

---

# Projet MOVI

## Modélisation, localisation, reconnaissance et interprétation en vision par ordinateur

---

**Localisation :** *Grenoble*

**Mots-clés :** bases d'images, calibration de caméra, commande référencée capteur, géométrie, indexation d'images, mise en correspondance, reconstruction 3D, séquence d'images, vision 3D, vision par ordinateur.

### 1 Composition de l'équipe

#### Responsable Scientifique

Roger Mohr, professeur, ENSIMAG

#### Secrétariat

Danièle Herzog, INRIA

#### Personnel des établissements partenaires (UMR GRAVIR-Imag)

Patrick Gros, chargé de recherche CNRS  
Radu Horaud, directeur de recherche CNRS  
Long Quan, chargé de recherche CNRS

#### Chercheur invité

Gerard McLean, professeur, Victoria University, Canada

#### Chercheurs post-doctorants

Gabriela Csurka, CIES  
Pascal Brand, post-doc industriel avec ITMI

#### Chercheurs doctorants

Nicolas Andreff, MENESR (en commun avec le projet BIP)  
Jérôme Blanc, MENESR  
Stéphane Christy, DRET  
Katja Daumüller-Claudé, TMR  
Bart Lamiroy, MENESR  
Zhong-Dan Lan, CIES  
Gustavo Olague, CONACYT  
Sylvaine Picard, DRET (à partir de sept. 96)

Andreas Ruf, MENESR (à partir de novembre 96)  
Cordelia Schmid, INRIA (jusqu'en juillet 96)  
Peter Sturm, MENESR  
William Triggs, INRIA

## 2 Présentation du projet

Le projet *Movi* est un projet commun entre le CNRS, l'INPG, l'UJF et l'Inria, localisé à l'Inria Rhône-Alpes et intégré dans le laboratoire GRAVIR de la fédération IMAG.

Identifier l'espace tridimensionnel et les objets qu'il contient à l'aide de caméras est notre axe de travail. Dans ce cadre, notre projet scientifique s'articule autour d'un thème majeur, la modélisation géométrique pour des objectifs de reconnaissance, de localisation, d'action. Dans ce cadre, nous approfondissons plus précisément les points suivants :

- l'étalonnage (détermination des modèles géométriques) de capteurs ;
- l'étude du couplage entre la vision et le contrôle d'un robot : calibration du couplage caméra-robot, asservissement visuel, manipulation d'objets et saisie d'outils guidés par la vision ;
- les mises en correspondance entre couples d'images afin d'appréhender finement l'espace tridimensionnel, entre image et modèle afin d'identifier et localiser des formes ;
- les algorithmes de reconstruction tridimensionnelle selon différentes hypothèses d'acquisition d'images (caméras étalonnées ou non, mouvement connu ou non) ;
- l'apprentissage de modèles visuels à des fins de reconnaissance ou de recherche dans des bases d'images.

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 des différents savoir-faire, cette intégration amène à reposer les problèmes fondamentaux dans des cadres nouveaux, comme par exemple celui de la commande par la vision de mouvement de bras de robots ou la création de bases d'images interrogeables par leur contenu.

## 3 Action de recherche

### 3.1 Géométrie

*Participants* : Jérôme Blanc, Stéphane Christy, Gabriella Csurka, Radu Horaud, Roger Mohr, Gustavo Olague, Long Quan, Peter Sturm, William Triggs

*Mots-clés* : calibration de caméra, géométrie, vision 3D, invariant géométrique, stéréovision.

Nous avons été présents depuis plus de quatre ans dans les travaux sur l'utilisation de la géométrie projective pour la vision tridimensionnelle. 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 paramètres des systèmes de vision ne sont que partiellement et/ou imparfaitement connus. Par ailleurs, la plupart des applications ont en dernier ressort besoin des connaissances euclidiennes (distances, angles, etc.). Nous nous attachons donc au passage projectif – euclidien, tout comme aux estimations dans cette dernière géométrie. Le démarrage du projet LTR CUMULI –dont nous sommes les coordinateurs– fixe pour les prochaines années nos orientations dans le domaine : prise en compte de primitives géométriques plus larges, robustification des méthodes, . . .

Nous nous étions plus particulièrement attachés les années précédentes au problème de précision dans la reconstruction. Cette activité s'est notablement réduite cette année où elle n'existe plus qu'à travers des travaux sur le positionnement optimal de caméras.

### 3.1.1 Étude des caméras affines

Le modèle de projection affine est une bonne approximation de la projection perspective dès lors que la taille de l'objet est sensiblement plus faible que la distance d'observation. L'avantage de ce modèle est sa simplicité vis-à-vis du modèle perspectif exact : la plupart des calculs restent linéaires.

Le principal résultat obtenu dans cette direction est un outil d'autocalibration général pour ce type de caméra. Ce modèle rend compte de l'ensemble des travaux antérieurs : approches orthographiques, para-perspectives et de perspective faible. Ce travail autorise aussi de prendre en compte les observations de droites comme celles de points.

### 3.1.2 Reconstruction 3D par itération affine

Dans le cadre d'une convention de recherche avec la société Aérospatiale, nous avons développé une méthode de reconstruction d'un objet ou d'une scène 3D à partir d'une séquence d'images obtenue avec une caméra en mouvement [566, 556]. Le principe de la méthode est d'effectuer une reconstruction affine, de convertir cette reconstruction affine en une reconstruction euclidienne et d'améliorer cette reconstruction dans le cadre d'un processus itératif. Le processus itératif modifie le modèle affine de la caméra de façon à ce qu'il converge vers un modèle projectif.

Les principales propriétés de cette méthode, par rapport à d'autres méthodes de reconstruction, sont sa rapidité de convergence (de l'ordre de 5 à 10 itérations), la simplicité des calculs (chaque itération utilise des techniques classiques d'algèbre linéaire) et la possibilité d'utiliser une caméra "mal étalonnée". En effet, concernant ce dernier point, le seul paramètre qui doit être connu avec précision est le rapport entre la taille verticale et la taille horizontale des pixels de la caméra.

## 3.2 Mouvements singuliers pour l'auto-calibrage et la reconstruction Euclidienne non-calibrée

L'auto-calibrage d'une caméra en mouvement, c'est-à-dire le calibrage à partir des seules images, sans information sur le mouvement ou la structure de l'environnement, est une tâche délicate. Une raison majeure des mauvais résultats est le fait que beaucoup de mouvements de caméras typiques sont singuliers pour cet auto-calibrage. Afin de mieux comprendre ce problème, nous avons défini la singularité en termes géométriques. Ceci nous a permis de dériver une caractérisation très simple et hiérarchique de *toutes* les séquences de mouvements de caméra qui sont singulières pour l'auto-calibrage, les reconstructions Euclidienne ou affine non-calibrées.

Ce résultat est en premier lieu théorique. Il permet de donner des consignes sur les mouvements à éviter lors de prises d'image afin d'obtenir une reconstruction stable. Cependant, il n'est pas toujours possible de contrôler la prise d'images, par exemple quand il s'agit d'images historiques ou récupérées sur Internet. Un système d'auto-calibrage applicable doit donc être capable de reconnaître si une situation singulière est rencontrée et réagir de manière adéquate. Actuellement, nous développons un tel système qui disposera d'algorithmes conçus spécialement pour des diverses situations singulières ; elles devront être en mesure de fournir des reconstructions ou auto-calibrages qui ne seront que partiels mais stables.

## 3.3 Approche tensorielle et la factorisation projective

Les calculs engendrés par la vision sont souvent lourds et parfois obscurs. Quand il y a plus de 2 images, le langage vectoriel-matriciel ne suffit plus pour exprimer les notions de base. Nous avons donc déve-

loppé un formalisme tensoriel pour la vision géométrique multi-images, couplé à une représentation sous forme de diagramme pour le rendre plus intuitif. Les avantages de ce formalisme sont :

- tous les objets géométriques d'intérêt dans une ou plusieurs images peuvent être très facilement manipulés (points, droites, coniques);
- les liens entre l'algèbre et la géométrie sont directs et transparents : les diagrammes tensoriels ont à la fois une interprétation géométrique, algébrique et numérique.

Dans le cadre de cette approche, nous avons créé un formalisme de "jointure d'images" (Joint Image) [581] pour la reconstruction visuelle d'une scène. L'acte de reconstruction est interprété par la recherche d'une cohérence des facteurs d'échelles projectives entre les points correspondants dans plusieurs images, et il se fait entièrement dans l'espace des coordonnées réunies des points images. Pour obtenir la structure 3D, on n'a plus qu'à interpréter les résultats.

Cette approche a déjà conduit à une compréhension quasi-complète de la structure des contraintes d'appariement entre plusieurs images [581], sur laquelle on a développé cette année deux nouvelles méthodes de reconstruction projective multi-images : l'approche *factorisation projective* [577, 582] qui "interprète" la structure de la jointure des images par une méthode de factorisation matricielle, et l'approche *clôture* [581] qui construit les matrices de projection directement à partir des matrices fondamentales. La première méthode est très stable mais exige que tous les points soient visibles dans toutes les images, la seconde n'a besoin que des appariements entre paires d'images.

### 3.3.1 Localisation 2D/3D

La localisation 3D d'un objet à partir d'une seule image est un élément de base de beaucoup d'algorithmes de reconnaissance d'objets et/ou d'asservissement visuel qui n'utilisent qu'une seule vue 2D. Dans ce cadre, nous avons développé deux classes de méthodes : une classe de méthodes privilégiant la précision et la robustesse aux éventuelles erreurs de mise en correspondance et une classe qui privilégie la vitesse et pouvant ainsi s'intégrer dans la boucle de commande d'un robot. Les méthodes que nous avons développées récemment font partie de la deuxième classe et utilisent aussi bien des points que des droites. Elles calculent la position et l'orientation d'un objet 3D par rapport à une vue, l'objet pouvant également être planaire. Ces méthodes consistent en des algorithmes itératifs qui "transforment" une caméra affine (perspective faible ou para-perspective) en une caméra perspective. Ces algorithmes sont de l'ordre de 50 fois plus rapides que les algorithmes non-linéaires.

### 3.3.2 Méthodes de reprojection

Si nous disposons de deux vues d'un même objet, nous pouvons, après une étape de mise en correspondance, calculer le modèle tridimensionnel de cette scène. Ce modèle peut ensuite être reprojété afin de visualiser la scène sous un autre angle.

Cette même technique peut être utilisée sans étalonnage de caméra, permettant ainsi de travailler caméra au poing. Pour ce faire, on exploite les relations algébriques liant les images, ce qui évite en outre de passer par une reconstruction qui peut s'avérer inexacte et surtout qui brise la structure des images initiales, rendant plus difficile des lissages par continuité par exemple.

Dans le cas le plus simple de la synthèse d'une troisième image à partir de deux vues antérieures, nous avons énoncé une paramétrisation simple et minimale permettant l'obtention du résultat [565].

### 3.4 Modélisation et base d'images

*Participants* : Bart Lamiroy, Zhongdan Lan, Roger Mohr, Sylvaine Picard, Cordelia Schmid

*Mots-clés* : base multimédia, classification, interprétation d'image, invariant géométrique, mise en correspondance d'images, reconnaissance des formes, vision par ordinateur.

Le travail mené depuis plusieurs années sur l'appariement d'images a été poursuivi dans plusieurs directions nouvelles. Il a aussi débouché sur un thème nouveau et très fécond : l'indexation des images par leur contenu et la recherche d'images dans de grandes bases pouvant contenir plusieurs milliers d'images.

#### 3.4.1 Algorithmes d'appariement d'images

Les travaux sur l'appariement se sont poursuivis dans quatre directions. L'appariement de segments est désormais possible dans des images où plusieurs objets ont des mouvements différents. L'algorithme est basé sur la comparaison de quasi-invariants locaux, puis sur l'utilisation d'une transformée de Hough. Une recherche locale et non plus globale dans cet espace de Hough permet de trouver des mouvements locaux entre les images.

Plus récemment nous avons découvert que l'appariement de points d'intérêt issus des images en niveaux de gris offrait une robustesse et une discrimination bien supérieure. Nous avons poursuivi le travail sur la caractérisation du signal local par des invariants, en particulier en introduisant une invariance au changement d'éclairage. La richesse de ces invariants permet un appariement par simple comparaison des valeurs, sans avoir besoin d'outils supplémentaires (cf. [573]). Les résultats sur des images aériennes donnent un taux de reconnaissance proche de 100%. Une introduction de la répartition statistique de ces invariants a permis d'améliorer encore de façon sensible cette technique de l'appariement.

#### 3.4.2 Indexation et recherche d'images

Le problème de l'indexation est une suite naturelle à celui d'appariement. Une manière simple de rechercher l'image la plus ressemblante à une image donnée dans un ensemble est d'apparier cette image avec toutes celles de l'ensemble. Le problème est de réaliser cette opération le plus rapidement possible, ce qui oblige le plus souvent à le faire de manière parallèle.

À partir du premier algorithme décrit pour les segments de droites, un système a été réalisé pour l'indexation. À partir de toutes les images de l'ensemble, on constitue une table de hachage de tous les invariants présents, codée sous forme d'un *kd-tree* pour optimiser la place occupée et le temps d'accès. La comparaison entre les invariants de l'image et ceux de l'ensemble des images est alors rapide et indépendante du nombre de ces images. On associe alors un espace de Hough à chacune des images, espaces qui permettent de ne prendre en compte que les appariements d'invariants compatibles du point de vue des transformations géométriques qu'ils définissent, et qui servent donc d'urnes pour un système de vote perfectionné. Avec ce système, les taux de reconnaissance obtenus sont compris entre 90% et 95%, que ce soit avec des images réelles ou issues de la CAO.

Pour les points d'intérêts et leurs invariants décrits plus haut, nous avons préféré développer une technique de hachage tolérante aux incertitudes de nos mesures. La structure est couplée avec un accès direct aux voisins des points identifiés, autorisant par là une vérification directe de la cohérence locale des appariements. Le gain en performance est significatif : un facteur 400 pour une base de plus de 1000 images avec 150000 points d'intérêt stockés.

La figure 1 donne un exemple de consultation de base avec les taux de réponse associée (pour ce cas la densité des invariants n'a pas été intégrée). On notera que l'image qui sert à l'interrogation correspond à un point de vue différent de celui des images stockées (environ 20 degrés de variation de l'angle de prise de vue).

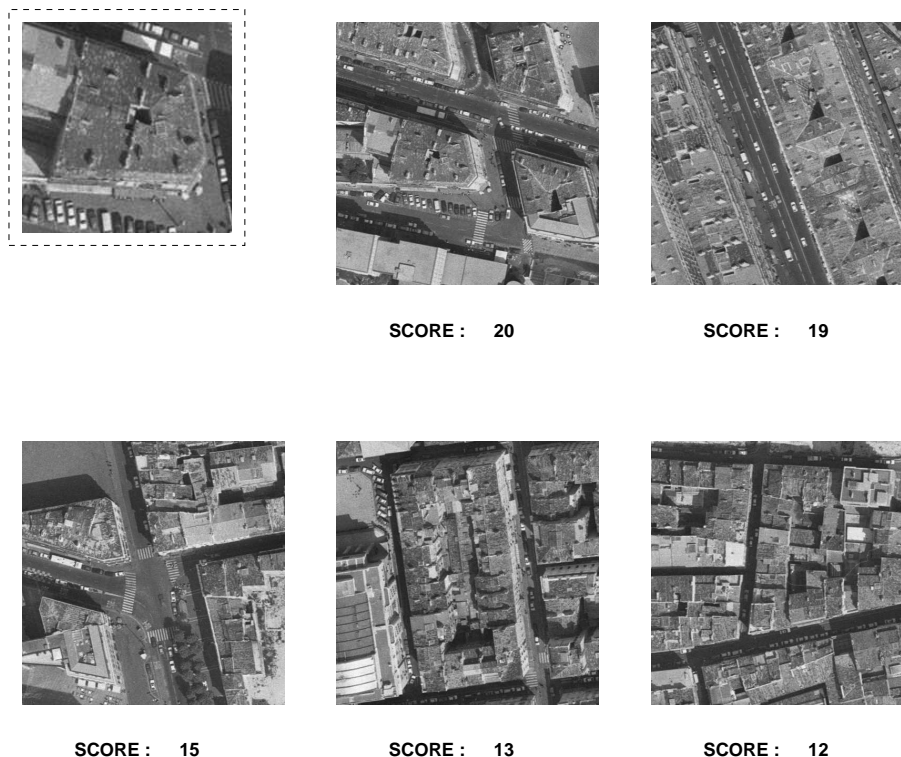


Figure 1: Exemple d'indexation : indexation de modèles pour des images structurées.

### 3.4.3 Mise en correspondance par corrélation

La mise en correspondance d'image à image se fait le plus souvent par corrélation dès lors que l'on veut de nombreux appariement. Cette technique échoue dans les régions proches des contours d'occultation. Pour ce faire nous avons développée une technique de corrélation partielle. Elle est basée sur une détermination robuste des parties d'images pouvant se correspondre et c'est sur cette partie que va s'appliquer la corrélation. Les expérimentations montrent un meilleur comportement que les méthodes robustes actuelles, mais à un coût de calcul supérieur.

Nous avons aussi étudié des techniques rapides permettant d'obtenir des corrélations subpixeliques. Le principe se base sur la régularité du signal image qui autorise une extrapolation linéaire pour des décalages subpixeliques du signal. L'estimation de ce décalage revient simplement à détecter les coefficients de l'approximation linéaire la meilleure. Les performances sont relativement moyennes (1 à 2/10e de pixel) face à des algorithmes plus complexes que nous avons mis en place les années précédentes, mais les coûts de calcul sont modestes ici.

## 3.5 Couplage perception-action

*Participants* : Nicolas Andreff, Pascal Brand, Gabriella Csurka, Katja Daumüller-Claudiel, Radu Horaud, Hervé Mathieu, Long Quan, Andreas Ruf

### 3.5.1 Asservissement visuel pour la saisie

A la suite du travail que nous avons effectué dans le cadre du projet Esprit-BRA SECOND (saisie d'un objet polyédrique), nous développons une nouvelle plate-forme expérimentale équipée d'une tête stéréoscopique et de plusieurs robots dont un robot cartésien à 5 axes et un robot à 6 axes. Cette plate-forme n'est que partiellement calibrée – la chaîne cinématique de chaque robot est connue mais le

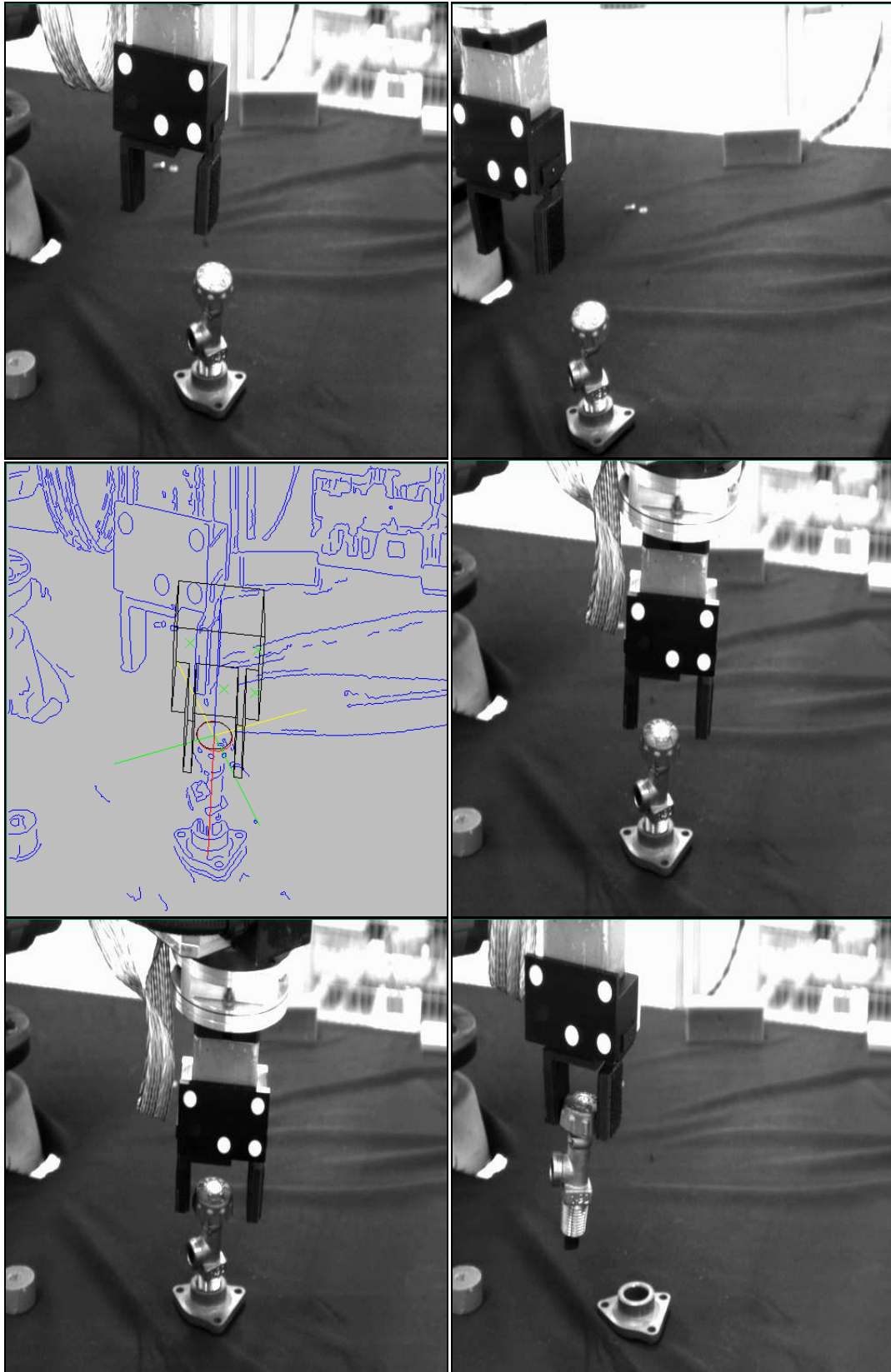


Figure 2: Un exemple de la saisie d'une valve à l'aide d'un système stéréoscopique, la méthode de reconstruction de coniques et la méthode d'asservissement visuel par une seule caméra.

capteur stéréoscopique n'est pas calibré. De plus on ne connaît pas la position relative des robots par rapport au capteur. On s'intéresse à plusieurs aspects : auto-calibration de la plate-forme (tête stéréo, robot-caméras, etc.), saisie d'objets courbes dont on ne possède pas les dimensions, asservissement à l'aide d'une tête stéréo, poursuite temps-réel de robots et objets en mouvement, etc. D'ores et déjà nous avons intégré une méthode de reconstruction de coniques, une méthode de planification de saisie et un algorithme d'asservissement visuel, comme montré sur la figure 2.

Afin d'assurer des caractéristiques temps-réel, nous développons un système d'acquisition et de traitement d'une paire d'images stéréo basé sur une architecture à base de circuits DSP.

## 4 Actions industrielles

Les aspects applicatifs de certains de nos travaux ont suscité des contacts pour des transferts que nous poursuivons. On pourra remarquer cependant que les actions de collaboration actuellement soutenues par nos partenaires sont plutôt fondamentales et d'amont.

### 4.1 Collaboration avec la société ITMI

*Participants* : Radu Horaud , Pascal Brand

Dans le cadre du projet CALVIN, financé par l'Agence Spatiale Européenne, nous réalisons un transfert industriel dans le domaine de la calibration d'une caméra (détection précise de cibles), de la calibration caméra-pince et de l'auto-calibration d'une tête stéréoscopique montée sur le bras d'un robot. Ce transfert est financé, en partie par l'Agence Spatiale Européenne et en partie par l'INRIA et ITMI dans le cadre d'une bourse post-doc industrie dont bénéficie Pascal Brand.

### 4.2 Collaboration avec la société Aérospatiale

*Participants* : Radu Horaud , Stéphane Christy

Nous sommes liés à la société Aérospatiale par un contrat de collaboration de recherche pour une durée de 4 ans et qui a débuté en 1994. Dans le cadre de cette collaboration nous étudions des algorithmes de reconstruction et de modélisation 3D.

### 4.3 Collaboration avec MATRA-CAP-SYSTÈME

*Participants* : Roger Mohr , Jérôme Blanc

Intéressé par nos travaux permettant la synthèse de nouvelles images à partir d'images existantes, le groupe MATRA a décidé de soutenir une action de recherche sur les possibilités de navigation virtuelle à travers un univers statique. Cette collaboration s'étend sur deux ans et doit aboutir à un démonstrateur sur la synthèse d'images sous un nouveau point de vue.

## 5 Actions nationales et internationales

### 5.1 Actions nationales

#### GDR-PRC Communication Homme-Machine

Notre projet assure avec le projet SHARP la responsabilité d'un consortium qui a été sélectionné lors d'un appel d'offre des PRC. L'objectif de cette action est l'intégration de capteurs dans la planification, le choix de l'espace géométrique le mieux adapté à la tâche, et la résolution dans des contextes incertains des problèmes de perception et planification. Dans ce cadre nous avons une collaboration marquée avec



le projet BIP sur la commande référencée vision, avec SHARP sur la planification des capteurs et avec ROBOTVIS sur la calibration progressive.

### **Action régionale Émergence**

Ce projet a été monté en réponse à un appel d'offre de la région Rhône-Alpes portant sur les technologies émergentes, et pouvant faire l'objet d'une priorité régionale dans le futur. Les partenaires sont, outre MOVI, l'équipe Bases de Données du laboratoire LSR, l'équipe Recherche d'Information du laboratoire CLIPS et la société Teamlog. Le projet concerne l'indexation et la recherche d'images par leur contenu, dans de grandes bases de données pouvant contenir plusieurs milliers d'images.

### **Autres collaborations.**

- projet iMAGIS : La communauté de synthèse d'images s'intéresse depuis peu à des techniques de reprojections comme celle décrite dans la partie 3.3.2. Nous démarrons une collaboration sur ce point avec un jeune étudiant en commun.
- projet BIP : dans le cadre des actions déjà mentionnées mettant en œuvre le couplage robotique et vision ; nous avons en outre un étudiant en commun sur le sujet du couplage de la commande pour des actions perceptions et de saisie assistée par la vision.
- projet TEMIS : nous travaillons avec ce projet pour la mise en place d'une action commune sur le problème de l'indexation de documents multimédia.

## **5.2 Actions internationales**

### **5.2.1 Europe de l'ouest**

#### **CALVIN.**

Ce projet (Calibration of vision systems) est financé par l'Agence Spatiale Européenne (ESA) et est mené en collaboration avec la société ITMI (coordinateur), la société Tecnospatio (Italie), l'Institut Royal de Technologie de Stockholm, ainsi que le projet Robotvis (Sophia-Antipolis). Le but du projet est d'intégrer le savoir-faire des instituts de recherche dans le domaine de la calibration et de réaliser un prototype industriel qui sera validé par l'ESA en vue d'une éventuelle utilisation dans l'espace.

#### **CUMULI.**

Le projet Esprit LTR CUMULI (Computational Understanding of Multiple Images) a débuté en septembre 96 pour 3 ans. C'est une collaboration avec l'équipe ROBOTVIS, l'université de Lund en Suède, l'institut Fraunhofer-IGD en Allemagne, et les sociétés IMETRIC en Suisse et INNOVATIV VISION en Suède. Le projet a pour but de raffiner et implanter dans l'industrie nos compétences sur la mesure précise industrielle à partir des images multiples.

#### **Visually-guided disassembly.**

Projet financé par le programme communautaire TMR (thèse de K. Daumüller-Claudé) pour réaliser une tâche de désassemblage avec un robot et une tête stéréo. D'une durée de trois ans, ce projet a débuté le 1er avril 1996.

## 5.2.2 Europe de l'est

Dans le cadre d'un programme TEMPRA, Tomáš Svoboda, thésard de l'université Polytechnique de Prague, a été accueilli pour un séjour de 5 mois. En collaboration avec P. Sturm et Roger Mohr, il a étudié l'influence d'erreurs dans le calibrage d'une caméra sur l'estimation du mouvement de la caméra ([583]).

## 5.2.3 Amérique

### Coopération avec l'université de Carnegie-Mellon

P. Gros (jusqu'à octobre) et L. Quan (jusqu'à août) ont fait un séjour d'un an au CMU. L. Quan a travaillé sur le problème de reconstruction tridimensionnelle des droites et P. Gros sur la caractérisation d'images dans une base d'images.

## 5.2.4 Asie et Océan Pacifique

### Coopération Chine-France

Avec le prof. C. Wu de l'université de Xi'an, nous poursuivons des travaux sur l'identification d'objets par des invariants projectifs. Nos orientations récentes portent sur l'usage des outils géométriques pour la compression d'images de scènes 3D semi rigides.

# 6 Diffusion des résultats

## 6.1 Enseignement

Le personnel du projet est naturellement impliqué dans les enseignements des établissements auquel il est rattaché. Il est en outre impliqué dans les enseignements plus spécifiques de son domaine dans les DEA de Grenoble (Signal-Image-Parole et Vision-Image-Robotique).

## 6.2 Participation à des colloques

Des membres de l'équipe ont participé à des conférences et *workshops* ; on se reportera à la bibliographie pour en avoir la liste.

## 6.3 Conférences invitées, tutoriels, cours, etc.

- Jurys de thèse : le projet participe aux jurys de thèse des établissements rattachés. Il a en outre participé à des jurys à Nancy, Clermont-Ferrand, Rennes, Toulouse, Paris, Nice-Sophia.
- Membres de comités de rédaction de revues : *International Journal of Pattern Recognition and Artificial Intelligence* (R. Mohr), *Traitement du Signal : Signal, Image et Parole* (R. Horaud), *Pattern Recognition Journal* (R. Mohr, R. Horaud), *Machine Vision and Applications* (R. Mohr), numéro spécial de *Artificial Intelligence Journal* sur la vision (R. Horaud) ;
- Participation à des comités de programme : ECCV96 (R. Horaud), AFCET96 (R. Horaud), CVPR97 (R. Mohr), CAIP97 (R. Mohr), ICCV98 (R. Horaud et R. Mohr) ;
- R. Mohr est membre du bureau du Conseil Scientifique de l'IMAG, président de la CSE de l'INPG, membre du GT4 du MESR, membre du bureau du GDR PRC C.H.M., membre du comité de pilotage de ECCV ;

- R. Mohr et B. Triggs ont donné un tutoriel sur la géométrie projective appliquée à la perception 3D au congrès ISPRS (Vienne, 1996) ;
- les membres du projet ont donné les conférences invitées suivantes :
  - à la création du groupe de travail sur les BD multimédia (R. Mohr)
  - aux journées nationales du GDR-PRC CHM (R. Mohr)
  - University of Maryland (L. Quan)
  - Siemens Inc. à Princeton (L. Quan)
  - Noblesse Workshop (C. Schmid).

## 6.4 Animations scientifiques

L'équipe était présente au CEBIT 96 à Hanovre, en Allemagne, où ont été présentés les travaux sur l'indexation d'images.

## 7 Publications

### Articles et chapitres de livre

- [555] S. CARLSSON, R. MOHR, L. MORIN, C. ROTHWELL, L. V. GOOOL, F. VEILLON, A. ZISSERMAN, «Semi-Local Projective Invariants for the Recognition of Smooth Plane Curves», *International Journal of Computer Vision* 19, 3, 1996, p. 211–236.
- [556] S. CHRISTY, R. HORAUD, «Euclidean Shape and Motion from Multiple Perspective Views by Affine Iterations», *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 10, October 1996.
- [557] P. GROS, R. HARTLEY, R. MOHR, L. QUAN, «How Useful is Projective Geometry?», *Computer Vision and Image Understanding*, 1996, to appear.
- [558] R. HARTLEY, P. STURM, «Triangulation», *Computer Vision and Image Understanding*, 1996, accepted.
- [559] R. HORAUD, F. DORNAIKA, B. LAMIROY, S. CHRISTY, «Object Pose: The Link between Weak Perspective, Paraperspective, and Full Perspective», *International Journal of Computer Vision*, 1997, to appear.
- [560] L. QUAN, «Conic Reconstruction and Correspondence from Two Views», *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 2, February 1996, p. 151–160.
- [561] L. QUAN, «Self-Calibration of An Affine Camera from Multiple Views», *International Journal of Computer Vision* 19, 1, May 1996, p. 93–105.
- [562] P. TAO, T. PHONG, R. HORAUD, L. QUAN, «Stability of Lagrangian Duality for Nonconvex Quadratic Programming Solution Methods and Applications in Computer Vision», *Mathematical Modelling and Numerical Analysis*, 1996, to appear.
- [563] C. ZHAO, R. MOHR, «Global Three-Dimensional reconstruction from Occluding Contours», *Computer Vision and Image Understanding* 64, 1, 1996, p. 62–96.

### Communications à des congrès, colloques, etc.

- [564] J. BLANC, R. MOHR, «Naviguer dans un monde virtuel», in : *Journées ORASIS 1996, Clermont-Ferrand (FR)*, p. 135–140, mai 1996.
- [565] P. BOBET, J. BLANC, R. MOHR, «Aspects cachés de la tri-linéarité», in : *Actes du 10ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle, Rennes (FR)*, LIFIA–IMAG–INRIA Rhône-Alpes, p. 137–146, janvier 1996.

- [566] S. CHRISTY, R. HORAUD, «Euclidean Reconstruction: from Paraperspective to Perspective», in : *Computer Vision – ECCV'96, Proceedings Fourth European Conference on Computer Vision, Cambridge, England*, Buxton-Cipolla (éd.), Springer Verlag, p. 129–140, April 1996.
- [567] G. FLOROU, R. MOHR, «What Accuracy for 3D Measurements with Cameras?», in : *Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria*, p. 354–358, 1996.
- [568] B. LAMIROY, P. GROS, «Rapid Object Indexing and Recognition Using Enhanced Geometric Hashing», in : *Proceedings of the 4th European Conference on Computer Vision, Cambridge, England, 1*, p. 59–70, April 1996.
- [569] B. LAMIROY, P. GROS, «Reconnaissance d'objets par indexation géométrique étendue.», in : *Journées ORASIS 1996, Clermont-Ferrand (FR)*, p. 19–24, mai 1996.
- [570] Z. D. LAN, R. MOHR, «Appariement robuste par corrélation partielle», in : *Journées ORASIS 1996, Clermont-Ferrand (FR)*, p. 99–104, 1996.
- [571] L. QUAN, T. KANADE, «A Factorization Method for Affine Structure from Line Correspondences», in : *Proceedings of the Conference on Computer Vision and Pattern Recognition, San Francisco, California, USA*, p. 803–808, 1996.
- [572] C. SCHMID, P. BOBET, B. LAMIROY, R. MOHR, «An Image Oriented CAD Approach», in : *Proceedings of the ECCV workshop on Object Representation*, 1996.
- [573] C. SCHMID, R. MOHR, «Combining Greyvalue Invariants with Local Constraints for Object Recognition», in : *Proceedings of the Conference on Computer Vision and Pattern Recognition, San Francisco, California, USA*, June 1996.
- [574] C. SCHMID, R. MOHR, «Image Retrieval Using Local Characterization», in : *Proceedings of the IEEE International Conference on Image Processing*, p. 781–788, September 1996.
- [575] C. SCHMID, R. MOHR, «Image Retrieval Using Local Characterization», in : *Proceedings of the 3rd International Conference on Communicating by Image and Multimedia*, p. 45–50, May 1996.
- [576] C. SCHMID, R. MOHR, «Mise en correspondance par invariants locaux», in : *Actes du 10ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle, Rennes (FR)*, p. 236–245, 1996.
- [577] 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*, B. Buxton, R. Cipolla (éd.), *Lecture Notes in Computer Science, 1065*, Springer Verlag, p. 709–720, April 1996.
- [578] P. STURM, B. TRIGGS, «Structure et mouvement projectifs à partir d'images multiples par factorisation», in : *Journées ORASIS 1996, Clermont-Ferrand (FR)*, p. 175–180, May 1996.
- [579] P. STURM, «Calibrage dynamique d'une caméra en mouvement par pré-calibrage», in : *Journées ORASIS 1996, Clermont-Ferrand (FR)*, p. 61–66, mai 1996.
- [580] P. STURM, «Self-Calibration of a Moving Camera by Pre-Calibration», in : *Proceedings of the seventh British Machine Vision Conference, Edinburgh, Scotland*, R. Fisher, E. Trucco (éd.), 2, British Machine Vision Association, p. 675–684, September 1996.
- [581] B. TRIGGS, «Matching Constraints and the Joint Image», in : *Proceedings of the 5th International Conference on Computer Vision, Cambridge, Massachusetts, USA*, E. Grimson (éd.), IEEE, IEEE Computer Society Press, p. 338–343, June 1995.
- [582] B. TRIGGS, «Factorization Methods for Projective Structure and Motion», in : *Proceedings of the Conference on Computer Vision and Pattern Recognition, San Francisco, California, USA*, p. 845–851, 1996.

## Rapports de recherche et publications internes

- [583] T. SVOBODA, P. STURM, «What can be done with a Badly Calibrated Camera in Ego-Motion Estimation ?», *rapport de recherche n°CTU-CMP-1996-01*, Center for Machine Perception, Czech Technical University, Prague, August 1996.

## 8 Abstract

Our research project is focused on the perception of three dimensional shapes using cameras. Within this general frame, we study more particularly the geometry of perception in order to be able to locate objects from a single or multiple views and to be able to recognize shapes or obstacles.

Our work on 3D reconstruction using non calibrated or partially calibrated cameras developed a general tensorial framework which allows a better understanding of the multiple image geometry and provide efficient computational schemes. Some basic new results were obtained in coupling the general camera model and affine camera model. Calibration is still active, more particularly in the direction of self-calibration and hand eye calibration. For the recognition purpose, we developed global matching techniques based on invariant features. They allow efficient indexing into databases for object or pattern retrieval.

Integration of all these aspects is done around a robotic manipulation application. Obstacle detection, object identification/location and grasping are combined in a global demonstrator.

