we proposed is a smoothed version of the original second-order div-curl regularization.

Once given a reliable description of the fluid motion, another important issue consists in extracting and characterizing structures of interest such as singular points or in deriving potential functions. The knowledge of the singular points is precious to understand and predict the considered flows, but it also provides compact and hierarchical representations of the flow [78]. Such a compact representation enables for instance to tackle difficult tracking problems. As a matter of fact, the problem amounts here to track high dimensional complex objects such as surfaces, level lines, or vector fields. As these objects are only partially observable from images and driven by non linear 3D laws, we have to face a tough tracking problem of large dimension for which no satisfying solution exists at the moment.

3.3. Object tracking with non-linear probabilistic filtering

Keywords: *data association, data fusion, importance sampling, multi-object tracking, particle filter, sequential Monte Carlo.*

Tracking problems that arise in target motion analysis (TMA) and video analysis are highly non-linear and multi-modal, which precludes the use of Kalman filter and its classic variants. A powerful way to address this class of difficult filtering problems has become increasingly successful in the last ten years. It relies on sequential Monte Carlo (SMC) approximations and on importance sampling. The resulting sample-based filters, also called particle filters, can, in theory, accommodate any kind of dynamical models and observation models, and permit an efficient tracking even in high dimensional state spaces. In practice, there is however a number of issues to address when it comes to difficult tracking problems such as long-term visual tracking under drastic appearance changes, or multi-object tracking.

The detection and tracking of single or multiple targets is a problem that arises in a wide variety of contexts. Examples include sonar or radar TMA and visual tracking of objects in videos for a number of applications (e.g., visual servoing, tele-surveillance, video editing, annotation and search). The most commonly used framework for tracking is that of Bayesian sequential estimation. This framework is probabilistic in nature, and thus facilitates the modeling of uncertainties due to inaccurate models, sensor errors, environmental noise, etc. The general recursions update the posterior distribution of the target state $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, also known as the filtering distribution, where $\mathbf{y}_{1:t} = (\mathbf{y}_1 \cdots \mathbf{y}_t)$ denotes all the observations up to the current time step, through two stages:

$$\begin{aligned} \text{prediction step:} \quad & p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})d\mathbf{x}_{t-1} \\ \text{filtering step:} \quad & p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}, \end{aligned}$$

where the prediction step follows from marginalisation, and the new filtering distribution is obtained through a direct application of Bayes' rule. The recursion requires the specification of a dynamic model describing the state evolution $p(\mathbf{x}_t|\mathbf{x}_{t-1})$, and a model for the state likelihood in the light of the current measurements $p(\mathbf{y}_t|\mathbf{x}_t)$. The recursion is initialised with some distribution for the initial state $p(\mathbf{x}_0)$. Once the sequence of filtering distributions is known, point estimates of the state can be obtained according to any appropriate loss function, leading to, e.g., Maximum *A Posteriori* (MAP) and Minimum Mean Square Error (MMSE) estimates.

The tracking recursion yields closed-form expressions in only a small number of cases. The most well-known of these is the Kalman Filter (KF) for linear and Gaussian dynamic and likelihood models. For general non-linear and non-Gaussian models the tracking recursion becomes analytically intractable, and approximation techniques are required. Sequential Monte Carlo (SMC) methods [80], [86], [85], otherwise known as particle filters, have gained a lot of popularity in recent years as a numerical approximation strategy to compute the tracking recursion for complex models. This is due to their efficiency, simplicity, flexibility, ease of implementation, and modeling success over a wide range of challenging applications.

The basic idea behind particle filters is very simple. Starting with a weighted set of samples $\{w_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}\}_{n=1}^N$ approximately distributed according to $p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})$, new samples are generated from a suitably designed proposal distribution, which may depend on the old state and the new measurements, i.e., $\mathbf{x}_t^{(n)} \sim q(\mathbf{x}_t|\mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t)$, $n = 1 \cdots N$. Importance sampling theory indicates that a consistent sample is maintained by setting the new importance weights to

$$w_t^{(n)} \propto w_{t-1}^{(n)} \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(n)})p(\mathbf{x}_t^{(n)}|\mathbf{x}_{t-1}^{(n)})}{q(\mathbf{x}_t^{(n)}|\mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t)}, \quad \sum_{n=1}^N w_t^{(n)} = 1,$$

where the proportionality is up to a normalising constant. The new particle set $\{w_t^{(n)}, \mathbf{x}_t^{(n)}\}_{n=1}^N$ is then approximately distributed according to $p(\mathbf{x}_t|\mathbf{y}_{1:t})$. Approximations to the desired point estimates can then be obtained by Monte Carlo techniques. From time to time it is necessary to resample the particles to avoid degeneracy of the importance weights. The resampling procedure essentially multiplies particles with high importance weights, and discards those with low importance weights.

In many applications, the filtering distribution is highly non-linear and multi-modal due to the way the data relate to the hidden state through the observation model. Indeed, at the heart of these models usually lies a data association component that specifies which part, if any, of the whole current data set is “explained” by the hidden state. This association can be implicit, like in many instances of visual tracking where the state specifies a region of the image plane. The data, e.g., raw color values or more elaborate descriptors, associated to this region only are then explained by the appearance model of the tracked entity. In case measurements are the sparse outputs of some detectors, as with edgels in images or bearings in TMA, association variables are added to the state space, whose role is to specify which datum relates to which target (or clutter).

In this large context of SMC tracking techniques, two sets of important open problems are of particular interest for Vista:

- selection and on-line estimation of observation models with multiple data modalities: except in cases where detailed prior is available on state dynamics (e.g., in a number of TMA applications), the observation model is the most crucial modeling component. A sophisticated filtering machinery will not be able to compensate for a weak observation model (insufficiently discriminant and/or insufficiently complete). In most adverse situations, a combination of different data modalities is necessary. Such a fusion is naturally allowed by SMC, which can accommodate any kind of data model. However, there is no general means to select the best combination of features, and, even more importantly, to adapt online the parameters of the observation models associated to these features. The first problem is a difficult instance of discriminative learning with heterogeneous inputs. The second problem is one of online parameter estimation, with the additional difficulty that the estimation should be mobilized only parsimoniously in time, at instants that must be automatically determined (adaptation when the entities are momentarily invisible or simply not detected by the sensors will always cause losses of track). These problems of feature selection, online model estimation, and data fusion, have started to receive a great deal of attention in the visual tracking community, but proposed tools remain ad-hoc and restricted to specific cases.
- multiple-object tracking with data association: when tracking jointly multiple objects, data association rapidly poses combinatorial problem. Indeed, the observation model takes the form of a mixture with a large number components indexed by the set of all admissible associations (whose enumeration can be very expensive). Alternatively, the association variables can be incorporated within the state space, instead of being marginalised out. In this case, the observation model takes a simpler product form, but at the expense of a dramatic dimension increase of the space in which the estimation must be conducted.

In any case, strategies have thus to be designed to keep low the complexity of the multi-object tracking procedure. This need is especially acute when SMC techniques, already often expensive for a single object, are required. One class of approach consists in devising efficient variants of particle filters in the high-dimensional product state space of joint target hypotheses. Efficiency can be achieved, to some extent, by designing layered proposal distributions in the compound target-association state space, or by marginalising out approximately the association variables. Another set of approaches lies in a crude, yet very effective approximation of the joint posterior over the product state space into a product of individual posteriors, one per object. This principle, stemming from the popular JPDAF (joint probabilistic data association filter) of the trajectography community, is amenable to SMC approximation. The respective merits of these different approaches are still partly unclear, and are likely to vary dramatically from one context to another. Thorough comparisons and continued investigation of new alternatives are still necessary.

3.4. Perceptual grouping for image and motion analysis

Keywords: *Helmoltz principle, a contrario decision, motion detection, number of false alarms, parameter-free method, perceptual grouping, shape matching.*

We have recently been interested in automatic detection problems in image and video sequence processing. A fundamental question is to know whether it is possible to automatically direct the attention to some object of interest (in the broad sense). We have been using a general grouping principle, asserting that conspicuous events are those that have a very small probability of occurrence in a random situation. We have applied this principle formalized within an a contrario decision framework to the detection of moving objects in an image sequence, to image comparison, and to the matching of shapes in images.

For the last few years, we have been interested in developing methods of image and video analysis with no complex *a priori* model. Of course, in this case the purpose is not to analyse and finely describe complex situations. On the contrary, we try to achieve very low-level vision tasks, with the condition that the methods must be very stable and provide a measure of validity of the detected structures. A qualitative principle, called the Helmholtz principle, was developed a few years ago in École Normale Supérieure de Cachan [79], and we used it in low-level motion analysis. This principle basically states that local observations have to be grouped with respect to some qualitative property, if the probability that they share this quality is very small, assuming that the quality is independently and identically distributed on these observations. In some sense, this can be related to a more classical hypothesis testing setting. Let us express it in the context of detection of motion in a given region R of the image. We would like to test the hypothesis H_0 “there is motion in R ” against H_1 “there is no motion”. The problem is that, usually, we do not have any precise model for H_0 . On the opposite, we model the background model H_1 (the absence of motion), by the fact that the pointwise observations are independent. This is sound since this hypothesis amounts to say that the observations are only due to noise. We then decide that H_0 is true whenever the probability of occurrence of the observed values in R are much improbable under the independence hypothesis H_1 . We call this methodology an *a contrario* decision framework.

More generally, assume given a set of local measures of some quantity Q . We also make local (pointwise or close to pointwise) observations on the images $(O_k)_{k \in K}$, where K is a set of spatial indices. Assume also that, for one reason or another, we can design some group candidates G_1, \dots, G_n , of the local observations, that is to say subsets of $\{O_k, k \in K\}$. We also consider an adequacy measure of $X_{G_i}(O_k)$ which we assume small when the quality Q is satisfied by O_k , relatively to G_i . As a simple example, we can consider as G_i the digital segments of the image, O_k a direction field defined at each position and $X_{G_i}(O_k)$ as the difference between the field at position k and the direction of the segment G_i . Finally, let u be an image. We ask the following question: “in u , is Q a good reason to consider G_i as a group?”

Helmholtz Principle. Assume that u is a realization of a random image U where it is assumed that, anything else being equal, the random variables $O_k(U)$ are independent and identically distributed. The group G_i is all the more conspicuous that the probability

$$\mathcal{P}(\forall k, X_{G_i}(O_k(U)) \leq X_{G_i}(O_k(u)))$$

is small.

From this qualitative perceptual principle, we can define the number of false alarms of a configuration, which is the expectation of its number of occurrence in the background model of independence. It can be proven, that this number is a very good and robust measure of the meaningfulness of a configuration. We have applied this principle to the detection of good continuations and corners, of straight lines trajectories of subpixel target, and more recently to the detection of moving objects in images, to image comparison and to shape matching.

4. Application Domains

4.1. Application Domains

Keywords: biological imagery, biology, defense, environment, experimental fluid mechanics, meteorological imagery, multimedia, vehicle navigation, video indexing, video processing.

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

- Content-aware video applications (Thomson, FT-RD, INA);
- Experimental fluid mechanics (Cemagref) and meteorological imagery (LMD). We are also leading the FET-IST European project Fluid (with see paragraph 7.6.1) and are in the Inria associate team FIM with the University of Buenos-Aires (see paragraph 8.3.1);
- Biological imagery (Inra, Curie Institute, Biology Dpt of University of Rennes 1)
- Surveillance (Onera, Thales, collaborations are nevertheless considered only from an academic viewpoint). The main addressed issues are search and surveillance, navigation, distributed tracking with a sensor network.

4.1.1. Content-aware video applications

The amount of video footage is constantly increasing due to the dissemination of video cameras, the broadcasting of TV programs by multiple means, the seamless acquisition of personal videos,...The exploitation of video material, whatever its usage, requires automatic (or at least semi-automatic) tools to process video contents. A wide range of applications can be envisaged dealing with editing, analyzing, annotating, browsing and authoring video contents. Video indexing and retrieval for audio-visual archives is, for instance, a major application, which is receiving lot of attention. Other needs include the creation of enriched videos, the design of interactive video systems, the generation of video summaries, and the development of re-purposing frameworks (specifically, for 3G mobile phones and Web applications). For most of all these applications, tools for segmenting videos, detecting events or recognizing actions are usually required.

We are mainly interested in the processing of videos which are shot (and broadcast) in the audiovisual domain, more specifically, sports videos but also TV shows or dance videos. Amateur videos of similar content can also be within our concern. On one hand, sports videos raise difficult issues, since the acquisition process is weakly controlled and content exhibits high complexity, diversity and variability. On the other hand, motion is tightly related to sports semantics. Besides, the exploitation of sports videos forms an obvious business target. We have developed several methods and tools in that context addressing issues such as shot change detection, camera motion estimation and characterization, object tracking, motion modeling and recognition, event detection, video summarization. Beside this main domain of applications, we are also investigating gesture analysis problems. An on-going project in particular aims at monitoring automatically car drivers' attention.

4.1.2. Experimental fluid mechanics and meteorological imagery

Concerning the analysis of fluid flows from image sequences, we focus mainly on the domains of experimental fluid mechanics and meteorological imaging. We aim at designing new methods allowing us to extract kinematic or dynamical descriptors of fluid flows from image sequences. We have to face an huge amount of high resolution image sequences. These data reveal in a more and more accurate way the spatio-temporal evolution of flow structures in a non intrusive way. The kinds of data involved in these applicative domains may be various, depending on the experimental imaging set-up and/or the image sensor used. Very specific applications may be tackled for some type of images, but general and common goals can nevertheless be defined in term of motion analysis. Image motion estimation aims at providing instantaneous measurements of the flow velocity and at bringing to physicists kinematic elements allowing them to analyze complex fluid flows. In both domains, the estimation of velocity flow fields from an image sequence is routinely performed with local methods which rely on the computation of average displacements by cross-correlation over small search windows. Despite sophisticated block-matching schemes have been designed in order to cope with intrinsic difficulties of particle-seeded images or atmospheric satellite images, these approaches can hardly cope with low contrast visualization techniques such as Schlieren images or images of the MSG (Meteosat Second Generation satellite) water vapor channel. These methods are not convenient also to get dense velocity fields accurate enough at different scales and for spatially varying motions in order to exhibit for instance the relevant flow features. Besides, the incorporation of fluid flow dynamic laws (almost inescapable in a near future with upcoming high time resolution image sequences) cannot be really handled with local correlation

methods. As a matter of fact, no spatial and temporal coherency can be handled with such processing techniques as they operate entirely in a data-driven way allowing no inclusion of physical prior knowledge (related to the basic equations of fluid mechanics). From that point of view, motion analysis techniques developed in computer vision are particularly relevant as they combine model-driven variational smoothness functions with data-driven terms.

On such a basis, as for the meteorological domain, the first objective we are pursuing consists in designing techniques for an accurate estimation of the atmospheric wind fields. Such a goal should require fine sophisticated schemes incorporating physical models of the atmosphere. The second goal is to propose methods for tracking cloud systems of importance, which are useful when one aims at monitoring potentially dangerous events such as convective clouds, hurricane, tornadoes, etc. These two issues have potentially a great impact on weather forecasting, risk prevention, or enhancement of global atmospheric circulation model assimilation.

As for experimental fluid mechanics, we are investigating new methods for the analysis of complex fluid flows from image sequences. A large range of applications is concerned for instance with turbulent flows in aerodynamics, aeronautics, heat transfer, etc. Applications involving flow control are of particular interest (flow separation delay, mixing enhancement, drag reduction,...). These applications need enhanced visualization and sound numerical techniques such as low-order modeling with reduced dynamical models. The processing of real data and the accuracy enhancement of spatio-temporal measurements may together bring improvements in the modeling of turbulent flows which is traditionally solely based on initial conditions captured through experimental conditions.

4.1.3. Biological imagery

Recent progresses in molecular biology and light microscopy make henceforth possible the acquisition of multi-dimensional data (3D + time) at one or several wavelengths (multispectral imaging) and the observation of intra-cellular molecular dynamics at sub-micron resolutions. Automatic image processing methods to study molecular dynamics from image sequences are therefore of major interest, for instance, for membrane trafficking involving the movement of small particles from donor to acceptor compartments within the living cell.

The challenge is then to track GFP tags (fluorescent proteins for labeling) with high precision in movies representing several gigabytes of image data. The data are collected and processed automatically to generate information on partial or complete trajectories. In our research work, we are developing methods to perform the computational analysis of these complex 3D image sequences since the capabilities of most commercial image analysis tools for automatically extracting information are rather limited and/or require a large amount of manual interactions with the user. In the applications we address, we mainly focus on the analysis of vesicles that deliver cellular components to appropriate places within cells.

Applications of the proposed image processing methods to biological problems should provide a new and quantitative way for interpreting the movement of fluorescently labeled membrane transport vesicles. Quantitative analysis of data obtained by fast 4D deconvolution microscopy allows one to enlighten the role of specific Rab proteins on HeLa human cell lines. The role of Rab proteins is viewed as to organize membrane platforms serving for protein complexes to act at the required site within the cell. Methods have been developed for specific Rab6a and Rab6a' proteins - involved in the regulation of transport from the Golgi apparatus to the endoplasmic reticulum. These small proteins act as molecular motors to move transport intermediates along polymers called microtubules. A second application concerns the CLIP 170 protein involved in the kinetochores anchorage (in the segregation of chromosomes to daughter cells, the chromosomes appear to be pulled via a so-called kinetochore attached to chromosome centromeres) using new fluorescent probes (Quantum Dots).

5. Software

5.1. Motion2d software - parametric motion model estimation

Participants: Fabien Spindler (Lagadic team), Patrick Boutheymy.

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 [90]. 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 (last release 1.3.11, January 2005) 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 downloaded at the same Web address.

5.2. d-Change software - motion detection

Participants: Fabien Spindler (Lagadic team), Patrick Boutheymy.

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. VTrack software - generic interactive visual tracking platform

Participants: Nicolas Gengembre, Patrick Pérez.

As part of the research contract with FT-RD (see paragraph 7.1)), we have developed an interactive tracking platform (Windows Visual C++ development with Microsoft MFC and Intel OpenCV). It includes both state-of-the-art generic tracking methods (template matching, feature tracking, kernel-based tracking with global color characterization, particle filtering) and original developments (see paragraph 6.3.2), as well as a number of visualization features for enhanced experimental and demonstration experiences. The flexible architecture and the rich HCI allow easy design, implementation and test of novel trackers.

5.5. safir-nd - image denoising software

Participants: Charles Kervrann, Jérôme Boulanger.

The SAFIR-ND software written in C++ enables to remove additive Gaussian and non-Gaussian noise in a still 2D or 3D image or in a 2D or 3D image sequence (with no motion computation). The method is unsupervised. It is based on a pointwise selection of small image patches of fixed size in (a data-driven adapted) spatial or space-time neighborhood of each pixel (or voxel). The main idea is to associate with each pixel (or voxel) the weighted sum of intensities within an adaptive 2D or 3D (or 2D or 3D + time) neighborhood and to use image patches to take into account complex spatial interactions. The neighborhood size is selected at each spatial or space-time position according to a bias-variance criterion. The algorithm requires no tuning of control parameters (already calibrated with statistical arguments) and no library of image patches. The method has been applied to real noisy images (old photographs, JPEG-coded images, videos, ...) and is exploited in different biomedical application domains (fluorescence microscopy, video-microscopy, MRI imagery, X-ray imagery, ultrasound imagery, ...). This algorithm outperforms most of the best published denoising methods for still images or image sequences.

6. New Results

6.1. Image sequence processing and modeling

6.1.1. Biological image sequence denoising

Participants: Charles Kervrann, Patrick Bouthemy, Jérôme Boulanger.

New video-microscopy technology enables to acquire 4-D data that require the design and the development of specific image denoising methods able to preserve details and discontinuities in both the $(x - y - z)$ space dimensions and the time dimension. Images are noisy due to the weakness of the fluorescence signal in time-lapse recording. Accordingly, we have developed an original and efficient spatio-temporal filtering method for significantly increasing the signal-to-noise ratio (SNR) in noisy fluorescence microscopic image sequences where small particles have to be tracked from frame to frame. The proposed “adaptive window approach” is conceptually simple, being based on the local estimation of a weighted average of the intensities (for the considered regression model) within an adaptively and locally selected space-time window size (neighborhood). We use statistical 4-D data-driven criteria to automatically select the size of the adaptive space-time neighborhood. At each pixel, we estimate the weighted average by iteratively increasing a space-time window to achieve an optimal compromise between bias and variance corresponding to the minimization of the pointwise L_2 risk of the local estimator. The method involves also a patch-based similarity step to fix the weights. The proposed algorithm complexity is actually controlled by simply limiting the size of the largest window and the patch size.

In addition, theoretical properties of the non-parametric estimator have been proved. Recently, we have shown that the proposed estimation procedure can be interpreted as a steepest descent algorithm related to the fixed point solution corresponding to the minimization of a global energy function involving non-local terms and local image contexts described by patches. We have applied this method to noisy synthetic and real 4-D images where a large number of small fluorescently labeled vesicles move in regions close to the Golgi apparatus within the cell. Preliminarily, the assumed Poisson noise is transformed into a Gaussian noise using an original variance stabilization procedure based on a generalized Anscombe transform. The SNR is shown to be drastically improved and enhanced images can then be correctly segmented. The objective is to report evidences about the lifetime kinetics of specific Rab3 for membrane transport. This novel approach can be further used for biological studies where dynamics have to be analyzed in molecular and subcellular bio-imaging. The patch-based method combined with Radon transform has been used to denoise images with a very low number of photon counts (< 1 photon/pixel). Finally, let us also point out that we have applied our adaptive patch-based denoising method to usual video sequences, and it was demonstrated that it outperforms other recent methods.

6.1.2. Motion-compensated temporal denoising of transparent image sequences

Participants: Patrick Bouthemy, Vincent Auvray.

Image sequence denoising is a key issue in biomedical imaging, for display as well as for processing and analysis purposes, especially if the acquisition modality supplies particularly noisy images like X-ray imaging. Motion compensation is crucial to allow for effective and accurate temporal noise reduction. We have addressed the intricate issue of motion compensation in transparent image sequences to improve the temporal denoising of X-ray angiographic or cardiac images. Indeed, since the X-rays are successively attenuated by different organs and tissues, the acquired images involve transparency phenomena. A novel hybrid motion-compensated temporal filter is introduced. Global transparent motion compensation allows for a better contrast preservation since it avoids blurring. However, it affects the noise reduction efficiency by increasing the noise of the predicted image, since the latter combines three previous images. We therefore propose to exploit the transparent motion compensation when appropriate only. We distinguish four local configurations related to the intensity homogeneity of the transparent layers forming the image. A filter is designed for each configuration and the overall filter is a data-driven weighted combination of them. The proposed method is able to preserve image contrast while significantly reducing the noise level. Its behaviour has been thoroughly evaluated on realistic simulated X-ray image sequences, and satisfactory results on clinical data have also been obtained.

6.1.3. Object detection and dynamical background estimation

Participants: Charles Kervrann, Patrick Bouthemy, Jérôme Boulanger.

In video-microscopy, the detection of moving objects is usually easier if a background subtraction procedure is applied. However, in the case of images depicting fluorescently tagged particles, the global background image intensity can vary slowly along time. This can be due to several physical phenomena such as photobleaching or diffusion of fluorescent proteins within the cell. Therefore, a stationary model for the background is too restrictive and the moving particles would not be successfully detected. However, we have conducted experiments showing that the intensity variation w.r.t. time can be captured by a parametric (e.g., linear) model for each pixel. This modeling provides a compact representation of a background intensity dynamics. It can be described by maps representing at each time instant the parameters of the parametric model for each spatial position. The parameters for each temporal intensity signal (e.g., for the intensity values along time at a given pixel) are robustly estimated using an asymmetric M-estimator, since vesicles appear as bright (small) blobs in the image. Moreover, the involved parameters are spatially correlated and this must be taken into account in the estimation process. We have introduced spatial coherence by designing a regularized estimation of the parameters. For each pixel, a set of nested space-time tubes is considered. More specifically, the diameter of the tube is increased while the point-wise L_2 risk of the estimator is not minimized (bias-variance trade-off). In contrast to Bayesian methods, our method is local and follows a point-wise adaptive estimation framework. We do not model the objects of interest as usually done, but we only assume that they correspond to bright (small) blobs against the background. This method is able to decompose the sequence into two components: the slowly varying background and moving vesicles.

Furthermore, we have proposed an original approach for automatically detecting moving objects in the reconstructed image for which the background has been removed. For each pixel, a collection of n models able to describe the n -sample signal is competing. The collection is obtained by progressively substituting 0 for observed values. The set of 0 values is assumed to be a satisfying representation of the temporal signal if no vesicles are present. We suppose that the number of 0 in the 1D temporal signal as well as their time positions are unknown. Finally, the “best” model is selected by considering a non-asymptotic penalized likelihood. It is worth noting that this “thresholding” method requires no optimization algorithm to detect the minimum of the objective criterion. This procedure is applied to each pixel in the image.

6.1.4. Motion detection in adverse conditions

Participants: Patrick Pérez, Aurélie Bugeau.

Detecting individual moving objects in videos that are shot by either still or mobile cameras is an old problem, which is routinely solved in a number of real applications such as tele-surveillance. There are, however, a number of interesting instances of this motion analysis problem that are not satisfactorily handled by existing techniques. One class of such problems is the extraction of certain types of moving regions. In the context of activity analysis in dynamically cluttered environments for instance, the problem is the one of separating out foreground moving objects from other moving objects. This might be addressed by characterizing the spatial and/or temporal content of surrounding clutter. This is an acute problem of this sort that we are facing in the Behaviour ACI project (see paragraph 8.2.3) where the detection and the tracking of driver’s moving hands is corrupted by the exterior view through the side car window. Various approaches to this challenging problem can be considered. One line of research, mainly investigated last year, relies on the spatio-temporal analysis of maps of dominant motion compliance (these maps are produced by Motion2D software, see paragraph 5.1), with the aim of extracting and tracking salient blobs of residual motions. Another line, which we are currently investigating, aims at estimating and analyzing sparse motion fields on interest points not belonging to the dominant motion of the car interior. Clustering of these interest points based on multi-dimensional motion-color features can then be performed with non-parametric techniques à la “mean shift”. This poses the generic problem of clustering within multi-dimensional mixed feature spaces, with the critical issue of determining automatically the multi-dimensional kernel. To this end, we have proposed a new variable bandwidth approach based on the so-called “balloon” density estimator. Experiments on the problem of space-range segmentation of static color images allowed us to validate this methodological development. The application of the automatic clustering method thus obtained to the original problem of sparse motion-color clustering is currently under study.

6.1.5. Conditional mixed-state Markov models

Participants: Patrick Bouthemy, Bruno Cernuschi-Frias, Jian-Feng Yao, Tomas Crivelli, Gwénaëlle Piriou.

Digital information present in or extracted from images may be expressed as numerical values or discrete (i.e., abstract label) ones. Beyond this common fact, these two types of variables more deeply reflect two different classes of information: continuous real values (either in one-dimensional or multi-dimensional spaces) *versus* symbolic values (one or several symbols). However, these two classes should not be necessarily viewed as two exclusive or consecutive (e.g., after a decision step) states. Indeed, a physical variable can take both continuous and discrete values, namely it can be a *mixed-state* variable. To give a simple example related to image motion analysis, a locally computed motion quantity can be either null or not. Then, it can be helpful to explicitly consider that it takes either a discrete value expressing the absence of motion, or continuous real values accounting for actual measurements. A discrete value may not be a specific real value only as aforementioned, but a pure symbolic value as well: for instance, for the optic flow case, it could be related to the presence of motion discontinuities while the corresponding continuous values are the velocity vectors. In that context, we have recently introduced so-called *mixed-state auto-models* (i.e., MRF models with two-site cliques), and we have used them for modeling and segmenting motion textures with convincing experimental results. Only generative mixed-state models were considered. New extensions have been defined this year.

First, we have theoretically demonstrated that the mixed-state auto-model can be *factorized* in an appealing way. The global mixed-state generalized pdf factorizes in the sense that, if it is written as a Gibbs distribution, then the global mixed-state potential is the sum, on one hand, of a term involving the discrete part of the variables and of the parametrization, and, on the other hand, of a term involving the continuous part. Only the normalizing factor comprises both parametrization components. The discrete potential is written using indicator functions for the isolated discrete values of the mixed-state model, the other potential being related to a probability density function (exponential family). The parameters corresponding to the discrete and the continuous potentials may be then taken functionally independent. This is a very interesting tool since it permits to independently design the discrete and continuous components, while they still interact in the global resulting modeling and the estimation. This factorization remains true for the *a posteriori* mixed-state model with respect to image observations. Then, it makes the definition of *conditional mixed-state Markov models* accessible. In particular, it means that we can design discriminative or descriptive mixed-state models with the conditional likelihood defined on any extended neighborhoods. We have thus defined a new approach for motion detection with background subtraction which can simultaneously estimate the background image and the binary map of moving objects in the scene. The mixed-state variable accounts for both outputs. The motion detection term can exploit an *a contrario* (discriminative) scheme, and the background estimation term an adaptive non parametric (exemplar-based) method. The regularization mixing two-level discrete values (presence/absence of motion) and real (even if they are quantized over $[0, 255]$) values (background intensity image) is provided by the mixed-state auto-model.

6.1.6. Motion texture modeling and segmentation

Participants: Patrick Bouthemy, Bruno Cernuschi-Frias, Tomas Crivelli, Jian-Feng Yao.

The aim of this work is to model the apparent motion in image sequences depicting natural dynamic scenes (rivers, sea-waves, smoke, fire, grass etc) where some sort of stationarity and homogeneity of motion is present. We adopt the mixed-state Markov Random Fields models recently introduced to represent so-called motion textures. The approach consists in describing the distribution of some motion measurements which exhibit a mixed nature: a discrete component related to absence of motion and a continuous part for measurements different from zero. We have developed several significative extensions to the first version of the model in order to capture more properties of the analyzed motion textures. Considered observations are (locally averaged) scalar normal flow measurements which take values on the whole real line. The use of an 8-nearest-neighbour system in the definition of conditional densities showed to be very efficient in capturing field orientation. We have implemented an 11-parameter mixed-state model where the conditional distribution of the continuous values is assumed to be Gaussian with mean and variance dependent on the neighbours, allowing us to model correlation properties between continuous motion values. Also, we have designed a new way for calculating the partition function in Gibbs distributions.

Then, we have defined a motion texture segmentation method exploiting this modeling. The representation of a motion texture with a relatively small set of parameters allows for a parsimonious characterization of the different parts of a natural dynamic scene. One important aspect that the model should fulfill for dynamic content analysis, is the ability of discrimination. In this context, the problem of segmentation, closely related to classification, aims at determining and locating regions in the image that correspond to a same motion texture or class. As we deal with discrete spatial schemes, the problem is equivalent to assign a label to each point in the image grid, indicating that it belongs to a certain motion texture. One important original aspect is that we do not assume conditional independence of the observations for each texture. The formulation of the segmentation problem we are dealing with, does not assume that the motion texture model parameters are known. Then, it is necessary to correctly estimate the mixed-state motion texture parameters for each class. As an initialization of the label field, we divide the image in square blocks of some fixed size, and for each block a set of motion-texture model parameters is estimated. Then, we apply a clustering technique exploiting the symmetrized Kullback-Leibler between the marginal mixed-state distributions, in order to obtain a first splitting of blocks. The minimization of the global segmentation energy function is performed using the technique of graph cuts [74], [87] to assign labels to points in the image grid. Results on real examples have demonstrated the accuracy and efficiency of our method for the two-class segmentation of motion textures.

6.2. Motion estimation and matching

6.2.1. Fluid motion estimation

Participants: Étienne Mémin, Anne Cuzol, Patrick Heas, Nicolas Papadakis, Guillermo Artana.

Fluid motion estimation is a difficult and important issue in experimental fluid mechanics or in environmental sciences for the study of geophysical flows. In this context, we have investigated different approaches for two important applications: velocity measurement for experimental fluid mechanics and atmospheric wind field estimation from meteorological satellite image sequences. We have addressed the problem of estimating the motion of fluid flows visualized with the Schlieren technique. Such an experimental visualization system is well known in fluid mechanics and it enables the visualization of unseeded flows. Since the resulting images exhibit very low intensity contrasts, classical motion estimation methods based on the brightness constancy assumption (correlation-based approaches, optical flow methods) are inefficient. The global energy function we have defined is composed of *i*) a specific data model accounting for the fact that the observed luminance is related to the gradient of the fluid density, and *ii*) a specific constrained div-curl regularization term. The recovery of the vertical component of fluid motion from a monocular image sequence is another very challenging problem. We have first proposed a dense motion estimator dedicated to the extraction of three-dimensional wind fields for the atmospheric upper layer. In order to derive an estimator for the complete three-dimensional space, we have extended the former to handle a multi-layer model involving a stack of layers interacting via the vertical wind components at the layer boundaries. Both are expressed as the minimization of a global energy function which classically includes a data-driven term and a smoothness term. A robust data-driven term relying on the integrated continuity equation revisited in the context of three-dimensional winds is proposed to fit the sparse cloud observations. The latter is combined with a spatial smoothing term preserving two-dimensional divergent and vorticity structures of the three-dimensional flow and enforcing regions of homogeneous vertical wind components.

We have also explored the problem of estimating mesoscale dynamics of atmospheric layers from satellite image sequences. Due to the intrinsic sparse 3-dimensional nature of clouds and to occluded areas between different cloud layers, the estimation of an accurate dense motion field is an intricate issue. Relying on a physically-sound vertical decomposition of the atmosphere into layers, we have proposed a dense motion estimator dedicated to the extraction of multi-layer horizontal (2D) wind fields. This estimator is expressed as the minimization of a global function including a data-driven term and a spatio-temporal smoothness term. A robust data term relying on shallow-water mass conservation model has been proposed to fit sparse observations related to each layer. A novel spatio-temporal regularizer derived from the shallow-water momentum conservation model has been considered to enforce temporal consistency of the solution along time. These constraints are combined with a robust second-order regularizer preserving divergent and vorticity structures of the flow. In addition, a two-level motion estimation scheme has been settled to overcome the limitations of the multiresolution incremental scheme when capturing the dynamics of fine mesoscale structures. This alternative approach relies on the combination of correlation and optical-flow observations. An exhaustive evaluation of the novel method has been first performed on a scalar image sequence generated by Direct Numerical Simulation of a turbulent bi-dimensional flow. Based on qualitative experimental comparisons, the method has also been assessed on a Meteosat infrared image sequence.

We have also worked on the definition of a low-dimensional fluid motion estimator. This estimator is based on the Helmholtz decomposition which consists in representing the velocity field as the sum of a divergence-free component and a curl-free one. In order to provide a low-dimensional solution, both components have been approximated using a discretization of the vorticity (curl of the velocity vector) and divergence maps through regularized Dirac measures. The resulting so-called irrotational (resp. solenoidal) field is then represented by a linear combination of basis functions obtained by a convolution product of the Green kernel gradient and the vorticity map (resp. the divergence map). The coefficient values and the basis function parameters are obtained by minimizing a function formed by an integrated version of the mass conservation principle of fluid mechanics. This fluid motion estimation method has also been applied to medical imagery in order to estimate the growing of multiple sclerosis lesions. This last study was done in cooperation with P. Hellier (Visages project-team).

6.2.2. *Motion estimation and spatial support segmentation of transparent layers*

Participants: Patrick Bouthemy, Vincent Auvray.

We have developed an original method for the estimation of the motions and the segmentation of the spatial supports of the different layers involved in transparent image sequences. Classical motion estimation methods fail on sequences involving transparent effects since they do not explicitly model this phenomenon. We assume that transparent images can be divided into regions containing at most two transparent layers, which is a reasonable assumption in most real situations. We call this configuration *bi-distributed transparency*. We have considered both medical image sequences (especially X-ray images) and video sequences involving special effects. The proposed method dedicated to image sequences in this configuration comprises three main steps: initial block-matching for two-layer transparent motion estimation, motion clustering with a 3D Hough transform (i.e., for a simplified 3-parameter affine model), and joint transparent layer segmentation and parametric motion (full 6-parameter affine model) estimation. The last step is solved by the iterative minimization of a MRF-based global energy function. The segmentation is improved by a mechanism detecting regions containing one single layer. The overall method is robust to high noise level and to low contrast typical of X-ray imaging. It was validated on various video image sequences and on clinical X-ray image sequences, and satisfactory experimental results have been obtained.

6.2.3. *Repetitive motion and video alignment*

Participants: Ivan Laptev, Patrick Pérez, Émilie Dexter.

Independent motion is a strong cue for detecting, segmenting and recognizing objects and activities in image sequences. Motion-based segmentation, however, is known to be a hard problem in scenes with motion parallax and in scenes with multiple moving objects. Non-parametric segmentation of independently moving objects using color-motion features is one of our research strands in this domain (see paragraph 6.1.4). Another strand concerns the exploitation of not only the presence but also the type of motion as an informative segmentation cue. As a distinctive characteristic, we focused in particular on *repetitiveness*. Motion repetition can take a number of different forms (possibly jointly present in certain applications). It is classically encountered when trying to learn actions and gestures from training videos containing various examples of a targeted action or gesture. Repetition also shows up at the intra-video level when some of the moving entities exhibit periodic, or more generally repetitive, motion patterns. As we demonstrated last year, it is for instance possible to jointly detect and segment periodic motion (e.g., walking persons in complex scenes with non-rigid backgrounds, moving camera and motion parallax) by casting the problem in terms of rough intra-sequence alignment (a sequence is matched to itself over one or more periods). The extension of this framework to more general repetitive motions (no fixed period) and to repetitiveness across videos (alignment of video sequences with similar type of activities, using appropriate local motion descriptors and global geometric constraints) is under study. Another new direction, which we are now taking in the context of the Ph.D. research of Émilie Dexter started in October with a DGA-CNRS funding, is the study of multi-camera video alignment. The ambition here is to address in its full generality the problem of aligning, both in space and time, multiple unsynchronized views of the *same dynamic scene* as provided by uncalibrated moving cameras. To this end, we shall try to rely as much as possible on motion information (whether sparse as with spatio-temporal interest points or interest points trajectories, or dense with optical flows possibly centered on tracked moving objects), while using a limited set of multi-view constraints.

6.3. Tracking

6.3.1. *Stochastic filtering for fluid motion tracking*

Participants: Étienne Mémin, Anne Cuzol, Nicolas Gengembre.

Classical motion estimation techniques usually proceed on pairs of two successive images, and do not enforce temporal consistency. This often induces an estimation drift essentially due to the fact that motion estimation is formulated as a temporal local process. No adequate physical dynamics law, or conservation law, related to the observed flow, is taken into account over long time intervals by the estimation process. Dynamical laws which are usually non linear and modeled up to some uncertainties can be naturally introduced and handled through a stochastic filtering framework. In that spirit, we have proposed a recursive Bayesian filter for tracking velocity fields of fluid flows. The filter combines an Itô diffusion process associated to 2D vorticity-velocity formulation of Navier-Stokes equation and discrete image error reconstruction measurements. In contrast to usual filters designed for visual tracking problem, our filter combines a continuous law for the description of the vorticity evolution with discrete image measurements. We resort to a Monte-Carlo approximation based on particle filtering. The designed tracker provides a robust and consistent estimation of instantaneous motion fields along the whole image sequence. In order to handle a state space of reasonable dimension for the stochastic filtering problem, the motion field is represented as a combination of adapted basis functions. The basis functions are derived from a mollification of Bio-Savart integral and a discretization of the vorticity and divergence maps of the fluid vector field. The output of such a tracking is a set of motion fields along the whole time range of the image sequence. As the time discretization is much finer than the frame rate, the method provides consistent motion interpolation between consecutive frames.

In order to reduce further the dimensionality of the associated state space when we are facing a large number of motion basis functions, we have explored a new dimensional reduction approach based on dynamical systems theory. The study of the stable and unstable directions of the continuous dynamics enables to construct an adaptive dimension reduction procedure. It consists in sampling only in the unstable directions, while the stable ones are treated deterministically. The difficulty is to determine adaptively the stable/unstable directions; one has to solve an eigenvalue problem. The cost of the eigenvalues must be taken into account to evaluate the practical gain brought by the reduction technique. The first results are encouraging, even if there is no computational gain. In fact, we can show that for a dimension reduction by a factor 2, the quality of the estimation obtained with the reduced technique is very similar to the one performed by sampling in the entire space. In order to cope with the tracking problem of complex flows, we are also working in collaboration with the Paris project-team to implement an efficient parallel version of the particle filter. This new version of the algorithm will allow us to handle state spaces of higher dimension and to test the capability of the method to track more complex fluid structures. We have also started to investigate the use of so-called ensemble Kalman filtering for fluid tracking problems. This kind of filters introduced for the analysis of geophysical fluids is based on the Kalman filter update equation. Nevertheless, unlike traditional Kalman filtering setting, the covariances of the estimation errors, required to compute the Kalman gain, rely on an ensemble of forecasts. Such a process gives rise to a Monte Carlo approximation for a family of stochastic non linear filters enabling to handle state space of large dimension.

6.3.2. Robust visual tracking without prior

Participants: Patrick Pérez, Nicolas Gengembre, Vijay Badrinarayanan, Kamel Aouichat.

In this research we are interested in the problem of tracking arbitrary entities along videos of arbitrary type and quality. Such a tracking cannot rely, as classically done, on a priori information regarding both the appearance of the entities of interest (shape, texture, key views, etc.) and their visual motion (kinematical constraints, expected dynamics relative to the camera, etc.). The first crucial step is then the definition and the estimation of the reference appearance model on which the tracking, no matter its precise form, will rely on. Roughly, two extreme types of representations are routinely used in the literature: detailed pixel-wise appearance models subject to rapid fluctuations (e.g, intensity template instantaneously refreshed) and rough color models very persistent over time (e.g., color histogram instantiated at initialization time and kept unchanged). They are both interesting and complementary. For these reasons, it is appealing to fuse them within the same probabilistic tracker. In order to address this fusion problem in a principled way, we are currently investigating, in the context of a Cifre convention with Thomson-RD, a unifying information-theoretic approach to the problem. Our first contribution lies in the unified computation of individual trackers entropies, whether posterior state distribution is approximated by samples (particle filter) or by grid discretization (correlation surface

of standard feature trackers). This permits the self-assessment of each individual tracker and, in the case of a bundle of feature trackers, the design of an automatic mechanism for generating births and deaths of such trackers. The second step, currently under study, will rely on cross-entropy computations to build more robust fused tracker. Another way to improve the robustness of generic visual tracking lies in the analysis of the target surrounding (whether local or not), as we demonstrated along three directions: (1) local color analysis of the surrounding background for improved discrimination, as well as for automatic diagnosis of occlusion and clutter situations; (2) robust dominant motion estimation based on KLT point trackers spread over the image and conditioning of target dynamics on the pan-tilt-zoom parameters thus estimated; (3) intermittent adaptation of reference color models based on outputs from both previous items (reference updating during zooms and in absence of occlusions). These three contributions have been implemented and validated on our tracking platform (see 7.1). Note also that the use of an evolution model conditioned on motion information extracted from the images appears as a particular instance of filtering with conditional Markov chains [13]. This opens up the more general perspective of using conditional Markov random fields for sequential state estimation.

Handling in a principled way a varying (and unknown) number of entities to be tracked is another difficult and important problem in visual tracking, which keeps receiving a great deal of attention from both trajectography and computer vision communities. Our recent contribution to these efforts lies in the definition of a generic particle filter framework where individual trackers (one particle filter per detected object of interest) interact indirectly via modified likelihood functions, which bears connections with JPDAF filters: each marginal filter only “sees” the others via a modification of its likelihood map that prevents it from snapping to neighboring objects of similar aspect while allowing superposition of trackers during occlusions among the targets. This approach offers a very effective alternative to costly trackers that maintain joint multi-object filtering distributions.

Robust generic tracking tools are of major interest for a wide range of applications dealing with editing, analyzing, annotating, browsing and authoring video contents (see paragraph 4.1). Even in applications where a strong prior is “available” (e.g., precise type of videos and/or type of objects of interest), such tools are crucial. On the one hand, they are useful complement to application-specific detection/tracking methods that are often more complex (both in terms of preliminary off-line learning and of on-line application) and sometimes less robust (e.g., false negatives of object-specific detectors). On the other hand, they can facilitate the semi-automatic extraction and labeling of data that application-specific learning modules will require. For these reasons, our contributions on this methodological front are useful for tracking and recognition problems addressed in the FT-RD contract (semi-automatic tracking of players for annotating team sport videos; see paragraph 7.1), and the Behaviour ACI project (fine-grain analysis of car drivers’ activity; see paragraph 8.2.3).

6.3.3. Variational data assimilation for visual tracking

Participants: Étienne Mémin, Nicolas Papadakis, Guillermo Artana.

We have proposed a variational framework for the tracking of features belonging to high dimensional spaces. This framework relies on variational data assimilation principles as developed in environmental sciences to analyse geophysical flows. We have first devised a data assimilation technique for the tracking of closed curves and their associated motion fields. The proposed approach enables a continuous tracking along an image sequence of both a deformable curve and its associated velocity field. Such an approach has been formalized through the minimization of a global spatio-temporal continuous cost functional, with respect to a set of variables representing the curve and its related motion field. The resulting minimization sequence consists in a forward integration of an evolution law followed by a backward integration of an adjoint evolution model. The latter PDE includes a term related to the discrepancy between the state variables evolution law and discrete noisy measurements of the system. The closed curves are represented through implicit surface modeling, whereas the motion is described either by a vector field or through vorticity and divergence maps according to the type of targetted application. The efficiency of the approach has been demonstrated on two types of image sequences showing deformable objects and fluid motions.

Secondly, we have proposed improvements to the construction of low order dynamical systems (LODS) for incompressible turbulent external flows. The reduced model is obtained by means of a Proper Orthogonal Decomposition (POD) basis extracted from experimental data. The POD modes are used to formulate an ordinary differential equations (ODE) system or dynamical system which contains the main features of the flow. This is achieved by applying a Galerkin projection to the Navier-Stokes Equations. Usually, the obtained LODS presents stability problems due to modes truncation and numerical uncertainties, specially when working on experimental data. We performed the model closure with a variational method, Data Assimilation, which refines the state variables within an iterative scheme. This technique allows us to correct the dynamic system coefficients as well as to identify and restore the noisy experimental data used to extract the POD basis. Finally, we have investigated the use of this framework to realize a temporal Bayesian smoothing of fluid flow velocity fields. The velocity measurements are assumed to be supplied by an optical flow estimator or PIV velocity measurements. These noisy measurements are smoothed according to the vorticity-velocity formulation of Navier-Stokes equation. As previously, following optimal control recipes, the associated minimization is conducted through an iterative process involving a forward integration of our dynamical model followed by a backward integration of an adjoint evolution law. Both evolution laws are implemented with a second order non-oscillatory scheme. The approach has been validated on a synthetic sequence of turbulent 2D flow provided by Direct Numerical Simulation (DNS) and on a real meteorological satellite image sequence depicting the evolution of a cyclone.

6.3.4. Image sequence simulation and trafficking analysis

Participants: Charles Kervrann, Patrick Bouthemy, Jérôme Boulanger, Thierry Pécot.

Image sequence analysis in video-microscopy for life sciences has now gained importance since molecular biology is presently having a profound impact on the way research is being conducted in medicine. However, complex interactions between a large number of small particles moving in a cell cannot be easily modeled, which limits the performance of classical object detection and tracking algorithms. This motivated our present research effort which is to develop a general estimation/simulation framework able to produce image sequences showing small moving particles in interaction and undergoing variable velocities, corresponding to intracellular dynamics and trafficking in biology. It is now well established that determining particles trajectories can play a role in the analysis of living cell dynamics, and simulating realistic image sequences is then of major importance. We have proposed a powerful benchmarking method for simulating complex video-microscopy image sequences. We have designed a realistic image sequence modeling framework describing the dynamical and photometric contents of video-microscopy image sequences. Unlike the biophysical approach which aims at describing the underlying physical phenomena, the proposed approach is only based on the analysis of original image sequences and uses graphical models to describe trafficking between source-destination pairs for moving vesicles on microtubule networks. While being quite general, the described method has been designed for analyzing the role of fluorescence-tagged proteins moving around the Golgi apparatus and participating in the intra-cellular traffic. These proteins are embedded into vesicles whose movement is supposed to be dependent on a microtubule network. These vesicles, propelled by motor proteins, move along these polarized “cables”. This mechanism explains the observed high velocities which could not be accounted by basic diffusions. We have demonstrated the potential of the proposed simulation/estimation framework in experiments, and shown that this approach can be also used to evaluate the performance of object detection/tracking algorithms in video-microscopy.

6.4. Dynamic event modeling, learning and recognition

6.4.1. Target behaviour analysis

Participant: Jean-Pierre Le Cadre.

The detection of target maneuvers is far from being a new subject, and has been widely investigated. However, operational requirements led us to consider a specific problem. The probability that a target maneuver be falsely detected, throughout a certain duration, must be upper bounded. This has a fundamental importance, especially for surveillance applications where the observer (e.g., a plane) has to monitor a large geographic area, including a high number of targets in its field of view. Since a few of them are maneuvering, it is necessary to develop a system ensuring both a good maneuver detection and a low level of false target maneuver detection *at the system level*. A comparison of performances of maneuver tests has shown the good performance of a simple test based on consecutive threshold excess. This test is clearly defined at an information processing level, since inputs are the outputs of a test. Robustness of this test is essentially inherited from the level of elementary P_{fa} required at the signal processing level. Thus, it is no longer necessary to use dubious assumptions about the tail distribution, a definite advantage for our problem. The success of this architecture is based on the exploitation of elementary detection trends. In order to analyze the performance of this test, a Discrete Time Markov Modeling (DTMC) is the convenient framework. It is then possible to take benefit of classical DTMC results for investigating the temporal behavior of this test (e.g., occupancy time-distribution, inter-visit times), and thus to adjust parameters (P_{fa}) so as operational requirements be satisfied.

6.4.2. Event detection in video

Participant: Ivan Laptev.

Detection of natural events in unconstrained video data is an unexplored research field. Detection and classification of non-rigid motion such as human actions is an obstacle in many applications including surveillance, content-based video retrieval, film industry and others. Among different types of non-rigid motion we here focus on *events* characterized by the limited spatio-temporal extent, the common structure and the well-defined semantic meaning. Examples of such event classes include short actions and interactions of people such as answering a phone, entering a room, hand shaking, drinking from a glass and smoking a cigarette. Modelling, recognition and detection of such events in common video data is a difficult problem due to the variations of view points, camera motion and illumination as well as the within-class variation of events in motion and appearance. We address the problems of dynamic view point changes within the framework of video alignment (see paragraph 6.2.3). Here, we particularly focus on the problem of within-class variability and build upon machine learning techniques to model the classes of events using annotated training data. One of the concerned issues is the definition of appropriate video features in terms of local spatio-temporal support regions and the combined cues of motion and appearance. To study this problem we initially focused on the related task of object-class detection in still images. We have shown that learning a boosted classifier from a set of local histogram features provides a state of the art method for object detection [59]. The performance of this method has been evaluated within the recent object detection challenge PASCAL VOC 2006. This work is currently being extended to event detection in video. We use boosting methods to select among shape, motion and joint motion-shape features to achieve the reliable and the discriminative cue combination. Using event training data from real movies we are addressing the natural variation of event classes in video.

6.4.3. Video content recognition using trajectories

Participants: Patrick Bouthemy, Jean-Pierre Le Cadre, Alexandre Hervieu.

We are investigating a new approach for dynamic event modeling, classification, and detection in video sequences based on the image trajectories of the moving objects of the scene. Indeed, trajectories form elaborated and meaningful space-time primitives. Besides, recent progress in object tracking in image sequences enable to supply accurate and reliable enough space-time trajectories to represent and model video events. This may be useful for different tasks, such as classification and recognition of events, or detection of abnormal events. To deal with trajectory classification, we first used classical statistical pattern recognition tools such as Support Vector Machine (SVM), or Gaussian Mixture Models (GMM) combined with Linear Discriminant Analysis (LDA) and Maximum Likelihood (ML). We also tried to perform event detection using different known distance metrics such as the Longest Common Subsequence (LCSS) distance. However, in both cases, the developed methods were limited in terms of applications and results, and only usable in restricted conditions (each trajectory captured from the same camera, having the same scale of size, no occlusions, no camera zooming or rotation).

Our aim is to develop novel methods for recognizing activities in video using object trajectories recovered in image sequences for unconstrained situations. A key issue is first to consider measurements exhibiting several invariance properties (such as scale, rotation or translation invariance), and being robust to noise, occlusions, and changes of camera. To achieve this goal, we focus on curvature (κ) and velocity (v) of trajectories, combined in a single measurement $\dot{\gamma}$ for each point of the processed trajectory, where $\dot{\gamma} = \kappa \times v$. Computation of $\dot{\gamma}$ is properly performed using a non-parametric approximation technique (the Nadaraya-Watson kernel method). Each trajectory is then characterized by a vector containing the successive values of $\dot{\gamma}$. We then resort to a statistical model based on Hidden Markov Chains (HMC). States are quantized values $\dot{\gamma}$. We train a HMC for each class of trajectory (using a mixture of splines as observation pdf). We can then compare two HMC using a distance involving mainly the respective sequence of estimated states of the two HMC [92]. It can be used to recognize a dynamic content in a video by comparing the HMC computed for the video segment to be processed and the HMCs learned for the different event classes of interest. We have tested the validity of our method by favourably comparing its performance with the first methods developed both on a set of synthetic examples and on real video sequences.

6.4.4. *Extracting information from a video sensor network*

Participants: Jean-Pierre Le Cadre, Adrien Ickowicz.

Recent trends lead to consider globally networks of video sensors. These networks can be relatively large, so we have to face specific problems. How can we use the data at the sensor level, how to represent the information collected at the sensor level, how to fuse? A first step consists in extracting spatio-temporal informations from video sensors. Of course, these sensors are generally uncalibrated and asynchronous. We have to consider rather rough informations. Roughness can go up to reduce the information to proximity and to a binary information about object motion (e.g., approaching or not). Considering multi-object tracking in this context leads to specific problems. Occultations are frequent while each sensor has only a partial view of the whole scene. Therefore, a large part of your present efforts is oriented toward the basic association problems, a preliminary and critical step for track initialization. A first idea consists in partitioning the observation space via elementary optimization (linear programming and linear regression). The aim of this approach is to drastically limit the combinatorial complexity. For the tracking step, particle filtering is the natural way since it can easily include complex priors, non-linear measurements as well as separation properties, within a hierarchical context.

6.4.5. *Hierarchical surveillance: a solution learning approach*

Participants: Jean-Pierre Le Cadre, Cécile Simonin.

Many real world applications require to optimize a hierarchical problem. These problems are often very hard to solve due to their intrinsic complexity. Indeed, it is difficult to model them so as they could be solved via classic optimization methods: at the upper level, objective functionals can be non-convex or implicitly defined as the result of an optimization algorithm defined at the lower level. A good way to overcome this difficulty is to resort to simulation methods. More precisely, we have developed a hierarchical learning approach based on rare event simulation (namely the Cross Entropy method) at the higher level, while optimization at the lower level is achieved via classical optimization. We have focused on particular two-level hierarchical search problems where the deployment of a set of sensors has to be optimized in order to facilitate the detection or the information gain about targets evolving in various zones of the search space. Various behaviours of the targets have also been considered in this way (one-sided or two-sided). It is noticeable that, despite the complexity of this problem, the algorithm performs very well, with a quite reasonable computation load. It is able to yield a subset of “best” feasible solutions. By this way, constraints are also naturally included.

7. Contracts and Grants with Industry

7.1. ft-rd contract : Probabilistic visual tracking for annotation of team sport broadcasts

Participants: Patrick Pérez, Nicolas Gengembre.

no. Inria 104C1114003, duration : 18 months.

The aim of this contract, started in December 2004, is to design probabilistic tracking tools to help operators annotate television broadcasts of teams sport (with a special emphasis on rugby games). The type of semi-automatic tracking that is targeted is especially challenging due to frequent occlusions, drastic changes of players' appearance within and across shots, joint presence of similar players (teammates) in the image, and diversity of shot types (viewpoint and camera motion). In the last six months of the contract, the platform for robust visual tracking of a single arbitrary object, developed in 2005, has been extended in various directions. First, different background analysis modules have been devised to improve the robustness of probabilistic color-based tracking (see paragraph 6.3.2). Second, an extension to joint multi-player tracking, which is both original and computationally effective, has been implemented. This contract has been concluded in June 2006, with a final report and the delivery of the final code.

7.2. Cifre grant with Thomson-rd

Participants: Patrick Pérez, Vijay Badrinarayanan.

no. Inria, duration 36 months.

This contract started in March 2006 is associated with the supervision of V. Badrinarayanan's thesis funded by a Cifre grant. It concerns the problem of robust tracking of arbitrary objects in arbitrary videos. The first goal is the design of novel probabilistic ingredients to improve the robustness of existing tracking tools, with a first contribution on information-theoretic uncertainty assessment in probabilistic tracking as a generic tool for multiple cue fusion and intermittent adaptation. The second goal concerns the application, and possibly the specialization, of proposed generic techniques to tasks of interest to Thomson. Two scenarios are especially targeted: blurring of selected objects (typically faces) in TV news for business unit Thomson Grass Valley, and object colorization in film post-production for business unit Technicolor. In both cases, robust tracking tools (tracking a bounding box in the first case and the precise object outline in the second case) allowing the partial automatization of painstaking tasks are sought.

7.3. Cifre grant with General Electric Healthcare

Participants: Patrick Bouthemy, Vincent Auvray.

no. Inria 103C19960, duration 36 months.

This contract started in January 2004. The x-ray medical exams present two main modalities, as far as image quality is concerned: high quality record or limited radiation fluoroscopy. When the clinician exploits fluoroscopy (mainly during interventions), image denoising is needed to maintain an acceptable signal-to-noise ratio in the displayed images. Temporal noise reduction filters are inefficient where the organs, the tissues or the devices are moving, since they induce motion blurring. Including a motion-compensation stage before filtering in the time dimension is then required. The goal of the Ph-D thesis work, carried out in partnership with General Electric Healthcare, was to design and evaluate motion estimation methods required for that type of image sequences, and how this information can be efficiently used to denoise. The main difficulty is coming from the particular nature of the X-ray images governed by the principle of superposition, which amounts to motion transparency issues (see paragraph 6.2.2). The Ph-D thesis has been defended in November 2006. The main results are the following. We have focused on the anatomical motion estimation. Two reasonable simplifying assumptions are made, about anatomical motion regularity (leading to the use of parametric motion models) and about the number of transparent layers that can be superimposed in a given spatial region (leading to the introduction of the concept of *bi-distributed transparency*). They enable to reformulate the difficult initial problem in a tractable one. As a result, we propose three transparent motion estimation methods (the more general one involving a joint motion estimation and segmentation process) that are particularly robust to noise. They allow to progressively address the general problem of anatomical transparent motion estimation. Their accuracy was demonstrated with numerous experiments on synthetic data, as well as on a large set of clinical image sequences. In a second stage of our work, we have explored the use of the estimated motions for denoising purposes. The transparency phenomenon enforces the development of a specific transparent motion

compensation procedure. We have studied the properties of the resulting denoising filters, and propose a hybrid method able to bypass some of their limitations (see paragraph 6.1.2).

7.4. cea-Cesta contract: Target tracking in a dense environment

Participants: Jean-Pierre Le Cadre, Adrien Ickowicz.

no. Inria 1542, duration 7 months.

The aim of the multitarget tracking is to associate elementary measurements corresponding to feasible trajectories. This association step is made jointly with a tracking step and both are completely entangled. This means that this problem is largely different from classical target tracking. There is fundamentally uncertainty about the origin of the measurements. To solve such problems, a wide variety of methods are available. Roughly, they can be divided in two categories: the probabilistic methods (e.g., JPDAF, PMHT), and combinatorial ones. However, a major problem remains for initializing multitarget algorithms. While integer programming or flow approaches have been developed for solving it rigorously, a basic tool is to limit the arborescence complexity via merging and pruning (the MHT). In contrast to the “elementary” target tracking framework, there is a strong need for defining convenient tool for the performance of data association. The probability of correct association and the track purity index are sensible tools. In this context, we have shown that a linear regression framework allows us to conduct explicit calculations. More precisely, the probability of correct association has been derived as an explicit function of the scenario parameters: scan number, mean track distance, measurement variance, probability of detection, etc. By this way, it is possible to derive a measure of track-to-track interaction. In a dense target environment, the problem becomes still more complicated since we have to investigate the effects of permutations.

7.5. Grant from Dupont De Nemours

Participant: Etienne Mémin.

University of Rennes contract, duration 24 months

In 2006, Dupont De Nemours company has provided an excellency unrestricted grant to E. Mémin to support his activities in the field of “Developing computational and visualization capabilities to extract object motion fields and fluid flow fields from high-speed imaging”.

7.6. European initiatives

7.6.1. fp6 Fluid project

Participants: Étienne Mémin, Nicolas Gengembre, Patrick Heas, Nicolas Papadakis, Anne Cuzol.

no. Inria 737, duration 36 months

The FLUID project is a FP6 STREPS project labeled in the FET-open program. It started in November 2004. E. Mémin is the scientific coordinator of the project. This 3-year project aims at studying and developing new methods for the estimation, the analysis and the description of complex fluid flows from image sequences. The consortium is composed of five academic partners (Inria, Cemagref, University of Mannheim, University of Las Palmas de Gran Canaria and the LMD, “Laboratoire de Météorologie Dynamique”) and one industrial partner (La Vision) specialized in PIV (Particle Image Velocimetry) systems. The project gathers computer vision scientists, fluid mechanics and meteorologists. The first objective of the project consists in studying novel and efficient methods to estimate and to analyze fluid motions from image sequences. The second objective is to guarantee the applicability of the developed techniques to a large range of experimental fluid visualization applications. To that end, two specific areas are considered: meteorological applications and experimental fluid mechanics for industrial evaluation and control. From the application point of view, the project particularly focuses on 2D and 3D wind field estimation, and on 2D and 3D particle image velocimetry. A reliable structured description of the computed fluid flow velocity field will further allow us to address the tracking of turbulent structures in the flows.

During the second year of this project we have developed methodologies that aim at introducing fluid dynamical model in the motion estimation schemes. This has been done either within a stochastic filtering framework or through variational data assimilation recipes. We have been involved also in the definition of a sound data model that fits at best a fluid flow observed with a given imaging device. We have in particular developed a data model dedicated to the Schlieren imaging technique and to meteorological satellite sequences. In the meteorological context we have investigated the issue of wind fields estimation for a stratified atmosphere into layers. We have studied also solutions to estimate the vertical component of wind fields. A collaboration with the University of Mannheim has also led us to propose mimetic discrete schemes for a high-order regularization term relying on the Helmholtz decomposition of vector fields. This type of scheme allows the discrete solution to fulfill basic properties related to continuous vector fields. Cooperations with the Cemagref Rennes and the LMD (Laboratoire de Météorologie Dynamique) have enabled to assess the relevance of the proposed methods either in the context of experimental fluid mechanics or for meteorological applications.

7.6.2. *fp6 Pegase project*

Participants: Jean-Pierre Le Cadre, Patrick Bouthemy.

no. Inria 1832, duration 36 months

PEGASE is a multidisciplinary European project involving 15 partners (industrial and academic) gathering the major actors of the domain. It is headed by Dassault Aviation. The kick-off-meeting was held in October 2006. For civil aviation, it is widely recognized that approaches, landings and take-offs, or more generally, maneuvers or navigation in the terminal zone, are among the most critical tasks in aircraft operation. PEGASE is a feasibility study of a new navigation system which should allow a three-dimensional truly autonomous approach and guidance for airports and helipads and improves the integrity and accuracy of GNSS differential navigation systems. The purpose of the PEGASE project is to prepare the development of an autonomous, all weather conditions, localization and guidance system based upon correlation between vision sensors output and a ground reference database. The work package WP6 will be led by Inria (Lagadic project-team) and is gathering academic labs (Inria - project-teams Icare, Lagadic and Vista -, CNRS, EPFL, ETHZ, ITJIS) and industrial partners (Dassav, EADS, Euroimage,...). Its aim is to develop new methods in image processing, visual tracking and visual servoing to implement the functionalities required in the PEGASE navid system, in connection with other WPs. Concurrent image processing methods will be implemented and tested on real image sequences, and synthesized image sequences.

7.6.3. *fp6-ist noe Muscle*

Participants: Ivan Laptev, Patrick Bouthemy, Alexandre Hervieu.

no. Inria 104A04950, duration 48 months

The Vista team is involved in the FP6 Network of Excellence MUSCLE (“Multimedia Understanding through Semantics, Computation and Learning”) started in April 2004. It gathers 41 research teams all over Europe from public institutes, universities or research labs of companies. Due to the convergence of several strands of scientific and technological progress, one is witnessing the emergence of unprecedented opportunities for the creation of a knowledge driven society. Indeed, databases are accruing large amounts of complex multimedia documents, networks allow fast and almost ubiquitous access to an abundance of resources and processors have the computational power to perform sophisticated and demanding algorithms. However, progress is hampered by the sheer amount and diversity of the available data. As a consequence, access can only be efficient if based directly on content and semantics, the extraction and indexing of which is only feasible if achieved automatically. MUSCLE aims at creating and supporting a pan-European Network of Excellence to foster close collaboration between research groups in multimedia datamining on one hand and machine learning on the other hand, in order to make breakthrough progress towards different objectives.

I. Laptev participated in the 4th and 5th MUSCLE scientific meetings held in Istanbul in February 2006 and in Paris in December 2006 respectively. He gave talks on “Detection and segmentation of periodic motion” as well as on “Improvements of object detection”. P. Bouthemy and I. Laptev organized a special session entitled “Content Analysis and Representation” within the International Workshop on Multimedia Content Representation, Classification and Security (MRCSS) with three talks given by MUSCLE participants. V. Auvray gave a talk in this session. Vista started collaboration with the Inria project-team Imedia within the e-team “Visual Saliency” of WP5 on the subject of video copy detection. Within this collaboration I. Laptev visited Imedia and stayed one week in Rocquencourt in June 2006. Within the e-team “Kernel Methods” of WP8, Vista initiated collaboration with K. Daoudi (Irit) on action recognition in video. We contributed to several WP reports.

7.6.4. *fp6-ist tn Visiontrain*

Participants: Patrick Bouthemy, Alexandre Hervieu.

no. Inria 850, duration 48 months

Visiontrain is a Marie Curie Research Training Network (belonging to the Computational and Cognitive Vision Systems chapter) which started in May 2005. Visiontrain addresses the problem of understanding vision from both computational and cognitive points of view. The research approach will be based on formal mathematical models and on the thorough experimental validation of these models. In order to achieve these ambitious goals, 11 academic partners plan to work cooperatively on a number of targeted research objectives: (i) computational theories and methods for low-level vision, (ii) motion understanding from image sequences, (iii) learning and recognition of shapes, objects, and categories, (iv) cognitive modelling of the action of seeing, and (v) functional imaging for observing and modelling brain activity. A. Hervieu has attended the first Visiontrain one-week thematic (winter) school held at Les Houches Physics School in March 2006 and devoted to “Optimization methods in computer vision”.

8. Other Grants and Activities

8.1. Regional initiatives

8.1.1. *prir Brittany council program : Imafuid project*

Participants: Étienne Mémin, Anne Cuzol, Patrick Heas, Nicolas Papadakis.

no. Inria 103C18930, duration 36 months.

This project funded by the Brittany council aims within a collaboration with the Cemagref Rennes at developing new methods for the estimation of dense motion fields of fluid flows. The purpose of this project is also to assess the accuracy of several estimation schemes on several known typical experimental flows observed through different image modalities. In this context we are working on methods allowing an effective collaboration of correlation techniques (PIV methods) and variational dense motion estimators. We investigate also the use of dynamical fluid priors to enforce a temporal consistency along time of velocity estimates. As a last issue we are studying how to incorporate within the estimation scheme the effect of small scales of the flow corresponding to unobservable sub-grid spatial resolution.

8.2. National initiatives

8.2.1. *aci Ministry grant : Assimage project*

Participants: Etienne Mémin, Anne Cuzol, Nicolas Papadakis.

no. Inria 103C18930, duration 36 months.

The ASSIMAGE project, labeled within the “Masse de données” ACI program, involves three Inria project-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 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 targeted applications are concerned with predictions of geophysical flows. Our contribution deals with the tracking of vortex structures in collaboration with Cemagref Rennes. Within this project two methods have been developed for vortex tracking. The first one relies on stochastic filtering, whereas the second one is based on variational data assimilation. Both techniques have been applied to meteorological images and to experimental fluid mechanics images. We refer to subsections 6.3.1 and 6.3.3 for further description of the designed methods.

8.2.2. *anr “Jeunes chercheurs” program, stands-msg project*

Participant: Etienne Mémin.

duration 36 months.

This project has been labeled by the ANR program “Jeunes chercheurs” in September 2006, and is entitled “Spatio-temporal analysis of deformable structures in Meteosat Second Generation images”. It aims at developing methods for the analysis of deformable structures in meteorological images. More precisely, within this project we will focus on two meteorological phenomena: the convective cells and the sea breeze circulation. The first type of cloud system is responsible of dangerous meteorological events such as strong showers. Their monitoring is thus very important. Sea breezes influence deeply the climate of coastal regions. The understanding of the daily and seasonal evolution of sea breeze fronts is of great importance for local weather forecasting. The goal of this project will be to propose tools based on appropriate physical evolution laws for the tracking and analysis of these events. This project involves computer vision scientists from different groups, climatologists and meteorologists. The partners are LETG (University of Rennes 2), GREYC (Caen), Inria (project-teams Perception in Grenoble and Vista), LMD (Palaiseau), Lasmea (Clermont-Ferrand).

8.2.3. *aci Ministry grant : Behaviour project*

Participants: Patrick Pérez, Aurélie Bugeau.

no. Inria 104C08130, duration 36 months.

The Behaviour project was granted in October 2004 by the collaborative ACI program on Security and Computer Science. It involves Compiègne University of Technology (Heudiasyc lab) as the prime, along with PSA-Peugeot-Citroën (Innovation and Quality group) and Vista. The main applicative goal is visual monitoring of car drivers, based of videos shot inside the car, such that hypo-vigilant behaviors (mainly drowsiness and distraction) can be detected. To this end, the project aims at providing new tools to perform automatically the recognition of a wide range of elementary behavioral items such as blinks and eye direction, yawn, nape of the neck, posture, head pose, interaction between face and hands, facial actions and expressions, control of the car radio, or mobile phone handling. Before trying to achieve such fine grain activity recognition, one has to select and extract relevant spatio-temporal features to apply subsequent learning on. While UTC is focusing on robust extraction and tracking of facial features in frontal views (shot through the wheel), we are attacking the complementary problem of detection and tracking of mobile items (especially head and hands) in arbitrary driver views. Although the problem seems classic, the specificity of videos under concern makes it very difficult (drastic changes of appearance and prolonged occlusions; low contrast of sequences shot at night; presence of very complex dynamic visual content through window in daylight). In this context, new motion detection, tracking and matching techniques have been studied last year. Further investigation of the first item (detection) has been conducted this year, with novel non-parametric tools for extracting interesting motion region within highly complex dynamical contents (see paragraph 6.1.4). The current focus is on the coupling of this first grid-based extraction step with state-of-art graph-cut techniques for more complete pixel-based extraction of moving regions of interest, while avoiding corruption by outside elements seen through the window.

8.2.4. aci Ministry grant : Modyncell5d project

Participants: Charles Kervrann, Patrick Bouthemey, Jérôme Boulanger.

no. Inria 104C08140, duration 36 months.

This project, labeled within the IMPBio ACI program, was contracted in October 2004. The Vista team is the prime contractor of the project MODYNCELL5D which associates the following other groups: MIA (Mathématiques et Informatique Appliquées) Unit from Inra Jouy-en-Josas, Curie Institute (“Compartimentation et Dynamique Cellulaires” Laboratory, UMR CNRS-144 located in Paris) and UMR CNRS-6026 (“Interactions Cellulaires et Moléculaires” Laboratory - “Structure et Dynamique des Macromolécules” team, University of Rennes 1). This project aims at extracting motion information on proteins dynamics from video-microscopy image sequences, using statistical methods. Methods are developed for two target proteins: CLIP 170 involved in the kinetochores anchorage (in the segregation of chromosomes to daughter cells, the chromosomes appear to be pulled via a so-called kinetochore attached to chromosome centromeres); Rab6a’ involved in the regulation of transport from the Golgi apparatus to the endoplasmic reticulum. Specific algorithms have been developed for processing 3D image sequences for various tasks such as spatio-temporal image denoising and particles detection. The work done this year is described in paragraphs 6.1.1 and 6.1.3.

8.2.5. aci Ministry grant : project related to the development of new fluorescent probes

Participant: Charles Kervrann.

CNRS contract, duration 36 months.

This project, labeled within the DRAB ACI program, was contracted in October 2004. It involves two other teams: UMR-CNRS 6026 (“Interactions Cellulaires et Moléculaires” Laboratory - “Structure et Dynamique des Macromolécules” team, University of Rennes 1) and UMR-CNRS 6510 (“Synthèse et Électrosynthèse Organiques” Laboratory - “Photonique Moléculaire” team, University of Rennes 1). The project aims at characterizing the +TIPs (plus-en tracking proteins) at the extremities “+” of microtubules and their dynamics using new fluorescent probes (Quantum Dots). New image analysis methods are developed for tracking fluorescent molecules linked to microtubules. We have focused on particles detection in images corrupted by Poisson noise.

8.2.6. Other involvements

- The Vista team is involved in the French network GDR ISIS, “Information, Signal and Images”.
- C. Kervrann participates in the network GDR 2588, “Microscopie Fonctionnelle du Vivant”.
- **Collaboration with Ifremer, Brest**

Participant: Charles Kervrann.

C. Kervrann is involved in the supervision of A. Chessel’s thesis (Ifremer Brest) in collaboration with R Fablet (Ifremer). The topic is the analysis of otolith images. Otoliths are small calcareous concretions that can be found in fishes inner ear. They presents rings (as tree trunks) which are representative of the growth of individuals. These images usually have a very low contrast and geometrical information is only partial. A problem is to reconstruct the orientation of the rings. Thus, we tried to find out what kind of interpolation method is sound for extending direction field from partial information. We used an axiomatic approach, similar to the work by Caselles *et al.* [76], [75]. The main novelty is that the data belong to the unit circle. The interpolation must be independent of the parameterization of the circle. We proved that the only second order operator satisfying this requirement is related to the scalar curvature of the level lines of the direction of the reconstructed field. Another interesting operator, which is only invariant with respect to affine parameterization is the so-called ∞ -Laplacian operator defined by $\Delta_\infty u = D^2u(Du, Du)$, for a real valued function u . It is by now a classical result that if u is solution of $\Delta_\infty u = 0$, then u is a local minimizer of $\|Du\|_\infty$, where Du is the gradient of u . Now if u assumes values in \mathbb{R}^2 , we remark that trying to minimize $\|Du\|_\infty$ with the constraint $|u| = 1$ everywhere, amounts to locally solve $\Delta_\infty \theta = 0$ where $\theta = \arg u$ is the direction of u . The interpoted orientation fields are now used either for

extracting subjective contour with low contrast in otolith images via minimal path implemented by a fast marching method, or for selective smoothing of images presenting geometrical properties via Line Integral Convolution.

- **Collaboration with Cesta, Bordeaux**

Participant: Jean-Pierre Le Cadre.

Target acquisition is a common problem for narrow-beam tracking radars. During the target acquisition stage, the radar must operate in a search mode over a limited volume of space. This limited volume corresponds to the prior uncertainty on the target location. Typically, a cued electronic beam scanning radar must seek the target in a 3-dimensional growing error basket. Therefore, the radar needs to fix a sequence of pulses or looks in successive appropriate directions. This sequence, determined over a fixed temporal horizon, should optimize the chances to detect the moving target, once or more times. There are classic acquisition search patterns for agile beam radars, such as rectangular raster scans, fence or ellipsoidal search patterns, which can be dedicated to various operational configurations. However, these semi-empirical patterns do not necessarily provide the best search. Other patterns could offer a higher probability of detection of the target or could require less resource or energy. The only way to determine a search pattern is to study a case where one must allocate integer search efforts into the cells. Then, the search consists of successive cell search moves which depend on the target probability of presence. Generally, B&B (Branch and Bound) methods are well-known exact optimization methods that consist in enumerating cleverly the solution space. Also called implicit enumeration methods, they aim at dividing the solution space in smaller and smaller subsets, most of them being eliminated by bound calculus before being constructed explicitly. In Hohzaki and Iida work, the B&B approach was developed above all in the conditionally deterministic target dynamic case, i.e., when the target dynamic is deterministic given the initial target dynamic state value, i.e., its position, speed, etc. We applied the Hohzaki B&B framework to the target acquisition search pattern issue. It was illustrated by the acquisition of a ballistic target by a narrow-beam sensor. The main assumption is effectively checked: the target dynamic is conditionally (to its ballistic coefficient) deterministic. In this way, optimized search patterns have been obtained for ballistic target acquisition.

- **Collaboration with DGA-CEP**

Participant: Jean-Pierre Le Cadre.

Path planning for navigation in a geographic information system. The problem we considered is the optimization of the navigation of an intelligent mobile in a real world environment, described by a map. The map is composed of features representing natural landmarks in the environment. The vehicle is equipped with sensors which allows it to obtain landmark parameter estimates during the execution. These measurements are correlated with the map so as to estimate the mobile position. The optimal trajectory must be designed in order to control a measure of the performance for the filtering algorithm used for mobile navigation. As the mobile state and the measurements are random, a well-suited measure can be a functional of the Posterior Cramer-Rao Bound. In many application, it is crucial to be able to estimate accurately the state of the mobile during the execution of the plan. Then, it seems necessary to couple the planning and the execution stages. A classical tool is the constrained Markov Decision Process (MDP) framework. However, our optimality criterion is based on the Posterior Cramer-Rao bound, and the nature of the objective function for path planning makes it impossible to perform complete optimization within the MDP framework. Indeed, the reward in one stage of our MDP depends on all the history of the trajectory. To overcome this problem, the Cross-Entropy method, originally used for *rare-events* simulation, is a valuable tool. Its principle is to translate a “classical” optimization method into an *associated stochastic problem* and then to solve it adaptively as the simulation of rare events. This approach has been tested on various (simple) geographic environments and performs satisfactorily. This work is related to the external supervision of F. Celeste’s thesis (DGA-CEP).

8.3. Bilateral international co-operation

8.3.1. Collaboration with University of Buenos-Aires, Inria associate team *fim*

Participants: Guillermo Artana, Bruno Cernuschi-Frias, Tomas Crivelli, Patrick Bouthemy, Étienne Mémin, Jian-Feng Yao, Anne Cuzol, Nicolas Papadakis, Gwénaëlle Piriou.

The Inria associate team FIM (“Fluidos e Imágenes de Movimiento”) is concerned with the analysis of fluid flow from image sequences. It was created in December 2004. This long-term and intensive cooperation involves two groups from the Engineering Faculty of the University of Buenos-Aires: the Signal processing group headed by Professor Bruno Cernuschi-Friàs and the Fluid Mechanics group headed by Professor Guillermo Artana. Two main themes are investigated. The first one deals with experimental visualization and embeds modeling, motion measurement and analysis of fluid flows. The second one is concerned with the modeling, segmentation and recognition of dynamic textures in videos of natural fluid scenes (sea-waves, rivers, smoke, moving foliage, etc...).

Concerning the first topic, we have continued our work on motion estimation in image sequences supplied by a Schlieren device. This imaging set-up enables the visualization of fluid flow without any seeding of particles. The images provided by this technique have poor contrast, and are therefore very difficult to analyze. By devising an adapted data model together with an appropriate smoothing function, we have proposed a first version of an accurate and efficient motion estimator for the Schlieren imaging technique. This first version nevertheless relies on strong assumptions about the derivative of the fluid density. In order to develop an improved estimator, we have worked on weakening these assumptions. Two other data models based on weaker assumptions have been proposed. G. Artana, J. D’Adamo, E. Mémin and N. Papadakis have designed a variational assimilation technique for the coefficient estimation of low-order dynamical systems. Such a system is obtained through a Galerkin projection of the Navier-Stokes equation on a modal representation of the flow under concern using a proper orthogonal decomposition. This decomposition is obtained from a set of noisy velocity measurements supplied by a PIV system. The proposed solution reveals very stable and efficient. It allows us to significantly improve results obtained by a state-of-the-art technique based on polynomial identification. With this new technique, we are able to accurately reconstruct for long time range the trajectories of temporal modes associated with 95% of the kinetic energy of the flow. G. Artana, J. D’Adamo and P. Heas have also worked on the improvement of a dedicated optical flow estimator for particle images of fluid flow. The proposed method combines the advantages of differential motion estimators and correlation-based methods. It avoids the use of multiresolution representation of data and enables in the same time the measurement of high range velocities. Enhanced dense motion fields were obtained on experimental flows.

P. Bouthemy, B. Cernuschi-Frias, T. Crivelli, G. Piriou and J.-F. Yao have been involved in the analysis of motion textures in video depicting natural dynamic scenes. This study is in particular conducted in the context of T. Crivelli’s Ph-D thesis within a “co-tutelle” program between University of Rennes 1 and UBA. Existing probabilistic models for describing motion deal either with continuous or discrete observations. However, real motion data implicitly involve both types of information. Therefore, a model that exploits this mixed-state nature has to be considered. We have continued the study of the so called mixed-state auto-models, for the modeling of dynamic textures in videos of natural dynamic scenes. Several extensions from previous formulations have been proposed. We have implemented an 11-parameter mixed-state model where the conditional distribution of the continuous values is assumed to be Gaussian (with positive or negative values) with mean and variance dependent on the neighbours, allowing us to model correlation properties between continuous motion values. The use of an 8-nearest-neighbour system in the definition of conditional densities showed to be very efficient in capturing field orientation. We have also mathematically shown that the global mixed-state auto-model can be factorized in two terms, one related to the discrete parametrization part and the other one to the continuous parametrization part (while the normalizing factor remains global). It should greatly facilitate the specification of mixed-state models and enlarge their possible use in different applications (see paragraph 6.1.5). A new motion texture segmentation approach was formulated. One important original aspect is that we do not assume conditional independence of the observations for each texture and normalizing factors (partition functions) are properly handled. We have adopted the MAP (Maximum *a posteriori*) criterion

which reduces to maximizing an energy function with respect to a label realization. Different methods were analyzed for this purpose, but in particular methods based on graph-cuts resulted to be the more efficient ones. Convincing accurate segmentation results have been obtained on real video sequences of dynamic natural scenes.

8.3.2. *Advanced Research Program (pra) France-China*

Participants: Jian-Feng Yao, Patrick Bouthemy.

This collaboration with the research group headed by Dr. Véronique Prinet at LIAMA (Sino-French Laboratory for Computer Sciences, Automation and Applied Mathematics, Beijing) is founded jointly by the French Ministry of Foreign Affairs and the Chinese Ministry of Science and Technology. It started in June 2006. It also involves the Ariana project-team (Inria Sophia-Antipolis, X. Descombes). The main objective of the collaborative research program is to build efficient Markov models and algorithms for modeling transformations of geometric structures in satellite image sequences. Potential applications include inspections of urban area or forest modifications from satellite images. V. Prinet spent one week in Rennes in December 2006. J.-F. Yao and X. Descombes spent a one-week stay in Beijing and gave tutorials in July 2006. In that context, P. Bouthemy and J.-F. Yao have organized a two-day workshop entitled “Images and Mathematical Models” in Rennes in December 2006, with the support of this PRA, Irmarm and Irisa. Two tutorials and five invited talks from leading researchers have been given.

8.3.2.1. *Visiting scientists*

- Bruno Cernuschi-Frias (Prof. University of Buenos-Aires) spent one month in our team in the context of the Inria Associate team FIM.
- Short visits by T.Aach (University of Aachen), V. Prinet (Liama, Beijing), C. Schnörr (University of Mannheim), A. Zisserman (University of Oxford).

9. Dissemination

9.1. Leadership within scientific community

- *Editorial boards of journals*
 - J.-P. Le Cadre is Area Editor of Journal of Advances in Information Fusion (ISIF);
 - P. Pérez is Associate Editor for the IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI).
- *Technical program committees of conferences*
 - P. Bouthemy: general co-chairman of ICCV’2007, co-chairman of the TPC of RFIA’2006, TPC member of ECCV’2006, ACIVS’2006, MRCS’2006, OTCBVS’2006, SAMT’2006, CBMI’2007, CIVR’2007, IBPRIA’2007, ICIAP’2007, WIAMIS’2007.
 - C. Kervrann: TPC member of TAIMA’2007, ORASIS’2007.
 - I. Laptev: TPC member of VISAPP’2006.
 - J.-P. Le Cadre: TPC member of Fusion’2006, COGIS’2006.
 - E. Mémin: TPC member of ECCV’2006, RFIA’2006, SSVM’2007.
 - P. Pérez: TPC member of CVPR’2006, ACVIS’2006, Eurographics’2006.
- *Ph.D. reviewing*
 - P. Bouthemy: V. Racine (Institut Curie, Paris).
 - J.-P. Le Cadre: B. Pannetier (Onera, Chatillon), F. Caron (Sequel project-team, Inria Futurs, Lille), O. Cappé (HDR, ENST, Paris).

- E. Mémin: L. Igual (University Pompeu Fabra, Barcelona), W. Reikik (University Paris-6); T. Isambert(Inria project-team Clime, Rocquencourt).
- P. Pérez: F. Pitié (Trinity College, Dublin), G. Delyon (Institut Fresnel, Marseille), A. Jacquot (INPG, Grenoble), P. Lanchantin (INT, Evry).
- *Project reviewing, consultancy, administrative responsibilities*
 - P. Bouthemy participates in the Board Committee of Technovision programme launched by the Ministry of Research and by DGA and aiming at supporting benchmarking projects of image processing and computer vision techniques. He is member of the committee of the EEA prize of the best (French) thesis in signal and image processing. He is member of the Board of the scientific association AFRIF (Association Française pour la Reconnaissance et l'Interprétation des Formes). P. Bouthemy was part of the evaluation committee (CEO, "Comité d'expertise et d'orientation") of the Information Processing and Modelling Department of Onera. He serves as a regular expert for the MRIS ("Mission pour la Recherche et l'Innovation Scientifique") of the French Defense Agency (DGA). He also served as a reviewer for the ANR program call on Multimedia. He is the Inria "main contact" in the preparation of the French-German Quaero program on multimedia indexing and retrieval.
 - P. Bouthemy and J.-P. Le Cadre are deputy members of the committee ("Commission de spécialistes") of the 61th section (Signal Processing and Automation) at University of Rennes 1.
 - C. Kervrann is member of the Scientific Council of the Biometry and Artificial Intelligence Department of Inra. He reviewed a research network project submitted to the High Council for Scientific and Technology Cooperation between France and Israel.
 - E. Mémin is member of the Computer Science committee (27th section, "Commission de spécialistes") at University of Rennes 1.
 - P. Pérez conducted one week of consultancy for Microsoft Research, Cambridge, UK, in March 2006. P. Pérez served as a reviewer for the ANR program call on massive data, and was part of the evaluation committee of Ifremer submarine systems department in Toulon. Until August 2006, P. Pérez was head of the personnel committee ("Commission Personnel") of Irisa, which oversees scientific non-permanent staff hiring at Irisa/Inria-Rennes. Since September 2006, P. Pérez is vice president of the Inria-Rennes project-team committee ("Comité des projets") and deputy member of Inria evaluation board ("Commission d'évaluation"). He is member of the direction team of Irisa/Inria-Rennes ("Équipe de direction") and member of the scientific and technological orientation council (COST, workgroup on large scale initiatives) of Inria. Also, in 2006, he was in charge of the post-doc campaign for Inria-Rennes.
 - J.-F. Yao is member of the executive committee of MAS, a section of the SMAI.

9.2. Teaching

- Master STI "Signal, Telecommunications, Images", University of Rennes 1, (E. Mémin : statistical image analysis, P. Bouthemy : image sequence analysis, J.-P. Le Cadre : distributed tracking, data association, estimation via MCMC methods, C. Kervrann : geometric modeling for shapes and images).
- Master of Computer Science, Ifsic, University of Rennes 1 (E. Mémin : motion analysis; P. Bouthemy : video indexing).
- DIIC INC, Ifsic, University of Rennes 1 (E. Mémin : motion analysis; E. Mémin : Markov models for image analysis; E. Mémin : PDEs for image processing; E. Mémin is in charge of the INC (Digital image analysis and communication) channel.)

- Insa Rennes, Electrical Engineering Dpt (J. Boulanger : computer vision).
- Master PIC and ENSPS Strasbourg, (P. Bouthemy : image sequence analysis).
- ENSAI Rennes, 3rd year (C. Kervrann, P. Pérez : statistical models and image analysis : particle filtering and target tracking).
- Graduate student trainees and interns :
 - B. Combes (Engineering Master in Computer Science, Ifsic, supervised by E. Mémin, work on fluid motion analysis).
 - D. Gayou (Ifsic and ENS Cachan-Bretagne, co-supervised by G. Piriou and P. Bouthemy, work on video indexing).
 - A. Ickowicz (ENSAI and Master Statistics, Rennes 1), supervised by J.-P. Le Cadre, work on tracking).
 - T. Pécot (Master STI, Rennes 1, supervised by C. Kervrann, work on MRF-based detection of local salient elements in images).
- External thesis supervision :
 - F. Celeste (DGA-CEP) supervised by J.-P. Le Cadre;
 - A. Lehuger (FT-RD, Rennes) supervised by P. Pérez;
 - A. Chessel (Ifremer Brest) and I. Bechar (Research Ministry grant - ACI IMPBio, MIA unit of Inra Jouy-en-Josas) co-supervised by C. Kervrann.

9.3. Participation in seminars, invitations, awards

- J. Boulanger received the best student paper prize at the International workshop “Statistical Methods in Multi-Image and Video Processing”, SMVP’2006 held in Graz, May 2006, in conjunction with ECCV’2006, for the paper “Adaptive space-time patch-based method for image sequence restoration” [40]. J. Boulanger gave a talk entitled “Statistical approach for the analysis of video-microscopy image sequences” at the IEEE EMBS Workshop on Microscopy and Medical Image Processing (May 2006) in Linz, Austria.
- P. Bouthemy was co-guest editor (with C. Djeraba and M. Gabbouj) of a special issue of the journal *Multimedia Tools and Applications* on “Content-based Multimedia Indexing” (Part I: Vol.30, Number 3, September 2006, and Part II: Vol.31, Number 1, October 2006).
- P. Bouthemy and I. Laptev were invited to organize a special session entitled “Content Analysis and Representation” within the International Workshop on Multimedia Content Representation, Classification and Security (MRCS’2006) held in Istanbul, September 2006.
- P. Bouthemy and J.-F. Yao have organized a two-day workshop entitled “Images and Mathematical Models” in Rennes in December 2006.
- A. Cuzol has been invited to give lectures at ITU, University of Copenhagen, on stochastic filtering (discrete and continuous time) and its application to continuous tracking in image sequences.
- C. Kervrann and J. Boulanger gave a talk at the Curie Institute seminar on “Méthodes statistiques pour l’analyse de séquences d’images en biologie intra-cellulaire” (Paris, March 2006). C. Kervrann gave a tutorial at the MiFoBio interdisciplinary school on “N-dimensional image analysis: descriptors, and modeling for live cells” (La Grande Motte, September 2006). He was invited speaker at the STAR workshop on “Non-parametric statistics” (Rennes, October 2006). He participates in a invited session at the Annual Meeting of the Institute of Mathematical Statistics (Rio, July 2006) with a contribution entitled “Adaptative estimation to variable smoothness for patch-based image denoising”. He demoed the image analysis tools developed for video-microscopy at the Inria-Industry meeting on Information and Communication Science and Technology for Medicine (Rocquencourt, January 2006).

- I. Laptev took part in the Workshop on “Category-Level Object Recognition” and made an invited poster presentation on “Improvements of object detection” (Siracusa, Italy, September 2006). He visited KTH and gave a seminar on “Object detection” (Stockholm, June 2006). He also visited the Inria project-team Imedia and gave an overview talk on periodic motion segmentation and object detection (Rocquencourt, June 2006).
- J.-P. Le Cadre gave a talk entitled “Analyse de performance en association de données” at the ENSIETA seminar (Brest, October 2006). He gave a presentation at the “Séminaire en navigation” (GDR-ISIS, Paris, February 2006) on “Planning optimal trajectories for navigation”.
- E. Mémin gave an invited talk entitled “Suivi du mouvement fluide: approche stochastique et variationnelle” at the AUM (University activities in Mechanics) conference on “Fluid - structure interactions and turbulent transfer”, held in La Rochelle in September 2006.
- P. Pérez was involved in the First Grets Summer school at Peyresq on Monte Carlo methods for signal and image analysis. He gave a course on Monte Carlo methods for image and video analysis. P. Pérez also gave a tutorial on Sequential Monte Carlo technique and visual tracking in the two-day seminar “Images and Mathematical Models” organized by Irmar and Irisa. P. Pérez gave the following invited talks: “Figure/ground tracking and conditional filters”, Microsoft Research Cambridge seminar (Cambridge, UK, March 2006), “15 years of computer vision: a visual tracking perspective”, IDIAP 15th Anniversary Workshop (Martigny, Switzerland, Sept. 2006); “Figure/ground tracking and conditional filters”, GDR ISIS, action Suivi et Analyse Dynamique (Paris, Nov. 2006); “Tracking, mixtures and particles”, Eurandom Workshop on Image Analysis and Inverse Problems (Eindhoven, the Netherlands, Dec. 2006).
- J.-F. Yao gave a presentation entitled “Multi-parameter auto-models with applications to cooperative systems and motion texture modeling” at the Building Models from Data - Cherry Bud Workshop, in Keio University, March 2006.

10. Bibliography

Major publications by the team in recent years

- [1] P. BOUTHEMY, M. GELGON, F. GANANSIA. *A unified approach to shot change detection and camera motion characterization*, in "IEEE Trans. on Circuits and Systems for Video Technology", vol. 9, n^o 7, October 1999, p. 1030-1044.
- [2] T. CORPETTI, E. MÉMIN, P. PÉREZ. *Dense estimation of fluid flows.*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 24, n^o 3, March 2002, p. 365-380.
- [3] F. DAMBREVILLE, J.-P. LE CADRE. *Detection of a Markovian target with optimization of the search efforts under generalized linear constraints*, in "Naval Research Logistics", vol. 49, n^o 2, February 2002, p. 117-142.
- [4] 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", vol. 25, n^o 12, December 2003, p. 1619-1624.
- [5] C. HUE, J.-P. LE CADRE, P. PÉREZ. *Tracking multiple objects with particle filtering*, in "IEEE Trans. on Aerospace & Electronic Systems", vol. 38, n^o 3, July 2002, p. 791-812.
- [6] C. KERVRANN. *Learning probabilistic deformation models from image sequences*, in "Signal Processing", vol. 71, 1998, p. 155-171.

- [7] P. MUSÉ, F. SUR, F. CAO, Y. GOUSSEAU. *Unsupervised thresholds for shape matching*, in "IEEE Int. Conf. on Image Processing, Barcelona", vol. Barcelona, September 2003.
- [8] E. MÉMIN, P. PÉREZ. *Dense estimation and object-based segmentation of the optical flow with robust techniques*, in "IEEE Trans. on Image Processing", vol. 7, n^o 5, May 1998, p. 703-719.
- [9] J.-M. ODOBEZ, P. BOUTHEMY. *Robust multiresolution estimation of parametric motion models*, in "Journal of Visual Communication and Image Representation", vol. 6, n^o 4, December 1995, p. 348-365.
- [10] P. PÉREZ, A. CHARDIN, J.-M. LAFERTÉ. *Non-iterative manipulation of discrete energy-based models for image analysis*, in "Pattern Recognition", vol. 33, n^o 4, April 2000, p. 573-586.

Year Publications

Doctoral dissertations and Habilitation theses

- [11] V. AUVRAY. *Estimation et segmentation des mouvements transparents dans des séquences d'images fluoroscopiques avec application au débruitage*, Ph. D. Thesis, Université de Rennes 1, Mention Traitement du Signal et des Télécommunications, November 2006.
- [12] A. CUZOL. *Analyse d'écoulements fluides à partir de séquences d'images: estimation de champs de vitesses et suivi temporel par filtrage stochastique non linéaire*, Ph. D. Thesis, Université de Rennes 1, Mention Mathématiques et Applications, December 2006.

Articles in refereed journals and book chapters

- [13] E. ARNAUD, E. MÉMIN. *Partial linear Gaussian models for tracking in image sequences using sequential Monte-Carlo methods*, in "International Journal of Computer Vision", to appear, 2006.
- [14] F. BAVENCOFF, J.-M. VANPEPERSTRAETE, J.-P. LE CADRE. *Constrained bearings-only target motion analysis via Monte Carlo Markov chain methods*, in "IEEE Trans. on Aerospace and Electronic Systems", vol. 42, n^o 4, October 2006.
- [15] J. BOULANGER, C. KERVRANN, P. BOUTHEMY. *An adaptive statistical method for 4D-fluorescence image sequence denoising with spatio-temporal discontinuities preserving*, in "Biophotonics for Life Sciences and Medicine", F. FAUGEL, P. SMIGIELSKI, A. BRANDENBURG, J. FONTAINE (editors)., Fontis Media SA, Lausanne, Switzerland, 2006, p. 97-113.
- [16] J. BOULANGER, K. KERVRANN, P. BOUTHEMY. *Space-time adaptation for patch based image sequence restoration*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", to appear, 2006.
- [17] P. BOUTHEMY, C. HARDOUIN, G. PIRIOU, J.-F. YAO. *Mixed-state auto-models and motion texture modeling*, in "Journal of Mathematical Imaging and Vision", vol. 25, n^o 3, October 2006, p. 387-402.
- [18] T. BRÉHARD, J.-P. LE CADRE. *Closed-form posterior Cramèr-Rao bound for bearings-only tracking*, in "IEEE Trans. on Aerospace and Electronic Systems", vol. 42, n^o 4, October 2006.
- [19] T. BRÉHARD, J.-P. LE CADRE. *Hierarchical particle filter for bearings only tracking*, in "IEEE Trans. on Aerospace and Electronic Systems", to appear, 2006.

- [20] F. CAO, J. DELON, A. DESOLNEUX, P. MUSÉ, F. SUR. *A unified framework for detecting groups and application to shape recognition*, in "Journal of Mathematical Imaging and Vision", to appear, 2006.
- [21] F. CAO, R. FABLET. *Automatic morphological detection of otolith nucleus*, in "Pattern Recognition Letters", vol. 27, n^o 6, 2006, p. 658-666.
- [22] T. CORPETTI, D. HEITZ, G. ARROYO, E. MÉMIN, A. SANTA-CRUZ. *Fluid experimental flow estimation based on an optical-flow scheme*, in "International Journal of Experiments in Fluid", vol. 40, n^o 1, 2006, p. 80-97.
- [23] A. CUZOL, P. HELLIER, E. MÉMIN. *A low dimensional fluid motion estimator*, in "Int. Journal on Computer Vision", to appear, 2006.
- [24] F. DAMBREVILLE, J.-P. LE CADRE. *Constrained minimax optimization of continuous search efforts for the detection of a stationary target*, in "Naval Research Logistics", to appear, 2006.
- [25] P. DODIN, F. MINVIELLE, J. -P. LE CADRE. *A branch-and-bound algorithm applied to optimal radar search pattern*, in "Aerospace Science and Technology", to appear, 2006.
- [26] R. DONATI, J.-P. LE CADRE. *Detection, target motion analysis and track association with a sensor network*, in "Journal of Information Fusion", vol. 7, n^o 3, 2006, p. 285-303.
- [27] C. HUE, J.-P. LE CADRE, P. PÉREZ. *Posterior Cramer-Rao bounds for multi-target tracking*, in "IEEE Trans. on Aerospace and Electronic Systems", vol. 42, n^o 1, January 2006, p. 37-49.
- [28] C. KERVRANN, J. BOULANGER. *Optimal spatial adaptation for patch-based image denoising*, in "IEEE Trans. on Image Processing", vol. 15, n^o 10, 2006, p. 2866- 2878.
- [29] A. KOKARAM, N. REA, R. DAHYOT, M. TEKALP, P. BOUTHEMY, P. GROS, I. SEZAN. *Browsing sports video (Trends in sports-related indexing and retrieval work)*, in "IEEE Signal Processing Magazine", vol. 23, n^o 2, March 2006, p. 47-58.
- [30] J.-P. LE CADRE. *Optimisation des émissions pour la trajectographie et la poursuite de cibles mobiles*, in "Traité IC2: Optimisation en traitement du signal et de l'image", P. SIARRY (editor). , chap. 7, Hermes, 2006.
- [31] P. MUSÉ, F. SUR, F. CAO, Y. GOUSSEAU, J.-M. MOREL. *An a contrario decision method for shape elements recognition*, in "International Journal of Computer Vision", vol. 69, n^o 3, 2006, p. 295-315.
- [32] P. MUSÉ, F. SUR, F. CAO, Y. GOUSSEAU, J.-M. MOREL. *Shape recognition based on an a contrario methodology*, in "Statistics and Analysis of Shapes", H. KRIM, A. YEZZI (editors). , Springer Verlag, 2006.
- [33] G. PIRIOU, P. BOUTHEMY, J.-F. YAO. *Recognition of dynamic video contents with global probabilistic models of visual motion*, in "IEEE Trans. on Image Processing", vol. 15, n^o 11, 2006, p. 3417-3430.
- [34] T. VEIT, F. CAO, P. BOUTHEMY. *An a contrario decision framework for region-based motion detection*, in "International Journal on Computer Vision", vol. 68, n^o 2, 2006, p. 163-178.

- [35] R. VENKATESH BABU, P. PÉREZ, P. BOUTHEMY. *Robust tracking with motion estimation and local Kernel-based color modeling*, in "Image and Vision Computing", to appear, 2006.
- [36] J. YUAN, C. SCHNÖRR, E. MÉMIN. *Discrete orthogonal decomposition and variational fluid flow estimation*, in "Journal of Mathematical Imaging and Vision", to appear, 2006.

Publications in Conferences and Workshops

- [37] E. ARNAUD, E. MÉMIN, R. SOSA, G. ARTANA. *A fluid motion estimator for Schlieren imaging velocimetry*, in "Proc. European Conf. Comp. Vision (ECCV'06), Graz, Austria", May 2006.
- [38] V. AUVRAY, P. BOUTHEMY, J. LIÉNARD. *Motion estimation in X-Ray image sequence with bi-distributed transparency*, in "Proc. Int. Conf. on Image Processing (ICIP'06), Atlanta, USA", October 2006.
- [39] V. AUVRAY, P. BOUTHEMY, J. LIÉNARD. *Motion-based segmentation of transparent layers in video sequences*, in "Proc. Int. Workshop on Multimedia Content Representation, Classification and Security, MRCS'2006, Istanbul, Turkey", September 2006.
- [40] J. BOULANGER, C. KERVRANN, P. BOUTHEMY. *Adaptive space-time patch-based method for image sequence restoration*, in "Proc. Int. Workshop on Statistical Methods in Multi-Image and Video Processing (SMVP'06), Graz, Austria", May 2006.
- [41] J. BOULANGER, C. KERVRANN, P. BOUTHEMY. *Estimation of dynamic background for fluorescence video-microscopy*, in "Proc. IEEE Int. Conf. on Image Processing (ICIP'06), Atlanta, USA", October 2006.
- [42] J. BOULANGER, C. KERVRANN, P. BOUTHEMY. *Simulation and estimation of fluorescence in microscopy image sequence*, in "Proc. MICCAI'06 workshop on microscopique image analysis and application to biology, Copenhagen, Denmark", October 2006.
- [43] T. BRÉHARD, J.-P. LE CADRE. *Closed-form posterior Cramér-Rao bound for active measurement scheduling*, in "Proc. Conf. on Cognitive Systems with Interactive Sensors (COGIS'06), Paris, France", March 2006.
- [44] T. BRÉHARD, J.-P. LE CADRE. *Closed-form posterior Cramér-Rao bound for a maneuvering target in the bearings-only tracking context using best-fitting Gaussian distribution*, in "Proc. Conf. on Information Fusion (Fusion'06), Florence, Italy", July 2006.
- [45] F. CAO, P. BOUTHEMY. *A general principled method for image similarity validation*, in "Proc. Int. Workshop on Adaptive Multimedia Retrieval (AMR'2006), Geneva, Switzerland", July 2006.
- [46] A. CHESSEL, F. CAO, R. FABLET. *Interpolating orientation fields: an axiomatic approach*, in "Proc. European Conf. Comp. Vision (ECCV'06), Graz, Austria", May 2006.
- [47] A. CHESSEL, R. FABLET, F. CAO, C. KERVRANN. *Orientation interpolation and applications*, in "Proc. IEEE Int. Conf. on Image Processing (ICIP'06), Atlanta, USA", October 2006.
- [48] T. CRIVELLI, B. CERNUSCHI-FRIAS, P. BOUTHEMY, J.-F. YAO. *Mixed-state Markov random fields for motion texture modeling and segmentation*, in "Proc. IEEE Int. Conf. on Image Processing (ICIP'06), Atlanta, USA", October 2006.

- [49] T. CRIVELLI, B. CERNUSCHI-FRIAS, P. BOUTHEMY, J.-F. YAO. *Segmentation of motion textures using mixed-state Markov random fields*, in "Proc. Conf. SPIE Math.of Data/Image Pattern Recognition, Compression, and Encryption with Applications, San Diego, USA", August 2006.
- [50] A. CUZOL, E. MÉMIN. *Suivi de mouvement fluide par filtrage stochastique*, in "Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA'06), Tours", January 2006.
- [51] F. CÉLESTE, F. DAMBREVILLE, J.-P. LE CADRE. *Optimal path planning using the cross entropy method*, in "Proc. Conf. on Information Fusion (Fusion'06), Florence, Italy", July 2006.
- [52] F. CÉLESTE, F. DAMBREVILLE, J.-P. LE CADRE. *Optimal strategies for mobile robots based on the Cross entropy Method*, in "Machine Learning for Signal Processing (MLSP'06), Maynooth, Ireland", September 2006.
- [53] P. DODIN, P. MINVIELLE, J.-P. LE CADRE. *A branch-and-bound algorithm applied to optimal radar search pattern*, in "Proc. Conf on Cognitive Systems with Interactive Sensors (COGIS'06), Paris, France", March 2006.
- [54] R. FABLET, S. PUJOLLE, A. CHESSEL, A. BENZINOU, F. CAO. *Variational level-set reconstruction of accretionary morphogenesis from images*, in "Proc. IEEE Int. Conf. on Image Processing (ICIP'06), Atlanta, USA", October 2006.
- [55] P. HEAS, E. MEMIN, N. PAPADAKIS. *Dense estimation of layer motions in the atmosphere*, in "Int. Conf. Pattern Recognition (ICPR'06), Hong-Kong, China", August 2006.
- [56] C. KERVRANN, J. BOULANGER. *Adaptation locale optimale pour la régularisation et la représentation d'image à base de motifs locaux*, in "Proc. Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA'06), Tours", January 2006.
- [57] C. KERVRANN, J. BOULANGER. *Adaptive estimation to variable smoothness for patch-based image denoising*, in "Proc. Annual Meeting of the Institute of Mathematical Statistics, Rio, Brazil", August 2006.
- [58] C. KERVRANN, J. BOULANGER. *Unsupervised patch-based image regularization and representation*, in "Proc. European Conf. Comp. Vision (ECCV'06), Graz, Austria", May 2006.
- [59] I. LAPTEV. *Improvements of object detection using boosted histograms*, in "Proc. British Machine Vision Conf. (BMVC'06), Edinburgh, U.-K.", September 2006.
- [60] A. LEHMANN, P. BOUTHEMY, J.-F. YAO. *Rank-test similarity measure between video segments for local descriptors*, in "Proc. Int. Workshop on Adaptive Multimedia Retrieval (AMR'2006), Geneva, Switzerland", July 2006.
- [61] A. LEHUGER, P. LECHAT, P. PÉREZ. *An adaptive mixture color model for robust visual tracking*, in "Proc. Int. Conf. on Image Processing (ICIP'06), Atlanta, USA", October 2006.
- [62] A. LEHUGER, P. LECHAT, P. PÉREZ. *An adaptive mixture color model for robust visual tracking*, in "Proc. Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA'06), Tours", January 2006.

- [63] N. PAPADAKIS, A. CUZOL, E. MÉMIN. *Suivi du mouvement fluide: approche stochastique et variationnelle*, in "Proc. Journées AUM/AFM, La Rochelle", September 2006, p. 117–122.
- [64] N. PAPADAKIS, E. MÉMIN. *Une approche variationnelle pour le suivi de contours*, in "Proc. Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA'06), Tours", January 2006.
- [65] V. RACINE, A. HERTZOG, J. JOUANNEAU, J. SALAMERO, C. KERVRANN, J. B. SIBARITA. *Multiple target tracking of 3D fluorescent objects based on simulated annealing*, in "Proc. Int. Symp. on Biomedical Imaging (ISBI'06), Arlington, USA", April 2006.
- [66] C. SIMONIN, J.-P. LE CADRE, F. DAMBREVILLE. *A solution-learning method for planning a hierarchical surveillance*, in "Proc. Conf. Rare Event Simulation (RESIM'06), Bamberg, Germany", Sept. 2006.
- [67] A. SZANTAI, P. HEAS, E. MÉMIN. *Comparison of atmospheric motion vectors and dense vector fields calculated from MSG images*, in "Proc. Int. Winds Workshop, Beijing, China", 2006.
- [68] J.-M. VANPEPERSTRAETE, J.-P. LE CADRE. *Robust detection of target maneuvers*, in "Proc. Conf. on Information Fusion (Fusion'06), Florence, Italy", July 2006.
- [69] R. VENKATESH BABU, P. PÉREZ, P. BOUTHEMY. *Kernel-based robust tracking for objects undergoing occlusion*, in "Proc. Asian Conf. on Computer Vision (ACCV'06), Hyderabad, India", January 2006.

Internal Reports

- [70] J. BOULANGER, C. KERVRANN, P. BOUTHEMY. *Space-time adaptation for patch-based image sequence restoration*, Technical report, n^o 1798, IRISA, Rennes, France, April 2006, <http://hal.inria.fr/inria-00071294>.
- [71] T. VEIT, F. CAO, P. BOUTHEMY. *An a contrario space-time grouping framework for the detection of coherent motions*, Research Report, n^o 6061, INRIA, 11 2006, <https://hal.inria.fr/inria-00115435>.

Miscellaneous

- [72] C. DJERABA, M. GABBOUJ, P. BOUTHEMY. *Multimedia indexing and retrieval: ever great challenges (guest editorial)*, vol. 30, n^o 3, 2006, Multimedia Tools and Applications, special issue.

References in notes

- [73] P. BOUTHEMY, E. FRANÇOIS. *Motion segmentation and qualitative dynamic scene analysis from an image sequence*, in "Int. Journal of Computer Vision", vol. 10, n^o 2, April 1993, p. 157-182.
- [74] Y. BOYKOV, O. VEKSLER, R. ZABIH. *Efficient approximate energy minimization via graph cuts*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 20, n^o 12, November 2001, p. 1222-1239.
- [75] V. CASELLES, L. IGUAL, O. SANDER. *An axiomatic approach to scalar data interpolation on surfaces*, in "Nümerische Matematik", 2006.
- [76] V. CASELLES, J.-M. MOREL, C. SBERT. *An axiomatic approach to image interpolation*, in "IEEE Trans. on Image Processing", vol. 7, n^o 3, March 1998, p. 376-386.

-
- [77] T. CORPETTI, E. MÉMIN, P. PÉREZ. *Dense estimation of fluid flows.*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 24, n^o 3, March 2002, p. 365-380.
- [78] T. CORPETTI, E. MÉMIN, P. PÉREZ. *Extraction of singular points from dense motion fields: an analytic approach*, in "Journal of Mathematical Imaging and Vision", vol. 19, n^o 3, 2003, p. 175-198.
- [79] A. DESOLNEUX, L. MOISAN, J.-M. MOREL. *A grouping principle and four applications*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 25, n^o 4, 2003, p. 508-513.
- [80] A. DOUCET, S. GODSILL, C. ANDRIEU. *On sequential Monte Carlo sampling methods for Bayesian filtering*, in "Statistics and Computing", vol. 10, n^o 3, 2000, p. 197-208.
- [81] R. FABLET, P. BOUTHEMY. *Non-parametric scene activity analysis for statistical retrieval with partial query*, in "Journal of Mathematical Imaging and Vision", vol. 14, n^o 3, May 2001, p. 257-270.
- [82] M. GELGON, P. BOUTHEMY. *A region-level motion-based graph representation and labeling for tracking a spatial image partition*, in "Pattern Recognition", vol. 33, n^o 4, April 2000, p. 725-745.
- [83] F. HEITZ, P. BOUTHEMY. *Multimodal estimation of discontinuous optical flow using Markov random fields*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 15, n^o 12, December 1993, p. 1217-1232.
- [84] F. HEITZ, P. PÉREZ, P. BOUTHEMY. *Multiscale minimization of global energy functions in some visual recovery problems*, in "CVGIP : Image Understanding", vol. 59, n^o 1, 1994, p. 125-134.
- [85] M. ISARD, A. BLAKE. *CONDENSATION—Conditional Density Propagation for Visual Tracking*, in "Int. Journal of Computer Vision", vol. 29, n^o 1, 1998, p. 5-28.
- [86] G. KITAGAWA. *Monte Carlo filter and smoother for non-Gaussian nonlinear state space models*, in "J. Computational and Graphical Stat.", vol. 5, n^o 1, 1996, p. 1-25.
- [87] V. KOLMOGOROV, R. ZABIH. *What energy functions can be minimized via graph cuts?*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 26, n^o 2, February 2004, p. 147-159.
- [88] E. MÉMIN, P. PÉREZ. *Hierarchical estimation and segmentation of dense motion fields*, in "Int. Journal of Computer Vision", vol. 46, n^o 2, February 2002, p. 129-155.
- [89] E. MÉMIN, P. PÉREZ. *Dense estimation and object-based segmentation of the optical flow with robust techniques*, in "IEEE Trans. on Image Processing", vol. 7, n^o 5, May 1998, p. 703-719.
- [90] J.-M. ODOBEZ, P. BOUTHEMY. *Robust multiresolution estimation of parametric motion models*, in "Journal of Visual Communication and Image Representation", vol. 6, n^o 4, December 1995, p. 348-365.
- [91] J.-M. ODOBEZ, P. BOUTHEMY. *Direct incremental model-based image motion segmentation for video analysis*, in "Signal Processing", vol. 6, n^o 2, 1998, p. 143-155.
- [92] L. RABINER. *A tutorial on hidden Markov models and selected applications in speech recognition*, in "Proceedings of IEEE", vol. 77, n^o 2, February 1989, p. 257-286.