
Projet ROBOTVIS

Robotique et Vision par Ordinateur

Localisation : *Sophia Antipolis*

Mots-clés : vision par ordinateur, traitement d'image, vision 3D, analyse du mouvement, reconstruction 3D, reconnaissance de forme, analyse de scène, vision active, stéréovision.

1 Composition de l'équipe

Responsable scientifique

Olivier Faugeras, directeur de recherche, Inria

Assistante de projet

Philippa Hook

Personnel Inria

Rachid Deriche, directeur de recherche

Luc Robert, chargé de recherche

Jean-Luc Szpyrka, ingénieur de recherche

Thierry Viéville, chargé de recherche

Zhengyou Zhang, chargé de recherche

Chercheurs invités

Marie Bysveen, depuis le 1er octobre

Angéla Lopez, de mars à mai 1996

Daniel Reifeld, jusqu'au 30 juin

Alfons Salden, à partir du 1er octobre

Andrew Zisserman, du 1 janvier au 1 juillet, HCM

Post-doctorants

Reyes Enciso, boursière du Ministère de l'Education d'Espagne, jusqu'au 1 Octobre 1996

Daniel Reifeld, jusqu'au 30 juin

Charly Rothwell, jusqu'au 31 août, HCM

Chercheurs doctorants

Didier Bondyfalat, boursier MESR, université de Nice-Sophia Antipolis, à partir du 1er octobre, co-encadré par le projet SAFIR
 Sylvain Bougnoux, action PRAXITELE
 Gabriella Csurka, boursier Inria, université de Nice-Sophia Antipolis, jusqu'au 30 avril
 Quentin Delamarre, boursier MESR, université de Nice-Sophia Antipolis, à partir du 1er octobre
 Frédéric Devernay, boursier Dret-X, École Polytechnique
 Cyrille Gauclin, boursier Inria
 Renaud Keriven, Cermics
 Pierre Kornprobst, boursier MESR, université de Nice-Sophia Antipolis
 Stéphane Laveau, boursier Dret-X, École Polytechnique, jusqu'au 30 juin
 Diane Lingrand, boursière MESR, université de Nice-Sophia Antipolis
 Laurence Lucido, laboratoire d'accueil IFREMER
 Nikos Paragios, depuis le 1er octobre, TMR
 Robert Stahr, boursier MESR, université de Nice-Sophia Antipolis
 Cyril Zeller, boursier Inria, jusqu'au 30 juin
 Imed Zoghلامي, boursier Inria

Stagiaires

Didier Bondyfalat, DEA Math-Info, 6 mois
 Soraya Arias, Eurecom et DEA ARAVIS, 6 mois
 Christophe Bouvin, DEA ARAVIS, 6 mois
 Quentin Delamarre, DEA ARAVIS, 6 mois
 Olivier Delayais, Stagiaire X, 6 mois
 François Gaspard, ESSI et DEA ARAVIS, 6 mois
 Veit Schenk, DEA ARAVIS, 6 mois

2 Présentation du projet

Les objectifs scientifiques et technologiques du projet sont de deux types :

- élaborer des théories de la perception visuelle par ordinateur et plus particulièrement de la perception tridimensionnelle dynamique,
- démontrer la validité de ces théories en les appliquant à des problèmes concrets, en développant et en implantant des algorithmes pour les résoudre et en travaillant en commun avec des industriels.

Nous accordons une importance centrale à la formalisation et tout particulièrement à la formalisation mathématique, en ayant cependant toujours recours à l'expérimentation, seule façon pensons-nous de valider cette théorie. De ce point de vue, il apparaît nettement deux grands axes théoriques qui guident notre approche.

Le premier est l'utilisation de méthodes algébriques et géométriques pour l'étude des systèmes multicaméras (voir les rapports des années précédentes et la section 3.1). Nous pensons que cet axe va se renforcer dans les années à venir, notamment au travers de collaborations avec le projet Safir (B. Mourrain) et grâce au projet européen LTR CUMULI.

Le second axe est l'utilisation de méthodes d'analyse, principalement issues de la théorie des équations aux dérivées partielles, pour l'étude des problèmes spécifiques du traitement d'images (amélioration, restauration, contours, stéréo, mouvement), voir la section 3.4. Là encore, nous pensons qu'il s'agit d'une orientation durable, que nous souhaitons renforcer, notamment au travers de collaborations avec l'université de Nice-Sophia Antipolis (Gilles Aubert, Michel Merle), l'université Paris IX Dauphine (Jean-Michel Morel) et le Cermics (Renaud Keriven).

Nous avons par ailleurs développé dans le passé un assez grand nombre de collaborations avec des industriels, avec lesquels nous avons travaillé surtout dans le cadre de projets européens et vers lesquels nous avons transféré beaucoup de notre savoir et de nos algorithmes (voir rapports d'activité des années 90 à 93). Nous continuons à rechercher activement ces collaborations, mais dans des domaines plus variés que celui de la robotique, notamment celui de la modélisation tridimensionnelle à partir de séquences d'images ou encore celui des "nouvelles images".

Dans ce contexte, nos efforts ont porté plus spécialement sur les points suivants :

- géométrie d'un système de caméras, calibration, autocalibration ;
- stéréoscopie et applications ;
- perception tridimensionnelle des formes et du mouvement par analyse de séquences d'images ;
- équations aux dérivées partielles, images et vision par ordinateur.

3 Actions de recherche

3.1 Géométrie d'un système de caméras, calibration, autocalibration

3.1.1 Estimation de la matrice fondamentale

Participant : Zhengyou Zhang

Mots-clés : reconstruction 3D, analyse du mouvement, stéréovision.

La géométrie épipolaire, nous le savons, décrit la relation entre deux images d'une même scène, qui est caractérisée par une matrice 3×3 singulière. Son estimation est donc très importante pour beaucoup d'applications comme la modélisation de scènes et le déplacement d'un robot mobile autonome. Nous avons fait une revue complète des techniques existantes d'estimation de la matrice fondamentale et de son incertitude [674]. Nous introduisons une méthode de mesure permettant de comparer avec précision ces techniques. La reconstruction projective est aussi passée en revue. Les logiciels que nous avons développés pour effectuer cette revue sont disponibles sur l'Internet, comme discuté dans ce document.

3.1.2 Estimation de la géométrie d'un système de N caméras

Participants : Stéphane Laveau, Olivier Faugeras

Mots-clés : reconstruction 3D, analyse du mouvement, stéréovision.

Nous avons achevé cette année le développement de méthodes d'estimation des objets représentant la géométrie d'un système de caméras, jusqu'au résultat final qui assure une cohérence entre les données, c'est à dire les matrices de projection. Ce travail fait partie d'une thèse soutenue en mai [636].

Le point de départ est toujours l'estimation de correspondances entre les images. C'est la partie la plus critique du système. Des bonnes correspondances permettent généralement d'obtenir une bonne géométrie du système. Toutefois, ces bonnes correspondances sont difficiles à obtenir.

L'estimation à base de matrices fondamentales souffre de défauts majeurs : elle ne permet pas l'obtention d'un résultat à coup sûr et elle est très dépendante de la qualité de l'estimation d'une des matrices fondamentales.

Les trilinearités sont aussi très sensibles au bruit. Un bruit faible peut facilement perturber l'estimation de ces dernières, même en prenant toutes les précautions nécessaires pour normaliser les données.

Au contraire, l'estimation directe des matrices de projection est assez robuste et permet l'estimation de la géométrie pour un nombre de caméras plus grand (6-8 au lieu de 3). Cette méthode ne marche absolument pas sans l'utilisation de méthodes statistiques robustes. En effet, les méthodes robustes servent non seulement à éliminer les faux appariements, mais aussi à choisir une paramétrisation des matrices

de projections optimales. La figure 1 montre le résultat de l'estimation de la géométrie épipolaire sur huit images de la place des arcades à Valbonne.

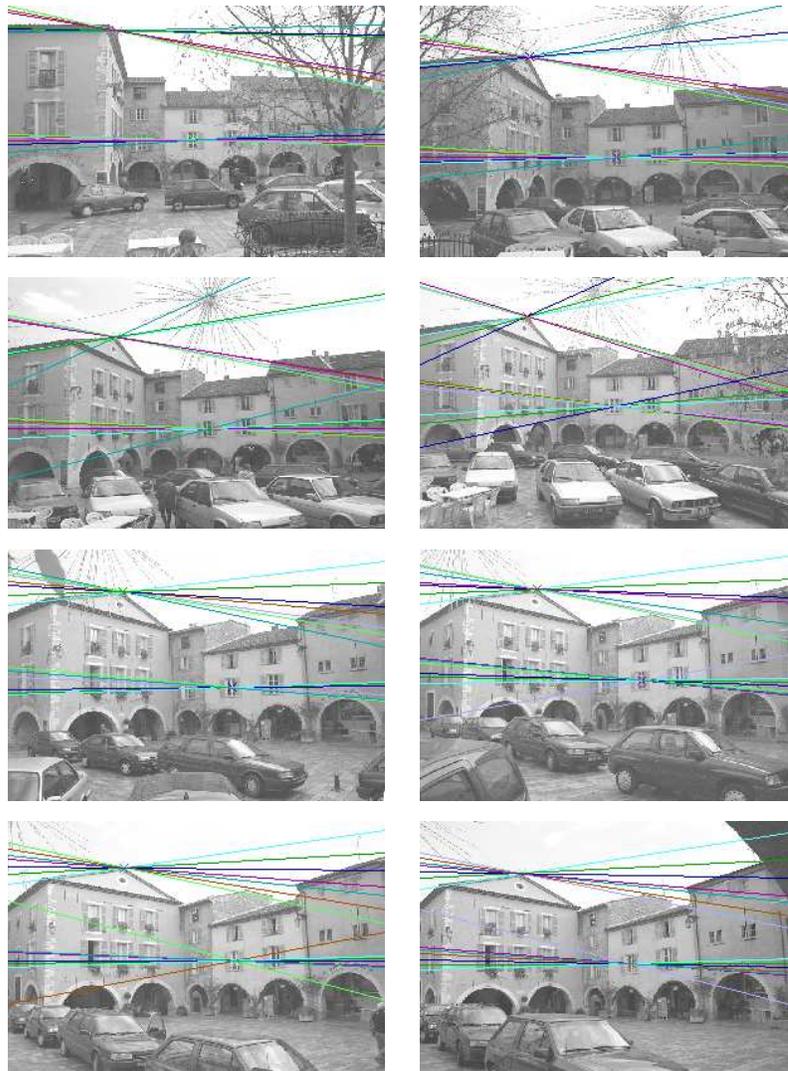


Figure 1: Vue d'ensemble de la place des arcades à Valbonne. Des droites épipolaires sont tracées en certains points. On remarque que les droites épipolaires se coupent en des points uniques. Ces points sont en correspondance entre les images.

Une estimation sur un petit nombre de caméras ne suffit pas forcément pour nos applications. Nous estimons donc alors la géométrie de caméras par sous-ensembles que nous recollons ensuite. Le changement de base est effectué en mettant en correspondance les points. Il est clair que l'algorithme de découpage en sous-ensembles doit dépendre de la scène observée et de la continuité des appariements dans cette scène. Plus de recherches devraient être faites pour déterminer un algorithme automatique.

La géométrie ainsi obtenue est raffinée par des méthodes itératives. Il en existe principalement deux classes, qui minimisent, l'une, une distance entre points, l'autre, une distance entre point et droite. Les deux sont mesurées dans les images.

Les résultats montrent clairement la supériorité des méthodes directes et robustes, même en l'absence de fausses correspondances. Ces méthodes ont toutefois un coût exponentiel par rapport au nombre

d'images. Il faut donc les associer à des méthodes de recalage entre sous-séquences pour pouvoir les utiliser sur plus de 5 à 10 images.

Pour le raffinement, l'ajustement de faisceaux est la méthode donnant des bons résultats le plus rapidement. De plus, il est facile d'implémenter des extensions de ces méthodes à des cas plus particuliers.

3.1.3 Géométrie projective orientée et applications

Participants : Stéphane Laveau, Olivier Faugeras

Mots-clés : reconstruction 3D, analyse du mouvement, stéréovision.

Nous avons introduit dans la vision par ordinateur un nouveau concept, celui d'orientation projective, décrit en détail dans [636]. A l'aide de modifications très simples (garder un signe) aux théories existantes, celles des espaces projectifs, nous obtenons des propriétés très utiles pour un certain nombre d'applications.

Les modifications de la théorie projective en vision entraînent des modifications de son interprétation. Le plan focal possède désormais un « devant » et un « derrière ». Les épipoles et les droites épipolaires ont une orientation, qui est reliée au sens du parcours de la lumière sur le rayon. Les plans de l'espace peuvent eux aussi être orientés, et ceci nous permet de définir des enveloppes convexes sans abandonner l'agréable formalisme de la géométrie projective.

Les applications sont multiples. Celles que nous avons développées ici sont au nombre de trois, mais on peut en imaginer d'autres :

- suppression de points aberrants lors d'une reconstruction projective (voir figure 2);
- organisation d'une géométrie épipolaire pour tenir compte des contraintes de visibilité (voir figure 3);
- raffinement des invariants à des fins de reconnaissance comme dans les travaux de Richard Hartley ou ceux de Charlie Rothwell.

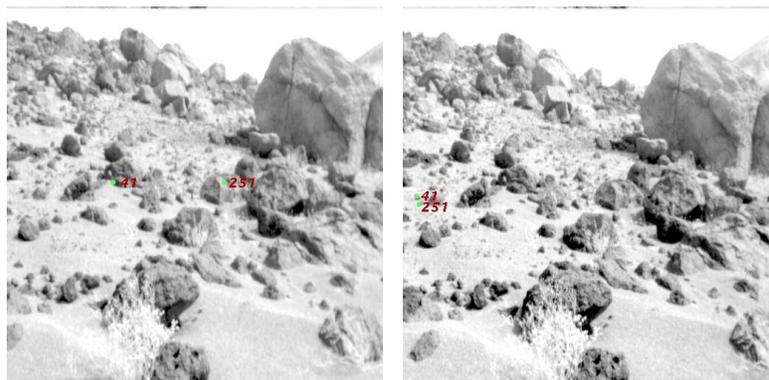


Figure 2: Faux appariements détectés grâce à l'orientation (point 251 par exemple). Le point correspondant dans l'espace est reconstruit derrière les plans focaux.

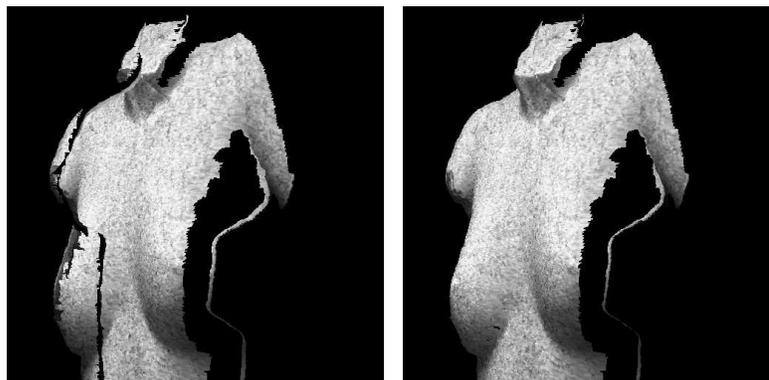


Figure 3: Deux images synthétisées avec des caméras différant uniquement par l'orientation. L'image de gauche présente des anomalies sur le côté qui sont dues au fait que le sein droit du mannequin est vu par derrière. L'objet n'est pas complet parce que les images sources ne le couvrent pas entièrement.

3.1.4 Autocalibration d'une caméra à partir d'une séquence d'images

Participants : Cyril Zeller, Olivier Faugeras

Mots-clés : reconstruction 3D, analyse du mouvement, stéréovision.

Nous avons achevé cette année le travail commencé il y a quelques années sur l'étalonnage euclidien de systèmes de caméras. Ceci a donné lieu à une thèse soutenue en février [637]. Nous avons exploré, en détail, plusieurs méthodes de calcul des paramètres intrinsèques d'une caméra.

Le calcul à partir de la matrice de l'homographie du plan à l'infini présente l'avantage d'avoir une formulation linéaire, mais certains auteurs (Hartley, Zisserman et al.) ont constaté qu'il est cependant très sensible à la précision de l'homographie du plan à l'infini et que des résultats satisfaisants ne sont obtenus qu'en utilisant plusieurs estimateurs de l'homographie du plan à l'infini, à partir de plusieurs mouvements.

Nous nous sommes surtout intéressés au calcul des paramètres intrinsèques à partir de la matrice fondamentale. Une étude théorique fouillée des équations correspondantes nous a permis d'exhiber en détail les cas de dégénérescence, ceci suivant la configuration spatiale des vues les unes par rapport aux autres.

Ces équations sont non-linéaires et la difficulté consiste à trouver une valeur initiale des paramètres intrinsèques suffisamment proche de la valeur optimale pour que l'algorithme de minimisation non linéaire converge correctement. Ceci est obtenu, là aussi, en augmentant le nombre de données, c'est-à-dire, le nombre de matrices fondamentales. En particulier, lorsque le nombre de vues est suffisamment élevé, le facteur d'aspect est, en général, très bien estimé, à tel point que le fixer à sa valeur initiale au cours de la minimisation conduit souvent à de meilleurs résultats. De plus, nous avons constaté que l'utilisation des matrices de covariance des matrices fondamentales dans la formulation du critère de minimisation améliore nettement les résultats.

L'estimation des paramètres extrinsèques, nécessaire à l'obtention de données tridimensionnelles, n'est satisfaisante que dans le cas de séquences dont les angles de vue par rapport à la scène sont nettement différents. Dans ce cas, les informations tridimensionnelles calculées, soit directement dans les images, soit en effectuant une reconstruction tridimensionnelle de la scène par la technique de l'ajustement de faisceaux, s'avèrent de bonne qualité.

Nous présentons en figure 4 des résultats de calculs d'invariants euclidiens (angles et rapports de longueurs), obtenus sur l'église de Valbonne à partir des résultats de l'étalonnage euclidien.

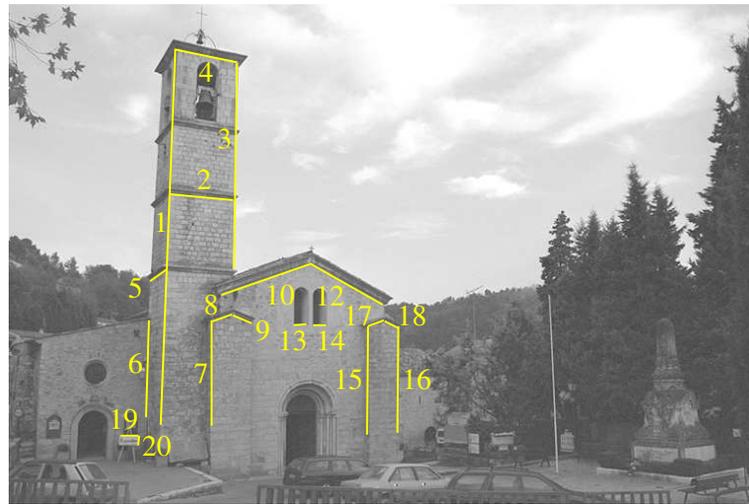


Figure 4: Les segments utilisés pour les mesures du tableau 1.

Angle	calculé	réel	Rapport	calculé	réel
$\cos(1, 3)$	1	1	2/4	0.993	1
$\cos(1, 2)$	0.054	0	4/5	0.999	1
$\cos(3, 5)$	0.032	0	10/12	1.05	1
$\cos(6, 7)$	0.999	1	8/9	1.004	1
$\cos(15, 16)$	1	1	17/18	0.966	1
$\cos(19, 20)$	0.101	0	13/14	0.902	1

Tableau 1: Mesures, relatives aux segments de la figure 4, calculées dans les images, en utilisant les paramètres intrinsèques résultant de l'étalonnage euclidien automatique.

3.1.5 Autocalibration en vision active et reconstruction 3D

Participants : Reyes Enciso, Thierry Viéville

Mots-clés : reconstruction 3D, analyse du mouvement, stéréovision.

Le présent travail a pour but l'auto-calibration d'un système binoculaire à trois degrés de liberté : rotation, élévation et vergence de la caméra gauche. La caméra droite dispose d'un mécanisme de mise au point et de zoom. Nous avons tout d'abord mis en évidence que seuls trois paramètres intrinsèques sont à prendre en compte, et que le changement de mise au point (et de focale dans le cas d'un zoom de haute qualité) revient à une transformation affine dans l'image, que l'on peut calculer à partir du suivi d'au moins deux points sur deux images.

Ensuite nous avons développé une méthode d'auto-calibration euclidienne du système permettant d'estimer des paramètres intrinsèques et extrinsèques variables entre les deux paires de vues utilisées. Par auto-calibration euclidienne, nous entendons donc une méthode automatique basée sur les informations extraites sur les images (les correspondances de points) et la géométrie épipolaire sous la forme de la matrice fondamentale. Finalement nous avons montré quelques exemples de reconstruction 3D basés sur des techniques de Vision Active.

En particulier, nous avons mis en œuvre un système de coopération entre les modules du zoom, de calcul de la profondeur par accommodation et de la fixation, ceci nous permettant la reconstruction 3D des scènes statiques.

Des exemples issus de [635] sont donnés figure 5 et figure 6.



Figure 5: Vue caméra de la reconstruction 3D euclidienne.

3.1.6 Géométrie épipolaire de caméras avec distorsion

Participant : Zhengyou Zhang

Mots-clés : reconstruction 3D, analyse du mouvement, stéréovision.

Pour obtenir une reconstruction 3D (euclidienne ou projective) avec une grande précision, on doit considérer la distorsion non linéaire de la caméra. Dans les travaux antérieurs sur les problèmes de caméras multiples, une caméra est modélisée comme un sténopé. La distorsion de la caméra a souvent été corrigée en différé (*off-line*). Nous avons étudié la possibilité de corriger cette distorsion en temps réel comme une partie intégrante de la caméra [667]. Dans ce cas, pour un point dans une image, son point en correspondance dans l'autre image doit se trouver, non plus sur une droite épipolaire, mais sur une courbe. Nous pouvons donc estimer à la fois la distorsion et la matrice fondamentale (qui correspond à la géométrie épipolaire entre deux images sans distorsion) en minimisant les distances entre les points et leurs *courbes épipolaires* correspondantes. Les résultats expérimentaux avec des données simulées montrent que nous pouvons effectivement estimer la distorsion, si elle est forte et si les points d'intérêt sont extraits avec une bonne précision (≤ 0.3 pixels). Sinon, il vaut mieux les considérer sans distorsion.

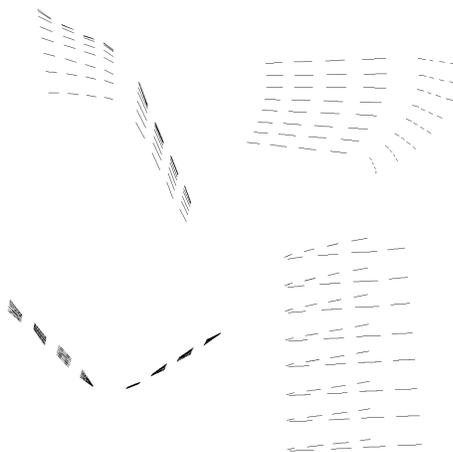


Figure 6: Vues de dessus et de face par deux méthodes concurrentes.

3.1.7 Recalibration d'un système de stéréovision dans le temps

Participants : Zhengyou Zhang, Veit Schenk

Mots-clés : reconstruction 3D, analyse du mouvement, stéréovision.

Durant l'exécution d'une tâche visuelle, la calibration des caméras peut ne plus être valide. Ceci est dû aux changements accidentels, comme les variations thermiques et les vibrations mécaniques, ou aux changements intentionnels comme les variations de zoom et de foyer. Il est donc indispensable de maintenir la calibration d'un système de stéréovision dans le temps. Dans le contrat CALVIN avec l'ESA [673], nous avons développé une technique qui permet de recalibrer dynamiquement le système de stéréovision, préalablement calibré, en utilisant uniquement les informations provenant de l'environnement et non des objets de référence comme une mire. L'idée est de maintenir la structure de l'environnement reconstruite dans le temps, qui est utilisée comme la référence pour la calibration. L'incertitude dans les paramètres de calibration, dans la structure reconstruite et dans les points extraits des images est prise en compte de manière systématique.

Un exemple est montré en figure 7. La paire d'images stéréo en haut est prise à l'instant t_1 , et calibrée par une technique classique avec une mire bi-plan. La paire d'images stéréo en bas est prise à l'instant t_2 . Entre les deux instants, le système stéréo a effectué une translation en avant et une petite rotation vers la droite et, en même temps, nous avons diminué le zoom pour garder les objets de la scène à peu près de la même taille.

Les résultats sont récapitulés dans le tableau 2. En effet, en comparaison, nous avons aussi calibré les deux images à l'instant t_2 avec la technique classique avec une mire bi-plan. Dans le tableau, les paramètres \mathbf{P}_i (pour $i = 1, 2, 3, 4$) sont obtenus avec la technique de calibration classique. Les indices 1 et 2 sont utilisés pour la paire d'images à t_1 et les indices 3 et 4, à t_2 . Les paramètres $\hat{\mathbf{P}}_3$ et $\hat{\mathbf{P}}_4$ sont obtenus avec notre nouvelle technique. Nous remarquons que les valeurs de a_u et a_v à t_2 sont effectivement plus petites qu'à l'instant t_1 . Une autre chose intéressante à savoir est que les paires d'images stéréo ont été prises par la même caméra déplacée, donc elles devront avoir les mêmes paramètres intrinsèques. Or, la technique classique donne des valeurs assez différentes entre les paramètres intrinsèques de \mathbf{P}_3 et ceux de \mathbf{P}_4 , mais la nouvelle technique donne des valeurs très cohérentes.

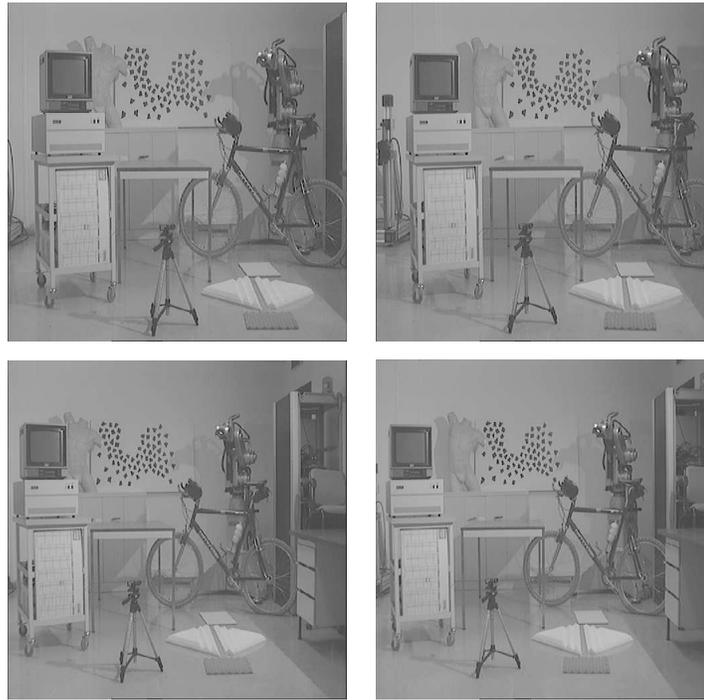


Figure 7: Deux paires d'images prises à deux instants différents : la paire en haut est calibrée, et celle en bas, non.

Tableau 2: Résultats de la recalibration d'un système de stéréovision dans le temps : comparaison entre la calibration classique (\mathbf{P}_i , pour $i = 1, 2, 3, 4$) et la recalibration ($\hat{\mathbf{P}}_3$ et $\hat{\mathbf{P}}_4$)

	α_u	α_v	u_0	v_0	\mathbf{r}	\mathbf{t}
\mathbf{P}_1	1465	1466	380	316	$[-0.576, 5.39, 0.278]$	$[-1482, -10.9, -1372]$
\mathbf{P}_2	1393	1391	310	274	$[-0.465, 5.55, 0.292]$	$[-1126, -0.24, -1634]$
$\hat{\mathbf{P}}_3$	1212	1211	396	314	$[-0.538, 5.33, 0.274]$	$[-1212, -2.09, -1181]$
$\hat{\mathbf{P}}_3$	1220	1208	311	308	$[-0.549, 5.40, 0.297]$	$[-1216, -0.24, -1198]$
\mathbf{P}_4	1102	1105	268	256	$[-0.327, 5.55, 0.270]$	$[-849, -10.0, -1406]$
$\hat{\mathbf{P}}_4$	1190	1203	304	280	$[-0.480, 5.51, 0.296]$	$[-955, -3.77, -1535]$

3.2 Stéréoscopie et applications

3.2.1 Reconstruction dense à partir de la vergence et la stéréo

Participants : Reyes Enciso, Zhengyou Zhang, Thierry Viéville

Mots-clés : reconstruction 3D, analyse de scène, stéréovision.

Dans ce travail, nous combinons la stéréo non calibrée, qui fournit une structure projective de l'environnement, et la vergence des caméras actives, qui fournit les coordonnées euclidiennes de quelques points caractéristiques de l'environnement, ceci pour obtenir une reconstruction euclidienne dense et précise d'une scène statique. L'algorithme se déroule en deux étapes. À partir de la connaissance des angles des degrés de liberté de notre robot, et de la cinématique inverse, nous pouvons obtenir les coordonnées euclidiennes d'au moins cinq points dans le repère attaché au robot. Puis, nous calculons la géométrie épipolaire à partir de deux images. En rectifiant les images, nous calculons la carte de disparité. A chaque point où la valeur de la disparité est connue nous pouvons associer des coordonnées projectives. Il suffit de calculer la matrice de distorsion projective qui nous fait passer de la base euclidienne à la base projective et d'appliquer son inverse aux points projectifs pour retrouver la reconstruction 3D de la scène.

Un exemple de résultat issu de [650] est donné figure 8.



Figure 8: Un exemple de reconstruction 3D dense.

3.2.2 Un système de stéréovision pour le véhicule planétaire

Participant : Zhengyou Zhang

Mots-clés : reconstruction 3D, analyse de scène, stéréovision.

Dans les dernières années, notre projet a développé un système de perception complet à base de stéréoscopie, constitué de quatre modules opérationnels et destiné aux applications planétaires :

calibration: une nouvelle technique efficace pour autocalibrer la stéréovision a été développée, qui utilise à la fois des points de référence et des informations provenant de l'environnement;

corrélation: un algorithme de stéréo par corrélation est utilisé pour produire des cartes de profondeur denses;

recalage: un algorithme a été développé pour recalibrer des cartes de profondeur. Il est basé sur l'hypothèse qu'une estimée initiale des déplacements est disponible, donc que l'on peut la raffiner en mettant en correspondance de manière itérative des points d'une vue avec les points les plus proches d'une autre vue. Une méthode fondée sur la statistique de la distribution des distances est utilisée pour éliminer les faux appariements.

fusion: les cartes 3D recalées sont fusionnées de manière incrémentale pour construire un modèle numérique de terrain de l'environnement. Les données erronées sont éliminées en vérifiant la cohérence.

Ce système a été décrit dans un article présenté à *IEEE Workshop on Planetary Rover Technology and Systems* [668].

3.2.3 Reconstruction de modèles de bâtiments à partir d'images non calibrées

Participants : Stéphane Laveau, Cyrille Gauclin, Luc Robert, Olivier Faugeras

Mots-clés : reconstruction 3D, analyse de scène, stéréovision.

L'objet du projet LTR REALISE est de développer les outils théoriques et logiciels permettant de produire un modèle tridimensionnel réaliste d'une scène à partir d'une séquence d'images, sans information préalable sur les paramètres ou le mouvement des caméras.

La géométrie épipolaire des vues est déterminée en utilisant les techniques décrites précédemment. Ceci nous permet de remonter à la structure de la scène à une transformation projective près. Afin de déterminer la structure euclidienne, nous utilisons une information a priori, qui porte soit sur la caméra, soit sur la scène :

- si les paramètres de la caméra sont constants entre les différentes vues, nous appliquons les résultats présentés précédemment,
- sinon, nous utilisons des informations a priori sur la scène. En pratique, ces informations sont fournies par l'utilisateur. En utilisant les relations de parallélisme entre droites visibles dans les images, nous déterminons la structure affine de la scène. Enfin, la connaissance de certains angles ou rapports de longueurs connus permet de remonter à la structure euclidienne.

Une fois les vues calibrées, il reste à construire le modèle géométrique de la scène. Nous adoptons une approche semi-automatique, où l'utilisateur définit un modèle polyédrique en interagissant directement avec les images, et reçoit l'assistance de processus de vision : extraction précise de coins ou de jonctions, mise en correspondance automatique de primitives entre les images. Des textures destinées à être plaquées sur les polygones sont extraites automatiquement des images. À l'aide des techniques de calcul de mosaïques (cf. paragraphe 3.3.2), nous combinons l'information de texture extraite des différentes images, ce qui nous permet d'éliminer automatiquement une grande partie des artefacts visuels (piétons, voitures).

Les procédures de vision sont regroupées par un serveur, auquel le programme de modélisation interactive développé pour REALISE par nos partenaires du Fraunhofer Institut (Darmstadt, DE) se connecte afin d'obtenir les informations extraites des images. De notre côté, nous avons développé, au-dessus de l'environnement *Target-Junior*, un modèleur interactif permettant de valider toutes les fonctionnalités offertes par le serveur vision. La figure 9 montre deux exemples de modèles obtenus à l'aide de notre interface. À gauche, nous montrons une des images utilisées, et le modèle filaire reconstruit. À droite, nous montrons deux vues du modèle texturé produit.

Une publication commune avec nos partenaires a été faite [655].

3.2.4 Synthèse d'images réalistes à partir d'autres images

Participants : Stéphane Laveau, Luc Robert, Olivier Faugeras

Mots-clés : reconstruction 3D, analyse de scène, stéréovision.

L'objet de ce travail est de produire de nouvelles vues, à partir de plusieurs images, sans passer par la reconstruction d'un modèle tridimensionnel explicite.

Nous considérons une paire d'images de géométrie épipolaire connue, pour laquelle une carte de disparité dense a été obtenue à l'aide d'un algorithme de mise en correspondance stéréoscopique. Nous spécifions ensuite une troisième vue, virtuelle, par sa géométrie relativement aux deux images de référence. Nous avons développé deux approches permettant de produire l'image de cette troisième vue à partir des deux autres.



Figure 9: Images et modèles tridimensionnels produits (voir texte).

Transfert pixel-à-pixel

Dans une première approche, nous produisons l'image pixel-à-pixel, par une technique inspirée du lancer de rayon. On détermine, pour chaque pixel de la vue à générer, le premier point de la scène traversé par le rayon optique correspondant. Ce calcul est en réalité effectué directement dans les images de référence, en examinant les pixels en correspondance le long des droites épipolaires orientées.

Transfert par polygones

L'approche précédente pouvant s'avérer lourde en temps de calcul, nous avons développé une autre méthode fondée sur le rendu de texture, qui peut mettre à profit la puissance des stations graphiques. Nous calculons une approximation de la carte de disparités, à base de triangles. Pour chacun de ces triangles, on détermine le triangle correspondant dans la nouvelle vue. La texture prise dans l'une des deux images de référence est attachée au triangle. L'ordre de dessin des triangles, du plus éloigné au plus proche, est déterminé à l'aide de l'orientation des droites épipolaires dans les images de référence. La figure 10 montre un exemple d'application de cette technique. Les triplets d'images pris de deux points distincts (à gauche et à droite du chemin) ont permis de confectionner deux mosaïques, à partir desquelles nous pouvons générer des images vues de points arbitraires, comme par exemple une vue du milieu de la route (en bas). Des séquences de synthèse sont visibles en : <http://www.inria.fr/robotvis/projects/Monaco>

3.3 Perception tridimensionnelle de formes et de mouvement par analyse de séquences d'images

3.3.1 Remplacement de motifs plans dans une séquence d'images en présence d'occlusions

Participants : Imed Zoghliami, Olivier Faugeras, Rachid Deriche

Mots-clés : reconstruction 3D, analyse de scène, analyse du mouvement.

Le problème étudié est le suivi et la modification d'objets dans des séquences d'images couleurs, disposant d'un modèle de l'objet suivi et de celui à insérer, ceci en deux parties :

- suivi de l'objet à modifier,



Figure 10: Images initiales, mosaïques et images de synthèse (voir texte).

- détection d'éventuelles occultations de l'objet suivi.

Deux hypothèses sont faites : l'objet suivi est plan et les objets occultants sont opaques.

- Suivi : l'objet étant plan, il suffit d'en suivre quatre points pour calculer la transformation géométrique (c'est à dire l'homographie) entre le modèle et l'objet suivi. Pour des raisons de robustesse, nous suivons un nombre plus important de points, puis nous calculerons l'homographie avec la méthode robuste des moindres carrés médians.
- Occultation : pour la détection, nous utiliserons une méthode basée sur la corrélation. Ainsi, via l'homographie, chaque point de l'objet suivi a son correspondant sur le modèle. Pour détecter les occultations, nous corrélerons l'objet suivi avec le modèle dans les trois composantes couleur; les points de faibles scores de corrélation sont classés comme points de l'objet occultant. Ce procédé est raffiné de trois manières : d'abord en gérant les frontières au niveau de la corrélation, ensuite en gérant l'évolution temporelle des différentes zones et enfin nous utiliserons des Snakes couleurs sur les contours de l'occultation pour corriger les bavures de la corrélation et pour avoir une qualité de détection subpixelique.
- Exemple : Ci-dessous figure 11 un exemple de résultat obtenu avec cette méthode : le panneau «ORANGINA» partiellement occulté par le joueur a été remplacé par le panneau «ROBOTVIS».

3.3.2 Calculs de mosaïques d'images et applications

Participants : Imed Zoghalmi, Olivier Faugeras, Luc Robert

Mots-clés : reconstruction 3D, analyse de scène, analyse du mouvement.



Figure 11: Image originale (à gauche), et image modifiée (à droite).

Si nous observons une scène statique avec une caméra en rotation pure autour de son centre optique, nous obtenons des images qui sont deux à deux en correspondance homographique. Cette propriété est aussi valable lorsqu'on observe une surface plane depuis plusieurs points de vue.

En s'appuyant sur des techniques de corrélation, on détermine les homographies entre les différentes images. La structure ainsi obtenue (images + homographies) est appelée *mosaïque*. Pour visualiser une mosaïque, il suffit de ramener toutes les images dans le même repère, en leur appliquant les homographies adéquates, puis de combiner les intensités aux pixels correspondants.

Nous avons utilisé les mosaïques dans les deux contextes suivants :

- La combinaison des images initiales en une image composite permet d'élargir l'angle de vue, tout en conservant les propriétés géométriques du modèle sténopé. En particulier, on peut appliquer les algorithmes de stéréo (calcul de la géométrie épipolaire, mise en correspondance, synthèse de nouvelles vues) aux mosaïques (cf paragraphe 3.2.4).
- En utilisant la cohérence entre les différentes images, on peut éliminer les objets qui ne satisfont pas la correspondance homographique : dans le cas d'une caméra en rotation autour de son centre optique, ce sont les objets en mouvement; dans le cas d'une surface plane observée de plusieurs points de vue, ce sont les objets qui ne se trouvent pas dans le plan, c'est-à-dire les obstacles visuels (cf paragraphe 3.2.3). L'exemple de la figure 12 montre une mosaïque de 5 images où nous avons effacé les personnages en mouvement.



Figure 12: Haut: cinq images d'une même scène. Bas: la mosaïque obtenue, avec élimination des personnages en mouvement.

3.3.3 Analyse des mouvements de la main dans des séquences d'images

Participants : Quentin Delamarre, Olivier Faugeras

Mots-clés : reconstruction 3D, analyse de scène, analyse du mouvement.

Détecter et estimer la position d'une main dans une séquence d'images est un problème difficile. Pour y parvenir, nous nous servons d'un modèle 3D (voir fig 13 à gauche), puis nous exécutons les étapes suivantes :

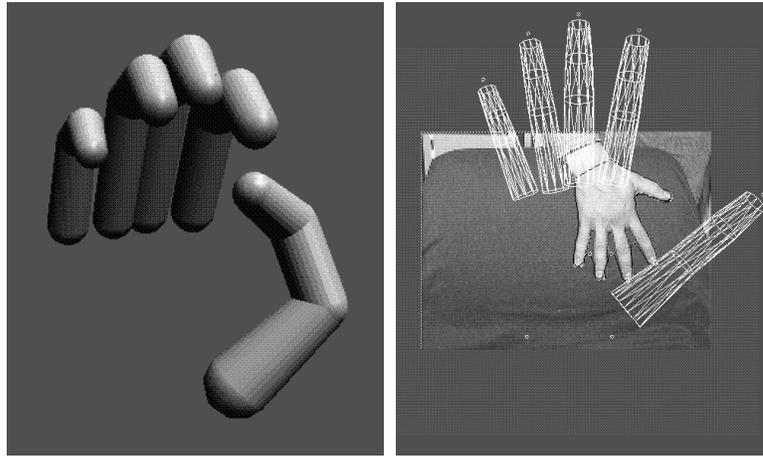


Figure 13: A gauche : une vue du modèle 3D de la main en utilisant les bibliothèques graphiques OpenGL. A droite : mise en correspondance automatique entre le modèle et l'image (ici les extrémités des doigts).

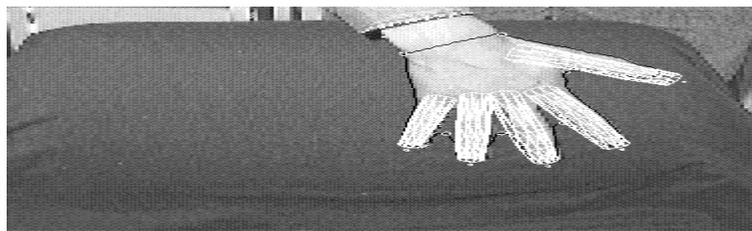


Figure 14: Superposition du modèle 3D avec la main de l'utilisateur.

- détecter les contours dans les images,
- établir des correspondances entre le modèle et ces contours (voir fig 13 à droite),
- en déduire la position dans l'espace de la main de l'utilisateur (voir fig 14),
- et enfin, prédire la position la plus probable de la main dans les images suivantes.

Le fait que la main soit un objet articulé rend son étude complexe. C'est pourquoi nous nous sommes limités, dans un premier temps, à la reconnaissance d'une main plane (aucun doigt plié) dans une image.

Utiliser la vision par ordinateur permet d'éviter à l'utilisateur de mettre un gant recouvert de capteurs de position, donc fragile. Cela permet aussi de travailler sur des séquences d'images prises à l'avance.

Les applications d'un tel système sont nombreuses : reconnaissance automatique et transcription en mots du langage des sourds-muets, interface homme-machine améliorée (« souris » 3D), manipulation d'objets à distance, etc ... On peut imaginer étendre ce système à d'autres objets connus (piétons, bras manipulateurs, ...).

Il est clair que le système final utilisera différentes méthodes complémentaires. Les contours actifs géodésiques fournissent déjà la silhouette de la main. L'étude de l'intérieur de ce contour fournira d'autres informations utiles. La stéréovision est une alternative intéressante.

Beaucoup de travail reste à faire dans ce domaine en pleine expansion.

3.3.4 Utilisation des singularités dans l'analyse du mouvement non calibré

Participants : Diane Lingrand, François Gaspard, Thierry Viéville

Mots-clés : reconstruction 3D, analyse de scène, analyse du mouvement.

Cette étude a permis l'analyse et l'expérimentation d'une méthode d'estimation de la structure d'une scène et des paramètres de calibration, ceci à partir d'un système visuel monoculaire non calibré. Dans un premier temps, des mouvements particuliers permettant la calibration du système ont été étudiés à partir d'un travail préliminaire sur la perception visuelle sans calibration. Ensuite, les différents modèles nécessaires à la calibration du système et à la perception de la structure ont été implémentés.

L'expérimentation pratique nous a ensuite permis d'améliorer le système initial et d'obtenir des résultats effectifs.

Le système robotique Argès

Afin d'expérimenter les algorithmes, nous utiliserons le système robotique Argès¹ (figure 16) comme plate-forme expérimentale. Ce système comporte une caméra à zoom et focus télécommandés et permet de générer des mouvements particuliers. En effet, il est constitué d'une tourelle montée sur un rail : figure 15. Nous avons donc une translation (le rail) et deux rotations (roulis et tangage). Le mouvement de roulis est excentré par rapport au centre optique de la caméra et ne permet pas de réaliser des rotations pures. Mais nous avons un axe de rotation pure : le mouvement de tangage.

Exemples de résultats obtenus.

Nous nous proposons de réaliser l'estimation des structures planaires de l'image et de la calibration en deux phases :

¹Toute documentation sur Argès est accessible via :
<http://www.inria.fr/robotvis/personnel/vthierry/acvis-demo/argès/argès.html>

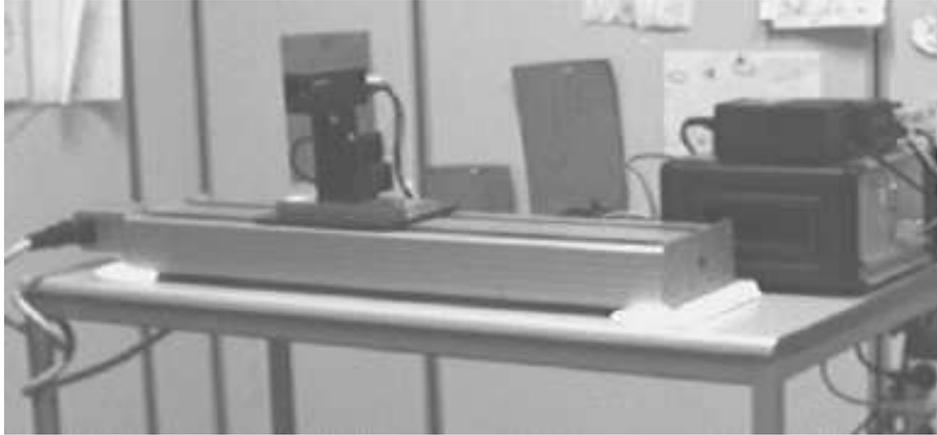


Figure 15: Le système robotisé Argès.

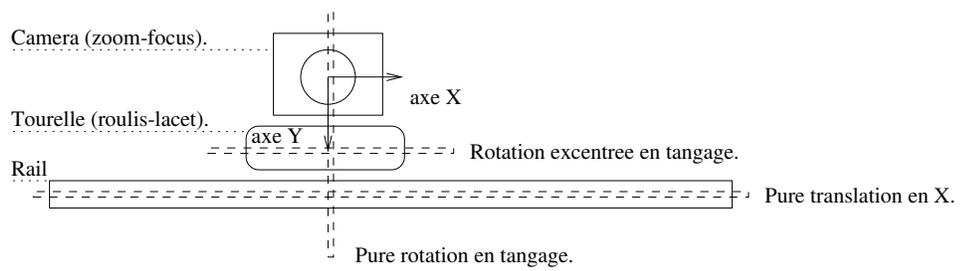


Figure 16: Schéma du système Argès.

- Nous effectuons un mouvement de translation, permettant d’estimer les différentes collinéations de la scène de manière plus stable. Nous recherchons alors, parmi ces collinéations, la collinéation fronto-parallèle la plus éloignée du centre optique de la caméra, qui correspondra à la collinéation du plan à l’infini.
- Un mouvement de tangage nous permet alors d’estimer les différents paramètres de calibration. Nous utilisons les appariements sélectionnés à l’étape précédente afin d’estimer \mathbf{H}_∞ . De plus nous utilisons un modèle adéquat permettant d’estimer \mathbf{H}_∞ avec un nombre de paramètres réduit.

Pour ce faire, nous supposons uniquement les hypothèses suivantes :

- nous observons des points de l’horizon;
- nous connaissons le sens de la translation;
- lors du mouvement de tangage, les paramètres intrinsèques de la caméra ne varient pas.

Nous pouvons alors estimer les structures planaires de la scène ainsi que les paramètres intrinsèques de la caméra.

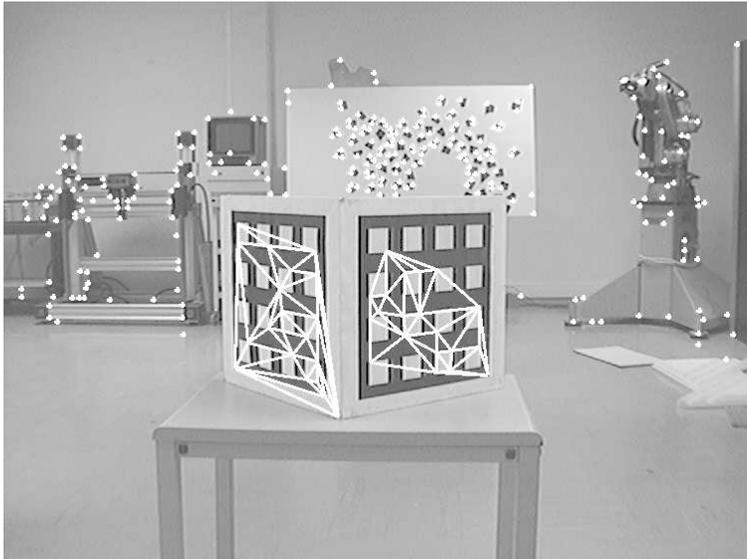


Figure 17: Exemples de structures planaires détectées par l’algorithme. permettant d’en déduire les paramètres intrinsèques de la caméra.

Utilisation de mouvements particuliers.

Certains mouvements particuliers donnent des matrices fondamentales ou des collinéations admettant un nombre réduit de paramètres. Etant donné que l’on s’intéresse à des systèmes robotisés, pour lesquels les mouvements ne sont pas toujours des mouvements rigides généraux, nous introduisons différents modèles particuliers.

Dans ces cas particuliers, l’estimation des matrices fondamentales ou des collinéations est plus simple et parfois les modèles choisis nous permettent d’obtenir des contraintes sur les paramètres de calibration. Pour une revue complète de ces mouvements, consulter [666].

3.3.5 Un nouvel algorithme pour l'estimation du mouvement et de la structure : des paramètres essentiels au mouvement euclidien à travers la matrice fondamentale

Participant : Zhengyou Zhang

Mots-clés : reconstruction 3D, analyse de scène, analyse du mouvement.

La technique classique pour estimer le mouvement et la structure à partir de deux projections perspectives se compose de deux étapes : (i) utiliser l'algorithme des 8-points pour estimer de manière linéaire les 9 paramètres essentiels définis à un facteur d'échelle près, (ii) raffiner l'estimation du mouvement à base d'un critère statistiquement optimal. Le problème (ii) est un problème d'estimation non linéaire sur un espace à 5 dimensions. Malheureusement, les résultats obtenus avec cette technique ne sont souvent pas satisfaisants, surtout quand le mouvement est petit ou quand les points observés sont proches d'une configuration dégénérée (par exemple une surface plane). Un problème important est que la deuxième étape est très sensible à l'estimée initiale et qu'il est très difficile d'obtenir une estimée initiale précise à partir de la première étape par projection. Nous proposons dans [675] une nouvelle technique introduisant une étape intermédiaire qui consiste en une estimation d'une matrice 3×3 définie à un facteur d'échelle près avec la contrainte de déterminant nul (donc cette matrice a 7 paramètres indépendants et est connue sous le nom de la matrice fondamentale). L'idée est de projeter de manière progressive les paramètres estimés dans un espace de dimension élevée sur un espace de dimension un peu plus faible. Pour notre application, nous passons progressivement de la dimension 8 à 5, en passant par la dimension 7. La technique proposée a été testée avec des données synthétiques et réelles, et une amélioration considérable a été observée dans les situations délicates mentionnées précédemment. À partir de ce travail, nous formons une conjecture sur la nécessité de l'estimation de la géométrie épipolaire non calibrée comme étape intermédiaire pour obtenir une estimation fiable du mouvement et de la structure euclidiens à partir d'images multiples calibrées.

Un exemple est montré en figure 18, avec deux images d'une scène de cailloux. Parce que la majorité de la scène observée est très plate et que le déplacement entre les deux images est faible, la technique classique n'a pas pu retrouver le mouvement et la structure. Avec notre nouvel algorithme, ils ont été bien recouverts. Une vue de face et une vue de dessus de la scène reconstruite sont montrés en figure 19, où l'on remarque bien, respectivement, le gros cailloux à droite et le cône de perception de la caméra. Une représentation plus visuelle mais moins objective est celle de stéréogrammes, qui sont montrés en figure 20.

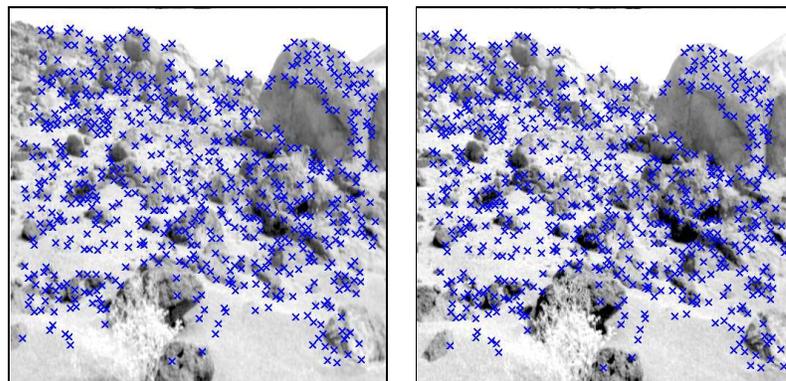


Figure 18: Deux images d'une scène de cailloux

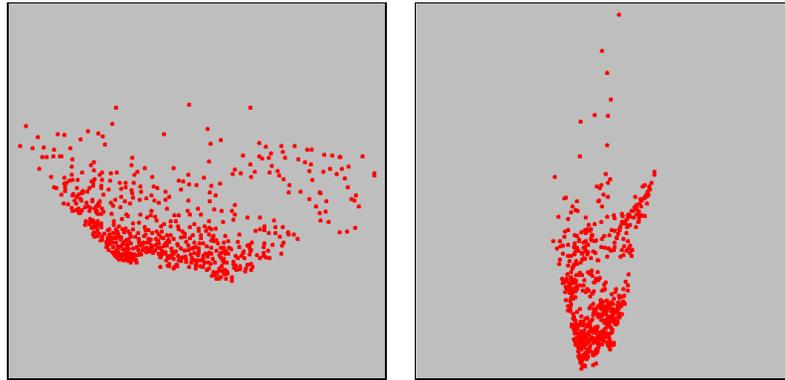


Figure 19: Les points 3D reconstruits de la scène de cailloux : vues de face et de dessus

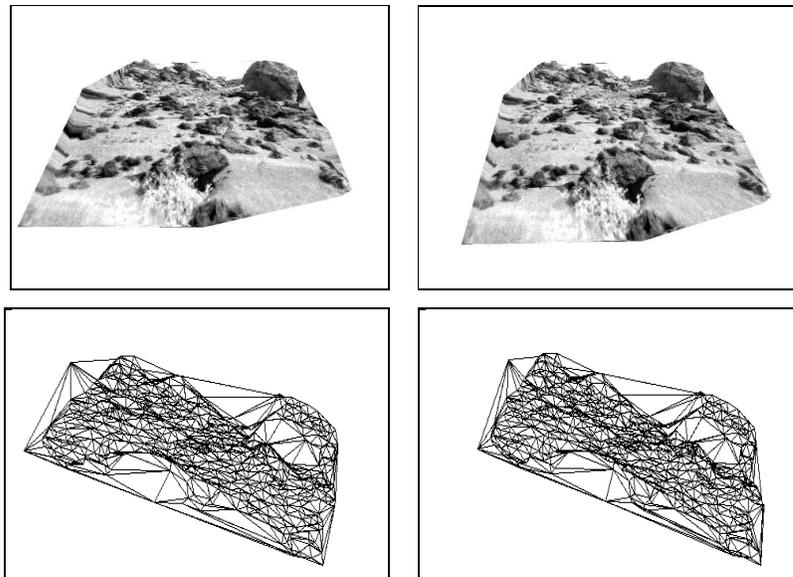


Figure 20: Stéréogrammes de la scène reconstruite.

3.4 Équations aux dérivées partielles, images et vision par ordinateur

3.4.1 EDP et restauration d'images

Participants : Pierre Kornprobst, Rachid Deriche

Mots-clés : traitement d'image, vision 3D.

Un certain nombre d'approches permettant de restaurer avec succès des images dégradées ont récemment vu le jour [638]. Ces approches, qui font référence maintenant en traitement des images, ont été proposées dans la littérature de manière indépendante, mais utilisent pour la plupart d'entre elles une méthodologie commune à base d'EDP [630]. A cause des fortes potentialités au niveau des applications qu'offre la restauration d'images dégradées, nous nous sommes tout naturellement intéressés à ce problème, aussi important en traitement d'images qu'en vision par ordinateur. Un certain nombre d'approches parmi les plus remarquables (Alvarez & Morel, Rudin & Osher, Malladi & Sethian ...) ont été mises en œuvre pour une évaluation et une comparaison expérimentale et quantitative précises de leur performances. Ce travail a été complété par le développement d'une nouvelle méthode de restauration d'images bruitées permettant de préserver au mieux les discontinuités présentes dans l'image originale. Cette méthode, qui fait intervenir un terme de régularisation d'attache aux données dans la fonctionnelle à minimiser, a été aussi mise en œuvre. Les premières conclusions des nombreuses évaluations expérimentales menées ainsi que le descriptif détaillé de toutes les méthodes mises en œuvre peuvent être consultés dans l'article [653].

3.4.2 EDP et stéréo

Participants : Rachid Deriche, Christophe Bouvin, Olivier Faugeras

Mots-clés : traitement d'image, vision 3D.

Faisant suite à notre publication [663] permettant de remonter par stéréoscopie à une carte de profondeur tout en préservant les discontinuités du relief, nous nous sommes intéressés à la mise en correspondance de courbes planes à partir d'une paire d'images stéréoscopiques.

On se donne une scène observée par un système stéréoscopique faiblement calibré, une courbe fermée englobant un ensemble d'objets dans une des images et un plan virtuel dans l'espace, spécifié de manière interactive par l'appariement de 3 points lui appartenant. La méthode développée permet alors de faire converger la courbe initiale, spécifiée par l'opérateur, vers un ensemble de courbes appariées dans les 2 images et appartenant au plan spécifié par l'opérateur.

L'idée principale est d'associer à la courbe initiale, ainsi qu'à la courbe homographiquement associée dans la seconde image, une loi d'évolution lui permettant de se déplacer sous l'action de forces internes et externes afin de minimiser une énergie appropriée, comprenant un terme de corrélation ainsi qu'un terme de régularisation sur les courbes appariées.

L'EDP associée à l'équation d'Euler-Lagrange de la fonctionnelle, similaire par certains aspects à celle issue des travaux de Caselles, Kimmel et Sapiro, est alors considérée comme un contour actif géodésique attirant les courbes vers un puit de potentiel de corrélation.

Les figures 21 et 22 illustrent la qualité de l'appariement obtenu sur des objets divers. Plus de résultats peuvent être trouvés dans les références [669] et [647].

3.4.3 EDP et flot optique

Participants : Pierre Kornprobst, Rachid Deriche

Mots-clés : traitement d'image, vision 3D.

Le principe de l'approche par minimisation et régularisation d'énergie, qui consiste à résoudre le problème :

$$\inf_{V \in \mathcal{E}} \int_{\Omega} (\nabla I \cdot V + I_t)^2 + \alpha(\phi(|Du|) + \phi(|Dv|)) d\Omega$$

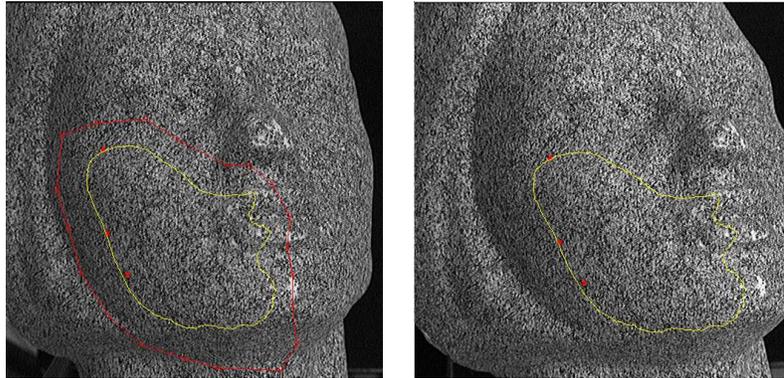


Figure 21: Appariement de courbes par EDP et contours actifs géodésiques sur un visage : remarquer l'initialisation qui peut être très lointaine

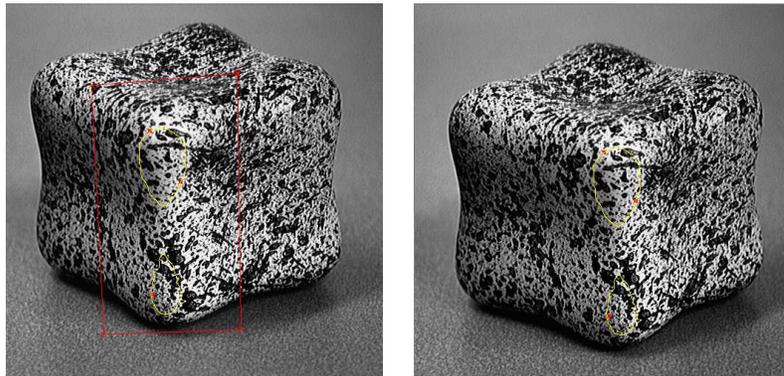


Figure 22: Appariement de courbes par EDP et contours actifs géodésiques sur un objet de forme cubique. Observer le changement de topologie de la courbe à la convergence

où \mathcal{E} est un espace fonctionnel adapté, nous a permis de retrouver un flot optique avec préservation des discontinuités (voir travaux antérieurs). Deux exemples de séquences synthétiques (voir figure 23) nous permettent de juger qualitativement les résultats obtenus. Dans ces exemples simples, on observe que seule la composante normale de la vitesse a pu être retrouvée dans le cas (1). De plus, des différences apparaissent, selon l'initialisation, dans les zones où $|\nabla I| = 0$. Ceci pose tout naturellement la question de l'existence et de l'unicité d'un tel problème.

Comme pour la restauration d'images, l'espace fonctionnel de travail adapté est l'espace $BV(\Omega)$, des fonctions à variations bornées. L'étude de l'existence et de l'unicité d'un tel problème devra prendre en compte les zones où $I_x = 0$. Cela nous conduira à ajouter des contraintes sur \mathcal{E} ou bien à proposer un nouveau modèle. Ensuite, nous étudierons le cas du flot optique 2-D pour lequel des spécificités apparaissent. En particulier, on s'intéressera à une nouvelle modélisation susceptible de mieux prendre en compte le problème d'ouverture. Des expériences numériques, qui compléteront celles déjà réalisées, seront ensuite menées pour tester la validation du modèle.

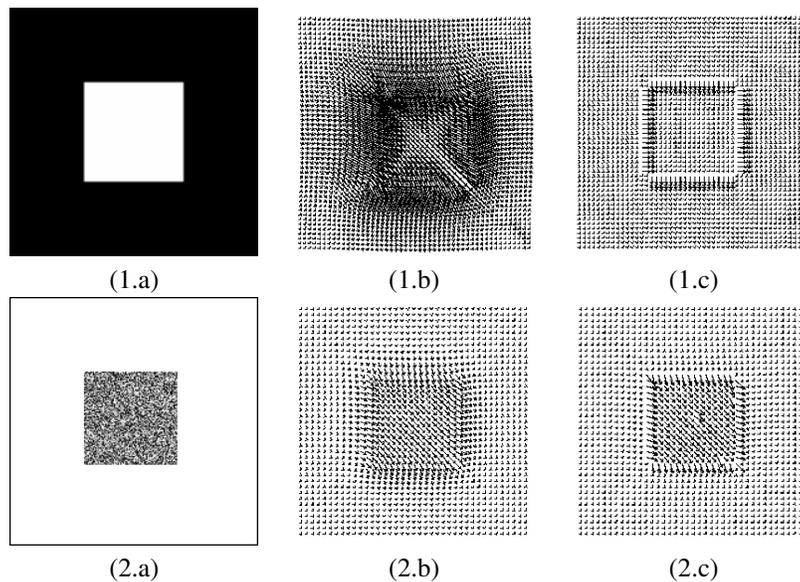


Figure 23: Les séquences étudiées (1.a) et (2.a) représentent un carré texturé ou non, en translation suivant la diagonale. (1.b) et (2.b) sont les résultats obtenus avec $\phi(x) = x^2$ (Horn et Schunck). (1.c) et (2.c) sont obtenus avec la fonction $\phi(x) = \log(1 + x^2)$ qui permet de préserver les discontinuités.

3.4.4 EDP et analyse multiéchelle

Participants : Laurence Lucido, Rachid Deriche, Zhengyou Zhang

Mots-clés : traitement d'image, vision 3D.

Ces travaux font l'objet d'une thèse effectuée dans le cadre d'une collaboration avec le Laboratoire de Robotique et d'Intelligence Artificielle (DITI - SM - RIA) d'IFREMER (centre de Toulon). Ils concernent la mise en correspondance de cartes multiéchelles pour la navigation autonome des engins sous-marins. Il s'agit de recalibrer une bathymétrie de forte résolution prise localement par un engin autonome ou remorqué et une carte de référence connue a priori mais de faible résolution.

Nous utilisons des données bathymétriques, préférées à l'imagerie sonar et l'algorithme de recalage qui a commencé à être développé se décompose ainsi :

- extraction de primitives bathymétriques de fortes pentes (cliffs maps, i.e. contours de falaises) et de points critiques (ou de forte courbure) sur ces cliffs;

- appariement et estimation de la position et de l'orientation du véhicule;
- analyse multiéchelle, approche issue des travaux du Ceremad pour son adéquation à notre problème en terme d'invariance morphologique et euclidienne.

À ce jour, les outils généraux de mise en correspondance ont été développés : analyse des profils bathymétriques, extraction d'attributs invariants à la rotation et à la position, appariement. Des tests statistiques réalisés en ajoutant à la carte locale un bruit gaussien de 10m sur la profondeur (ce qui est considérable par rapport à la qualité physique des sondeurs multifaisceaux utilisés) montrent que, sur 500 réalisations de ce bruit, la qualité du positionnement est de 1.1 pixels en x et 1.5 pixels en y.

Nous nous orientons désormais sur une analyse multiéchelle de ces profils afin que la mise en correspondance soit aussi invariante à l'échelle de ces profils. Un point important sera l'application de notre algorithme à des séquences d'images réelles issues du sonar panoramique installé sur le robot VORTEX du laboratoire RIA (Robotique et Intelligence Artificielle) de l'Ifremer, ainsi que l'intégration des informations provenant des capteurs d'estime (vitesse, etc.).

Pour plus d'information, on trouvera dans les références [659, 661, 658, 660], les détails sur les différentes parties de ce travail.

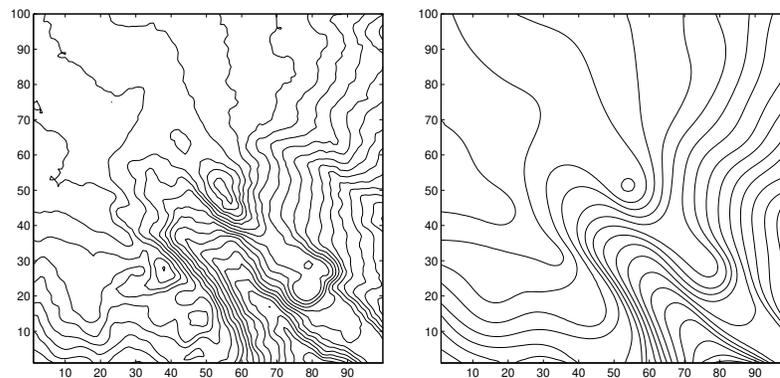


Figure 24: Application de l'EMSS à une image bathymétrique réelle

3.4.5 EDP, évolution de courbes, de surfaces et stéréoscopie

Participants : Renaud Keriven, Olivier Faugeras

Mots-clés : traitement d'image, vision 3D.

Nous avons poursuivi nos travaux sur l'évolution des courbes projectives qui nous ont permis de clarifier les relations entre notre approche intrinsèque et celle de Olver-Sapiro-Tannenbaum. Nous avons en particulier obtenu l'équation générale de toutes les évolutions invariantes projectives [652].

Nous avons aussi abordé une étude plus proche des applications qui s'appuie, d'une part, sur l'expérience acquise lors de nos travaux sur l'évolution affine et projective des courbes et, d'autre part, sur les travaux rapportés dans la section précédente. L'idée consiste, à partir d'une formulation variationnelle de la stéréoscopie, à exprimer le problème comme un problème d'évolution de surface (ou de courbe en deux dimensions). Nous avons écrit les équations d'évolution d'une manière intrinsèque et développé une implémentation préliminaire en deux dimensions de ces idées qui peut travailler avec un nombre arbitraire de caméras. Les résultats semblent prometteurs.

L'implémentation dont nous disposons suppose que les caméras sont des caméras linéaires mais leur nombre est arbitraire. L'objet analysé est un cercle dont la texture est représentée en figure 25.

La figure 3.4.5 montre la géométrie de la scène : le cercle, entouré des caméras, et l'évolution de la surface de l'objet calculé par stéréo au fil des itérations.

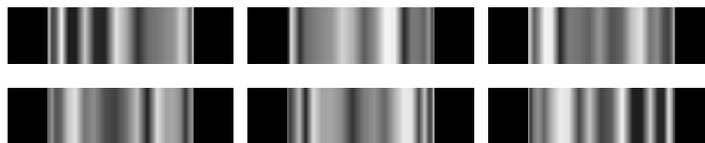
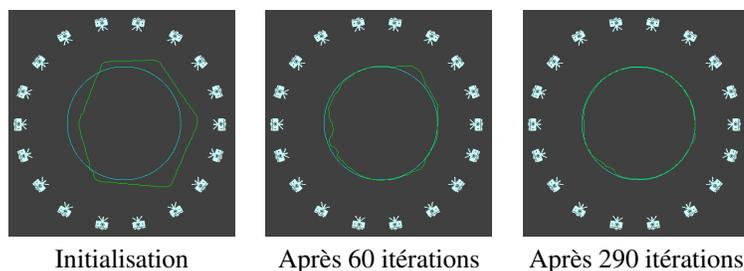


Figure 25: Images du cercle vu tous les 60 degrés



Initialisation

Après 60 itérations

Après 290 itérations

4 Actions industrielles

- Dans le cadre du **projet Esprit REALISE**, nous collaborons avec la société SYSECA en France, le Fraunhofer de Darmstadt (DE) et l'université d'Utrecht (NL). Ce projet a pour but d'automatiser partiellement la construction de modèles de bâtiments, de type CAO, à partir de séquences d'images.
- Dans le cadre du **projet Esprit LTR CUMULI**, nous entamons une collaboration avec la société Innovativ-Vision (SE), et Imetrics (CH).
- Dans le cadre du **projet Esprit HPCN EPSIS**, nous poursuivons notre collaboration avec les sociétés Symah-Vision et Matra Cap Systèmes et entamons une collaboration avec Digitale Vi-deosysteme (DE), Mondial Promotions SA (GR) et Radamec Broadcast Systems LTD (GB).
- Dans le cadre du **projet Esprit LTR IMPROOFS**, nous entamons une collaboration avec Forensic Science Service (GB), Nationaal Instituut voor Criminalistiek en Criminologie (BE). Nous espérons aussi approfondir à cette occasion nos relations avec les sociétés NOESIS (FR), Matra Cap Systèmes (FR) et Cognitek (US).
- Dans le cadre du **projet ESA Calvin**, nous collaborons avec l'ESA, ITM APTORI, KTH/NADA (Stockholm, SE) TECHNOSPAZIO et le projet INRIA Rhône-Alpes (Movi). Le thème est l'étalonnage de caméras. Les résultats sont le transfert chez nos partenaires de notre expertise en étalonnage automatique de caméras et l'amélioration de techniques d'étalonnage d'un système de stéréovision.
- Dans le cadre de **l'action de recherche et de développement Praxitèle**, nous participons à la construction de modèles tridimensionnels de parkings et la localisation visuelle de véhicules électriques automatiques.

5 Actions nationales et internationales

5.1 Actions nationales

- T. Viéville et S. Arias collaborent avec le projet ICARE (E. Coste-Manière) sur les problèmes de génération automatique de code et de gestion de modules de perception réactifs.

- R. Deriche est membre du comité de rédaction de la revue scientifique TS *Traitement du Signal* et a été membre du comité de programme du *Dixième Congrès sur la Reconnaissance Des Formes Et l'Intelligence Artificielle* (Rennes, 1996). Il est membre du comité des professeurs invités et du comité des bourses à l'UR de Sophia. Il est coanimateur du GDR Mathématiques des systèmes perceptifs et cognitifs.
- Z. Zhang a participé à GDR-PRC Communication Homme-Machine, Groupe de travail sur la perception 3D et celui sur l'indexation d'images.

5.2 Actions Européennes

L'ensemble de nos actions européennes concerne l'Europe de l'Ouest.

5.2.1 Projets de Recherche Communautaires

Nous avons trois projets européens de type LTR en cours (dont deux obtenus cette année) et deux autres projets :

- REALISE : commencé début 1994, se termine début 97. L'objet est la reconstruction de modèles de bâtiments pour la réalité virtuelle à partir de séquences d'images non calibrées. Nos partenaires sont IGD-Fraunhofer Institut (DE), Flux-University of Utrecht (NL). Le coordinateur est Thomson-Syseca (FR). Page web :
<http://www-uk.research.ec.org:80/esp-syn/text/8878.html>.
- CUMULI : partenaires INRIA Rhône-Alpes (Projet Movi, Coord. Roger Mohr), université de Lund (SE), Fraunhofer Darmstadt, IMETRICS (CH), Innovativ-Vision (SE). Durée : 1996-1999. Modélisation tridimensionnelle par vision par ordinateur, analyse de multiséquences d'images, automatisation du raisonnement géométrique. Page web :
<http://www.inrialpes.fr/CUMULI/>.
- IMPROOFS : IMAge PROcessing Operations for Forensic Support. Partenaires : universités d'Oxford (GB) et catholique de Louvain (BE), Royal Institute of Technology (KTH, SE), Forensic Science Service (GB), Nationaal Instituut voor Criminalistiek en Criminologie (BE). Application de techniques de traitement et d'analyse des images aux problèmes de criminalité.

Nous avons obtenu un projet Esprit HPCN :

- EPSIS : partenaires : Symah-Vision (FR), Digitale Videosysteme (De), Matra Cap Systèmes (FR), Mondial Promotions SA (Grèce), Radamec Broadcast Systems LTD (GB). Durée : 1996-1998. Analyse en temps réel de séquences d'images vidéo.

Nous poursuivons le projet Esprit KIT :

- ARTVIS : partenaires : l'Instituto de Automatica Industrial-Madrid (CSIC-IAI), university of Vitoria - UFES (Brésil). Durée : 3 ans à partir de janvier 1995. Organisations de cours, séminaires et ateliers de travail sur la Vision ainsi que des visites entre les partenaires.

5.2.2 Réseaux d'excellence

Nous participons aux deux réseaux d'excellence IT ECVNET-8212 et VIRGO :

- ECVNET European Computer Vision Network. Partenaires : pratiquement tous les centres de recherche en vision par ordinateur européens. Réseau d'excellence européen pour promouvoir la Recherche, l'Enseignement et les applications de la vision par ordinateur. Page web :
<http://afrodite.lira.dist.unige.it/fullservice.html>.

- VIRGO Vision-Based Robot Navigation Research Network. Partenaires : ICS-Forth (Grèce, Coord. Stelios Orphanoudakis), universités d'Aalborg (DK), de Gênes (It.), de Graz (Autriche), KTH (Suède), GMD (D), de Zürich (Suisse), de Copenhague (DK), de Bonn (D). Durée : 1996-1999. Coordination de la recherche européenne dans le domaine du développement de systèmes robotiques intelligents capables de naviguer dans des environnements partiellement inconnus et éventuellement variables. Page web : <http://www.ics.forth.gr/virgo/>.

5.2.3 Divers

- R. Deriche est, depuis Janvier 1993, chargé des relations scientifiques européennes au sein de l'UR de Sophia. En particulier, il a créé cette année une page html sur les activités européennes de l'Institut, en <http://www.inria.fr/robotvis/personnel/der/ec/ec.html>.
- Nous avons obtenu un financement d'un montant de 200kEcus de la CE pour des chercheurs postdoctorants avec les projets Epidaure et Pastis dans le cadre de bourses HCM institutionnelles. La sélection des candidats est en cours. Nous avons depuis un boursier HCM individuel, Charlie Rothwell, pour une durée de deux ans.

5.3 Actions internationales

L'ensemble des actions internationales concerne l'Amérique du Nord, le Japon et la Chine.

- Le projet participe à IUE (Image Understanding Environment) de plusieurs façons. O.D. Faugeras est membre du «IUE board» qui supervise l'avancement du projet et T. Viéville est responsable du comité d'ingénierie du logiciel dans le réseau d'excellence ECVNET financé par la CE. Dans ce cadre, Thierry Viéville et Diane Lingrand ont organisé un cours international de vision industrielle intitulé «Utilisation d'IUE et de Target-Junior». du 28 juillet au 3 août 1996 à Sophia Antipolis.
Nous participons à l'évaluation des environnements logiciels pour la vision par ordinateur, et à l'expérimentation de portage de modules logiciels sous IUE de AAI/USA et Target-Junior.
- O.D. Faugeras est membre du conseil des sages de la conférence *European Conference on Computer Vision* (Cambridge, 1996). Il a présidé le comité scientifique de la conférence *12th International Conference on Analysis and Optimization of Systems. Images, Wavelets and PDE's*, ICAOS'96 (Paris, 1996). Il est membre de plusieurs comités de rédaction de revues scientifiques, il est notamment co-éditeur en chef, avec Takeo Kanade, de l'*International Journal of Computer Vision*.
- R. Deriche a été membre du comité de programme des conférences *European Conference on Computer Vision* (Cambridge, 1996). Il est membre du comité d'organisation de la conférence *12th International Conference on Analysis and Optimization of Systems. Images, Wavelets and PDE's*, ICAOS'96 (Paris, 1996) et est chargé des relations scientifiques internationales pour l'UR de Sophia depuis avril 1996.
- Z. Zhang est «Guest Research Professor» à l'académie des sciences de Chine depuis 1994 et professeur à temps partiel à l'université de Northern Jiatong, Beijing, Chine, depuis 1996. Il a été membre du comité de programme de la conférence américaine CVPR'96. Il est invité pour un an, à partir de novembre, au centre de recherche ATR à Osaka, Japon
- T. Viéville est membre du comité de programme de CVPR'97 et conférencier invité de ICARCV'96.

6 Diffusion des résultats

6.1 Enseignements

Les membres du projet ont assuré au total 300 heures d'enseignement :

6.1.1 Enseignement universitaire

Université de Nice-Sophia Antipolis : DEA Aravis : Rachid Deriche est responsable du module de techniques avancées pour l'imagerie et la vision. Olivier Faugeras et Luc Robert sont co-responsables du module de Vision 3D, Thierry Viéville et Zhengyou Zhang sont co-responsables du module de Vision dynamique. Chaque module est d'une durée de 15 heures. Rachid Deriche et Olivier Faugeras sont membres du comité scientifique de ce DEA. Rachid Deriche en est de plus le responsable des stages proposés aux élèves.

Université de Paris XI Orsay : DEA d'Informatique commun à Paris XI, ENS Cachan, Supelec : Rachid Deriche est responsable du module Perception 3D qu'il enseigne en commun avec Nicholas Ayache (projet Epidaure) et Luc Robert (24 heures).

Institut National des Télécommunications- Evry : École d'Ingénieurs - 3ème année - Option: Traitement et Applications de l'Image. Rachid Deriche enseigne au sein des modules de « Vision par ordinateur » et « Imagerie dynamique » (24 heures).

EERIE : École pour les Études et la Recherche en Informatique et Electronique - Nimes - École d'Ingénieurs - 3ème année. Option: Imagerie Numérique. Rachid Deriche est responsable du module « Vision par Ordinateur » (24 heures).

École Polytechnique : Olivier Faugeras y a été Maître de Conférence jusqu'au 31/08/96. Il y a développé un cours de Vision algorithmique (32 heures) et un cours de méthodes variationnelles pour la vision par ordinateur en commun avec Jean-Michel Morel (32 heures). Il a enseigné en commun avec Rachid Deriche et Luc Robert un module de 21 heures de Perception 3D dans le DEA IMA, puis Algorithme, commun à l'ENS Cachan, Ulm et aux universités Paris VI, VII, et XI. Jusqu'en 1996 il a été responsable de la filière Géométrie, images et robotique du DEA Algorithmique.

École Supérieure des Sciences de l'Informatique (ESSI) : Thierry Viéville est responsable de l'enseignement de Vision 3D et de l'enseignement de Vision Temps-Réel depuis 1993 (80heures/an).

ISIA : Thierry Viéville y est responsable du cours de Calcul Symbolique (6 heures).

École des Mines de Paris : Thierry Viéville y est responsable du cours de Vision Robotique (30 heures dont 6 enseignées par lui, 6 par Rachid Deriche).

6.1.2 Thèses

Le projet est équipe d'accueil de doctorants des formations doctorales DEA Algorithme (Paris VI, VII, XI), DEA Aravis (Nice-Sophia Antipolis), DEA Informatique (Paris XI, VI, VII).

Thèses en cours :

- Thierry Blaszk, *Contributions et applications des approches par modèles en vision par ordinateur*, université de Nice-Sophia Antipolis,
- Didier Bondyfalat, *Démonstration automatique de théorèmes de géométrie et vision par ordinateur*, université de Nice-Sophia Antipolis.

- Sylvain Bougnoux, *Modélisation de l'environnement pour la localisation visuelle de véhicules automatiques*, université de Nice-Sophia Antipolis,
- Quentin Delamarre, *Analyse des mouvements du corps humain par vision par ordinateur*, université de Nice-Sophia Antipolis,
- Frédéric Devernay, *Stéréoscopie surfacique et applications*, École Polytechnique,
- Cyrille Gauclin, *Construction et maintenance d'une base de données géométriques tridimensionnelles à partir d'une séquence d'images non calibrées*, université de Nice-Sophia Antipolis,
- Renaud Keriven, *Évolution de courbes et de surfaces par des équations de la chaleur intrinsèques*, École Nationale Supérieure des Ponts et Chaussées,
- Pierre Kornprobst, *EDP et vision par ordinateur*, université de Nice-Sophia Antipolis,
- Diane Lingrand, *Étude du mouvement et de la structure au sein d'une séquence d'images non calibrées : approche continue*, université de Nice-Sophia Antipolis,
- Laurence Lucido, *Navigation autonome référencée terrain par mise en correspondance de cartes multiéchelles*, université de Nice-Sophia Antipolis,
- Robert Stahr, *Analyse du mouvement*, université de Nice-Sophia Antipolis,
- Imed Zoghliami, *Analyse de mouvements particuliers dans des séquences d'images couleur*, université de Nice-Sophia Antipolis.

Thèses soutenues en 1996 :

- Gabriella Csurka, *Modélisation projective des scènes tridimensionnelles en vision par ordinateur*, université de Nice-Sophia Antipolis (3/04/96),
- Reyes Enciso, *Auto-calibration des capteurs visuels actifs. Reconstruction 3D active*, université de Paris-Sud (20/12/95).
- Stéphane Laveau, *Géométrie d'un système de N caméras. Théorie. Estimation. Applications*, École Polytechnique (31/05/96),
- Cyril Zeller, *Calibration projective, affine et euclidienne en vision par ordinateur. Application à la reconstruction tridimensionnelle*, École Polytechnique (2/02/96).

6.2 Participation à des conférences et colloques

Les membres du projet ont participé à ECCV'96, ICAOS'96, ACCV'95, RFIA'96, ICRA'96 et ICPR'96.

6.3 Diffusion de produits

- Le programme `image-matching` développé par Z. Zhang, C. Zeller, R. Deriche, O. Faugeras, Q.-T. Luong a été déposé à l'**Agence pour la Protection des Programmes (APP)**, et l'exécutable est disponible par ftp `krakatoa.inria.fr`. Ce programme a pour but de mettre en correspondance de manière robuste deux images non calibrées et d'estimer leur géométrie épipolaire. 305 copies en ont été faites entre le 23 juillet 1995 et le 3 septembre 1996 à l'intention, entre autres, de Yale, CMU, MIT AI-lab, MIT Media-lab, Microsoft, Alcatel-Alsthom, Cornell, Cemagref, USC, JPL NASA, Tech. U Graz, K U Leuven, U. Tokyo, Siemens, Applet Realise (LUC), ...
- Le logiciel `sfm` développé par Z. Zhang a été déposé à l'APP (Agence pour la Protection des Programmes) en 1995 sous le numéro `IDDN.FR.001.120021.00.R.X.1995`. Il a été mis sur ftp (`krakatoa.inria.fr`) en juin 1996. Il permet d'estimer de manière robuste le mouvement et la structure 3D à partir de deux images perspectives avec paramètres internes connus.

- Le logiciel `FMatrix` développé par Z. Zhang a été mis sur ftp (krakatoa.inria.fr) en juillet 1996. Il a été développé pour évaluer différentes techniques proposées dans la littérature pour estimer la géométrie épipolaire (la matrice fondamentale) entre deux ensembles d'appariements de points 2D non calibrés éventuellement aberrants.
- Le package d'extension `Mascotte` de Maple5.3, développé au sein du projet RobotVis, est diffusé sous licence Gnu sur le Web en :
<http://www.inria.fr/robotvis/personnel/vthierry/mascotte/main.html>.
Développé pour mettre en œuvre des outils de génération de code en vision et robotique, il compte aujourd'hui une vingtaine d'utilisateurs.

7 Publications

Livres et monographies

- [630] M. BERGER, R. DERICHE, I. HERLIN, J. JAFFRÉ, J. MOREL (éd.), *ICAOS'96 : 12th International Conference On Analysis and Optimization of Systems - Images, Wavelets and PDE's, Lecture Notes in Control and Information Sciences*, 219, Springer - Verlag, juin 1996.
- [631] O. D. FAUGERAS, *Recherche de bords dans des images et évolution de courbes*, Ecole Polytechnique, 1996.
- [632] S. MA, Z. ZHANG, *Computer Vision: Theory and Applications*, Chinese Academy of Sciences, 1996.
- [633] G. XU, Z. ZHANG, *Epipolar Geometry in Stereo, Motion and Object Recognition: A unified approach*, Kluwer Academic Publishers, 1996.

Thèses

- [634] G. CSURKA, *Modélisation projective des objets tridimensionnels en vision par ordinateur*, thèse de doctorat, université de Nice Sophia Antipolis, avril 1996.
- [635] R. ENCISO, *Auto-Calibration des Capteurs Visuels Actifs. Reconstruction 3D Active.*, thèse de doctorat, université Paris XI Orsay, décembre 1995.
- [636] S. LAVEAU, *Géométrie d'un système de N caméras. Théorie. Estimation. Applications.*, thèse de doctorat, École Polytechnique, mai 1996.
- [637] C. ZELLER, *Calibration Projective Affine et Euclidienne en Vision par Ordinateur*, thèse de doctorat, École Polytechnique, février 1996.

Articles et chapitres de livre

- [638] R. DERICHE, O. FAUGERAS, «Les EDP en Traitement des Images et Vision par Ordinateur», *Traitement du Signal*, 1996, Accepté pour publication.
- [639] R. ENCISO, T. VIÉVILLE, «Self-calibration from four views with possibly varying intrinsic parameters», *Image and Vision Computing Journal*, 1996, In Press.
- [640] O. FAUGERAS, L. ROBERT, «What can two images tell us about a third one?», *The International Journal of Computer Vision* 18, 1, avril 1996, p. 5–20.
- [641] Q.-T. LUONG, O. D. FAUGERAS, «The Fundamental Matrix: Theory, Algorithms and Stability Analysis», *The International Journal of Computer Vision* 1, 17, janvier 1996, p. 43–76.
- [642] L. ROBERT, C. ZELLER, O. FAUGERAS, M. HÉBERT, «Applications of Non-Metric Vision to Some Visually-Guided Robotics Tasks», in : *Visual Navigation: From Biological Systems to Unmanned Ground Vehicles*, Y. Aloimonos (éd.), Lawrence Erlbaum Associates, 1996, to appear, also INRIA Technical Report 2584.

- [643] T. VIÉVILLE, O. D. FAUGERAS, Q.-T. LUONG, «Motion of Points and Lines in the Uncalibrated Case», *The International Journal of Computer Vision* 17, 1, janvier 1996, p. 7–42.
- [644] T. VIÉVILLE, O. FAUGERAS, «The First Order Expansion of Motion Equations in the Uncalibrated Case», *CVGIP: Image Understanding* 64, 1, juillet 1996, p. 128–146.
- [645] Z. ZHANG, Q.-T. LUONG, O. FAUGERAS, «Motion of an uncalibrated stereo rig: self-calibration and metric reconstruction», *IEEE Transactions on Robotics and Automation* 12, 1, février 1996, p. 103–113.
- [646] Z. ZHANG, «Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting», *Image and Vision Computing Journal*, 1996, In Press.

Communications à des congrès, colloques, etc.

- [647] R. DERICHE, O. D. FAUGERAS, «A Level-set Based Approach for Stereo», in: *Investigative Image Processing, Proc. SPIE*, novembre 1996.
- [648] F. DEVERNAY, O. FAUGERAS, «From Projective to Euclidean Reconstruction», in: *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, B. Bhanu (éd.), IEEE, p. 264–269, San Francisco, CA, juin 1996.
- [649] R. ENCISO, T. VIÉVILLE, O. FAUGERAS, «Simplifier la calibration du mécanisme de mise au point», in: *ORASIS'96*, C.-F. université Blaise Pascal (éd.), p. 55–60, mai 1996.
- [650] R. ENCISO, Z. ZHANG, T. VIÉVILLE, «Dense Reconstruction using Fixation and Stereo Cues», in: *World Automation Congress, ISIAE Symposia*, Montpellier, mai 1996.
- [651] R. ENCISO, A. ZISSERMAN, T. VIÉVILLE, «An affine solution to the Euclidean calibration while using a zoom lens», in: *Workshop ALCATECH, 21-27, Danemark*, juillet 1996.
- [652] O. FAUGERAS, R. KERIVEN, «On projective plane evolution», in: *ICAOS'96*, M.-O. Berger, R. Deriche, I. Herlin, J. Jaffré, J.-M. Morel (éd.), *Lecture Notes in Control and Information Sciences*, 219, Springer-Verlag, p. 66–73, 1996.
- [653] P. KORNPROBST, R. DERICHE, «Image Restoration via PDE», in: *Proceedings of the Conference Investigative Image Processing, part of SPIE's First Annual Symposium on Enabling Technologies for Law Enforcement and Security*, SPIE, Boston, MA, novembre 1996.
- [654] S. LAVEAU, O. FAUGERAS, «Oriented Projective Geometry for Computer Vision», in: *Proceedings of the 4th European Conference on Computer Vision*, B. Buxton (éd.), p. 147–156, Cambridge, UK, avril 1996.
- [655] F. LEYMARIE, A. DE LA FORTELLE, J. KOENDERINK, A. KAPPERS, M. STAVRIDIS, B. VAN GINNEKEN, S. MULLER, S. KRAKE, O. FAUGERAS, L. ROBERT, C. GAUCLIN, S. LAVEAU, C. ZELLER, «REALISE: Reconstruction of Reality from Image Sequences», in: *Proceedings of the International Conference on Image Processing*, (éd.), 3, p. 651–654, Vienna, septembre 1996.
- [656] D. LINGRAND, T. VIÉVILLE, «Dynamic Foveal 3D Sensing Using Affine Models», in: *13th ICPR*, 1, p. 810–814, août 1996.
- [657] D. LINGRAND, T. VIÉVILLE, «Perception tridimensionnelle dynamique au niveau de la fovéa. Utilisation de modèles affines.», in: *Journées ORASIS'96*, 1, p. 181–186, mai 1996.
- [658] L. LUCIDO, J. OPDERBECKE, V. RIGAUD, R. DERICHE, Z. ZHANG, «An Integrated Multiscale Approach for Terrain Referenced Underwater Navigation», in: *Proc. IEEE International Conference on Image Processing*, Lausanne, septembre 1996.
- [659] L. LUCIDO, J. OPDERBECKE, V. RIGAUD, R. DERICHE, Z. ZHANG, «A Multiscale Approach for Terrain Referenced Underwater Navigation», in: *Proc. IEEE Int. Symp. on Time-Frequency and Time-Scale Analysis*, Paris, juin 1996.
- [660] L. LUCIDO, J. OPDERBECKE, V. RIGAUD, R. DERICHE, Z. ZHANG, «A terrain referenced underwater positioning using sonar bathymetric profiles and multiscale analysis», in: *Proc. OCEANS 96 MTS-IEE*, Fort Lauderdale, Florida (US), septembre 1996.

- [661] L. LUCIDO, J. OPDERBECKE, V. RIGAUD, R. DERICHE, Z. ZHANG, «Underwater Navigation by Terrain Matching», in : *Proc. World Automation Congress, ISRAM'96, Montpellier*, mai 1996.
- [662] T. PAPADOPOULO, O. FAUGERAS, «Computing Structure and Motion of General 3D Curves from Monocular Sequences of Perspective Images», in : *Proceedings of the 4th European Conference on Computer Vision*, B. Buxton (éd.), p. 696–708, Cambridge, UK, avril 1996.
- [663] L. ROBERT, R. DERICHE, «Dense Depth Map Reconstruction: A Minimization and Regularization Approach which Preserves Discontinuities», in : *Proceedings of the 4th European Conference on Computer Vision*, B. Buxton (éd.), Cambridge, UK, avril 1996.
- [664] C. A. ROTHWELL, J. STERN, «Understanding the Shape Properties of Trihedral Polyhedra», in : *Proceedings of the 4th European Conference on Computer Vision*, B. Buxton (éd.), p. 175–185, Cambridge, UK, avril 1996.
- [665] C. A. ROTHWELL, «Reasoning about Occlusions during Hypothesis Verification», in : *Proceedings of the 4th European Conference on Computer Vision*, B. Buxton (éd.), p. 599–609, Cambridge, UK, avril 1996.
- [666] T. VIÉVILLE, D. LINGRAND, «Using Singular Displacements for Uncalibrated Monocular Visual Systems», in : *4th ECCV*, 2, p. 207–216, avril 1996.
- [667] Z. ZHANG, «On the Epipolar Geometry Between Two Images With Lens Distortion», in : *Proceedings International Conference on Pattern Recognition, I*, p. 407–411, Vienna, Austria, août 1996.
- [668] Z. ZHANG, «A Stereovision System for a Planetary Rover: Calibration, Correlation, Registration, and Fusion», in : *Proceedings IEEE Workshop on Planetary Rover Technology and Systems*, Minneapolis, Minnesota, USA, avril 1996.

Rapports de recherche et publications internes

- [669] C. BOUVIN, *Contours Actifs Géodésiques : Application à la Stéréovision*, Mémoire, université Nice-Sophia Antipolis, France, 1996.
- [670] F. GASPARD, *Perception Visuelle Hiérarchique sans Calibration*, Mémoire, DEA ARAVIS, université de Nice-Sophia Antipolis, juin 1996.
- [671] T. PAPADOPOULO, «Motion Analysis of 3D Rigid Curves from Monocular Sequences», *rapport de recherche n°2779*, INRIA, janvier 1996.
- [672] C. ZELLER, O. FAUGERAS, «Camera Self-Calibration from Video Sequences: the Kruppa Equations Revisited», *rapport de recherche n°RR2793*, INRIA, février 1996.
- [673] Z. ZHANG, V. SCHENK, «Improvement of Online Calibration Methods: Self-maintaining calibration over time», *rapport de recherche n°2H5065/CALVIN/TN/WPB3-T3-001*, ESA Calvin Project, 1996.
- [674] Z. ZHANG, «Determining the Epipolar Geometry and its Uncertainty: A Review», *rapport de recherche n°RR2927*, INRIA, juillet 1996.
- [675] Z. ZHANG, «A New Multistage Approach to Motion and Structure Estimation: From Essential Parameters to Euclidean Motion Via Fundamental Matrix», *rapport de recherche n°RR2910*, INRIA, juin 1996.

8 Abstract

The scientific and technological objectives of the project fall in two categories:

- to develop theories of machine visual perception and in particular of three-dimensional perception,
- to demonstrate the validity of these theories by applying them to practical problems, typically in robotics, and implementing algorithms for solving them.

Formalisation and in particular mathematical formalisation is a central motivation for us and is always checked by experiments which we believe to be the only way to validate a theory. Our theoretical work has been concentrated in two main directions. The first one is the investigation of the use of algebraic and geometric methods for studying systems with many cameras (see the section on geometry and the reports from previous years). We will continue to invest in that direction in the years to come in particular through collaborations with the Safir group (B. Mourrain) and through the Cumuli ESPRIT LTR project. The second direction is the investigation of the use of analysis methods, mainly from the theory of partial differential equations, to study the problems of image processing (enhancement, restoration) and computer vision (edges, stereo, motion), see section on EDP. This is also a lasting investment which we would like to reinforce in particular through collaborations with the Nice-Sophia Antipolis University (Gilles Aubert, Michel Merle), the Paris IX Dauphine University (Jean-Michel Morel), and the Cermics (Renaud Keriven).

We also have developed in the past a fairly large number of industrial collaborations in the framework of european projects. We have transferred to industrials a lot of our know-how and algorithms (see activity reports for the years 90 to 93). We continue to seek these collaborations but try to broaden them to a wider class of applications than Robotics, in particular to the area of 3-D modelling from image sequences and to the area of the processing of video sequences for such applications as the virtual studio.

Within this framework we have particularly pursued the following directions:

- geometry of systems of cameras, calibration, self-calibration;
- stereo and applications;
- three-dimensional shape and motion perception by analysis of monocular sequences of images;
- partial differential equations, image processing and computer vision.