

Rapport INRIA 1994 — Programme 4
Traitement, exploitation et modélisation
d'images séquentielles

Projet TEMIS

3 mai 1995

Projet TEMIS

Traitement, exploitation et modélisation d'images séquentielles

Localisation : *Rennes*

Mots-clés : analyse du mouvement (5, 22), asservissement visuel (17), classification (1, 22), codage (1, 22), estimation paramétrique (5), modèle markovien (5), modélisation statistique (1), mouvement (1), multigrille (5), multimédia (22), parallélisme (5, 22), robotique (1, 17), séquence d'images (1, 5, 17, 22), stratégie de perception (17), télésurveillance (1), télévision (1, 22), TEMIS (1), vision 3D (17), vision active (17), vision par ordinateur (5, 17).

Temis est un projet commun Inria/CNRS (URA 227).

1 Composition de l'équipe

Responsable scientifique

Claude Labit, DR Inria

Secrétaire

Édith Blin-Guyot, TR Inria

Isabelle Mescam, en remplacement d'Édith Blin-Guyot, d'avril
à novembre

Personnel Inria

Patrick Bouthemy, DR

François Chaumette, CR
Fabrice Heitz, CR, jusqu'au 31 août 1994
Patrick Pérez, CR, à partir du 15 septembre 1994
Noëlle Adam, TR (Atelier)
Fabien Spindler, IR (Atelier), à partir du 1er novembre 1994

Personnel Ura 227

Luce Morin, maître de conférences, université de Rennes 1
Ahmed Tamtaoui, Ater, université de Rennes 1, jusqu'au 31 août 1994
Thierry Fauconnier, IR CNRS (Atelier)

Chercheurs post-doctorant

Jean-Pierre Leduc, chercheur post-doc HCM
Demin Wang, chercheur invité à partir du 1er octobre 1994

Chercheurs doctorant

Lila Abdessemed, Ater, IUT de Vannes
Laurent Bonnaud, bourse Inria-Région
Farabi Bensalah, bourse Inria-Région
Nicolas De Maistre, bourse Inria-Région, à partir du 1er octobre 1994
Victor Garcia-Garduño, bourse CIES
Charles Kervrann, bourse MESR
Jean-Marc Laferté, bourse MESR
Hervé Lefloch, bourse MESR, à partir du 1er octobre 1994
Eric Marchand, bourse MESR
Marianne Maurizot, bourse MESR, à partir du 1er octobre 1994
Eric Nguyen, bourse MESR
Victor Nzomigni, bourse Inria-Cies
Jean-Marc Odobez, bourse Inria-Région
Vincent Ricordel, bourse MESR
Yann Ricquebourg, bourse BDI CNRS-Région

Collaborateurs extérieurs

Brigitte Escofier, professeur, IUT Vannes
Etienne Mémin, Ater puis maître de conférences, IUT Vannes
Veronique Haese-Coat, maître de conférences, Insa
Joseph Ronsin, professeur, Insa
Fabrice Heitz, professeur, ENSPS Strasbourg, à partir du 1er septembre 1994

In memoriam

Brigitte Escofier nous a quitté le 15 juillet 1994 après une longue et éprouvante maladie. La statistique française perd un chercheur et un professeur d'un rayonnement et d'une créativité exceptionnelles. Professeur à Rennes 1 puis à Rennes 2, Brigitte Escofier a créé et dirigé depuis 1988 le Laboratoire d'Application des Méthodes Statistiques et Informatiques, le LAMSI, qui regroupe des enseignants chercheurs de l'IUT de Vannes. Les travaux scientifiques de Brigitte Escofier ont surtout porté sur l'analyse des correspondances, ses propriétés, ses extensions. Dès 1965, sa thèse (l'Analyse factorielle des correspondances, faculté des Sciences de Rennes, publiée en 1969 dans les cahiers du B.U.R.O, n°13) est, avec le cours de linguistique mathématique à l'université de Rennes de Jean-Paul Benzecri qui dirigea ses premiers travaux, la référence de base sur cette méthode. Tous les thèmes de recherche importants de ce que l'on devait parfois appeler par la suite l'*analyse des données à la française* ont par la suite reçu des contributions originales de Brigitte Escofier. Citons dans l'ordre chronologique, mais sans être exhaustif :

- les problèmes de stabilité des résultats des analyses vis-à-vis des perturbations du tableau de données ou d'agrégation de lignes ou de colonnes de ce tableau (cf. *Etudes de trois problèmes de stabilité en analyse factorielle. Pub. de l'ISUP, vol. 11, 1972 avec B. Le Roux, Stabilité et Approximation en analyse factorielle*, thèse d'Etat soutenue à Paris 6 en 1979) ;
- la notion d'équivalence distributionnelle et les diverses métriques qui satisfont cette propriété dans *La Revue de Statistique Appliquée, vol 26, 1978* ;
- la représentation des variables (et non seulement des modalités) en analyse des correspondances multiples (*Revue de Stat. App. vol 27, 1979*) ;
- la généralisation de cette dernière méthode d'analyse au mélange de variables quantitatives et qualitatives dans *les cahiers de l'Analyse des Données, n°2, 1979* ;
- les analyses factorielles multiples dans des éditions successives d'un ouvrage rédigé avec Jérôme Pagès, *Analyses factorielles simples et multiples publié chez Dunod, 1988* ;

- diverses généralisations de l'analyse des correspondances permettant d'analyser les tableaux évolutifs ou les tableaux d'échange (*les Cahiers d'Anal. de Don.*, n° 3 et 4, 1983 avec D Drouet) ;
- les méthodes d'analyse conditionnelle prenant en compte des informations a priori et dans un cadre similaire les analyses intégrant les notions de contiguïté, méthodes utiles notamment en traitement d'images dans *Statistique et Analyse des Données*, 15, n° 3, 1991, thèse de H. Benali et tout récemment (1994) de L. Abdessemed

Ces quelques références ne constituent que des jalons destinés à donner une idée des domaines couverts. Brigitte Escofier avait avec enthousiasme, ces dernières années, au sein du projet Temis, entrepris d'appliquer ses outils fondamentaux d'analyse de données au problème de la classification contextuelle d'images texturées. Sa curiosité scientifique et sa passion intellectuelle resteront comme un bel exemple d'une recherche réellement interdisciplinaire.

2 Présentation générale et objectifs du projet

Les grandes orientations scientifiques du projet sont les suivantes :

- la prise en compte systématique de la dimension temporelle dans les algorithmes étudiés de traitement de séquences d'images, tant pour le thème de vision dynamique et de vision active que celui de compression d'images ;
- l'emploi d'un cadre mathématique unifié de type statistique pour les diverses étapes algorithmiques en analyse du mouvement (détection, estimation, segmentation, interprétation, déformation) ; les outils et modèles manipulés en analyse et compression tendent de plus en plus à se rejoindre, notamment concernant les schémas de compression orientée objet et l'indexation vidéo ;
- le traitement de séquences longues d'images par l'introduction de méthodes de suivi temporel de primitives et de l'identification de trajectoires ;
- la définition d'une modélisation markovienne et de sa technique d'optimisation associée nécessairement multiéchelle afin d'améliorer l'efficacité et la convergence des algorithmes d'analyse d'image ;

- l'exploration du thème de la vision active selon une approche en boucle fermée de commande référencée vision ;
- la liaison algorithme-architecture ; en relation avec le projet API, nous poursuivons la recherche de parallélisation de classes génériques d'algorithmes (relaxation, estimation de mouvement, ...)
- enfin, le thème de la focalisation d'attention et de traitement est devenu -notamment pour les applications en télésurveillance active- un thème fédérateur pour notre projet et un champ d'investigation nouveau associant les aspects de perception, d'analyse, de communication et de commande.

L'année 1994 a vu, sur la base de ces orientations scientifiques, se concrétiser des résultats particulièrement intéressants sur les sujets suivants :

- Projet Vidac : lancement, au sein d'une action de contrat de plan Etat-Région, d'un cadre fédérateur d'études regroupant nos recherches en analyse, communication et contrôle à partir de l'observation de phénomènes dynamiques (séquences d'images) ;
- Vision dynamique : développement de techniques d'estimation robuste de mouvement 2D et extension aux cas de mouvements déformables ;
- Vision active : stratégies optimales de perception et de reconstruction 3D d'une scène par vision active ;
- Compression de séquences d'images : spécification d'un cadre générique de compression orientée objet avec techniques adaptatives de focalisation.

3 Actions de recherche

3.1 Analyse statistique de l'image et Vision dynamique

L'analyse d'une scène à partir d'un capteur vidéo est qualifiée de dynamique lorsque le mouvement en constitue l'un des éléments clé, qu'il en soit l'objet de mesure même ou le moyen d'accéder à une information donnée. On parle également de *vision dynamique*. L'équipe possède sur ce sujet un acquis important.

Notre approche de la vision dynamique se singularise d'une part par une formulation du problème dans l'espace même de la séquence d'images. Les avantages suivants en découlent : exploitation complète du "signal" d'image (au sens large), richesse et puissance des outils de modélisation et d'estimation disponibles dans ce cadre, économie des calculs de structures de représentation 3D. Cela s'est traduit par un intérêt particulier porté à l'analyse du mouvement 2D apparent.

L'analyse de scène dynamique gagne d'autre part à s'appuyer sur une phase préalable de modélisation statistique lorsqu'il s'agit de concevoir des méthodes générales et robustes. Nous avons ainsi obtenu des avancées significatives dans le traitement des problèmes se posant en vision dynamique (détection, mesure et segmentation du mouvement, suivi temporel, analyse des déformations) par le développement d'approches statistiques s'appuyant sur des modèles markoviens 2D associés à une estimation bayésienne. Ces travaux ont d'ailleurs une portée générale, pour ce qui est notamment des modèles markoviens multi-échelles que nous avons définis, ou des schémas de parallélisation d'algorithmes de relaxation que nous avons conçus puis validés sur machines parallèles. Nous avons aussi récemment développé une technique d'estimation robuste pour l'identification de modèles paramétrés de mouvement.

3.1.1 Modélisation markovienne et analyse statistique hiérarchique de l'image

Participants : Fabrice Heitz, Jean-Marc Laferté, Eric Fabre, Patrick Pérez

Le besoin de hiérarchiser l'information spatio-temporelle nous a conduits à développer ces dernières années des versions multi-échelles de modèles markoviens non-causaux, permettant d'obtenir des gains importants, tant au niveau de la vitesse de convergence des algorithmes que de la qualité des solutions finales.

L'étude de représentations statistiques hiérarchiques a pris cette année une nouvelle orientation, à travers le développement et l'évaluation, dans le cadre de l'analyse multi-image, de modèles markoviens *causaux* définis sur des graphes hiérarchiques (quad-arbres ou graphes pyramidaux). Cette étude a été menée avec la collaboration d'É. Fabre du projet AS. Les avantages de ces nouveaux modèles sont multiples. Tout d'abord, ils offrent la possibilité de représenter de nouvelles classes de processus

aléatoires et ainsi, de modéliser de nouveaux types d'interactions entre pixels de l'image. Ensuite, contrairement aux champs markoviens non-causaux, ces représentations conduisent à des algorithmes non-itératifs qui allient rapidité et optimalité. Enfin la structuration hiérarchique qu'ils induisent sur l'information-image les rend particulièrement attractifs lorsqu'il s'agit de fusionner des données hétérogènes provenant de capteurs différents.

Une évaluation approfondie de ces modèles a donc été entreprise, dans le cadre d'une application de fusion d'images multicapteurs-multirésolutions, en collaboration avec l'Inra et l'Ensar. Un des objectifs de ce projet est de modéliser les dépendances statistiques inter-échelles entre images satellitaires, images aériennes et mesures obtenues par des capteurs disposés au sol, dans le cadre d'études menées par l'Inra en bioclimatologie.

Une première phase de l'étude a porté sur :

- la définition de modèles hiérarchiques sur des structures de graphe simples (quad-arbres, graphes pyramidaux) ;
- le développement d'algorithmes non-itératifs (de type Viterbi) pour la mise en œuvre de ces modèles ;
- la mise en place, en parallèle, d'algorithmes d'estimation de leurs paramètres (correspondant à des versions hiérarchiques des algorithmes EM et SEM).

La validation a été menée, dans un premier temps, dans un problème de classification d'images multirésolutions et multicapteurs en imagerie satellitaire, sur données synthétiques et réelles. Cette première validation a clairement mis en évidence les intérêts de ce type d'approche : bonne capacité de représentation de l'information-image, robustesse au bruit apportée par la fusion et rapidité de traitement même lorsque des techniques statistiques d'identification des paramètres (réputées pour leur lenteur dans le cas des modèles non-causaux) sont introduites.

3.1.2 Algorithmes et architectures pour la parallélisation des algorithmes de relaxation

Participants : Fabrice Heitz, Etienne Mémin

L'étude de la mise en œuvre parallèle des algorithmes de relaxation multi-échelle s'est poursuivie cette année par le développement d'algorithmes parallèles pour les modèles faisant intervenir des variables continues dans

la représentation de l'image. Ces modèles (modèles gaussiens et leurs déclinaisons non-linéaires) sont particulièrement attractifs pour la mesure du mouvement, comme l'avaient montré des études précédentes.

L'effort a donc porté sur le développement de modèles multiéchelles continus de mesure du mouvement dans une séquence d'images, intégrant une gestion implicite des discontinuités par le biais de potentiels non-linéaires (quadratiques tronquées, Φ fonctions, ...). Les modèles qui ont été retenus permettent d'obtenir une mesure multiéchelle de champs denses continus de vecteurs-vitesse respectant les discontinuités de mouvement et ils se prêtent bien à une parallélisation sur des architectures simples et embarquables constituées de réseaux linéaires ou bidimensionnels de processeurs identiques et localement interconnectés. Ce dernier point est étudié en collaboration avec F. Charot du projet Api.

Système temps-réel embarqué - projet Prometheus

Participants : Fabrice Heitz

Par ailleurs, l'étude engagée depuis trois années dans le cadre du projet européen Eureka Prometheus (projet d'aide à la conduite automobile), a connu cette année son aboutissement à travers l'intégration par le Lasmea de l'université de Clermont-Ferrand, sur le véhicule de démonstration Prolab2, d'un ensemble de logiciels de détection et d'interprétation du mouvement développés à l'Irisa et s'appuyant sur des modèles markoviens. Le système est constitué de deux caméras embarquées connectées à la machine parallèle de vision "Transvision" conçue au Lasmea. Il s'intègre dans un système de vision plus vaste, aux fonctionnalités multiples, développé par les différents partenaires du projet Pro-Lab2, sous la direction des constructeurs automobiles PSA et Renault.

Ce système temps-réel embarqué, maintenant opérationnel, est en cours d'évaluation sur circuit et sur le réseau routier. Il permet, au niveau d'un carrefour, de donner au conducteur des informations qualitatives et quantitatives sur le comportement cinématique des véhicules latéraux (sens de parcours, distance) avec si besoin, déclenchement d'alarmes. Rappelons que le carrefour est reconnu statistiquement comme une zone à fort risque d'accidents, en raison d'une surveillance parfois insuffisante des espaces latéraux par le conducteur.

3.1.3 Détection, mesure et segmentation du mouvement

L'analyse du mouvement dans une séquence d'images représente un des sujets centraux en vision par ordinateur et est motivée par de nombreuses applications. Des situations difficiles se rencontrent ainsi dans des problèmes de navigation d'un véhicule en environnement extérieur quelconque ou de télésurveillance de sites, dans le domaine des transports par exemple, où des mouvements de personnes doivent être pris en compte.

Estimation robuste de mouvement et régularisation

Participants : Patrick Bouthemy, Jean-Marc Odobez

Nous avons défini l'an passé une méthode originale et performante d'estimation robuste multi-résolution de modèles paramétrés de mouvement. Cet estimateur a été mis à profit dans la définition d'un schéma de détection du mouvement dans une scène. Plus précisément, l'objectif est de détecter les régions dans l'image correspondant à des éléments mobiles d'une scène observée par une caméra elle-même placée sur un support mobile. Le principe de la méthode est dans un premier temps de compenser le mouvement induit dans l'image par le déplacement de la caméra. Ceci est réalisé à l'aide de l'estimateur robuste multi-résolution évoqué plus haut, qui permet de mesurer le mouvement dominant dans l'image sans être affecté par la présence d'autres mouvements. Le problème posé se réduit alors à la détection des régions non compensées (c.à.d non conformes au mouvement dominant) dans la séquence d'images ainsi reconstruite, soit un problème de décision binaire. Afin d'obtenir un taux de fausses alarmes très faible et afin d'extraire des masques des zones mobiles les plus précis et complets possible, nous avons traité ce problème comme un problème d'étiquetage statistique dans un cadre markovien. La définition d'une fonction d'énergie de type robuste quant au terme d'attache aux données, la prise en compte d'une intégration temporelle des observations et des cartes d'étiquettes et le recours à un modèle markovien multi-échelle ont permis d'obtenir des résultats très satisfaisants notamment sur des exemples de scène d'extérieur complexes correspondant à la base de données Prometheus.

Une extension assez directe de ce schéma a permis de passer d'une méthode de détection du mouvement (c.à.d. gestion de deux classes), à une méthode de segmentation de l'image en régions de mouvement homo-

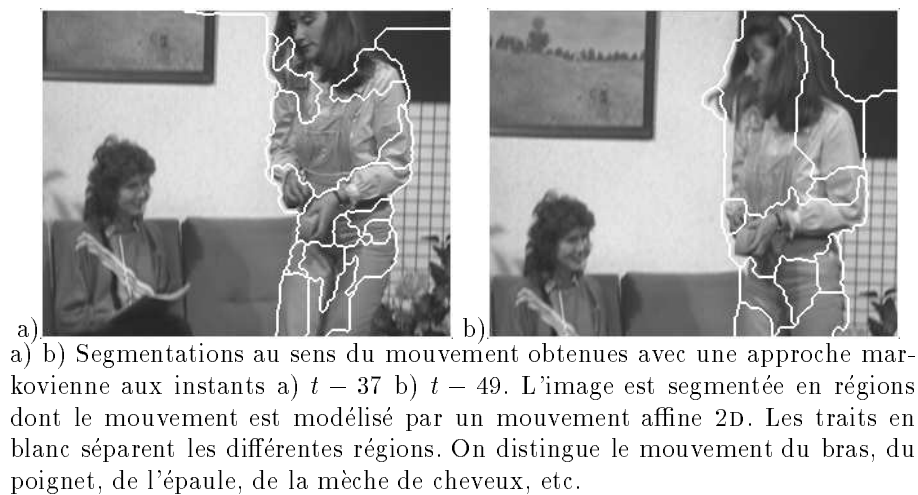
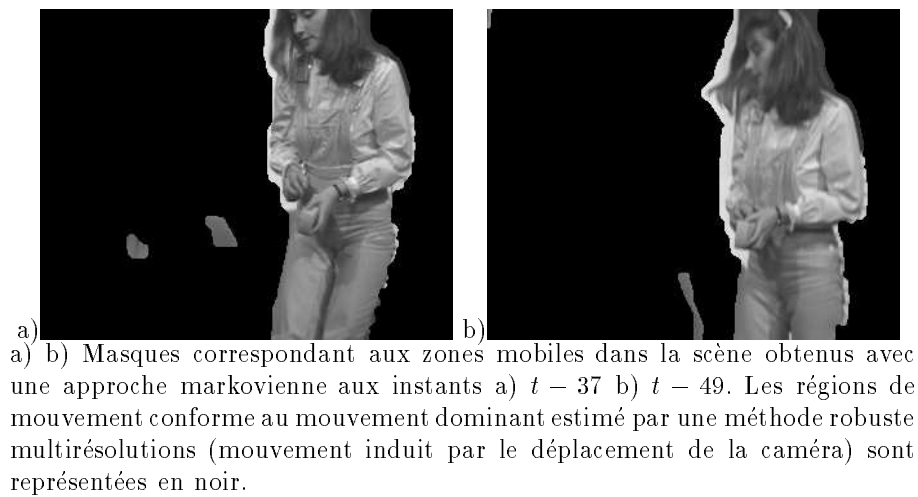
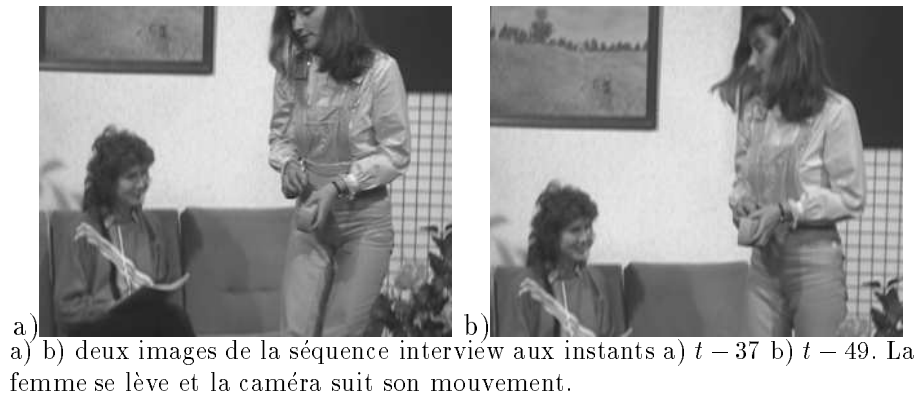


Figure 1 : Détection et segmentation régularisées du mouvement

gène (c.à.d. détermination de n composantes connexes, n étant identifié en ligne). Ces deux méthodes ont été testées sur des exemples relevant de domaines d'applications très variés. Nous pensons qu'elles peuvent aussi trouver une utilisation dans des schémas de codage "orienté-objets" à très bas débit.

Nous avons par ailleurs entrepris une évaluation de cet algorithme (et de ses variantes) dans le cadre de deux applications très différentes présentant des situations complexes quant aux formes et mouvements contenus dans les images. Des résultats prometteurs ont été obtenus. La première application concerne la classification de couches nuageuses en fonction de leur mouvement apparent mesuré dans des séquences d'images vidéo prises du sol. Cette étude s'insère dans le projet de télé-observation météorologique Tetrina mené avec Meteo-France, Timeat, Thomson-Ler et l'ENSTBr. La seconde application traite d'images sous-marines vidéo du fond marin et s'effectue en liaison avec l'Ifremer-Toulon. Il s'agit de mesurer le mouvement dans l'image induit par le déplacement du véhicule sous-marin, l'objectif étant d'asservir la caméra sur une zone d'intérêt lors d'une campagne d'étude scientifique. Plusieurs mouvements secondaires ou parasites peuvent venir compliquer cette phase d'estimation (poissons en mouvement, présence de phénomènes de bulles ou de "fumées", etc ...). De plus, dans ces deux exemples d'application, les images présentent de nombreuses zones d'intensité uniforme, peu de contrastes marqués et peu de formes nettes et stables.

Contours et mouvement

Participants : Patrick Bouthemy, Yann Ricquebourg

Une nouvelle étude a été lancée concernant l'analyse et le suivi de mouvement articulé dans une séquence d'images. En pratique il s'agira de mouvement de personnes, avec en particulier un contexte d'application en télésurveillance (de parkings) dans le cadre du projet Praxitèle. Une première étude de ce type avait été menée sur la période 89-92 à travers la thèse de Juhui Wang. Cette dernière visait plus à l'analyse du geste en tant que tel, en particulier du geste sportif, (mouvement du bras ou des jambes), et manipulait des modèles de structure et de mouvement 3D identifiés directement à partir des variations spatio-temporelles des intensités dans la séquence d'images. Cette fois, l'objectif est plutôt de différencier et de localiser dans l'image un ou des mouvement(s)

articulé(s) parmi des mouvements rigides pouvant correspondre au mouvement relatif entre la scène statique et la caméra mobile, et/ou aux mouvements propres d'objets rigides tels que des véhicules, ... La voie d'investigation actuellement suivie privilégie la mesure d'informations locales de mouvement éparses, sans segmentation explicite de l'image en régions homogènes au sens du mouvement, et l'identification du mouvement articulé, sans recours à des modèles 3D. Cette année, l'effort de recherche a évidemment porté sur la phase de mesure. Nous avons choisi de restreindre l'estimation du mouvement 2D aux contours préalablement détectés dans l'image, et de la limiter à la composante de vitesse perpendiculaire au contour. Nous avons repris la méthode, que nous avons définie en 85-86, basée sur des tests de vraisemblance, de détermination d'éléments de contour en mouvement modélisés comme des portions de surfaces spatio-temporelles dans une séquence d'images. Nous en avons proposé une refonte dans un cadre markovien, qui permet d'atteindre une mesure de précision sous-pixelique, d'éviter les phases de recherche exhaustive sur un intervalle prédéfini, et de fournir une régularisation de la solution par la prise en compte du contexte local. Des évaluations sur des séquences d'images réelles représentatives sont en cours.

3.1.4 Étude du mouvement déformable

L'analyse du mouvement déformable que l'on peut décomposer en trois grandes classes, mouvements articulé, élastique et fluide, représente un sujet de première importance de par la place croissante prise par les applications dans le domaine de l'environnement, de la santé, et du multi-média (interfaces gestuelles par exemple ou reconnaissance du mouvement des lèvres). Nous nous sommes intéressés à deux types de situations, l'une où la forme de l'objet d'intérêt reste structurée, l'autre où la forme est floue et le mouvement fluide. Dans les deux cas, nous avons privilégié une approche statistique des problèmes posés.

Modèles statistiques hiérarchiques déformables

Participants : Charles Kervrann, Fabrice Heitz, Patrick Pérez

Cette étude vise à développer une approche générale (de nature statistique) pour la segmentation, le suivi et la caractérisation d'objets

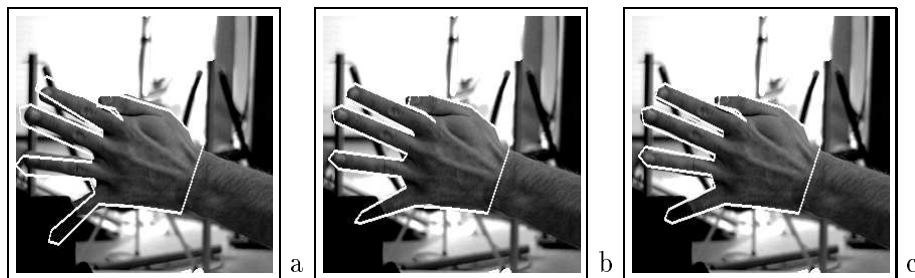


Figure 2 : Segmentation et suivi temporel de modèles stochastiques déformables : Segmentation basée mouvement et suivi d'une main au moyen d'une procédure d'estimation récursive des paramètres de déformation du modèle polygonal par filtrage de Kalman a) prédiction des paramètres de déformation ; a) prédiction des paramètres de déformation ; b) ajustement des paramètres de déformation pour extraire l'objet en mouvement ; c) estimation des paramètres filtrés.

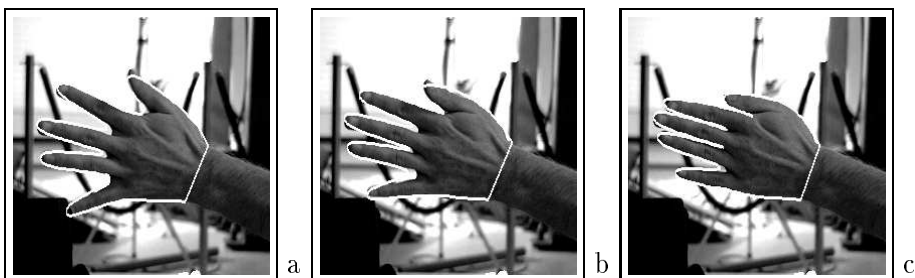


Figure 3 : Segmentation et suivi temporel de modèles stochastiques déformables : Segmentation basée mouvement et suivi d'une main modélisée par une B-spline dans une séquence d'images longue.

déformables structurés dans une séquence d'images. Un modèle statistique hiérarchique a ainsi été conçu l'an passé pour rendre compte des différents types de déformations observées dans l'évolution dynamique de formes structurées. La représentation s'appuie sur un modèle markovien gaussien des déformations locales d'une forme prototype ("template"), couplé à une représentation modale, par décomposition de Karhunen-Loeve, des déformations globales de la structure. Les modes dominants de déformation sont obtenus dans une phase d'apprentissage effectuée hors-ligne. Cette représentation hiérarchisée de l'information s'est révélée particulièrement parcimonieuse puisqu'un nombre réduit

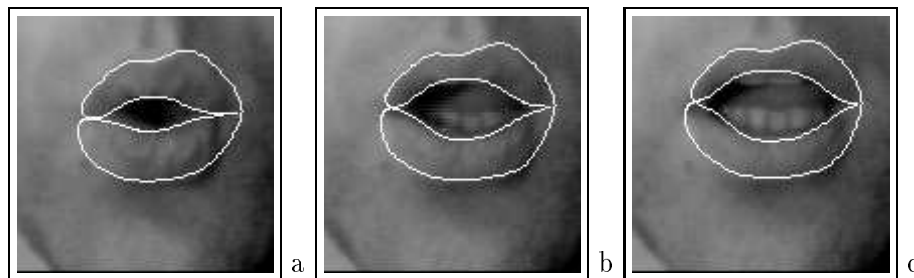


Figure 4 : Segmentation et suivi temporel de modèles stochastiques déformables : Segmentation basée contour et suivi d'une bouche en mouvement modélisée par une B-spline dans une séquence d'images longue.

de paramètres globaux (de l'ordre d'une dizaine) a permis de représenter avec une bonne précision les principales déformations de la structure. Le gain en vitesse de traitement, par rapport aux modèles statistiques classiques purement locaux, est alors considérable. Rappelons également que l'adoption d'algorithmes d'optimisation stochastique discrets (rapides, puisque portant sur un faible nombre de variables dans ce cas) a permis de s'affranchir des phases manuelles d'initialisation du modèle, qui sont souvent pénalisantes dans les applications.

Les développements autour de ce modèle ont porté cette année sur quatre points complémentaires.

Tout d'abord, le modèle de la forme prototype a été raffiné, en adoptant une représentation à base de splines plutôt qu'une simple approximation polygonale pour définir la silhouette des objets d'intérêt. Un raffinement a également été apporté dans la partie algorithmique, des versions continues des algorithmes d'optimisation (algorithmes SAS et de diffusion) ayant été implantées. Une meilleure précision est ainsi obtenue dans l'ajustement entre le modèle et les frontières de l'objet d'intérêt.

Par ailleurs, des algorithmes originaux de suivi de la structure déformable, s'appuyant sur le formalisme du filtrage de Kalman ont été développés. Un filtrage récursif temporel des paramètres de déformation, couplé à une technique de détection de rupture de modèle permet ainsi de prédire et de remettre à jour le modèle de façon fiable sur une séquence longue. Il en résulte une diminution des temps de calcul (par une meilleure initialisation de la forme dans l'image suivante) et

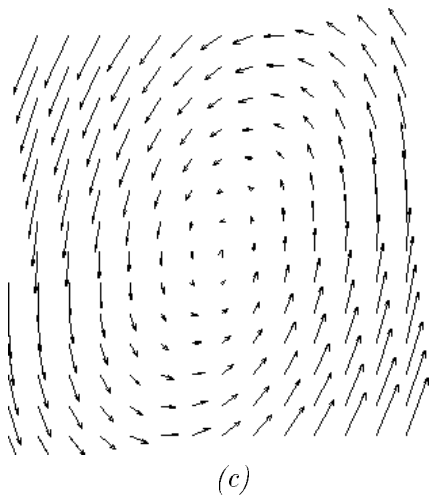
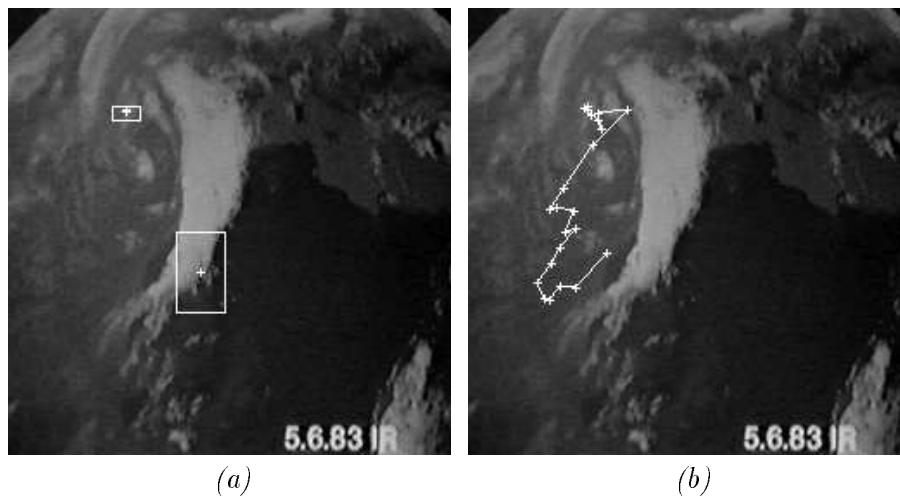
un meilleur “accrochage” du modèle lorsque les déplacements sont de grande amplitude. Le suivi fournit également des informations filtrées sur l'évolution dynamique des paramètres de déformation qui pourront ultérieurement être exploitées dans une interprétation ou une caractérisation des déformations au cours du temps. Le modèle cinématique utilisé dans le suivi intègre explicitement le modèle structural de l'objet et de ses déformations, et conduit donc à un véritable suivi de la structure (contrairement aux approches usuellement développées dans ce domaine qui s'appuient plutôt sur le suivi de primitives individualisées - points ou contours). Deux modèles cinématiques (modèle à vitesse instantanée et modèle à vitesse constante) ont été implantés et comparés. Ils conduisent à des performances comparables. Une diminution d'un facteur de trois du temps de traitement global par rapport à une version sans suivi a été observée (on passe de 6 mn. à moins de 2 mn. sur une station Sun Sparc IPX). Des suivis fiables d'une main en mouvement sur fond texturé, avec occlusions partielles, ont ainsi été obtenus sur des séquences composées de plus d'une centaine de trames.

Enfin, cette technique a été validée avec succès dans d'autres domaines d'applications (modélisation et suivi du mouvement des lèvres dans des séquences d'images de visages, représentation - en cours - de la structure et du mouvement du muscle cardiaque dans des séquences d'images médicales).

Points singuliers et caractérisation locale des déformations

Participants : Patrick Bouthemy, Mariette Maurizot, Jean-Marc Odobez

La caractérisation du mouvement et des déformations peut s'appuyer sur l'interprétation géométrique et qualitative des équations différentielles ordinaires, qui possède l'attrait de mettre en évidence de manière simple et explicite l'ensemble des solutions ou configurations particulières d'intérêt (“portraits”) et leurs points singuliers associés, c.à.d., le point où le vecteur vitesse est nul, (nœud, nœud spiralé, centre, point selle, ...). Cela présente un intérêt particulier dans l'étude du mouvement de fluides. Tout champ de vitesses 2D analytique peut être approché dans un voisinage convenable d'un point singulier, par un champ au premier ordre (c.à.d., affine: $\underline{v} = A\underline{x} + \underline{b}$) qui en partage les mêmes propriétés. La caractérisation du mouvement, à travers l'étude de la matrice A et de ses valeurs propres, nécessite donc à la fois de localiser de tels points



(a) : Les croix localisent les points caractéristiques, les rectangles la région d'incertitude. En fait seuls les points ayant une incertitude petite seront conservés.

(b) : Illustration de la convergence de l'algorithme à partir d'une initialisation lointaine. le chemin matérialise la suite des positions du centre de la fenêtre d'estimation pour le point caractéristique situé dans le coin supérieur gauche.

(c) : Partie du modèle affine de la convergence champ des vitesses estimé autour du point caractéristique situé dans le coin supérieur gauche.

(rotation spiralée)

Figure 5 : Localisation de points caractéristiques dans des séquences d'images météorologiques (cas d'une dépression)

singuliers, de définir le voisinage sur lequel l'approximation au premier ordre est valide, et d'identifier la matrice A . Ces problèmes sont bien sûr couplés.

Nous avons défini une méthode itérative pour localiser de tels points singuliers dans une séquence d'images (avec la collaboration de A. Iouditski et B. Delyon du projet AS). Elle met à profit un algorithme proposé par A. Iouditski. La taille de la fenêtre servant de support à l'estimation des paramètres du modèle affine 2D de mouvement est considérée comme un paramètre de nuisance. Un test "emboîté" permet de choisir, parmi un jeu prédéfini, la taille adéquate de la fenêtre de calcul (suffisante pour une estimation correcte en présence de bruit, pas trop grande pour ne pas introduire de biais par la perte du caractère linéaire du champ des vitesses sous-jacent). Une stratégie adéquate fournit la position du centre de la fenêtre à l'itération suivante. Les points singuliers sont localisés sans estimation préalable du champ des vitesses, les paramètres du modèle affine du mouvement étant déterminés à partir des variations spatio-temporelles des intensités. D'autre part, la convergence de l'algorithme a été obtenue pour un éventail assez large d'initialisations. Cette méthode a eu une première validation sur des séquences d'images satellitaires météorologiques. Des applications en mécanique des fluides (étude d'écoulements par exemple) sont également envisagées.

3.2 Vision active

La vision active consiste à élaborer des stratégies de perception afin d'en améliorer la qualité par rapport à une approche passive où l'on se restreint à observer, mesurer et interpréter les données issues du capteur sans agir sur celui-ci. En ce qui concerne les recherches développées dans le projet sur ce thème et en liaison avec les travaux effectués en vision dynamique, nous nous intéressons à la commande du mouvement de la caméra pour la réalisation de tâches robotiques (commande référencée vision, également appelée asservissement visuel) et aux améliorations que peut apporter cette commande dans la résolution de problèmes classiques en vision par ordinateur, tels le suivi et l'estimation du mouvement 3D d'objets mobiles (thèse de F. Bensalah) ou la reconstruction 3D de scènes complexes (thèse de E. Marchand).

3.2.1 Commande référencée vision

Participants : Farabi Bensalah, François Chaumette

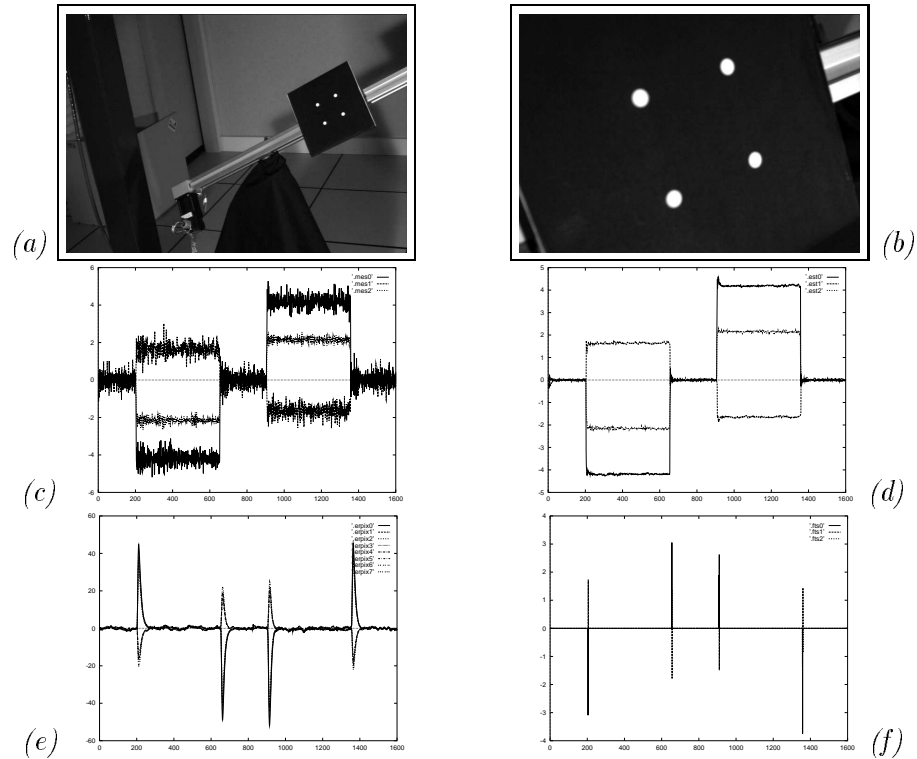


Figure 6 : Détection de changement de vitesse

- (a) Position initiale du carré dans l'image
- (b) Position désirée du carré dans l'image
- (c) Mesure du mouvement du banc
- (d) Estimation du mouvement du banc (par filtrage de Kalman et détection de rupture de modèle - GLR)
- (e) Erreur de poursuite du banc en pixels
- (f) Sauts de vitesse détectés

La commande référencée vision consiste à utiliser les informations fournies par une caméra mobile (par exemple embarquée sur un bras manipulateur ou un robot mobile) afin de réaliser des tâches élémentaires (positionnement, suivi d'objets mobiles) en contrôlant le mouvement de la caméra par rapport à son environnement. L'approche déjà dévelop-

pée depuis quelques années dans le projet repose sur la modélisation de fonction de tâches appropriées et consiste à spécifier le problème en termes de régulation dans l'image. Elle permet, en outre, de compenser les imprécisions des modèles, aussi bien du capteur que du porteur de la caméra, par des lois de commande robustes en boucle fermée sur les informations visuelles extraites de l'image.

Cette année a vu se concrétiser l'application de ces techniques dans le domaine de la robotique nucléaire par l'intermédiaire d'un transfert de technologie auprès de la société Edixia dans le cadre d'une convention avec la Direction des études et recherche d'EDF. L'application visée consiste à réaliser la préhension d'outils dont la position 3D est inconnue en utilisant une caméra embarquée solidairement à la pince du robot et en mettant à profit les techniques de commande référencée vision. Les expérimentations effectuées à l'Irisa sur la cellule de vision robotique équipée d'une pince et d'un porte-outils de type CCMOP ont permis de valider cette tâche de préhension. Des recherches sont actuellement en cours afin de rendre plus robuste cette approche, notamment pour prendre en compte d'éventuelles pertes momentanées ou définitives d'indices visuels (en raison d'occultations notamment).

Nous avons également poursuivi l'étude de l'estimation de la vitesse d'un objet mobile afin d'intégrer cette estimation dans des lois de commande permettant une poursuite de cet objet sans erreur de traînage. Cette année, nous nous sommes focalisés sur la détection et l'estimation des ruptures de modèles possibles du mouvement de l'objet afin de les compenser et d'atteindre un suivi minimisant les erreurs de traînage. L'algorithme retenu, basé sur le rapport de vraisemblance généralisé (GLR), a été implanté sur la cellule de vision robotique pour les tâches de focalisation sur une cible mobile par le contrôle de l'orientation de la caméra. Il a également été testé dans le contrôle des six degrés de liberté du robot afin pour l'obtention d'une liaison rigide entre la caméra et un objet mobile. Les résultats obtenus montrent la possibilité de détecter, estimer et compenser correctement des sauts en vitesse de l'objet. L'étude de mouvements plus complexes (sauts sur l'accélération par exemple) est actuellement en cours. Par ailleurs, nous avons débuté l'étude de l'estimation 3D du mouvement de l'objet par vision active. Le principe consiste à générer, par asservissement visuel et à l'aide de tâches secondaires compatibles, des mouvements de la caméra permet-

tant, au cours du suivi, d'observer l'ensemble des composantes de ce mouvement.

L'étude de l'utilisation d'informations visuelles dynamiques (et non plus seulement géométriques) dans les boucles de commande par asservissement visuel a également été poursuivie. Cette étude devrait élargir grandement le champ d'application de la commande référencée vision par la réalisation de nouvelles tâches, la suppression de connaissances géométriques a priori sur la scène et son applicabilité à des scènes complexes grâce à des traitements de séquences d'images fondés sur l'analyse du mouvement (segmentation, estimation robuste), étudiés par ailleurs dans le projet. Après l'étude et la validation expérimentale d'une tâche consistant à aligner l'orientation de la caméra sur sa direction inconnue de translation, nous avons cette année réalisé une tâche de compensation du mouvement permettant le suivi d'un objet mobile en utilisant la mesure de son mouvement 2D (et non plus sa position mesurée dans l'image). Cette réalisation s'inscrit dans le cadre des projets Praxitèle et Vidac, la tâche considérée étant une brique de base des systèmes de télésurveillance active. Ne disposant pas actuellement d'algorithmes de segmentation au sens du mouvement fonctionnant à une cadence proche de la cadence vidéo, signalons que, pour cette tâche de suivi, le mouvement 2D de l'objet est obtenu à l'aide des mesures successives de sa position. Des études seront menées prochainement afin d'adapter un des algorithmes de segmentation développés dans le projet de manière à disposer d'une version temps réel sur une carte de traitements d'images spécialisée. De plus, des études sont en cours afin de généraliser cette nouvelle approche, de manière à caractériser et modéliser l'ensemble des informations visuelles dynamiques utilisables en commande référencée vision et l'ensemble des tâches que cette nouvelle approche permet de réaliser.

Par ailleurs, en collaboration avec l'équipe Prima du Lifa Grenoble et dans le cadre d'une action commune du GDR-PRC CHM (pôle vision), nous avons traité le problème de la commande redondante du point de fixation d'un système stéréoscopique à huit degrés de liberté. L'introduction du contrôle de la profondeur du point de fixation dans la boucle de commande et la modélisation de tâches secondaires adéquates (symétrie des axes de vergence, éloignement par rapport aux butées mécaniques) ont permis d'aboutir à une solution élégante et efficace, s'intégrant par-

faitement dans l'approche fonction de tâche. La validation expérimentale de ces algorithmes a été réalisée sur la cellule robotique du Lifa.

3.2.2 Reconstruction 3D par vision active

Participants : Eric Marchand, François Chaumette

Cette année, nous avons poursuivi l'étude de la reconstruction 3D d'une scène par vision active. Le schéma de base proposé consiste à focaliser successivement la caméra sur les différents objets de la scène. Les techniques d'asservissement visuel sont utilisées pour effectuer cette focalisation et pour commander les mouvements de la caméra afin d'obtenir une reconstruction et une localisation optimales (au sens de la suppression de biais dans l'estimation et de la robustesse vis à vis des erreurs de mesure). La méthode de reconstruction que nous avons développée s'appliquait jusqu'à présent aux primitives géométriques paramétrables simples (points, droites, cercles, sphères et cylindres). Nous avons cette année débuté la généralisation de cette approche au cas des quadriques quelconques (ellipsoïdes, paraboloides et hyperboloides), de même qu'une étude sur la reconstruction de formes complexes, c'est-à-dire non représentables globalement par une équation géométrique paramétrable.

Parallèlement, nous avons poursuivi l'étude sur les stratégies de perception dans un contexte de vision active. Ce travail, initialisé dans le cadre d'une action commune du GDR-PRC (pôle Vision) avec le projet Icare de l'Inria Sophia-Antipolis, fait maintenant partie intégrante du projet PRC VIA (vision intentionnelle et action) traitant de la perception et du raisonnement géométrique pour le mouvement dans le monde réel. Le but de ce travail consiste à enchaîner les différentes tâches nécessaires à la reconstruction précise et complète de la scène observée. Pour l'instant, les tâches suivantes ont été réalisées :

- création d'une base de données 2D comportant l'ensemble des segments initialement présents dans l'image fournie par la caméra ;
- sélection successive des différents segments en fonction de leur taille et de leur position dans l'image ;
- reconnaissance des primitives 3D (segments ou cylindres) basée sur un test de rapport de vraisemblance à l'issue d'une première phase de reconstruction 3D par vision active ;

- reconstruction optimale des primitives 3D en utilisant la méthode de focalisation active décrite précédemment ;
- reconstructions 3D grossières effectuées en parallèle d'une reconstruction optimale ;
- estimation de la longueur des primitives 3D par l'observation successive de leurs points d'extrémité, en générant par asservissement visuel un mouvement le long de la primitive considérée ;
- création éventuelle de polygones à partir des segments 3D reconstruits.

Nous avons également développé un réseau d'automates permettant d'enchaîner correctement ces différentes phases. Un travail important, mené avec la collaboration d'Eric Rutten du projet EP-ATR, a consisté cette année à appliquer l'approche synchrone pour la programmation de ce réseau et les différentes tâches le constituant. Plus précisément, nous avons utilisé le langage synchrone SIGNAL développé à l'Irisa. Les intérêts de SIGNAL sont multiples pour notre application. D'une part, son caractère flot de données est parfaitement adapté aux tâches d'asservissement visuel et aux algorithmes de reconstruction 3D. De plus, ce langage permet un parallélisme implicite entre plusieurs actions (par exemple, asservissement visuel effectué en parallèle à des reconstructions optimale et grossières). Les opérateurs de SIGNAL permettent aussi une programmation aisée des algorithmes utilisant des valeurs mesurées à des instants passés (dans notre cas, mesure du déplacement de la caméra, mesure du déplacement des primitives entre plusieurs images successives, filtres moyennés, etc.). Finalement, l'introduction récente de la notion d'intervalle dans SIGNAL a permis une programmation du réseau d'automates proche de sa spécification et dans le même cadre de développement.

Sur le thème des stratégies de perception, nos perspectives de recherche portent maintenant sur la résolution du problème de l'exploration des zones de la scène qui n'ont pas encore été observées par la caméra, ceci afin d'obtenir la complétude de la reconstruction 3D.

3.3 Compression de séquences d'images

Ce thème de recherche vise à l'élaboration de méthodes algorithmiques efficaces d'analyse et de compensation de mouvement dans une ou

plusieurs séquences d'images afin de tirer profit, dans un but de compression, des redondances interimages présentes au sein de celles-ci. Ces travaux s'inscrivent dans le cadre des contributions pour l'élaboration des futures normes de compression du signal vidéo animé (famille MPEG, HDTV numérique, TV3D...) et la conception de nouveaux services de vidéo-communications (télésurveillance, télémanipulation, ou consultation de bases de données image, ...).

Plusieurs critères d'évaluation conditionnent ces recherches :

- une qualité de reconstruction des images après traitement (de codage-décodage ou d'interpolation). Jusqu'à présent, seuls les critères de qualité visuelle tolérant une reconstruction numérique non exacte étaient abordés. Dans le contexte de l'imagerie médicale, les critères de reconstruction numérique exacte doivent être désormais pris en compte (thèse de V. Nzomigni menée en collaboration avec le partenaire GEMS-Europe, anciennement société GE-CGR). Dans un contexte de télésurveillance active, les schémas de compression avec focalisation d'attention introduisent également la notion de qualité de reconstruction adaptative selon les régions d'intérêt (thèses de E. Nguyen et N. De Maistre) ;
- une faible complexité algorithmique et/ou un caractère parallélisable des méthodes retenues, ceci afin de répondre aux contraintes fortes de temps réel présentes dans les applications visées. Une étude menée avec le projet Api (P. Quinton, F. Charot) concerne la parallélisation d'algorithmes d'estimation de mouvement non réguliers (thèse de P. Lemonnier, voir le rapport d'activité du projet Api) ;
- une compression plus sémantique du signal d'images par l'introduction de modèles de mouvements paramétriques, d'un suivi spatiotemporel de cartes de segmentation et de l'adaptation du critère de qualité selon les régions d'intérêt. Ce thème fait l'objet de thèses déjà bien engagées (thèses de V. Garcia-Garduno et E. Nguyen).

D'autres études de doctorat sont également lancées concernant les schémas d'interpolation temporelle (L. Bonnaud) ou spatiale (H. Lefloch) et les approches par quantification vectorielle arborescente (V. Ricordel).

3.3.1 Compression sans pertes de séquences d'images médicales

Participants : Victor Nzomigni, Claude Labit

En collaboration avec la société GEMS-Europe, nous avons entrepris une étude sur les techniques de compression sans pertes faisant intervenir un module d'estimation de mouvement pour la compression réversible de séquences d'images biomédicales. Le schéma général utilisé consiste à décrire la technique de compression sans pertes comme une technique de compression avec pertes (étape de décorrélation par codage prédictif par exemple) à laquelle est adjoint un procédé de codage sans pertes (de l'erreur de prédiction par exemple).

Dans ce cadre, deux phases d'études distinctes sont donc à mener. La première concerne le choix d'une méthode de décorrélation du signal. Au vu de la nature fortement bruitée spatiotemporellement et des mouvements de nature essentiellement déformable dans les séquences d'images médicales manipulées, les algorithmes d'estimation de mouvement usuels (mise en correspondance par blocs ou pel-récurrents) échouent lorsqu'ils sont utilisés pour réaliser une prédiction locale avec compensation de mouvement. Le schéma de décorrélation actuellement testé est une prédiction hybride spatiotemporelle avec compensation de mouvement. Cette compensation est réalisée sur la base d'une approche entropique adaptative mesurant, pour un modèle de mouvement donné, sur un bloc d'informations de taille fixée, le coût global entropique pour quantifier les caractéristiques de mouvement identifiées et l'information résiduelle de prédiction compensée en mouvement qui en résulte.

La seconde partie que nous avons déjà plus particulièrement développée les années antérieures concerne le choix d'un algorithme de compression sans pertes numériques, efficace et bien adapté à la source de symboles issus de la phase de décorrélation précédente. Les algorithmes classiques tels que Huffman, Huffman adaptatif, Lempel-Ziv et algorithme du Context (travaux de Rissanen et al.) ont été développés et comparés. L'algorithme du Context portant sur une source n -aire ou sur n sources binaires issues d'une décomposition préalable en arbre de Huffman semble être le plus efficace.

Sur les séquences testées (coronarographie, ventriculographie), un taux de compression moyen de 3.5 à 4 est obtenu ce qui représente un gain appréciable par rapport aux bas taux de compression (de l'ordre de

1.5 ou 2) traditionnellement atteints (avec Lempel-Ziv par exemple). Ces travaux se poursuivent par une comparaison avec les approches traditionnelles basées JPEG ou MPEG et par l'introduction de la notion de focalisation d'attention (segmentation externe décomposant le signal en plusieurs sources d'informations).

3.3.2 Compression par analyse sous-bandes et focalisation d'attention

Participants : Eric Nguyen, Claude Labit, Jean-Pierre Leduc, Nicolas De Maistre

Contrairement aux méthodes désormais standards (ISO JPEG, MPEG) pour la compression d'images fixes ou animées, l'approche par "focalisation d'attention", qui est une méthode basée région, permet d'appréhender le problème de la compression à bas débit de manière *sélective* au sens d'une qualité de reconstruction inhomogène spatialement. L'étude est fondée sur la notion de *région(s) d'intérêt* interactivement désignée(s) interactivement ou automatiquement sélectionnées par une segmentation dédiée selon des critères dépendant de l'application (texture, mouvement). L'opération de compression pour un débit de transmission/stockage global donné conduit à privilégier la qualité de reconstruction sur les différentes régions définies selon la hiérarchie d'intérêt.

La méthode développée repose sur la représentation du signal en sous-bandes associée à des bases de projection à la fois bien localisées en espace et en "fréquence". Cette localisation espace-fréquence permet naturellement le codage sélectif des zones spatiales d'intérêt en prenant en compte l'analyse en fréquence pour le rendu du signal reconstruit après synthèse. La quantification opérant la compression proprement dite est réalisée par l'optimisation numérique du critère débit-distorsion (R-D) sur la base d'un choix fini de conditions de quantification (scalaire ou vectorielle) admissible par sous-bande associée à chaque région. La représentation peut sous certaines conditions être rendue adaptative en utilisant les techniques récentes de recherche de "meilleure base" de représentation sous critère R-D. Cette méthode étend la notion d'allocation optimale R-D en codage sous-bandes usuel, à la notion de codage région/sous-bandes. Une mesure de distorsion quadratique (EQM) pondérée spatialement et/ou fréquemment est utilisée. La pondération

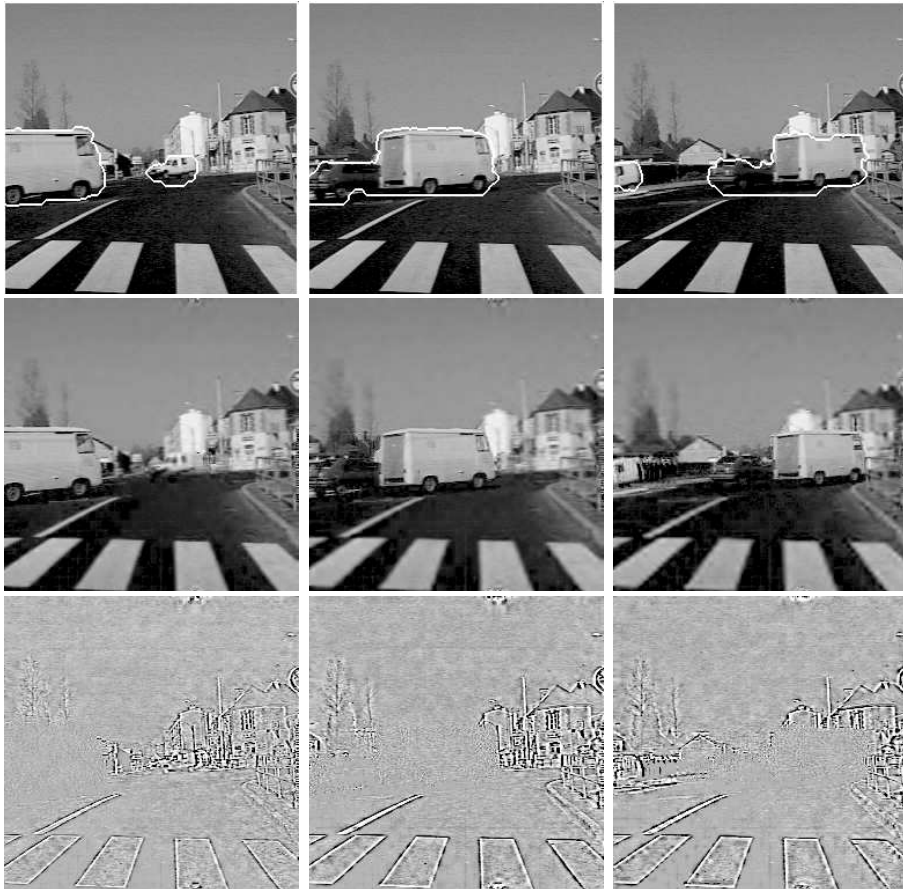


Figure 7 : Compression avec focalisation d'attention

Images à $t=24, 46$ et 65 d'une séquence de trafic routier: originales avec surimposition de la segmentation basée mouvement (la région d'intérêt (ROI) est constituée par le(s) véhicule(s) $V1$ se déplaçant de gauche vers la droite), reconstruites après codage hybride avec focalisation visuelle pour un débit entropique de $0.1bpp$, et d'erreur (amplifiées par 5 avec offset de 128).

spatiale définie selon la hiérarchie d'intérêt permet de quantifier les contributions relatives des erreurs de reconstruction de chacune des régions. La pondération en "fréquence" portant sur les sous-bandes et étudiée précédemment permet dans une certaine mesure de prendre en compte les données psychophysiques de sensibilité en fréquence du sys-

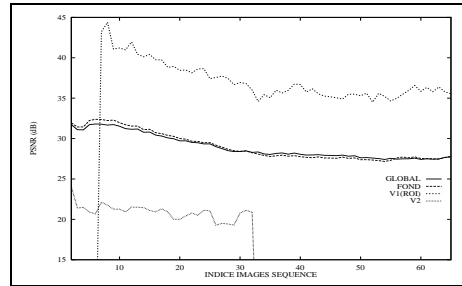


Figure 8 : Distorsions (EQM) mesurées au cours de la séquence (globale et par région).

tème visuel humain (SVH). La sensibilité peut être définie de manière différente pour chacune des régions considérées dans l'encodage comme indépendantes. Dans ce contexte, un schéma de codage hybride (couleur introduit dans un schéma temporel DPCM avec prédiction au sens du mouvement) de séquences d'images animées a été développé sur la base d'une segmentation basée mouvement développée dans l'équipe. En amont, l'étape de segmentation fournit l'information des régions homogènes au sens du mouvement décrit par un modèle affine qui fournit une information dense pour la compensation sous une forme compacte. Cette information peut alors être utilisée à la fois pour la phase de sélection de la (des) région(s) d'intérêt et pour la phase de compensation dans le cas d'un codeur hybride classique en vue de la compression sélective du signal d'erreur de prédiction. Dans l'hypothèse d'une région d'intérêt unique suivi du regard par l'observateur au cours de la séquence, la pondération spatio-fréquentielle est définie en fonction des vitesses apparentes moyennes des régions par rapport à la région d'intérêt. On définit alors explicitement un système d'attention visuelle pour le rendu des images reconstruites prenant en compte la dépendance en vitesse visuelle de la réponse du SVH. La validation d'un tel schéma est en cours sur des séquences d'images réelles de scènes de trafic routier ou de vidéoconférence. L'étude se poursuit par l'étude de différentes configurations de modèles hybrides pouvant inclure le couplage entre la phase d'analyse-segmentation et la phase de compression, celles-ci étant supposées en premier lieu indépendantes et séquentielles.

3.3.3 Filtrage compensé en mouvement et approche multirésolution

Participants : Jean-Pierre Leduc, Claude Labit

Ce travail concerne le filtrage spatio-temporel ($2D+t$) des signaux vidéo (séquence d'images). Ce type de traitement s'avère efficace pour de multiples applications qui s'étendent des signaux de haute définition (TV et HDTV) ceux de très bas débits. En effet, le filtrage spatio-temporel est un outil essentiel aux applications suivantes : la prédiction spatio-temporelle d'image (codage), l'interpolation temporelle (codage et changement de format), le filtrage réducteur de bruit (amélioration de la qualité de documents bruités) et l'analyse multirésolution et spatio-temporelle (codage). Les axes de recherche retenus vont se concentrer essentiellement sur l'analyse et l'interpolation. Les outils mis en œuvre sont la segmentation spatio-temporelle et le filtrage sous-bande par ondelettes discrètes.

Les filtres tridimensionnels linéaires usuels sont bien souvent implantés de manière similaire selon les trois dimensions. Toutefois, la dimension temporelle diffère des deux autres dimensions spatiales du point de vue de l'échantillonnage et de la dépendance des variations temporelles vis-à-vis des détails spatiaux et des vitesses des objets. Il en résulte des contraintes sur le spectre local spatio-temporel qui rendent le filtrage temporel classique significativement sous-optimal. L'analyse temporelle va donc être effectuée par des filtres d'ondelettes biorthogonaux appliqués le long des trajectoires de mouvement. Cette méthode est intitulée filtrage compensé en mouvement. L'avantage de la technique est de capturer la majeure partie du spectre dans la bande de basse fréquence lors d'une décomposition dyadique et de ne laisser que du bruit dans la bande de haute fréquence. Les retombées intéressantes de cette technique sont donc une décorrélation accrue du contenu temporel du signal et une approche multirésolution. L'analyse du signal qui en résulte consiste ainsi en une projection locale et adaptée de l'opérateur de décomposition le long des trajectoires du mouvement.

La mise en œuvre de cette technique requiert une analyse préalable du signal vidéo en termes de segmentation spatio-temporelle au sens du mouvement. Chaque image de la scène est ainsi segmentée en un ensemble de régions bidimensionnelles lesquelles sont caractérisées par un modèle propre de mouvement affine. L'ajout d'un module de suivi