

Projet MOVI

Modélisation pour la vision par ordinateur

Rhône-Alpes

THÈME 3B

R *apport*
d'Activité

1999

Table des matières

| | | |
|----------|--|-----------|
| 1 | Composition de l'équipe | 2 |
| 2 | Présentation et objectifs généraux | 3 |
| 3 | Domaines d'applications | 4 |
| 3.1 | Panorama | 4 |
| 3.2 | Vision et robotique | 4 |
| 3.3 | Indexation de bases d'images | 4 |
| 3.4 | Vision et défense nationale | 5 |
| 4 | Résultats nouveaux | 5 |
| 4.1 | Géométrie multi-images et multi-caméras | 5 |
| 4.2 | Couplage vision/robotique | 9 |
| 4.3 | Indexation d'images et reconnaissance d'objets | 12 |
| 4.4 | Structuration de vidéos | 14 |
| 5 | Contrats industriels (nationaux, européens et internationaux) | 16 |
| 5.1 | Reconnaissance et indexation de vues aériennes | 16 |
| 5.2 | Indexation de vidéo | 17 |
| 5.3 | Projet RNRT "AGIR" | 17 |
| 6 | Actions régionales, nationales et internationales | 17 |
| 6.1 | Actions régionales | 17 |
| 6.2 | Actions nationales | 18 |
| 6.3 | Actions financées par la Commission Européenne | 18 |
| 6.4 | Relations bilatérales internationales | 19 |
| 6.4.1 | Amérique | 19 |
| 6.4.2 | Asie | 19 |
| 7 | Diffusion de résultats | 19 |
| 7.1 | Animation de la communauté scientifique | 19 |
| 7.2 | Enseignement universitaire | 20 |
| 7.3 | Participation à des colloques, séminaires, invitations | 21 |
| 8 | Bibliographie | 21 |

1 Composition de l'équipe

Responsable scientifique

Radu Horaud [(depuis le 18 octobre 1999) directeur de recherche au CNRS, en détachement à l'INRIA depuis le 1 octobre 1998]

Roger Mohr [(avant le 18 octobre 1999) professeur, ENSIMAG-INPG, chercheur invité à l'Univ. de Melbourne du 1 octobre 98 au 31 juillet 1999]

Assistante de projet

Danièle Herzog [TR INRIA, jusqu'au 31 mai 1999]

Personnel INRIA

Cordelia Schmid [chargée de recherche]

Peter Sturm [chargé de recherche, depuis le 1 novembre 1999]

Personnel CNRS

Patrick Gros [chargé de recherche, jusqu'au 1 juin 1999]

Long Quan [chargé de recherche]

Personnel universitaire

Edmond Boyer [maître de conférences à l'université Joseph Fourier]

Ingénieurs experts

Bart Lamiroy [en collaboration avec le projet BIP]

Chercheurs doctorant

Marc-André Ameller [allocataire ENS Cachan, thèse ayant débuté le 1 octobre 1999]

Nicolas Andreff [allocataire MENESR puis ATER, en commun avec le projet BIP]

Cristian Sminchisescu [allocataire CIES depuis le 1 octobre 1999]

David Demirdjian [allocataire MENESR]

Yves Dufournaud [boursier CIFRE avec la Société Aérospatiale]

Ryad Hammoud [allocataire INRIA, sur contrat avec la Société Alcatel]

Maxime Lhuillier [allocataire MENESR]

Frédéric Martin [allocataire MENESR à partir du 1 octobre 1999]

Krystian Mikolajczyk [allocataire CIES sur contrat depuis le 1 octobre 1999]

Fabien Pelisson [boursier de la région Rhône-Alpes]

Sylvaine Picard [allocataire DRET jusqu'au 30 octobre 1999]

Andreas Ruf [boursier TMR (bourse Marie-Curie)]

William Triggs [allocataire INRIA sur contrat européen]

Chercheurs post-doctorants

Alberto Aguado [boursier TMR, université de Southampton]

Christophe Biernacki [post-doc INRIA jusqu'au 31 août 1999]

Professeur invité

Chengke Wu [université de Xidian, Xi'an, Chine, du 15 septembre au 15 novembre 1999]

2 Présentation et objectifs généraux

Le projet MOVI est un projet commun entre le CNRS, l'INPG, l'UJF et l'INRIA, localisé à l'INRIA Rhône-Alpes et appartenant au laboratoire GRAVIR de la fédération IMAG.

Comprendre l'espace tridimensionnel perçu par une caméra, identifier les objets qu'il contient, se localiser et agir par la commande forment un premier ensemble d'activités qui peut se regrouper schématiquement sous le vocable de « géométrie de la vision 3D ». Un second groupe d'activités, plus récent et plus modeste à ce jour, s'intéresse à la recherche d'objets ou d'images dans une base de référence très large par des techniques d'indexation. Plus précisément, sur ces deux thèmes, nous développons les points suivants :

- étalonnage des capteurs visuels ;
- couplage de la vision d'un robot et de son contrôle : asservissement visuel, manipulation guidée par la vision, étalonnage du couple caméra-robot ;
- mise en correspondance d'images (stéréovision multiple) ;
- apprentissage de modèles visuels à des fins de reconnaissance et de localisation ;
- indexation de large base d'images, structuration de vidéos.

Outre les approfondissements des aspects mentionnés ci-dessus, notre projet vise à développer des démonstrateurs intégrant différents aspects de ce savoir-faire. Plus que la juxtaposition de différentes techniques, cette intégration amène à reposer les problèmes fondamentaux dans des cadres nouveaux, comme par exemple celui de la commande par la vision de mouvements de bras de robots ou la création de bases d'images interrogeables par leur contenu.

3 Domaines d'applications

3.1 Panorama

Les domaines d'application habituels de la vision tridimensionnelle par ordinateur ont été liés à la robotique ou à la défense. Notre projet continue à être bien présent dans ces domaines. Il est cependant de nouveaux domaines en émergence et liés aux médias. Dans ces créneaux nous nous positionnons dans la problématique de la synthèse d'images à partir d'images, et de façon plus conséquente, sur le problème de l'indexation de bases d'images à partir de leur contenu, ou avec des outils de la même famille, sur la structuration de vidéos en terme d'objets présents.

3.2 Vision et robotique

Résumé : *Nous abordons plusieurs aspects du couplage de la vision et de la robotique : étalonnage et auto-étalonnage d'ensembles de caméras et de robots, qu'ils soient liés rigidement ou non, le guidage visuel de robots manipulateurs, la modélisation d'un robot dans un espace visuel calibré ou non calibré.*

Les domaines d'applications de l'intégration vision-robotique sont les suivants :

- l'auto-étalonnage d'une tête stéréoscopique à deux degrés de liberté (deux rotations) ; ce problème a été abordé dans le cadre du contrat européen Reactive LTR VIGOR 26247 (1998-2001) ;
- le guidage visuel d'un robot à plusieurs degrés de liberté à l'aide d'une ou plusieurs caméras non étalonnées ; ce problème est abordé dans le cadre du projet VIGOR en collaboration avec Odense Steel Shipyard (chantiers navals). L'application visée est le positionnement d'une torche de soudure par rapport à une pièce de bateau avec une précision inférieure à 1mm ;
- l'auto-étalonnage de la chaîne cinématique d'un mécanisme articulé a de nombreuses applications dès lors qu'il s'agit d'estimer les paramètres d'un robot "in situ".

3.3 Indexation de bases d'images

Résumé : *Les secteurs de la presse et de l'audiovisuel, ceux de l'industrie (imagerie scientifique), de la médecine ou encore ceux de la propriété industrielle collectent des quantités impressionnantes d'images qu'il faut pouvoir gérer. Souvent le processus d'acquisition est plus rapide et simple que celui de l'indexation, ce qui fait naître un besoin urgent d'indexation automatique par le contenu.*

Les applications des bases d'images peuvent être divisées comme suit :

- agence de presse et de l'audiovisuel : ce segment se caractérise par un énorme volume de données d'images, de plusieurs millions pour les images fixes à des centaines de milliers de vidéos pour les images animées ; les besoins en consultation sont complexes, les requêtes sont souvent de haut niveau et elles requièrent une forte interactivité avec l'utilisateur ;
- les secteurs médicaux, scientifiques et industriels ont des besoins plus spécifiques, très liés aux différents domaines concernés ; les volumes dépassent parfois les centaines de milliers d'images et les types d'interrogation sont très variables ;
- les demandes de la propriété industrielle correspondent à une interrogation bien plus précise : « quelles sont les archives similaires à tout ou partie d'un motif présenté? » Pour cette classe d'application, la réponse exacte, si elle existe, ne doit pas être manquée. Le volume d'images traité peut atteindre celui des agences de presse.

A la différence de l'indexation automatique des textes, les images n'apportent pas directement d'information conceptuelle de haut niveau sémantique. Il faut donc développer, pour les différentes classes d'application, des index qui soient pertinents pour permettre une recherche efficace et une interaction proche des concepts de l'utilisateur.

3.4 Vision et défense nationale

Notre projet a développé depuis plusieurs années un partenariat de recherche avec la Division Missiles de la Société Aérospatiale. Après avoir collaboré dans le domaine de la reconstruction tri-dimensionnelle nous abordons depuis 1998 le problème de la comparaison de deux images prises dans des conditions de vue très différentes. L'application envisagée est la mise en correspondance de données visuelles obtenues avec un "drone"¹ de reconnaissance d'une part et un missile opérationnel d'autre part (voir section "contrats industriels").

4 Résultats nouveaux

4.1 Géométrie multi-images et multi-caméras

Participants : Edmond Boyer, David Demirdjian, Radu Horaud, Maxime Lhuillier, Long Quan, Andreas Ruf, Peter Sturm, William Triggs.

Mots clés : géométrie, étalonnage, reconstruction 3-D, vision 3-D, séquence d'images.

Résumé : *Nous avons été présents depuis plusieurs années dans les études sur l'utilisation de la géométrie projective pour la vision tri-dimensionnelle. L'avantage essentiel de cette approche est de permettre de s'affranchir de l'étalonnage des systèmes de vision, offrant ainsi un cadre de calcul rigoureux et exact lorsque les*

1. un engin volant sans pilote

paramètres des systèmes de vision ne sont que partiellement connus ou pas connus du tout.

Le projet Esprit Reactive LTR 21914 CUMULI (Computational Understanding of Multiple Images) dont nous sommes contractant principal nous permet par ailleurs une collaboration étroite avec d'autres équipes européennes universitaires et industrielles.

Reconstruction minimale avec des données manquantes Les données minimales nécessaires pour la reconstruction projective à partir des correspondances de points images est un problème déjà bien connu quand les points sont tous visibles dans toutes les images. Dans ce travail, nous avons d'abord formulé et ensuite résolu une nouvelle famille de problèmes de reconstruction à partir de données minimales qui peuvent contenir les données manquantes dans certaines images. La possibilité de pouvoir traiter les cas minimaux avec des données manquantes a une grande importance tant théorique que pratique: ces méthodes sont inévitables pour pouvoir utiliser les méthodes robustes d'estimation telle que RANSAC et LMS et pour initialier l'estimation optimale telle que l'ajustement des faisceaux.

Nous avons d'abord développé un formalisme pour paramétrer l'ensemble d'images multiples pour traiter les cas de données manquantes. Ensuite nous avons présenté une solution pour le cas minimal de 8 points dans 3 images dans lequel l'un des points est manquant dans l'une des trois images. Nous avons démontré qu'il peut y avoir au plus 11 solutions algébriques pour ce cas minimal. En plus, tous les cas minimaux avec des données manquantes pour 3 et 4 images ont été catalogués.

Ce travail [39] a été réalisé en collaboration avec l'Université de Lund en Suède dans le cadre de notre projet européen commun Cumuli.

Reconstruction d'environnements urbains À la suite d'un projet initié lors d'un séjour post-doctoral d'Edmond Boyer à l'université de Cambridge, et concernant la reconstruction d'environnements urbains, nous avons poursuivi dans cette direction [31]. Nous avons proposé une approche basée sur l'utilisation de contraintes fortes présentes dans la scène : parallélisme et orthogonalité. À l'aide de ces contraintes, il est possible de calibrer les caméras utilisées et de déterminer des modèles Euclidiens de la scène observée. L'originalité de l'approche repose sur l'utilisation de parallélépipèdes pour définir les contraintes. La projection dans une image d'un parallélépipède permet en effet de déterminer plusieurs paramètres de la caméra utilisée. Lorsque le parallélépipède est présent dans plusieurs images, il est alors possible de déterminer les matrices de projections correspondant à ces images. Ces matrices sont utilisées ensuite pour calculer un modèle tridimensionnel partiel de la scène, ce modèle permettant, par exemple, de visualiser la scène suivant de nouveaux points de vue. La méthode qui est proposée ne nécessite aucune information a priori sur les caméras utilisées.

Ambiguïté intrinsèque de la reconstruction projective 2D à partir de 3 images non-calibrées 1D Suite à nos travaux précédents sur la modélisation, l'autocalibration de la caméra projective mono-dimensionnelle, nous avons démontré que la reconstruction projective 2D à partir de trois images non-calibrées 1D a toujours une ambiguïté double indépendamment

du nombre de correspondances de points images qu'on peut avoir [40]. Plus précisément, nous avons prouvé que les deux reconstructions projectives différentes sont exactement reliées par une transformation quadratique avec les trois centres de caméra comme les points fondamentaux de la transformation quadratique. En revanche, la reconstruction unique est possible si et seulement si les trois centres de caméra sont alignés.

Synthèse d'images par interpolation Créer de nouvelles images par interpolation des images existantes a de nombreuses applications en simulation visuelle. Suite à nos travaux précédents, nous avons développé une nouvelle méthode d'interpolation automatique qui s'attaque aux deux problèmes les plus difficiles sachant que l'on n'a pas d'informations de profondeur : la mise en correspondance de pixels et le traitement de la visibilité. Nous avons d'abord proposé un algorithme d'appariement quasiment dense reposant sur une méthode de croissance de région et utilisant une stratégie "meilleur d'abord". Ensuite nous décrivons un algorithme de construction robuste de portions d'images appariées à l'aide de contraintes géométriques locales représentées par une homographie. Enfin, nous introduisons une nouvelle représentation -la triangulation jointe- pour modéliser les portions visibles et partiellement occultées de deux vues et traiter le problème de la visibilité pendant la création de nouvelles vues [49, 37, 38, 36]. La méthode a été démontrée sur de nombreuses paires d'images réelles et est présentée dans la page Web Demo du projet MOVI. Un exemple est illustré sur la figure 1.

Contraintes différentielles d'appariement Les contraintes d'appariement multi-images sont celles dues à la géométrie 3D, qui lie les différentes images d'une seule primitive 3D. Cette année nous avons poursuivi nos travaux sur ces contraintes, avec une étude sur le cas différentiel, où plusieurs des images sont voisines entre elles [47, 22]. Notre approche se base sur un développement par différences finies projectives, ce qui simplifie énormément le problème en comparaison avec les approches antérieures basées sur les séries de Taylor. On détermine systématiquement la forme des contraintes différentielles dans tous les cas où un seul ensemble d'images sont voisines, et on considère comment utiliser ces contraintes pour le « suivi » des correspondances et de la géométrie des caméras le long d'une séquence d'images. On lie ce dernier cas à l'estimation non-linéaire des contraintes non-différentielles.

Mouvements critiques de l'auto-calibrage Depuis une reconstruction projective d'une scène 3D inconnue, on peut retrouver sa structure euclidienne si on dispose de plusieurs connaissances supplémentaires sur les calibrages internes des caméras. On appelle ceci l'« auto-calibrage » si les connaissances ne relèvent pas d'un calibrage numérique classique. Nous avons déjà classifié l'année dernière les cas critiques de l'auto-calibrage à partir de la *constance* des paramètres internes inconnues entre images. Cette année – en collaboration avec l'université de LUND dans le cadre de notre projet européen commun CUMULI – nous avons étudié [35] plusieurs cas où certains paramètres sont connus pour avoir leurs valeurs par défaut (*p.e.* skew 0 et rapport d'échelle 1), mais d'autres sont autorisées à varier entre images (*p.e.* focale).

Calcul de pose et calibrage partiel Pour initialiser les caméras dans nos problèmes de reconstruction 3D, nous avons développé plusieurs méthodes de calcul de la pose, et parfois

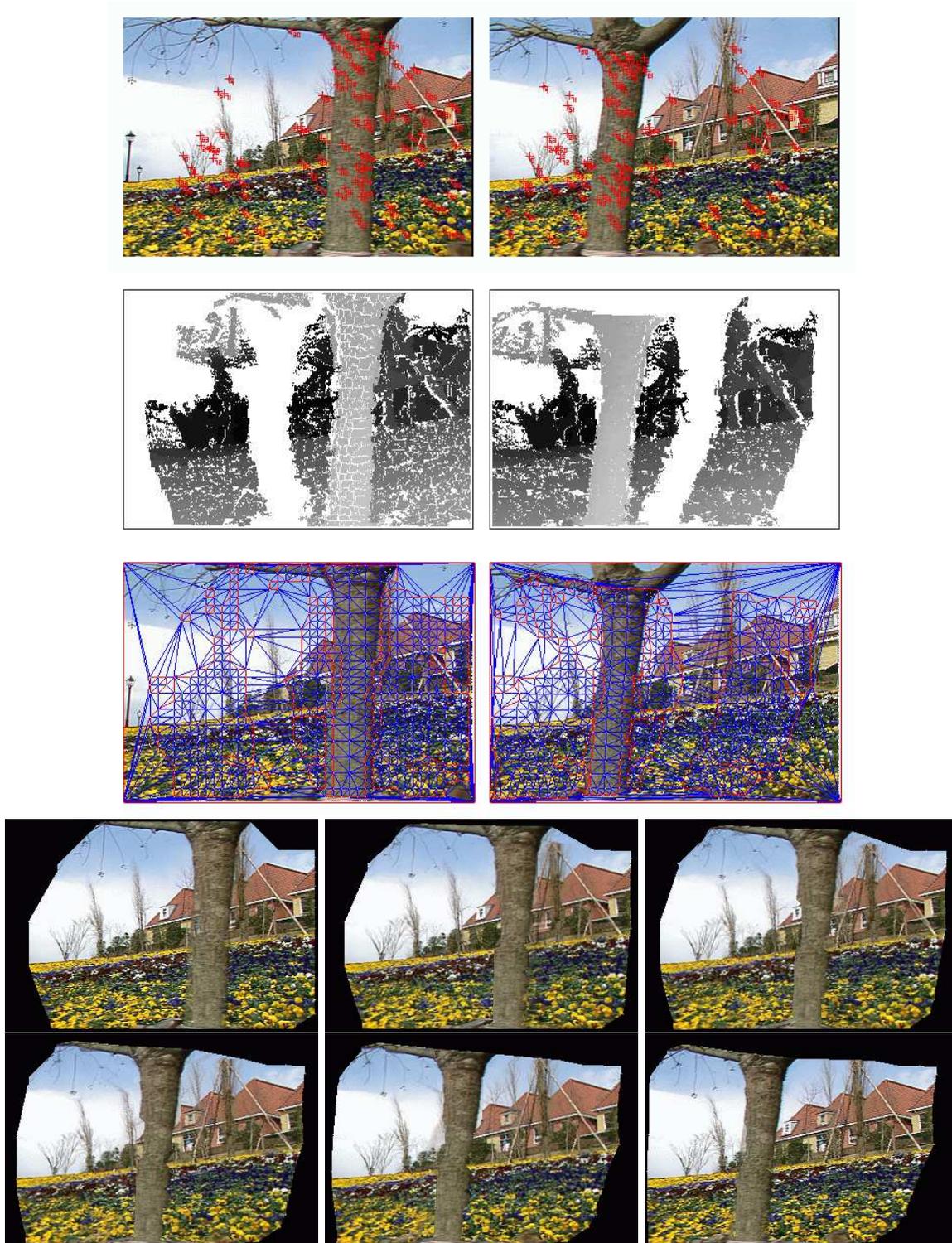


FIG. 1 – De haut en bas: Une paire d'images et l'appariement initial de quelques points ; l'appariement de régions autour de ces points, la triangulation et le résultat final d'interpolation entre les deux images initiales. Le processus d'interpolation brièvement décrit ici est entièrement automatisé.

aussi du calibrage interne, d'une caméra à partir de quelques primitives 3D connues. Ces méthodes font partie d'une bibliothèque en MATLAB, distribuée aux partenaires de notre projet européen CUMULI. Ici, nous ne mentionnons que l'estimation quasi-linéaire de la pose et de plusieurs paramètres internes d'une caméra à partir de $n \geq 4$ points ou de $n \geq 5$ points 3D connus [46, 22]. (4 donnent seulement la focale, 5 la focale et le point principal). La méthode se base sur la théorie des matrices de résultant multi-variables, et généralise la méthode classique de « transformation directe linéaire » au cas où il y a des connaissances partielles sur le calibrage interne de la caméra.

Auto-étalonnage stéréoscopique Ce thème est abordé dans le projet depuis 1997. En 1998 nous avons obtenu d'importants résultats théoriques et expérimentaux. Nous avons développé un nouvel outil mathématique permettant de caractériser d'un point de vue algébrique et géométrique un système formé de deux caméras rigidement liées l'une à l'autre et se déplaçant dans l'espace. Nous avons montré que dans le cas d'un mouvement général la solution est linéaire et il faut alors deux mouvements pour calibrer le système complètement. Nous avons également étudié plusieurs cas particuliers comme les mouvements plans et les mouvements de translation. En 1999 nous avons abordé le cas particulier des rotations pures et nous avons proposé une solution complète pour l'auto-étalonnage projectif et/ou euclidien d'une paire de caméras fixée sur un dispositif rotatif à deux degrés de liberté [41].

Un des problèmes clés concernant l'auto-étalonnage stéréoscopique est l'estimation de l'homographie reliant deux reconstructions projectives obtenues avec le système stéréo à partir de deux positions différentes. Le problème est difficile car, ne connaissant pas a priori le type de mouvement, il n'est pas possible de paramétrer l'homographie. Il faut alors mettre au point des outils statistiques afin d'obtenir la meilleure estimation possible. Nous avons proposé dans un premier temps des méthodes linéaires ainsi qu'une méthode non-linéaire pour estimer cette transformation alors que le système stéréoscopique observe un seul mouvement [24]. Ensuite nous avons étendu cette approche au cas où le système stéréoscopique effectue un mouvement alors qu'il observe plusieurs objets en mouvement [32]. Ceci est illustré sur la figure 2.

Calcul de pose avec des points et des droites Nous avons mis au point des nouveaux algorithmes de pose à partir de correspondances 2-D/3-D de droites [23], de points et de droites [45] ainsi qu'un algorithme linéaire de calcul de pose à partir de 4 ou 5 correspondances de points [26, 25].

Cet algorithme calcule une pose affine qu'il modifie d'une façon itérative pour converger vers une pose perspective. Le nombre minimal de droites requises par la méthode est de 4 dans le cas général et de 3 lorsque les droites se trouvent dans un plan. On montre également que la seule situation dégénérée est celle de droites formant un faisceau.

4.2 Couplage vision/robotique

Participants : Radu Horaud, Bart Lamiroy, Andreas Ruf, Nicolas Andreff, Frédéric Martin.



FIG. 2 – Un capteur stéréoscopique observe plusieurs objets en mouvement alors qu'il bouge lui-même. Sur cette figure on peut voir deux paires d'images extraites d'une séquence de 10 paires d'images, les points extraits de la scène statique (en bas à gauche) ainsi que les deux objets en mouvement (en bas à droite).

Mots clés : asservissement visuel, auto-étalonnage, étalonnage caméra-pince, modélisation de robots.

Résumé : *Les approches classiques en asservissement visuel considèrent le cas d'une caméra étalonnée intervenant dans la boucle d'asservissement d'un robot. Nos travaux, menés en collaboration avec le projet BIP (projet européen VIGOR) et avec les projets BIP, VISTA et ICARE (Action de recherche coopérative ARC/AVEC), s'intéressent au cas de caméras non-étalonnées en posant la question suivante : peut-on faire de l'asservissement visuel sans un étalonnage préalable des caméras ? On étudie l'élargissement du paradigme "asservissement visuel" au cas de deux caméras liées rigidement (couple stéréoscopique). Ces travaux sont intimement liés au problème d'auto-étalonnage d'une paire de caméras. Les travaux théoriques que nous développons montrent qu'on peut représenter la cinématique d'un robot avec des transformations projectives. Il apparaît donc possible de modéliser un robot et ses actions aussi bien dans l'espace projectif qu'eulidien et le passage du projectif à l'eulidien n'est, ni plus ni moins, une forme d'auto-étalonnage de l'ensemble robot-caméras.*

Les expérimentations sont menées en utilisant un robot portique à cinq degrés de liberté sur lequel est fixé une tête stéréoscopique. On dispose également d'un robot manipulateur à six degrés de liberté (cf. figures 3 et 4).

Auto-étalonnage caméra-pince. Les techniques classiques d'étalonnage caméra-pince utilisent une mire d'étalonnage et des mouvements spécifiques de la pince afin que les rotations associées à ces mouvements soient bien conditionnées. Lorsque l'on veut intégrer cet étalonnage dans le processus boucle-fermée de l'asservissement, on ne dispose ni d'une mire ni de la possibilité de choisir les mouvements de la pince. Nous avons intégré la reconstruction 3-D dans le processus d'auto-étalonnage et nous avons développé une méthode qui accepte de faibles déplacements. L'implémentation de cette méthode sur la plate-forme expérimentale du projet donne des résultats comparables à ceux obtenus avec un étalonnage classique [28].

Asservissement visuel avec des droites Dans le cadre de l'asservissement visuel, nous nous sommes penchés sur la modélisation des droites de l'espace. Nous avons montré que les modélisations présentes dans la littérature n'étaient pas idéales. Nous avons dérivé, des coordonnées euclidiennes de Plücker d'une droite, une notion de coordonnées euclidiennes de Plücker normées. Elle permet de définir une nouvelle notion d'alignement des droites aussi bien dans l'image, comme dans les approches précédentes, qu'en orientation dans l'espace. À partir de cet alignement mixte (2D-3D), nous avons pu bâtir une loi de commande explicite et partiellement découplée en rotation et translation. Nous avons ainsi pu faire une analyse globale de la convergence de cette commande et en exhiber les singularités [20]

Cette loi de commande a été appliquée, dans le cadre du projet VIGOR, au positionnement d'une caméra par rapport à un trièdre orthogonal. Cette configuration de droites étant invariante par translation le long de son centre, nous avons proposé une méthode originale d'observation de la profondeur. Elle se fait, par similarité avec la stéréovision, par adjonction

d'un pointeur laser non étalonné. Ce pointeur étant fixe par rapport à la caméra, la projection du point laser dans l'image est contraint à se déplacer le long d'une droite épipolaire. L'étude de ces déplacements permet de définir la loi de commande en profondeur.

Géométrie projective et asservissement visuel. On étudie en détail les propriétés algébriques associées à l'observation d'un mouvement rigide (tel que le mouvement de la pince d'un robot) par un système stéréo non-étalonné. Ces propriétés algébriques permettent de paramétrer l'homographie reliant deux reconstructions projectives obtenues avant et après un mouvement rigide et d'extraire des propriétés euclidiennes dans le cas d'un mouvement général ou planaire et des propriétés affines dans le cas d'un mouvement de translation. Ces travaux sont fortement liés aux travaux fondamentaux dans le domaine de la géométrie pour la vision. En particulier nous avons montré cette année qu'il est possible de modéliser un robot manipulateur couplé à une paire stéréo dans un espace non métrique. Ceci permet de représenter des mécanismes tels que le robot dans l'espace visuel projectif associé avec une paire de caméras faiblement calibrées (seule la géométrie épipolaire est connue). Nous avons complètement caractérisé les propriétés algébriques et différentielles des transformations projectives agissant sur cet espace et nous avons mis en évidence la structure de groupe de Lie de ces transformations.

Ces travaux ont donné lieu à quatre publications [42], [50], [27] et [43].

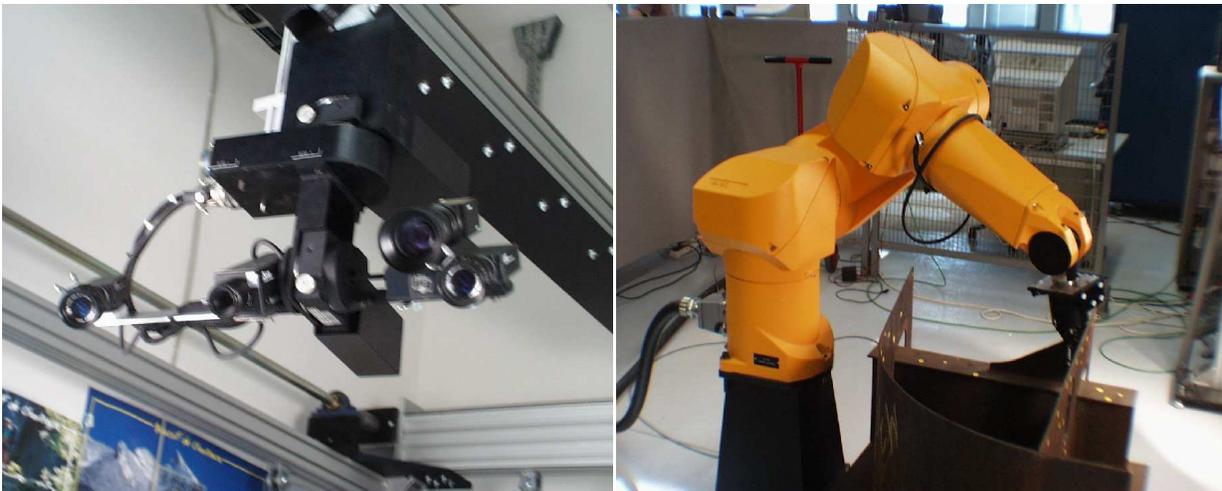


FIG. 3 – La plate-forme expérimentale du couplage vision-robotique comportant une tête stéréo montée sur un robot portique à 5 degrés de liberté (gauche) et un robot manipulateur à 6 degrés de liberté (droite). La pièce métallique visible dans l'image de droite est une pièce de bateau fabriquée par les chantiers navals d'Odense.

4.3 Indexation d'images et reconnaissance d'objets

Participants : Christophe Biernacki, Yves Dufournaud, Patrick Gros, Riad Hammoud, Radu Horaud, Krystian Mikolajczyk, Roger Mohr, Fabien Pelisson, Sylvaine Picard, Cordelia Schmid.

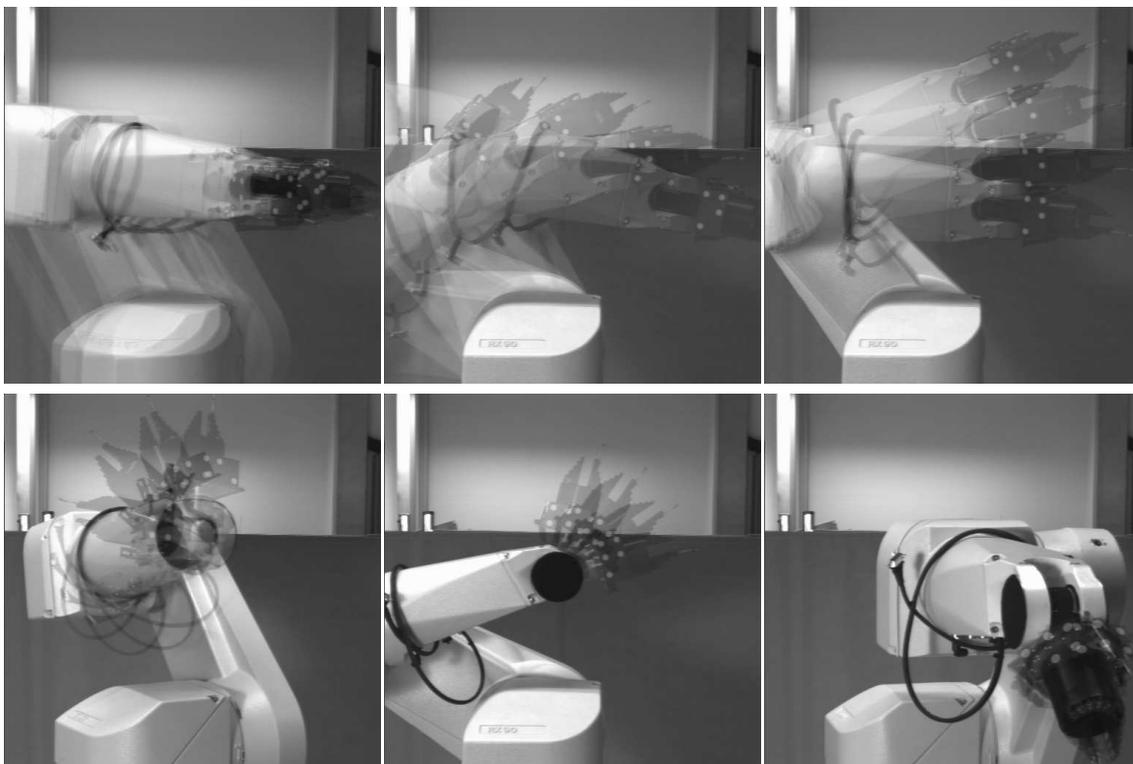


FIG. 4 – Cette figure montre une séquence d’auto-étalonnage du robot. Celui-ci effectue des mouvements articulaires (une articulation à la fois) devant le capteur stéréoscopique. On obtient ainsi un modèle de “cinématique projective” du robot à partir duquel on peut extraire les paramètres euclidiens du robot ainsi que les paramètres du capteur.

Mots clés : base multimédia, classification, indexation d’images, interprétation d’images, invariant géométrique et photométrique, mise en correspondance d’images, reconnaissance d’objets, vision par ordinateur.

Résumé : *L’appariement et l’indexation des images sont l’un des axes importants de recherche du projet. Cette activité s’est poursuivie selon plusieurs directions : appariement entre images prises de points de vue différents, développement de descripteurs robustes, robustification de l’indexation par l’utilisation de modèles statistiques, enfin la généralisation des descripteurs pour pouvoir reconnaître non seulement des objets, mais aussi des classes d’objets.*

Appariement entre images prises de points de vue différents L’appariement entre deux images en présence de changements d’échelle importants et de changements perspectifs est un problème difficile [29]. Une première approche consiste à utiliser une approche multi-échelle pour l’ensemble des descripteurs et pour l’extraction de caractéristiques, endroits où sont calculés les descripteurs. L’utilisation des contraintes géométriques entre images et l’estimation

robuste des ces contraintes permet un rejet important des faux appariements. L'approche développée permet d'estimer le rapport d'échelle entre deux images sans avoir recours à une estimation initiale. Une étude expérimentale - grâce aux données fournies par l'Aérospatiale - montre d'excellents résultats (voir Figure 5).

Développement de descripteurs robusts Nous avons poursuivi nos travaux sur les descripteurs non paramétriques. Ces descripteurs visent à s'affranchir des changements d'illumination tout en ne nécessitant pas de modèles de ces changements: il suffit que ceux-ci soient monotones croissants. Suivant une idée originale de Zabih et Woodfill, nous avons proposé une série de descripteurs invariants à la rotation, aux changements d'échelle et d'illumination. Les premiers sont une adaptation du descripteurs rank pour le rendre invariants aux rotations, les autres sont basés sur des accumulations sous forme d'histogrammes de code de comparaison entre pixels. Une étude expérimentale poussée a été menée par les valider, et les résultats obtenus sont très positifs [21].

Robustification de l'indexation à l'aide d'une modélisation statistique L'ajout d'un modèle probabiliste à l'algorithme d'indexation permet d'améliorer de façon significative les résultats de reconnaissance [44], [48]. Ce modèle tient compte de la variabilité des descripteurs, de leur pertinence et également de la probabilité de la cohérence spatiale entre descripteurs. Il respecte en outre les différents niveaux d'abstraction des étapes du processus de reconnaissance.

Descripteurs pour la recherche de classes d'objets Jusqu'à présent le but de nos recherches était de retrouver pour une image requête l'objet correspondant dans une base d'images comportant un grand nombre d'images (1000). Il s'agit maintenant d'étendre ces travaux à des classes d'objets. On souhaite retrouver des objets comme par exemple des visages, des chevaux, des chaises etc. Ceci est important pour des applications telles qu'il en existe par exemple dans les agences de presse. Ces agences souhaitent pouvoir répondre à des requêtes comme trouver quelques images d'enfants qui jouent.

Pour définir de telles classes d'objets nous avons étudié la ressemblance entre des contours pour des objets d'une même classe. Nous avons obtenu de bons résultats pour des classes d'objet telles que les visages. En outre, notre méthode est invariante aux transformations de similitude et robuste aux occultations.

Nous avons également commencé à définir des descripteurs génériques à partir des invariants locaux. À partir d'un ensemble de visages, nous avons construit des descripteurs génériques pour des caractéristiques tel que bouche, nez et oeil. Ces descripteurs génériques capturent la variabilité des descripteurs pour l'ensemble des images. Pour ce faire on utilise des mélanges de Gaussiennes. La détection de visages à partir de ces descripteurs montre de bons résultats. Reste à déterminer les endroits caractéristiques de façon automatique.

4.4 Structuration de vidéos

Participants : Christophe Biernacki, Patrick Gros, Riad Hammoud, Krystian Mikolajczyk, Roger Mohr, Cordelia Schmid.

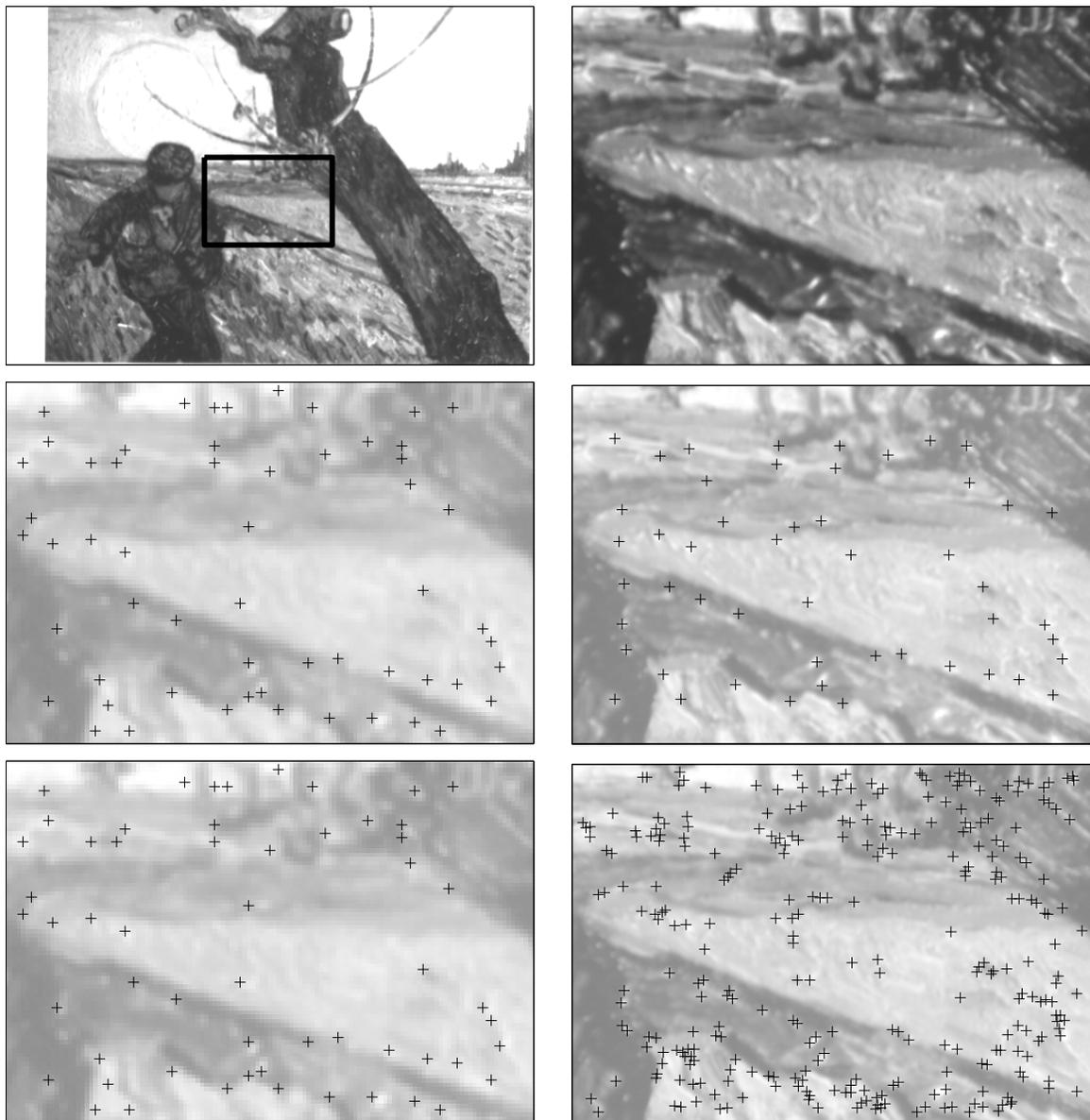


FIG. 5 – Résultats d’adaptation du détecteur de points d’intérêt sur la séquence Van Gogh. On peut voir la paire d’images utilisées (en haut), les points détectés par la version adaptée (au milieu), les points détectés par la version standard (en bas). Les images de gauche (milieu et bas) correspondent à la partie encadrée dans l’image du haut, agrandie pour une meilleure visualisation. Le facteur d’échelle est ici de 4.17.

Résumé : *Pour pouvoir manipuler l'information du contenu d'une vidéo, il faut la structurer. Ces structures permettent l'interactivité de l'utilisateur, la recherche d'informations pertinentes, ou plus simplement un parcours adapté au besoin de l'utilisateur. Si structurer automatiquement les vidéos en leur contenu sémantique est actuellement hors de portée de nos programmes, il reste en revanche possible d'établir des liens entre les objets détectés, de découper les objets de la vidéo de façon semi-automatique, et à partir de ces éléments d'autoriser l'éditeur à créer la structure recherchée.*

Structuration de la vidéo Deux grands types d'applications requièrent la structuration des vidéos ; ces structurations sont actuellement réalisées avec des outils rudimentaires et elles demandent des dizaines heures de travail pour seulement quelques minutes de vidéo ; ces types d'applications sont :

- la création de vidéo interactive, tant dans le domaine éducatif que dans celui de la documentation assistée. Les objets doivent être délimités, et à chaque objet doit être assignée une action qui va s'exécuter lors de la désignation à la visualisation.
- la recherche dans les vidéos : cette recherche se fera par une consultation dans un système documentaire qui contiendra les informations extraites des vidéos, mais dans un deuxième temps la structure sera utilisée pour permettre des parcours rapide comme la recherche de tous les personnages, ou du prochain plan où cette scène apparaît.

Nous avons développé une interface pour la structuration de la vidéo en coopération avec l'IRISA, projet VISTA qui travaille sur la partie mouvement de cette application ; ce travail est soutenu par la société Alcatel et par le RNRT [30].

Descripteurs adaptés à la vidéo Pour mettre en correspondance des objets entre différents plans, il faut choisir des descripteurs adaptés à ce problème. Pour cela on détermine les descripteurs les mieux adaptés à chaque classe d'objet. Puis, on modélise leurs variations temporelles à l'intérieur d'un plan pour mieux pouvoir les mettre en correspondance [34, 33].

5 Contrats industriels (nationaux, européens et internationaux)

5.1 Reconnaissance et indexation de vues aériennes

Participants : Radu Horaud, Cordelia Schmid, Yves Dufournaud.

Nous collaborons avec la Société Aérospatiale depuis 1993. Depuis 1998 et pour une durée de trois ans nous avons un contrat de collaboration dans le domaine de l'interprétation d'images aériennes.

Yves Dufournaud effectue une thèse de doctorat dans le cadre d'une convention CIFRE. L'objectif de ce travail est d'étudier, expérimenter et réaliser un logiciel permettant de comparer des images prises dans des conditions de vue très différentes en termes de la focale utilisée.

Les travaux effectués à ce jour sont très encourageants et nous avons mis en place un transfert de savoir-faire méthodologique.

5.2 Indexation de vidéo

Participants : Patrick Gros, Riad Hammoud, Roger Mohr, Cordelia Schmid.

La société Alcatel, à travers sa division Alcatel-CRC, a lancé une action de partenariat avec l'INRIA sur le sujet de l'indexation vidéo (autres partenaires de l'institut : le projet VISTA). L'objet de cette action est la réalisation d'outils de base permettant d'indexer et structurer de façon automatique des bandes de vidéo, en se limitant aux images.

Dans cette opération, notre projet est principalement acteur sur le calcul de signatures d'indexation, et de la réalisation de liens entre plans de la vidéo. Une démonstration est en cours de réalisation.

Par ailleurs, Alcatel finance la bourse CIES de R. Hammoud, thèse qui sera soutenue en 2000.

Divers contacts ont été noués avec des PME dans le domaine de l'indexation des images fixes : Alma, Caldera et SGBI.

5.3 Projet RNRT "AGIR"

Participants : Patrick Gros, Riad Hammoud, Krystian Mikolajczyk, Cordelia Schmid.

Le projet AGIR est un projet RNRT précompétitif dont le début a été fixé au 1er septembre 1999. Il regroupe le centre de recherche d'Alcatel, une PME (Arts Video Interactive), l'INA, et des équipes de recherches (IRISA, INRIA Rhône-Alpes, IRIT, CLIPS-IMAG, INT, LIP6). Ce projet est financé pour 18 mois, avec un renouvellement possible de 12 mois.

Le but du projet est de développer un système complet permettant d'indexer des vidéos, éventuellement accompagnée de textes, puis d'utiliser cette indexation pour rechercher des documents par leur contenu, ceci en mélangeant les divers médias. Ce projet prend donc naturellement la suite du travail fait dans le cadre du projet européen DIVAN, avec un accent plus fort mis sur l'aspect multimédia des données, des index et des requêtes, sur l'aspect architecture globale du projet et sur les liens avec les efforts de normalisation actuels (MPEG7).

Les travaux développés par MOVI concerneront la fourniture de descripteurs d'images. Ces descripteurs, calculés à partir de diverses primitives (points, régions, visages) doivent permettre l'identification d'objets ou de scènes. Un effort spécifique sera fourni pour adapter les techniques utilisées au cas de la vidéo où la redondance des images permet une robustification des descripteurs. Cela sera fait en reprenant les travaux de R. Hammoud sur le calcul de modèles de descripteurs par mélange de gaussiennes. Sera aussi étudié le problème de la composition spatiale des descripteurs élémentaires pour obtenir un langage de requête plus riche.

6 Actions régionales, nationales et internationales

6.1 Actions régionales

Le projet est impliqué dans une action régionale ACTIV qui fédère onze groupes de recherche régionaux en vision par ordinateur et qui aborde les thèmes de catégorisation et segmentation

d'images, de mesure et rendu de la couleur, et d'indexation d'images. R. Mohr est coordinateur de ce dernier thème.

6.2 Actions nationales

- Le projet VISIR est une action incitative du ministère de la recherche qui regroupe 5 équipes de recherche (VISTA, MOVI, SIGMA2, IHMPT de l'IRIT, et MRIM de CLIPS-IMAG) et l'INA autour du thème de l'indexation multimédia. Elle est dirigée par P. Gros. Cette action, qui ne regroupe donc que des partenaires impliqués dans le projet AGIR, sera consacrée à développer des techniques plus en amont de celles qui seront implémentées dans le projet AGIR qui a une visée plus applicative. Seront particulièrement concernées les techniques permettant de faire collaborer descripteurs et techniques issus de médias différents pour améliorer les résultats de recherche par le contenu.
- L. Quan participe à l'animation du GT5 dans le GDR ISIS.
- Le projet participe également à l'ARC (Action de recherche coopérative) AVEC (Asservissement visuel en environnements complexes). Cette action a débuté en 1998 et se prolonge jusqu'en 2000. Les autres partenaires sont les projets BIP, VISTA (Rennes) et ICARE (Sophia-Antipolis). Voir (<http://www.irisa.fr/AVEC/>)

6.3 Actions financées par la Commission Européenne

CUMULI. Nous sommes coordinateur du projet LTR-21914 européen CUMULI (<http://www.inrialpes.fr/CUMULI>).

Ce projet s'exécute en liaison avec nos partenaires de l'université de Lund, de l'institut Fraunhofer de Darmstadt, du projet ROBOTVIS de Sophia-Antipolis, et de deux partenaires industriels, les sociétés IMETRIC en Suisse et INNOVATIV VISION IMAGE SYSTEMS en Suède. Le but de cette collaboration est d'affiner et d'implanter dans l'industrie nos compétences sur la mesure précise à partir d'images multiples, depuis l'auto-étalonnage jusqu'à la reconstruction tridimensionnelle ou l'estimation de trajectoires 3D.

VIGOR. Nous sommes coordinateur du projet européen Reactive LTR VIGOR. (<http://www.inrialpes.fr/VIGOR/>). Ce projet auquel participe également le projet BIP s'exécute en partenariat avec deux équipes universitaires (University of Cambridge et Hebrew University of Jerusalem), un institut de recherche (IITB) et deux partenaires industriels (Sinters SA et Odense Steel Shipyard). D'une durée de trois ans et ayant commencé le 1 février 1998, ce projet a comme objectifs d'étendre le concept de caméras non-calibrées au guidage visuel de robot et de proposer un nouveau concept où les caméras et les robots sont considérés ensemble pour des applications navales et aéronautiques.

VIMINI. Nous sommes partenaire du projet européen Open LTR VIMINI (<http://www.dist.unige.it/vimini/>). La première phase de ce projet s'exécute en partenariat avec une équipe universitaire (Universita di Genoa) et un industriel (ELSAG Spa). Ce projet a comme objectif l'étude de faisabilité d'une "souris interactive", deux caméras disposées de part et d'autre

d'un écran d'ordinateur observent et mesurent les mouvements de la tête et les mouvements des yeux. Ces mouvements pourraient ensuite être utilisés pour agir sur un pointeur écran. Applications possibles : aide aux personnes handicapées et travail "mains libres". D'une durée de 6 mois, la demande de prolongation de ce projet a été déposée le 30 septembre 1999.

Bourses TMR. MOVI accueille :

- Adreas Ruf (bourse TMR de type 20) pour une durée de trois ans (1997–2000),
- Alberto Aguado (bourse TMR de type 30) pour une durée de deux ans (1997–1999).

6.4 Relations bilatérales internationales

6.4.1 Amérique

Dans le cadre du programme de la collaboration CNRS/UIUC (The Univ. of Illinois, U.S.A.), nous menons un projet avec l'équipe de J. Ponce sur le rendu d'images à base d'images. En 1999 les professeurs Jean Ponce et David Kriegman nous ont rendu visite pendant une semaine. David Demirdjian et Maxime Lhuillier ont effectué une visite de trois jours à l'Université d'Illinois.

6.4.2 Asie

Dans le cadre du laboratoire Franco-chinois, nous avons un projet de collaboration avec l'académie de science de Chine sur la sélection de modèles pour la reconstruction à partir d'images.

Une collaboration de plusieurs années se poursuit avec l'université Xidian de Xi'an (professeur Wu Chengke). Elle porte sur l'utilisation des outils géométriques pour la compression d'images et pour la synthèse d'images à partir d'images. En 1998 L. Quan y a fait un séjour de trois semaines. En 1999 le professeur Wu Chengke séjourne à Montbonnot pour une durée de deux mois.

7 Diffusion de résultats

7.1 Animation de la communauté scientifique

Les membres du projet font partie des comités de rédaction de revues suivantes :

- *Traitement du Signal: Signal, Image et Parole* (R. Horaud et P. Gros),
- *Pattern Recognition Journal* (R. Mohr, R. Horaud),
- *Machine Vision and Applications* (R. Mohr),
- *International Journal of Robotics Research* (R. Horaud),
- *Computacion y Sistemas* (R. Horaud),

- *Journal of Images and Graphics* (L. Quan est "overseas associated editor").

Les membres du projet font partie des comités de programme des conférences suivantes :

- IJCAI'99 (R. Horaud)
- ICCV'99 (R. Horaud et R. Mohr),
- CVPR'99 (R. Horaud, C. Schmid, L. Quan, W. Triggs)
- VISION INTERFACE'99 (R. Mohr),
- VISION ALGORITHMS'99 (W. Triggs, chairman et L. Quan) [19],
- CAIP'99 (R. Mohr),
- Workshop on the Integration of Appearance and Geometric Methods in Object Recognition—CVPR99 (C. Schmid),
- P. Gros était responsable du comité d'organisation et de programme des journées ORASIS'99 réunissant les jeunes chercheurs en vision par ordinateur et perception visuelle,
- R. Mohr était membre du comité de programme et P. Gros membre du comité d'organisation du workshop international CBMI'99 qui a eu lieu en octobre à Toulouse.

7.2 Enseignement universitaire

- Outils mathématiques pour la vision et la robotique, DEA IVR, INPG, 12h, P. Gros, L. Quan et R. Mohr
- Vision 3D, DEA IVR, INPG, 18h, R. Horaud
- Applications de la vision à la robotique, DEA SIP, INPG, 30h, R. Horaud.
- Introduction à la vision par ordinateur, magistère informatique, université Joseph Fourier, 15h, R. Horaud.
- Vision 3D INPG - ENSIEG, 20h, L. Quan
- École Polytechnique, Palaiseau, 2e et 3e année (P. Gros : algorithmique, imagerie tridimensionnelle)
- ENSTA, Paris, 3e année (P. Gros : géométrie projective et vision 3D)

7.3 Participation à des colloques, séminaires, invitations

Les membres du projet ont été invités à faire des présentations aux manifestations suivantes :

- Radu Horaud a été conférencier invité (the computer vision invited talk) à IJCAI'99, Stockholm, Août 1999,
- Long Quan a été nommé “overseas assessor” pour la Chinese Academy of Sciences pour une durée de trois ans,
- Patrick Gros a été invité au séminaire « Content-Based Image and Video Indexing » organisé par le centre de conférence de Dagstuhl en Allemagne, décembre 1999,
- Cordelia Schmid a fait une présentation invitée à Microsoft, mars 1999,
- William Triggs a fait des présentations invitées à Microsoft (2 fois), au workshop Mview'99 et à KTH en Suède.

8 Bibliographie

Ouvrages et articles de référence de l'équipe

- [1] S. CHRISTY, R. HORAUD, « Euclidean Shape and Motion from Multiple Perspective Views by Affine Iterations », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 11, November 1996, p. 1098–1104, <ftp://ftp.inrialpes.fr/pub/movi/publications/rec-affiter-long.ps.gz>.
- [2] P. GROS, O. BOURNEZ, E. BOYER, « Using Local Planar Geometric Invariants to Match and Model Images of Line Segments », *Computer Vision and Image Understanding* 69, 2, 1998, p. 135–155.
- [3] R. HARTLEY, P. STURM, « Triangulation », *Computer Vision and Image Understanding* 68, 2, 1997, p. 146–157.
- [4] L. HÉRAULT, R. HORAUD, « Figure-ground discrimination: a combinatorial optimization approach », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15, 9, September 1993, p. 899–914.
- [5] R. HORAUD, F. DORNAIKA, B. ESPIAU, « Visually Guided Object Grasping », *IEEE Transactions on Robotics and Automation* 14, 4, August 1998, p. 525–532.
- [6] R. HORAUD, F. DORNAIKA, B. LAMIROY, S. CHRISTY, « Object Pose: The Link between Weak Perspective, Paraperspective, and Full Perspective », *International Journal of Computer Vision* 22, 2, March 1997, p. 173–189.
- [7] R. HORAUD, F. DORNAIKA, « Hand-Eye Calibration », *International Journal of Robotics Research* 14, 3, June 1995, p. 195–210.
- [8] R. HORAUD, O. MONGA, *Vision par ordinateur: outils fondamentaux*, Editions Hermès, Paris, 1995, *Deuxième édition revue et augmentée*.

- [9] R. MOHR, B. BOUFAMA, P. BRAND, « Understanding Positioning from Multiple Images », *Artificial Intelligence* 78, 1995, p. 213–238.
- [10] L. QUAN, T. KANADE, « Affine Structure from Line Correspondences with Uncalibrated Affine Cameras », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 8, août 1997, p. 834–845.
- [11] L. QUAN, « Invariants of Six Points and Projective Reconstruction from Three Uncalibrated Images », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17, 1, janvier 1995, p. 34–46.
- [12] L. QUAN, « Conic Reconstruction and Correspondence from Two Views », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 2, février 1996, p. 151–160.
- [13] L. QUAN, « Self-Calibration of An Affine Camera from Multiple Views », *International Journal of Computer Vision* 19, 1, mai 1996, p. 93–105.
- [14] C. SCHMID, R. MOHR, « Object Recognition Using Local Characterization and Semi-Local Constraints », *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 5, mai 1997, p. 530–534.
- [15] P. STURM, B. TRIGGS, « A Factorization Based Algorithm for Multi-Image Projective Structure and Motion », *in: Proceedings of the 4th European Conference on Computer Vision, Cambridge, England*, p. 709–720, Avril 1996.
- [16] P. STURM, « Critical Motion Sequences for Monocular Self-Calibration and Uncalibrated Euclidean Reconstruction », *in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Puerto Rico*, p. 1100–1105, Juin 1997.
- [17] B. TRIGGS, « Matching Constraints and the Joint Image », *in: IEEE Int. Conf. Computer Vision*, E. Grimson (éditeur), p. 338–43, Cambridge, MA, juin 1995.
- [18] B. TRIGGS, « Autocalibration and the Absolute Quadric », *in: IEEE Conf. Computer Vision & Pattern Recognition*, Puerto Rico, 1997.

Livres et monographies

- [19] B. TRIGGS, R. SZELISKI, A. ZISSERMAN (éditeurs), *Vision Algorithms: Theory and Practice*, Workshop held during ICCV'99 in Kerkyra (Corfu), Greece, INRIA, September 1999. Final proceedings will appear in Springer LNCS in early 2000.

Thèses et habilitations à diriger des recherches

- [20] N. ANDREFF, *Asservissement visuel à partir de droites et auto-étalonnage pince-caméra*, thèse de doctorat, Institut National Polytechnique de Grenoble, November 1999.
- [21] S. PICARD, *Etude de descripteur non paramétriques pour l'indexation d'images par le contenu*, Thèse de doctorat, Institut National Polytechnique de Grenoble, Grenoble, November 1999.
- [22] B. TRIGGS, *Géométrie d'images multiples*, thèse de doctorat, Institut National Polytechnique de Grenoble, November 1999.

Articles et chapitres de livre

- [23] S. CHRISTY, R. HORAUD, «Iterative Pose Computation from Line Correspondences», *Computer Vision and Image Understanding* 73, 1, January 1999, p. 137–144.
- [24] G. CSURKA, D. DEMIRDJIAN, R. HORAUD, «Finding the Collineation Between two Projective Reconstructions», *Computer Vision and Image Understanding* 75, 3, September 1999, p. 260–268, <ftp://ftp.inrialpes.fr/pub/movi/publications/GabrielaDavidRadu-cviu99.ps.gz>.
- [25] Y. OHTA, G. INOUE, T. KOBAYASHI, L. QUAN, «Vision-Based Geometric Registration of Virtual and Real Worlds», in : *Mixed Reality—Merging Real and Virtual Worlds*, Y. Ohta et H. Tamura (éditeurs), Springer-Verlag, 1999, ch. 5, p. 85–100.
- [26] L. QUAN, Z. LAN, «Linear N-Point Camera Pose Determination», *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 8, August 1999, p. 774–780.
- [27] A. RUF, R. HORAUD, «Visual Servoing of Robot Manipulators, Part I: Projective Kinematics», *International Journal of Robotics Research* 18, 11, November 1999, p. 1101–1118, <ftp://ftp.inria.fr/INRIA/publication/RR/RR-3670.ps.gz>.

Communications à des congrès, colloques, etc.

- [28] N. ANDREFF, R. HORAUD, B. ESPIAU, «On-line Hand-Eye Calibration», in : *Second International Conference on 3-D Digital Imaging and Mode ling (3DIM'99)*, p. 430–436, Ottawa, October 1999.
- [29] C. BAILLARD, C. SCHMID, A. ZISSERMAN, A. FITZGIBBON, «Automatic line matching and 3D reconstruction of buildings from multiple views», in : *ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, IAPRS Vol.32, Part 3-2W5*, p. 69–80, 1999.
- [30] P. BOUTHEMY, Y. DUFOURNAUD, R. FABLET, R. MOHR, S. PELEG, A. ZOMET, «Video Hyper-link Creation for Content-Based Browsing and Navigation», in : *Workshop on Content-Based Multimedia Indexing*, October 1999.
- [31] R. CIPOLLA, D. ROBERTSON, E. BOYER, «Photobuilder – 3D models of architectural scenes from uncalibrated images», in : *IEEE International Conference on Multimedia Computing and Systems, Firenze, I*, p. 25–31, June 1999.
- [32] D. DEMIRDJIAN, R. HORAUD, «A Projective Framework for Scene Segmentation in the Presence of Moving Objects», in : *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society Press, p. 2–8, Fort Collins, Colorado, June 1999, <ftp://ftp.inrialpes.fr/pub/movi/publications/David-cvpr99.ps.gz>.
- [33] R. HAMMOUD, L. CHEN, «A spatiotemporal approach for semantic video macro-segmentation», in : *European Workshop on Content-Based Multimedia Indexing*, p. 195–201, IRIT-Toulouse FRANCE, Octobre 1999.
- [34] R. HAMMOUD, R. MOHR, C. BIERNACKI, «Robustification des signatures de couleurs par modélisation de le urs variabilités intra -plan vidéo », in : *Dix-septième Colloque Gretsi sur le Traitement du Signal et d es Images*, p. 279–282, Vannes-France, Septembre 1999.
- [35] F. KAHL, B. TRIGGS, «Critical Motions in Euclidean Structure from Motion», in : *Proceedings of the Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA*, p. 366–372, 1999.

-
- [36] M. LHUILLIER, L. QUAN, R. MOHR, « Automatic View Interpolation », *in: Proceedings in ISPRS Commission V, International Workshop on Vision-Based Techniques in Visualization and Animation*, The Japan Society of Photogrammetry and Remote Sensing, p. 107–113, October 1999.
- [37] M. LHUILLIER, L. QUAN, « Image Interpolation by Joint View Triangulation », *in: Proceedings of the Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA, 2*, p. 139–145, 1999.
- [38] M. LHUILLIER, « Joint View Triangulation for Two Views », *in: Proceedings of the 12th Conference on Vision Interface, Trois Rivières. Québec*, may 1999.
- [39] L. QUAN, A. HEYDEN, F. KAHL, « Minimal Projective Reconstruction with Missing Data », *in: Proceedings of the Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA*, IEEE Computer Society, p. 210–216, June 1999.
- [40] L. QUAN, « Inherent Two-Way Ambiguity in 2D Projective Reconstruction from Three Uncalibrated 1D Images », *in: Proceedings of the 7th International Conference on Computer Vision, Corfu, Greece*, p. 344–349, 1999.
- [41] A. RUF, R. HORAUD, « Projective Rotations Applied to a Pan-tilt Stereo Head », *in: IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society Press, p. 144–150, Fort Collins, Colorado, June 1999, <ftp://ftp.inrialpes.fr/pub/movi/publications/Andreas-cvpr99.ps.gz>.
- [42] A. RUF, R. HORAUD, « Rigid and Articulated Motion Seen with an Uncalibrated Stereo Rig », *in: IEEE Seventh International Conference on Computer Vision*, IEEE Computer Society Press, p. 789–796, Corfu, Greece, September 1999, <ftp://ftp.inrialpes.fr/pub/movi/publications/RufHoraud-iccv99.ps.gz>.
- [43] A. RUF, F. MARTIN, B. LAMIROY, R. HORAUD, « Visual Control Using Projective Kinematics », *in: Proceedings of the Ninth International Symposium on Robotics Research (ISRR 1999)*, p. 66–73, Snowbird, Utah, USA, October 1999, <ftp://ftp.inrialpes.fr/pub/movi/publications/ruf99g.ps.gz>.
- [44] C. SCHMID, « A Structured Probabilistic Model for Recognition », *in: Proceedings of the Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA, II*, p. 485–490, June 1999.
- [45] P. STURM, S. MAYBANK, « On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications », *in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, USA*, p. 432–437, Juin 1999.
- [46] B. TRIGGS, « Camera Pose and Calibration from 4 or 5 known 3D Points », *in: Proceedings of the 7th International Conference on Computer Vision, Kerkyra, Greece*, p. 278–284, Kerkyra, Greece, September 1999.
- [47] B. TRIGGS, « Differential Matching Constraints », *in: Proceedings of the 7th International Conference on Computer Vision, Kerkyra, Greece*, IEEE press, p. 370–376, Kerkyra, Greece, September 1999.

Rapports de recherche et publications internes

- [48] C. BIERNACKI, R. MOHR, «Indexation et Appariement d'Images par Modèles de Mélanges Gaussien des Couleurs», *rapport de recherche*, INRIA, January 1999, <http://www.inria.fr/RRRT/RR-3600.html>.
- [49] M. LHUILLIER, «Towards Automatic Interpolation for Real and Distant Image Pairs», *rapport de recherche*, Inria, february 1999, <http://www.inria.fr/RRRT/RR-3619.html>.
- [50] A. RUF, R. HORAUD, «Visual Servoing of Robot Manipulators, Part I: Projective Kinematics», *rapport de recherche n° RR-3670*, INRIA, April 1999, To appear in International Journal of Robotics Research, <ftp://ftp.inria.fr/INRIA/publication/RR/RR-3670.ps.gz>.