



HAL
open science

Détection et Positionnement 3D pour le Guidage par Vision : Contributions et applications à la robotique médicale

Doignon Christophe

► **To cite this version:**

Doignon Christophe. Détection et Positionnement 3D pour le Guidage par Vision : Contributions et applications à la robotique médicale. Traitement du signal et de l'image [eess.SP]. Université Louis Pasteur - Strasbourg I, 2007. tel-00644737

HAL Id: tel-00644737

<https://theses.hal.science/tel-00644737>

Submitted on 25 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Détection et Positionnement 3D pour le Guidage par Vision

Contributions et applications à la robotique médicale

Mémoire

présenté publiquement le 9 Octobre 2007 pour l'obtention de

l'Habilitation à Diriger des Recherches
de l'Université Louis Pasteur

Discipline : Electronique, Génie électrique - Spécialités : Robotique et Vision

par

Christophe DOIGNON

Composition du jury :

Rapporteurs :

François CHAUMETTE, Directeur de Recherche, IRISA/INRIA de Rennes
Michel DHOME, Directeur de Recherche CNRS, LASMEA, Clermont-Ferrand
Christophe COLLET, Professeur à l'Université Louis Pasteur (ENSPS), Strasbourg

Examineurs :

El Mustapha MOUADDIB, Professeur à l'Université de Picardie-Jules Verne, Amiens
Nassir NAVAB, Professeur à l'Université Technique de Munich, Allemagne
Michel de MATHELIN, Professeur à l'Université Louis Pasteur (ENSPS), Strasbourg

REMERCIEMENTS

En premier lieu, je tiens à remercier l'ensemble des membres du jury pour m'avoir fait l'honneur d'accepter d'évaluer ces travaux.

Je remercie Michel de Mathelin, pour ses conseils, ses encouragements, son optimisme et sa générosité. Il a été un acteur majeur dans la nouvelle orientation de mes activités de recherche.

Mes remerciements s'adressent également à Naoufel, Alexandre, Florent, Benjamin et Chadi, que j'ai eu le plaisir de voir évoluer durant leur thèse de doctorat.

Je remercie de tout coeur l'ensemble des personnes qui, par leurs compétences ou simplement par leur bonne humeur m'ont aidé à évoluer sereinement dans le milieu universitaire : Pierre, Jacques, Dominique, Eric, Philippe, Bernard, Iuliana, Edouard, Mohamed, Joana, Laurent, Loïc... et aussi ceux qui sont maintenant sous d'autres horizons : Gabriel, Guillaume, Kanako, Frédéric, Xavier,...

Merci au Pr. Marescaux, à Luc Soler et à tous les collègues de l'IRCAD¹ pour leur enthousiasme et leur soutien sans réserve dans nos recherches.

Enfin, je remercie mon épouse, Patricia, mes enfants, Guillaume et Juliette, ainsi que mes parents, pour leur amour et leur soutien inconditionnels. Je leur dédie ce mémoire.

¹ Institut de Recherche contre les Cancers de l'Appareil Digestif, Strasbourg

TABLE DES MATIÈRES

Remerciements	5
Table des matières	8
Avant-propos	9
partie I <i>Dossier individuel</i>	11
I. Curriculum Vitae	15
II. Activités administratives et responsabilités collectives	19
III. Activités d'enseignement (résumé)	23
IV. Encadrements et jurys de thèse de doctorat	27
V. Transferts et valorisations technologiques	35
VI. Publications Scientifiques	41
partie II <i>Synthèse des activités de recherche</i>	51
1. Introduction générale	55
1.1 GÉNÉRALITÉS	56
1.1.1 Contexte	56
1.1.2 Thèmes abordés	57
1.1.3 Thèmes non abordés	59
1.1.4 Motivations	59
1.2 PRÉSENTATION DU LSIIT ET DE IRMC	61
1.3 TRAVAUX ANTÉRIEURS	62
1.3.1 La période post-doctorale (1994-1996)	62
1.3.2 De fin 1996 à 1999	63
1.4 STRUCTURE DU DOCUMENT	64
2. Problématiques communes au guidage par vision	65
2.1 LE RECALAGE 2-D/3-D	66
2.1.1 La modélisation	66
2.1.2 Formulation du recalage rigide	68
2.1.3 Recalage basé sur les descripteurs locaux	70
2.1.4 Recalage basé sur les descripteurs globaux	74

2.2	LES TECHNIQUES D'ASSERVISSEMENTS VISUELS	78
2.2.1	Les différentes approches	78
2.2.2	Formalisme de la fonction de tâche	81
2.2.3	De l'intelligence artificielle à la vision robotique dynamique	81
2.3	LE SUIVI VISUEL	85
2.3.1	Suivi de primitives géométriques 3-D	86
2.3.2	Suivi de caractéristiques apparentes	86
2.4	LA SEGMENTATION D'IMAGES EN TEMPS RÉEL	92
3.	Synthèse des travaux	95
3.1	INTRODUCTION	96
3.2	SEGMENTATION D'INDICES VISUELS GÉOMÉTRIQUES NON LINÉAIRES	98
3.2.1	Filtrage	98
3.2.2	Détection des discontinuités et classification	100
3.3	SEGMENTATION DANS DES ENVIRONNEMENTS NON CONTRÔLÉS	103
3.3.1	Segmentation en temps réel des instruments de laparoscopie	103
3.4	POSITIONNEMENT DES INSTRUMENTS EN LAPAROSCOPIE ET APPLI- CATIONS	109
3.4.1	Cas des ellipses	110
3.4.2	Cas des sphères	110
3.4.3	Cas des cylindres	115
3.4.4	Contrainte de mouvement en chirurgie à invasion minimale	118
3.5	GUIDAGE VISUEL D'UN INSTRUMENT DE LAPAROSCOPIE	121
3.5.1	Guidage et navigation avec un instrument marqué	121
3.5.2	Assistance à la suture	130
3.5.3	Guidage d'un laparoscope articulé	136
3.6	RECALAGE STÉRÉOTAXIQUE EN RADIOLOGIE INTERVENTIONNELLE	145
3.6.1	Guidage par imagerie TDM	145
3.6.2	La stéréotaxie	147
3.6.3	Méthodologie proposée pour le recalage stéréotaxique	149
3.7	RECONSTRUCTION 3-D PAR CODAGE DE LA LUMIÈRE STRUCTURÉE	158
3.8	BILAN	161
4.	Perspectives de recherche	163
4.1	PERSPECTIVES	164
4.2	CONCLUSION GÉNÉRALE	167
	Bibliographie	168
	Publications choisies	193

AVANT-PROPOS

Ce mémoire retrace mes activités de recherche et d'enseignement depuis ma thèse de Doctorat soutenue en janvier 1994. J'y décris plus spécifiquement mes travaux de recherche menés entre 1996 et 2007 en tant que Maître de Conférences à l'Université Louis Pasteur de Strasbourg (ULP). La majeure partie a été effectuée au Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT), dans le Groupe de Recherche en Automatique et Informatique Industrielle (GRAII, puis GRAVIR, dirigés par Eric Ostertag) jusqu'en 2002, puis dans l'équipe Automatique, Vision et Robotique (AVR - dirigée par Michel de Mathelin). Parallèlement à cela, depuis l'an 2000, j'ai été impliqué dans des activités de valorisation de la recherche pour l'industrie au sein de l'Équipe de Recherche Technologique (ERT) en enroulement et transport à haute vitesse de bande flexible, dirigée par Dominique Knittel.

Il ne m'est pas possible dans ce manuscrit d'évoquer l'ensemble des travaux durant un peu plus d'une décennie. Souhaitant offrir un document synthétique et cohérent au lecteur, j'ai volontairement préféré mettre en avant certaines activités par rapport à d'autres qui ont été diffusées par ailleurs, d'autres encore ont été interrompues faute de perspectives attrayantes. Que ceux qui y étaient impliqués, et qui donc n'en voient pas mention ici, veuillent bien me le pardonner.

Treize ans depuis la soutenance de thèse de Doctorat, cela peut sembler long pour préparer l'Habilitation à Diriger les Recherches. J'ai remarqué qu'il est usuel de faire preuve d'un peu plus de diligence. Pour décrire mes activités de toutes ces années, j'ai souhaité considérer ici deux époques : la première s'étendant jusqu'à 1999 avec une forte implication dans les activités liées à l'enseignement (mis à part la période post-doctorale), la seconde depuis 2000 en participant conjointement à deux projets de recherche d'envergure nationale, nés à quelques mois d'intervalle.

Aujourd'hui, et après que mes collègues mentionnés ci-dessus m'aient permis, par leur confiance, de développer un travail personnel, d'encadrer ou de co-encadrer des stagiaires et des doctorants et par voie de conséquence m'aient donné la possibilité de valoriser leur travail et le mien, je me sens prêt à assumer la direction de recherche. Cependant, je préfère le terme de co-direction, car comme vous le découvrirez au cours de cette lecture, les projets auxquels j'ai contribué sont à caractères multi-disciplinaires. D'ailleurs, les travaux d'ampleur collective fournissent des résultats qui dépassent bien souvent le cadre des objectifs initiaux d'un projet de recherche collaborative. C'est pour les mêmes raisons que j'emploierai le "nous" pour exhiber cette pluralité.

Première partie

DOSSIER INDIVIDUEL

I. CURRICULUM VITAE

1. Etat civil

Prénom : Christophe
Nom patronymique : DOIGNON
Date et lieu de naissance : Né le 14 janvier 1965 (42 ans), au Mans (Sarthe)
Situation de famille : Marié, 2 enfants
Adresse personnelle : 8, rue de la forêt, 67140 Barr
Téléphone (mobile) : 06.07.73.75.29

Fonctions actuelles : Maître de Conférences à l'Institut Professionnel des Sciences et Technologies (IPST) de l'Université Louis Pasteur de Strasbourg, 6^{ème} échelon de la classe normale.

2. Adresses et coordonnées professionnelles

Recherche

Équipe Automatique, Vision et Robotique (AVR)
Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT),
Unité Mixte de Recherche ULP-CNRS 7005,
Bureau C 139 - Pôle API
Bd Sébastien Brant, BP 10413 - 67412 Illkirch Cedex

Tel : 03.90.24.44.71 (pôle API) et 03.88.11.91.11 (IRCAD)
Fax : 03.90.24.44.80
Courriel : doignon@lsiit.u-strasbg.fr

Enseignement

Bureau 13 - Institut Professionnel des Sciences et Technologies
15, rue du Maréchal Lefèbvre 67100 Strasbourg Cedex

Tél : 03.90.24.49.48 Fax : 03.90.24.49.72

3. Diplômes universitaires

1994 : Doctorat en Sciences de l'Université Louis Pasteur
Spécialité : Robotique et traitement d'images
Mention : très honorable

Thèse effectuée au LSIT, sous la direction d'Eric Ostertag (Prof. ULP/ENSPS)
et de Gabriel Abba (Prof. ENI Metz) et soutenue le 20 janvier 1994.

Titre : *Etude et réalisation d'un système de vision monoculaire pour la reconnaissance et la localisation d'objets en robotique*

1989 : DEA Automatique et Productique de l'Université de Haute Alsace

1989 : Diplôme d'ingénieur de l'ENSPS

1987 : Licence de Physique de l'ULP

1986 : Deug A de l'Université des Sciences d'Angers, major
Admission sur concours à l'ENSPS

1983 : Baccalauréat E au lycée polyvalent des Lombards, à Troyes

4. Cursus professionnel

Depuis 1996 : Maître de Conférences à l'ULP (IPST) et chercheur au LSIT
Section CNU 61 (génie informatique, automatique, traitement du signal)

1995-1996 : Post-doctorant au Département Electronique et Informatique
de l'université de Padoue (Italie) : programme européen HCM-MAVI

1993-1995 : Attaché temporaire d'enseignement et de recherche (ATER) à l'ENSPS

1990-1993 : Allocataire de recherche (bourse MRES) et moniteur de l'enseignement
supérieur à l'UFR de Physique de l'ULP.

1989-1990 : Service national scientifique effectué au centre informatique de
l'Ecole Supérieure d'Application du Matériel (ESAM) de Bourges.

5. Domaines de recherche

- Recalage et positionnement 3D à partir de nombreux types de dispositifs d'imagerie
- Vision active, endoscopie active
- Suivi visuel et asservissements visuels en robotique
- Traitement des images (pré-traitement et segmentation) en temps réel, filtrage.

6. Publications scientifiques

- 5 articles parus ou acceptés dans des revues internationales et 3 en cours d'évaluation,
- 3 contributions à un ouvrage collectif,
- 40 conférences internationales dont 3 en tant qu'invité,
- 7 communications lors de conférences nationales avec actes et comité de lecture.

II. ACTIVITÉS ADMINISTRATIVES ET RESPONSABILITÉS COLLECTIVES

Après avoir été durant 5 ans (1997-2001) responsable pédagogique de la Licence Ingénierie Electrique, mention réseaux (formation continue, par alternance) à l'IPST, il m'a été proposé en 2001 la responsabilité de la Licence Réseaux Informatiques et Applications (également une formation continue, par alternance). Cette charge comprenait notamment le suivi pédagogique, l'évaluation des dossiers lors du recrutement, la coordination des enseignants et des enseignements (plusieurs UFR de l'ULP, l'IUT de Colmar, le Lycée Couffignal et des intervenants provenant de diverses entreprises en Alsace), l'organisation des emplois du temps et des jurys. Il m'a été demandé également de participer à l'élaboration du dossier d'habilitation de cette formation. J'ai quitté la responsabilité de cette formation au moment de la mise en place de la réforme LMD en 2005 et correspondant également à l'obtention d'une délégation au CNRS.

J'ai par ailleurs été responsable de la commission pédagogique de l'IPST de 1998 à 2000. J'ai été sollicité à de nombreuses reprises par la commission de validation des acquis professionnels/d'expériences de l'ULP pour expertiser des dossiers de demande d'équivalence et de dispenses d'enseignement (1998-2004).

J'ai accepté la prise en charge d'autres responsabilités et la participation à d'autres activités collectives, telles que :

1. des fonctions électives d'intérêt général :

- Membre de la commission de spécialistes mixte 61-63 de l'ULP depuis 1997,
- Membre nommé suppléant de la commission (multidisciplinaire) de spécialistes de l'ENSAIS de 1997 à 1999, puis titulaire en 2000,
- Membre nommé titulaire de la commission de spécialistes (section 61) de l'INPL de Nancy,
- Vice-président de la commission de spécialistes mixte 61-63 de l'ULP de 2001 à 2003.
- Membre du conseil du LSIIT (collège B) depuis 2001.

2. des activités collectives liées à l'enseignement :

- Participation à de nombreux jurys de fin d'année, de semestre, de soutenances et de diplômes (entre autre l'IUP GSI, le DESS TSI Systèmes Automatisés puis le Master IT-IISA à l'IPST, le diplôme d'ingénieur ENSPS, le Master ISTI et la FIP à l'ENSPS).

- Prise en charge et gestion de 22 machines (partie du parc informatique de l'IPST) de 1996 à 2001 (dossiers de demandes d'investissement pédagogique, suivi des achats, installation des machines, installation de logiciels, installation de comptes pour les enseignants et les étudiants).
- Conception, rédaction et mise en oeuvre de TP en informatique industrielle (12 sujets et maquettes de TP), en vision industrielle (10 sujets de TP) et en automatique.
- Participation et membre du groupe de travail sur l'innovation animé par l'ITII (Institut des Techniques de l'Ingénieur de l'Industrie). L'ITII Alsace est partenaire de la formation des ingénieurs en partenariat (FIP) de l'ENSPS.
- Responsable pédagogique de la FIP à partir de la rentrée universitaire 2007.

3. des activités liées à la recherche :

- Participation à l'élaboration du contrat de plan état-région *Enroulement et transport à haute vitesse de bande flexible* pour le compte de l'Équipe de Recherche Technologique (ERT n°8, CPER 2003-2007).
- Participation à l'élaboration du dossier d'évaluation de l'ERT n°8. Présentation lors de l'évaluation nationale (avril 2005).
- Participation à la rédaction du rapport d'activités du LSIIT (axes vision et robotique de l'équipe GRAViR) en 2001 (mi-parcours), en 2003 et en 2007 pour l'équipe AVR.
- L'équipe AVR du LSIIT a installé sa plate-forme robotique médicale à l'IRCAD depuis le début de l'année 2007. Cette mission m'a été confiée par Michel de MATHELIN, responsable de l'équipe AVR. Durant 15 mois environ, j'ai coordonné un nombre important de tâches entre mes collègues et des intervenants extérieurs divers (bureau d'étude, architecte, société de déménagement, les services techniques et informatiques de l'IRCAD, les entreprises d'électricité et de mobiliers retenues après les appels d'offres,...). Cette installation a demandé un grand nombre de travaux d'aménagement pour permettre l'installation de 7 cellules robotiques et de 24 bureaux au deuxième étage du bâtiment de l'IRCAD (hôpital civil de Strasbourg). L'inauguration de ce nouveau site a eu lieu le 29 Mars 2007.
- Responsable de l'organisation du congrès francophone des jeunes chercheurs en vision par ordinateur (Orasis), du 4 au 8 juin 2007.
- Co-porteur du projet intitulé "Développement d'un dispositif endoscopique miniaturisé pour la vision robotique par lumière structurée en chirurgie mini-invasive". Obtention

d'un Bonus Qualité Recherche (BQR) en 2006 qui correspond à un appel à projets du conseil scientifique de l'ULP. Le LSP (Laboratoire des Systèmes Photoniques de Strasbourg) et l'IRCAD sont partenaires de ce projet.

- Responsable scientifique et porteur du projet intitulé "Endoscopie active pour l'assistance médicale robotisée". Rédaction du dossier de demande de financement ANR dans le cadre de l'appel à projets "Systèmes Interactifs et Robotique" 2007 (actuellement en soumission). L'INRIA de Rennes, le CRAN (Centre de Recherche en Automatique de Nancy) et l'IRCAD sont partenaires de ce projet.
- Relecteur de plusieurs articles (environ 30) de conférences internationales dont MICCAI 2006, ICRA 2006 & 2007, IROS 2006 & 2007, ACC 2006 et de revues internationales IJRR (numéro spécial), CVIU et MVA, entre autres.

III. ACTIVITÉS D'ENSEIGNEMENT (RÉSUMÉ)

J'ai effectué tous mes enseignements à l'Université Louis Pasteur.

Mes activités d'enseignement sont réparties selon trois périodes. D'une part, en tant que moniteur lors de ma préparation de thèse de Doctorat (1990-1993), puis en tant qu'ATER (1993-1995) et finalement en tant que Maître de Conférences à l'ULP, rattaché à l'Institut Professionnel des Sciences et Technologies (1996-).

Durant mon monitorat de l'enseignement supérieur à l'UFR de Physique (1990-1993), j'ai effectué des TD de Physique (thermodynamique statistique et relativité restreinte), des TD d'électronique analogique en DEUG A (2ème année), ainsi que des cours préparatifs au concours ENSI DEUG (matières préparées : optique, électromagnétisme et thermodynamique).

J'ai obtenu de septembre 1993 à août 1995 un poste d'attaché temporaire d'enseignement et de recherche (ATER) à l'ENSPS. J'y ai enseigné principalement l'automatique en deuxième année de la formation d'ingénieur (84 h de TD et 356 h de TP), et j'ai également assuré des encadrements de projets en traitement d'images (ENSPS 2A - 40 h) et en robotique (ENSPS 3A - 38 h). Durant cette période où l'ENSPS s'est installée au pôle API à Illkirch (février 1994), j'ai participé à l'installation des salles de TP d'automatique et des manip avec Yves Dossmann et Gabriel Abba. J'ai également défini et rédigé une série de nouveaux TP d'automatique avec eux.

Depuis ma nomination à l'Université Louis Pasteur en septembre 1996, j'ai dispensé des enseignements à l'Institut Professionnel des Sciences et Technologies (IPST) de l'ULP (59 %), à l'École Nationale Supérieure de Physique de Strasbourg (ENSPS) (28 %) et, à un degré moindre, au Département d'Informatique (DI) de l'UFR de Mathématique et d'Informatique (13 %). A part une parenthèse de 18 mois (de septembre 2005 à février 2007) où j'étais en délégation CNRS, les enseignements ont été dispensés dans ces trois composantes aussi bien dans des filières en formation initiale qu'en formation dite continue (en alternances). Une particularité de mes enseignements est qu'une part significative concerne un engagement dans les formations par alternances (39 % de l'ensemble des enseignements).

Plus précisément, depuis 1996, je suis intervenu dans les filières suivantes :

- **La Licence Ingénierie Electrique (IPST)**, mention Réseaux Informatiques. C'est une formation à temps partagé sur 18 mois. J'y ai enseigné essentiellement l'algorithmique et la programmation en C++ (60 h/an) et le système d'exploitation UNIX (60 h/an) de 1996 à 2001.
- **La Licence Réseaux et Applications (DI)**. La Licence Ingénierie Electrique n'étant plus un diplôme délivrée par l'IPST à partir de 2001, il a été proposé par le Département d'Informatique de l'ULP la Licence Réseaux et Applications (également en formation

- à temps partagé). De 2001 à 2005, j'y ai enseigné l'algorithmique et la programmation (50 h/an).
- La **Maîtrise Réseaux et Applications (DI)**. J'y ai enseigné la programmation système (32 h/an), pendant 3 ans.
 - L'**IUP GSI** (Génie des Systèmes Industriels, IPST) en 2ème année (Licence - TD et TP de traitement du signal, d'automatique et TP de Mathématiques !) de 1996 à 1998 et surtout en 3ème année (Maîtrise - Cours, TD et TP d'informatique industrielle - 62 h/an), de 1996 à 2004.
 - Le **DEUST PIE** (Procédés Industriels Electro-mécaniques, IPST) en 2ème année (Cours et TP de programmation en C - 40 h/an) de 1997 à 1999.
 - Le **DESS TSI-SA** (Technologies et Stratégies Industrielles, option Systèmes Automatisés, IPST). J'y ai enseigné la programmation avancée et l'informatique temps réel (Cours et TP - 50 h/an) de 2001 à 2004.
 - La **formation d'ingénieur ENSPS**, en 2ème année (TD d'automatique 44 h/an, TP UNIX ponctuellement) de 1996 à 2000, et en 3ème année (projets et cours de vision par ordinateur - 25 h/an) de 1999 à 2007.
 - Le **DEA PI** (Photonique, Image), puis **PIC** (Photonique, Image et Cybernétique - ENSPS). De 1999 à 2004, j'ai enseigné la première partie (15 h/an) du cours de vision et commande.
 - Le **Master ISTI** (Master Recherche Images, Sciences et Technologies de l'Information, ENSPS), parcours Automatique et Robotique. J'y enseigne la première partie (15 h/an) du cours de vision et commande depuis 2005.
 - La **Licence de Physique et Applications (IPST/UFR Physique)**, parcours Ingénierie. J'enseigne aux étudiants de 3ème année le cours intitulé "signal, image et vision" (cours et TP - 40 h/an), depuis 2005.
 - La **FIP** (Formation d'Ingénieur en Partenariat) de l'ENSPS. Depuis 2006, j'y enseigne le traitement du signal déterministe (75 h/an), en 2ème année.
 - Le **Master IT** (Master Professionnel - Ingénierie et Technologie, IPST), parcours IISA (Informatique Industrielle et Systèmes Automatisés). J'y enseigne depuis 2005 la vision industrielle (M1 - Cours et TP, 33 h/an) et je partage la responsabilité du cours d'Informatique Industrielle (M2).

En résumé, les trois grands domaines d'intervention de mes enseignements sont les suivants :

- l'informatique industrielle et la programmation informatique (46 %),
- la vision par ordinateur et la vision industrielle (25 %),
- le traitement du signal et l'automatique (20 %).

Ils représentent, cumulés, 91 % de l'ensemble de mes enseignements (3 300 h eq. TD) depuis décembre 1990.

IV. ENCADREMENTS ET JURYS DE THÈSE DE DOCTORAT

1. Encadrements de stages avant 1996

Durant ma thèse de Doctorat jusqu'à la fin de mon contrat d'ATER (août 95), j'ai été amené à co-encadrer des stagiaires de diverses filières et d'horizons très différents. La plupart de ces encadrements concerne également mon co-directeur de thèse Gabriel Abba. En voici la liste :

- juin 91 à septembre 92 : **Mustapha Hafdi**. Stage d'ingénieur CNAM en informatique (et co-encadrement avec G. Abba) sur la *calibration pince-robot et la détection en temps réel dans les images*.
- janvier à juillet 92 : **Frédéric Imbert**. Stage de DEA de l'université de Mulhouse (UHA) (et co-encadrement avec G. Abba) portant sur l'*asservissement 2D d'un robot par caméra*.
- janvier à juillet 93 : **Jean-Luc Moron**. Stage de DEA de l'université de Mulhouse (UHA) portant sur l'*Etude des déformations dans l'image d'un objet en mouvement, estimation de la vitesse apparente*.
- mai 93 à juillet 94 : **René Roeckel**. Stage d'ingénieur CNAM en électronique (et co-encadrement avec G. Abba) sur le *Transfert rapide des images dans un réseaux de Transputers*, en partenariat avec l'ANVAR Alsace.
- de septembre 93 à mars 94 : **Ludwig Albrecht**. Stage de Diplomarbeit (et co-encadrement avec G. Abba). Les résultats de son travail ont fait l'objet d'une communication intitulée *Multi-controller design for robot control with visual feedback and switching strategy* lors de l'European Control Conference, ECC'95 [ADAO95].
- juillet à septembre 94 : **C. Furhmann**. Stage de studienarbeit (et co-encadrement avec G. Abba) intitulé *Methodologies for geometric robot calibration by the use of a monocular vision system*.
- février à août 95 : **H. Jiang**. Stage de DEA PI portant sur la *Reconnaissance d'objets - recherche de contraintes géométriques pour la réduction de l'espace des correspondances*.

2. Encadrements durant la période post-doctorale

Lors de mon séjour post-doctoral à l'université de Padoue en Italie, j'ai eu l'opportunité de co-encadrer 3 étudiants préparant le "Tesi di laurea" (équivalent à un projet de fin d'études) d'octobre 1995 à mars 1996. Ces co-encadrements ont été effectués sous la tutelle de Guido Cortelazzo (pour les deux premiers stagiaires) et Ruggiero Frezza (concernant C. Saviane), tous les deux professeurs au DEI (Dipartimento di Elettronica et Informatica) de Padoue. Il s'agit de :

- **Michele Michielin.** Tesi di Laurea de l'université de Padoue : *Alignement projectif d'images : application à la mosaïque d'images.*
- **Andrea Boscaini.** Tesi di Laurea de l'université de Padoue : *Filtrage et extraction sub-pixellique d'indices visuels.*
- **Claudio Saviane.** Tesi di Laurea de l'université de Padoue : *Calibration de caméras : évaluation de plusieurs méthodes géométriques.*

3. Encadrements de stages depuis 1996

J'ai encadré 9 stagiaires, dont 5 stagiaires de DEA/Master depuis que je suis Maître de Conférences à l'ULP. Il s'agit de stagiaires accueillis dans l'équipe GRAII, puis GRAViR et AVR du LSIIIT et dans l'ERT n°8 de l'ULP. En voici la liste :

- mars à juillet 97 : **Thomas Mohr.** Stage de diplomarbeit (Technische Universität Illmenau, Institut für Allgemeine und Theoretische Elektrotechnik, Prof. U. Seidel) intitulé *Feature Tracking by the means of a dynamic management of areas of interest.*
- février à juin 98 : **Emmanuel Jacquinet.** Stage de DEA PI (Photonique et Image) intitulé *Elaboration d'une commande dans le plan image.*
- mars à fin juillet 2001 : **Alain Ochs.** Stage de DEA PI (et co-encadrement avec Dominique Knittel) intitulé *Estimation par vision et observation robuste appliquées à un système d'entraînement de bande.*
- juin à août 2003 : **Didier Schelcher.** Stage de 2ème année ENSPS portant sur l'*Implémentation d'un algorithme de détection d'instruments de laparoscopie par contours actifs paramétriques.*
- février à juillet 2004 : **Christophe Delimal.** Stage de DEA PIC (Photonique, Image et Cybernétique). Stage intitulé *Détection par vision des vibrations et des défauts d'une bande d'enroulement large à l'aide de la lumière structurée.* Communication au congrès des jeunes chercheurs en vision par ordinateur (Orasis 2005) [DDK05].

- juin à août 2004 : **Hanitriniaina Ranaivoson**. Stage de 2ème année ENSPS intitulé *Asservissement d'un zoom à l'aide d'un objectif motorisé à focale variable*.
- juin à août 2005 et juillet 2006 : **Xavier Maurice**. Stage de fin d'études de l'IUT d'Informatique de l'Université Robert Schuman de Strasbourg (première période) et stage de fin de 1ère année du cycle ingénieur de l'INSA Lyon (seconde période). Stage intitulé *Estimation en temps réel du déport latéral de bande flexible en mouvement par vision stéréoscopique et lumière structurée*. Ce travail réparti sur deux années a fait l'objet de la rédaction d'un article qui a été communiqué lors de la conférence IEEE IMTC 2007 [MDK07].
- mars à juillet 2006 : **Frédéric Bourger**. Stage de Master ISTI, parcours Automatique et Robotique (et co-encadrement avec Ph. Zanne) intitulé *Asservissement visuel 2D sans modèle : application à la robotique médicale*. Ce travail qui comprenait la détection automatique de zones d'occultations a fait l'objet de la rédaction d'un article qui a été communiqué lors de la conférence IEEE ICRA 2007 [BDZM07].
- mars à juillet 2007 : **André Lesaux**. Stage de Master ISTI, parcours Image et Vision. Le travail en cours porte sur la *Détection et caractérisation des mouvements des souris par vision artificielle pour l'analyse comportementale*. Ce travail constitue une première collaboration entre l'équipe AVR du LSIIT et l'Institut des Neurosciences Cellulaires et Intégratives de l'ULP.

4. Co-encadrements de thèses de doctorat

A. Thèse de Naoufel WERGHI (septembre 1993 - décembre 1996, encadrement de 50 %)

J'ai débuté mon premier co-encadrement de thèse en septembre 1993 (co-encadrement partagé avec Gabriel Abba). L'un des axes thématiques de l'équipe concernait la "Commande dynamique de robots par vision" et c'est au sein de ce thème que s'est inscrit le sujet de thèse de Naoufel WERGHI intitulé

*Analyse et interprétation d'images pour la reconnaissance
et la localisation d'objets en vision robotique*

Sujet : Afin de pouvoir faciliter la mise en correspondance de primitives géométriques lors d'un calcul de recalage rigide (ou pose), nous nous sommes tournés vers l'extraction et la modélisation d'indices visuels géométriques non-linéaires tels que les coniques [Wer96]. Comme il est connu que de telles structures géométriques contraignent jusqu'à 5 degrés de liberté d'un point de vue mais sont très délicates à segmenter avec précision, plusieurs outils de filtrage et de classification de contours ont été développés, du pré-traitement des images jusqu'à la localisation 3D des objets présentant de tels éléments. Ce travail a donné lieu à la rédaction de 4 articles dans des conférences internationales avec actes [DAWO95, WD95, WDA96b, WDA96a], dont

deux très reconnues dans ce domaine ([WD95] à ACCV'95 pour la partie filtrage et [WDA96a] à IEEE ICIP'96 pour la classification par l'emploi des ondelettes et la localisation). Aujourd'hui, Naoufel WERGHI est Professeur à l'Université d'Abù Dhabi (Emirats Arabes Unis).

Composition du jury :

- Jack-Gérard Postaire (rapporteur externe), Prof. à l'USTL (Lille),
- Jean Gallice (rapporteur externe), Prof. à l'université Blaise Pascal de Clermont-Ferrand (LASMEA),
- Marcel Nougaret (examineur), Prof. à l'IUT de Grenoble,
- Eric Ostertag (directeur de thèse), Prof. à l'ULP/ENSPS (LSIIT),
- Gabriel Abba (co-directeur de thèse), Maître de Conférences à l'ULP/ENSPS (LSIIT).

B. Thèse d'Alexandre KRUPA (septembre 1999 - juillet 2003, encadrement de 33 %)

Lorsqu'Alexandre KRUPA a commencé sa thèse de doctorat, nous débutions également notre première collaboration avec l'IRCAD en 1999. C'est ainsi que répondant à un appel à projets ACI "jeunes chercheurs" en 2000, Guillaume Morel, alors Maître de Conférences à l'ENSPS et porteur de ce projet, proposa à Jacques Gangloff et à moi-même d'y participer. Guillaume Morel fut le co-encadrant principal de ce travail de thèse la première année, sous la co-direction de Michel de Mathelin (co-tutelle avec l'INPL de Nancy). Le départ de Guillaume Morel en septembre 2000 nécessita une implication plus forte de ma part dans le projet et j'ai assuré le co-encadrement de la thèse d'Alexandre avec Jacques Gangloff et Michel de Mathelin, thèse dont le sujet avait pour titre

Commande par vision d'un robot de chirurgie laparoscopique

Sujet : Il s'agissait d'une recherche collaborative consacrée à l'aide au geste chirurgical en laparoscopie opératoire assistée par un robot. Dans ce travail ont été réalisés l'étude, le développement et la validation expérimentale d'un système de vision permettant de déplacer automatiquement, par retour visuel, un instrument chirurgical marqué en vue d'assister le chirurgien qui télé-opère à l'aide d'un système robotique. Deux tâches ont été proposées : la première consistait à récupérer un instrument chirurgical initialement hors du champ visuel de la caméra pour l'amener au centre de l'image endoscopique, la deuxième tâche contrôlait automatiquement les déplacements de l'instrument pour le guider vers une position 3D préalablement choisie par le chirurgien.

Il s'agissait de ma première participation à un jury de thèse et un nombre important de communications est en rapport avec ce travail. En premier lieu, des communications lors de congrès internationaux parmi lesquels IEEE ICRA 2002 [KJM⁺02], IEEE/RSJ IROS 2002 [KDGM02], MICCAI (2001 et 2002) [KMD⁺01, KMD⁺02], et surtout un article dans la revue IEEE Transactions on Robotics and Automation [KGD⁺03]. Alexandre KRUPA est maintenant chargé

de recherche à l'INRIA de Rennes dans l'équipe-projet Lagadic.

Composition du jury :

- Jocelyne Troccaz (rapporteur externe et présidente), Directrice de Recherche CNRS (TIMC/IMAG), université Joseph Fourier de Grenoble,
- Philippe Martinet (rapporteur externe), Prof. à l'université Blaise Pascal de Clermont-Ferrand (LASMEA),
- Didier Wolf (directeur de thèse), Prof. à l'INPL Nancy (CRAN),
- Michel de Mathelin (co-directeur de thèse), Prof. à l'ULP/ENSPS (LSIIT),
- Jacques Gangloff (examinateur), Maître de Conférences à l'ULP (LSIIT),
- Christophe Doignon (examinateur), Maître de Conférences à l'ULP (LSIIT),
- Guillaume Morel (invité), Maître de conférences à Paris 6 (LRP6)),
- Luc Soler (invité), Docteur en informatique à l'IRCAD et Prof. associé au CHU de Strasbourg,
- Didier Mutter (invité), Prof. chirurgien à l'IRCAD.

C. Thèse de Florent NAGEOTTE (sept. 2001 - nov. 2005, encadrement de 50 %)

Dans la thèse de Florent Nageotte que j'ai co-encadré avec Michel de Mathelin, nous avons souhaité passer du stade de la navigation et du guidage d'instruments par vision dans la cavité abdominale à la réalisation d'une tâche particulière en interaction avec les tissus : la suture. Le mémoire est intitulé

Contribution à la suture assistée par ordinateur en chirurgie mini-invasive

Sujet : La suture est un geste courant mais difficile en chirurgie coelioscopique. Les mouvements possibles des instruments sont limités en raison du passage par un point fixe et le retour visuel fourni par une caméra endoscopique est indirect et bidimensionnel. Il est donc difficile pour les chirurgiens de planifier les mouvements de l'aiguille de suture. En pratique, le passage de l'aiguille dans les tissus est réalisé par essais successifs et fait intervenir des déformations indésirables des tissus.

Nous avons proposé d'étudier la cinématique des mouvements possibles du couple d'instruments (aiguille circulaire, porte-aiguille) en chirurgie coelioscopique. Nous avons recherché tout d'abord dans quels cas il était possible de trouver des chemins sans déformation entre deux points à la surface des tissus. Nous avons montré en formalisant le problème qu'il existait des conditions simples sur la prise d'aiguille et sur le placement du trocart permettant de garantir l'existence de tels chemins. Nous avons alors proposé une méthode pratique de planification permettant de générer des chemins dits à déformation minimale.

La planification de chemin nécessite la connaissance de données spatiales comme la position de l'aiguille dans le porte-aiguille et la position du trocart par rapport aux tissus. Ces informations ont été obtenues en utilisant une caméra endoscopique couleur. Des méthodes classiques

de traitements d'images, dont certaines ont été adaptées pour des applications temps réel, ont permis l'extraction rapide des informations géométriques. La reconstruction des données spatiales est basée sur une optimisation itérative des erreurs de reprojection par asservissement visuel virtuel. Enfin, nous avons proposé des outils de réalité augmentée permettant d'assister le chirurgien durant une suture. Nous avons montré également la faisabilité d'une aide robotisée à la suture en réalisant un passage d'aiguille, basé sur des asservissements visuels référencés image, dans des conditions de laboratoire.

Composition du jury :

- Michel de Mathelin (directeur de thèse), Prof. à l'ULP/ENSPS (LSIIT),
- Nicolas Chaillet (rapporteur externe), Prof. à l'université de Franche-Comté (LAB),
- Philippe Poignet (rapporteur externe), Maître de Conférences Habilité, à l'université de Montpellier (LIRMM),
- Christophe Collet (rapporteur interne), Prof. à l'ULP/ENSPS (LSIIT),
- Christophe Doignon (examineur), Maître de Conférences à l'ULP (LSIIT),
- Nikolaos Papanikolopoulos (examineur), Prof. à l'université du Minnesota,
- Philippe Zanne (invité), Ingénieur de Recherche CNRS au LSIIT.

Parmi les publications en rapport avec ce travail, je retiens ici des communications à des workshops internationaux tels que MRNV 2004 [NMD⁺04], WDV 2006 [DNM06b], les conférences IEEE ICRA 2005 [NZMD05], IEEE/RSJ IROS 2004 [DNM04] et 2006 [NZDM06], MICCAI 2006 [DNM06a] et la contribution à un ouvrage collectif [DNM07]. Un article très récemment soumis à une revue internationale est en cours d'évaluation. Florent Nageotte est actuellement Maître de Conférences à l'ULP, chercheur dans l'équipe AVR en robotique médicale.

D. Thèse de Benjamin MAURIN (sept. 2001 - nov. 2005, encadrement de 25 %)

La thèse de Benjamin Maurin s'est déroulée parallèlement à celle de Florent Nageotte. Je n'en ai donc pas été l'encadrant principal (Olivier Piccin). Cependant, Benjamin ayant montré des aptitudes et un intérêt certain pour les problématiques liées à la vision par ordinateur, nous nous sommes intéressés ensemble au recalage stéréotaxique automatique par imagerie tomodensitométrie à rayons X, une étape nécessaire dans l'accomplissement de son travail. En effet, le travail de thèse porte en premier lieu sur la conception et de la réalisation d'un robot hybride et léger - appelé CT-Bot - pour la radiologie interventionnelle et destiné au positionnement d'aiguille. Ce travail s'insère dans le projet IRASIS (Insertion Robotisée d'Aiguille Sous Imagerie Scanner) du programme national ROBEA du CNRS et qui a consisté en un partenariat entre le LAAS, l'INSA de Strasbourg, le département de radiologie B, l'IRCAD et le LSIIT. Le mémoire est intitulé

*Conception et réalisation d'un robot d'insertion d'aiguille
pour les procédures percutanées sous imageur scanner*

Contribution en vision : Pour placer automatiquement l'aiguille dans une configuration souhaitée sur l'abdomen et afin de pointer des cibles à partir d'une image tomodynamométrique, il est essentiel de localiser précisément le dispositif de positionnement par rapport aux coupes du scanner.

La solution que nous avons retenue pour déterminer la position et l'orientation entre le dispositif de positionnement de l'aiguille et un référentiel lié au scanner était d'employer des marqueurs tridimensionnels. Ces marqueurs, comparables à ceux utilisés en stéréotaxie dans le domaine de la neurochirurgie, ont permis d'effectuer une estimation de la pose complète (à six degrés de liberté) mais également de calibrer le scanner.

En modélisant de manière géométrique le recalage rigide entre un marqueur de type droite et une image, nous avons fourni des solutions algébriques et numériques au problème de pose avec un minimum de quatre indices visuels visibles simultanément. Les conditions, quant aux choix des marqueurs, ont été détaillées grâce à la représentation compacte obtenue [Mau05]. Les algorithmes présentés sont nouveaux en stéréotaxie assistée par ordinateur : les travaux jusqu'alors publiés utilisaient au mieux entre six et neuf indices visuels pour reconstruire la pose. Par ailleurs, nous avons fourni une méthode de calcul des bornes d'erreur sur la solution par linéarisation au premier ordre.

L'ensemble du volet correspondant au recalage stéréotaxique a permis la rédaction d'articles dans des congrès nationaux (Surgetica'05 [MDG⁺05], Orasis'05 [MDMG05] et RFIA'06 [DMBM06]), dans des conférences internationales, dont IEEE CVPR'03 [MDMG03] et MIC-CAI'04 [MGB⁺04]. Un article de revue a été soumis à IEEE Transactions on Medical Imaging et est à l'heure actuelle en révision (seconde révision, mais pas encore accepté). Un autre article rédigé pour la revue internationale Robotics and Autonomous Systems est en cours d'évaluation.

Composition du jury :

- Michel de Mathelin (directeur de thèse), Prof. à l'ULP/ENSPS (LSIIT),
- Olivier Piccin (co-encadrant), Maître de Conférences à l'INSA de Strasbourg (LGECO),
- Jocelyne Troccaz (rapporteur externe), Directrice de Recherche CNRS (TIMC/IMAG), université Joseph Fourier de Grenoble,
- François PIERROT (rapporteur externe), Directeur de Recherche CNRS au LIRMM,
- Yves Rémond (rapporteur interne), Prof. à l'ULP (IMFS),
- Wisama Khalil (examineur), Prof. à l'Ecole Centrale de Nantes (IRCCYN),
- Afshin Gangi (invité), Prof. à l'ULP et radiologue au CHU de Strasbourg.

Pour l'ensemble de sa thèse, Benjamin a obtenu le prix du conseil scientifique de l'ULP.

Benjamin Maurin est maintenant ingénieur roboticien, pour le compte de la société Cerebellum Automation à Chavanod (Annecy).

E. Thèse de Chadi ALBITAR (septembre 2005 - , encadrement de 50 %)

Depuis septembre 2005, je co-encadre avec Pierre Graebing, le travail de thèse de Chadi Albitar. Nous avons sollicité auprès du conseil scientifique de l'ULP une aide spécifique pour financer la conception et la réalisation d'un prototype de dispositif endoscopique à lumière structurée. Cette demande qui est également un partenariat entre le LSIIT, le LSP (Laboratoire des Systèmes Photoniques de Strasbourg) et l'IRCAD a été acceptée. C'est dans le cadre de ce projet 2006-2007 que se situe la thèse de Chadi Albitar.

Sujet : Nous nous intéressons plus particulièrement à la conception de dispositifs capables de fournir des reconstructions 3D des structures internes de l'abdomen en temps réel et avec une grande précision, dans le but d'intégrer ces données dans des systèmes d'assistance médicale. En effet, disposer de l'information tridimensionnelle, et en temps réel, relative à l'environnement dans lequel opère un chirurgien permettrait d'envisager de nouvelles perspectives dans ce domaine. Par exemple, dans le cadre d'interventions mini-invasives, cela permettrait de restituer visuellement la profondeur et ainsi offrir au chirurgien un plus grand confort et une plus grande précision dans son geste. D'autre part, la connaissance en temps réel de la géométrie 3D de la cavité abdominale pourrait aussi permettre de piloter des systèmes robotisés dans le but d'assister le chirurgien et d'aboutir à une plus grande sécurité pour le patient.

Afin de réduire la complexité algorithmique inhérente à la mise en correspondance entre les éléments qui composent le motif projeté et leurs homologues dans l'image, nous avons associé un codage au motif, de manière à pouvoir identifier rapidement ses éléments. Le problème lié au choix du motif utilisé et au codage qui lui est associé n'est pas nouveau, mais nous avons proposé ici une amélioration substantielle par rapport à ce qui existe par ailleurs, amélioration qui concerne le taux d'erreurs (données manquantes ou aberrantes) acceptable et la possibilité de retrouver des éléments manquants. Des premiers résultats ont été obtenus et un motif structurant, composé de 783 éléments (3 primitives différentes), a été réalisé grâce à la technologie de l'optique diffractive et fabriqué par micro-gravure.

Un article a ainsi pu être présenté au congrès des jeunes chercheurs en vision par ordinateur (Orasis'07) [AGD07a]. Deux autres articles portant respectivement sur la segmentation des composants du motif [AGD07b] et sur le décodage en temps réel et la reconstruction 3D [AGD07c] ont été acceptés pour être présenté lors des conférences internationales IEEE ICIP'07 et ICCV'07.

Tableau récapitulatif des co-encadrements de thèses de doctorat.

Doctorants	Taux encadrement	Période	Durée (mois)
Naoufel WERGHI	50 %	sept. 93 - déc. 96	40
Alexandre KRUPA	33 %	sept. 2000 - juil. 2003	34
Florent NAGEOTTE	50 %	sept. 2001 - nov. 2005	50
Benjamin MAURIN	25 %	sept. 2001 - nov. 2005	50
Chadi ALBITAR	50 %	sept. 2005 -	24

V. TRANSFERTS ET VALORISATIONS TECHNOLOGIQUES

1. Valorisation par la vision industrielle

Depuis que j'enseigne à l'université, j'ai pu assister, puis participer à l'essor de la vision industrielle. De plus en plus d'applications dans la production industrielle font appels à l'inspection visuelle automatique. C'est un moyen de contrôler, sans contact, des dimensions, de repérer des défauts d'aspects, d'évaluer un positionnement relatif de pièces assemblées. La fiabilité de plus en plus élevée pour ce type de mesure, la souplesse d'utilisation avec les capteurs et les logiciels associés, l'intégration de plus en plus rapide grâce à des interfaces de programmation de plus en plus conviviales et à des caméras embarquant des calculateurs de plus en plus sophistiqués (caméras "intelligentes"), tout ceci allant de pair avec le choix crucial du type d'éclairage et du confinement de la cellule de mesure, ont permis à la vision industrielle d'être de plus en plus évoquées, puis incorporées dans les formations universitaires.

Je citerai ici à titre d'exemples quelques applications auxquelles j'ai pu apporter une expertise voire être sollicité pour encadrer certains projets d'étudiants au sein d'entreprises, souvent locales. Ainsi, l'entreprise Osram (Molsheim, 1998) a souhaité se doter d'une mesure par vision de la hauteur de métallisation sur les ampoules. Lilly France (Fegersheim) utilise une chaîne de mesures par automates de vision pour contrôler le niveau de liquide dans les capsules à insuline, la présence de bulles de gaz et la mise en place correcte du bouchon de fermeture. Le contrôle de plusieurs points de soudure pour la conception d'une ligne de fabrication des pots d'échappement (Roser et Duprat, Bouxwiller) se fait à l'aide d'une caméra embarquée sur un robot à trajectoire pré-programmée. Le décodage correcte des informations renfermées dans les codes Data matrix (Solu-Tech, Illkirch) gravés sur les injecteurs de moteurs Diesel est essentiel pour la programmation correcte des calculateurs électroniques embarqués. La sélection des mirabelles dans la société VegaFruits (Meurthe et Moselle) est effectuée par un contrôle d'aspect et de couleur (et 25 autres paramètres), par une caméra fixée au-dessus d'un convoyeur. Ce sont des manipulations inutiles voire néfastes pour le produit qui sont évitées grâce à l'utilisation de la vision industrielle. De plus, la répétabilité très élevée obtenue dans un environnement contrôlé, la définition de multi-critères et la précision de plus en plus fine en fait un outil très efficace dans la définition d'un process.

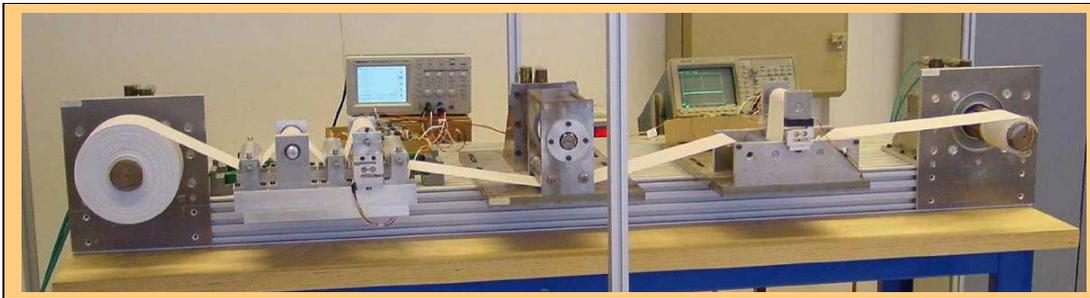
C'est pourquoi je me suis intéressé à développer, partager, puis transmettre certaines techniques simples aux étudiants. C'est aussi, je dois l'avouer, un moyen d'aborder le traitement numérique du signal à un public souvent réfractaire aux fondements mathématiques de cette discipline. Dans la section suivante, je vais décrire brièvement les travaux que nous avons entrepris en partenariat, qui relèvent de l'inspection visuelle dynamique, pour les systèmes industriels d'enroulement de bandes.

Il est à noter que ces partenariats m'ont permis de proposer, dans les cours que j'anime en licence et master professionnels, des interventions régulières d'industriels.

2. Travaux au sein de l'ERT Enroulement

C'est en 1997 que fut initié, par trois collègues, le projet transdisciplinaire sur le système de production qu'est l'enroulement et le transport de bandes flexibles. Patrick Bourgin (Prof. à l'ULP jusqu'en 1999), Christian Gauthier et Dominique Knittel (Maîtres de Conférences à l'IPST, lors de la création de l'Équipe de Recherche Technologique - ERT) ont développé un sujet technologique transversal dans le domaine des systèmes industriels pour adosser l'enseignement de l'IUP GSI (Génie des Systèmes Industriels, à l'IPST) à une recherche technologique en phase avec les centres d'intérêt de l'IUP. Le projet, soutenu par la Fédération de Recherche en mécanique, avait également pour objectif de mettre en oeuvre les compétences existantes dans les laboratoires d'accueil des enseignants-chercheurs de l'IPST, et de créer un sujet multidisciplinaire qui permette à chacun d'enrichir les applications de son thème de recherche.

En s'appuyant sur des activités de recherche antérieures, notamment en mécanique des fluides, en tribologie et en automatique, la demande de reconnaissance d'ERT a été obtenue en 1999, à mi-parcours du contrat quadriennal (1997-2000). Intitulée "Enroulement et transport à haute vitesse de bande flexible", l'ERT n°8 a été évaluée en 2001, reconduite pour le quadriennal 2001-2004 et le quadriennal 2005-2008. L'ERT a été adossée à plusieurs laboratoires dont l'ICS (Institut Charles Sadron, UPR CNRS 22, Strasbourg), le LR2P (Laboratoire de Recherche Pluridisciplinaire en Plasturgie, Ecole Supérieure de Plasturgie d'Oyonnax), le MIPS (Modélisation, Intelligence, Processus, Systèmes, équipe MIAM (Modélisation et Identification en Automatique et Mécanique), EA 2332, UHA de Mulhouse) et le LSIT. Depuis 2005, l'ERT n°8 est une équipe de recherche autonome.

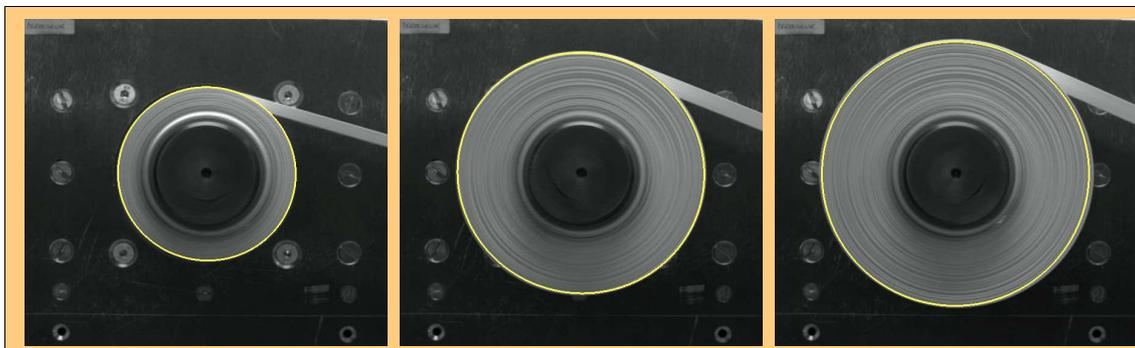


Premier banc à l'ULP/IPST composé d'un dérouleur, d'un tracteur et d'un enrouleur .

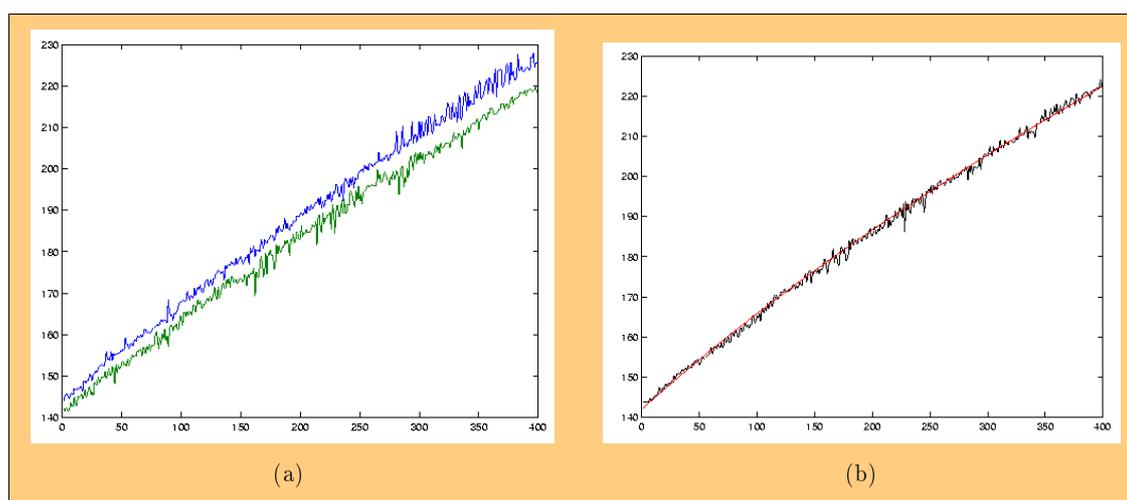
C'est durant l'année universitaire 1999-2000 que j'ai proposé un projet technologique de l'IUP GSI en relation avec les activités de l'ERT. Depuis janvier 2001, j'ai officiellement intégré cette équipe, j'ai mené des activités plus technologiques encore, en m'intéressant à l'apport de la vision industrielle dans le processus d'enroulement de bandes. J'apportais donc un domaine de compétences supplémentaire à l'ERT, celui du traitement des images, plus spécifiquement celui du contrôle visuel dynamique.

En développant un outil de détection de l'excentricité des bobines de matériaux flexibles durant la phase de déroulement et d'enroulement, nous avons proposé en 2001 une amélioration

significative des méthodes usuelles de classification de contours dans les images (emploi des ondelettes continues) [Mey92]) et d'identification des primitives elliptiques (amélioration de la méthode de Pilu, [PFR96]).



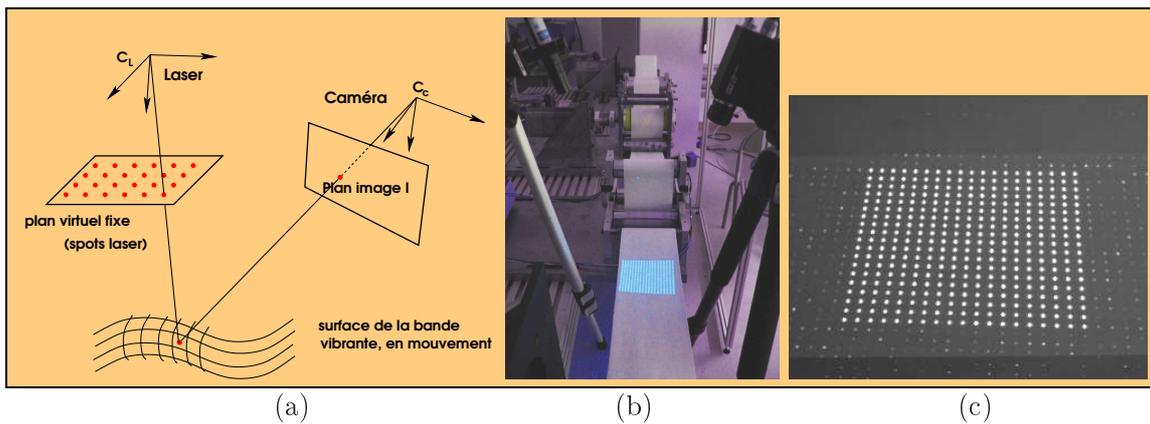
Approximation elliptique du faux-rond (en jaune), superposée au contour extérieur de la bobine, durant la phase d'enroulement (3 images extraites de la séquence).



(a) Mise en évidence du faux-rond par le calcul du grand axe (en bleu) et du petit axe (en vert) de l'ellipse durant la phase d'enroulement. (b) Comparaison entre le rayon moyen mesuré de la bobine et le modèle cinématique proposé pour son évolution (en rouge).

Ce travail, qui reprend certains développements réalisés durant la thèse de Naoufel Werghi en 1996 puis certains résultats obtenus lors du stage de DEA d'Alain Ochs en 2001, a été communiqué lors de la conférence QCAV'01 [DK01], puis a fait l'objet d'un long article en 2002 dans la revue internationale Applied Signal Processing (Eurasip) [DK02].

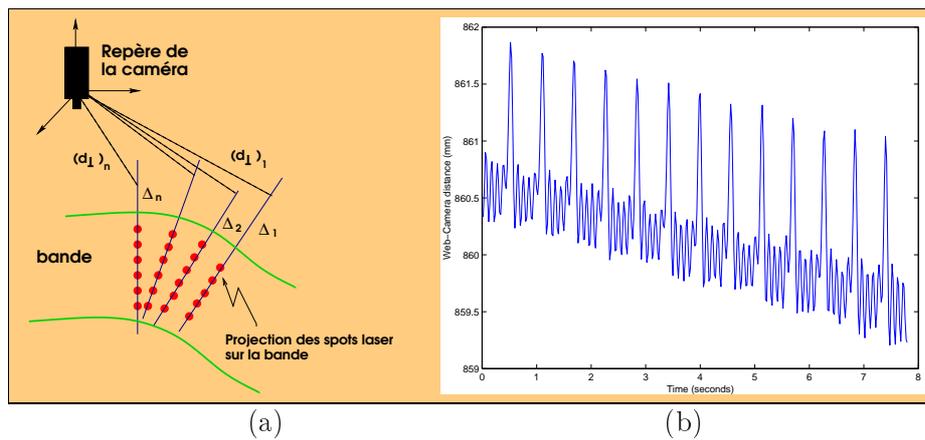
Nous nous sommes également intéressés au transport de bande et nous avons mené une réflexion sur l'amélioration de celui-ci. La bande (de polymère ou de tissu) ne présentant généralement pas une texture décelable à l'échelle macroscopique, nous avons proposé d'employer la lumière structurée (spots lasers), ce qui nous a permis de détecter et de caractériser



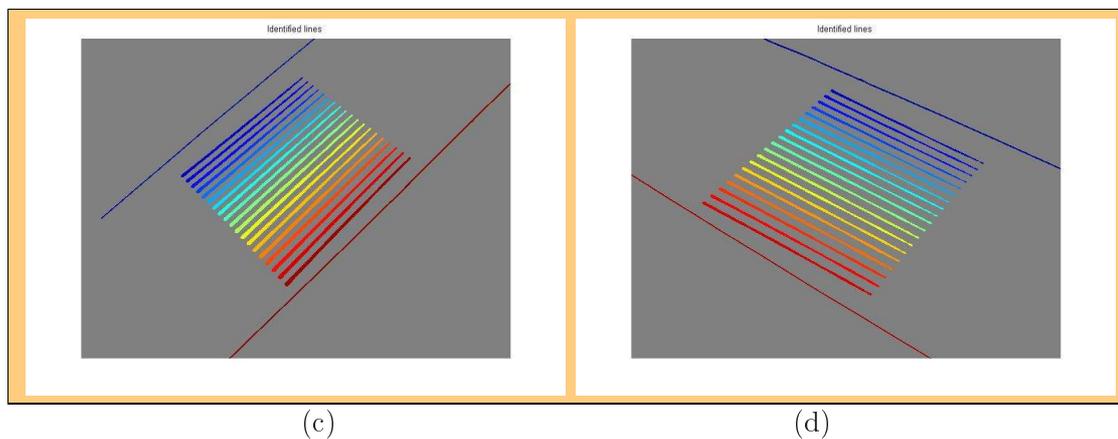
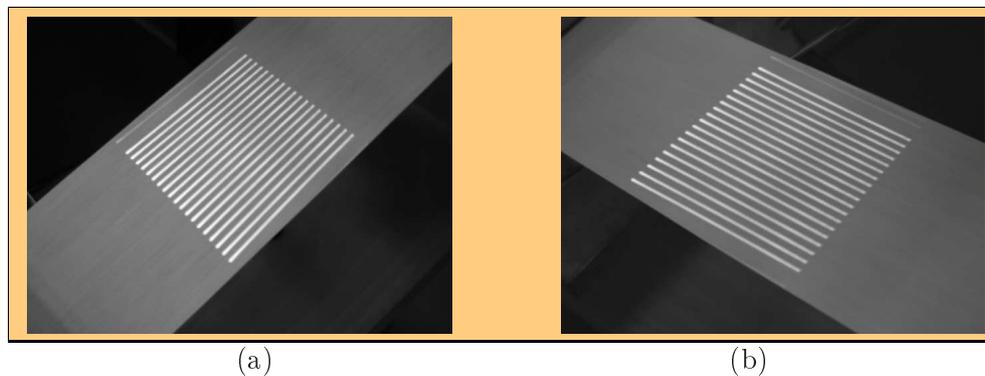
(a) Schéma représentant le dispositif de mesure par lumière structurée. (b) Mise en place sur le système d'enroulement. (c) Projection des spots laser sur la bande.

les fréquences de vibrations hors-plan de bandes flexibles lors du transport (100 à 300 m/min). A travers cette application, nous avons pu valider une technique d'alignement projectif par homographies 1-D. Les résultats obtenus ont été communiqués lors de la conférence mondiale (biennale) de l'enroulement qui se tient à Stillwater (Oklahoma State University, USA) en 2003 [DK03], puis ont fait l'objet d'un article en 2005 dans la revue internationale IAPR Machine Vision and Applications [DK05].

Bien que ce travail ait permis d'introduire un peu de vision industrielle dans un domaine très fortement peuplé de mécaniciens, et de montrer la faisabilité d'une mesure des vibrations sans contact, il s'est avéré que l'emploi de caméras classiques permettant l'acquisition de 25 i/s à la norme CCIR (50 champs dé-entrelacés/seconde) était trop limitatif et qu'il fallait disposer d'un autre système. Une caméra numérique Dalsa permettant d'acquérir 80 i/s puis deux caméras Firewire rapides Basler (> 100 i/s chacune) ont permis l'acquisition de modes propres plus élevés et ont également permis de valider la mesure de l'amplitude des vibrations par reconstruction 3D. L'utilisation de plans laser plutôt que de spots a été retenue pour faciliter la mise en correspondance entre images. Lors du stage INSA (Lyon) de Xavier Maurice, nous avons pu mettre en place ce dispositif et acquérir des données à plus de 400 i/s. Le déport latéral et le déplacement hors-plan de la bande ont ainsi pu être estimés par stéréovision rapide avec lumière structurée. Au vu de ces résultats, un article a été récemment présenté lors de la conférence IEEE IMTC'07 (Instrumentation et Mesures) [MDK07].



(a) Distances orthogonales, $^j d_{\perp}$, entre le centre de projection de la caméra C_c et les droites $\{\Delta_j | j = 1, \dots, n\}$ supportant les ensembles des m (presque collinéaires) spots laser projetés. (b) Distance mesurée (en mm) pour le premier ensemble de spots alignés, durant 8 secondes (50 i/s).



Durant la phase d'enroulement, des plans laser sont projetés sur la surface de la bande en mouvement qui, à leur tour, sont capturés par le système de vision stéréo rapide. (a-b) Un couple d'images capturé par les deux caméras. (c-d) Le suivi des contours par la méthode du moving edge (ME) [Bou89] permet de n'effectuer qu'une seule fois la classification et la mise en correspondance des droites (intersections des plans laser avec une bande de section supposée plane).

VI. PUBLICATIONS SCIENTIFIQUES

Articles dans des revues internationales (4 parus, 1 accepté, 3 en cours d'évaluation)

- C. Doignon et D. Knittel,
A structured light vision system for out-of-plane vibration frequencies location of a moving web, **IAPR Int'l Journal of Machine Vision and Applications**. vol. 16, n° 5, pages 289-297, Décembre 2005.
- C. Doignon, P. Graebbling et M. de Mathelin,
Real-time segmentation of surgical instruments inside the abdominal cavity using a joint hue saturation color feature, **Real-Time Imaging**, vol. 11, pages 429-442, Elsevier, Novembre 2005.
- A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel, L. Soler, J. Leroy et J. Marescaux,
Autonomous retrieval and 3D positioning of surgical instruments in robotized laparoscopic surgery, **IEEE Transactions on Robotics and Automation**, vol. 19, n° 5, pages 842-853, Octobre 2003.
- C. Doignon et D. Knittel,
Detection of non-circularity and eccentricity of a rolling winder by artificial vision, **European Journal of Applied Signal Processing**, pages 714-727, Hindawi Publishing, Juillet 2002.
- C. Doignon, B. Maurin, B. Bayle et M. de Mathelin,
A Visual 3D Tracking and Positioning Technique for Stereotaxy with CT scanners, **Robotics and Autonomous Systems**, Elsevier (accepté, à paraître).
- B. Maurin, C. Doignon, M. de Mathelin et A. Gangi,
A Fast and Automatic Stereotactic Registration with a Single CT-slice, **IEEE Transactions on Medical Imaging** (révision, 2e relecture).
- C. Doignon, D. Knittel et X. Maurice,
A Vision-Based Technique for the Edges Displacement and Vibrations Estimations of a Moving Flexible Web, **IEEE Transactions on Instrumentation and Measurement** (soumis).

- F. Nageotte, C. Doignon et M. de Mathelin,
A Multiple Endoscopic View-Based Technique for Entry Ports Location in Computer-aided/Robot-assisted Laparoscopic Surgery, *IAPR Int'l Journal of Machine Vision and Applications* ([soumis](#)).

Contribution à un ouvrage collectif

- C. Doignon,
Chapitre 20 : An Introduction to Model-Based Pose Estimation and 3D Tracking Techniques, Scene Reconstruction, Pose Estimation and Tracking, livre édité by Rustom Stolkin et publié par *International Journal of Advanced Robotic Systems*, ISBN 978-3-902613-06-6.
- C. Doignon F. Nageotte et M. de Mathelin,
Segmentation and Guidance of Multiple Rigid Objects for Intra-operative Endoscopic Vision, Dynamical Vision, LNCS 4358, pages 314-327, édité par R. Vidal, A. Heyden and Y. Ma, et publié par Springer-Verlag, ISBN 978-3-540-70931-2, 2007, (articles issus des workshops *WDV'05 (ICCV)* et *WDV'06 (ECCV)*).
- C. Doignon, F. Nageotte, B. Maurin et A. Krupa,
Pose Estimation and Feature Tracking for Robot Assisted Surgery with Medical Imaging, Unifying Perspectives in Computational and Robot Vision, livre édité par D. Kragic et publié par Springer-Verlag (accepté, à paraître).

Articles dans des congrès internationaux avec comité de lecture et actes (40, dont 3 invités - 13 conférences IEEE)

2007

- C. Albitar, P. Graebing et C. Doignon,
Robust Structured Light Coding for 3D Reconstruction, IEEE Int'l Conf. on Computer Vision, *ICCV'07*, Rio de Janeiro, Brésil, (accepté pour présentation orale).
- C. Albitar, P. Graebing et C. Doignon,
Design of a monochromatic pattern for a robust structured light coding, Proceedings of the IEEE Int'l Conf. on Image Processing, *ICIP'07*, San Antonio, USA, (accepté).
- X. Maurice, C. Doignon et D. Knittel,
On-line Vibrations Characterization of a Moving Flexible Web with a Fast Structured Lighting Stereovision System, Proceedings of the IEEE Instrumentation and Measurement Technology Conference, *IMTC'07*, Varsovie, Pologne, Mai 2007.

- C. Doignon, F. Nageotte, B. Maurin et A. Krupa, *Model-based 3-D Pose Estimation and Feature Tracking for Robot Assisted Surgery with Medical Imaging*, Proceedings of the Workshop : From features to actions - Unifying perspectives in computational and robot vision, in conjunction with the IEEE Int'l Conf. on Robotics and Automation, Rome, Italie, Avril 2007.
- F. Bourger, C. Doignon, P. Zanne et M. de Mathelin, *A Model-free Vision-based Robot Control for Minimally Invasive Surgery using ESM Tracking and Pixels Color Selection*, Proceedings of the IEEE Int'l Conf. on Robotics and Automation, **ICRA'07**, Rome, Italie, Avril 2007.
- C. Doignon et M. de Mathelin, *A Degenerate Conic-Based Method for a Direct Fitting and 3-D Pose of Cylinders with a Single Perspective View*, Proceedings of the IEEE Int'l Conf. on Robotics and Automation, **ICRA'07**, Rome, Italie, Avril 2007.

2006

- F. Nageotte, P. Zanne, C. Doignon et M. de Mathelin, *Visual Servoing-Based Endoscopic Path Following for Robot-Assisted Laparoscopic Surgery*, Proceedings of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems, **IROS'06**, Pékin, Chine, Octobre 2006.
- C. Doignon, F. Nageotte et M. de Mathelin, *The role of Insertion Points in the Detection and Positioning of Instruments in Laparoscopic Surgery for Robotic Tasks*, Proceedings of the Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention, **MICCAI'06**, Copenhagen, Danemark, Octobre 2006.
- C. Doignon, F. Nageotte et M. de Mathelin, *Segmentation and Guidance of Multiple Rigid Objects for Intraoperative Endoscopic vision*, Proceedings of the 2nd Int'l Workshop on Dynamical Vision, in conjunction with ECCV'2006, **WDV'06**, Graz, Autriche, Mai 2006.

2005

- F. Nageotte, P. Zanne, M. de Mathelin et C. Doignon, *A circular needle path planning method for suturing in laparoscopic surgery*, Proceedings of the IEEE Int'l Conf. on Robotics and Automation, **ICRA'05**, Barcelone, Espagne, Avril 2005.
- B. Maurin, C. Doignon, J. Gangloff, B. Bayle, M. de Mathelin, O. Piccin et A. Gangi, *Needle tip positioning on the abdomen with a novel stereotaxic robotic assistant*, Proceedings of Surgetica'2005 : Computer-Aided Medical Interventions : tools and applications, pp. 157-164, Chambéry, France, Janvier 2005.

- F. Nageotte, P. Zanne, M. de Mathelin, C. Doignon, L. Soler, J. Leroy et J. Marescaux,
Needle path planning for suturing in laparoscopic surgery, Proceedings of Surgetica'2005 : Computer-Aided Medical Interventions : tools and applications, pp. 387-395, Chambéry, France, Janvier 2005.
- F. Nageotte, C. Doignon, M. de Mathelin, P. Zanne et L. Soler,
Circular needle and needle-holder localization for computer-aided suturing in laparoscopic surgery, Proceedings of the SPIE Int'l Symposium on Medical Imaging, volume 5744, pp. 87-98 San Diego (CA), USA, Février 2005.
- B. Maurin, C. Doignon, J. Gangloff, B. Bayle, M. de Mathelin, O. Piccin et A. Gangi,
CT-Bot : A stereotactic-Guided Robotic Assistant for Percutaneous Procedures of the Abdomen, Proceedings of the SPIE Int'l Symposium on Medical Imaging, volume 5744, pp. 241-250 San Diego (CA), USA, Février 2005.

2004

- C. Doignon, F. Nageotte et M. de Mathelin,
Detection of grey regions in color images : application to the segmentation of a surgical instrument in robotized laparoscopy, Proceedings of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems, **IROS'04**, Sendai, Japon, Octobre 2004.
- B. Maurin, J. Gangloff, B. Bayle, M. de Mathelin, O. Piccin, P. Zanne, C. Doignon, L. Soler et A. Gangi,
A parallel robotic system with force sensors for percutaneous procedures under CT-guidance, Proceedings of the Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention, **MICCAI'04**, St Malo, France , Septembre 2004.
- F. Nageotte, M. de Mathelin, C. Doignon, L. Soler et J. Leroy,
Computer-aided suturing in laparoscopic surgery, Proceedings of the Computer Assisted Radiology and Surgery Conference, **CARS'04**, USA, Chicago , Juin 2004.
- C. Doignon et D. Knittel,
A structured light vision system for out-of-plane vibrations estimation of a moving web, Workshop on Applications of Computer Vision, **ACV'04**, in conjunction with the European Conference on Computer Vision, Prague, République Tchèque, Mai 2004.
- F. Nageotte, M. De Mathelin, C. Doignon, L. Soler, J. Leroy et J. Marescaux,
Computer-aided suturing in laparoscopic surgery, Int'l Workshop on Medical Robotics, Navigation and Visualization, **MRNV'04**, Remagen, Allemagne, Mars 2004.

2003

- B. Maurin, C. Doignon, M. de Mathelin et A. Gangi,
Pose reconstruction with an uncalibrated CT imaging device, Proceedings of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition, **CVPR'03**, Madison, USA, Juin 2003.
- C. Doignon et D. Knittel,
Detection and characterization of web vibrations by artificial vision, Proceedings of the seventh Int'l Conf. on Web Handling, **IWEB'03**, Stillwater, Oklahoma, USA, Juin 2003.

2002

- A. Krupa, C. Doignon, J. Gangloff et M. de Mathelin,
Combined image-based and depth visual servoing applied to robotized laparoscopic surgery, Proceedings of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems, **IROS'02**, vol. 1, pages 323-329, Lauzanne, Suisse, Octobre 2002.
- A. Krupa, M. de Mathelin, C. Doignon, J. Gangloff, G. Morel, L. Soler, J. Leroy et J. Marescaux,
Automatic 3-D positioning of surgical instruments during robotized laparoscopic surgery using automatic visual feedback, Proceedings of the fourth Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention, **MICCAI'02**, LNCS 2488, pages 9-16, Tokyo, Septembre 2002.
- A. Krupa, C. Doignon, J. Gangloff, M. de Mathelin, G. Morel, L. Soler, J. Leroy et M. Ghodoussi,
Towards semi-autonomy in laparoscopic surgery : first live experiments, Proceedings of the Int'l Symp. on Experimental Robotics, **ISER'02**, Sant'Angelo d'Ischia, Italie, Juillet 2002 (*invité*).
- A. Krupa, M. de Mathelin, J. Gangloff, C. Doignon et G. Morel,
A vision system for automatic 3D positioning of surgical instruments for laparoscopic surgery with robots, Proceedings of the 12th Int'l Symposium on Measurement and Control in Robotics, Bourges, France, Juin 2002.
- A. Krupa, J. Gangloff, M. de Mathelin, C. Doignon, G. Morel, L. Soler, J. Leroy et J. Marescaux,
Autonomous retrieval and positioning of surgical tools in robotized laparoscopic surgery using visual servoing and lasers pointers, Proceedings of the Int'l Conf. on Robotics and Automation, **ICRA'02**, Washington D.C., USA, pages 3769-3774, Mai 2002.

2001

- A. Krupa, M. de Mathelin, C. Doignon, J. Gangloff, G. Morel, L. Soler et J. Marescaux, *Development of semi-autonomous control modes in laparoscopic surgery using visual servoing*, Proceedings of the fourth Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention, **MICCAI'01**, LNCS 2208, pages 1306-1307, Utrech, Pays-Bas, Octobre 2001.
- C. Doignon et D. Knittel, *Non-circularity detection of a rolling winder by artificial vision*, Proceedings of the Int'l Conf. on Quality-Control by Artificial Vision, **QCAV'01**, Le Creusot, Mai 2001.

2000

- A. Krupa, C. Doignon, J. Gangloff, M. de Mathelin, L. Soler et G. Morel, *Towards semi-autonomy in laparoscopic surgery through vision and force feedback control*, Proceedings of the seventh International Symposium on Experimental Robotics, **ISER'00**, Hawaï, USA, Décembre 2000 (*invité*).

1999

- C. Doignon et G. Abba, *A practical multi-plane method for a low-cost camera calibration technique*, Proceedings of the fifth European control conference, **ECC'99**, Karlsruhe, Allemagne, Septembre 1999 (*invité*).

1996

- N. Werghi, C. Doignon et G. Abba, *Ellipses fitting and three-dimensional localisation of objects based on elliptic features*, Proceedings of the IEEE Int'l Conf. on Image Processing, **ICIP'96**, vol 1, pages 57-60, Lausanne, Suisse, Septembre 1996.
- N. Werghi, C. Doignon et G. Abba, *Pose Estimation of Object based on Circular Patterns in Monocular Computer Vision*, Proceedings of the Int'l SPIE Symposium on Optical Science, Engineering and Instrumentation, Denver, USA, Août 1996.
- C. Doignon, G. Cortelazzo et R. Frezza, *A robust camera calibration technique for high resolution image acquisition*, Proceedings of the European Conference on Multimedia Applications, Services and Techniques, **ECMAST'96**, Louvain, Belgique, Mai 1996.
- G. Cortelazzo, C. Doignon et R. Frezza, *High resolution visualisation of paintings and frescoes*, Proceedings of the Int'l Congress on Environment/Climate, **ICEC'96**, UNESCO, Rome, Italie, Mars 1996.

1995

- N. Werghi et C. Doignon,
Contour decomposition with wavelet transform and parametrisation of elliptic curves with an unbiased extended Kalman filter, Proceedings of the Second Asian Conference on Computer Vision, **ACCV'95**, vol. 3, pages 186-190, Singapour, Decembre 1995.
- L. Albrecht, C. Doignon, G. Abba et E. Ostertag,
Multi-controller design for robot control with visual feedback and switching strategy, Proceedings of the European Control Conference, **ECC'95**, pages 3480-3485, Rome, Italie, Septembre 1995.
- C. Doignon, G. Abba, N. Werghi et E. Ostertag,
Reconnaissance et localisation d'objets en robotique, Proceedings of the Second Int'l Conf. on Industrial Automation, **AI'95**, Nancy, France, Juin 1995.
- C. Doignon, N. Werghi, G. Abba et E. Ostertag,
Localization of objects by a monocular vision system, Proceedings of the fourth Int'l Conf. on Intelligent Autonomous Systems, **IAS'4**, pages 513-520, Karlsruhe, Allemagne, Mars 1995.

1994

- C. Doignon, G. Abba et E. Ostertag,
Recognition and localization of solid objects by a monocular vision system for robotic tasks, Proceedings of IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems, **IROS'94**, Munich, Allemagne, pages 2007-2014, Septembre 1994.

1992

- C. Doignon, F. Imbert, G. Abba et E. Ostertag,
Attitude of known objects from a single view for robot feedback control, Proceedings of the first Japanese-French congress of Mechatronics, Besançon, France, Octobre 1992.

Articles dans des congrès nationaux avec comité de lecture et actes
(9, dont 2 en soumission)

- C. Albitar, P. Graebing et C. Doignon,
Reconstruction 3D en temps réel de scènes intra-abdominales : applications à la robotique et à la réalité augmentée, 16^e congrès francophone AFRIF-AFIA sur la Reconnaissance des Formes et l'Intelligence Artificielle, **RFIA'08**, Amiens, France, Janvier 2008 (soumis).

- C. Albitar, C. Doignon et P. Graebing,
Décodage en temps réel de la lumière structurée codée pour des applications robotiques dans les environnements complexes, Actes (poster) des Journées Nationales de Recherche en Robotique, **JNRR'07**, Octobre 2007 (*soumis*).
- C. Albitar, P. Graebing et C. Doignon,
Conception d'un motif monochrome pour le codage robuste de la lumière structurée, Actes des 11^e journées des Jeunes Chercheurs en Vision par Ordinateur, **Orasis'07**, Obernai, Bas-Rhin, France, Juin 2007.
- C. Doignon, B. Maurin, B. Bayle et M. de Mathelin,
Vers un asservissement visuel en tomodensitométrie à rayons X, Actes du 15^e congrès francophone AFRIF-AFIA sur la Reconnaissance des Formes et l'Intelligence Artificielle, **RFIA'06**, Tours, France, Janvier 2006.
- C. Delimal, C. Doignon et D. Knittel,
Localisation de fréquences de vibration d'une bande en mouvement à l'aide de la vision à lumière structurée, Actes des 10^e journées des Jeunes Chercheurs en Vision par Ordinateur, **Orasis'05**, Fournols, Puy-de Dome, France, Mai 2005.
- B. Maurin, C. Doignon, M. de Mathelin et A. Gangi,
Algorithmes d'estimation automatique de la pose pour la radiologie interventionnelle assistée par ordinateur, Actes des 10^e journées des Jeunes Chercheurs en Vision par Ordinateur, **Orasis'05**, Fournols, Puy-de Dome, France, Mai 2005.
- A. Krupa, M. de Mathelin, C. Doignon, J. Gangloff, G. Morel, L. Soler, J. Leroy et J. Marescaux,
Automatic positioning of surgical instruments during laparoscopic surgery with robots using automatic visual feedback, Actes de la Conference on Modelling and Simulation for Computer aided Medecine and Surgery, **MS4CMS'02**, Ed. ESAIM-Proceedings, Rocquencourt, France, Novembre 2002.
- N. Werghi, C. Doignon et G. Abba,
Localisation tridimensionnelle d'objets basée sur des primitives elliptiques, Colloque AGI'96 (Automatique Génie informatique Image), **AGI'96**, Tours, France, Juin 1996.
- N. Werghi, C. Doignon et G. Abba,
Localisation d'objets présentant des primitives circulaires, 5^e journées des Jeunes Chercheurs en Vision par Ordinateur, **Orasis'96**, GDR-PRC Communication Homme Machine, Clermont-Ferrand, France, Mai 1996.
- C. Doignon, G. Abba et E. Ostertag,
Reconnaissance et localisation d'objets en robotique par un système de vision monoculaire, Colloque AGI'96 (Automatique Génie informatique Image), **AGI'94**, Poitiers, France, Juin 1994.

Articles dans des GDR (4)

- C. Doignon, C. Delimal et D. Knittel,
Détection de fréquences de vibrations transverses d'une bande flexible en mouvement à l'aide de la lumière structurée, Journée TSI éclairage structurée pour la vision active, Le Creusot, France, 27 Mai 2004.
- A. Krupa, C. Doignon, J. Gangloff, M. de Mathelin, G. Morel et L. Soler,
Réalisation de tâches semi-autonomes avec un robot de chirurgie laparoscopique par retour d'effort et asservissement visuel, Journées du GDR-ISIS, GT 5 (Vision), INRIA Rennes, Décembre 2000.
- C. Doignon et P. Graebbling,
L'apport de contraintes contextuelles dans l'appariement d'indices visuels géométriques, Journées du GDR-ISIS, GT 5 (Vision), Université d'Amiens, Juillet 2000.
- C. Doignon et G. Abba,
Calibrage d'un système de vision monoculaire, GRECO-Robotique, Ecole Centrale de Nantes, Avril 1993.

Thèses des doctorants co-encadrés (5)

- Naoufel Werghi,
Analyse et interprétation d'images pour la reconnaissance et la localisation d'objets en vision robotique, Thèse de Doctorat, Université Louis Pasteur, 5 Décembre 1996.
- Alexandre Krupa,
Commande par vision d'un robot de chirurgie laparoscopique, Thèse de Doctorat de l'INP Lorraine, 4 Juillet 2003.
- Florent Nageotte,
Contributions à la Suture Assistée par Ordinateur en Chirurgie Mini-Invasive, Thèse de Doctorat, Université Louis Pasteur, 24 Novembre 2005.
- Benjamin Maurin,
Conception et réalisation d'un robot d'insertion d'aiguille pour les procédures percutanées sous imageur scanner, Thèse de Doctorat de l'Université Louis Pasteur, 28 Novembre 2005.
- Chadi Albitar (thèse en cours),
Titre provisoire : *Codage robuste de la lumière structurée pour la reconstruction 3D de surfaces en temps réel : application à la réalité augmentée et au guidage de robots*

de chirurgie. Doctorant en thèse depuis septembre 2005 à l'Université Louis Pasteur.

Thèse de doctorat (1)

- C. Daignon, *Etude et réalisation d'un système de vision monoculaire pour la reconnaissance et la localisation d'objets en robotique*. Thèse de doctorat, Université Louis Pasteur, 20 janvier 1994.

Deuxième partie

SYNTHÈSE DES ACTIVITÉS DE RECHERCHE

Chapitre 1

Introduction générale

Les travaux que je présente dans cette partie portent sur la problématique du guidage par la vision monoculaire artificielle. L'objectif visé est de pouvoir déterminer le déplacement, voire l'attitude d'un objet d'intérêt (la cible) dans l'espace 3-D en exploitant conjointement des informations visuelles et des données relatives à l'objet d'intérêt ou à son environnement. En fonction de la connaissance a priori sur la structure de cette cible, ceci peut être traité en employant des techniques de recalage ou certaines approches du suivi des mouvements apparents. Cette problématique ayant de multiples connexions avec les asservissements visuels, les applications que l'on trouve en robotique sont nombreuses.

Ce rapport s'appuie pour une très large part sur les travaux de thèses de doctorat de Naoufel Werghi (1994-1997), d'Alexandre Krupa (2000-2003), de Benjamin Maurin (2001-2005) et de Florent Nageotte (2001-2005) que j'ai eu le plaisir de co-encadrer. Il concerne également les travaux de la thèse en cours de Chadi Albitar (2005-) et de stagiaires en Master Recherche ISTI.

1.1 GÉNÉRALITÉS

1.1.1 Contexte

Le contexte applicatif et la finalité de ces travaux concernent l'analyse des images provenant de dispositifs d'imagerie utilisés en médecine et en chirurgie en vue de réaliser des tâches de positionnement d'instruments, tâches qui peuvent être réalisées, partiellement ou intégralement, par la mise en oeuvre d'une commande ou par un opérateur assisté (assistance informatique pouvant s'exprimer de plusieurs manières, typiquement par un simple retour visuel ou par augmentation de la réalité : affichage de valeurs numériques, augmentation d'un retour d'effort, affichage par superposition d'indices visuels,...). Il m'est donc apparu plus approprié d'employer le terme "guidage" plutôt que "commande" dans la mesure où, lorsqu'il s'agit d'apporter une assistance au praticien, celui-ci fait appel également à son expérience de la réalisation du geste et à ses connaissances de l'environnement (anatomie du patient, comportement des tissus, temps d'exposition acceptable,...).

Plusieurs types d'imageurs ont servi de support à ces travaux : les caméras classiques, les endoscopes, les scanners à imagerie tomographique, la lumière structurée classique, ainsi que, plus récemment, la lumière structurée codée.

Comme je l'ai indiqué en préambule de ce document, ma carrière d'universitaire est décomposée en deux époques : la première s'étendant de 1994 à 1999 avec une forte implication dans les activités liées à l'enseignement et qui inclut également le séjour post-doctoral à l'université de Padoue (1995-1996), la seconde depuis 2000 en participant aux projets en vision et robotique médicale. C'est au cours de ces 8 dernières années que j'ai pu construire un projet de recherche, m'amenant à présenter ce mémoire. J'ai en effet eu l'opportunité de participer à partir de la fin 1999 à plusieurs projets d'envergure nationale me permettant de mettre en oeuvre une production scientifique plus importante et de valoriser des développements technologiques. Durant cette dernière période, j'ai également obtenu une délégation du CNRS pour une période de 18 mois (septembre 2005 à février 2007) auprès du LSIIT et dont le programme scientifique concerne le guidage intra-opératif d'instruments de chirurgie par vision.

Les activités présentées ici contribuent au développement de la thématique intitulée « Robotique et Vision » de l'équipe AVR du LSIIT. Certains sujets de recherche transversaux m'ont amené cependant à travailler conjointement avec des collègues de l'équipe MIV (Modèle, Image et Vision) du même laboratoire, en mobilisant des compétences complémentaires en analyse d'images. Depuis 2000, la recherche que j'ai menée dans l'équipe AVR a eu pour cadre des projets et des programmes ayant de multiples intersections. Il s'agit notamment de :

1. L'action incitative concertée (ACI) "jeune chercheur" obtenue en 2000 et initiée par Guillaume Morel, programme intitulé "Automatisation du geste chirurgical par asservisement visuel" en (premier) partenariat avec l'IRCAD (Institut de Recherche contre les Cancers de l'Appareil Digestif de Strasbourg).
2. La participation au programme national ROBEA (ROBots et Entités Artificielles) du CNRS, dans le projet IRASIS (Insertion Robotisée d'Aiguille Sous Imagerie Scanner) en

collaboration avec le LGECO (ex-LICIA à l'INSA de Strasbourg), le LAAS et l'IRCAD.

3. La participation au programme intitulé «Imagerie et Robotique Médicale et Chirurgicale» (IRMC). C'est un des trois programmes mobilisateurs du LSIIT, programme soutenu par la Région Alsace et qui sera développé dans le paragraphe 1.2.
4. L'obtention d'un Bonus Qualité Recherche (BQR) en 2006 qui correspond à un appel à projets du conseil scientifique de l'ULP. En tant que co-responsable de ce projet intitulé "Développement d'un dispositif endoscopique miniaturisé pour la vision robotique par lumière structurée en chirurgie mini-invasive", je co-encadre avec Pierre Graebbling (Prof. ULP/ENSPS) un doctorant en thèse. Le LSP (Laboratoire des Systèmes Photoniques de Strasbourg) et l'IRCAD sont partenaires de ce projet.

1.1.2 Thèmes abordés

Alors qu'au milieu des années 90, j'achevais une thèse de doctorat sur l'étude et la réalisation d'un système de vision monoculaire dans le but de réaliser un asservissement visuel 3-D d'un robot manipulateur à 6 degrés de liberté [Doi94], la plupart des travaux développés dans ce domaine depuis lors, aussi bien à Strasbourg qu'ailleurs, ont, pour une grande majorité, plutôt mis en oeuvre la commande référencée image (encore appelée asservissement visuel 2-D) réputée plus précise et nécessitant moins de connaissances *a priori* de la scène.

A l'heure actuelle, aussi bien l'exploitation des images médicales que les perspectives de leur utilisation tendent un peu à équilibrer l'intérêt des deux approches. Mais au-delà de ces distinctions, les différentes approches pré-citées partagent la nécessité de *relier les informations visuelles 2-D à des grandeurs géométriques tridimensionnelles et leurs évolutions (3-D+t)* ou bien de *relier l'évolution des informations visuelles 2-D à des grandeurs cinématiques tridimensionnelles (2-D+t)*. Ceci englobe les techniques de recalage 2-D/3-D que l'on nomme usuellement *estimation de pose* dans le cas fréquent où on modélise ce recalage par à une transformation euclidienne 3-D (on parle aussi de recalage rigide) et qui peut être intégré dans la problématique plus générale du suivi 3-D. Mais cela surtout distingue les travaux relevant du suivi visuel 2-D d'une part de ceux qui concernent le suivi visuel 3-D d'autre part.

Alors que le suivi visuel 2-D consiste à décrire l'évolution de caractéristiques dans une séquence d'images bidimensionnelles (mouvement global apparent, occlusion réelle apparente ou virtuelle, déformation apparente, variation d'un attribut chromatique,...), la formulation obtenue n'intègre pas forcément une modélisation du mouvement tridimensionnel, parfois même la préoccupation principale n'est pas de savoir si ces variations peuvent provenir d'une modification réelle dans la scène telle que celles émanant du déplacement ou d'une déformation d'un objet d'intérêt, du mouvement de l'arrière-plan, d'une occultation réelle ou d'un changement d'illumination de la scène. Seules certaines approches du suivi visuel 2-D, et plus précisément celles qui permettent de retrouver le déplacement 3-D de l'objet d'intérêt, autorisent la réalisation d'un guidage par la mise en oeuvre d'un asservissement visuel.

En ce qui concerne le suivi 3-D, l'évolution des informations visuelles 2-D sont toujours induites

par des variations spatio-temporelles dans la scène et de la scène vis-à-vis des capteurs visuels ¹.

La commande (et plus généralement le guidage) de robot par vision emprunte de nombreuses techniques relevant du suivi visuel qui est, à l'heure actuelle, une des problématiques les plus étudiées en vision par ordinateur. De nombreux organismes de recherche dans le domaine des STIC ont investi beaucoup de moyens pour solutionner des problèmes liés à l'activité économique des pays industrialisés. C'est qu'il y a évidemment un grand nombre de secteurs économiques qui s'y intéressent. Jugez plutôt, pêle-mêle : qu'il s'agisse de la vidéo-surveillance des zones à fort trafic (aéroport, métro, carrefours routiers, tunnels), du guidage d'engins sans pilote, aériens, terrestres ou sous-marins, de la capture de mouvement pour la décomposition d'une figure sportive ou pour la production cinématographique, de l'insertion de fibrobronchoscopes ou de colonoscopes pour la détection de tumeurs, du guidage des instruments de chirurgie pour la suture ou l'ablation d'organes ou pour effectuer des biopsies, du suivi des mouvements des yeux, des lèvres ou des doigts de la main pour les personnes handicapées, de la reconnaissance et du suivi des visages pour la visio-conférence ou la surveillance des foules, du guidage des paramécies par micro-manipulation ou de la sélection de cellules par microscopie active, de l'observation du comportement locomoteur des animaux soumis à des substances pharmacologiques, du suivi de chemin pour l'automatisation des machines agricoles.... il y a là un spectre extrêmement étendu d'applications qui suscite de l'intérêt mais qui nécessite surtout que nous circonscrivions les travaux de recherche à une thématique spécifique au sein de cette problématique. En effet, selon que l'objectif se trouve être l'obtention d'une mesure dense ou paramétrique du mouvement, le suivi visuel peut être abordé de différentes manières. Etant donné que les applications auxquelles nous nous sommes intéressés concernent spécifiquement le guidage de robot par vision, c'est donc plus précisément les techniques du *suivi par estimation paramétrique du mouvement 3-D* qui nous intéressent afin d'associer le suivi (dans les images ou dans la scène) au déplacement de systèmes articulés ou d'une caméra. Cependant, plusieurs techniques développées dans le cadre général du suivi visuel fournissent des résultats convaincants dans la détection de caractéristiques géométriques. Certaines d'entre elles, en lien avec la segmentation d'images, seront évoquées dans ce rapport.

Les différentes approches que nous allons rappeler sont analysées en considérant tout d'abord le capteur visuel monoculaire le plus répandu - la caméra - afin par la suite d'aborder le contexte médical avec des capteurs visuels dédiés pour des environnements et des problèmes spécifiques. En particulier, la sensibilité du suivi vis-à-vis des problèmes d'occultations partielles, de changement d'illumination ou de couleur ou la modification de l'arrière-plan sera examinée car il s'agit de perturbations communément rencontrées lors d'un guidage par vision.

Plus exhaustivement, nos contributions scientifiques et technologiques concernent plusieurs aspects du guidage, dans un spectre assez large, que je propose de décomposer de la manière suivante :

- la modélisation géométrique de capteurs visuels en imagerie médicale, le recalage et le positionnement 3-D,
- les techniques de suivi visuel permettant la mise en oeuvre d'asservissements visuels 2-D, 3-D ou hybrides,

¹ Les variations provenant d'une modification des paramètres des capteurs ne sont pas considérées ici.

- le choix d'attributs visuels pertinents et d'une représentation appropriée pour la mise en oeuvre d'un suivi robuste,
- le filtrage pour le (pré-)traitement d'images couleur en temps réel,
- la formulation spécifique des mouvements contraints pour le guidage en chirurgie à invasion minimale,
- la conception de dispositifs endoscopiques pour la vision active dans les environnements complexes.

1.1.3 Thèmes non abordés

Délibérément, nous n'allons pas nous intéresser à l'aspect prédictif du suivi. Cet aspect est délaissé car nous considérons que la fonction de prédiction rend tous les processus de suivi plus robustes [TV98]. Ce n'est donc pas un élément discriminant en soi. De plus, la loi de transition (de l'état) dans la modélisation dynamique est bien souvent non accessible, ce qui amène alors à proposer des hypothèses commodes pour l'emploi des estimateurs linéaires ou Bayésien. Quand la fréquence d'acquisition est élevée par rapport aux modifications des mouvements apparents, la prédiction peut être réalisée à l'aide du filtrage de Kalman [Kal60] (original, étendu [Pic91] ou sans parfum (*unscented*) [JUDW95, JU97, WM00]) ou d'un filtrage particulière (ou *bootstrap filter*) [GSS93]. Les différences entre ces deux estimateurs récursifs optimaux concernent surtout la modélisation de la propagation de l'état (plus restrictives pour le filtre de Kalman) et la modélisation des perturbations, mono-modale et gaussienne pour le premier ou multi-modales pour le second. Ce dernier point a tout de même permis de montrer la supériorité du filtrage particulière vis-à-vis du filtrage de Kalman, par exemple lors du suivi de contours en présence d'occultations (Michael Isard et Andrew Blake, [IB98]) ou lors du suivi de multiples objets d'intérêt (Patrick Pérez *et al.* [CP02]). D'autres travaux ont également montré la supériorité du filtre particulière pour les systèmes fortement non-linéaires et quant à la convergence de l'estimation vis-à-vis de l'initialisation. Cependant, je souhaiterai nuancer mon propos (et pas mettre au rancart le filtrage de Kalman que j'ai employé avec succès à plusieurs occasions) car les performances du filtrage particulière sont pour beaucoup dues au nombre élevé et à la stratégie de (re-)échantillonnage des particules (typiquement 200 particules) et de la communication des particules éventuellement (par modèle de graphes [SOD04]). Il s'ensuit que les algorithmes implémentés nécessitent beaucoup plus de temps de calculs que ceux correspondant au filtrage de Kalman.

1.1.4 Motivations

Ce n'est pas un mémoire traitant spécifiquement des asservissements visuels à partir d'images médicaux qui est proposé ici (bien qu'il s'agisse d'une des finalités envisagées), mais plus largement de la "perception visuelle" pour décrire des contributions de natures différentes (des données images au positionnement) pour un objectif commun : le guidage par vision. D'ailleurs, l'obtention d'un guidage robotisé ne constitue pas l'unique objectif ; la plupart des enjeux économiques qui sont visés ici concernent l'aide au geste chirurgical tout autant par l'apport d'une assistance informatique avec interface homme-machine que par celui d'une assistance robotique semi-autonome.

A travers ces activités, des motivations scientifiques personnelles ont également pu s'exprimer. En premier lieu, j'avais le souhait, à l'issue de la thèse de doctorat, de développer de nouvelles techniques ou simplement d'adapter des techniques récentes en traitement des images pour les rendre compatibles avec les contraintes inhérentes à la commande de robots. Ceci fait référence à deux aspects essentiels qui particularisent la vision robotique. D'une part, les traitements envisagés sur les informations vidéo doivent être *automatiques* ce qui sous-entend de parvenir à un haut niveau de fiabilité. D'autre part, dans le cas de la génération automatique de mouvements, les temps de calculs doivent être compatibles avec les équipements existants. On aborde alors la notion de système *temps réel* qui se traduit ici par l'emploi d'algorithmes déterministes. Ce sont selon moi ces raisons qui ont, pendant longtemps, contraint l'emploi de techniques simples et éprouvées de segmentation (comme la binarisation et les multiples traitements des images binaires proposés [Shi87, HS92, Cor96]) et de reconnaissance [DH73, Hor86, CD86] applicables à des environnements très structurés (scènes du type "blocks world" [Eji72, SI73, Agi77]) voire marqués. Si cela a suffi pour valider des approches nouvelles en commande par vision, cela a également conduit à sous-exploiter l'information visuelle en délaissant d'autres informations capturées. Au bout du compte, cela a abouti à limiter le champ d'applications pour finalement se retrouver démuné face à des environnements complexes². Parallèlement, des avancées notables en vision par ordinateur étaient réalisées mais les algorithmes proposés nécessitaient des temps de calculs jugés prohibitifs au milieu des années 80 telles que la transformation de Paul Hough [Hou62], la détection optimale des contours de John Canny [Can86], l'algorithme d'estimation robuste RANSAC de Martin Fischler et Robert Bolles [FB81]. Ces algorithmes (ou leurs variantes) sont à l'heure actuelle employés dans de très nombreux domaines, incluant la robotique.

En second lieu, je souhaitais que des tâches plus élaborées, qui puissent prendre en compte le suivi de paramètres locaux, la présence d'artefacts ou d'informations aberrantes, soient envisagées et incorporées dans un schéma de commande par vision. Il s'agit par exemple du suivi des caractéristiques statistiques d'une région (évolution temporelle de l'intensité moyenne, de la couleur moyenne, de leurs fluctuations), de la cohérence temporelle de la forme pour les objets rigides, de la cohérence du mouvement (tels que les mouvements contraints en chirurgie à invasion minimale) ou encore la cohérence spatio-temporelle des couleurs (pour la gestion des occultations, la détection des spéculaires). En apportant des solutions pour un grand nombre de ces aspects, les tâches robotiques que nous avons mises en oeuvre ont permis entre autre d'améliorer les conditions d'obtention d'un diagnostic ou de faciliter la thérapie en assistant le chirurgien : il s'agit par exemple de la récupération d'instruments de laparoscopie, de l'aide à la suture d'organes dans la cavité abdominale ou encore du positionnement et du guidage d'aiguille de radio-fréquence en radiologie interventionnelle. Ce sont des réalisations qui ont nécessité l'implication de la majorité des membres de l'équipe AVR, en collaboration avec d'autres organismes de recherche et qui ont soulevé des problèmes de précision, de fiabilité et d'implémentation efficace des algorithmes de vision.

² Si une scène contient plusieurs objets d'intérêt, éventuellement en mouvement, interagissant différemment avec la lumière, mais également un arrière-plan en mouvement ou bien non-uniformément éclairé ou encore présentant des variations d'illumination ou de couleur, cet environnement peut être qualifié d'environnement complexe en référence à la difficulté d'en dégager les informations pertinentes par les techniques connues à ce jour.

1.2 PRÉSENTATION DU LSIIT ET DE IRMC

Le LSIIT est un laboratoire multidisciplinaire de l'ULP, une unité mixte de recherche ULP-CNRS (7005). Les recherches qui y sont menées concernent des thématiques des STIC et la majorité d'entre elles s'articulent autour des sciences de l'image. Au CNRS, il dépend du département ST2I (Sciences et Technologies de l'Information et de l'Ingénierie) et secondairement du département MPPU (Mathématiques, Physique, Planète et Univers). Il est composé de 7 équipes regroupant chacune plusieurs thèmes de recherche fondamentale et appliquée. La structuration actuelle des équipes est construite sur des thématiques, elles-mêmes décomposées en opérations de recherche. C'est ainsi que l'équipe AVR (Automatique, Vision et Robotique) anciennement appelée GRAViR (Groupe de Recherche en Automatique et Vision Robotique) jusqu'en 2002, est porteuse des thématiques "Robotique et Vision" et "Commande et observation des systèmes complexes".

Le laboratoire a également développé des collaborations entre ses équipes, grâce à trois programmes transversaux "mobilisateurs" dont le programme Imagerie et Robotique Médicale et Chirurgicale (IRMC), sous la responsabilité de Michel de Mathelin. IRMC a pour ambition de développer des systèmes complets d'assistance aux gestes médicaux et chirurgicaux (GMCAO) s'articulant autour de trois thèmes :

- le traitement d'images médicales,
- la modélisation, la visualisation et la simulation,
- la réalité augmentée et la robotique.

Ce programme fédère en interne des activités des équipes MIV (Modèle, Image et Vision), IGG (Informatique Géométrique et Graphique) et AVR. Depuis sa reconnaissance 2002 en tant qu'équipe-projet multi-laboratoire (EPML n°9) du département STIC du CNRS, il inclut des partenaires extérieurs tels que l'Institut de Physique Biologique (IPB - UMR CNRS-ULP 7004), le LRPS (Laboratoire de Recherche en Productique - EA 3434, INSA de Strasbourg) et l'IRCAD. IRMC repose sur de nombreuses collaborations avec des médecins, des radiologues, des chirurgiens et des physiciens des hôpitaux de Strasbourg et fait partie des programmes de recherche soutenus par la Région Alsace.

En 2005, IRMC est devenu un programme pluriformations (PPF) du ministère de la recherche et de l'enseignement supérieur (DSPT 9). Outre les trois équipes du LSIIT pré-citées, il regroupe des activités du LGeCo (EA 3938 et ex-LRPS, Laboratoire de Génie de la Conception) de l'INSA de Strasbourg, de l'IMFS (Institut de Mécanique des Fluides et des Solides - UMR CNRS-ULP 7507, Département Ingénierie) du laboratoire de Neuro-imagerie in-vivo (UMR CNRS 7004, Département Ingénierie) au sein de l'Institut de Physique Biologique, de l'IRCAD, de nombreux services des hôpitaux universitaires, l'unité de recherche en psychopathologie et pharmacologie de la cognition de l'INSERM (U 405) et le Laboratoire des Systèmes Photoniques de l'ULP (LSP - EA 3426). Pour ses actions de recherche et d'innovations technologiques, IRMC collabore avec des partenaires industriels tels qu'Alcatel, France Telecom, Siemens, Karl Storz et Laennext. L'équipe-projet IRMC a obtenu une reconnaissance internationale par son implication majeure dans les activités du pôle de compétitivité «Innovations Thérapeutiques» à vocation mondiale (<http://irmc.u-strasbg.fr>).

1.3 TRAVAUX ANTÉRIEURS

1.3.1 La période post-doctorale (1994-1996)

La période post-doctorale, bien que très courte, est elle-même constituée de deux époques, correspondant à deux situations professionnelles différentes. La première partie (février 1994 - septembre 1995) de 18 mois couvre mes activités de recherche en tant qu'ATER au LSIIT. La seconde (septembre 1995 - juillet 1996) correspond à un séjour post-doctoral.

Durant la première époque, mes activités ont concerné d'une part la valorisation de ma thèse de doctorat, essentiellement par l'obtention de la qualification aux fonctions de Maître de Conférences et surtout par la présentation d'articles lors des conférences internationales IEEE/RSJ IROS 1994 [DAO94], IAS-4 [DAWO95] et ECC en 1995 [ADAO95].

D'autre part, j'ai effectué une partie du co-encadrement de la thèse de Naoufel Werghi (co-encadrement partagé avec G. Abba). Les activités de l'équipe ERIAV (Equipe de Recherche en Images et en Automatisation par Vision) du LSIIT étaient à ce moment-là structurées autour des axes "Commandes robustes et adaptatives" (resp. E. Ostertag) d'une part et "Commande dynamique de robots par vision" (resp. G. Abba) d'autre part. C'est au sein de ce second axe que se positionnent les travaux de Naoufel Werghi qui concernaient plus spécifiquement le recalage 3D rigide de primitives géométriques non-linéaires et leur segmentation dans les images. L'estimation paramétrique des primitives coniques ainsi que l'estimation de l'attitude des cercles dont elles sont l'image ont été abordées avec l'approche optimale du filtre de Kalman étendu (voir la partie IV-4.A pour les détails et les publications associées à ce travail).

Le Département d'Électronique et d'Informatique (DEI) de l'université de Padoue (Italie, province de Vénétie) m'a accueilli en post-doc durant l'année universitaire 1995-1996. Le travail correspondant à cette position était placé sous la responsabilité de Guido Maria Cortelazzo (Prof. à l'université de Padoue) et se situait dans le cadre du programme communautaire européen HCM-MAVI (Capital Humain et Mobilité - Model-based Analysis of Video Information, contrat n° CHRX-CT94-0625).

L'obtention d'une mosaïque par alignement projectif automatique d'images de petites dimensions constituait le sujet de mon travail. Le projet du DEI, en collaboration avec d'autres universités et dans un vaste programme national de numérisation des oeuvres culturelles, était de mettre en place les éléments permettant la visite virtuelle de musées par l'acquisition et la modélisation des oeuvres culturelles italiennes (pas toutes!). En ce qui me concernait, j'avais pour mission de concevoir un programme informatique qui devait aligner automatiquement des images de fresques figurant sur des surfaces planes, sphériques ou cylindriques. Ces données devaient par la suite être utilisées pour une visualisation interactive devant permettre au visiteur de distinguer les détails. J'ai par conséquent longuement étudié les méthodes de warping, d'alignement d'images et d'interpolation des couleurs. J'ai proposé une technique basée à la fois sur des descripteurs locaux (extraction de points) et semi-globaux (statistiques du premier et second ordre) pour parvenir à résoudre rapidement la mise en correspondance de plusieurs centaines de points. J'ai été secondé dans cette tâche par trois étudiants en stage de fin d'études (cf. partie IV-2).

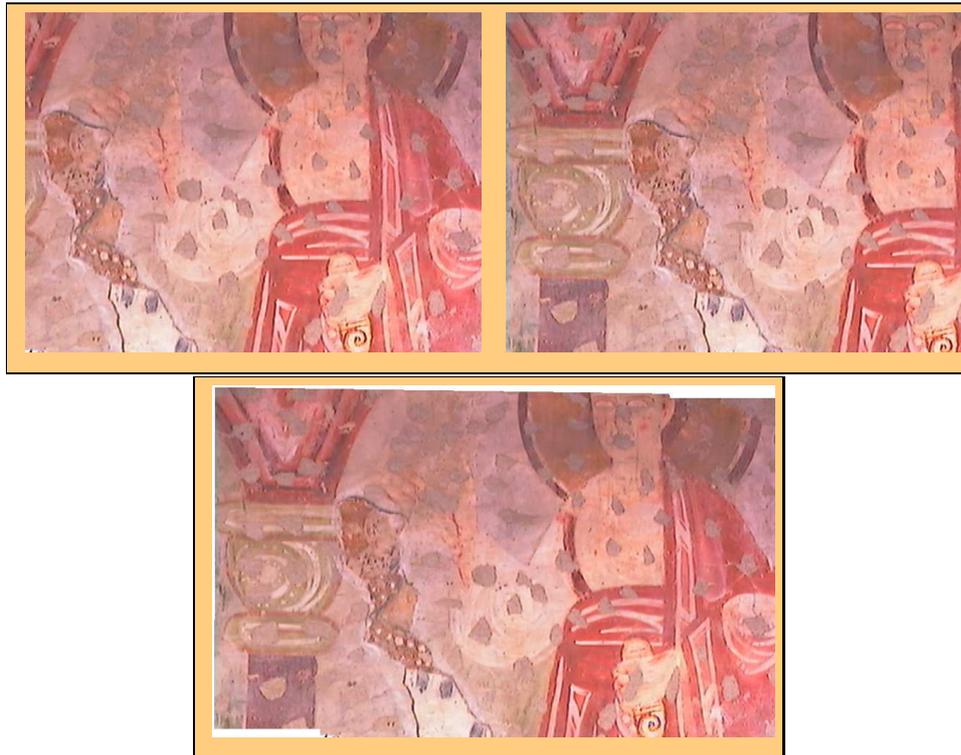


Fig. 1.1: Alignement d'images de fresques et construction d'une mosaïque par "forward warping" et interpolation bilinéaire des couleurs.

Deux communications sont en rapport avec ce travail, l'une porte sur la mosaïque d'images [CDF96], l'autre concerne une technique de calibration de caméras par multi-plans [DCF96], bien antérieure à la parution de l'article de Zhang en 1999 à ICCV [Zha99] et qui emprunte une démarche analogue à celle de Ravn *et al.* [RAS93].

1.3.2 De fin 1996 à 1999

Durant cette période débutant au moment de mon recrutement en tant que Maître de Conférences à l'université Louis Pasteur, j'ai investi la plupart de mon temps à concevoir de nouveaux cours et de nouveaux travaux pratiques, à la responsabilité de la licence réseaux électrique et à la direction de la commission pédagogique de l'IPST (cf. partie II). j'ai par conséquent eu une production scientifique très faible, la seule communication qui est à noter concerne une invitation à ECC (European Control Conference) en 1999 pour présenter la correction de la distorsion des objectifs des caméras lors de la calibration par multi-plans [DA99].

1.4 STRUCTURE DU DOCUMENT

En se voulant synthétique, ce rapport ne peut pas rendre compte de tous les détails et aspects techniques que nous avons rencontrés. Cependant, je m'efforcerai de décrire par endroits quelques points particuliers que je considère comme significatifs, qu'il s'agisse du point de vue de l'originalité que de celui de la mise en oeuvre.

Le chapitre suivant aborde un ensemble de techniques existantes qui composent un système de guidage par vision. Sans être exhaustif, il s'agit de mentionner et de comparer des techniques très diverses qui concernent le recalage, les asservissements visuels, le suivi visuel et la segmentation d'images.

En particulier, les méthodes usuelles de recalage rigide 2-D/3-D seront discutées en fonction de différents types de descriptions et pour différents types de capteurs visuels. Je passerai également en revue un ensemble de méthodes de suivi visuel où je distingue les méthodes locales basées sur des primitives géométriques des méthodes globales basées sur la distribution de l'intensité et de la couleur. Bien que les approches locales du suivi sont privilégiées ici, j'évoque aussi à plusieurs reprises les représentations statistiques auxquelles nous avons de plus en plus recours aussi bien pour tenir compte des données aberrantes que pour représenter globalement les objets d'intérêts par leur apparence visuelle.

Enfin, des techniques de pré-traitement d'images de ces contenus dynamiques sont abordées.

Le chapitre III présente l'ensemble des travaux que nous avons menés sur la problématique du guidage, en mettant en perspective nos contributions scientifiques qui ont concerné, en majeure partie, la détection et le guidage d'instruments chirurgicaux.

Je décris des techniques qui ont notablement contribué à améliorer la segmentation des images pour les applications temps-réel, afin de mieux appréhender la perception des environnements dynamiques et complexes pour la robotique. C'est ainsi que des travaux concernant la détection de caractéristiques visuelles utilisées pour le suivi et le recalage sont également détaillées. Le chapitre se termine par un bilan.

Dans le chapitre IV, je discute des perspectives scientifiques qui me semblent attrayantes tant d'un point de vue disciplinaire que de l'impact qu'elles pourraient avoir dans le domaine du guidage par imagerie médicale. C'est en m'appuyant sur des challenges technologiques en vision dynamique médicale que je propose un certain nombre de pistes à explorer, en particulier autour de la vision endoscopique active.

Chapitre 2

Problématiques communes au guidage par vision

Ce chapitre est consacré à la présentation des éléments majeurs composants un système de guidage par vision. Etant donné la diversité des approches et des cadres de travail possibles, ce sujet est volontairement limité à quatre problématiques : le recalage 2-D/3-D, les asservissements visuels, le suivi visuel et la segmentation d'images en temps réel. D'autres problématiques auraient pu également être abordées, comme le recalage 3-D/3-D, voire multi-modal ou encore la planification de trajectoire. Cette limitation est surtout le reflet des problématiques pour lesquelles nous avons contribué dans nos travaux et que nous présentons dans le chapitre suivant.

2.1 LE RECALAGE 2-D/3-D

Le recalage consiste à déterminer une transformation géométrique entre des données homogènes (comme le recalage d'images) ou entre des données hétérogènes, par exemple entre une image et un patient, plus généralement entre des données mesurées et un modèle. Il peut s'exprimer génériquement par la minimisation d'un critère de similarité f

$$\min_{t \in \mathcal{T}} f(I, t(J)) \quad (2.1)$$

où I et J représentent les deux ensembles de données à recaler, t est la transformation géométrique recherchée appartenant à une classe particulière de transformations \mathcal{T} examinées (euclidienne, affine, projective, ...) [Bro92b].

Il existe plusieurs types de recalage. Le recalage peut être mono-modale (ne concernant que des données provenant d'un seul et même type de système d'imagerie), multi-modale, 2-D/2-D, 2-D/3-D ou 3-D/3-D, voire de dimensions supérieures. Le recalage d'images est très utilisé en imagerie médicale [SF00] ou dans le domaine des géosciences [FB81, GS99] par exemple.

Le recalage 2-D/3-D rigide (ou euclidien positif) permet d'aborder le principe général sur lequel reposent les tâches de positionnement et de suivi 3-D d'un capteur de vision monoculaire par rapport à son environnement. Sous ces conditions, un modèle des objets d'intérêt de la scène est nécessaire pour permettre leur recalage vis-à-vis du capteur visuel. La vision monoculaire étant simplement un capteur 2-D, c'est la projection des informations 3-D sur un plan - le plan image - qui est enregistrée à chaque prise de vue. Une connaissance *a priori* de la structure des objets d'intérêt et un processus de reconnaissance viendront combler cette réduction de la dimension en évaluant une transformation géométrique prenant en compte les degrés de liberté contraints par la forme des objets perçus [DH73, FGL87]. Le suivi par modèle d'objets (*model-based object tracking*) supposés opaques et rigides ici convient aux applications où un modèle géométrique de l'objet à suivre (CAO, par exemple) est envisageable et où la pose (position et orientation) de ces objets doit être déterminée avec précision [VSGA05].

2.1.1 La modélisation

Une modélisation géométrique, à la fois des objets d'intérêt mais également du capteur visuel est nécessaire pour réaliser un suivi 3-D à l'aide d'un capteur monoculaire. Bien que nous allons préférentiellement utiliser la caméra comme capteur visuel, nous verrons d'autres modélisations géométriques, notamment lorsqu'un dispositif d'imagerie tomodynamométrique à rayons X sera employé [Hsi03]. Pour une caméra monoculaire classique, les relations géométriques entre la scène et l'image sont couramment exprimées par la projection perspective [DH73], bien que d'autres modèles, comme celui de Daniel Huttenlocher [HU90] aient été proposés et validés. Ces autres modèles sont généralement des approximations plus ou moins fortes du modèle perspectif (projection faible, projection ortho-perspective et projection para-perspective) [Alo90, DD92, HS92, Alt94]. Ce sont des modèles tout à fait pertinents lorsque l'ensemble de la zone d'intérêt est peu profonde, quand elle se trouve éloignée du capteur, (et d'une manière générale lorsque la distance objet-capteur est grande vis-à-vis de la longueur focale de l'objectif) et quand elle reste proche de l'axe optique. La projection perspective est elle-même une approximation et émane de l'optique géométrique des lentilles minces, pour de faibles

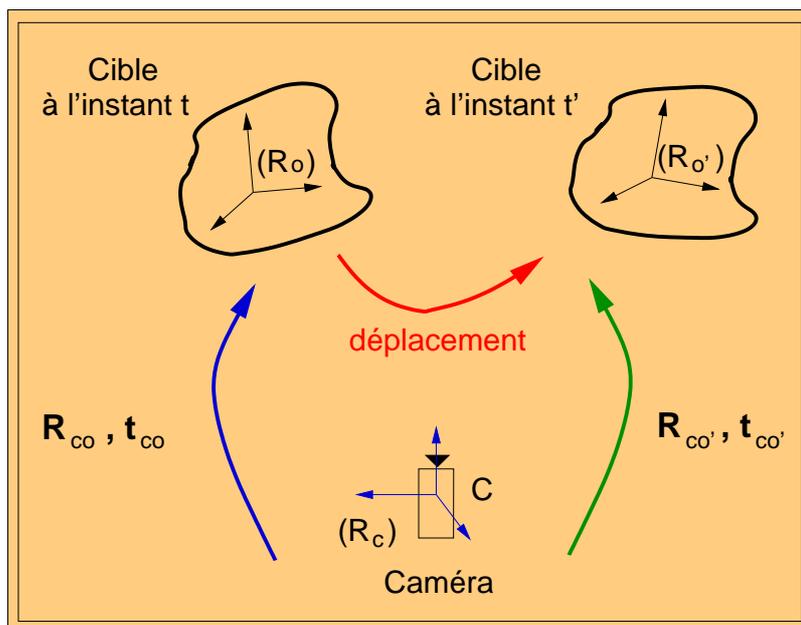


Fig. 2.1: Le suivi visuel 3-D basé sur la succession de recalages rigides 2-D/3-D. Ici, on considère le déplacement d'une cible mobile par rapport à une caméra stationnaire.

inclinaisons des rayons lumineux incidents (conditions de Gauss) et pour une lumière monochromatique [Doi94]. L'utilisation d'une approximation de la projection perspective peut nuire à une interprétation correcte des résultats et il est beaucoup moins pertinent maintenant de passer par une approximation, car de nombreux problèmes inverses ont déjà été traités et résolus à l'aide de la projection perspective exacte [Hor86, HS92, HZ00].

L'obtention d'un modèle de l'objet d'intérêt permet d'établir un critère de similarité entre les entités géométriques qui le représentent et celles issues de la segmentation des images, par l'obtention d'une description commune. Selon la complexité de la forme de chaque entité - que l'on nommera par la suite, primitive géométrique - de son caractère local ou global, cette similarité se traduira par l'élaboration d'un appariement (*feature matching*) ou bien par une comparaison globale à l'aide de descripteurs statistiques ou fréquentiels le plus souvent. Alors qu'une description globale permet d'établir rapidement cette similarité en demandant peu de pré-traitements sur l'image, qu'elle autorise en outre l'incorporation d'informations de texture, cette description souffre d'une forte sensibilité aux occultations, éventuellement aux changements d'illumination et l'obtention du recalage n'est pas immédiat. A l'opposée, la description locale nécessite des traitements sophistiqués permettant d'extraire des primitives géométriques simples (points, coins, droites, coniques) de manière précise ([Pav77, Bal81, Low85, MM86, HS88, HS92, SB97, SMB00, DRLR89, PFR96, WDA96a], par exemple) généralement par une approche de la segmentation basée sur les contours. Elle est moins sensible aux occultations. Le recalage est alors une opération qui est simple lorsque les primitives sont classées et appariées [Zha93, Gro98]. Or, les deux problèmes (recalage et appariement) sont intimement liés et les schémas de résolution relèvent le plus souvent du domaine de la reconnaissance d'objets.

Apparier les primitives (entre les primitives géométriques du modèle et celles extraites de

l'image) est une opération simple et directe lorsque le recalage est déjà établi. De même, la détermination d'une transformation géométrique exprimant le recalage se ramène presque toujours à la résolution d'un système linéaire avec contraintes si les appariements ont déjà été affectés. A l'opposé, l'obtention de l'un et de l'autre simultanément est un problème difficile (dans un cadre général), car à complexité algorithmique élevée [Zha93]. C'est ainsi que de nombreux travaux motivés par l'obtention d'une recherche rapide des appariements ont été menés [FB81, Gri90, Jur99] et que d'autre part, des primitives géométriques contraignant fortement le recalage (le point de vue), saillantes et que l'on peut extraire avec précision, ont été privilégiées. L'étude des invariants projectifs [FMZ⁺91, May92b, Mor93] et différentiels [GB92, Bro92a] prend tout son sens ici (la projection perspective est une transformation projective particulière) en intégrant tout ou partie des paramètres du recalage mais leur évaluation précise se heurte souvent à des problèmes de sensibilité au bruit de mesure et à la discrétisation.

2.1.2 Formulation du recalage rigide

Le recalage 2-D/3-D rigide consiste à déterminer, à partir d'une seule vue, la transformation Euclidienne $(R_{co}, \mathbf{t}_{co}) \in SE(3) = SO(3) \times \mathbb{R}^3$ entre le repère du capteur visuel (R_c) et un repère attaché à l'objet d'intérêt (R_o) (ou cible) comme illustré sur la Fig. 2.1. R_{co} est une matrice de rotation 3-D et \mathbf{t}_{co} est un vecteur de position 3-D. Le suivi tridimensionnel est alors obtenu par deux recalages successifs à deux instants t et t' et convient bien aux asservissements visuel en situation (*look and move*) [Agi77, Mau05]. Le déplacement de la cible entre les repères (R_o) et ($R_{o'}$) pour ces deux instants exprimé dans le repère en mouvement (R_o), est défini par :

$$\begin{cases} R_{oo'} &= R_{co}^T R_{co'} \\ \mathbf{t}_{oo'} &= \mathbf{t}_{co'} - \mathbf{t}_{co} \end{cases}, \quad (2.2)$$

et exprimé dans le repère fixe (R_c) :

$$\begin{cases} R_{oo'} &= R_{co'} R_{co}^T \\ \mathbf{t}_{oo'} &= R_{co} (\mathbf{t}_{co'} - \mathbf{t}_{co}) \end{cases}. \quad (2.3)$$

$\mathbf{t}_{oo'} = R_{co} \mathbf{OO}'$ est la translation entre les origines des repères, exprimée dans (R_c). L'orientation entre les repères peut être représentée par la matrice de rotation $R_{oo'}$ ou bien par un vecteur $\mathbf{r}_{oo'} = \theta \mathbf{u}$ où θ et \mathbf{u} sont respectivement l'angle et le support unitaire de la rotation. Lorsque les composantes de la matrice de rotation sont notées r_{ij} ($i, j = 1, 2, 3$), on a :

$$\begin{aligned} \theta &= \arccos\left(\frac{1}{2} (r_{11} + r_{22} + r_{33} - 1)\right), \\ \mathbf{u} &= \begin{cases} \operatorname{sgn}(r_{32} - r_{23}) \sqrt{\frac{|r_{11} - \cos \theta|}{1 - \cos \theta}} \\ \operatorname{sgn}(r_{13} - r_{31}) \sqrt{\frac{|r_{22} - \cos \theta|}{1 - \cos \theta}} \\ \operatorname{sgn}(r_{21} - r_{12}) \sqrt{\frac{|r_{33} - \cos \theta|}{1 - \cos \theta}} \end{cases} \end{aligned} \quad (2.4)$$

La rotation dans l'espace 3-D peut aussi être exprimée à l'aide d'un quaternion unitaire. Les quaternions sont des nombres hyper-complexes 4-D qui peuvent être écrits comme une combinaison linéaire $a + bi + cj + dk$, avec $i^2 = j^2 = k^2 = ijk = -1$. Un quaternion unitaire q peut alors représenter une rotation d'angle θ et de vecteur unitaire \mathbf{v} quand il est mis sous la forme

$$q = (\cos(\theta/2), \sin(\theta/2) \mathbf{v}).$$

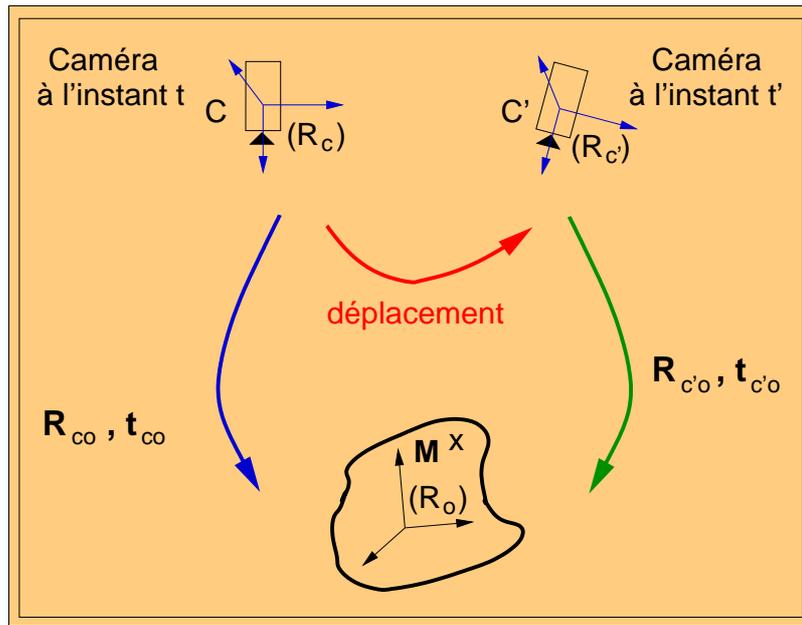


Fig. 2.2: Le suivi visuel 3-D basé sur la succession de recalages rigides 2-D/3-D. Ici, le déplacement de la caméra mobile par rapport à une cible fixe est considéré.

Ainsi, pour orienter un point 3-D M , nous l'exprimons d'abord sous la forme du quaternion $p = (0, \mathbf{M})$, et son homologue orienté $p' = (0, \mathbf{M}')$ par la rotation précédente est obtenu par la composition $p' = q \cdot p \cdot \tilde{q}$, où \cdot est l'opérateur de multiplication des quaternions et $\tilde{q} = (\cos(\theta/2), -\sin(\theta/2) \mathbf{v})$ est le quaternion conjugué de q . D'autres représentations peuvent être employées, comme la projection exponentielle (ou la formule de Rodrigues) [Gan84].

Il est à noter que le déplacement déterminé de cette manière peut être important car il n'est pas nécessaire que les mêmes primitives géométriques de l'objet d'intérêt soient visibles dans les deux vues, et par conséquent aucun recouvrement entre les deux images n'est requis. A un instant donné, t^* , correspond une pose désirée (R_{co^*}, t_{co^*}) entre le capteur et la cible, le suivi consiste alors à minimiser les deux vecteurs t_{oo^*} et r_{oo^*} . La détermination du déplacement par cette formulation autorise aussi bien un déplacement de l'objet d'intérêt qu'un déplacement du capteur. Une analyse équivalente peut donc être réalisée en considérant que c'est l'objet d'intérêt uniquement qui est en mouvement. Seul le déplacement de l'un vis-à-vis de l'autre n'a de sens ici. En tout état de cause, si l'erreur représentée par le vecteur à 6 composantes $e_{oo^*} = (t_{oo^*}^T, r_{oo^*}^T)^T$ est exprimée dans le repère caméra (R_c) (fixe) comme c'est le cas pour la relation (2.3) ou bien si elle est exprimée dans le repère (en mouvement) de la cible (R_o) à partir de (2.2) les trajectoires résultantes dans les images seront différentes.

Dans les applications de suivi 3-D, le mouvement de la caméra entre deux instants proches peut souvent être supposé faible. Ainsi, dans le cas de petits déplacements entre les deux instants t et $t' = t + dt$ et par exemple en supposant que c'est le capteur qui est en mouvement, le suivi est alors caractérisé par les valeurs successives des vitesses instantanées de rotation et de translation du capteur. Par rapport à un repère fixe (repère galiléen ou monde) (R_w) d'origine O_w , supposons que le repère (R_c) lié à la caméra et d'origine C soit animé d'un mouvement de translation $(\mathbf{V}_{R_w}^C)_{R_c}$ et d'un mouvement de rotation $(\Omega_{cw})_{R_c}$. Si un point

géométrique \mathbf{M} de l'objet est fixe dans ce repère monde (voir la figure 2.2) - environnement statique ($\mathbf{V}_{R_w}^M / (R_w) = 0$), alors on peut prendre le repère de la cible comme repère monde ($o = w$). La composition des vitesses nous fournit la relation

$$(\mathbf{V}_{R_o}^M) / (R_c) = \mathbf{R}_{co} (\mathbf{V}_{R_o}^M) / (R_o) = 0 = (\mathbf{V}_{R_c}^M) / (R_c) + (\mathbf{V}_{R_o}^C) / (R_c) + (\boldsymbol{\Omega}_{co}) / (R_c) \times \mathbf{CM}$$

c'est-à-dire, dans le repère (R_c) :

$$\mathbf{V}_{R_c}^M = -\mathbf{V}_{R_o}^C + [\mathbf{CM}]_{\times} \boldsymbol{\Omega}_{co} . \quad (2.5)$$

Cette dernière expression sera également utilisée pour établir le lien entre le mouvement du capteur par rapport à la scène observée et le champ apparent du mouvement du point \mathbf{M} dans les images, quand on y ajoutera le modèle du capteur visuel (modèle perspectif, projection orthogonale, plan d'intersections ou de coupes,...). Un jacobien image (ou matrice d'interaction) issu de la dérivée temporelle du modèle de projection permettra d'établir le lien entre le torseur cinématique représentant le mouvement du capteur et les variations dans l'image des primitives géométriques [Cha90, FLM91, CRE93].

Finalement, si les expressions précédentes d'un déplacement 3-D sont simples à obtenir à partir de recalages successifs, la difficulté est reportée sur l'obtention de chaque recalage 2-D/3-D et l'estimation des paramètres, à intervalles de temps rapprochés, est considérablement facilitée par les positions successives estimées antérieurement. Il existe cependant des situations qui demandent une estimation de paramètres sans *a priori*, c'est notamment le cas lors de l'initialisation du suivi, lors d'occultations, ou de la disparition du champ de vision d'un certain nombre de primitives. Nous allons aborder ce problème en considérant deux catégories d'approches, celles qui sont construites à partir d'une description locale d'une part et celles qui nécessitent le recalage à l'aide d'une description globale d'autre part.

2.1.3 Recalage basé sur les descripteurs locaux

La détermination d'un recalage 2-D/3-D par une description locale nécessite la recherche des appariements des primitives géométriques du modèle avec celles de l'image [HCWS84, Low87, Ike87, TM87, Sug88, BAK91, GH91, Jur99, DDDS02]. Pour les primitives très simples comme des points (sommets d'un objet polyédrique, centres, coins, point d'inflexions,...) ou des droites (arêtes, tangentes d'inflexion,...), le point de vue ¹ est très peu contraint en considérant une seule de ces primitives. Divers schémas de reconnaissance d'objets ont été proposés tels que la recherche arborescente (arbre d'interprétation [Gri90]), le hachage géométrique [LW88], les graphes d'aspect [Ike87, HH89], les primitives focales [BC82, BH86], le clustering de pose et le "Soft Assignment" [GRL⁺98, DDDS02] et l'alignement [FH86, HU90].

Les méthodes d'alignement sont très répandues. Elles permettent d'obtenir directement la transformation géométrique représentant le recalage. A cette fin, il faut apparier des sous-ensembles de primitives géométriques (ordonnées dans un des deux ensembles *modèle* ou *image*) pour contraindre le point de vue et obtenir un nombre fini de solutions pour le recalage. La taille de ces sous-ensembles dépend à la fois de la transformation géométrique

¹ "Le point de vue" correspond à la détermination de l'attitude ou la pose en considérant le problème du recalage rigide comme celui de l'estimation des paramètres extrinsèques d'une caméra étalonnée.

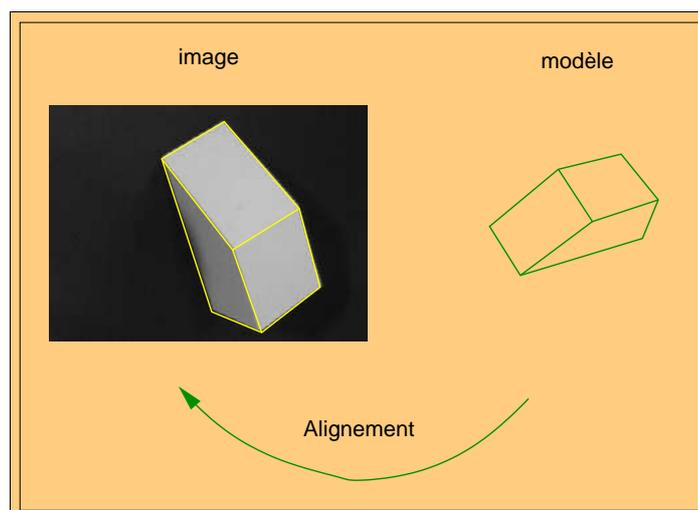


Fig. 2.3: Le recalage par la technique d'alignement d'un modèle sur l'image.

recherchée, de la méthode de résolution et du nombre de degrés de liberté contraints par les primitives. Par exemple, déterminer une homographie linéique (homographie 1-D) entre des points collinéaires requiert des regroupements de sous-ensembles de trois points, car une base projective mono-dimensionnelle est définie par 3 points². Calculer tous les alignements possibles en utilisant un nombre minimum de correspondances entre les primitives du modèle et celles des données visuelles est prohibitif. Ainsi pour m points 3-D dans le modèle et n points image, il y a $A_m^3 C_n^3$ alignements possibles en les groupant par triplets pour rechercher une transformation affine (alignement de Daniel Huttenlocher [HU90]) ou le recalage rigide à partir du modèle perspectif et d'une caméra étalonnée (*The perspective 3-points problem* [Hor87, Har89, DD92, Alt94]). Par exemple, pour le simple objet polyédrique de la figure 2.3 apparaissant seul dans l'image, on a typiquement $m = 8$ sommets et $n = 7$ points visibles au plus dans l'image, ce qui correspond à rechercher 11 760 alignements possibles³. Un seul point image supplémentaire (par exemple un point aberrant provenant d'une modification du rapport signal/bruit ou du contraste dans l'image) fait passer le nombre d'appariements possibles à 18 816, deux à 28 224. Non seulement, il faut réduire l'espace de recherche pour en pratique pouvoir recaler le modèle sur l'image, mais il y a lieu de le rendre moins sensible, d'un point de vue combinatoire, à la présence de points aberrants (*outliers*), autrement dit, il faut réduire la complexité.

La mise en correspondance robuste basée sur la technique des moindres carrés médian [RL87, Ros99] ou sur RANSAC [FB81, HZ00] permet de sélectionner plus ou moins aléatoirement un nombre réduit de sous-ensembles d'échantillons (typiquement de 2 à 5 primitives) dans chaque ensemble de primitives, de supposer qu'ils sont appariés, de calculer la transformation pour le recalage candidat, d'apparier le plus possible de paires de primitives *modèle-image* et au bout du compte de retenir le recalage dont la transformation calculée rassemble le maximum

² Si un quatrième point, collinéaire aux trois autres, est ajouté, on obtient un invariant projectif bien connu : le birapport. Il est équivalent à la coordonnée de ce dernier point dans la base pré-définie.

³ Chaque alignement correspond à 2 solutions pour le recalage avec le modèle para-perspectif et 4 avec le modèle perspectif exact, le capteur étant supposé calibré.

d'appariements (*l'ensemble consensuel*). Il est également possible d'insérer des contraintes (de rigidité, contextuelles) et d'opérer des groupements perceptuels pour éliminer *a priori* de mauvais appariements parmi les appariements hypothétiques [Low85, BFM88]. D'autres types de primitives géométriques peuvent participer au processus de recalage, telles que des coniques qui contraignent plus fortement le point de vue mais qui en revanche sont plus délicates à estimer. Les méthodes existantes permettant de résoudre le recalage rigide à partir d'une seule vue, pour les primitives linéaires (points, droites) sont brièvement décrites dans les deux prochains paragraphes. Le recalage à partir de certaines primitives non linéaires sera décrit plus en détails dans le prochain chapitre, l'estimation de pose à partir de primitives quadratiques faisant partie de nos travaux sur le sujet.

Recalage à partir de primitives linéaires

Le recalage à l'aide de points ou de droites caractéristiques et à partir d'une seule vue fait l'objet de la présente partie. Il est bien connu que dans un plan les points et les droites sont des entités géométriques duales et sont toujours représentées par deux paramètres indépendants. Dans l'espace 3-D ce n'est pas le cas et alors qu'un point géométrique est représenté par 3 degrés de liberté, une droite est une variété de dimension 4.

Recalage à partir des points C'est en fonction de la disposition relative des points (appartenant au même plan ou non), en fonction de leur nombre et de l'approche retenue (analytique ou numérique, prise en compte de la redondance des données) que des solutions ont été proposées depuis environ 25 ans, à partir du modèle perspectif ou d'une de ses approximations et nous renvoyons le lecteur intéressé vers les diverses méthodes proposées dans [Low80, Hor86, FT87, Hor87, Low87, Tsa87, LHD88, HCLL89, DRLR89, Har89, Yua89, LHF90, HLON91, DD95, Ros99, QL99, Tri99, HZ00, DDDS02, AQT02] et à l'état de l'art dans [Chr98]. Généralement, l'extension à des droites est possible dès lors qu'un point correspond à une intersection réelle de droites concourantes (on se ramène alors à des points) et également lorsque les droites sont coplanaires (objet 2-D). Dans le cas d'un objet correspondant à une courbe plane non paramétrique, l'utilisation des inflexions présentes le long de la courbe permet de se ramener à un recalage projectif à partir de primitives linéaires si un nombre suffisant d'entre elles est disponible (4). En effet, si les points d'inflexion sont des primitives très délocalisées, ce n'est pas le cas des tangentes d'inflexion et il est alors possible d'effectuer un recalage rigide entre une courbe et son image en détectant les passages à zéro de courbure qui sont, dans le cas général, des invariants projectifs [MM86, RDL87, RDLR91, Doi94].

Recalage à partir des droites Dans le cas d'un objet polyédrique, un modèle peut être décrit par un ensemble de droites 3-D qui d'une manière générale ne sont pas coplanaires. La dualité droite-point dans un plan ne peut plus être considérée et une méthode spécifique pour les droites doit alors être envisagée. C'est un problème alors plus difficile à aborder que celui des points car les équations reliant toutes les représentations des droites 3-D et leurs projections dans l'image sont des fonctions quadratiques des paramètres de pose. Les méthodes proposées par Dhome [DRLR89], Chen [Che91], Liu [LHF90], Navab [NF96] et Ansar [AD03], entre autres, fournissent des algorithmes d'estimation de pose d'un faisceau de droites. Dans [DRLR89], l'orientation est obtenue par résolution, dans le cas général, d'un polynôme de degré 8 à partir d'au moins 3 droites. Dans un deuxième temps, un système linéaire fournit la

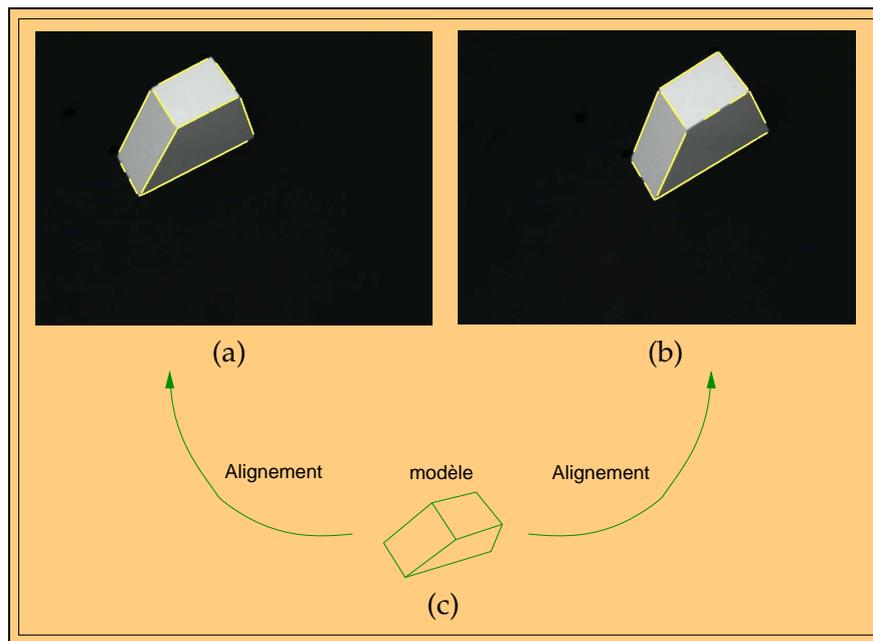


Fig. 2.4: (a-b) Images d'un polyèdre. Le suivi 3-D consiste à déterminer les attitudes successives par des alignements entre les droites du modèle (c) et les droites images de la séquence.

solution. Chen [Che91] met en avant des cas particuliers importants (droites concourantes, 2 droites parallèles sur 3, droites coplanaires, droites mutuellement perpendiculaires,...) afin de diminuer l'ordre du polynôme. De manière synthétique, les diverses approches se distinguent par l'emploi de contraintes géométriques [LHF90, Che91], par l'utilisation éventuelle de la représentation Plückerienne [NF96], (normalisée) [And99], des droites.

Comme tout ou partie des droites (entre autres les extrémités) peuvent être occultés lors d'un suivi 3-D dans des environnements non contrôlés, la résolution du problème d'estimation de pose de droites à partir de la représentation Plückerienne est un problème sur lequel nous nous sommes également penchés car elle a été employée dans plusieurs contributions que nous verrons par la suite.

La représentation de Plücker

Les coordonnées Plückeriennes sont construites à partir de deux vecteurs. Un vecteur directeur de la droite \mathbf{r} et un vecteur \mathbf{w} normal au plan d'interprétation (plan π qui passe par Δ et l'origine du repère O), comme illustré sur la figure 2.5. Ainsi pour tout point $P \in \Delta$, on a : $\mathbf{w} = \mathbf{r} \times \mathbf{OP}$. Il s'ensuit que \mathbf{w} est perpendiculaire à \mathbf{r} et qu'il n'est pas nécessaire de disposer d'une origine sur la droite Δ pour définir le couple (\mathbf{r}, \mathbf{w}) . On remarque que comme $\mathbf{r}^T \mathbf{w} = 0$, si $\|\mathbf{r}\| = 1$, seuls 4 paramètres sont indépendants dans cette représentation ce qui correspond bien au nombre de degrés de liberté d'une droite dans l'espace 3-D euclidien. On a l'habitude de regrouper les coordonnées Plückeriennes dans un vecteur $\mathcal{L} = (\mathbf{r}^T, \mathbf{w}^T)$ à six composantes ou dans une matrice appelée matrice de Plücker L construite à partir du support de deux points 3-D. Cette matrice, ainsi que la matrice duale L^* correspondant à l'intersection de deux plans,

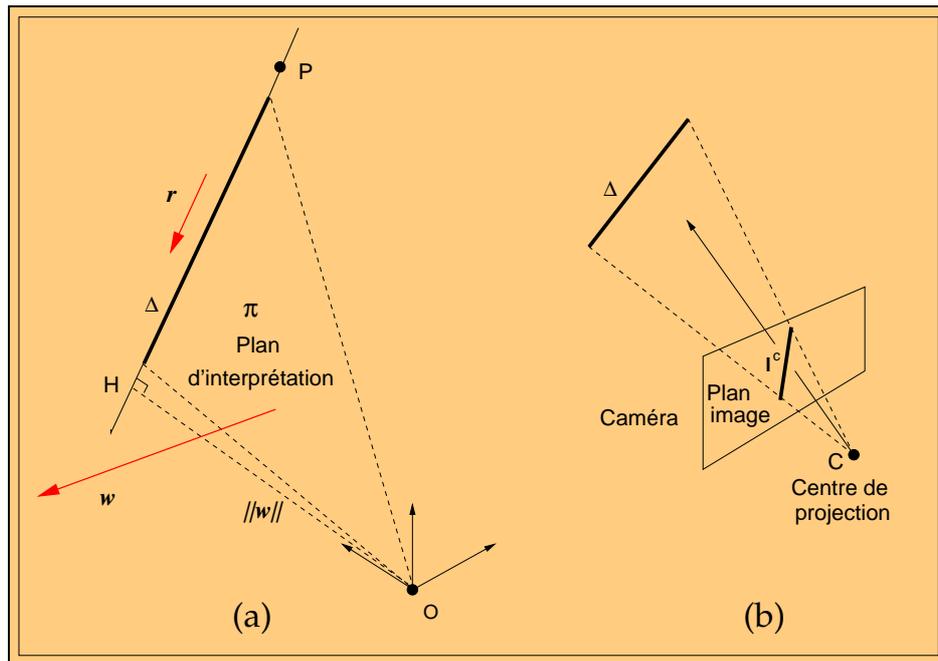


Fig. 2.5: (a) Représentation Plückerienne d'une droite 3-D par le couple de vecteurs (\mathbf{r}, \mathbf{w}) . (b) : Une droite 3-D Δ et son image I^c par projection perspective de centre C .

ont pour expression

$$\mathbf{L} = \begin{bmatrix} [\mathbf{w}]_{\times} & -\mathbf{r} \\ \mathbf{r}^T & 0 \end{bmatrix} \quad \mathbf{L}^* = \begin{bmatrix} [\mathbf{r}]_{\times} & -\mathbf{w} \\ \mathbf{w}^T & 0 \end{bmatrix} \quad (2.6)$$

et $\mathbf{L} \mathbf{L}^* = 0_{4 \times 4}$. L'obtention de \mathbf{L}^* à partir de \mathbf{L} se fait par une règle simple [HZ00] et l'une et l'autre des représentations seront employées par la suite. Ces deux matrices (4×4) sont anti-symétriques et singulières, de rang 2 précisément et la valeur de ce rang exprime la contrainte d'orthogonalité existant entre les vecteurs \mathbf{r} et \mathbf{w} . Par un simple calcul du polynôme caractéristique de la matrice de Plücker, on peut facilement montrer que les valeurs propres de \mathbf{L} sont de la forme $\{\pm i \lambda_1, 0, 0\}$ où $\lambda_1 = \sqrt{\mathbf{r}^T \mathbf{r} + \mathbf{w}^T \mathbf{w}}$ est un nombre réel arbitraire (non nul) qui peut être choisi pour normaliser la matrice.

2.1.4 Recalage basé sur les descripteurs globaux

La résolution de l'équation 2.1 au début de cette section peut être envisagée par l'analyse des distributions globales des données à recalcer. Ainsi, la distribution de l'intensité d'une image ou d'une partie de celle-ci peut être représentée sous différentes formes statistiques pour élaborer un critère de similarité qui permettra de comparer cette zone avec une celle d'une autre image ou d'un modèle. Différentes techniques sont évoquées ci-dessous, dans le cadre du recalage 2-D/3-D et plus spécifiquement dans le calcul de pose.

Dans la littérature, plusieurs méthodes ont été proposées pour résoudre le problème du calcul de pose en utilisant les moments par exemple [LD89, CO93, FS88, FS93]. Dans le cas où on ne dispose que d'une image à l'aide d'une caméra monoculaire, une méthode de calcul

de pose a été proposée quand un mouvement de rotation est considéré entre deux positions de la caméra [Muk92, CR80]. Cette méthode utilise les moments invariants au mouvement de rotation autour de l'axe optique pour estimer le mouvement de rotation autour des autres axes. La valeur de cette rotation est déterminée à partir d'une table de valeurs. Dans [MR96], une méthode qui considère une transformation affine composée de mouvements de translation et de rotation est proposée dans le cas d'un objet plan. Si ceci est exact pour une projection orthographique, cela correspond à une approximation pour une projection perspective.

Pour résoudre le problème du calcul de pose pour des objets de forme complexe, Omar Tahri [Tah04] a proposé deux méthodes utilisant les moments invariants, la première est basée sur les asservissements visuels virtuels [SB98, MC02a], la seconde, s'appuie sur un calcul exhaustif en utilisant une table de valeurs pour la pose. Pour les deux techniques proposées, le calcul est effectué en deux étapes ; en premier lieu, un alignement permet d'estimer numériquement la rotation seule entre les repères, la translation est par la suite calculée analytiquement.

Luca Lucchese s'est intéressé depuis plus de dix ans au recalage entre images par une approche fréquentielle. Bien qu'il ne s'agit pas exactement d'un calcul de pose, il existe de nombreuses similitudes avec l'estimation d'une transformation entre deux images, quand la scène observée est plane ou peut être considérée comme telle. Dans les situations fréquentes où l'effet perspectif n'est pas très prononcé, la transformation homographique H reliant les deux images peut être approchée par une transformation affine, telle que

$$H = \begin{bmatrix} A_{2 \times 2} & \mathbf{v} \\ \mathbf{d}^T & 1 \end{bmatrix} \approx \begin{bmatrix} A'_{2 \times 2} & \mathbf{v}' \\ 0 & 0 & 1 \end{bmatrix} \quad (2.7)$$

La stratégie développée par Lucchese dans [Luc01] a consisté à déterminer les paramètres de la sous-matrice 2×2 $A'_{2 \times 2}$ de H par une minimisation non linéaire basée sur un critère énergétique entre les deux images. Dans un second temps, les composantes du vecteur \mathbf{v}' sont estimés par un calcul standard de corrélation de la phase.

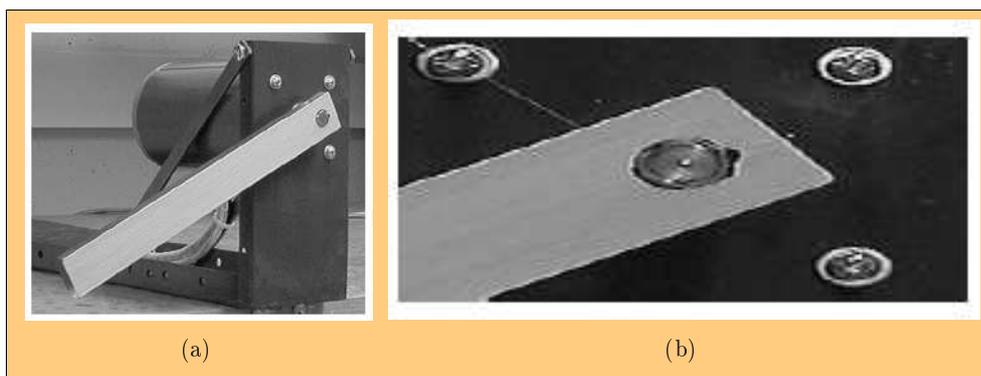


Fig. 2.6: (a) Position angulaire de la pale d'une tourelle à asservir. (b) Image de référence associée à une position angulaire. (Rapport de DEA de Chadi Albitar, 2005 [Alb05])

Koichiro Deguchi, d'une part, Shree Nayar, d'autre part, se sont intéressés à l'analyse (ou décomposition) en composantes principales de l'image (ACP ou transformation de Karhunen-Loève) pour effectuer un recalage par l'apparence visuelle. Le premier a proposé une application

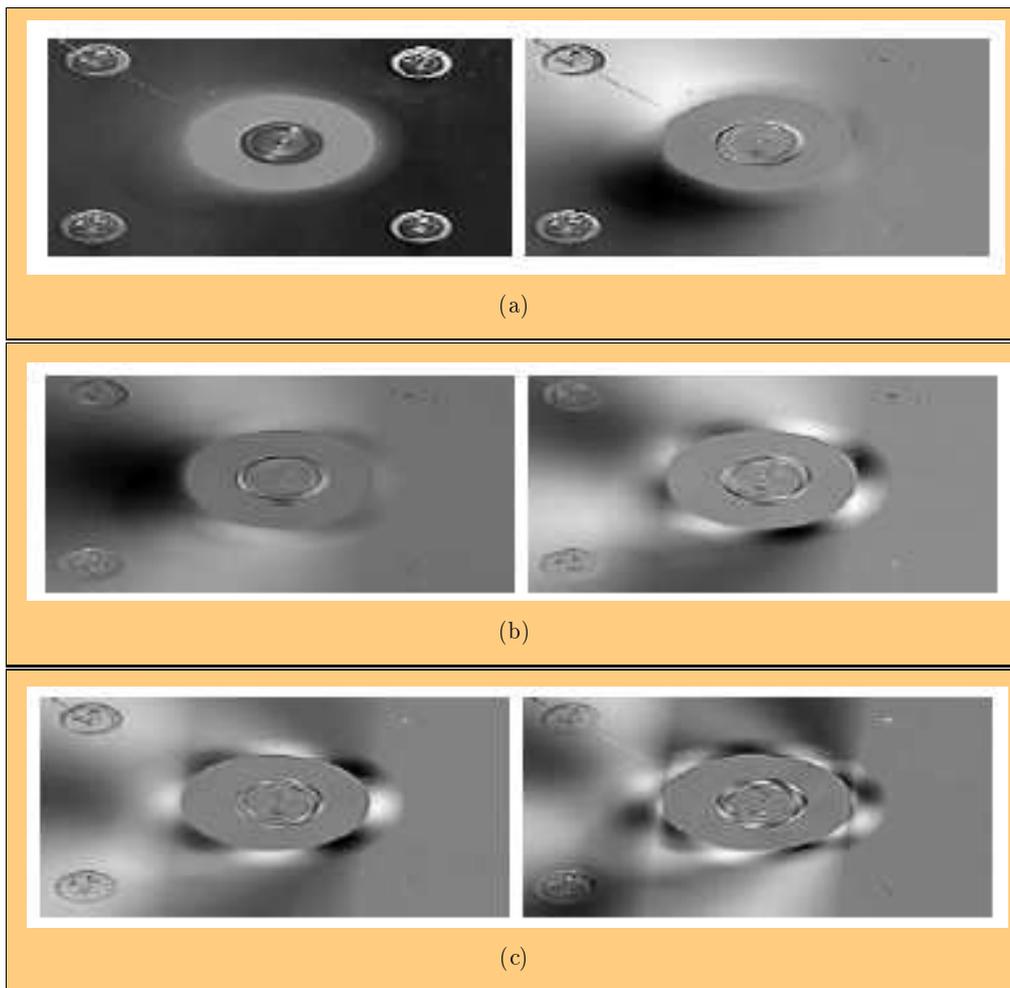


Fig. 2.7: Décomposition en composantes principales de l'image 2.6-b. (a) Image moyenne et première valeur propre. (b) Valeurs propres 2 et 3. (c) Valeurs propres 4 et 5. (Rapport de DEA de Chadi Albitar, 2005 [Alb05])

au guidage de robot dans [DN96] par l'estimation d'une matrice d'interaction et l'élaboration d'un asservissement visuel dans l'"eigenspace" alors que le second dans [NNM96] a proposé une commande par retour visuel basée sur un alignement avec une trajectoire de l'espace cartésien, pour 3 degrés de liberté. L'idée qui était également développée par Schuurman et Capson [SC04] était de trouver un sous-espace optimal à partir des images corrélées. Le calcul de pose revient le plus souvent à associer une vue avec des conditions particulières de prises de vue (attitude, illumination, objet) (voir figures 2.6 et 2.7). Un critère de similarité construit sur une métrique entre images, comme la corrélation ou la somme des carrés des différences (SSD) permet à la fois d'élaborer une stratégie d'apprentissage, puis par la suite d'effectuer une identification [Alb05].

Les désavantages des méthodes de recalage ou d'estimation de pose reposant sur des descripteurs globaux sont identiques pour les approches pré-citées (les moments, le domaine fréquentiel, l'ACP ou la corrélation), à savoir que toutes ces méthodes sont sensibles aux

changements d'illumination, aux occultations, à des variations de l'arrière-plan et considèrent le plus souvent qu'un seul objet d'intérêt est imagé. Des stratégies particulières doivent alors être élaborées pour rendre ces approches moins sensibles à l'une ou l'autre des perturbations. Gregory Hager et Peter Belhumeur ont par exemple proposé un modèle d'illumination en conjonction avec un modèle de transformation géométrique du mouvement dans [HB98] pour aligner des motifs par corrélation. Ales Leonardis et Horst Bischof ont proposé une méthode robuste (RANSAC) pour détecter les variations de l'arrière-plan et les occultations survenant lors d'une reconnaissance par l'apparence visuelle [LB00].

2.2 LES TECHNIQUES D'ASSERVISSEMENTS VISUELS

Un asservissement visuel consiste à commander un dispositif physique ou virtuel, le plus souvent actionné mécaniquement, à partir d'un capteur visuel quelconque, le plus répandu étant la caméra matricielle. Le dispositif d'acquisition/traitement est un système capable d'analyser l'information visuelle numérisée, autrement dit un ordinateur. Souvent, c'est le même système qui est chargé du traitement de l'image et de la commande. Cette commande, qui résulte du retour visuel, est envoyée au dispositif.

Comme pour tout asservissement, les données issues du capteur sont comparées avec la consigne. Pour exprimer cette comparaison, on a l'habitude de distinguer plusieurs espaces de représentation des données [Cha90, HHC96]. C'est cette classification que nous avons retenue pour la suite. Les asservissements visuels peuvent aussi être distingués par le fait que le capteur est soit embarqué sur le dispositif (l'oeil est dans la main et observe la scène - *eye-in-hand*), soit déporté (l'oeil regarde la scène (et éventuellement la main) *eye-to-hand*). Cette distinction est illustrée sur la figure 2.8 ci-dessous.



Fig. 2.8: (A gauche) Configuration eye-to-hand. (A droite) Configuration eye-in-hand.

Le choix de la configuration dépend de la tâche à exécuter ou de l'encombrement du capteur visuel. On peut également distinguer les asservissements visuels par leurs comportements dynamiques. Ainsi, la commande d'un déplacement du dispositif peut être effectuée, par séquence de mouvements discontinus, à l'issue de l'analyse des images (*look and move*). Ce type de commande ne nécessite généralement pas de modéliser le comportement dynamique des éléments entrant en jeu et a été mise en place à l'origine. La commande cinématique [Agi77, SW80] a permis plus tard d'élaborer une commande en vitesse pour obtenir une séquence continue des mouvements du dispositif. Plus récemment, les asservissements visuels rapides ont été mis en oeuvre, grâce aux capacités grandissantes des calculateurs, des capteurs, mais aussi des technologies des actionneurs [HN01, Gan04, Has04].

2.2.1 Les différentes approches

Dans le domaine des asservissements visuels, il existe essentiellement trois approches qui sont communément employées. Loin d'être un état de l'art complet, nous en rappelons brièvement les aspects principaux. Le lecteur intéressé trouvera dans [Cha90, Mar04, Gan04] les formulations mathématiques concernant les méthodes évoquées ici.

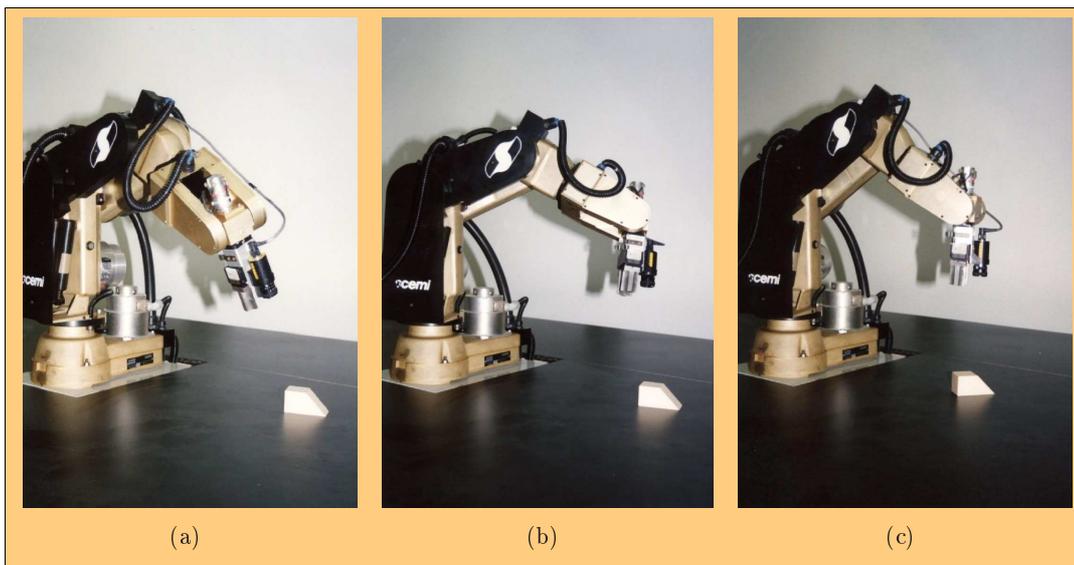


Fig. 2.9: Mise en oeuvre d'un asservissement visuel 3-D à six degrés de liberté sur le robot-manipulateur Scemi 6P01. L'attitude de la pièce polyédrique est estimée en temps réel (80 ms) grâce à la détection des arêtes [Doi94]. (a) Position initiale. (b) Position finale. (c) Position finale, après déplacement de l'objet.

Sanderson et Weiss [SW80] ont introduit tout d'abord une classification concernant les deux approches les plus utilisées : l'approche basée sur la position et l'approche référencée (basée sur l'image). Dans la première, des indices visuels géométriques sont extraits de l'image et, en introduisant un modèle géométrique de l'objet d'intérêt de la scène, la position et l'orientation de cet objet sont estimées à chaque prise de vues [Doi94]. La définition d'une fonction de coût permet alors de formuler l'erreur à minimiser en vue de réaliser la commande qui permettra d'atteindre les positions et orientations souhaitées (figure 2.9). Dans cette approche, des variantes ont été proposées, par exemple, en distinguant le référentiel dans lequel s'exprime les variations 3-D de la position et de l'orientation (référentiel objet, référentiel caméra ou autre) [CM99]. Ces variantes ont un impact sur la trajectoire (et par conséquent sur la visibilité) des indices visuels [TMCG02] et sur le comportement de la loi de commande [Cha98].

Dans la deuxième approche [Wei84, FM99, ECR92, HHC96], l'erreur à minimiser est estimée dans le plan image directement à partir des variations des indices géométriques extraits. Dans le cas d'objet d'intérêt rigide, la relation vision-commande est traduite par le calcul d'une matrice d'interaction [Cha90] qui relie les variations de positions des indices visuels et un torseur cinématique. Cette matrice n'est pas nécessairement re-estimée à chaque prise de vue. En effet, seule l'information de profondeur, souvent fortement entachée de bruit, peut être re-estimée ce qui alors ne nécessite qu'une estimation partielle de la pose [Gan99]. Cette formulation permet également de traduire plus aisément les conditions de stabilité. En suivant cette approche, des travaux concernant l'analyse du mouvement de droites 3-D dans l'image et faisant appel à la représentation de Plücker ont déjà été menés, d'une part dans un contexte général [Mit94], en considérant un alignement projectif de droites [BS01], pour un asservissement visuel classique

[RE87, Cha90], ou encore en découplant l'asservissement d'orientation de l'asservissement de position [AEH00].

La troisième approche est plus récente. Mentionnée sous les termes d'asservissement 2D 1/2 ou hybride, il s'agit d'inclure dans le signal d'erreur à minimiser à la fois les variations de position dans l'image des indices visuels mais également une autre composante qui est souvent issue de la profondeur [Mal98, MC00]. Le découplage entre la commande en position et la commande en orientation s'exprime facilement avec cette approche et sous certaines conditions énoncées par Ezio Malis, la stabilité de ce type d'asservissement a été prouvée. Dans [CH01], un autre type de partitionnement de la commande est proposé où il y a une séparation entre les composantes autour et le long de l'axe z du repère caméra et les composantes autour et le long des axes x et y de ce même repère.

La présence de ces différentes méthodes montrent bien qu'il n'existe pas une seule manière d'aborder les asservissements visuels. Chacune des approches mentionnées ont leurs inconvénients. Ainsi, la commande 3-D souffre d'un manque de précision provenant de l'estimation en ligne de la pose et des imprécisions des valeurs des paramètres intrinsèques au capteur. Comme la correction s'exprime dans l'espace par la compensation des erreurs de position 3-D et d'orientation 3-D, elle ne garantit pas la visibilité des indices visuels dans l'image tout au long de l'asservissement. On peut choisir pour la représentation de la rotation n'importe quelle décomposition [Spo87] (angles d'Euler, roulis tangage lacet, angle/axe, ...), mais il s'avère que certains choix sont plus judicieux que d'autres. Par exemple, en exprimant la position de la cible dans le repère courant de la caméra et non dans un repère lié à l'environnement, on évite d'introduire dans le calcul de la pose le modèle géométrique du robot et donc le bruit de quantification dû aux codeurs incrémentaux [Gan04]. De plus, en faisant ce choix, on garantit une trajectoire rectiligne de l'origine du repère lié à la cible dans l'image de la caméra lorsqu'on utilise une commande cinématique [MGK96]. Pour la décomposition de la rotation, on préférera une représentation qui ne soit pas singulière dans la zone de travail. La décomposition angle/axe est souvent choisie car sa seule singularité est pour une rotation de 2π [Mal98]. A l'opposée, l'asservissement visuel référencé image permet de garantir cette visibilité (ou de l'exprimer dans une stratégie de commande planifiée [SM06]). Moins sensibles aux erreurs de calibration (qui modifient la trajectoire dans le plan image), elle est plus précise que l'approche précédente. Une loi de commande cinématique est exponentiellement stable avec cette approche. Toutefois, la trajectoire dans l'espace 3-D n'est pas rectiligne, ce qui peut poser des problèmes d'évitement d'obstacles, voire aboutir à une trajectoire impossible à réaliser (problème d'avance/retrait). C'est pour contourner ce genre de problème qu'ont été imaginées les méthodes hybrides.

En ce qui concerne l'approche hybride, l'estimation en ligne d'une homographie à partir de scènes planes ou du choix d'un plan virtuel peut influencer les performances de l'asservissement [MC00]. Elle nécessite une mise en correspondance qui peut s'avérer délicate quand la situation courante et la situation désirée sont éloignées. Néanmoins, cette approche semble être un excellent compromis des deux premières, et permet d'appréhender les environnements difficilement modélisables.

2.2.2 Formalisme de la fonction de tâche

Dans [Cha90], une classification des tâches d'asservissement est réalisée, en fonction du nombre et du type de degrés de liberté contraints et l'asservissement visuel est exprimé à l'aide du formalisme de la fonction de tâches [SBE91], un outil mathématique bien commode pour établir l'influence mutuelle de plusieurs tâches, quel que soit le capteur ou la stratégie de commande associés. L'intégration de l'asservissement visuel dans l'approche générale de la fonction de tâche permet de résoudre de manière efficace et élégante le problème de tâche hybride, combinant plusieurs objectifs à atteindre, comme par exemple des opérations de suivi de trajectoires pour des applications d'inspection, ou d'évitement des butées et singularités du robot, d'évitement d'obstacles [Mar04]. Dans ce contexte, la tâche visuelle peut alors être considérée comme principale et prioritaire. Les autres objectifs s'expriment sous la forme d'une tâche secondaire. Cette tâche secondaire peut être définie, par exemple, comme le gradient d'une fonction à minimiser sous la contrainte que la tâche principale soit réalisée (voir aussi dans [Doi94, Man06]). Ceci est possible car, si la pseudo-inverse fournit une commande de norme minimale permettant de réguler l'erreur, il existe en fait une infinité de solutions permettant d'assurer cette minimisation. Toutes les autres solutions se situent dans le noyau du jacobien de la tâche. Si la tâche secondaire est projetée sur ce noyau, elle n'aura alors aucun effet sur la tâche principale.

2.2.3 De l'intelligence artificielle à la vision robotique dynamique

Il n'est pas nécessaire au lecteur de s'attarder sur cette partie pour la compréhension de la suite du manuscrit. Comme il est rare de disposer dans les publications scientifiques de suffisamment d'espace (de pages !) pour s'exprimer, j'ai simplement souhaité, d'une part avec cette courte parenthèse historique, rendre hommage à ceux qui ont inspiré et contribué à l'évolution de la vision robotique par des concepts novateurs ou des réalisations remarquables en rappelant quelques faits qui ont marqué cette discipline et ont permis son émergence. D'autre part, j'estime que cela est une manière simple d'appréhender les enjeux à venir qui sont en lien de près ou de loin avec ce domaine. Les travaux mentionnés ici correspondent à un choix personnel des nombreux acteurs de la recherche en vision par ordinateur, en intelligence artificielle, en robotique et en automatique. J'ai donc conscience que si cette section avait été rédigée par une tierce personne, elle aurait vraisemblablement mis en relief d'autres événements.

L'utilisation de la vision en robotique a été motivée, semble-t-il, par la volonté d'étendre la flexibilité et l'autonomie, et par conséquent de couvrir un champ d'applications plus large de la robotique. La perception visuelle de son environnement permet donc à un robot de réagir, si possible de manière autonome, aux variations notables de celui-ci. Pour appréhender un environnement partiellement ou totalement inconnu, d'autres types de capteurs extéroceptifs peuvent également être employés avec des motivations identiques. Alors que des capteurs de proximité, tactiles ou d'efforts tiennent un rôle important dans les capacités d'un système robotique à percevoir son environnement par une mesure essentiellement locale voire ponctuelle, la perception visuelle fournit, sans contact, des informations plus globales et plus diversifiées de l'environnement et de ses changements. Cette richesse d'informations distingue la vision des autres mesures, car d'une part les capteurs visuels, les équipements et les traitements associés sont bien plus sophistiqués, et d'autre part une fraction seulement de l'information capturée

est exploitée.

"La vision robotique traite des procédés permettant d'extraire, de caractériser et d'interpréter l'information contenue dans les images de l'environnement tridimensionnel" [FGL87].

Cette définition assez ancienne a été énoncée à une époque où la vision par ordinateur et la vision robotique n'étaient pas clairement démarquées l'une de l'autre [Hor86, HS92]. Cette définition ne montre pas non plus les liens existants entre la vision et la robotique, car elle ne prend pas en compte les préoccupations principales de la robotique. L'autonomie, par exemple, est une des caractéristiques d'un robot qui nécessite que les procédés évoqués précédemment soient automatisés. Alors que de nombreuses techniques développées en reconnaissance des formes utilisent l'image comme support expérimental par la recherche d'une information maximale, un robot n'a besoin de voir que ce qui lui est nécessaire pour accomplir sa tâche. Par ailleurs, les contraintes temporelles inhérentes à la commande des déplacements d'un robot dans un environnement dynamique requièrent le plus souvent des processus informatiques déterministes, quelle que soit la complexité de la scène perçue. L'analyse d'image en temps réel est alors impérative car c'est cette analyse qui va permettre au robot de progresser vers la réalisation de la tâche qu'on lui a imposée. De plus, la génération automatique des mouvements par les techniques d'asservissement visuel fait appel à l'extraction de l'information géométrique et de ses variations contenues dans une séquence d'images. On peut alors considérer deux points de vues diamétralement opposés pour situer la vision robotique autant par le contexte que par les objectifs à atteindre. Les deux points de vues peuvent être énoncés ainsi [CR80] :

- dès la conception d'un système robotique, on est en mesure de connaître quelles informations visuelles on souhaite extraire pour l'exécution d'un ensemble de tâches. Même s'il existe plusieurs méthodes d'accès à cette information, dans tous les cas celle-ci sera particularisée, amenant à l'utilisation d'un traitement spécialisé à l'aide d'une représentation *optimale* vis-à-vis du but poursuivi.
- en s'inspirant des performances du système de vision de l'homme, on peut s'intéresser à l'obtention d'une représentation *universelle* de l'information contenue dans les images. L'utilisation de cette représentation par un robot ne sera alors qu'une application particulière.

Ne considérer qu'un seul de ses points de vue est trop réducteur et ne reflèterait pas la diversité des thématiques étudiées aussi bien robotique qu'en vision par ordinateur. Les deux points de vues doivent être partagés, et même si à l'aube de la vision robotique, le second point de vue a dominé les esprits, la recherche d'une représentation universelle a été abandonnée. En effet, les premiers projets pour doter les robots d'une vision artificielle, vers 1966, ont révélé que la problématique soulevée dans ce domaine avait été sous-estimée. Il est notamment admis que cette discipline est complexe, cependant dans les années 60, presque personne ne réalisait l'ampleur des difficultés d'élaborer un système de vision permettant à un robot d'interpréter tout ce qui l'entoure. La conséquence fut l'échec de plusieurs tentatives pour doter la machine d'un sens de vision. Selon David Marr, spécialiste de la vision humaine, les déconvenues des premiers chercheurs viennent du fait d'avoir voulu passer directement de l'énoncé d'un

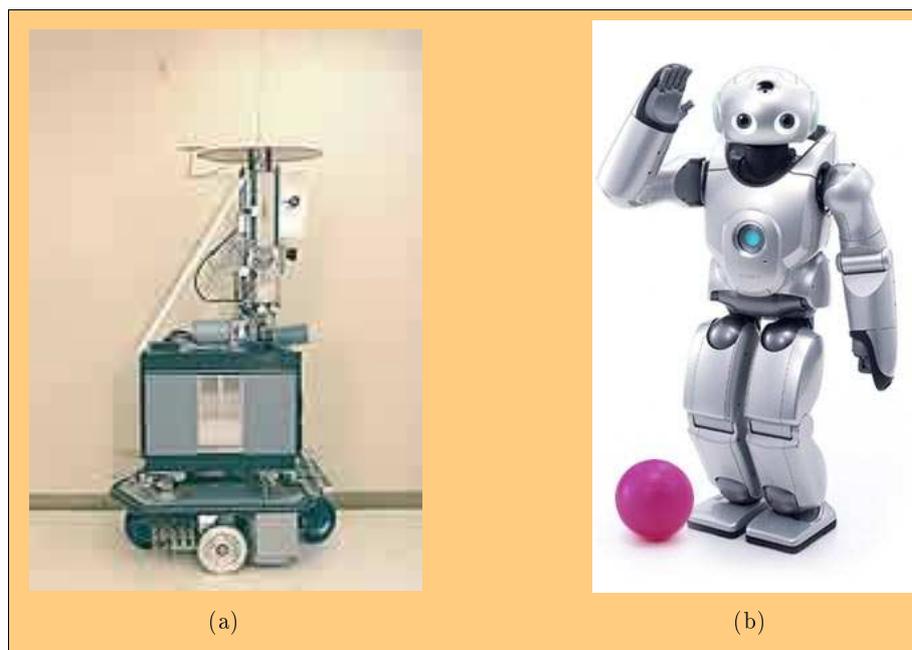


Fig. 2.10: (a) Le robot expérimental SHAKEY du Stanford Research Institut (SRI) est le premier robot mobile muni d'une caméra TV, d'un détecteur de choc et de proximités (1970). La logique embarquée lui permettait de modéliser son environnement en quelques minutes. (b) QRIO, robot humanoïde japonais de SONY corp. est équipé en 2004 d'une vision stéréoscopique, lui permettant, en temps réel vidéo, de prendre en compte et d'éviter des obstacles, de monter et descendre des escaliers,....

problème à sa solution, sans disposer d'une connaissance scientifique préalable leur permettant de concevoir des algorithmes efficaces [Mar82]. La théorie que Marr a développé a beaucoup influencé la pensée des chercheurs en vision par ordinateur dans les années 80.

La première génération de robots dit "intelligents" car munis de plusieurs capteurs extéroceptifs et capables de réagir aux modifications de l'environnement est apparue vers 1970. A l'aide des informations visuelles, ils pouvaient reconnaître des objets de forme simple dans un environnement uniforme et invariant. Le robot SHAKEY (figure 2.10-a), élaboré par l'Institut de Recherche de Stanford (SRI) [RN67, Nil69] en novembre 1967 par l'équipe de Charles Rosen, fut le premier robot mobile muni d'une caméra TV (et muni également d'un détecteur de choc et de plusieurs proximités) pour percevoir et modéliser un environnement intérieur, à l'aide d'un programme informatique. Dans le même institut, l'utilisation conjointe de la vision et de la robotique a servi initialement à réaliser des démonstrateurs pour valider des algorithmes issus de l'intelligence artificielle [FPB⁺71]. Peu de temps après, le projet japonais ETL (Electrotechnical Laboratory) a donné le jour à un robot d'assemblage automatique exploitant exclusivement les informations visuelles provenant d'une caméra montée sur lui (*eye-in-hand*) pour se recalibrer par rapport à une scène constituée de cubes [Eji72].

En 1973, Y. Shirai et H. Inoue ont développé une méthode de saisie d'un objet cubique à l'aide d'un robot et d'une caméra fixe [SI73]. Ils ont montré l'intérêt d'une boucle de retour (*visual feedback*) à la fois pour corriger la position du robot (en commandant directement les

variables articulaires) mais également pour en améliorer la précision. Un schéma itératif en deux étapes était apparu : "static look and move". C'est en 1977 qu'apparaît, semble-t-il pour la première fois, le terme d'asservissement visuel (*visual servoing*) avec les travaux de Agin [Agi77, Agi79] sur l'assemblage de pièces mécaniques. Il s'agit en fait de ce que l'on nomme le "dynamic look and move" qui permet de considérer des objets de la scène en mouvement. Cela est rendu possible par la prise en compte des informations visuelles dès qu'elles sont disponibles, les deux étapes "look" and "move" étant alors fusionnées. Dans les travaux de Hill et Park en 1979, la position de l'organe terminal d'un robot est commandée en fonction de l'attitude relative entre le capteur visuel et la cible : c'est l'asservissement visuel basé sur la position 3-D (*Position-Based Visual Servoing*) [HP79, Doi94, MDGD97]. Sans doute inspirés par cela, Sanderson et Weiss [SW80, SW82, WSN87] ont étudié différents schéma de commande pour les asservissements visuels, se distinguant par les grandeurs à asservir : variables articulaires (*joint-level*), variables opérationnelles (espace de la tâche ou *recognition-level*) ou positions des indices visuels dans l'image (*feature-level*). Ils ont proposé pour ce dernier cas, un schéma de commande où une matrice jacobienne relie les variations dans l'image des informations visuelles et les mouvements de la caméra. Cette approche, nommée asservissement visuel 2-D (*Image-Based Visual Servoing*) permet de choisir les informations visuelles capables de représenter une tâche et, d'autre part, d'élaborer des lois de commande à partir de celles-ci.

Le premier ouvrage consacré à la vision par ordinateur est celui de Duda et Hart, datant de 1973 [DH73]. A une première partie consacrée à la reconnaissance des formes fait suite une deuxième partie qui introduit les bases théoriques d'une approche géométrique de l'interprétation d'une image par sa relation avec la scène. Les travaux de Marr et Poggio sur la vision humaine et sur le processus stéréoscopique en particulier ont permis de mieux appréhender la fusion des données entre images. Ils ont proposé en 1976 de subdiviser le problème de la reconstruction 3-D à partir de deux points de vues en deux phases distinctes : la mise en correspondance un à un des indices visuels (appariement) d'une part et la détermination de la disparité d'un couple d'indices appariés d'autre part [MP76]. Cinq ans plus tard, Christopher Longuet-Higgins fournissait un algorithme de reconstruction 3-D à partir de 8 points images appariés (droites et gauches) [LH81]. La même année, Fischler et Bolles [FB81] ont proposé une méthode devenue populaire lorsqu'ils ont introduit la méthode RANSAC (RANDOM SAMPLE CONSENSUS) pour la détection des valeurs aberrantes dans les données initiales. Appliquée à la vision, cette méthode permet simultanément d'apparier des sous-ensembles d'indices visuels sélectionnés aléatoirement (image/modèle ou image/image) en suivant le schéma classique "test d'hypothèses, puis vérification" et de déterminer une transformation géométrique. De nombreuses améliorations de cette méthode ont été proposées par la suite. Cette approche est encore très employée de nos jours, par exemple dans le calcul d'homographies [HZ00].

Le formalisme de fonction de tâches en général [SBE91], puis pour les tâches visuelles en particulier [Cha90], les asservissements visuels sans modèle du robot [MGM02], ou de la scène [MB04], l'estimation de l'environnement en ligne (*SLAM - Simultaneous Localisation and Mapping* [SSC90, Bur04], *SfM - Structure from Motion* [TK92, ST96, MSKS03]) sont des exemples d'avancées importantes qui, à travers la vision monoculaire, stéréo (figure 2.10-b) ou panoramique par exemple, contribuent à renforcer les liens privilégiés qui unissent la robotique et la vision.

2.3 LE SUIVI VISUEL

Il existe une bibliographie très riche sur le suivi visuel. De nombreuses méthodes de la littérature concernent le suivi des objets précis, 2-D ou 3-D, en temps réel. A l'opposée des méthodes qui traitent une séquence vidéo dans son ensemble avec des stratégies non causales, le suivi en temps réel requiert une estimation de paramètres basée sur un historique (à horizon plus ou moins étendu, éventuellement Markovien [KH98]) et une méthodologie de la prédiction de ces mêmes paramètres ou d'une partie seulement de ces paramètres. Le *tracking* 2-D a pour objectif de suivre la projection dans les images d'objets ou de parties d'objets dont les déplacements 3-D induisent un mouvement ou une déformation qui peut être modélisé sous la forme d'une transformation 2-D. Un modèle adaptatif est alors requis pour rendre compte des changements d'apparence dus aux effets perspectifs ou à une déformation 3-D. Ce modèle permet de formaliser la position de l'objet d'intérêt sous la forme d'un centre de gravité, d'une aire ou d'une transformation affine [WADP95, CRM03, JFEM03]. De manière alternative, des modèles plus sophistiqués à base de splines [IB98], de masques déformables [YHC92], de mailles déformables [SI98] ou encore de multi-articulations [CR99] peuvent être employés. Cependant, aucune de ces méthodes n'implique l'estimation de la position actuelle dans l'espace.

Comme cela a déjà été mentionné dans l'introduction générale, nous évoquons ici surtout les techniques de suivi qui s'intéressent à l'estimation paramétrique d'un mouvement dans la scène observée. Le suivi visuel 3-D a pour objectif d'estimer les degrés de liberté qui définissent la position et l'orientation du capteur visuel par rapport à la scène, ou les déplacements de l'objet d'intérêt vis-à-vis du capteur visuel. Malgré tout, au sein d'un système de guidage par vision, on est également amené à employer des méthodes de suivi d'entités spécifiques dans l'image telles que celles référencées ci-dessus pour diminuer le volume de données à traiter, d'une part, durant la phase de segmentation et d'autre part pour réduire considérablement l'espace de recherche des appariements de primitives entre image, un traitement sur toutes les données vidéo étant généralement trop prohibitif en termes de temps de calculs. On trouve donc fréquemment des procédures de suivi de points d'intérêt [ST94, PKK93], de droites [DF90], de contours [Bou89], de motifs [HB98, JD02, MB04] au sein d'un système de guidage par vision [MC04].

Les méthodes de suivi peuvent être classées selon qu'il s'agit de suivre des entités dans la scène ou dans les images, ces dernières pouvant être encore distinguées selon que l'on s'intéresse à des données éparées représentées par des descripteurs géométriques locaux (points, droites,...) ou bien des données denses (régions, motifs) plutôt représentées par des descriptions statistiques globales. Cette classification correspond également à considérer un volume croissant de données : d'une description symbolique vers une représentation iconique en requérant le plus souvent un effort de segmentation inversement proportionnel. Le cas des contours apparents est particulier dans cette classification. On peut en effet considérer tout autant une représentation locale des contours (approximation polygonale [Shi78, Low87, IK88, WB91] ou conique [Sam82, Low87, RW90, HS92, WD95, Vin01],...) qu'une représentation globale (snake [KWT87, BI98], descripteurs de Fourier [PF77, KG82], moments [Hu62, FS93],...) plutôt adaptée aux contours fermés.

2.3.1 Suivi de primitives géométriques 3-D

De plus en plus d'applications requièrent un module spécifique de *tracking*. Avec les possibilités de plus en plus répandues d'accéder à des services en ligne via Internet, des outils de la réalité augmentée sont développés ici et là pour faciliter la compréhension de telles ou telles techniques de maintenance, pour promouvoir un produit ou un process, dans un but pédagogique (apprentissage, entraînement, expertise) ou commercial (navigation, visite virtuelle). De par son aspect non-invasif, le tracking visuel 3-D est un moyen technologique approprié pour répondre, de plus en plus finement, à des programmes de visualisation ou d'interfaçage homme-machine. Il peut être comparé à d'autres technologies ; ainsi, les traqueurs mécaniques sont doués d'une grande précision en général, mais dans un espace de travail le plus souvent très limité. Dans des applications médicales, on trouve des traqueurs optiques ou magnétiques. Si les premiers sont assez fiables et utilisent, entre autres, la stéréovision par lumière infra-rouge (Systèmes Polaris et Optotrack de NDI), les derniers sont sensibles à la présence d'objets métalliques dans le volume de mesures comme les instruments de chirurgie et ne sont pas adaptés à certains dispositifs d'imagerie médicale. Les traqueurs à ultra-sons, bien que des progrès significatifs aient été réalisés ces dernières années, souffrent encore d'un rapport signal sur bruit peu élevé, et peuvent voir leur précision diminuer durant un examen prolongé, car ils sont sensibles aux variations de la température. Toutefois, l'évolution extrêmement rapide des dispositifs à ultra-sons pourraient bien remettre en question les arguments avancés ici dans quelques années.

On voit donc qu'il y a un réel intérêt, grâce aux applications potentielles et aux capacités intrinsèques, à développer des solutions faisant appel au suivi visuel 3-D, pour qui veut bien se donner la peine de développer des algorithmes suffisamment robustes, avec l'apport éventuellement de marqueurs artificiels.

Une étude très complète sur ce sujet a été publiée par Vincent Lepetit et Pascal Fua dans [LF05]. Le problème d'estimation de la pose est intimement lié au suivi tridimensionnel et on trouvera dans les travaux de Markus Vincze *et al.* [VSGA05], de Luca Vachetti [VLF04], de Muriel Pressigout [Pre06] et de Danika Kragic [KC03] entre autres, de nombreuses expériences de suivi 3-D qui, la plupart du temps, sont basées sur plusieurs types de primitives. Plus de détails pourront être trouvés dans ces publications et également dans ma participation à un ouvrage [Doi07].

2.3.2 Suivi de caractéristiques apparentes

Suivi de contours

Le suivi de contours a été étudié par de nombreux auteurs. Renfermant une information géométrique importante, il a été envisagé plusieurs applications au suivi [Bou89, IB98] et aux asservissements visuels [CMCK93, CMC00, CC00, GM02]. Dans [CPK93] par exemple, la silhouette des objets d'intérêt est capturée. Les contours actifs servent alors à représenter l'évolution de celle-ci au cours du temps par le déplacement de points de contrôle, à partir desquels un asservissement visuel est finalement mis en oeuvre.

Dans un article souvent mentionné, Patrick Bouthemy [Bou89] a développé une stratégie locale pour la recherche des contours en mouvement (*moving edge*). La recherche s'effectue par déplacement d'un masque de convolution dans la direction perpendiculaire au contour. La réponse maximale de la convolution avec l'image nouvellement acquise permet d'actua-

liser la localisation du contour. Nous avons mis en oeuvre cette technique pour le suivi des droites apparentes d'un instrument de chirurgie dans [Nag05], qui est illustré sur la figure 2.11.

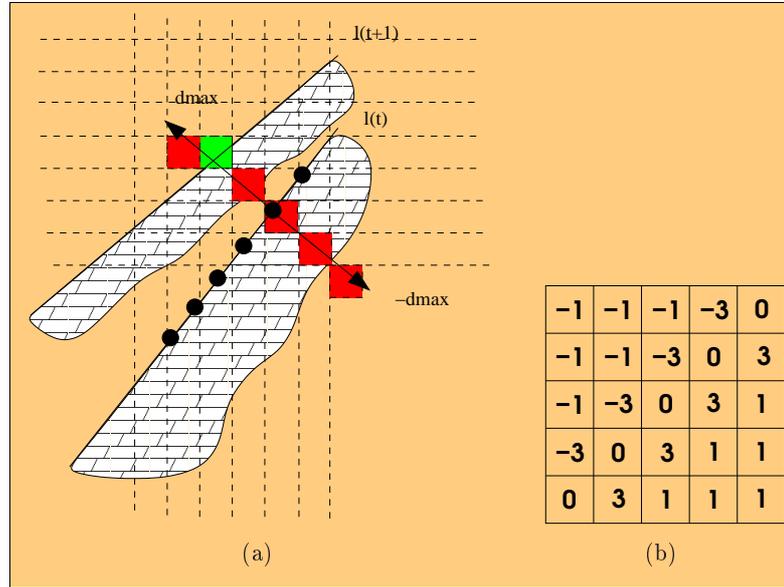


Fig. 2.11: (a) Les droites image I aux instants t et $t+1$. La droite $I(t)$ est échantillonnée. Pour chaque échantillon le nouveau contour est recherché dans la direction normale à $I(t)$. (b) Le masque 5×5 utilisé pour les contours orientés à 45° . Les pixels rouges indiquent les pixels testés, le pixel vert correspond à la nouvelle position trouvée.

Le suivi de contours peut être vu comme un processus dynamique, qui, contrairement à l'approche de Bouthemy, peut être modélisé. Michael Isard et Andrew Blake ont proposé il y a une dizaine d'année de modéliser l'évolution temporelle des contours par application du filtre à particules [IB98]. La modélisation de l'évolution peut également être déterministe et paramétrique. A cette fin, on peut rappeler la notion de *contour actif* qui peut être représentée sous la forme d'une courbe, fermée ou non, et en évolution. A partir d'une position initiale, le contour va évoluer, soit dans une même image (segmentation), soit dans des images successives d'une séquence (suivi spatio-temporel). Une paramétrisation de la courbe à l'aide d'une variable scalaire $s = s(t)$ pourra être exprimée génériquement par

$$\begin{aligned} v : [0, 1] &\rightarrow \mathbb{R}^2 \\ s &\rightarrow v(s) = [x(s), y(s)] \end{aligned} \quad (2.8)$$

La courbe est déformée de manière itérative afin de minimiser une fonctionnelle d'énergie. Cette fonction peut être définie dans le domaine continu comme

$$E(v) = \int_0^1 [\alpha_{int} E_{int}(v(s)) + \alpha_{ext}(v(s))] ds, \quad (2.9)$$

où E_{int} et E_{ext} représentent les énergies internes et externes. Les coefficients α_{int} et α_{ext} sont utilisés pour donner plus ou moins d'influence à l'une ou l'autre des énergies, composées elles-mêmes de différentes contributions énergétiques particulières. Les énergies sont liées à

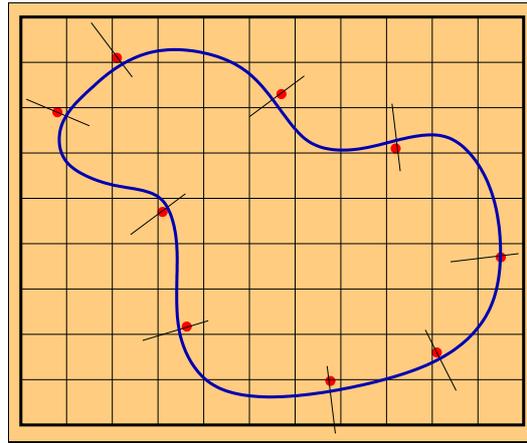


Fig. 2.12: Lignes perpendiculaires (en noir) au contour actif pour la recherche locale des minima d'énergie. Les points de contrôle (en rouge) sont échantillonnés à équi-distance initialement. En fonction des courbures rencontrées, le nombre et la répartition de ces points peut être modifié durant le processus d'évolution.

une force (module du gradient local de l'image, GVF, force ballon, continuité, courbure), dont le comportement permettra de comprendre l'intérêt et l'influence de l'énergie dans l'évolution du contour actif. L'évolution du contour actif est déterminée par son équation de mouvement, dérivée à partir de l'équation 2.9 et les forces sont appliquées en des points de contrôle, qui à chaque itération, peuvent se déplacer dans un voisinage donné (typiquement, dans une fenêtre 7x7 centrée sur chaque point de contrôle).

Si les termes énergétiques internes contrôlent la bonne évolution du contour en représentant ces propriétés physiques, ils ne renferment aucune information sur le contour lui-même. La force associée à l'énergie de continuité influe sur le rayon de courbure du contour en conduisant les points du contour à se positionner de manière à être équidistants, comme l'illustre la figure 2.12. La forme d'un contour fermé tend alors vers un cercle. L'énergie de courbure a pour but d'éviter que le contour contienne des points isolés qui ne seraient pas cohérents avec la forme globale et régulière supposée du contour.

Plusieurs approches ont été proposées pour l'implémentation du modèle de contour actif : le calcul variationnel [KWT87], la programmation dynamique [AWJ90], ou encore l'algorithme greedy [WS92]. Dans une étude comparative [DN95], il a été montré que l'algorithme greedy est bien plus rapide que les méthodes par calcul variationnel ou programmation dynamique.

Les contours actifs sont particulièrement intéressants quand on ne peut pas modéliser la forme des frontières de l'objet que l'on cherche à suivre. Utilisés en segmentation d'image et en suivi d'objet [BI98], différents modèles de contours actifs ont été proposés dans la littérature. Ainsi, on doit à Kass et al. [KWT87] les travaux originaux dans ce domaine. Le modèle utilisé a cependant montré différentes limites, liées notamment à l'initialisation, au paramétrage, et à l'impossibilité de changer la topologie du contour actif. Certains auteurs ont donc proposé d'autres modèles. Parmi eux, les contours géodésiques [Par00] (de la famille des approches *level set* - courbes de niveaux - [Set96, Set99]), préconisés actuellement dans la littérature, qui permettent de gérer des topologies quelconques, mais se traduisent par des temps de calcul importants [ABF99].

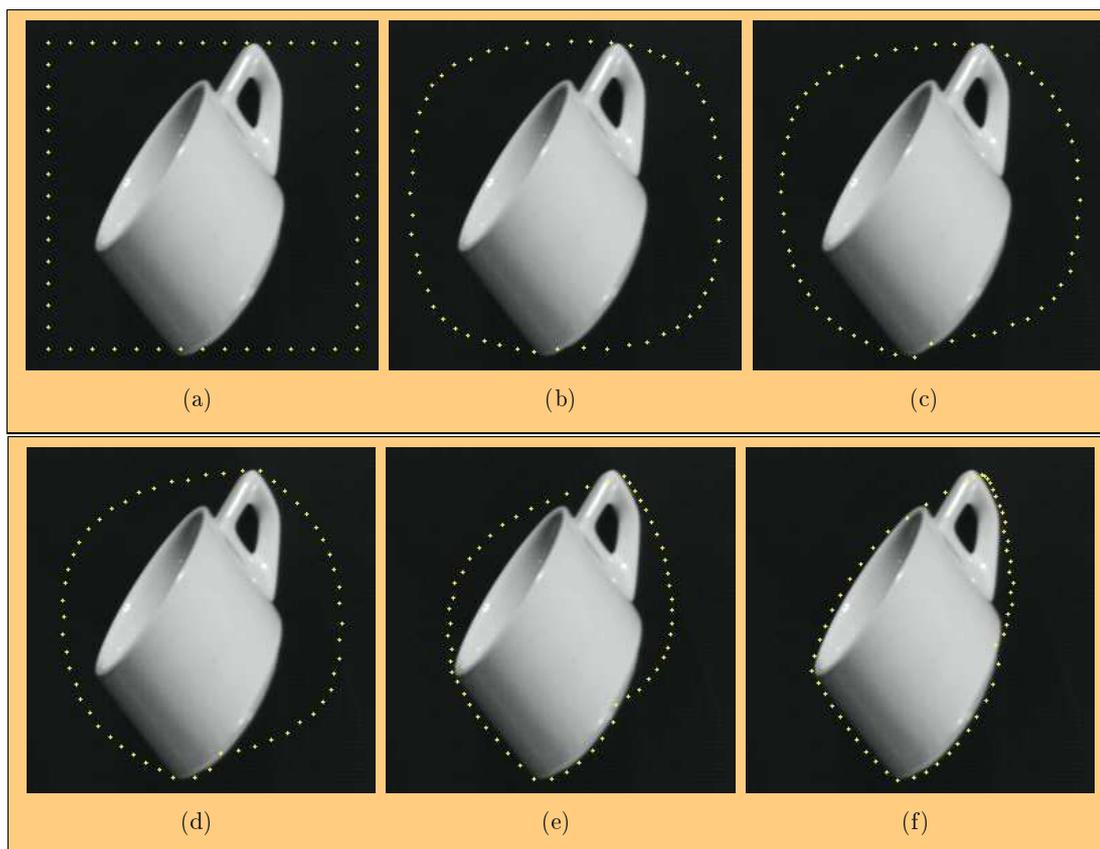


Fig. 2.13: Evolution d'un contour actif pour extraire le contour externe de l'image d'une tasse de café. Ces résultats ont été obtenus par l'algorithme greedy et un filtrage de Canny (fenêtre 7×7). Noter que l'image n'est constituée que d'un objet d'intérêt et que le fond est uniforme. Un résultat similaire aurait pu être obtenu par la recherche du plus petit polygone contenant la silhouette de l'objet dans une image binaire.

Le suivi des objets en temps réel à l'aide des contours actifs présente donc des limitations importantes. Ainsi, afin d'obtenir un résultat correct de convergence, il est souvent nécessaire d'initialiser le contour actif proche du contour recherché. De plus, le modèle du contour actif comporte de nombreux paramètres, souvent difficiles à régler. Enfin, ils restent difficiles à exécuter en temps réel, toutefois des travaux récents ont contribué à améliorer la rapidité d'exécution. Par exemple, dans [Lef02], une technique pour effectuer un suivi d'objet en temps réel est proposé avec la prise en compte d'une scission éventuelle du contour.

En pratique, il est nécessaire, la plupart du temps, de faire une redistribution des points de contrôle en opérant un nouvel échantillonnage le long de la courbe, afin d'éviter qu'après quelques images dans la séquence ou après un certain nombre d'itérations, ces points deviennent arbitrairement groupés en amas ou séparés sur la courbe. D'autre part, pour les contours présentant des coins ou des zones de courbure élevée, il est nécessaire d'augmenter le nombre de points de contrôle sur ces zones, afin de les localiser plus finement. Un échantillonnage adaptatif semble alors la réponse la plus adéquate pour résoudre ces problèmes.

Quand l'objet à suivre a un modèle connu, en particulier, s'il a une forme structurée, comme un rectangle ou un polygone, des contraintes peuvent être imposées au modèle du contour actif, afin de le déformer d'une façon convenue [BI98, FCH06]. Des relations entre les points de contrôle peuvent alors être incorporées sous la forme de termes d'énergie interne.

Suivi de régions/motifs

Les méthodes de suivi visuel de régions sont utiles pour appréhender les environnements non maîtrisés. Elles requièrent en général que peu de pré-traitements des images par rapport aux méthodes précédentes. En contrepartie, il n'est pas facile d'extraire certains paramètres du mouvement, ceux qui sont fortement liés aux effets perspectifs (distance et rotations hors plans) et sont de plus, sans l'apport de traitements particuliers, sensibles aux occultations. Quand la scène est complexe (environnement fortement texturé ou contours difficilement détectables) d'autres méthodes sont à envisager. Une solution consiste à considérer directement l'intensité des pixels dans les images et à effectuer un alignement, plutôt qu'une mise en correspondance, sur une partie de l'image (*template-based matching*) sans passer par une phase d'extraction de primitives. L'objectif de tels algorithmes est d'estimer un ensemble de paramètres qui décrivent au mieux la transformation ou le mouvement de la partie de l'image considérée en optimisant un certain critère de corrélation. Il existe des techniques de minimisation capables de résoudre ce problème en prenant en compte des modèles de transformation relativement complexes (transformations affines ou homographiques).

L'algorithme de Bruce Lucas et Takeo Kanade a été conçu à l'origine pour calculer le flot optique [LK81, BM04] mais il est également approprié pour recalculer un masque bidimensionnel dans une image quand on lui fait subir des déformations. Il n'est donc pas exclusivement restreint à des primitives locales ou à la segmentation par le mouvement mais peut être employé dans un processus de suivi de régions, en considérant par exemple un motif d'intérêt à suivre. Cette méthode, très populaire, convient bien au traitement d'objet complexe, difficilement modélisable à l'aide de primitives locales. Dans [HB98], Gregory Hager et Peter Belhumeur ont proposé une approche basée sur la corrélation (exprimée par la somme des différences des carrés de l'intensité - *SSD*) en considérant la variation des paramètres du modèle de mouvement comme une fonction linéaire des différences d'intensité dans la séquence d'images. Ils définissent une matrice jacobienne qui relie les variations des paramètres de la transformation aux variations de l'intensité lumineuse. Bien qu'il requiert de nombreux calculs, il a été montré dans [HB98] que cet algorithme pouvait être implémenté de manière efficace. Depuis lors, il a été amélioré par d'autres auteurs et appliqué au suivi 3-D [JD02] et à la commande de robot [PKK93, MB04, BM06].

Dans le souci de relier les déformations d'un motif dans l'image à un déplacement dans l'espace, les travaux de Tom Drummond et Roberto Cipolla [DC00] et ceux menés par Selim Benhimane et Ezio Malis [MB04] ont apporté une contribution importante dans ce domaine. Suivant le schéma proposé dans [HB98], ils ont, grâce à la conjonction de l'expression des transformations homographiques (entre motifs d'images) dans une base de l'espace de Lie et à une minimisation numérique judicieuse (l'algorithme ESM - *Efficient Second-Order Minimization*), permis d'obtenir l'ensemble des paramètres d'un mouvement 3-D d'un capteur visuel lors de l'obser-

vation de scènes planes, sans l'apport de modèle géométrique de celle-ci. Par application de l'algorithme ESM, un développement au second ordre des composantes de l'homographie, est réalisé sans calculer explicitement une matrice Hessienne, ce qui réduit significativement les calculs en ligne. Nous reviendrons, dans le chapitre suivant, sur l'application de cette technique à la commande de robot sous contraintes de mouvement [BDZM07].

Emanuele Trucco et Alessandro Verri notent dans [TV98], que lorsque les intensités des motifs sont normalisées, la distance euclidienne dans l'espace engendré par les vecteurs propres (*eigenspace*) est équivalent à la corrélation d'image, et que par conséquent minimiser cette distance revient à maximiser la corrélation. On peut donc affirmer qu'à travers la décomposition d'une image par la transformation de Karhunen-Loève, l'apparence visuelle est équivalente (aux composantes non principales près) à la corrélation. C'est pourquoi, on retrouve également le suivi de motifs traité sous l'angle de l'apparence, comme dans les travaux de Frédéric Jurie et Michel Dhome [JD02] où une phase d'apprentissage est nécessaire pour estimer deux matrices d'interaction. La première matrice lie les variations d'intensité lumineuse du motif à son déplacement fronto-parallèle (qui ne change pas son apparence) alors que la seconde matrice relie les variations d'apparence suite à un changement d'orientation.

D'autres approches du suivi basées sur des noyaux, comme le Mean-Shift [Che95, CRM03] ou sur l'estimation du champ des vitesses ont été également proposées. Un algorithme de suivi de régions a été proposé par François Meyer et Patrick Bouthemy [MB94] qui permet d'obtenir les trajectoires complètes d'objets en mouvement. Les auteurs utilisent des modèles affines de champs de vitesses qui autorise, d'une part, l'obtention d'une modélisation robuste du processus de suivi, et d'autre part l'accès à la trajectoire dense des projections de l'objet dans l'image. Il en résulte alors la possibilité d'une interprétation du mouvement tridimensionnel à partir d'une séquence d'images monoculaires. Ferrari *al.* [FTG01] a proposé une méthode de suivi de régions issues d'une décomposition de surfaces par plans. La méthode de suivi, qui associe texture et contours des régions, peut retrouver également les paramètres d'un modèle affine de transformations entre images lors du suivi simultané de plusieurs régions.

2.4 LA SEGMENTATION D'IMAGES EN TEMPS RÉEL

Le but de la segmentation d'image est de subdiviser l'image en éléments qui ont une forte corrélation avec les objets d'intérêts contenus dans la scène 3-D imagée. Il s'agit généralement de classer ces éléments entre eux, d'en fusionner certains ou d'en éliminer d'autres. Il n'est pas du tout dans mon intention d'inventorier les (trop nombreux peut-être) algorithmes de segmentation existants dans la littérature, je m'y perdrais. La plupart des techniques sont dédiées à un type d'imageurs, pour extraire à l'aide de modèles, paramétriques ou non-paramétriques, des caractéristiques, pour une application donnée. Je n'ai ici que la volonté de mettre en avant quelques contributions faisant appel de manière récurrente à des traitements similaires, compatibles avec des contraintes temporelles fortes, à l'aide d'algorithmes dont l'exécution est de si possible déterministe, et pouvant être destinés à un processus de guidage. Pour initier un suivi visuel, ou pour ré-initialiser un asservissement visuel (quand la cible est perdue!), il est généralement nécessaire de mettre en oeuvre des techniques rapides et efficaces, prioritairement à l'obtention d'une haute précision, pour pallier de telles situations.

L'extraction des contours ou des régions sont des subdivisions possibles, qui à travers la définition de descripteurs, vont permettre d'opérer le recalage. Ceux-ci renferment des informations utiles à la classification, à la mise en correspondance, à la localisation, avec parfois des facteurs de pondération qui reflètent un degré de confiance ou de précision. Les travaux de Cordelia Schmid et de David Lowe sur la définition de descripteurs locaux, tels que SIFT [Low04], pour l'alignement d'images s'inscrivent dans cette problématique. L'extraction de points caractéristiques comme les coins par le détecteur de Stephen-Harris, le détecteur SUSAN de Stephen Smith et Michael Brady [SB97], plus précis, le détecteur binaire de coins proposé récemment par Saeedi *et al.* [SLL06], la localisation délicate des points d'inflexion par le passage à zéro (et changement de signe) de la courbure [RDLR91, WD95], ou des droites par la transformée de Hough sont des outils qui font partie la segmentation en temps réel. Toute une chaîne de traitements doit alors être élaborée pour mettre en oeuvre et choisir les traitements appropriés, pour passer d'une représentation iconique avec un volume de données important à une représentation symbolique, compacte, des données image.

Qu'il s'agisse de la recherche de régions, de points d'intérêt, de contours particuliers, il existe certaines représentations de l'image qui ont permis d'élaborer des algorithmes rapides de segmentation d'un objet (ou les frontières d'un objet) par rapport au fond et des algorithmes de comparaison des images. La pyramide d'images [JM92] et l'histogramme [HS92] sont de telles représentations pour effectuer une analyse à résolution multiple ou une analyse statistique globale des données. Il est possible, via l'une ou l'autre de ces représentations de classer les données, de mesurer une "distance" entre deux images et finalement de réduire le volume de données. Dans le cas d'images couleur, la mesure de distance entre deux images peut être calculée comme la plus grande des différences entre deux histogrammes de même couleur (teinte et pureté). D'autres mesures de distance peuvent être utilisées, telles que la mesure de similarité, le test du χ^2 , la distance de Kullback-Liebler [Lef02], de Hausdorff [HR93] ou encore la distance de Bhattacharyya [DSG90].

Dans [MT00] une technique de reconnaissance très rapide de motif (< 50 ms), basée sur la corrélation normalisée est proposée. L'originalité vient de l'implémentation où le calcul du

gradient de la corrélation est mis en oeuvre en conjonction avec une pyramide de l'image. Dans le même ordre d'idée, une sélection robuste de point d'intérêt est réalisée à travers une approche multi-échelle dans [EMR02]. Les points d'intérêt retenus pour effectuer un asservissement visuel hybride dans un environnement naturel sont ceux qui apparaissent dans plusieurs niveaux de la pyramide (3 niveaux sont proposés dans [EMR02]). Dans [GH95], l'histogramme de l'image est utilisé dans un algorithme rapide de filtrage non linéaire médian. Un tel filtrage est nécessaire pour éviter la sur-segmentation en permettant de lisser la distribution des intensités dans une région de l'image tout en préservant le profil de transition de ses contours.

Chapitre 3

Synthèse des travaux

Ma principale thématique concerne l'assistance par vision artificielle, avec comme champ d'applications privilégié la robotique médicale. Par conséquent, ce chapitre décrit les activités liées à ce domaine dans lequel j'ai été fortement impliqué. En particulier, il s'agit du développement de méthodes de recalage pour la mise en oeuvre de tâches de positionnement, le suivi et le contrôle des déplacements d'instruments de chirurgie par asservissements visuels à l'aide d'imageurs médicaux, tels que la vision endoscopique en laparoscopie assistée et l'imagerie tomodensitométrique en radiologie interventionnelle. Il s'agit également, mais à un degré moindre, du développement de techniques ou d'extension de techniques existantes de filtrage et de segmentation d'indices visuels associés aux objets d'intérêts, en temps réel, et pour des images en couleurs.

3.1 INTRODUCTION

En premier lieu je me suis intéressé avec Naoufel Werghi à la segmentation et l'exploitation d'indices visuels géométriques non linéaires, en particulier les coniques. Cette étude balaie les différents problèmes rencontrés pour ce type de primitives, dans un environnement contrôlé, de la détection à l'estimation de la pose [Wer96]. Lorsqu'elle est présente dans une scène 3-D, cette famille de primitives visuelles a l'avantage de pouvoir contraindre fortement le point de vue, et donc de faciliter les appariements entre les primitives d'un modèle et celles extraites d'une image. Cependant, les coniques présentes dans les images sont notoirement délicates à extraire avec précision. C'est l'objet de la section suivante. Par la suite, je présente la segmentation des instruments de laparoscopie dans un environnement non contrôlé à l'aide d'images en couleurs.

La formulation géométrique et cinématique d'un asservissement visuel référencé image et spécifique au contrôle du déplacement d'un instrument de laparoscopie a été étudiée avec de nombreuses expériences *in vivo*. Pour parvenir à cela, il nous a fallu surmonter de nombreux problèmes que je décris dans la quatrième et la cinquième section, et qui englobent les travaux de thèse d'Alexandre Krupa [Kru03] et de Florent Nageotte [Nag05].

Ce projet a marqué un tournant important, à la fois en ce qui concerne l'équipe AVR mais également au plan personnel. Pour l'équipe AVR, ceci correspondait à l'obtention de premiers résultats quant à l'assistance robotisée au geste chirurgical par asservissement visuel. D'autres travaux, dont certains en cours, ont repris les résultats et les développements obtenus lors de ce travail exploratoire. Nous avons, grâce à cela, obtenu les premières reconnaissances de nos compétences en GMCAO.

L'estimation de pose de primitives géométriques quadratiques est abordée en premier lieu dans la section 4 pour résoudre le problème du positionnement de l'instrument sans marqueurs, pour estimer la position 3-D du point d'incision sur l'abdomen du patient et pour localiser une aiguille circulaire.

Dans la section cinq, les guidages par asservissements visuels permettant la réalisation d'une tâche de récupération d'instruments, de centrage dans l'image et finalement en interaction avec les tissus, sont présentés. Nous avons souhaité progresser dans l'assistance à la thérapie en robotique médicale. Ainsi, dans la thèse de Florent Nageotte [Nag05], je me suis intéressé avec lui à la modélisation géométrique et cinématique du couple d'instruments articulés (aiguille circulaire, porte-aiguille), montés sur un robot de chirurgie laparoscopique, afin de formaliser les déplacements et d'étudier l'assistance informatique et robotique de la suture. Le guidage est réalisé ici par la mise en oeuvre d'un asservissement visuel dynamique 2-D. La section se termine par la description d'une expérience de suivi et de guidage d'un endoscope embarqué.

Dans le cadre du projet IRASIS du programme ROBEA, je me suis intéressé avec Benjamin Maurin au recalage stéréotaxique à l'aide d'un scanner à imagerie tomographique et de marqueurs artificiels. Ce problème de recalage 2-D/3-D est essentiel pour évaluer le positionnement 3-D et le guidage automatique d'un instrument de chirurgie tel qu'une aiguille de radio-fréquence utilisée par les techniques mini-invasives en radiologie interventionnelle. Dans la sixième section, je présente la modélisation du recalage, en séparant, comme il est coutume de le faire pour d'autres modalités d'images mais pas en stéréotaxie, la modélisation de la formation de l'image de celle des marqueurs externes. Cette partie s'achève par l'application au positionnement 3-D à l'aide du robot CT-Bot.

Des travaux en cours sont présentés dans la dernière section, où l'emploi d'un capteur monoculaire est délaissé au profit d'un capteur visuel 3-D actif. Cette nouvelle orientation émane autant de l'impossibilité de localiser les objets imagés à l'aide d'un capteur monoculaire sans un apport de connaissances *a priori*, que de la volonté d'aborder le guidage au sein d'environnements complexes, en recherchant en quelque sorte à maîtriser la complexité algorithmique liée au processus de mise en correspondance de descripteurs.

3.2 SEGMENTATION D'INDICES VISUELS GÉOMÉTRIQUES NON LINÉAIRES

Nous avons vu que la complexité algorithmique du processus de mise en correspondance d'indices visuels dépendait, entre autre, du choix de ces indices et plus précisément du nombre de degrés de liberté qu'ils peuvent contraindre individuellement. Afin de pouvoir faciliter la mise en correspondance lors du calcul du recalage 2-D/3-D élaboré à partir de descripteurs géométriques locaux, nous nous sommes tournés vers des primitives contraignantes comme les coniques. Cependant, parmi les descripteurs locaux, cette famille de primitives est bien connue pour poser des obstacles à la détection.

Dans ce contexte, nous avons développé plusieurs outils de filtrage en vue d'extraire les coniques des images issues d'environnements structurés, en commençant par classer les fragments de contours des images en tant que fragments linéaires ou non [RW90], puis en approchant les fragments classés par des fonctions polynomiales.

3.2.1 Filtrage

En abordant la segmentation par la détection des contours par le calcul d'un gradient (approché) de l'image, il est usuel d'analyser la phase du gradient $\theta(s)$, signal qui représente l'orientation du contour en fonction de l'abscisse curviligne s . Il est en effet connu que les dérivées successives de l'orientation, par exemple, permettent de définir des invariants différentiels vis-à-vis d'un changement d'échelle ou d'un déplacement du contour [GB92]. Les discontinuités de l'orientation locale (dérivée seconde) marquent les limites entre les fragments élémentaires au sein d'un contour et leurs variations (dérivée troisième) reflètent les transitions entre eux. Le problème majeur apparaissant dans les méthodes de détection de ces fragments est d'accéder correctement aux informations contenues dans les dérivées supérieures, en subissant le moins possible la sensibilité au bruit, à la numérisation des données et aux troncatures dans les calculs. L'application d'un filtre de lissage classique engendre souvent des effets néfastes sur la fonction orientation qui se manifestent par une atténuation des discontinuités, une distorsion de la forme ou un effet de flou. Les conséquences sont alors fâcheuses pour la détection des discontinuités associées aux points singuliers. Un lissage sévère peut conduire à une forte délocalisation, voire à la perte de ces points importants.

Pour montrer simplement l'effet d'un filtre de lissage, considérons le développement de Taylor $f(x) = \sum \frac{f^{(n)}}{n!} x^n$ d'une fonction $f(x)$ représentant un signal d'entrée. La convolution de f avec la réponse impulsionnelle du filtre $h(x, \sigma)$, où σ est le paramètre de lissage, donne :

$$(f \star h)(x) = \sum \frac{f^{(n)}}{n!} (x^n \star h)(x)$$

La convolution de x^n ajoute d'autres termes d'ordre inférieur. A titre d'exemple, le lissage de x^2 avec une gaussienne $g(x, \sigma) = e^{-x^2/2\sigma^2} / \sigma\sqrt{2\pi}$ donne $x^2 \star g(x, \sigma) = x^2 + \sigma^2$. Dans le cas de x^3 , on obtient $x^3 + 3\sigma^2 x$, mais par contre $x \star g(x, \sigma) = x$. Ainsi, c'est seulement avec des fonctions affines que la gaussienne n'ajoute pas d'autres termes. Nous remarquons donc que des termes supplémentaires sont introduits, causant des changements dans la forme de la

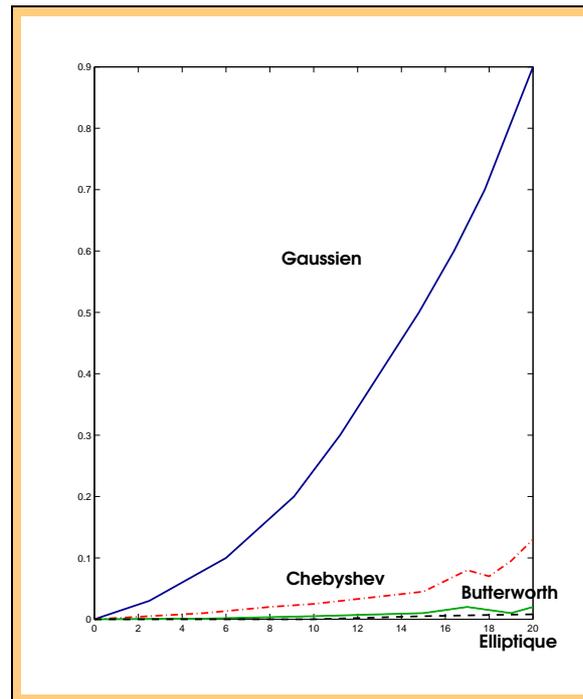


Fig. 3.1: Le critère de distorsion de Weiss en fonction de la largeur de la bande passante pour les filtres linéaires suivants : filtre gaussien (—) Chebyshev (-.-.), Butterworth (—) et elliptique (- - -).

fonction filtrée, et ces changements s'amplifient davantage lorsque la valeur du paramètre de lissage augmente. De plus, ce sont les termes issus de la convolution avec les termes de faible degré qui contribuent le plus à la distorsion de la forme.

Le critère de distorsion établi par Isaac Weiss [Wei93] nous a servi de référence pour synthétiser et comparer des filtres linéaires polynômiaux et optimaux préservant autant que possible la forme du contour [PD94], les filtres gaussiens n'étant pas appropriés pour préserver la forme des contours non linéaires. L'objectif est d'aboutir à un compromis entre l'atténuation du bruit et la préservation de la forme du contour dans sa représentation par la fonction orientation et ses dérivées successives. Le critère de Weiss s'exprime ainsi :

$$\mathcal{C}_{Weiss} = \frac{\|m_{l+1}\|}{(l+1)!} \frac{\sigma^{l+1}}{s_0} . \quad (3.1)$$

Plus \mathcal{C}_{Weiss} est faible, meilleur est le filtre. σ est le paramètre de lissage, s_0 est l'échelle de la fonction lissée (dans notre cas, la longueur du plus petit fragment du contour), m est le moment du filtre et l est son ordre. Un filtre h est dit d'ordre l s'il préserve tous les termes de degrés inférieurs ou égal à l , c'est-à-dire que la réponse impulsionnelle $H_l(x, \sigma)$ "conserve" la puissance x^n jusqu'à l'ordre l pour l'opération de convolution. Les moments, qui sont définis par

$$m_n = \int_{-w}^w x^n H_l(x, \sigma) dx ,$$

sont nuls jusqu'à l'ordre l excepté pour le moment d'ordre 0 qui vaut 1.

Finalement, pour préserver la forme du contour, la phase des filtres linéaires passe-bas doit être quasiment linéaire dans la bande passante. Parmi les filtres linéaires classiques, nous avons calculé et comparé le critères de Weiss pour les filtres polynomiaux (Butterworth et Chebyshev) et elliptiques (Fig. 3.1).

3.2.2 Détection des discontinuités et classification

Grâce à ses propriétés qui permettent une analyse temps-échelle (ou spatio-échelle) des signaux non stationnaires, la transformée en ondelette s'est avérée être un outil très efficace notamment pour réduire le bruit [CD95] et pour la détection de structures isotropes à différentes échelles. En vision par ordinateur, les ondelettes ont été utilisées pour la déconvolution, en restauration et en compression d'images [Mal89], pour le seuillage des images [Oli94], pour l'analyse de texture [Mey92, Auj04], pour la détection des contours [Mar82, FRD00] (même si on se tourne à l'heure actuelle plutôt vers des variantes qui permettent de rechercher des structures anisotropiques), pour ne citer que quelques exemples. Une démarche telle que l'approche multi-échelle permet d'aboutir à une extraction robuste de primitives géométriques, qu'elle soit effectuée au niveau des coins [CLS95], des contours [Mal89, Wer96, FRD00] ou au niveau iconique par la construction d'une pyramide d'images [EMR02].

Nous avons montré qu'il était possible de détecter les discontinuités des contours de manière efficace en convoluant le signal de phase (filtré) avec une ondelette. La classification des fragments est alors réalisée simplement en comptant le nombre de passages à zéro du signal obtenu en sortie (dérivée de la courbure). La robustesse d'une telle segmentation des contours peut être appréciée sur la figure 3.2 où les fragments linéaires sont affichés en vert, les autres en rouge, séparés par les discontinuités notées (i, j, k, l, m) sur les Fig. 3.2-b,d. Une étape ultérieure de chaînage des fragments voisins [HU90] est mise en oeuvre afin de fusionner des pixels contribuant aux mêmes fragments non linéaires (par exemple, de part et d'autre de la discontinuité i sur la Fig. 3.2-d).

L'obtention de fragments non linéaires significatifs permet d'envisager une recherche paramétrique d'une courbe quadratique (ou d'ordre supérieur) entre deux points d'inflexion. C'est ainsi que nous avons élaboré un algorithme d'estimation elliptique non biaisé, par un filtrage de Kalman. Les mesures étant basées sur la localisation des pixels des fragments de contours, le vecteur d'état étant constitué des paramètres de l'ellipse (5 paramètres), le lien entre les mesures et l'état correspond à une formulation non linéaire à partir d'un critère de minimisation non-algébrique. Pour pallier aux problèmes rencontrés par Maurizio Pilu et John Porril [Por90, PFR96] avec un critère algébrique lorsque l'excentricité elliptique est fortement prononcée [TV98, Sam82], nous avons proposé un critère de minimisation basé sur la distance orthogonale et un schéma d'estimation récursive à partir du filtre de Kalman étendu [WD95]. Un état de l'art des principales méthodes de recherche paramétrique des coniques a par la suite été réalisé par Zhengyou Zhang [Zha97].

Ce travail a donné lieu à la rédaction de 4 articles dans des conférences internationales avec actes [DAWO95, WD95, WDA96b, WDA96a], dont deux très reconnues dans ce domaine ([WD95] à ACCV'95 pour la partie filtrage et [WDA96a] à IEEE ICIP'96 pour la classification par l'emploi des ondelettes et pour la localisation).

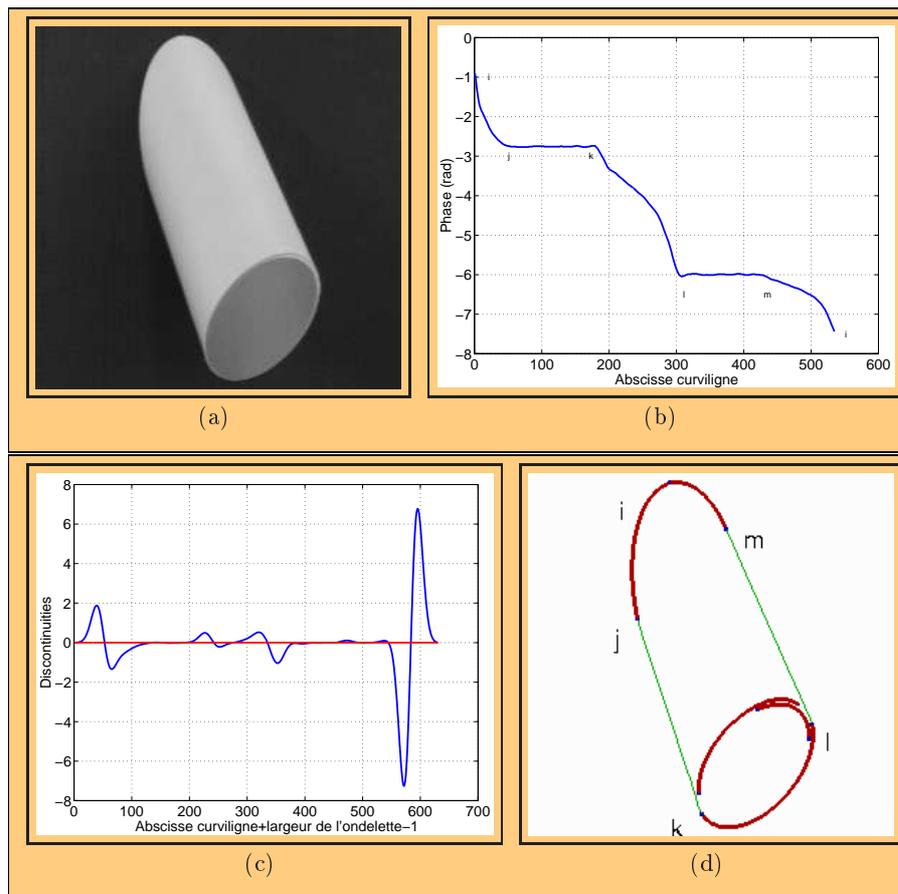


Fig. 3.2: (a) Image d'une pièce cylindrique de révolution à bases latérales circulaires et elliptiques. (b) Orientation d'un fragment de contour (ici, de la silhouette) en fonction de l'abscisse curviligne : différentes discontinuités (i, j, k, l, m) sont observées. (c) Détection des discontinuités par transformée en ondelettes (paramètre d'échelle ($\sigma = 2$)). (d) Classification de l'ensemble des fragments en deux classes.

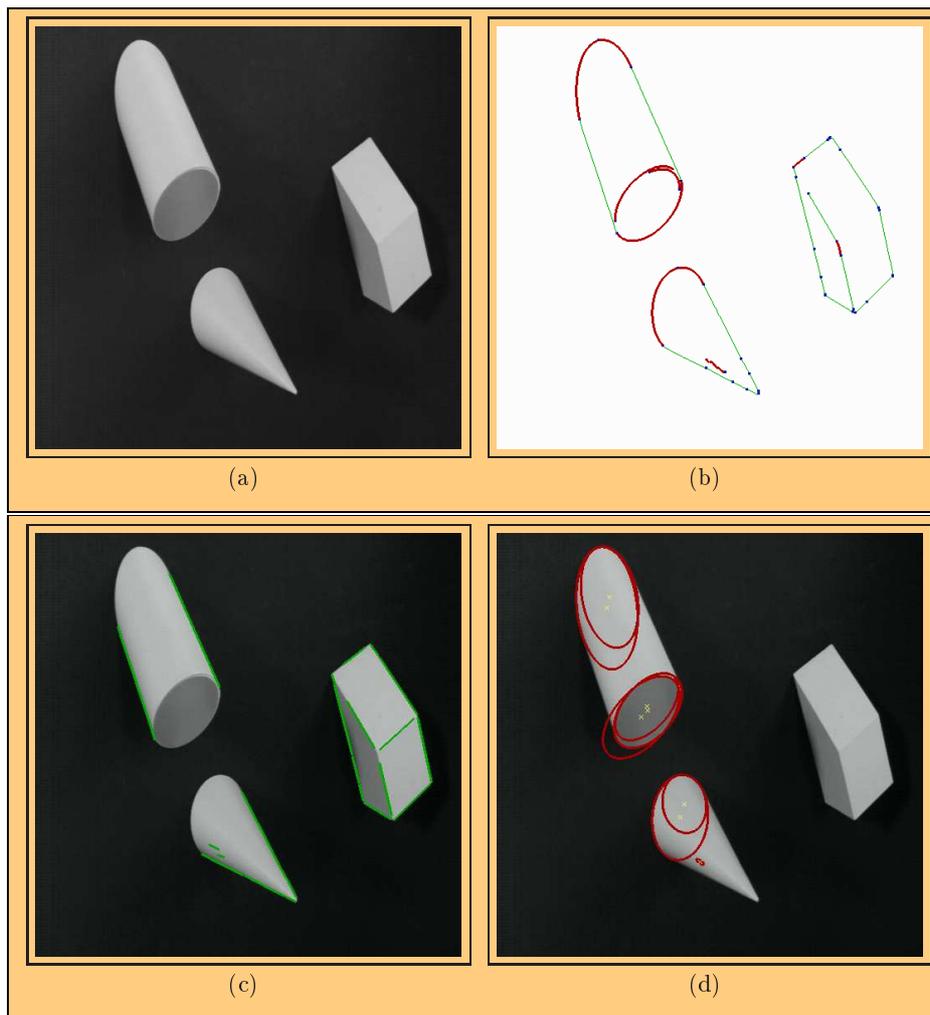


Fig. 3.3: (a) Image contenant plusieurs pièces (cylindre, cone, polyèdre). (b) Classification des fragments des contours. (c) Approximation rectiligne des fragments linéaires (droites). (d) Approximation elliptique des fragments non linéaires.

3.3 SEGMENTATION DANS DES ENVIRONNEMENTS NON CONTRÔLÉS

Quand on aborde la segmentation d'images issues d'environnements non contrôlés, il n'est pas risqué de dire qu'il n'existe pas d'approche de la segmentation préférable à une autre. Ce n'est déjà pas le cas pour les environnements contrôlés même si les images provenant de ceux-ci peuvent être interprétées à l'aide d'une seule et même approche durant une séquence temporelle. Pour les environnements non contrôlés, l'extraction d'indices visuels peut être envisagée avec succès dans la mesure où il est possible d'introduire des contraintes contextuelles, traduisant des connaissances *a priori* de l'environnement perçu.

C'est ainsi que pour extraire des descripteurs locaux dans des images endoscopiques de laparoscopie, nous avons fait plusieurs hypothèses qui n'ont de sens que pour l'application visée, l'adaptation à d'autres environnements non contrôlés ne pouvant être envisagée raisonnablement qu'après avoir modifié voire substitué l'ensemble des hypothèses formulées ici. Il n'empêche que, généralement, lorsque la segmentation d'images s'appuie sur des critères englobant plusieurs paramètres physiques, voire hétérogènes, géométriques et photométriques par exemple, on constate que les résultats obtenus sont bien plus robustes aux incertitudes qu'avec des critères construits à partir d'un seul paramètre. On a déjà rencontré cela avec les méthodes récentes de tracking qui associent avec succès intensité, couleur, texture ou contours. De plus, lors de la mise en oeuvre, il faut privilégier autant que possible les algorithmes adaptatifs au détriment de ceux qui requièrent des paramètres de seuil (de décision) dont les valeurs sont pré-définies, ces dernières étant toujours délicates à choisir.

C'est en suivant ces remarques que nous allons présenter ci-dessous la segmentation des instruments de laparoscopie dans la cavité abdominale. En dehors de l'application elle-même, deux aspects particularisent cette segmentation ; d'une part l'aspect temps réel des traitements informatiques, d'autre part l'emploi de la couleur.

3.3.1 Segmentation en temps réel des instruments de laparoscopie

Comme nous l'avons indiqué au chapitre précédent, l'objectif de la segmentation d'images est de subdiviser celle-ci en différentes parties qui correspondent aux objets d'intérêts contenus dans la scène 3-D imagée. En laparoscopie, les premières difficultés de segmentation proviennent de la complexité d'interprétation de cette scène due, entre autres, aux variations spatio-temporelles de l'éclairage, aux variations de l'arrière-plan émanant de la respiration du patient et des contacts avec les instruments. A cela s'ajoute la présence éventuelle de sang sur des parties d'instruments et des spécularités sur l'image des organes et des instruments dues à la configuration relative des surfaces éclairées vis-à-vis du laparoscope et des propriétés de diffusion/absorption de la lumière.

Pour surmonter ces problèmes, plusieurs auteurs ont par le passé, et pour l'application à la laparoscopie, employés déjà la couleur. Ainsi l'équipe d'Alicia Casals a utilisé en 1995 des instruments marqués (montés sur un robot industriel) pour réaliser une tâche de suivi 2-D, la projection perspective des marqueurs dans l'image étant représentée alors par des lignes droites. Le système de guidage avait une période d'échantillonnage de 5 Hz et nécessitait la présence

d'un assistant [CAPL95]. Un peu plus tard, l'équipe d'Hirzinger a fait appel à un système de vision stéréoscopique pour effectuer une tâche de suivi 3-D d'instruments marqués, mais cette fois-ci avec un endoscope monté sur un bras robotique. Contrairement aux travaux de Casals, il s'agissait ici de marqueurs de couleur, où, après avoir identifié la (zone spectrale de) couleur la moins fréquemment rencontrée lors d'une phase d'apprentissage (par une analyse de l'histogramme de la teinte), celle-ci a été employée pour marquer artificiellement l'instrument. Plusieurs techniques judicieuses sont à mettre au crédit de ces travaux, dont notamment le choix de l'espace de couleur retenu (HSI) qui montrent, lors d'une segmentation par la couleur, une certaine robustesse vis-à-vis des variations de l'intensité. Bien qu'il n'y ait pas une complète indépendance entre les composantes chromatiques (H et S) et celle de l'intensité lumineuse (I), ceci nous a incité à nous orienter vers une analyse chromatique pour mettre en oeuvre notre segmentation. Nous avons rencontré d'autres travaux sur la segmentation d'images couleurs endoscopiques, comme ceux d'Ascari *et al.* pour la détection de l'épine sur-arachnoïde [ABL⁺04] et les techniques de segmentation et de recalage dédiées à la laparoscopie et développées à l'université Johns Hopkins [BCD⁺04].

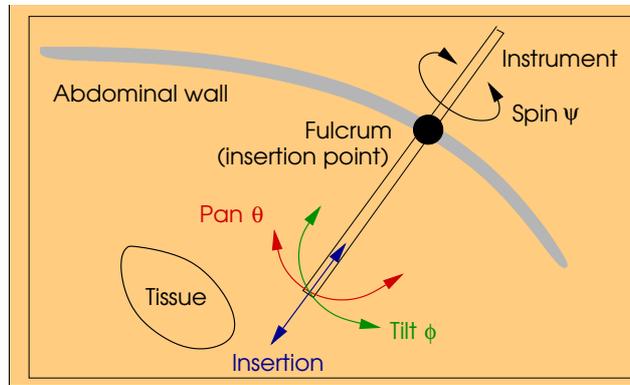


Fig. 3.4: Les mouvements possibles des instruments de laparoscopie et la définition des angles de rotation et de la profondeur d'insertion. Un guide, appelé trocart, est inséré au point d'insertion pour contraindre la mobilité afin d'éviter les frottements et les déchirements.

Nous nous sommes donc inspirés de cela pour concevoir des algorithmes adaptatifs et proposer une méthodologie pour la segmentation d'images couleur en laparoscopie qui supposent que l'arrière-plan (la cavité abdominale) est majoritairement de couleur rouge, que les instruments de chirurgie étant métalliques sont supposés être gris et que d'autre part, la contrainte cinématique des instruments de laparoscopie se traduit par la présence (invisible) d'un point d'insertion, une zone qui, théoriquement, ne permet pas de mouvements tangents à la surface de l'abdomen comme le montrent la figure 3.4.

La saturation S est un attribut chromatique qui mesure la pureté de la couleur, ou, en d'autres termes, la quantité de gris contenue dans une couleur. Ceci nous est apparu être une bonne manière d'aborder l'analyse des images de laparoscopie. Cette grandeur est définie dans la représentation HSI par rapport aux composantes RGB dans [BB82] par :

$$S_{HSI} = 1 - 3 \frac{\min\{R, G, B\}}{R + G + B} \quad (3.2)$$

et dans HSV par

$$S_{HSV} = 1 - \frac{\min\{R, G, B\}}{\max\{R, G, B\}} \quad (3.3)$$

A priori, les systèmes de coordonnées comme HSI et HSV, qui se rapprochent de la perception psychologique des couleurs par l'oeil humain, semblent plus appropriés pour effectuer une analyse des couleurs que l'espace RGB, qui présente une évolution non uniforme de la chromaticité en fonction de l'intensité [YR97]. L'espace RGB devient totalement inadéquat pour discriminer deux distributions de couleur par les techniques de segmentation qui ont besoin d'une mesure de similarité qui soit insensible aux variations de l'intensité [LSM⁺05]. Dans le plan chromatique, la teinte (H - hue) et la saturation (S) sont souvent représentées respectivement par une orientation et par une distance. Si cela est vrai pour certains espaces de (représentations des) couleurs, on voit clairement que les formules 3.2 et 3.3 n'obéissent pas à cela. Ces définitions montrent une certaine immunité vis-à-vis des variations de l'intensité, sauf à proximité du gris ($R = G = B$) justement et une discontinuité proche du noir ($R = G = B = 0$). Une saturation nulle correspond à l'absence d'une composante de couleur dominante (ce qui pour notre application signifie "région non rouge" la plupart du temps) alors qu'une saturation élevée (S proche de 1) correspond à une composante de couleur très fortement prédominante (ce qui peut correspondre à "une région rouge" pour notre application). On s'aperçoit que la segmentation des régions grises correspond avec ce critère à la recherche des régions peu colorées, c'est-à-dire contenant des pixels achromatiques. L'imprécision de la teinte pour une saturation au voisinage de 0 pose des problèmes lors des transitions de couleurs et, surtout ces deux définitions de la saturation ne sont pas des métriques (voir également [IPV99]), ce qui n'en fait pas un bon candidat pour établir une comparaison.

En 2003, le LSIIT a accueilli Jean Serra dans le cadre de ses séminaires. Parallèlement à ses travaux en morphologie mathématique, il a mené des recherches avec Allan Hanbury sur l'analyse des couleurs. Lors d'une présentation très pédagogique, il a clairement démontré la déficience de certains critères comme celui de la saturation et il proposé de modifier la définition précédente de la saturation pour l'espace HSV en lui substituant la formule suivante

$$S'_{HSV} = \max\{R, G, B\} - \min\{R, G, B\} \quad (3.4)$$

Sans rentrer dans plus de détails que l'on trouvera dans [HS02], nous avons immédiatement adopté ce critère, les démonstrations, les résultats obtenus et la simplicité de la définition de la saturation nous ont convaincu d'employer cette définition.

Il demeure tout de même qu'en présence de régions présentant des spéularités, un critère basé uniquement sur la pureté de la couleur sélectionnera naturellement ces régions comme étant constituées de pixels achromatiques. Dans [WGG03], les auteurs ont proposé, lors de la recherche des contours d'une image couleur, d'utiliser un critère, que nous noterons \mathcal{H} , construit à partir du produit des deux attributs chromatiques H et S . Les auteurs ont également montré que la dérivée première de \mathcal{H} (à S constant) était directement reliée à un pseudo-invariant photométrique, noté \mathbf{H}_x^c et dont la norme vaut $|\mathbf{H}_x^c| = S \cdot H_x$ (H_x est le gradient spatial de la teinte H).

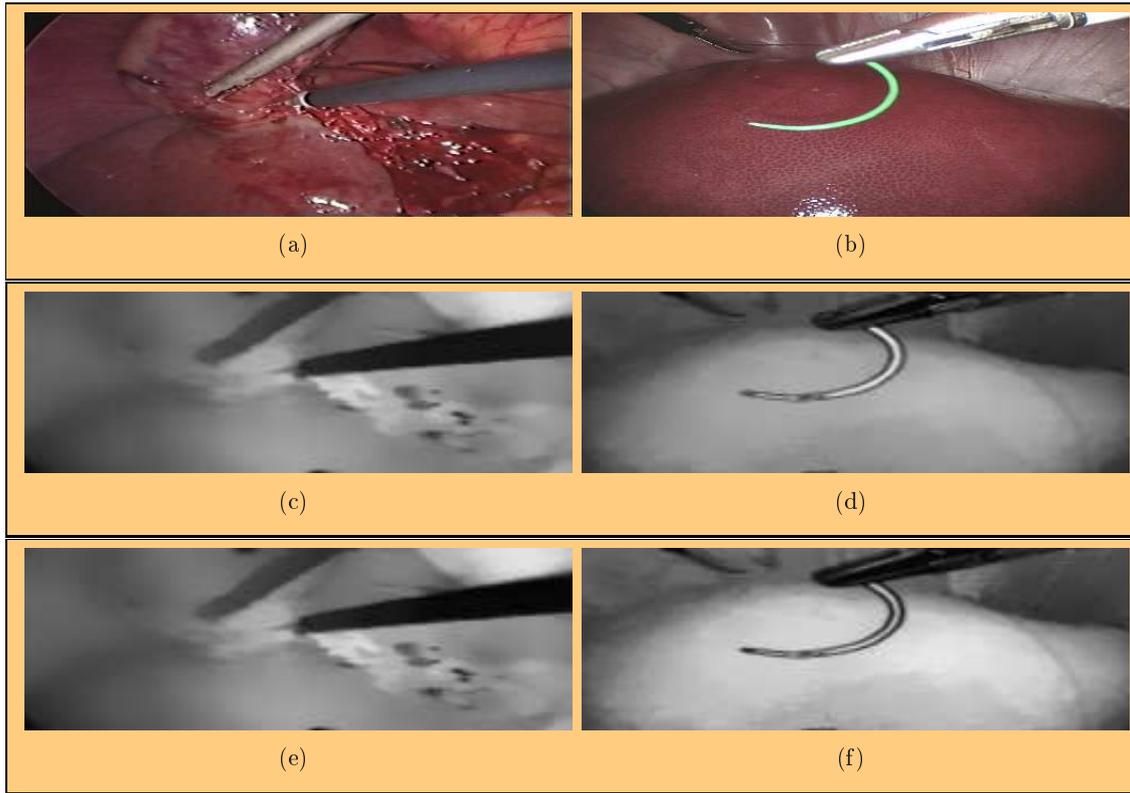


Fig. 3.5: Discrimination des couleurs avec le critère \mathcal{H}' conjoint à la saturation et à la teinte, et sa robustesse vis-à-vis des spécularités et des changements d'intensité. (a-b) images extraites d'une séquence vidéo. (c-d) L'application du filtre Sigma sur l'image de la saturation S'_{HSV} . (e-f) L'application du filtre Sigma sur l'image formée par le critère \mathcal{H}' . Les différences entre (c) et (e) ne sont pas très significatives dans les régions très colorées de l'arrière-plan présentant peu de spécularités, alors que les différences sont visibles entre les images (d) et (f) où des spécularités sont produites sur la surface (grise) de l'instrument de chirurgie. On peut noter aussi une meilleure détection des contours de l'aiguille, correspondant à des transitions de la teinte.

La teinte peut être définie à partir des composantes RGB par

$$H = \arccos \left(\frac{\frac{1}{2} [(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right). \quad (3.5)$$

La teinte H n'est pas définie pour des pixels achromatiques. Ainsi, un inconvénient bien connu, réside dans le fait que pour des valeurs faibles de la saturation S , de faibles variations autour de l'axe de luminance (ou axe des gris) génèrent des modifications importantes de la direction du pseudo-invariant \mathbf{H}_x^c [WGG03] et par conséquent que la teinte H n'est pas bornée. L'aspect le plus pertinent de ces travaux est que par contre la norme de \mathbf{H}_x^c reste bornée, donc son intégrale l'est aussi. Il s'ensuit que \mathcal{H} reste borné. Van de Weijer *et al*, ont montré que ce critère était efficace pour discriminer les couleurs des régions contenant des spécularités.

Finalement, nous avons donc adopté le critère conjoint suivant

$$\mathcal{H}' = S'_{HSV} \cdot H \quad (3.6)$$

qui est illustré sur les images de la figure 3.5. Sur cette figure, un filtrage non linéaire du type filtre *Sigma* de Lee [Lee83, HS92] a été exécuté sur les images dont les pixels correspondent à la valeur (entière et normalisée) du critère \mathcal{H}' , avec une implémentation temps réel. Le filtrage non linéaire a pour objectif d'éviter une sur-segmentation des images [PV90]. Le filtrage non linéaire Sigma de Lee fournit des résultats très voisins de ceux obtenus à l'aide du filtre à diffusion anisotrope de Malik et Perona [PSM94, VESS04], mais permet de plus de mettre en oeuvre, moyennant quelques approximations, une exécution très efficace. Ces filtres, de manière analogue au filtre médian permettent de lisser les pixels des régions tout en préservant les hautes fréquences donc les transitions (avec tout de même souvent quelques distorsions observées). Pour être aussi efficace, le filtre médian doit être itéré, ce qui augmente sensiblement le temps de calculs, même si des implémentations efficaces de celui-ci existent dans la littérature (voir [GH95], par exemple). Les détails de l'ensemble de la segmentation et du filtrage non linéaire en temps réel sont rassemblés dans la publication [DGM05].

Pour être complet, un algorithme récursif d'expansion de régions a été employé après la recherche des germes potentiels [AB94] à la périphérie de l'image, les régions dans l'image correspondant aux instruments de laparoscopie ayant toujours une intersection avec le bord de celle-ci [DNM04]. Des résultats sont illustrés sur la figure 3.6. Ils montrent que les régions correspondant aux instruments sont détectées, même si parfois d'autres régions sont retenues. Sur cette figure, les croix noires indiquées sur les courbes bleues correspondent à la localisation des germes potentiels aux niveaux des minima locaux.

A ce stade, à la fois la forme cylindrique des instruments n'est pas pris en compte, ni la présence d'un point d'insertion fixe. Afin de ne retenir finalement que les régions correspondant aux instruments et pour paramétrer celles-ci, l'estimation de pose, la modélisation des mouvements et les contraintes des mouvements en chirurgie laparoscopique seront présentées dans la section suivante.

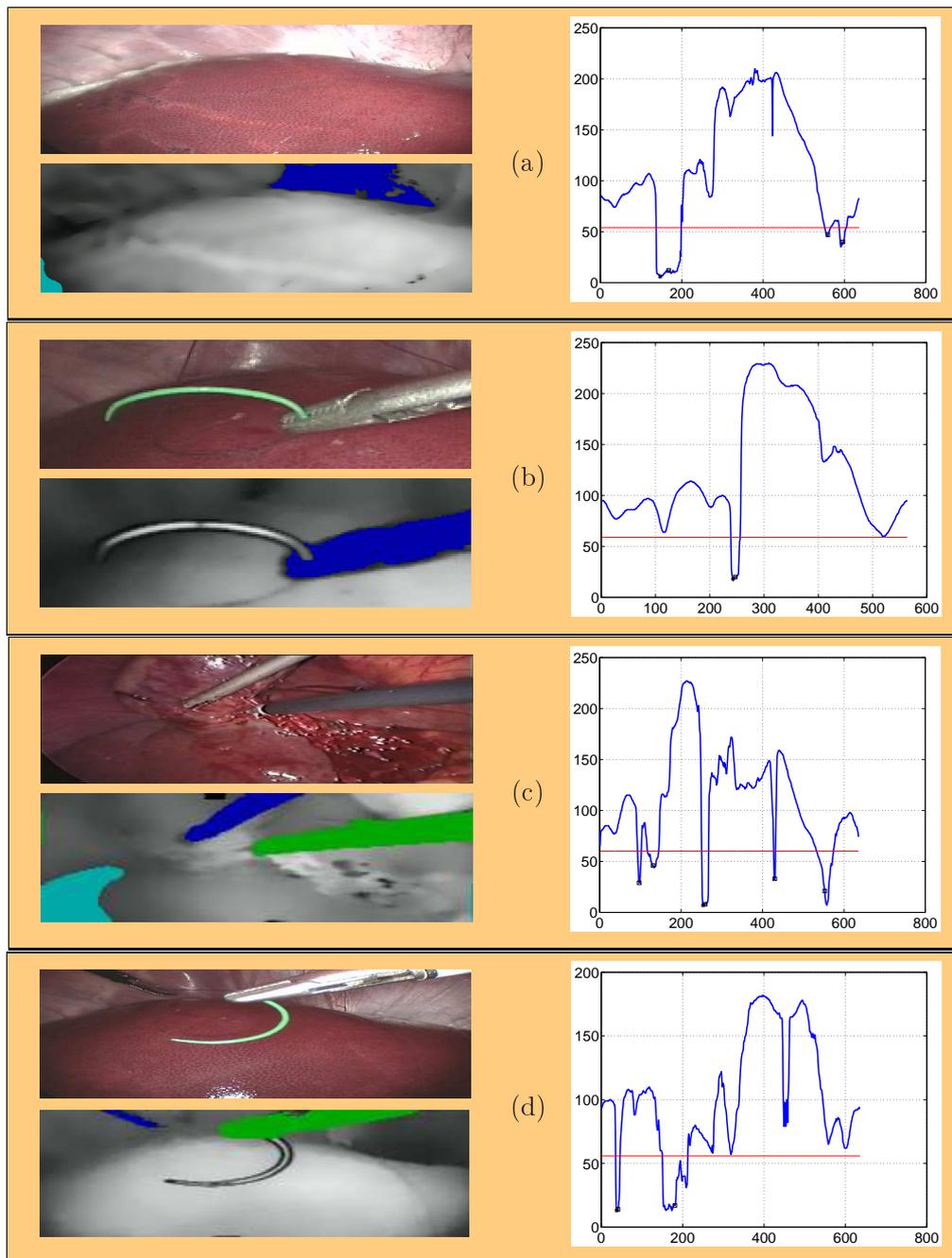


Fig. 3.6: Quatre exemples (a-d) de la segmentation basée région d'images en couleurs. (A gauche) Les régions qui sont été segmentées (colorées) et correspondant à l'image des instruments sont toujours détectées. Les autres régions retenues sont dues à une non-uniformité de la lumière ou à des zones spéculaires. (A droite) Le critère \mathcal{H}' le long de la périphérie de l'image (le coin supérieur gauche de l'image est pris comme origine de l'axe horizontal, le parcours s'effectue dans le sens des aiguilles d'une montre). La ligne rouge correspond à la valeur moyenne de \mathcal{H}' diminuée du double de son écart-type.

3.4 POSITIONNEMENT DES INSTRUMENTS EN LAPAROSCOPIE ET APPLICATIONS

Le recalage 2-D/3-D est étudié ici pour le positionnement d'instruments de chirurgie à invasion minimale, marqués ou non marqués. Dans ce dernier cas, il s'agit surtout d'étudier l'estimation de la pose à partir de primitives géométriques quadratiques, dans un contexte général et un environnement structuré, puis finalement à l'intérieur du corps humain.

Tout d'abord, il n'est pas inutile de noter que lors de la projection dans les images de formes géométriques différentes, des contours apparents similaires peuvent être obtenus. Il en est ainsi, par exemple, du cercle, de l'ellipse et de la sphère. Il est alors nécessaire d'inclure des données du modèle ou de s'intéresser à la distribution des intensités de lumière pour déterminer l'algorithme d'estimation de pose qu'il convient d'employer. On retrouve cela pour d'autres formes, comme pour un couple de droites et les génératrices des cylindres droits de révolution.

Bien que certaines tâches de positionnement ne requièrent pas une calibration complète du capteur, il est néanmoins nécessaire de pouvoir corriger les effets non linéaires de la projection de la lumière à travers un endoscope. Des techniques bien connues peuvent être mises en oeuvre pour corriger surtout les effets de la distorsion radiale [Tsa87, RDD194, HYM95, DA99, DF01], c'est, entre autre, ce que propose les outils logiciels mis en ligne par Jean-Yves Bouguet [Bou], que l'on peut utiliser aussi bien pour la vision à travers un laparoscope ou une webcam (figure 3.7).



Fig. 3.7: Les endoscopes, au même titre que les caméras fish-eye, webcams ou tout dispositif optique à champ de vision large et à focale faible, provoquent une distorsion de l'image qu'il convient de corriger avant d'effectuer un recalage basé sur le modèle perspectif. (a) Exemple d'une scène intérieure dont l'image est distordue. (b) Image corrigée selon le modèle de la distorsion radiale de Robert Tsai [Tsa87].

3.4.1 Cas des ellipses

Dans la l'assistance à la suture, la localisation d'une aiguille circulaire (c'est en fait une forme torique de révolution, mais de section suffisamment faible pour la réduire à son axe) a été possible grâce à l'emploi des techniques de la littérature. Il n'y a pas à proprement parler de contributions scientifiques de notre part sur ce sujet, et parmi les méthodes existantes [SA89b, DLRR90, FMV93, SRTSB92], nous avons emprunté celle de Dhome *et al.* [DLRR90] pour l'estimation de pose d'un cercle (5 ddl, figure 3.8).



Fig. 3.8: Approximation elliptique et localisation 3-D d'une aiguille de suture (en bleue).

3.4.2 Cas des sphères

Je me suis intéressé également à la pose à partir d'objets de forme sphérique depuis longtemps, pour proposer durant ma thèse de doctorat une méthode numérique de calibration de caméras [Doi94]. Il n'existe pas, à ma connaissance, d'instruments de chirurgie de cette forme, mais on peut la rencontrer pour localiser des marqueurs passifs employés pour la capture des mouvements humains mais aussi pour permettre aux trackers optiques dédiés au domaine médical de localiser des postures de chirurgiens [ZP02, NHK⁺03] ou des instruments [NIN⁺06]. Nous les avons employés à maintes reprises avec le dispositif Polaris de l'équipe AVR (figure 3.9) et aussi directement avec un système de vision monoculaire classique. Enfin, la modélisation géométrique développée ci-dessous pour les sphères servira, dans le prochain paragraphe, à modéliser les cylindres. C'est lors d'un projet de robotique mobile (Coupes de robotique E=m6) avec les élèves-ingénieurs ENSPS que j'ai proposé le schéma de résolution suivant.

Il est bien connu que la projection d'une sphère par la projection centrale est un cône de révolution dont le sommet est confondu avec le centre de projection. L'intersection de ce cône avec la surface de la sphère est appelée le contour générateur (Γ) et l'intersection avec le plan image fournit le contour apparent (γ). Ces intersections sont des courbes planes, de forme elliptique en générale (figure 3.10). A notre connaissance, la formulation mathématique et la solution du problème de pose à partir d'une vue d'un objet sphérique a été à l'initiative de Shin et Ahmad [SA89a]. Safaee-Rad *et al.* [SRTSB92] ont étudié également ce problème dans le contexte de la robotique mobile et ils ont mis en avant des limitations pratiques restreignant la précision de la localisation 3-D, comme la localisation et la paramétrisation des contours par les ellipses, l'influence des facteurs d'échelle de la caméra et la distorsion de l'objectif.

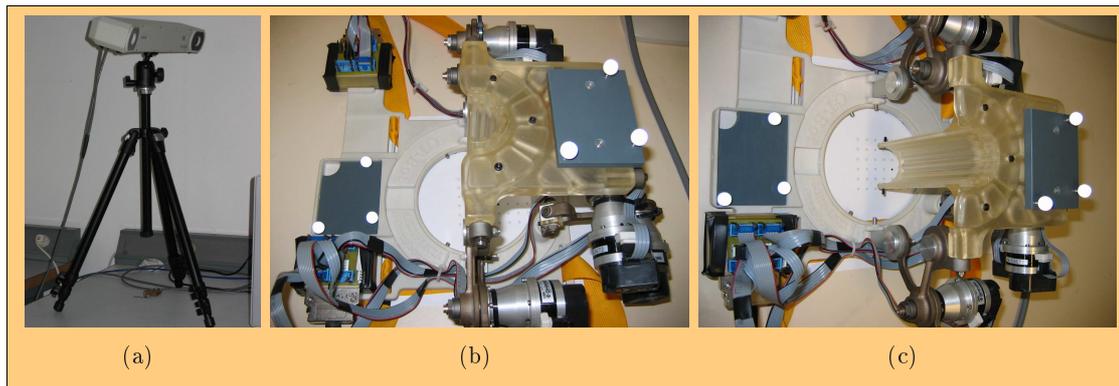


Fig. 3.9: (a) Le tracker optique Polaris de NDI. (b-c) Utilisation de marqueurs sphériques pour calibrer le robot CT-Bot.

Une méthode de calibration de caméras avec ce type de primitives a été aussi proposée par Daucher *et al.* [DLRR90, DDL94] où à peu près les mêmes problèmes ont été soulevés et en remarquant que l'axe principal de l'ellipse approximant le contour apparent passe par le centre de projection. Plus récemment Teramoto et Xu [TX02], Agrawal et Davis [AD02] ont proposé des solutions élégantes respectivement au problème de la pose et à la calibration à partir de 3 sphères.

Nous allons ci-dessous rappeler la méthode proposée par Teramoto, à partir de laquelle nous avons apporté une légèrement modification.

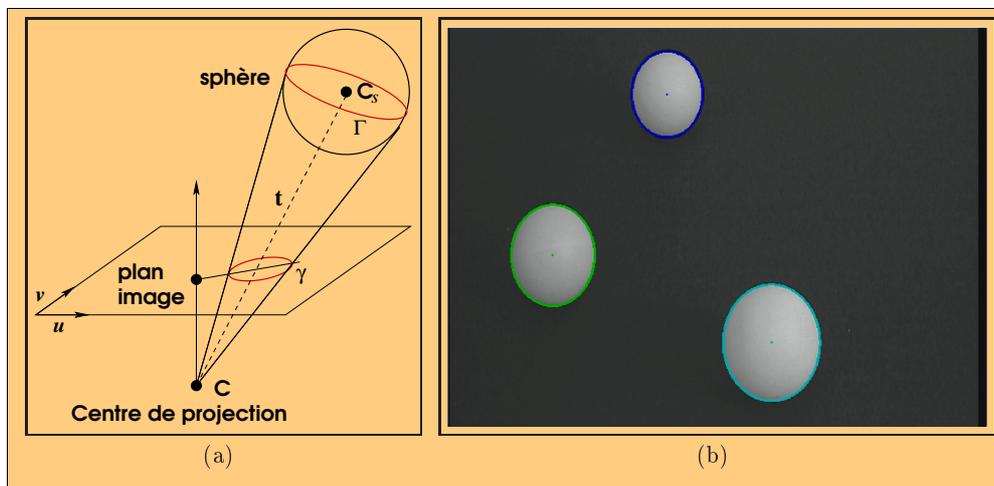


Fig. 3.10: L'image d'une sphère par projection perspective est une ellipse dont l'axe principal intersecte le centre de projection. (a) Une sphère, son contour générateur Γ et son contour apparent (γ) dans l'image. (b) Chaque ellipse détectée est représentée par une matrice symétrique E^* qui est reliée aux coordonnées 3-D du centre de la sphère.

Une sphère S est une quadrique, et pour tout point M de la surface exprimé par ses coordonnées homogènes, elle peut être représentée par une matrice (4×4) symétrique réelle S de la manière

suivante :

$$\mathbf{M}^T \mathbf{S} \mathbf{M} = 0 \quad (3.7)$$

La matrice \mathbf{S} dépend du rayon r et de la position du centre de la sphère représenté par le vecteur $\mathbf{t} = (t_x, t_y, t_z)^T$, exprimé dans le repère monde \mathcal{F}_w . Quand le repère de la caméra \mathcal{F}_c coïncide avec le repère monde, t_z est généralement choisi perpendiculairement au plan image et de direction pointant vers la scène (figure 3.10-a). La matrice \mathbf{S} exprimée dans \mathcal{F}_c est alors de la forme

$$\mathbf{S}(r, \mathbf{t}) = \begin{bmatrix} \mathbf{I} & -\mathbf{t} \\ -\mathbf{t}^T & \mathbf{t}^T \mathbf{t} - r^2 \end{bmatrix} = \mathbf{H}_{cs}^{-T} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^T & -r^2 \end{bmatrix} \mathbf{H}_{cs}^{-1}$$

où \mathbf{I} est la matrice identité et \mathbf{H}_{cs} est la matrice de transformation euclidienne entre le repère de la caméra et le centre de la sphère (la rotation \mathbf{R} peut donc être choisie de manière arbitraire)

$$\mathbf{H}_{cs} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3.8)$$

\mathbf{R} est une matrice 3×3 de rotation ($\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}$, $\det(\mathbf{R}) = 1$). Comme la quantité $\mathbf{t}^T \mathbf{t}$ correspond au carré de la distance entre le centre de la sphère et le centre de projection C de la caméra, le scalaire $\mathbf{t}^T \mathbf{t} - r^2$ doit être positif lorsqu'on considère un objet sphérique devant la caméra (la sphère n'englobe pas la caméra). La figure duale de la sphère \mathcal{S} est aussi une sphère représentée par la matrice adjointe de \mathbf{S} , \mathbf{S}^* ($\mathbf{S}^* = \mathbf{S}^{-1}$). Les deux représentations équivalentes seront utilisées par la suite.

Avec le modèle de la projection perspective, un point \mathbf{M} de l'espace 3-D est projeté en un point \mathbf{m} tel que les coordonnées homogènes vérifient

$$\lambda \mathbf{m} \equiv \mathbf{P} \mathbf{M} \quad (3.9)$$

λ est un scalaire non nul. La matrice \mathbf{P} de dimensions (3×4) représente la projection perspective canonique et est appelé matrice caméra. Elle peut s'écrire

$$\mathbf{P} = \mathbf{K} \left[\mathbf{I} \mid \mathbf{0} \right], \quad (3.10)$$

dans le repère \mathcal{F}_c . \mathbf{K} est la matrice des paramètres intrinsèques. C'est une matrice triangulaire supérieure normalisée ($K_{33} = 1$) qui est souvent décomposée selon

$$\mathbf{K} = \mathbf{K}_1 \mathbf{K}_f = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.11)$$

(u_0, v_0) sont les coordonnées inhomogènes du point principal (intersection de l'axe optique de l'objectif sur le plan image, qui n'est pas forcément confondue avec le centre de l'image). α_u et α_v sont des facteurs d'échelle qui dépendent de la taille des éléments photo-sensibles, des dimensions de l'image selon les deux directions horizontale, u , et verticale v . s est un paramètre de déformations (cisaillement) entre les deux axes. Avec des capteurs modernes, les axes sont très proches de l'exacte perpendicularité avec une "horloge pixel" et ce paramètre est souvent négligé. Avec des systèmes de vision à vidéo composite, il doit être pris en considération

[Tsa87, Doi94]. f est la longueur focale effective. C'est le seul paramètre intrinsèque à l'objectif lorsque les distorsions ont été corrigées au préalable.

Avec ce modèle de projection perspective, l'image du contour générateur d'une sphère \mathcal{S} est une ellipse \mathcal{E} qui peut être représentée par la matrice (3×3) symétrique E selon l'équation suivante qui fait intervenir les matrices adjointes :

$$E^* = P S^* P^T . \quad (3.12)$$

Par substitution des équations (3.8) et (3.10) dans (3.12), il a été montré dans [AD02] que :

$$\mu E^* = K \left(I - \frac{1}{r^2} \mathbf{t} \mathbf{t}^T \right) K^T \quad (3.13)$$

pour tout nombre réel μ non nul. L'algorithme de Teramoto permet, à partir de cette représentation des sphères, de calculer directement les composantes du vecteur \mathbf{t} .

L'algorithme de Teramoto

Dans [TX02], la position relative d'un objet sphérique est déterminée à partir d'une seule vue d'une caméra calibrée. Etant donné les matrices K et E^* , on note la direction de vue par le vecteur unitaire $\mathbf{t}_u = \mathbf{t}/s$, ($s = \pm \|\mathbf{t}\|$). A partir de l'équation (3.13), nous obtenons alors :

$$\mu Q^* = I - \frac{1}{r^2} \mathbf{t} \mathbf{t}^T \quad (3.14)$$

$$= I - \frac{s^2}{r^2} \mathbf{t}_u \mathbf{t}_u^T \quad (3.15)$$

où $Q^* = K^{-1} E^* K^{-T}$. Comme $\|\mathbf{t}\|/r$ est toujours supérieur à 1, le membre de droite de cette équation correspond à une matrice de rang plein (rang 3). Il peut s'écrire sous la forme

$$[\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_u] \text{diag}(1, 1, 1 - \frac{s^2}{r^2}) [\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_u]^T \quad (3.16)$$

et le membre de gauche de la même équation peut être décomposé selon

$$U \text{diag}(\lambda_1, \lambda_2, \lambda_3) U^T \quad (3.17)$$

où $U = [U_1, U_2, U_3]$ est une matrice orthonormale. Il est donc clair que nous avons les équivalences suivantes :

$$1 - \frac{s^2}{r^2} = \lambda_3/\lambda_1 \quad (3.18)$$

$$\mathbf{t} = s U_3 \quad (3.19)$$

La solution est unique car le signe positif de t_z impose celui de s . Le principal avantage de cette solution réside dans sa simplicité à déterminer le vecteur de position à partir de la décomposition en valeurs singulières de Q^* .

Une version modifiée

Il convient de remarquer que, d'une part, la matrice symétrique E^* a 5 paramètres indépendants (elle représente une conique) et que d'autre part, la quantité $I - \frac{1}{r^2} \mathbf{t} \mathbf{t}^T$ a deux valeurs singulières égales à 1 ($\mathbf{t} \mathbf{t}^T$ est une matrice de rang 1) et représente une ellipse dans l'espace image "calibrée". Par conséquent l'information géométrique renfermée dans l'équation (3.15) n'a pas été pleinement exploitée. Cela signifie que soit le vecteur de position \mathbf{t} ne peut être déterminé qu'à partir d'un système surdéterminé (redondance des données) de rang non plein ou que d'autres paramètres (comme des paramètres intrinsèques) peuvent être simultanément déterminés à partir d'une seule sphère. Agrawal et Davis [AD02] ont montré que l'ensemble des composantes de K pouvait être déterminé théoriquement avec seulement 3 sphères de position respective inconnue.

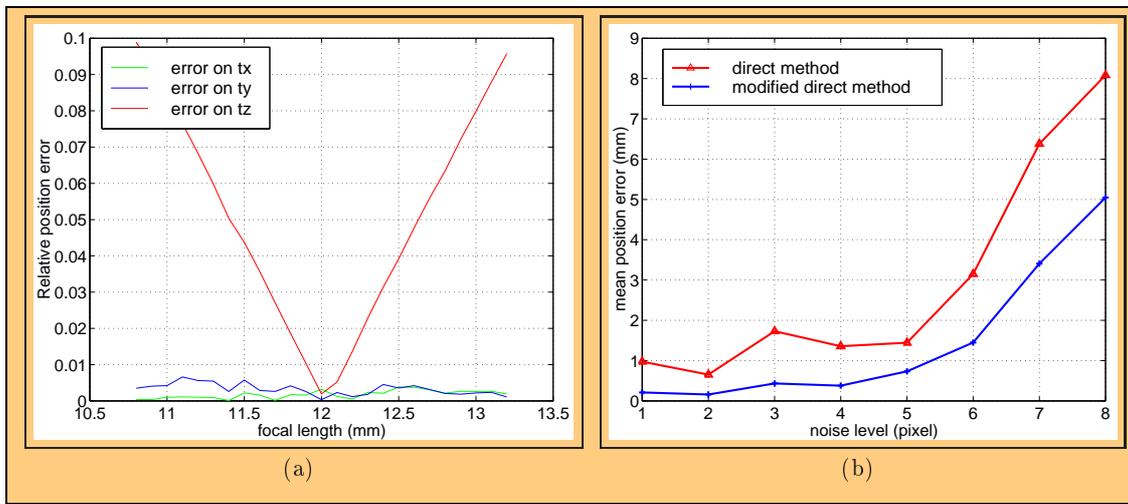


Fig. 3.11: (a) Erreurs relatives de positions sur les trois composantes du vecteur de position \mathbf{t} lors d'un changement de la longueur focale (la valeur exacte est $f = 12$ mm, le bruit gaussien sur les pixel-contours de l'ellipse vaut 0.5 pixel). (b) Erreur moyenne de position en fonction du niveau de bruit pour la méthode directe (Teramoto, en rouge) et la méthode modifiée (en bleu).

Cette remarque nous a conduit à améliorer la méthode de Teramoto vis-à-vis du bruit. Elle consiste à légèrement modifier l'équation (3.18). Comme les deux valeurs singulières λ_1 et λ_2 sont égales, elle doivent demeurer égales en présence de bruit sur les données. En remplaçant (3.18) par

$$1 - \frac{s^2}{r^2} = 2\lambda_3 / (\lambda_1 + \lambda_2) \quad (3.20)$$

c'est-à-dire quand λ_1 est remplacée par la moyenne des deux valeurs singulières, la dispersion des résultats en présence de bruit est forcément plus faible comme en attestent les résultats de simulations numériques de la figure 3.11-b. La matrice résultante Q_m^* est alors simplement

$$Q_m^* = U \text{diag}((\lambda_1 + \lambda_2)/2, (\lambda_1 + \lambda_2)/2, \lambda_3) U^T. \quad (3.21)$$

en forçant les deux valeurs singulières λ_1 et λ_2 . La matrice (4×4) \hat{S} correspondant à Q_m^* est

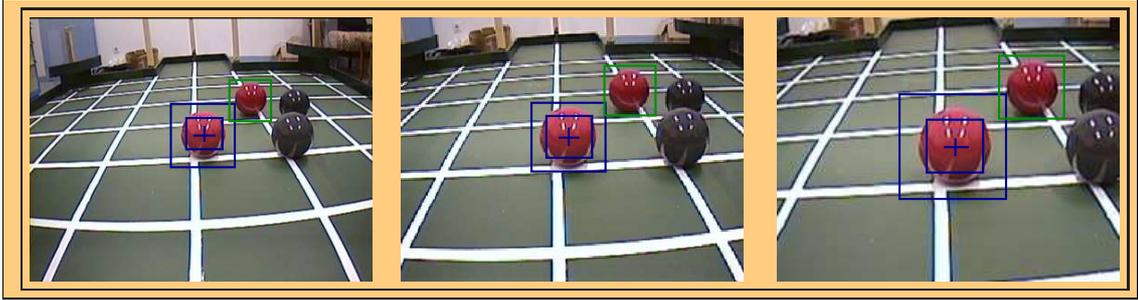


Fig. 3.12: Suivi 3-D de balles rouges. Un robot à roues embarque une caméra et se déplace sur le terrain. Il doit se saisir le plus rapidement possible des balles et les lancer dans un panier (de basket). Les régions d'intérêt 3-D sont délimitées par des fenêtres rectangulaires dans l'image et servent à définir l'espace de recherche du suivi. Seule, la balle la plus proche est prise en compte. Les deux rectangles bleus correspondent (à un facteur près) à l'incertitude (majorée) sur la distance estimée de la balle. La croix bleue correspond à la position estimée de l'image du centre.

reliée au vecteur de position $\hat{\mathbf{t}}$ qui s'obtient par

$$\mu \mathbf{Q}_m^* = \mathbf{I} - \frac{1}{r^2} \hat{\mathbf{t}} \hat{\mathbf{t}}^T. \quad (3.22)$$

3.4.3 Cas des cylindres

Il existe une littérature abondante sur l'estimation de pose de cylindres. Nous avons effectué un état de l'art cette année dans [DM07] où nous proposons également une nouvelle méthode, directe, et qui considère le problème de l'estimation de la position de l'axe de symétrie d'un cylindre homogène, droit et de révolution (SHCC - *Straight Homogeneous Circular Cylinder*). La représentation géométrique d'un cylindre, en tant que quadrique dégénérée, est modélisée par des intersections de sphère et de plans, pour aboutir à une matrice singulière qui dépend du rayon du cylindre r_c , et des coordonnées de Plücker \mathbf{r} et \mathbf{w} de son axe de symétrie

$$\mathbf{M}^T \underbrace{\begin{bmatrix} [\mathbf{r}]_{\times} [\mathbf{r}]_{\times}^T & [\mathbf{r}]_{\times} \mathbf{w} \\ \mathbf{w}^T [\mathbf{r}]_{\times}^T & \|\mathbf{w}\|^2 - r_c^2 \end{bmatrix}}_{\mathbf{Q}_c} \mathbf{M} = 0. \quad (3.23)$$

La projection perspective d'une quadrique est une conique \mathbf{C} qui est usuellement calculée par l'intermédiaire de la matrice duale \mathbf{Q}_c^* de \mathbf{Q}_c et grâce au fameux théorème de Cayley-Hamilton [HZ00]. Malheureusement dans le cas d'une matrice dégénérée, la matrice duale \mathbf{Q}_c^* n'est pas unique $\mathbf{Q}_c (\mathbf{Q}_c^*)^T = 0$. Une solution alternative consiste à prendre directement les équations de la projection perspective qui à tout point $\mathbf{m} = (u, v, 1)^T$ de l'image correspond un ensemble infini de points $\mathbf{M} = (x, y, z, 1)$ dans l'espace

$$\mathbf{M} = z \begin{bmatrix} \mathbf{K}^{-1} \\ \mathbf{0}^T \end{bmatrix} \mathbf{m} + [0 \ 0 \ 0 \ 1]^T \quad (3.24)$$

quand l'axe z est choisi selon direction de l'axe optique. La substitution de l'équation précédente dans (3.23) fournit une équation polynômiale du second degré $A z^2 + 2B z + C = 0$, avec

$$\begin{cases} A = \mathbf{m}^T \mathbf{K}^{-T} [\mathbf{r}]_{\times} [\mathbf{r}]_{\times}^T \mathbf{m} \\ B = \mathbf{m}^T \mathbf{K}^{-T} [\mathbf{r}]_{\times} \mathbf{w} \\ C = \|\mathbf{w}\|^2 - r_c^2 \end{cases} \quad (3.25)$$

Le contour apparent (γ) d'un cylindre est un ensemble de points qui intersecte la ligne de vue avec le plan image en une seule localisation [Cha90]. Ceci signifie que c'est la racine double du polynôme qui fournit la solution recherchée, en exprimant que le discriminant $B^2 - AC$ est nul. Cela nous donne

$$\left(\frac{r_c B}{\sqrt{C}} + \mathbf{w}^T \mathbf{K}^{-1} \mathbf{m}\right)^T \left(\frac{r_c B}{\sqrt{C}} - \mathbf{w}^T \mathbf{K}^{-1} \mathbf{m}\right) = 0. \quad (3.26)$$

Si la quantité $C \leq 0$, le centre de projection $C_c = (0, 0, 0, 1)^T$ est localisé à l'intérieur du cylindre, (ou sur sa surface, pour $C = 0$) et n'amène pas l'obtention d'une solution réelle pour z . Ceci peut être facilement observé à partir de l'équation (3.23) car $C_c^T Q_c C_c \leq 0$ dans ce cas. Nous ne considérerons pas ce cas particulier par la suite, nous nous focaliserons sur des situations rencontrées en pratique avec les solutions réelles. Pour un cylindre droit de hauteur infinie (ou non-définie) et de rayon constant, la décomposition de l'équation (3.26) montre que le contour apparent est un ensemble de deux droites (dites droites apparentes) (voir figure 3.13) représenté par le couple de vecteurs \mathbf{l}^- et \mathbf{l}^+ satisfaisant

$$\begin{cases} (\mathbf{l}^-)^T \mathbf{m} \equiv \{\mathbf{K}^{-T} (\mathbf{I} - \alpha [\mathbf{r}]_{\times}) \mathbf{w}\}^T \mathbf{m} = 0 \\ (\mathbf{l}^+)^T \mathbf{m} \equiv \{\mathbf{K}^{-T} (\mathbf{I} + \alpha [\mathbf{r}]_{\times}) \mathbf{w}\}^T \mathbf{m} = 0 \end{cases} \quad (3.27)$$

avec $\alpha = r_c / \sqrt{\|\mathbf{w}\|^2 - r_c^2}$, ou de manière équivalente par la matrice symétrique réelle (3×3) $\mathbf{C} = \mathbf{l}^- \mathbf{l}^{+T} + \mathbf{l}^+ \mathbf{l}^{-T}$ satisfaisant $\mathbf{m}^T \mathbf{C} \mathbf{m} = 0$ de rang 2 et définie à un facteur d'échelle près. Les deux représentations sont équivalentes pour définir le contour apparent à l'aide de 4 paramètres indépendants.

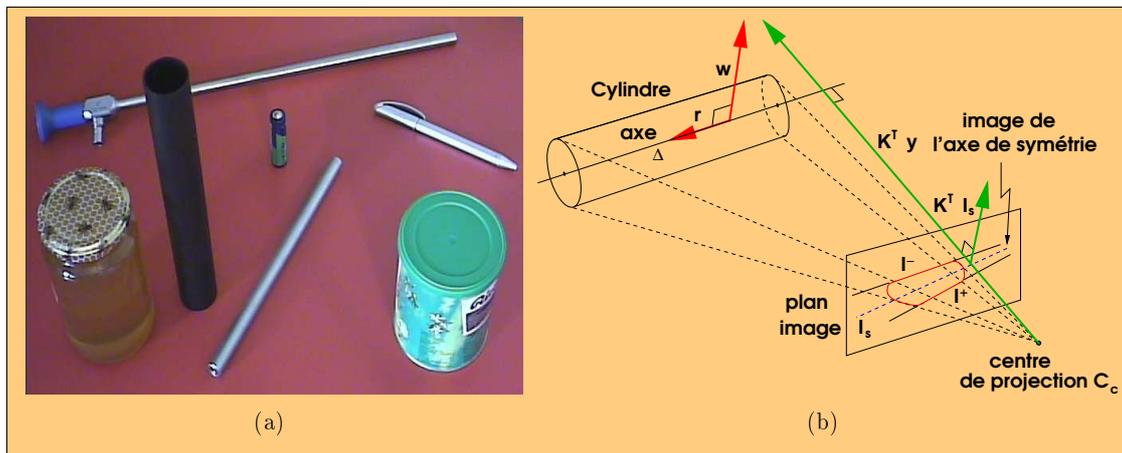


Fig. 3.13: (a) Quelques objets de forme cylindrique : laparoscope, stylo, pot de miel, tube de plastique noir, boîte de thé vert, tube métallique, pile. (b) Un cylindre de révolution et sa projection perspective $\mathbf{P}^c = \mathbf{K} \mathbf{P}$. La projection inverse des droites apparentes ($\mathbf{l}^-, \mathbf{l}^+$) est une paire de plans $(\mathbf{P}^c)^T \mathbf{l}^-$ et $(\mathbf{P}^c)^T \mathbf{l}^+$, passant par le centre de projection. L'image de l'axe du cylindre Δ est l'axe \mathbf{l}_s d'une homologie harmonique \mathbf{H} reliant les deux droites apparentes.

Nous énonçons quelques propriétés géométriques intéressantes entre les deux droites apparentes dans [DM07], comme le fait que les deux vecteurs \mathbf{l}^- et \mathbf{l}^+ sont reliés par une transformation géométrique homologique. Plusieurs solutions pour déterminer la matrice \mathbf{C} sont proposées (solution directe, optimisation numérique récursive). Les techniques sont inspirées de la géométrie épipolaire car la matrice \mathbf{C} partage de nombreuses propriétés avec la matrice fondamentale (exceptée la symétrie). L'intérêt de rechercher directement la matrice \mathbf{C} au lieu des deux droites séparément évite également de classer les pixels, selon qu'ils contribuent à une droite ou à une autre.

L'estimation des coordonnées de Plücker (\mathbf{r}, \mathbf{w}) de l'axe du cylindre est également détaillée dans [DM07]. Nous n'en rappelons ici que les aspects les plus importants. Étant donné une caméra calibrée et la matrice \mathbf{C} , la matrice de Plücker peut être estimée linéairement de la façon suivante. À partir des équations (3.27), la matrice \mathbf{C} peut être reliée aux paramètres de la pose par sa décomposition en valeurs singulières. D'une part, nous avons

$$\begin{aligned}
 \mathbf{K}^T \mathbf{C} \mathbf{K} &\equiv \mathbf{K}^T \left(\mathbf{l}^- \mathbf{l}^{+T} + \mathbf{l}^+ \mathbf{l}^{-T} \right) \mathbf{K} \\
 &\equiv (\alpha [\mathbf{r}]_{\times} - \mathbf{I}) \mathbf{w} \mathbf{w}^T (\alpha [\mathbf{r}]_{\times}^T + \mathbf{I}) \\
 &+ (\alpha [\mathbf{r}]_{\times} + \mathbf{I}) \mathbf{w} \mathbf{w}^T (\alpha [\mathbf{r}]_{\times}^T - \mathbf{I}) \\
 &\equiv \alpha^2 [\mathbf{r}]_{\times} \mathbf{w} \mathbf{w}^T [\mathbf{r}]_{\times} + \mathbf{w} \mathbf{w}^T \\
 &\equiv [\mathbf{r}]_{\times} \left(\alpha^2 \frac{\mathbf{w} \mathbf{w}^T}{\|\mathbf{w}\|^2} + \frac{[\mathbf{w}]_{\times} [\mathbf{w}]_{\times}^T}{\|\mathbf{w}\|^2} \right) [\mathbf{r}]_{\times}^T \\
 &\equiv [\mathbf{r}]_{\times} \left(\mathbf{I} - (1 - \alpha^2) \frac{\mathbf{w} \mathbf{w}^T}{\|\mathbf{w}\|^2} \right) [\mathbf{r}]_{\times}^T \\
 &\equiv \left(\mathbf{I} - \mathbf{r} \mathbf{r}^T - \mathbf{z} \mathbf{z}^T \right) \\
 &= \begin{bmatrix} \mathbf{a} & \mathbf{z}_u & \mathbf{r} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 - \sigma^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{a}^T \\ \mathbf{z}_u^T \\ \mathbf{r}^T \end{bmatrix}
 \end{aligned} \tag{3.28}$$

avec $\mathbf{z} = \frac{\sqrt{1-\alpha^2}}{\|\mathbf{w}\|} [\mathbf{r}]_{\times} \mathbf{w}$ et le vecteur unitaire $\mathbf{z}_u = \mathbf{z}/\sigma$. De l'autre, la décomposition SVD fournit

$$\mathbf{K}^T \mathbf{C} \mathbf{K} = \mathbf{U} \mathbf{D} \mathbf{U}^T = \mathbf{U} \text{diag}(\lambda_1, \lambda_2, 0) \mathbf{U}^T. \tag{3.29}$$

Il est facile d'observer que $\mathbf{U} = \begin{bmatrix} \mathbf{a} & \mathbf{z}_u & \mathbf{r} \end{bmatrix}$ et que

$$\sigma = \|\mathbf{z}\| = \sqrt{1 - \alpha^2} = \sqrt{1 - \frac{\lambda_2}{\lambda_1}}.$$

Finalement $\mathbf{w} \equiv \mathbf{z}_u \times \mathbf{r} = \mathbf{a}$ et en introduisant le rayon r_c du cylindre, pour l'obtention d'une reconstruction métrique, on a

$$\|\mathbf{z}\| = \sqrt{1 - \frac{r_c^2}{\|\mathbf{w}\|^2 - r_c^2}} \Rightarrow \|\mathbf{w}\| = r_c \sqrt{1 + \frac{\lambda_1}{\lambda_2}}. \tag{3.30}$$

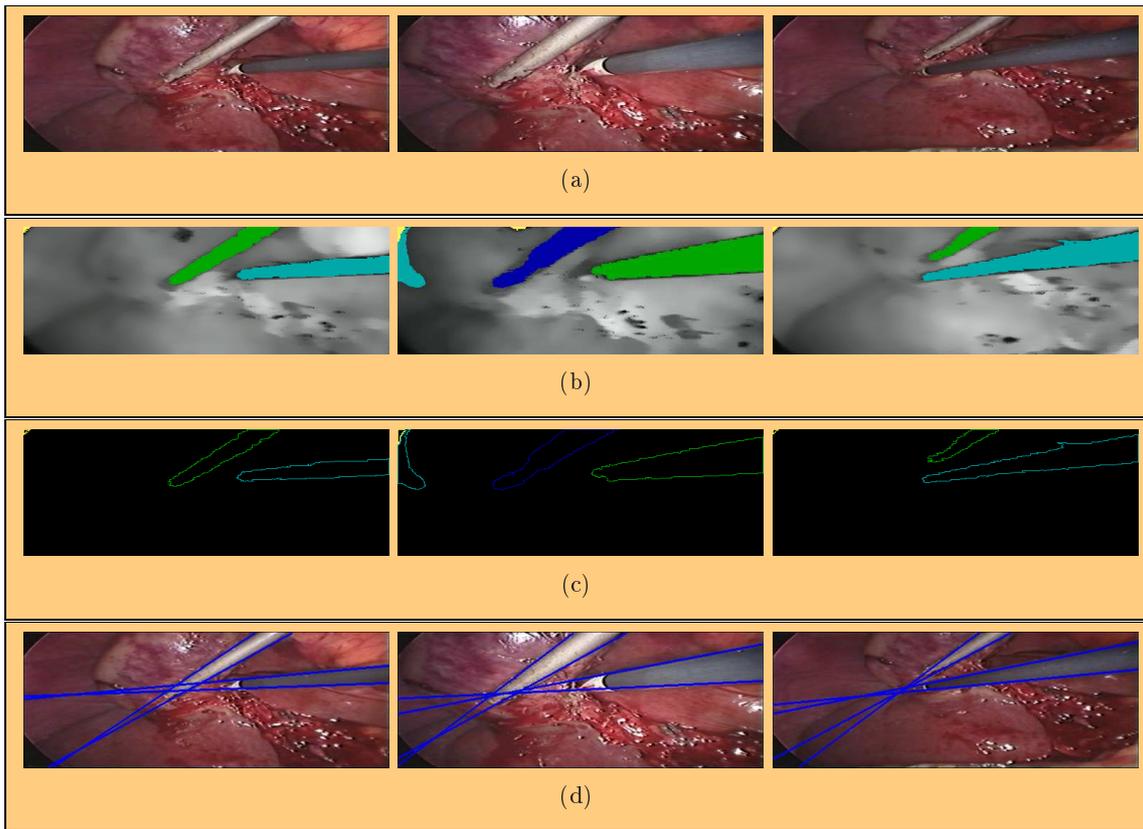


Fig. 3.14: Segmentation des contours apparents des instruments de laparoscopie à partir de trois images endoscopiques de la cavité abdominale. (a) Images originales (dé-entrelacées). (b) Résultats de l'expansion de régions. (c) Contours des régions retenues. (d) Approximation des contours apparents par une quadrique dégénérée représentée par la matrice C décrite dans ce paragraphe, pour les deux instruments.

Des résultats de simulation et à partir de d'images réelles dans un environnement structuré et finalement dans la cavité abdominale sont fournis dans [DM07]. Une étude comparative avec d'autres méthodes y est également décrite.

3.4.4 Contrainte de mouvement en chirurgie à invasion minimale

La forme cylindrique des instruments de laparoscopie nous a permis de vérifier la formulation des contraintes des mouvements en chirurgie à invasion minimale (figure 3.15). Une méthode décrite en détails dans [DNM06b] est très brièvement présentée ici. Une approche multi-vues est considérée pour estimer la position du point d'insertion, supposé immobile, et son image. Cette méthode fait appel à la localisation partielle ou complète de l'axe de symétrie de l'instrument, selon la technique décrite dans le paragraphe précédent.

Géométriquement, tout point \mathbf{X} appartenant à une droite représentée par sa matrice duale de Plücker L^* vérifie l'équation $L^* \mathbf{X} = \mathbf{0}$. Etant donné n positions successives $\{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n\}$ de l'axe de symétrie et correspondant à l'ensemble des matrices duales de Plücker $\{L_1^*, L_2^*, \dots, L_n^*\}$, l'intersection de ces droites convergentes est donnée par le noyau de (la transformation linéaire

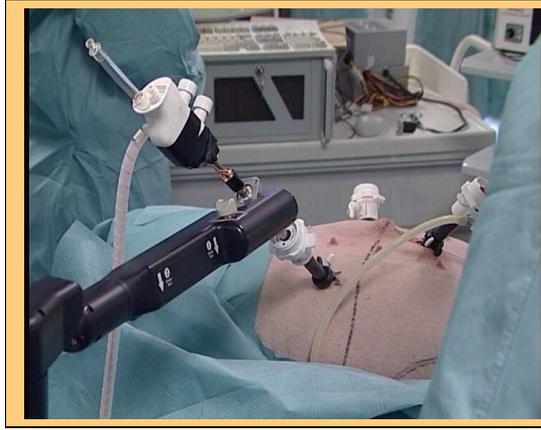


Fig. 3.15: Le robot AESOP 3000 dans la salle d'opération de l'IRCAD. Les trocars sont insérés dans les incisions faites sur la surface de l'abdomen pour guider un instrument ou un laparoscope (stationnaire ici) et éviter les frottements directs avec la peau.

représentée par) la matrice $(4n \times 4)$ G_n^T de rang 3, définie par

$$G_n = [L_1^*, L_2^*, \dots, L_n^*] . \quad (3.31)$$

Autrement dit, le noyau de la transformation G_n^T doit être de dimension 1 et l'intersection peut être calculée à partir de n positions 3-D différentes de l'instrument ($n \geq 2$). En décomposant en valeurs singulières la matrice G_n^T , on peut obtenir cette intersection qui est donnée par le vecteur singulier associé à la valeur singulière nulle.

Avec le modèle de la projection perspective, la projection dans l'image de l'axe de l'instrument est une droite représentée par le vecteur \mathbf{l}_s défini par rapport à une des coordonnées de Plücker ([HZ00])

$$[\mathbf{l}_s]_x = KP L (KP)^T = [K^{-T} \mathbf{w}]_x \Rightarrow \mathbf{l}_s \equiv K^{-T} \mathbf{w} . \quad (3.32)$$

Comme le vecteur \mathbf{l}_s est ainsi défini à un facteur d'échelle près, il ne dépend pas de l'amplitude du vecteur \mathbf{w} . Par conséquent, pour n positions successives, les n vecteurs $\mathbf{l}_{s1}, \mathbf{l}_{s2}, \dots, \mathbf{l}_{sn}$ correspondant aux n droites convergentes doivent satisfaire la relation

$$\begin{pmatrix} \mathbf{l}_{s1}^T \\ \mathbf{l}_{s2}^T \\ \vdots \\ \mathbf{l}_{sn}^T \end{pmatrix} i = \underbrace{\begin{pmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_n^T \end{pmatrix}}_{W_n} K^{-1} i = \mathbf{0} \quad (3.33)$$

et i est l'image du point d'insertion I . Il s'ensuit qu'un ensemble de n droites de l'espace est projeté selon n droites convergentes dans l'image si la matrice $(n \times 3)$ W_n est de rang 2. Ceci est seulement une condition nécessaire qui n'assure pas du tout que les droites dans l'espace sont convergentes mais simplement que leurs projections le sont. L'image du point d'insertion, supposé immobile (simplement perturbé par la respiration, voir la figure 3.17) joue un rôle important dans la détection et le suivi des instruments, qui s'en trouve contraint, donc

simplifié. L'apparition ou la réapparition d'un instrument après disparition du champ de vue doit être consistant avec la localisation de l'un des points d'insertion estimés. Nous proposons dans [DNM06b], une méthode robuste qui ne retient par la suite que les positions les plus cohérentes (voir la figure 3.16). Des expériences, à l'aide d'une "training box" et *in vivo* ont permis de valider l'approche proposée, sur un faible nombre d'images.

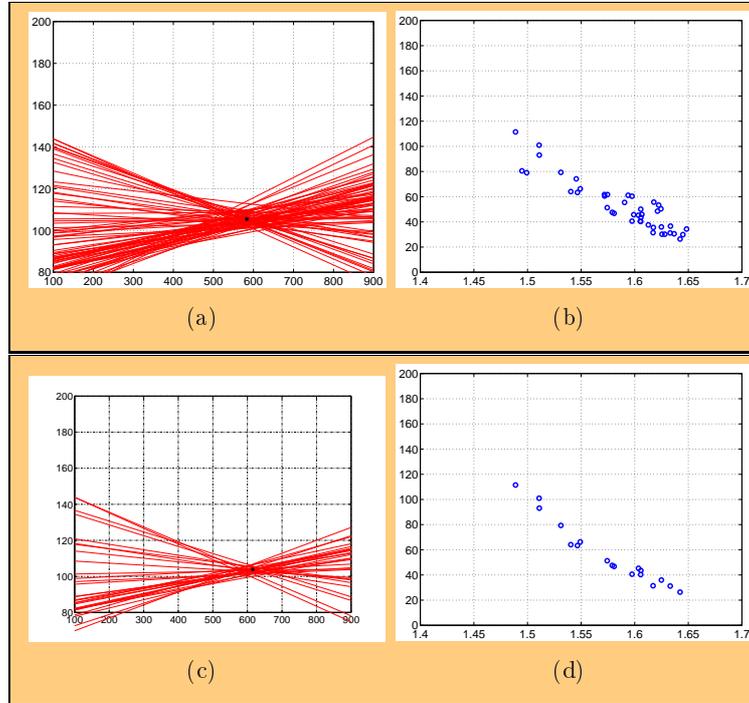


Fig. 3.16: (a) L'intersection des images de l'axe de symétrie d'un instrument en mouvement et estimation de l'image du point d'insertion en (593.4; 105.5) (croix noire) par les moindres carrés (SVD) et à partir d'expériences *in vitro*. (b) Dans l'espace des paramètres (θ, ρ) , les "points" (bulles bleues) doivent être colinéaires. (c) Estimation de l'image du point d'insertion en (615.5; 103.9) (croix bleue) et l'espace des paramètres (d) avec une méthode robuste qui rejette 50 % des différentes positions de l'instrument.

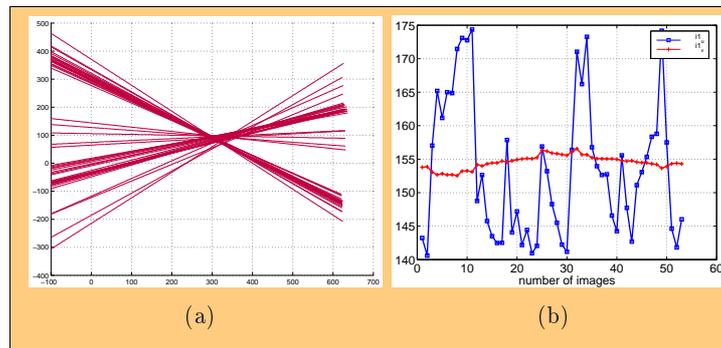


Fig. 3.17: (a) Deux des trois coordonnées du point d'insertion durant le guidage d'un des instruments. (b) Les variations temporelles des coordonnées de l'image du point d'insertion *in vivo* et pour une séquence brève de 52 images (≈ 2 secondes).

3.5 GUIDAGE VISUEL D'UN INSTRUMENT DE LAPAROSCOPIE

Dans cette section, nous présentons trois applications du guidage par vision dans un environnement de laparoscopie. Les deux premières techniques mettent en jeu des marqueurs artificiels et considèrent la configuration *eye-to-hand*. La dernière technique, récente, présente un guidage où l'endoscope rigide est monté sur le bras du robot (configuration *eye-in-hand*) et s'intéresse à la navigation par asservissement visuel sans modèle de la scène abdominale.

3.5.1 Guidage et navigation avec un instrument marqué

Nous présentons ici une application du recalage limité dans un premier temps à l'estimation d'une distance (euclidienne) entre deux objets indépendants et en mouvement. Il s'agit plus précisément de la distance entre un instrument de chirurgie laparoscopique et un organe interne dans l'abdomen (d'un porc). Ceci nous permet d'une part d'illustrer une technique de guidage dans un milieu complexe, en mettant en oeuvre des éléments structurants et des méthodes dédiées à la navigation visuelle et la manipulation contrôlée à l'intérieur du corps humain. D'autre part, cela nous permet d'aborder ce type de retour visuel particulier et ses spécificités pour l'élaboration d'un asservissement visuel en synthétisant une partie des travaux de thèse d'Alexandre Krupa [Kru03].



Fig. 3.18: En chirurgie laparoscopique robotisée, plusieurs instruments sont montés sur des bras robotiques. Ici, le robot AESOP 3000 de l'ex société Computer Motion (Intuitive Surgical). Le laparoscope (à droite) est monté sur une caméra qui peut être fixe ou montée sur un robot porte-endoscope.

Comme la chirurgie laparoscopique est une technique à invasion minimale, des petites incisions sont réalisées sur le corps du patient pour atteindre le siège de l'opération. Un trocart est alors inséré : c'est un accessoire (figure 3.19-a) qui permet aux instruments et au laparoscope de ne pas être en contact avec la paroi, de servir de guide et d'assurer l'étanchéité (quand

ils sont insérés dans la cavité abdominale, celle-ci est gonflée au dioxyde de carbone pour décoller légèrement les organes, la pression du gaz est alors réglée pour maintenir un volume constant). Un laparoscope est un tube cylindrique métallique droit ou coudé, d'environ 10 mm de diamètre et d'environ 25-35 cm de long (voir la figure 3.19b-c). Cet endoscope rigide est employé à la fois pour acheminer la lumière qui illuminera la scène (la cavité abdominale dans notre cas) et d'autre part comme système optique pour délivrer une vision monoculaire, voire binoculaire ; l'extrémité externe est donc montée sur une caméra immobile (voir la figure 3.18) ou bien embarquée sur un robot.

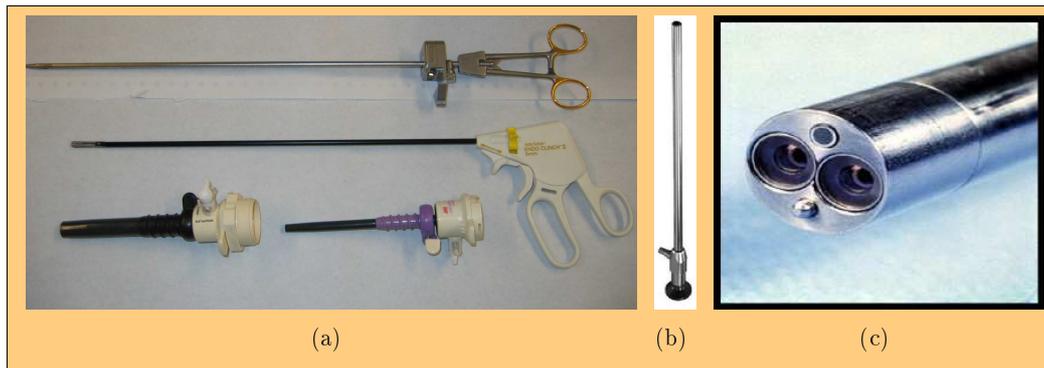


Fig. 3.19: (a) Exemple d'instruments manuels de laparoscopie (en haut). Deux trocars de diamètres 5 et 10 mm (en bas). (b) Laparoscope droit. (c) Le stéréo-laparoscope du système Da Vinci (Intuitive Surgical).



Fig. 3.20: Un instrument de chirurgie de laparoscopie composé de trois marqueurs optiques et projetant 4 faisceaux laser parallèles. Ce dispositif permet d'estimer la distance entre l'extrémité de l'instrument et l'organe pointé.

Quand on aborde le problème de l'estimation en temps réel d'une position 3-D entre un organe humain et un instrument de chirurgie à invasion minimale à l'aide d'un capteur visuel endoscopique, plusieurs difficultés doivent être surmontées. Dans le contexte de la coelioscopie, nous avons affaire à une scène très peu structurée, soumise à des variations de

lumière, et dont l'arrière-plan est en mouvement et présente des déformations. De par la nature des tissus, des spécularités peuvent apparaître également dans l'image, rendant difficile l'interprétation de celle-ci en vue d'effectuer un guidage. Deux pistes ont été explorées, d'une part le développement de méthodes de segmentation spécifiques à l'emploi des images laparoscopiques (voir aussi dans [Gan04]) et d'autre part la possibilité d'ajouter des marqueurs afin de simplifier la détection des régions d'intérêt. Pour l'application de navigation que nous avons proposé dans la thèse d'Alexandre Krupa [Kru03], nous nous sommes focalisés surtout sur la seconde issue. Ainsi, un instrument de laparoscopie a été adapté pour intégrer 4 faisceaux laser et a de plus été marqué à l'aide d'un ensemble de trois LEDs alignées et fixées sur sa surface (figure 3.20). Si les faisceaux laser nous permettent d'estimer l'orientation de l'instrument (sauf la rotation propre, bien sûr !) et d'offrir au chirurgien la possibilité de réintégrer l'instrument dans le champ de vision, l'apport des LEDs concerne la possibilité d'asservir également la distance entre un organe et l'extrémité de l'instrument.

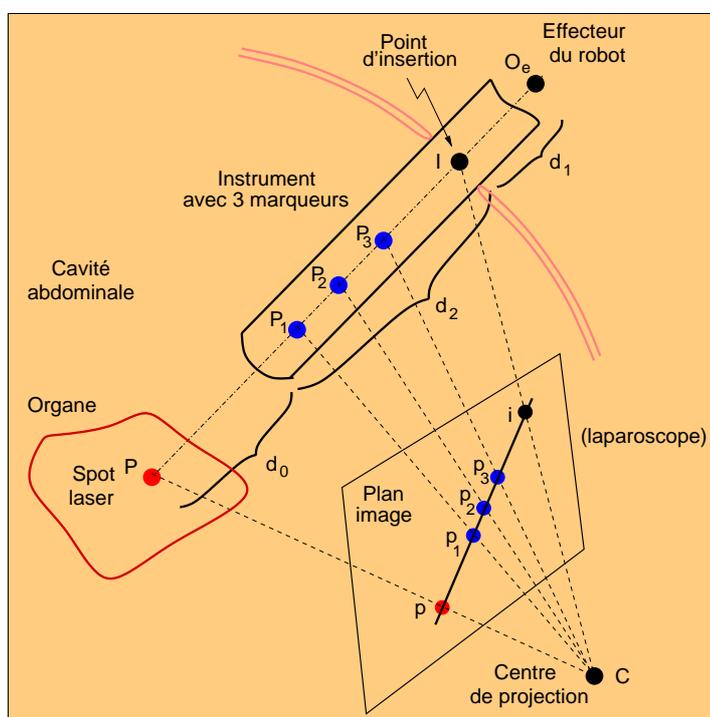


Fig. 3.21: Modélisation géométrique l'instrument de chirurgie marqué et de l'organe pointé par les faisceaux lasers embarqués.

Un tel dispositif peut être schématisé selon l'illustration de la figure 3.21, où nous avons représenté conjointement la configuration des marqueurs optiques, la projection de la lumière sur un organe et leurs projections respectives dans le plan image. Plusieurs hypothèses sont faites pour aboutir à cette simple modélisation. En premier lieu, nous ne nous intéressons qu'aux centres des marqueurs optiques que nous supposons plans, alignés entre eux et alignés avec le centre du quadrilatère formé par les projections des faisceaux laser sur l'organe ciblé. En second lieu, alors que sur un plan métallique, les projections des 4 faisceaux forment 4 taches bien distinctes (figure 3.20), dans l'abdomen il en est autrement. Avec la longueur d'onde choisie (autour de 670 nm), l'interaction de la lumière monochromatique rouge avec la surface

des organes provoque une diffusion qui rend délicate voire impossible la distinction entre les projections (voir les figures 3.22 et 3.23). Il s'ensuit à la fois une perte d'information et un manque de précision de localisation. Cet effet est amoindri avec d'autres longueurs d'onde, qui sont donc plus appropriées. Le choix de la couleur est à rechercher au niveau de l'émission et de l'aspect économique : les lasers rouge sont moins onéreux que les autres d'une manière générale, et nous en avons 4 ici. Nous supposons en outre que la surface sur laquelle sont projetés les faisceaux laser est localement plane. Cette hypothèse est assez contraignante et rarement vérifiée. Elle devrait aboutir à la visibilité de l'ensemble des faisceaux (ou d'aucun d'entre eux). Mais du fait de la présence de courbures fortes et des (auto-)occultations des organes, un ou plusieurs faisceaux peut aboutir sur une surface éloignée des autres.

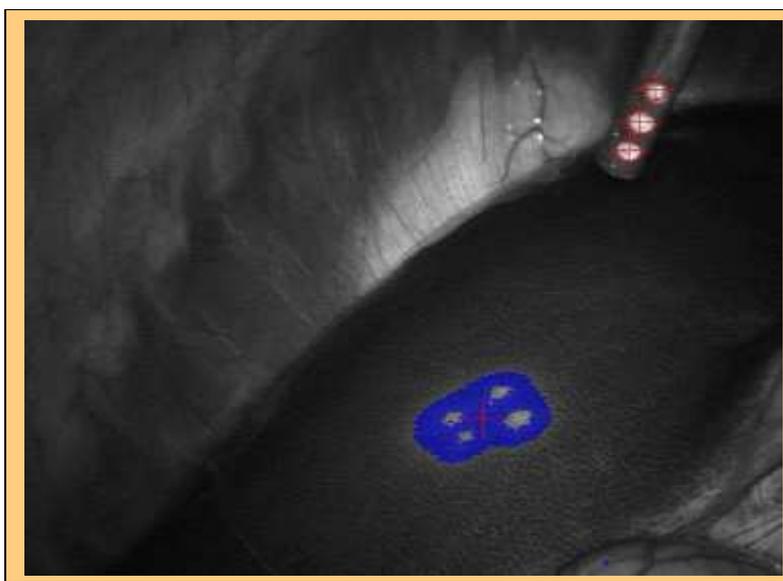


Fig. 3.22: Localisation du motif laser et des marqueurs lumineux dans les images à niveaux de gris [Kru03].

Néanmoins, l'alignement supposé entre les 3 centres des marqueurs implique l'alignement des projections perspectives dans le plan image. Cela correspond à mettre en correspondance deux bases d'un espace projectif mono-dimensionnel. L'alignement supplémentaire avec le centre de la "tache" laser signifie que son image est également alignée avec celles des marqueurs. Etant donné la transformation projective reliant les deux bases (homographie de droites), la distance entre la projection de l'organe pointé et l'origine de la base projective formée par les centres des marqueurs est reliée à la coordonnée projective de l'image de l'organe pointé dans la base projective formée des images des centres des marqueurs. Cette distance peut donc être calculée à l'aide du birapport (ou rapport anharmonique), un invariant projectif bien connu (voir par exemple dans [Hor86, HS92, FMZ⁺91, May92a, Mor93]). Finalement, en asservissant la coordonnée projective dans l'image, la distance entre l'organe et l'instrument peut être contrôlée.

Une méthode robuste pour l'extraction des informations visuelles de l'image a été mise en oeuvre. Cette technique, originale, permet de détecter le motif laser et les marqueurs lumineux. Elle repose sur le principe de la synchronisation de l'allumage et l'extinction du laser et des

marqueurs lumineux avec l'acquisition des trames paires et impaires de l'image (signal vidéo CCIR, entrelacé). Cette synchronisation a pour effet de créer dans l'image entrelacée des motifs striés correspondant au laser et aux marqueurs lumineux qui sont aisément détectables en appliquant un filtrage passe-haut sur l'image. La position dans l'image du centre du motif laser est ensuite obtenue en calculant le barycentre du motif laser détecté. Les coordonnées des marqueurs lumineux sont obtenues en utilisant un algorithme d'approximation elliptique des contours (voir la figure 3.22).

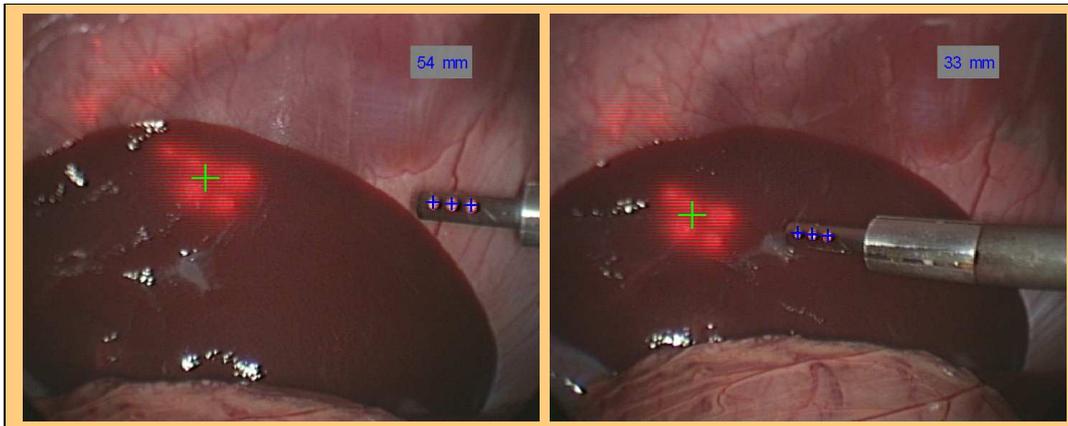


Fig. 3.23: La projection des 4 faisceaux laser sur l'organe ciblé (croix jaune) et les 3 marqueurs ajoutés sur la surface de l'instrument (croix bleues) permettent, en conjonction avec leurs projections perspectives dans l'image, de calculer un birapport. L'invariance projectif du birapport est utilisée pour estimer la distance entre l'extrémité de l'instrument pointé (valeur affichée en haut à droite).

Récupération et centrage par asservissement visuel

L'objectif de l'asservissement visuel ici est d'orienter l'instrument chirurgical autour des axes perpendiculaires au trocart de manière à amener la projection image du motif laser à une position indiquée sur l'image endoscopique. Le point P (voir figure 3.21) qui correspond au point d'intersection entre l'axe de l'instrument et la scène et sa projection p sur le plan image de la caméra sont utilisés à cette fin. Nous posons l'hypothèse que les coordonnées du barycentre de l'image du motif laser correspondent aux coordonnées image du point d'intersection entre l'axe de l'instrument et la scène. L'asservissement consiste alors à amener le point p sur une cible spécifiée dans l'image endoscopique. Pour établir la loi de commande, nous avons utilisé l'approche par fonction de tâche introduit par Samson *et al.* [SBE91]. Les primitives visuelles à asservir correspondent aux coordonnées dans l'image $\mathbf{s}_p = [u_p \ v_p]^T$ du point p . Le vecteur d'état $\mathbf{s}(t)$ contenant les informations visuelles courantes est par conséquent directement défini par les coordonnées dans l'image \mathbf{s}_p du barycentre du motif laser

$$\mathbf{s}(t) = \mathbf{s}_p = [u_p \ v_p]^T . \quad (3.34)$$

La consigne est $\mathbf{s}_p^* = [u_p^* \ v_p^*]^T$ où u_p^* et v_p^* sont les coordonnées image de la cible à atteindre. Dans notre système, la caméra endoscopique n'est pas solidaire de l'instrument chirurgical dont nous cherchons à commander les déplacements. Nous nous trouvons dans la configuration où

la caméra est déportée. Dans l'hypothèse où la caméra reste immobile durant l'asservissement visuel, les variations des coordonnées dans l'image du barycentre du motif laser dépendent de la vitesse de l'instrument et du mouvement propre de la scène dû à la respiration du patient. La variation du vecteur de mesure \mathbf{s}_p qui en résulte peut être mise sous la forme :

$$\dot{\mathbf{s}}_p = \mathbf{L}_s \mathbf{w}_{op} + \frac{\partial \mathbf{s}_p}{\partial t} \quad (3.35)$$

où, en utilisant le formalisme de la fonction de tâche :

- \mathbf{L}_s est une matrice de dimension 2×4 qui relie la variation des informations visuelles à la vitesse opérationnelle de l'instrument chirurgical $\mathbf{w}_{op} = [v_z \ \omega_x \ \omega_y \ \omega_z]^T$ et contenant les vitesses de pénétration v_z et le vecteur vitesse instantanée de rotation $\boldsymbol{\Omega} = [\omega_x \ \omega_y \ \omega_z]^T$.
- $\frac{\partial \mathbf{s}_p}{\partial t}$ est la variation propre de \mathbf{s}_p introduite par le mouvement de l'organe sur lequel est projeté le motif laser, et un éventuel mouvement de l'endoscope. Les mouvements de l'organe et de l'endoscope sont principalement dus à la respiration du patient.

Etant donné que le centre de la projection du motif laser correspond au point d'intersection P de la droite prolongeant l'axe de l'instrument avec la surface de l'organe, la translation et la rotation de l'instrument le long et autour de l'axe de rotation propre du trocart ne modifient pas la position du point P . Par conséquent la variation de \mathbf{s}_p est indépendante des composantes v_z et ω_z de la vitesse opérationnelle de l'instrument. Ce découplage se traduit par la présence de zéros sur la première et la quatrième colonne de la matrice \mathbf{L}_s

$$\mathbf{L}_s = \begin{bmatrix} 0 & J_{\omega 11} & J_{\omega 12} & 0 \\ 0 & J_{\omega 21} & J_{\omega 22} & 0 \end{bmatrix} \quad (3.36)$$

Seuls les 2 degrés de liberté d'orientation l'instrument chirurgical autour des axes x et y du repère du trocart permettent de positionner le motif laser dans l'image. La variation du vecteur de mesure \mathbf{s}_p peut alors être mise sous la forme :

$$\dot{\mathbf{s}}_p = \mathbf{J}_\omega \begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix} + \frac{\partial \mathbf{s}_p}{\partial t} \quad (3.37)$$

où $\mathbf{J}_\omega = \begin{bmatrix} J_{\omega 11} & J_{\omega 12} \\ J_{\omega 21} & J_{\omega 22} \end{bmatrix}$ est la matrice Jacobienne qui permet d'exprimer la variation de \mathbf{s}_p en fonction seulement des vitesses de rotation ω_x et ω_y de l'instrument chirurgical.

Pour aligner le motif laser sur une cible définie dans l'image, l'asservissement consiste à réguler vers zéro la fonction de tâche $\mathbf{e} = \mathbf{C} (\mathbf{s}_p - \mathbf{s}_p^*)$ de dimension $m = 2$, où \mathbf{C} est une matrice de combinaison permettant de prendre en compte plus d'informations que de degrés de liberté nécessaires pour réaliser la tâche. Etant donné que le vecteur de mesure \mathbf{s}_p est de dimension $k = 2 = m$, la matrice de combinaison \mathbf{C} est choisie égale à l'identité de manière que le comportement de \mathbf{s}_p soit similaire à celui de \mathbf{e} . Nous avons alors $\mathbf{e} = (\mathbf{s}_p - \mathbf{s}_p^*)$ et en utilisant la commande cinématique, décrite dans [Kru03], qui garantit la décroissance exponentielle de la fonction de tâche \mathbf{e} , nous obtenons la loi de commande suivante :

$$\begin{bmatrix} \dot{\omega}_x^* \\ \dot{\omega}_y^* \end{bmatrix} = -\hat{\mathbf{J}}_\omega^{-1} (\lambda \mathbf{e} + \frac{\partial \mathbf{e}}{\partial t}) \quad (3.38)$$

où :

- \widehat{J}_ω est la valeur approchée de la matrice d'interaction J_ω .
- λ est un gain positif définissant l'inverse de la constante de temps de la décroissance exponentielle de e .
- $\frac{\widehat{\partial e}}{\partial t}$ est une estimation de la variation propre de la fonction de tâche, qui, ici peut se mettre sous la forme

$$\frac{\widehat{\partial e}}{\partial t} = \frac{\widehat{\partial s_p}}{\partial t} - \dot{s}_p^* , \quad (3.39)$$

avec

- $\frac{\widehat{\partial s_p}}{\partial t}$, une estimation de la variation propre de s_p introduite par le mouvement de l'organe et un éventuel mouvement de la caméra,
- \dot{s}_p^* est la dérivée temporelle du signal de consigne.

Finalement, nous pouvons exprimer la commande cinématique en fonction des informations visuelles par l'équation suivante :

$$\begin{bmatrix} \omega_x^* \\ \omega_y^* \end{bmatrix} = \widehat{J}_\omega^{-1} \left(\lambda (s_p^* - s_p) + \dot{s}_p^* - \frac{\widehat{\partial s_p}}{\partial t} \right) \quad (3.40)$$

Nous avons considéré $\dot{s}_p^* = 0$ pour centrer le motif laser sur une cible fixe dans l'image par asservissement visuel 2-D. Généralement si aucune estimation du mouvement propre de

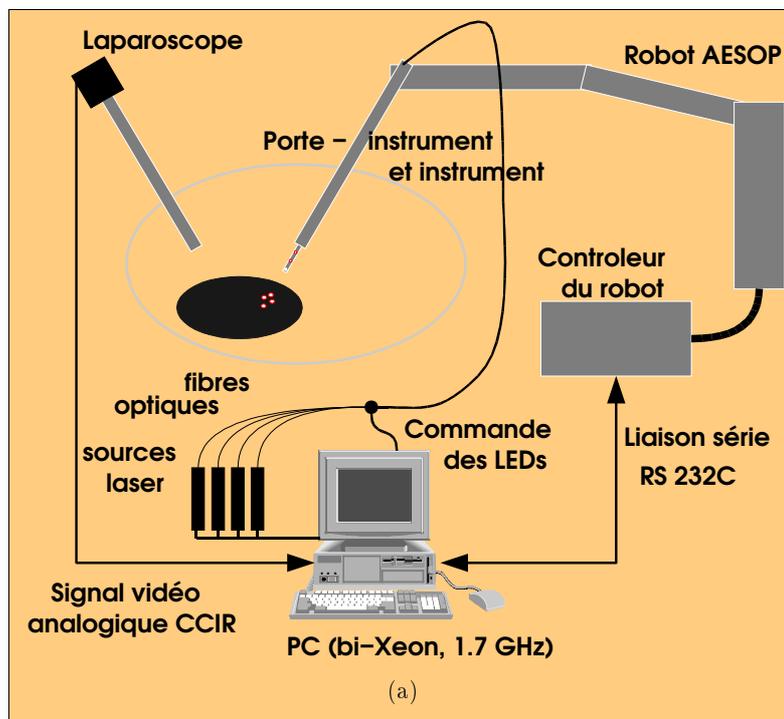


Fig. 3.24: Composants du système robotique réalisant l'asservissement visuel d'un instrument de laparoscopie dans la cavité abdominale.

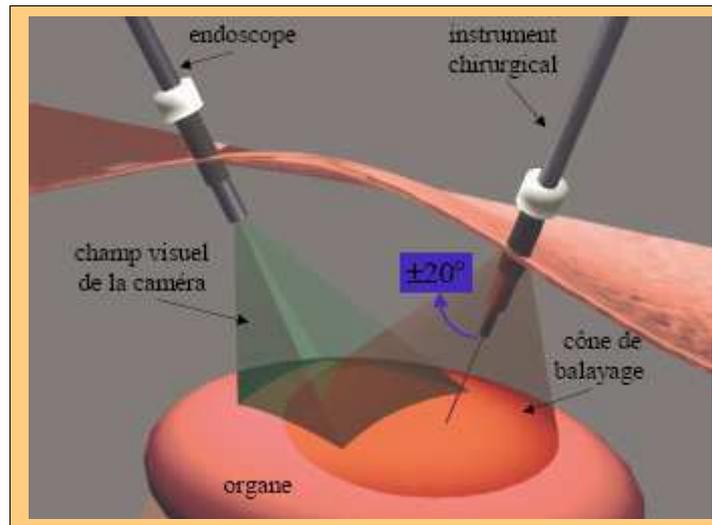


Fig. 3.25: Cône de balayage durant la procédure de récupération de l'instrument, antérieure à celle du centrage dans l'image.

l'instrument est disponible, le terme $\frac{\partial s_p}{\partial t}$ n'est pas pris en compte dans la loi de commande. De plus, pour réaliser l'asservissement, il est nécessaire de déterminer une approximation de la matrice J_ω . Deux méthodes d'estimation numérique (en boucle ouverte et adaptative tout au long de l'asservissement) sont présentées dans [Kru03].

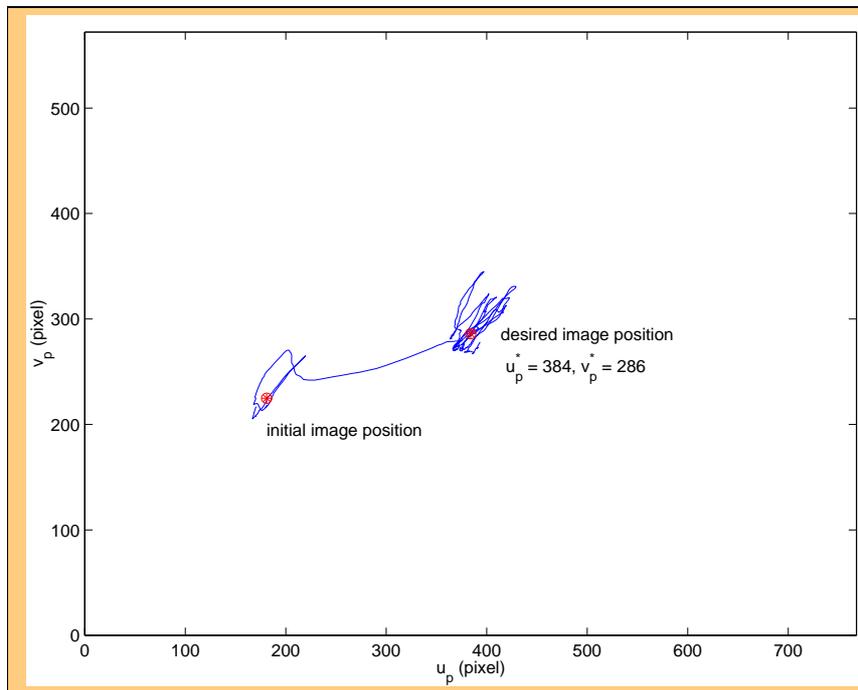


Fig. 3.26: Trajectoires dans l'image du centre des spots laser durant l'asservissement visuel effectué en salle d'opération à l'IRCAD. Les perturbations constatées aux voisinage de la position désirée proviennent de la respiration du porc.

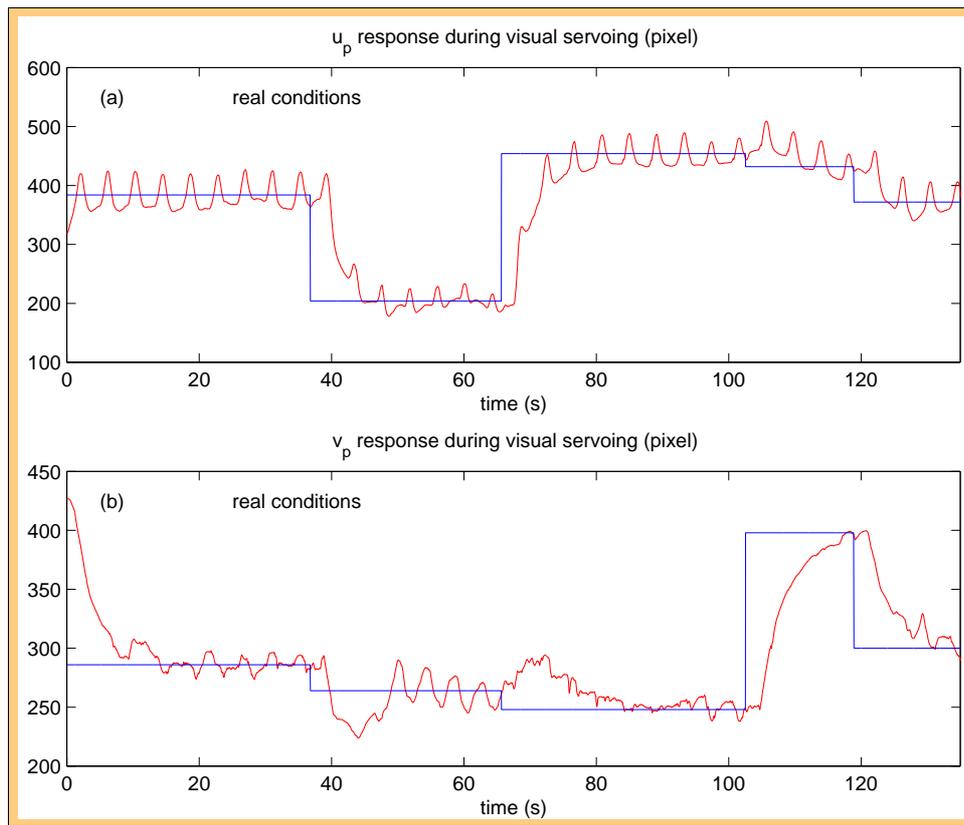


Fig. 3.27: Réponses temporelles des coordonnées $(u_p(t), v_p(t))$ du centre du spot laser dans l'image durant des asservissements visuels successifs (pour différentes positions à atteindre, sélectionnées par le chirurgien).

Les figures 3.26 et 3.27 présentent les résultats expérimentaux, quand l'asservissement visuel est mis en oeuvre sur un proc, dans la salle d'opération de l'IRCAD. Ces résultats attestent du fonctionnement correct de l'asservissement 2-D et montrent également comment la précision de localisation est perturbée par la respiration.

Tous les résultats concernant l'organigramme de séquençage pour la récupération de l'instrument, le schéma de commande pour le centrage présenté ici, la prise en compte et l'asservissement de la distance organe-instrument et l'étude de la stabilité en supposant un modèle ellipsoïdal de la cavité abdominale sont présentés dans [Kru03]. Le schéma de commande décrit ici pour seulement deux degrés de liberté permet de se rendre compte du nombre important de termes incertains ou exprimant les perturbations possibles avec un tel environnement.

Un nombre important de publications est en rapport avec ce travail. En premier lieu, des communications lors de congrès internationaux parmi lesquels IEEE ICRA 2002 [KJM⁺02], IEEE/RSJ IROS 2002 [KDG⁺02], MICCAI (2001 et 2002) [KMD⁺01, KMD⁺02], et surtout un article dans la revue internationale IEEE Transactions on Robotics and Automation [KGD⁺03].

3.5.2 Assistance à la suture

En s'attaquant à la réalisation d'une tâche particulière en interaction avec les tissus, nous avons souhaité progresser dans l'assistance à la thérapie en robotique médicale. Dans la thèse de Florent Nageotte [Nag05], je me suis intéressé avec lui à la modélisation géométrique et cinématique du couple d'instruments articulés (aiguille circulaire, porte-aiguille), montés sur un robot de chirurgie laparoscopique, afin de formaliser les déplacements et d'étudier l'assistance informatique et robotique de la suture.

Une tâche de suture robotisée par asservissement visuel a ainsi pu être mise en oeuvre *in vitro*. Celle-ci a été décomposée en une phase d'approche (vers un point d'entrée à la surface des tissus) et la phase de d'enfoncement et de passage de l'aiguille à travers les tissus (*stitching*). La phase d'approche a été réalisée par un calcul de pose basé sur un asservissement visuel virtuel 2-D combinant différentes primitives géométriques, provenant des marqueurs collés sur l'instrument, des contours du porte-aiguille et de l'aiguille. Dans [Nag05], la segmentation de ces différents éléments a été effectuée, en temps réel vidéo. Nous présentons ici brièvement les techniques mises en oeuvre pour la segmentation et la phase d'approche du guidage.

Segmentation de l'aiguille

L'information visuelle que nous avons proposé d'extraire pour permettre un calcul de pose est le squelette de l'image de l'aiguille. Dans le cas idéal où l'aiguille n'est pas tenue par le porte-aiguille, bien visible et pas affectée par des spécularités, une seule région sera extraite comme dans le cas des deux images de test (figures 3.28-a,b et 3.29-a,b). Mais, la plupart du temps, et

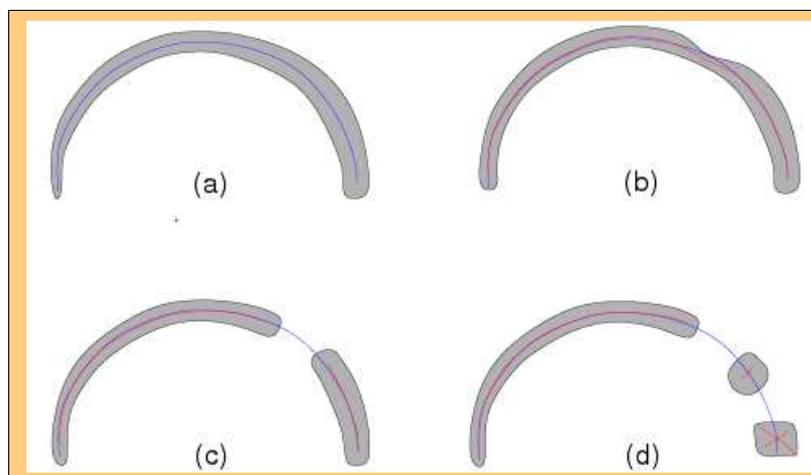


Fig. 3.28: Quatre exemples de segmentation et les squelettes associés. En bleu le squelette idéal et en rouge le squelette réel (a) l'aiguille extraite et son squelette dans le cas idéal. (b) La segmentation arrondit l'extrémité de l'aiguille et déforme le corps. Le squelette est tronqué au bout et déformé. (c) L'extrémité est arrondie et deux régions sont extraites. Deux squelettes tronqués sont obtenus. (d) Comme pour (c) plusieurs squelettes sont obtenus, mais deux régions de petite taille donnent des squelettes dont l'information n'est pas pertinente et doivent donc être rejetées.

notamment dans les cas pour lesquels l'aiguille est tenue par le porte-aiguille, plusieurs régions déconnectées sont extraites. Nous avons calculé le squelette morphologique de chacune des régions obtenues, à condition que leur taille soit significative. En effet, les régions de petite taille ne sont que peu représentatives de la forme globale de l'aiguille. Les effets de bord dans le calcul du squelette sont alors importants et le squelette obtenu ne contient pas d'information pertinente, comme le montre la figure 3.28-d. A partir des squelettes partiels obtenus et conservés, nous avons recherché une ellipse approchante globale. Afin de rendre l'ellipse finale robuste aux erreurs de segmentation, nous avons utilisé une méthode d'estimation de paramètres robuste basée sur les moindres carrés médian [RL87] ; des sous-ensembles de 5 points ont été sélectionnés aléatoirement parmi les points constituant les squelettes calculés. A partir de ces points nous avons pu calculer les paramètres de la meilleure ellipse approchante [Zha97].

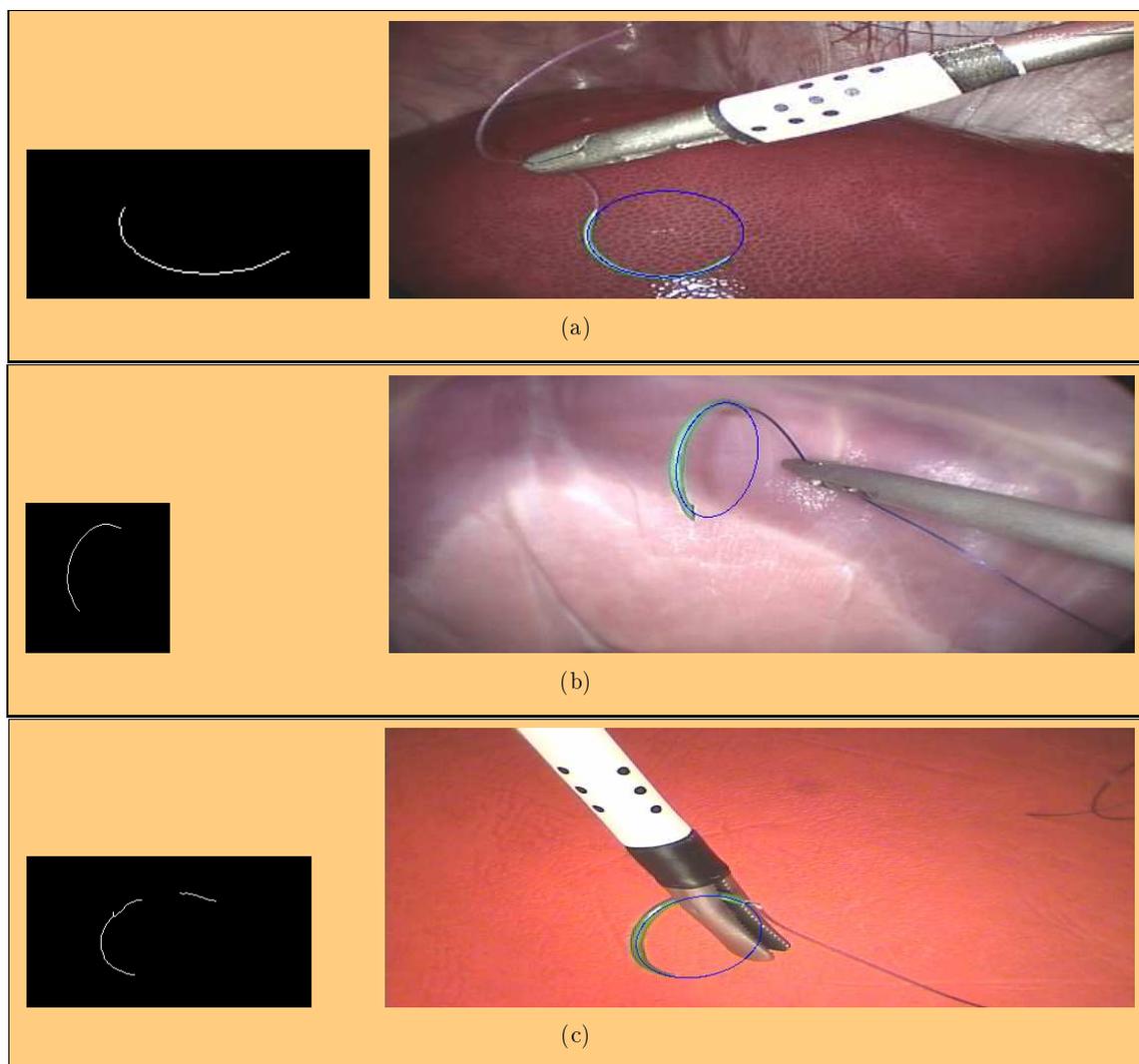


Fig. 3.29: (A gauche) Le squelette extrait à partir des zones vertes segmentées et (à droite) l'ellipse obtenue par estimation robuste dans des cas sans occultation (images de test) et avec occultation (image de laboratoire).

Dans le but de calculer la pose complète de l'aiguille (six degrés de liberté), il est nécessaire de pouvoir déterminer la rotation de l'aiguille autour de son axe. Pour cela nous avons utilisé la position de la pointe de l'aiguille. Celle-ci a été déterminée en recherchant les discontinuités de couleur le long de l'ellipse complète extraite. A cette fin, tous les types de transitions ont été répertoriés [Nag05].

On peut apprécier sur la figure 3.29 les résultats de la segmentation de l'aiguille circulaire (peinte en vert) et utilisée pour la suture en chirurgie laparoscopique.

Segmentation du porte-aiguille

Par segmentation du porte-aiguille, nous entendons l'extraction de points de marquage visibles et des droites apparentes du cylindre (ce dernier point a déjà été abordé). Le marqueur utilisé pour cet instrument est composé de points noirs sur un fond blanc. Les méthodes proposées pour la détection sont donc directement dérivées des techniques de segmentation d'images en niveaux de gris. Les marqueurs noirs étant de taille modeste dans les images endoscopiques, l'idée traduite ici est de commencer par la détection du fond de marquage blanc dont les di-

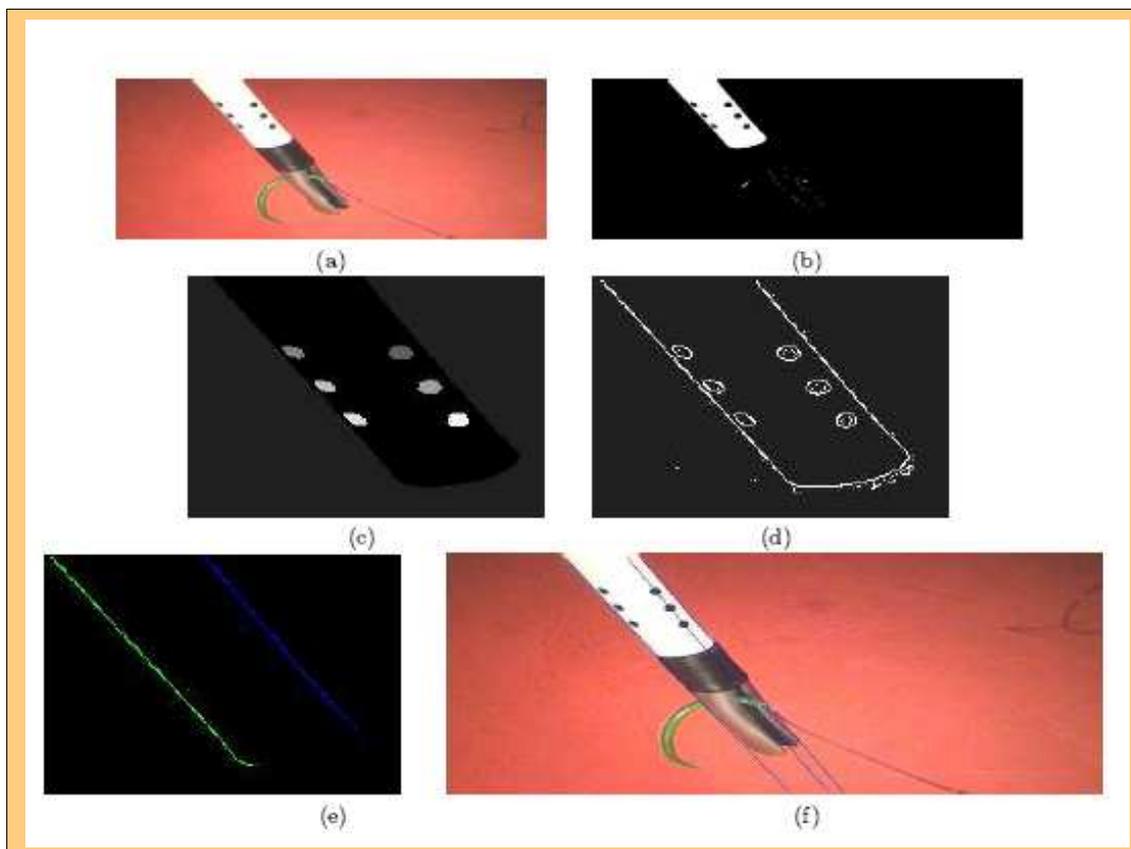


Fig. 3.30: Étapes de la segmentation du porte-aiguille. (a) Image originale de laboratoire. (b) Image binaire après seuillage de l'intensité. (c) Marqueurs après classification. (d) Contours obtenus par filtrage de Sobel, (e) Contours conservés sur le pourtour de la zone blanche et classifiés par rapport au triplet de marqueurs, (f) Droites approchantes après estimation robuste.

mensions permettent de le différencier aisément des autres zones indésirables éventuellement segmentées. L'analyse de l'histogramme de luminance des images permet de choisir assez simplement un seuil pour segmenter les zones alors quasi saturées. En pratique nous procédons par un seuillage par hystérésis qui rend la segmentation de la zone de marquage blanc peu sensible au choix des seuils. Après l'opération de seuillage, outre le marquage blanc, les éventuelles zones de pixels saturés correspondant aux spécularités ont également été extraites de l'image. En pratique, il est suffisant de conserver uniquement la région de plus grande taille qui correspond presque toujours à la zone de marquage. Cependant, si deux zones sont de tailles comparables, le calcul des moments d'ordre deux des régions détectées permet de discriminer la zone de marquage (rectangulaire) des zones spéculaires qui présentent le plus souvent une autre forme. Lorsque la zone du marqueur est détectée, on recherche à l'intérieur les marqueurs circulaires. Cette opération est réalisée par simple seuillage, le contraste étant généralement suffisant. Les régions noires extraites sont ensuite étiquetées et classifiées et leurs centres d'inertie sont calculés.

Positionnement 3-D du porte-aiguille par AVV

Lorsque l'aiguille n'est pas tenue par le porte-aiguille, l'estimation de la pose de chacun des instruments doit être effectuée séparément. A partir de l'ellipse image, il est possible d'obtenir deux solutions de l'attitude réelle du cercle supportant l'aiguille (5 degrés de liberté reconstruits). La position du centre de l'aiguille est identique pour ces deux poses, mais l'orientation est différente. Des données supplémentaires sont nécessaires pour lever l'indétermination. Cette information peut être apportée soit par un mouvement connu de la caméra (dans l'esprit de ce que proposent Taylor *et al.* pour estimer la profondeur dans les images endoscopiques [TFE⁺95]), soit par l'utilisation de deux caméras dont la position relative est connue (vision stéréoscopique), soit par un mouvement connu de l'aiguille. La dernière solution n'est pas envisageable puisque l'aiguille n'est pas commandée. La première peut s'appréhender de différentes façons : si l'endoscope est tenu par un bras robotique, le mouvement peut être mesuré à l'aide des codeurs donnant la position des axes ; si le mouvement de l'endoscope ne peut pas être mesuré directement, il faut disposer d'un objet connu dans la scène pour l'estimer. Si l'aiguille et l'instrument sont visibles simultanément, l'instrument peut servir de référence pour estimer le mouvement de la caméra et ainsi obtenir la solution valide de la pose de l'aiguille. La pose de l'instrument (4 degrés de liberté) peut être calculée à partir des droites apparentes et d'un point connu (3 ddl contraints) à la surface du cylindre. Le calcul est alors direct à partir des droites apparentes (voir le paragraphe 3.4.3) et une solution analytique existe pour retrouver les 6 degrés de liberté avec le point supplémentaire [Nag05].

Avec les marquages utilisés, on dispose, sauf en cas d'occultation, de trois à six points dont la position est connue. Cette redondance d'information peut être utilisée pour améliorer la précision de l'estimation de la pose. Pour notre application, nous ne chercherons pas à déterminer la pose de l'aiguille lorsqu'elle n'est pas tenue par le porte-aiguille. Nous avons considéré la situation où l'aiguille est tenue dans le porte-aiguille, et nous nous sommes intéressés à l'estimation de la pose du porte-aiguille seul et à la reconstruction des paramètres de saisie, c'est-à-dire la pose relative du porte-aiguille et de l'aiguille.

Le problème de l'estimation de la pose d'un objet à partir d'une image peut être vu comme la minimisation de l'erreur entre les informations de l'image (l'observation) et la projection

dans l'image des objets à la pose estimée, que nous appellerons erreur de reprojection. Ce problème de minimisation peut alors être traité en utilisant des méthodes non-linéaires classiques et des algorithmes numériques itératifs, comme la méthode de Newton-Raphson, ou les algorithmes habituellement utilisés pour les problèmes de moindres carrés non-linéaires comme ceux de Newton-Gauss ou de Levenberg-Marquardt. Cependant, ces algorithmes ont souvent des minima locaux qui empêchent la convergence vers les minima globaux. Eric Marchand et François Chaumette [MC02a] ont proposé d'utiliser les asservissements visuels pour résoudre les problèmes d'estimation de pose en développant un processus de minimisation non-linéaire itératif appelé "asservissement visuel virtuel" (AVV ou - VVS pour *Virtual Visual Servoing*).

L'objectif d'un asservissement visuel virtuel est d'amener une caméra virtuelle, dont la position initiale est connue, à la même position que la caméra réelle en utilisant l'image fournie par la caméra réelle et le principe des asservissements visuels basés image (2-D), qui, en pratique, peut être mis en oeuvre sans estimer la pose de l'objet d'intérêt. Dans une première phase, l'algorithme est initialisé. La pose de la caméra réelle par rapport à l'objet est estimée grossièrement et sert de pose initiale pour la caméra virtuelle. La position désirée de la caméra virtuelle par rapport à l'objet est celle de la caméra réelle. Elle est donc inconnue (c'est ce que l'on cherche à estimer), mais, en revanche, l'image désirée est connue : c'est l'image réelle à partir de laquelle on veut estimer la pose de l'objet. La partie importante de ce processus se situe au niveau du calcul de la matrice d'interaction. Un aspect intéressant des asservissements visuels virtuels est de tirer parti de tous les travaux déjà réalisés pour les asservissements visuels réels. En effet, l'expression des matrices d'interaction a été calculée pour des primitives élémentaires (points, droites, cylindres, cercles, sphères) [Cha90]. En référence à cela, nous avons proposé une matrice globale J_{pa} prenant en compte conjointement les marqueurs et les droites apparentes extraits, qui exprime sous la forme suivante

$$\begin{pmatrix} \dot{\mathbf{i}}^+ \\ \dot{\mathbf{i}}^- \\ \dot{p}_1 \\ \vdots \\ \dot{p}_n \end{pmatrix} = \underbrace{\begin{pmatrix} L_{line}(\mathbf{l}^+) \\ L_{line}(\mathbf{l}^-) \\ L_{pt}(p_1) \\ \vdots \\ L_{pt}(p_n) \end{pmatrix}}_{J_{pa}} \underbrace{\begin{pmatrix} {}^cV_{c/i} \\ {}^c\omega_{c/i} \end{pmatrix}}_{\boldsymbol{\tau}}. \quad (3.41)$$

le lien entre les variations des coordonnées des marqueurs dans l'image, \dot{p}_i , et celles des droites apparentes $\dot{\mathbf{i}}^-$, $\dot{\mathbf{i}}^+$ d'une part, et le torseur cinématique $\boldsymbol{\tau}$ représentant le mouvement de l'instrument d'autre part. Un schéma de commande cinématique a donc été élaboré, très similaire à celui déjà présenté pour le centrage dans le paragraphe précédent. Des expériences ont été menées au laboratoire à l'aide d'une "training box" où les instruments sont contraints de passer par un point d'incision sans effort [FTE⁺96, Kru03]. Sur la figure 3.31, nous avons extraits trois images durant l'asservissement, vu comme un processus itératif de la pose. Comme la plupart des optimisations numériques, cette approche fournit des résultats moins sensibles aux bruits de mesure que les techniques analytiques.

Il est intéressant de remarquer que ce schéma de résolution aurait pu contenir aussi la matrice d'interaction du cercle supportant l'aiguille. Un bref contact de l'aiguille avec un organe ou tout autre objet se trouvant dans la cavité abdominale pourrait suffire à la ré-orienter, les mâchoires du porte-aiguille ne pouvant pas supporter un effort perpendiculaire au plan de l'aiguille. Il

pourrait donc être intéressant de considérer à l'avenir le couple {aiguille, porte-aiguille} comme un objet rigide articulé, des techniques ont été développées récemment en ce sens par ailleurs.

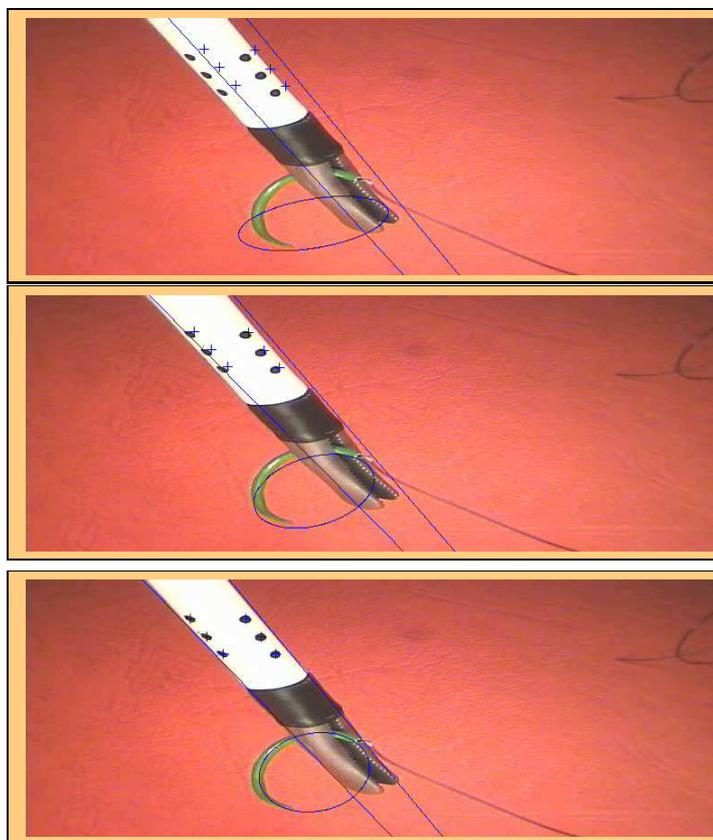


Fig. 3.31: L'estimation de la pose des instruments de chirurgie par la mise en oeuvre d'un asservissement visuel virtuel combinant différentes primitives géométriques (points et droites apparentes) modélisant un porte-aiguille. Les courbes bleues sont les projections de ces primitives pour la position initiale (virtuelle) de la caméra endoscopique (haut), pour une position intermédiaire (centre), et pour la position finale (bas).

Dans la thèse de Florent Nageotte [Nag05], la planification du chemin de l'aiguille lors de son passage à travers les tissus a été longuement étudiée et un algorithme a été développé, contraignant le choix du positionnement du trocart de l'instrument et permettant de fournir un ensemble de consignes successives à la phase de passage de l'aiguille à travers les tissus. Des outils d'assistance par réalité augmentée ont également été proposés.

Parmi les publications en rapport avec ce travail, je retiens ici des communications à des workshops internationaux tels que MRNV 2004 [NMD⁺04], WDV 2006 [DNM06b], les conférences IEEE ICRA 2005 [NZMD05], IEEE/RSJ IROS 2004 [DNM04] et 2006 [NZDM06], MICCAI 2006 [DNM06a] et la contribution à un ouvrage collectif [DNM07] venant de paraître (Dynamical Vision, 2007 chez Springer) qui est issu d'une sélection de communications présentées lors des deux premiers workshops (WDV'05 à ICCV et WDV'06 à ECCV) sur la vision dynamique [DNM07]. Deux articles pour des revues internationales sont en cours d'évaluation.

3.5.3 Guidage d'un laparoscope articulé

La dernière technique présentée ici dans le cadre de la laparoscopie assistée propose un guidage de l'endoscope embarqué sur le robot AESOP. L'idée qui a été développée l'an dernier consistait à réaliser un suivi d'une zone d'intérêt, préalablement sélectionnée par le chirurgien, zone pouvant correspondre aux organes et non plus aux instruments eux-mêmes. Les méthodes de suivi visuel présentées dans le chapitre II, en conjonction avec les asservissements visuels hybrides permettent de réaliser un tel suivi en prenant en compte les déformations d'un motif dans l'image et en offrant la possibilité d'estimer un mouvement 3-D dans l'espace, quand celles-ci émanent d'un déplacement d'un objet d'intérêt.

La méthode *ESM* de minimisation numérique, appliquée au suivi de zones d'organes supposées localement planes permet, par le biais d'un asservissement visuel, d'effectuer le guidage de l'endoscope [Ma198, MC00, MB04]. En appliquant ce schéma méthodologique à la laparoscopie robotisée, nous avons dû surmonter deux problèmes. Le premier concerne la détection des occultations (dus par exemple à la proximité d'autres organes ou à celle des instruments) et est un problème récurrent pour les approches de suivi par descripteurs globaux, du type *template-based matching*. Le second est inhérent à la chirurgie à invasion minimale : pour ce type de chirurgie, la mobilité est réduite à 4 degrés de liberté (voir la figure 3.4) et tous les degrés de liberté d'une cible rigide ne peuvent pas être considérés pour la commande, ce qui amène à élaborer une stratégie particulière.

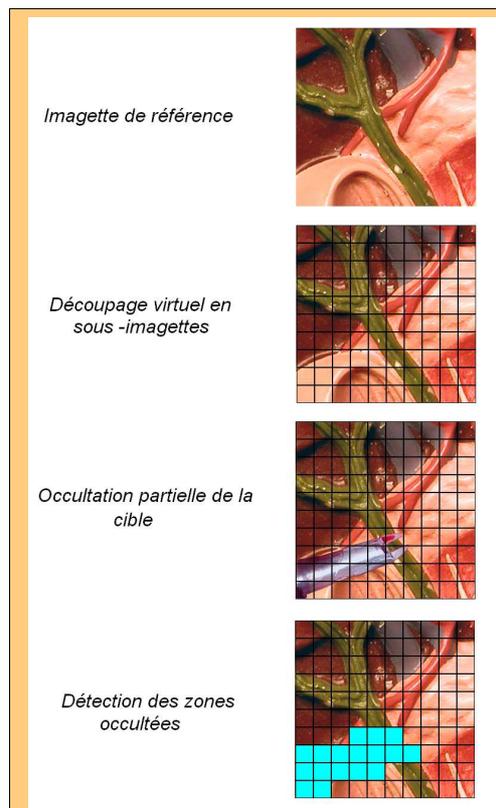


Fig. 3.32: Processus de découpage en sous-imagettes et détection des occultations.

Détection des occultations

Pour résoudre le problème lié aux occultations, nous avons échantillonné chacune des zones d'intérêt de l'image en sous-imagettes (voir figure 3.32). Une des difficultés relève alors de la forme non rectangulaire de la zone d'intérêt quand le mouvement 3-D n'est pas fronto-parallèle (effet perspectif). Pour cela l'estimation en ligne de l'homographie pour calculer la zone rectangulaire équivalente (virtuelle) par "warping" est nécessaire.

Chaque sous-imagette transformée de l'image courante est comparée à son correspondant dans l'image de référence afin de détecter une éventuelle occultation. Diverses méthodes de détection de changements basées sur la comparaison des niveaux de gris ont été testées, telles que :

- Comparaison des niveaux de gris moyens entre sous-imagettes,
- Comparaison des écarts-types,
- Méthode Geo-Pixel [JMN77],
- Comparaison d'histogrammes au moyen de la distance de Bhattacharyya [Sco04, LLLP05].

L'inconvénient de toutes ces méthodes qui ont été appliquées principalement sur des images à niveaux de gris réside dans la sensibilité des critères aux variations de luminosité. En effet, ces méthodes ne font guère de différences entre une réelle occultation et un simple changement de luminosité. Il est donc apparu essentiel d'introduire la couleur pour une détection plus fiable. L'idée retenue a été de représenter les couleurs dans l'espace HSV (teinte, pureté, luminosité - *Hue, Saturation, Value*, cf. paragraphe 3.3.1) et d'utiliser un des attributs de couleur pour détecter les occultations. L'attribut qui a été retenu est (une variante de) celui qui a été décrit au paragraphe 3.3.1, à savoir

$$\mathcal{H} = S'_{HSV} \cdot \cos(H) . \quad (3.42)$$

Les histogrammes de \mathcal{H} pour chaque sous-imagette de l'imagette de référence et son correspondant dans l'imagette actuelle sont calculés et comparés en considérant une décomposition en plusieurs classes. La détermination du nombre de classes d'un histogramme est généralement délicate et il n'existe pas de règle absolue. Un trop faible nombre de classes fait perdre de l'information et aboutit à gommer les différences pouvant exister entre deux sous-imagettes. En revanche, un trop grand nombre de classes aboutit à des graphiques incohérents où certaines classes deviennent vides ou presque car la taille de la sous-imagette est finie. Les meilleurs résultats sont obtenus en utilisant la règle de Sturges qui propose comme limite maximale du nombre de classes la formulation suivante

$$M = \left\lceil 1 + \frac{10}{3} \log(W \times L) \right\rceil \quad (3.43)$$

où $\lceil \cdot \rceil$ représente la partie entière, et $(W \times L)$ le nombre de pixels de la sous-imagette. Ainsi, pour une sous-imagette de taille 10×10 , le nombre de classe est typiquement égal à 7.

Les deux distributions discrètes non-paramétriques associées respectivement à la sous-imagette a_i dans l'image de référence et à la sous-imagette a_j dans l'image actuelle sont comparées à travers un critère de similarité basé sur le coefficient de Bhattacharyya, B_{ij} [DSG90]. Comme relevé par Dorin Comaniciu dans [CRM03], ce coefficient est une variante de la distance de

Bhattacharyya qui n'est pas une métrique *stricto sensu* dans sa forme initiale. Pour la sous-image a_i par exemple, la densité discrète est définie par

$$\hat{\mathbf{a}}_i = \{\hat{a}_i^{(l)}\}_{l=1,\dots,m} \quad \text{avec} \quad \sum_{l=1}^m \hat{a}_i^{(l)} = 1 \quad (3.44)$$

où $\hat{a}_i^{(l)}$ correspond à la $l^{\text{ème}}$ classe de l'histogramme. La distance statistique entre les deux distributions a_i et a_j est alors simplement définie par

$$d_{ij} = \sqrt{1 - B_{ij}} \quad (3.45)$$

avec

$$B_{ij} = \sum_{l=1}^m \sqrt{\hat{a}_i^{(l)} \hat{a}_j^{(l)}}. \quad (3.46)$$

Minimiser la distance (3.45) est équivalent à maximiser le coefficient de Bhattacharyya (3.46). La sous-image courante de la région d'intérêt est considérée comme présentant une occultation si d excède une valeur de seuil, τ_b , pré-définie (typiquement 50 %).

En fonction des capacités disponibles de calculs, la comparaison peut s'effectuer une seule fois à chaque acquisition d'image ou plusieurs fois, par exemple à chaque itération de l'algorithme *ESM*.

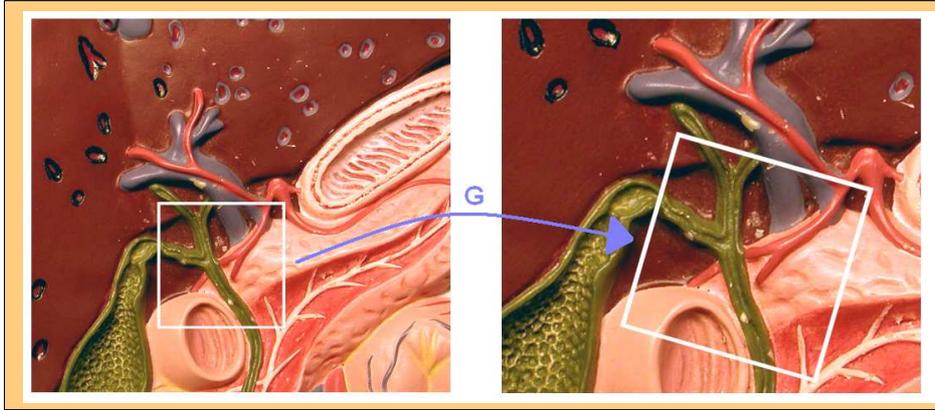


Fig. 3.33: Transformation géométrique entre l'image de référence et l'image actuelle. Quand la différence entre ces deux images provient d'un déplacement 3-D d'une scène plane, une transformation homographique G relie les deux plans images.

Une évaluation quantitative est nécessaire pour attester des performances de notre méthode. A cette fin, nous avons défini une erreur, $e_{occl}(t)$, basée sur la norme de Frobenius, afin de comparer avec une cible fixe, les composantes de l'homographie actuelle (entre la région d'intérêt actuelle et la région d'intérêt correspondante dans l'image de référence) en présence d'occultations, quand celles-ci sont détectées, et quand elles ne le sont pas.

Les résultats obtenus concernant la prise en compte des occultations montrent que l'utilisation de la couleur dans le processus de détection permet de réduire de plus de 60% l'erreur d'estimation de l'homographie (figure 3.35) par rapport à la méthode originelle, tout en conservant

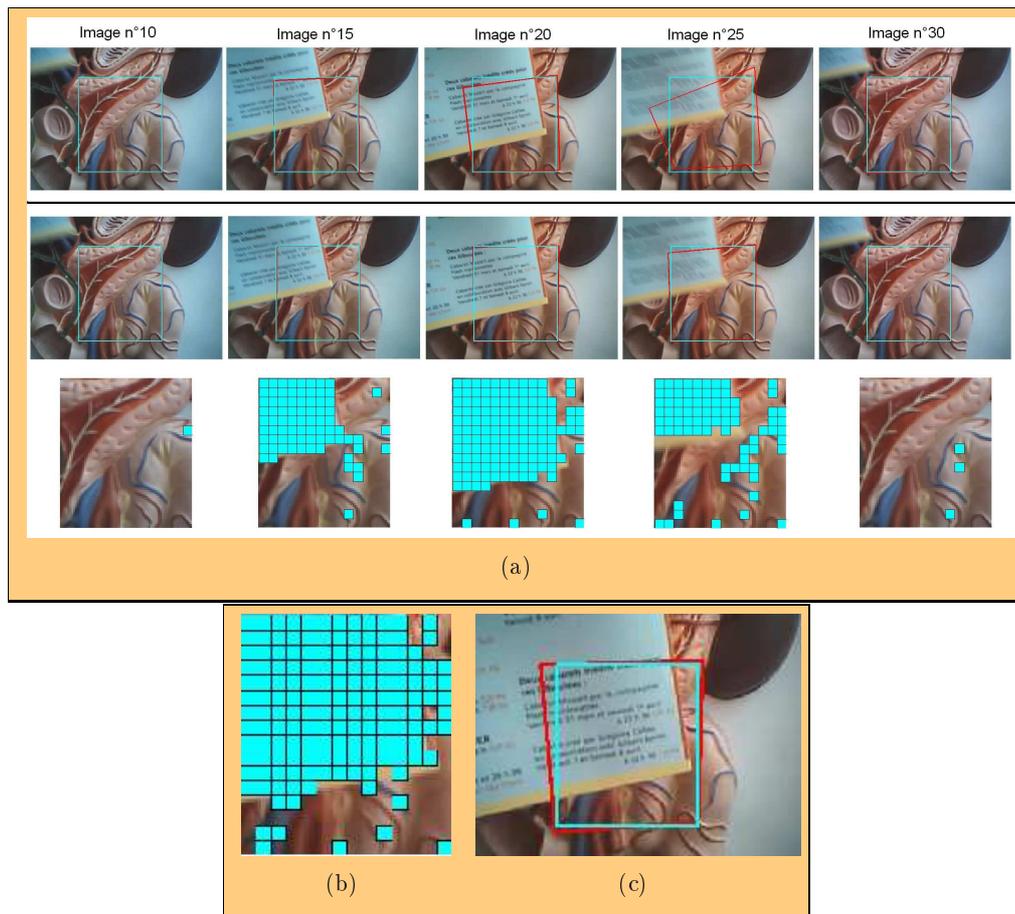


Fig. 3.34: (a) Comparaison du suivi (ESM) sans détection des occultations (en haut) et avec détection des occultations (au centre). En bas, sont reportées (en bleu) les sous-imagettes correspondantes où la distance de Bhattacharyya était supérieure à 50 %. (b-c) En particulier, pour l'image 23 de la séquence, 155 des 196 zones ont été rejetées. On peut observer (qualitativement) la précision de l'homographie estimée en comparant les différences de formes des quadrilatères délimitant l'imagette transformée (le rectangle de référence est en bleu, le quadrilatère rouge illustre l'estimation de l'homographie).

la robustesse de l'algorithme vis-à-vis des variations de luminosité [BDZM07]. On constate également qu'un petit nombre de sous-imagettes bruitées mais ne présentant pas d'occultations sont néanmoins rejetées. Cela n'est pas gênant pour le calcul de l'homographie, qui ne tient compte que des pixels des sous-imagettes retenues et s'avère être précis même avec une faible fraction de la zone d'intérêt (voir figure 3.34,b-c).

La méthode proposée, en l'état actuel, est limitée à des temps de calculs encore trop élevés (> 150 ms) pour prétendre être incorporée dans un processus de guidage robotique. Néanmoins, des efforts se portent en ce moment sur une meilleure implémentation des algorithmes proposés ici, ceux-ci ne présentant pas (dans leur conception) de fortes variations de temps de calculs en fonction de la complexité de la scène.

Ce travail a fait l'objet d'un article qui a été présenté lors de la conférence IEEE ICRA 2007.

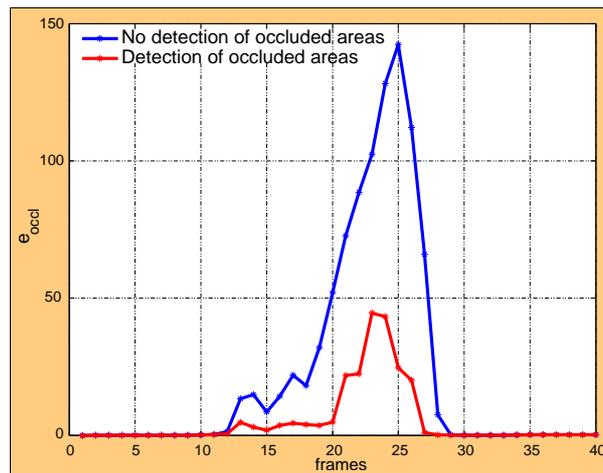


Fig. 3.35: L'erreur algébrique e_{occl} durant la même séquence vidéo que celle décrite sur la figure 3.34 avec (courbe rouge) et sans (courbe bleue) la détection des sous-images occultées.

Asservissement visuel de l'endoscope en chirurgie laparoscopique

Comme toutes les procédures percutanées, la chirurgie laparoscopique induit des mouvements des instruments chirurgicaux qui sont contraints de passer à travers un point d'incision sans qu'il y ait des efforts le long de la paroi abdominale, réduisant la mobilité à 4 degrés de liberté seulement¹. Au même titre que les autres instruments, le laparoscope (E) est contraint de passer par ce point d'incision, noté O_A sur la figure 3.36, et supposé immobile, qu'il soit embarqué sur un robot ou tenu par un assistant.

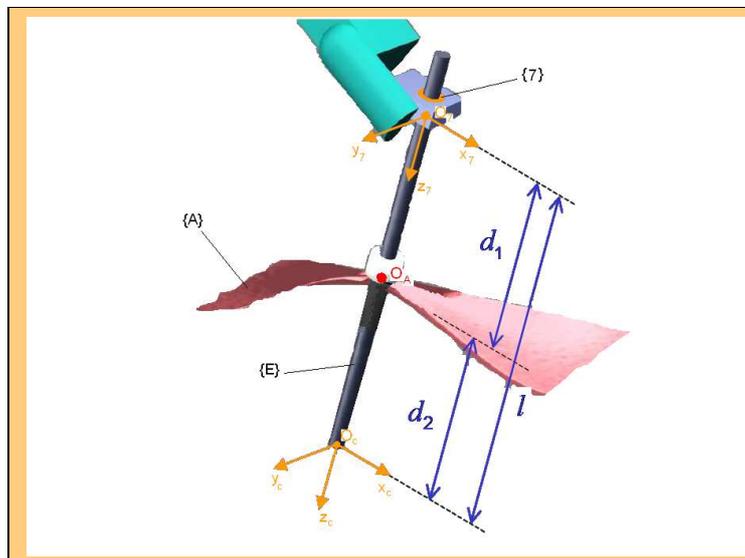


Fig. 3.36: Modélisation géométrique et définition des notations et des repères pour une configuration eye-in-hand en laparoscopie robotisée.

¹ Certains instruments de laparoscopie possèdent des articulations à l'une des extrémités pour augmenter la dextérité (*microwrist*)

Pour commander un robot chirurgical par un retour visuel de ce type, une fonction de tâches $\mathbf{e} = (\mathbf{e}_\nu, \mathbf{e}_\omega) \in \mathbb{R}^6$ a été définie, de manière identique à celle proposée par Selim Benhimane [BM06]. A l'aide des images du laparoscope corrigées des distorsions radiales, la fonction de tâches peut s'exprimer de la façon suivante

$$\mathbf{e}_\nu = (\mathbf{H} - \mathbf{I}_{3 \times 3}) \mathbf{m}^* \quad \text{et} \quad [\mathbf{e}_\omega]_\times = \mathbf{H} - \mathbf{H}^\top \quad (3.47)$$

où \mathbf{m}^* et $\mathbf{H} \mathbf{m}^*$ sont des points géométriques en correspondance respectivement dans la région d'intérêt de l'image de référence et dans celle de l'image actuelle ($[\mathbf{v}]_\times$ est la matrice anti-symétrique du vecteur \mathbf{v}). Avec un dispositif de vision complètement calibré, de matrice de paramètres intrinsèques \mathbf{K} , la matrice \mathbf{H} est calculée à travers l'algorithme de suivi ESM qui est alors reliée à l'homographie de plans $\mathbf{G} = \mathbf{K} \mathbf{H} \mathbf{K}^{-1}$. Le découplage de la commande est possible grâce au découplage de l'erreur de position \mathbf{e}_ν avec l'erreur d'orientation \mathbf{e}_ω et aboutit à un comportement indépendant et de profils en décroissance exponentielle de la loi de commande du torseur cinématique du mouvement du laparoscope qui s'exprime par

$$(\boldsymbol{\tau}_c)_{/O_C}^{\text{esm}} = \begin{bmatrix} \boldsymbol{\nu} \\ \boldsymbol{\omega} \end{bmatrix} = - \begin{bmatrix} \lambda_\nu \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \lambda_\omega \mathbf{I}_{3 \times 3} \end{bmatrix} \begin{bmatrix} \mathbf{e}_\nu \\ \mathbf{e}_\omega \end{bmatrix}, \quad (3.48)$$

où λ_ν et λ_ω sont des nombres réels positifs. Ezio Malis et Selim Benhimane ont fourni la preuve de la stabilité de cette loi dans [BM06].

Pour assurer qu'aucun mouvement de translation n'est généré le long de la paroi abdominale (tangent à la surface, au point d'incision), le torseur cinématique, exprimé en ce point, doit être limité aux 4 degrés de liberté suivants

$$(\boldsymbol{\tau}_c)_{/O_A}^{4 \text{ dof}} = [0 \quad 0 \quad V_z \quad \Omega_x \quad \Omega_y \quad \Omega_z]^\top. \quad (3.49)$$

Ceci signifie que seuls les mouvements d'orientation et d'insertion le long de l'axe de l'incision sont autorisés. L'expression du torseur donnée par l'équation (3.48) au point d'incision O_A est (voir les notations sur la figure 3.36)

$$(\boldsymbol{\tau}_c)_{/O_A}^{\text{esm}} = [v_x - d_2 \omega_x \quad v_y + d_2 \omega_x \quad v_z \quad \omega_x \quad \omega_y \quad \omega_z]^\top \quad (3.50)$$

où d_2 représente la profondeur de pénétration. Par identification des deux équations (3.49) et (3.50), il vient

$$\begin{cases} v_x - d_2 \omega_x = 0 & \omega_x \equiv \Omega_x \\ v_y + d_2 \omega_x = 0 & \omega_y \equiv \Omega_y \\ v_z \equiv V_z & \omega_z \equiv \Omega_z \end{cases} \quad (3.51)$$

Alors que l'obtention des expressions de la vitesse de rotation propre Ω_z et de la vitesse de pénétration V_z sont immédiates, il n'y a par contre aucune garantie que les deux équations $v_x - d_2 \omega_x = 0$ et $v_y + d_2 \omega_x = 0$ soient vérifiées. En effet, l'objectif de commander les derniers paramètres Ω_x and Ω_y afin d'éviter tout mouvement de translation perpendiculaire à l'axe d'incision et de simultanément suivre la région d'intérêt en compensant les déformations du motif est impossible à atteindre dans le cas général puisque la région d'intérêt peut se déplacer sans contrainte de mouvement. Seul, un compromis, inclus dans la stratégie de commande, peut être envisagé ici. La solution que nous proposons tient compte de considérations pratiques.

Etant donné les forts grossissements mis en jeu et l'effet de levier important avec un laparoscope, les mouvements de translation induisent des variations dans l'image plus importantes que les mouvements de rotation (ces derniers provoquent en fait des distorsions perspectives). Ceci signifie que le suivi visuel du centre de la région d'intérêt est à considérer prioritairement pour conserver celle-ci (ou une partie de celle-ci) dans le champ de vision. Par conséquent, la solution retenue est d'exprimer les composantes Ω_x et Ω_y du torseur cinématique comme des fonctions de v_x et v_y , c'est-à-dire

$$\begin{bmatrix} V_z \\ \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & k_d & 0 & 0 & 0 \\ 0 & -\frac{k_p}{d_2} & 0 & 0 & 0 & 0 \\ \frac{k_t}{d_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_s \end{bmatrix} (\tau_c)_{/O_C}^{\text{esm}} \quad (3.52)$$

Finalement, en exprimant le torseur cinématique à l'origine du repère de l'organe terminal du robot, O_7 , on a :

$$(\tau_c)_{/O_7}^{4 \text{ dof}} = \begin{bmatrix} \frac{-k_t}{l-d_1} d_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{k_p}{l-d_1} d_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_d & 0 & 0 & 0 \\ 0 & \frac{-k_p}{l-d_1} & 0 & 0 & 0 & 0 \\ \frac{k_t}{l-d_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_s \end{bmatrix} (\tau_c)_{/O_C}^{\text{esm}} \quad (3.53)$$

où $l = d_1 + d_2$ est une quantité constante (dans les travaux de thèse d'Alexandre Krupa [KGD⁺03], nous avons proposé deux méthodes pour l'estimation de d_1).

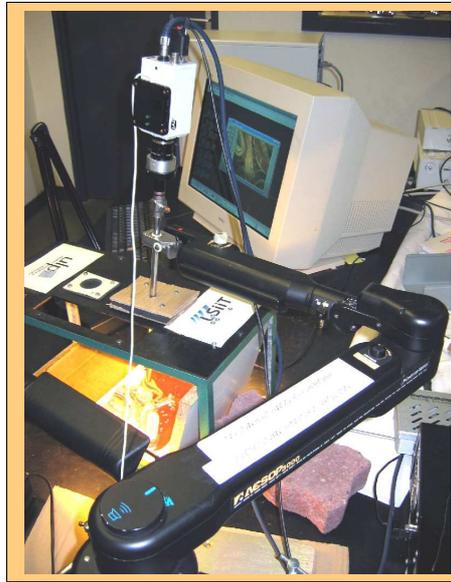


Fig. 3.37: Le robot porte-endoscope AESOP et la "training box" utilisés durant les expérimentations du suivi et de l'asservissement sans modèle au laboratoire, en incorporant les contraintes du mouvement dues au passage par le trocart.

Afin de valider cette stratégie de commande avec compromis, des expériences ont été conduites au laboratoire avec le robot de chirurgie AESOP conçu à l'origine aussi bien pour télé-manipuler les instruments de chirurgie que comme porte-endoscope (figure 3.37). Le système est composé également d'une caméra CCD couleur connectée au laparoscope, d'une "training box" avec deux points d'insertion et d'une maquette en céramique d'organes internes à l'abdomen. La région d'intérêt désirée (rectangle vert sur la figure 3.38) est sélectionnée au moyen d'une interface graphique conçue au laboratoire. Le suivi ESM sans asservissement est alors activé, puis la maquette est déplacée à la main. Les articulations du robot AESOP sont alors déplacées vers une autre configuration (quadrilatère rouge) en modifiant sur l'interface graphique les positions des 4 axes actifs par l'intermédiaire des barres de défilement que l'on peut voir sur les deux figures 3.38-a,b. Finalement, l'asservissement visuel est activé (clic de souris sur le bouton jaune) puis exécuté (figure 3.38-b).

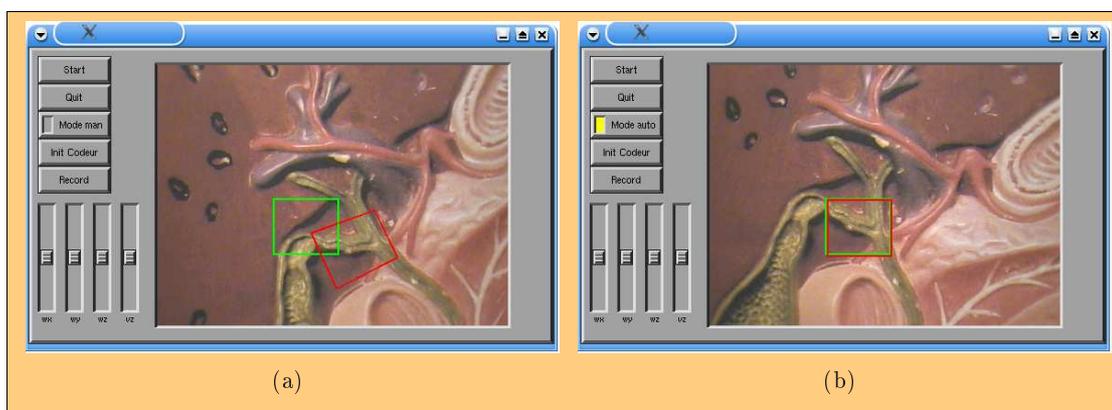


Fig. 3.38: (a) La région d'intérêt est sélectionnée (rectangle vert), puis le suivi visuel ESM est activé (sans asservissement) et la maquette et le robot sont déplacés vers une autre configuration (quadrilatère rouge). (b) L'asservissement visuel est activé, et le quadrilatère rouge converge (complètement pour son déplacement, la forme dans une moindre mesure) vers le carré vert.

Les réponses temporelles des vitesses de rotation ($\Omega_x, \Omega_y, \Omega_z$) et d'insertion V_z durant la phase de l'asservissement sont reportées sur les figures 3.39-a et 3.39-b. Les valeurs des gains choisies pour la correction proportionnelle des erreurs sont $(k_p, k_t, k_s, k_d) = (0.3, 0.3, 0.3, 20)$ (les indices p, t, s , et d correspondent aux termes anglais *span, tilt, spin* et *descent*). L'évolution des coordonnées du centre de la région d'intérêt, $(\mathbf{m}^* = (u^*, v^*))$ et les trajectoires dans le plan image durant l'asservissement sont reportées sur les figures 3.39-c et 3.39-d respectivement.

Malgré le bruit significatif constaté sur les mesures, la situation du robot converge vers la situation désirée en approximativement 16 secondes (400 images) avec les valeurs des gains indiquées précédemment. Cela est relativement lent, mais il semble difficile d'améliorer le comportement dynamique du système global sans en modifier le schéma de commande, car un phénomène de cycle limite est observé au voisinage de la position désirée (voir la figure 3.39-d). Des non-linéarités telles que des frictions ou des zones mortes sur les axes du robot ont tendance à provoquer des petits déplacements rapides qui modifient significativement certaines composantes de l'homographie. Deux autres facteurs limitant les performances sont, d'une part, le temps de communication important entre l'ordinateur effectuant les traitements

et le contrôleur de robot et, d'autre part, ce contrôleur qui est aussi très lent à élaborer la commande électrique sur les axes du robot AESOP : le délai total a été estimé à 160 ms [Kru03], c'est-à-dire 4 périodes d'échantillonnage de la cadence vidéo PAL.

De plus, les expériences présentées pour ce guidage ont été accomplies en considérant une valeur constante de la quantité d_1 (distance entre l'organe terminal et le point d'insertion) alors qu'il s'agit d'un paramètre variable.

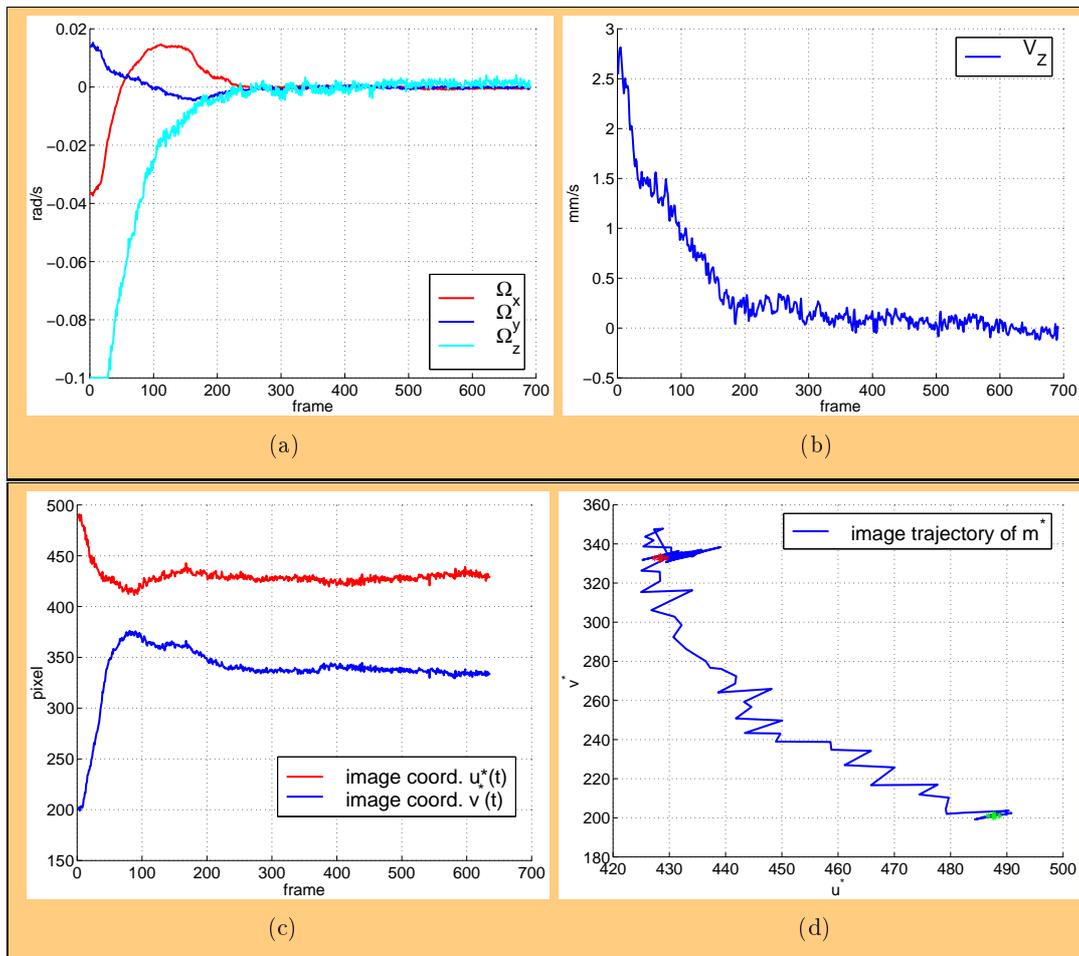


Fig. 3.39: Comportement dynamique de l'asservissement visuel sans modèle avec contraintes du mouvement (4 degrés de liberté) pour l'assistance en laparoscopie. (a-b) Réponses temporelles des grandeurs commandées. (c) Réponses temporelles (des coordonnées) du centre de la région d'intérêt $\mathbf{m}^ = (u^*(t), v^*(t))$. (d) Trajectoires dans l'image de \mathbf{m}^* .*

3.6 RECALAGE STÉRÉOTAXIQUE EN RADIOLOGIE INTERVENTIONNELLE

Nous présentons ici une nouvelle méthode de recalage par descripteurs locaux destinée à localiser un objet radio-opaque par rapport à un capteur visuel particulier : un scanner tomodensitométrique (TDM) à rayons X [GD94, Hsi03]. Ce problème de recalage 2-D/3-D est essentiel pour évaluer le positionnement et le guidage automatique d'un instrument de chirurgie [SWA⁺98], tel qu'une aiguille de radio-fréquence utilisée par les techniques peu invasives en radiologie interventionnelle.

3.6.1 Guidage par imagerie TDM

Les progrès réalisés sur les systèmes d'imagerie et l'apparition des systèmes de GMCAO ont participé au développement de la chirurgie dite à invasion minimale pour laquelle au lieu de réaliser des opérations lourdes pour le patient, le chirurgien effectue des incisions de taille réduite. Pour visualiser le siège de l'intervention, il utilise les informations fournies par divers capteurs et en particulier les images médicales. Ainsi une stratégie chirurgicale peut être définie seulement sur des images pré-opératoires ou bien ces images peuvent être utilisées de concert avec des images per-opératoires, qui souvent sont issues d'une autre modalité. Il y a généralement lieu de recalibrer les unes avec les autres (recalage multi-modales) ou de recalibrer un modèle issu des images per-opératoires avec le patient. Les images per-opératoires peuvent également être utilisées en conjonction avec d'autres capteurs (localiseur optique, guide magnétique, électrocardiogramme,...).

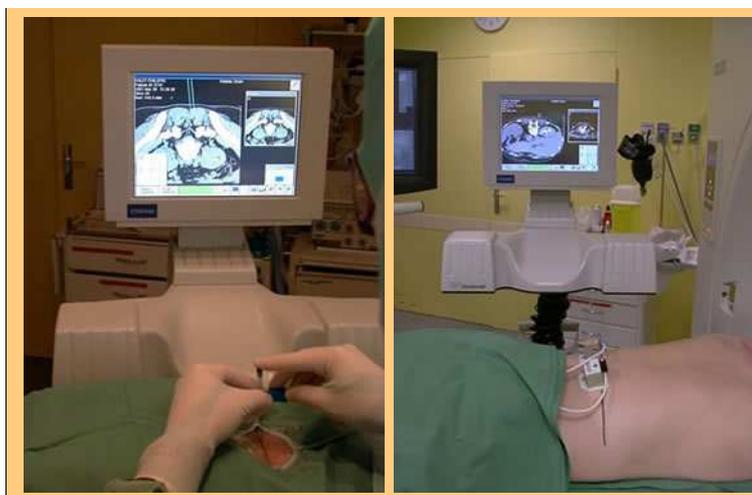


Fig. 3.40: . Visualisation par superposition sur des images pré-opératoires de la position en temps-réel d'un système de pointage électromagnétique (CT-Guide).

Dans la plupart des cas, le retour visuel est employé pour suivre manuellement la manipulation en appréciant l'évolution d'une insertion percutanée (figure 3.40). Certains prototypes de maintien et de guidage de l'instrument nécessitant un retour visuel ont été conçus dans le passé pour être employés avec l'imagerie échographique [ASZ⁺02, HDH⁺02, VMO05] et l'imagerie à résonance magnétique (IRM) [MKM⁺95, TAVA⁺04].

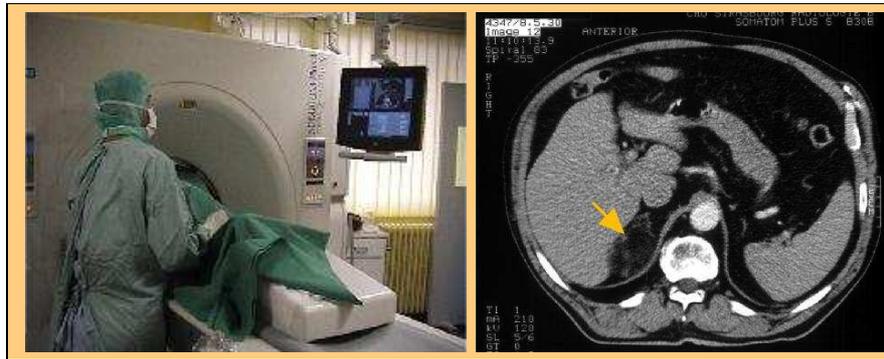


Fig. 3.41: (a) Guidage de l'insertion par suivi visuel manuel à l'aide du mode fluoroscopique des récents scanners est encore peu répandu. (b) image tomодensitométrie d'une tumeur.

La tomодensitométrie apporte des avantages non négligeables par rapport à ces techniques d'imagerie, même si son fonctionnement est à temps différé. Le scanner à rayons X offre une très bonne différenciation des tissus organiques, encore inférieure à l'IRM, tout en autorisant la visualisation de tous les matériaux possibles (voir la figure 3.41(b)). La TDM n'a pas de zones aveugles comme c'est le cas pour l'échographie en présence des parois osseuses et elle peut fournir de surcroît des images en quasi temps-réel pour les systèmes les plus récents (16 images par seconde en mode fluoroscopie à faible puissance). Ce type d'imagerie est également reconnu pour la quasi absence de distorsions dans l'image, au contraire de l'IRM. Elle reste la modalité d'images la plus courante pour la radiologie interventionnelle en raison d'un coût plus faible que l'IRM. Ces informations visuelles peuvent alors être exploitées par les robots pour assister ou guider le chirurgien grâce à des systèmes de guidage passifs, semi-actifs, actifs, synergiques ou télé-opérés [CBBB95, TCB⁺05].

Ashin Gangi, radiologue expert à l'Hôpital Civil de Strasbourg, nous ayant exposé les problèmes d'irradiations subis par le corps médical lors de gestes répétés sous imagerie à rayons X, nous sommes orientés vers une solution d'assistance qui concerne aussi bien le maintien de l'instrument que son positionnement (projet IRASIS : Insertion Robotisée d'Aiguille Sous Imageur Scanner) dédiée à la tomодensitométrie interventionnelle. En collaboration avec lui et les membres de son service, nous avons travaillé sur une assistance au guidage. Plus précisément, lorsque le praticien souhaite choisir un point d'insertion sur le corps du patient pour accéder à une zone interne qu'il observe en effectuant (ponctuellement) une coupe scanner, il serait souhaitable de lui fournir une évaluation de la position et de l'orientation de l'instrument par rapport au scanner. A cet effet, il existe de nombreux systèmes en développement ou déjà commercialisés. Ceux-ci sont répertoriés dans le mémoire de thèse de Benjamin Maurin [Mau05]. On peut citer ici à titre d'exemple le système RCM PAKY ou AcuBot de l'université John Hopkins, le système Neuromate de TIMC à Grenoble, Minerva de l'EPFL, PathFinder de Armstrong Healthcare, entre autres.

Un facteur clé dans le guidage intraopératif par vision d'un robot médical est de parvenir à localiser précisément un instrument de chirurgie dans l'espace opératoire [SAJ93]. Dans le cas présent, le recalage d'une aiguille de radio-fréquence doit être envisagé par rapport au scanner. Ceci est habituellement réalisé à l'aide d'un ensemble de marqueurs artificiels externes ou ana-

tomiques [RRAC05] qui peuvent être facilement détectés. Des motifs de forme simple peuvent être fixés sur le patient [BRO80], implantés dans le corps [JFW⁺97] ou rigidement liés à une plateforme porte-aiguille. Le scanner à rayons X délivre des images correspondant à des coupes (plan de coupes) effectuées sur un patient, le recalage est résolu par une technique appelée stéréotaxie basée sur l'utilisation d'un ensemble de marqueurs tridimensionnels radio-opaques, appelés marqueurs stéréotaxiques.

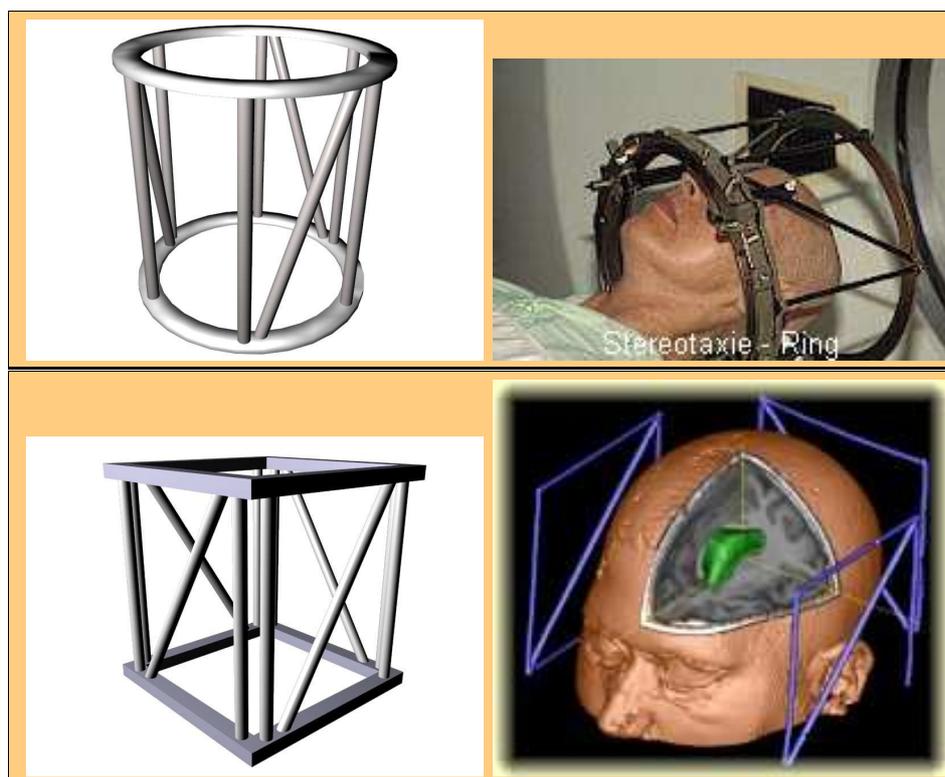


Fig. 3.42: Différents types de marqueurs (et "casques") stéréotaxiques.

3.6.2 La stéréotaxie

La stéréotaxie est généralement définie comme une méthode qui permet de repérer physiquement les structures internes d'un patient dans les trois plans de l'espace, en utilisant des informations visuelles fournies par un marqueur tridimensionnel. Historiquement, c'est en neurochirurgie que fut utilisée une armature (cadre ou casque stéréotaxiques) vissée sur le crâne d'un patient et embarquant des marqueurs rigides absorbant les rayons X (radio-opaques) tels que ceux illustrés sur la figure 3.42 ([BRO80], [WMFM96], [LJ94]). Un point de la tête du patient peut ainsi être facilement localisé, point qui peut désigner une tumeur à atteindre afin d'effectuer une biopsie ou une résection. C'est surtout dans ce domaine de la chirurgie que l'on trouve la source bibliographique la plus abondante sur la stéréotaxie. Dès lors, le nom a gardé une connotation de neuro-navigation. Aujourd'hui, le terme général de stéréotaxie concerne également les techniques de guidage par reconstruction de pose entre un repère d'une image

médicale et un repère attaché au patient [GCK⁺03]. Plus tard, pour les opérations robotisées, les marqueurs furent placés le plus proche possible de l'organe terminal ou de l'instrument. Néanmoins cette proximité est limitée car la présence des marqueurs ne doit pas aboutir à une réduction significative de l'espace de travail du robot ou gêner le praticien. Le choix de la taille, du diamètre, et de la configuration des marqueurs a un impact important sur l'espace disponible et sur la précision de la localisation.

Un grand nombre de systèmes stéréotaxiques existent, tous employant ce même principe pour localiser une cible. Nous pouvons citer comme exemple le système de repérage de Brown-Roberts-Wells (BRW) [BRO80] (figure 3.42). Bien que le modèle de projection soit ici différent du modèle sténopé utilisé pour une caméra classique, le recalage stéréotaxique peut être abordé en s'inspirant des méthodes développées dans le cas de la projection perspective. C'est du moins ce que nous avons souhaité faire en proposant une nouvelle formulation du recalage stéréotaxique, très différente de ce qui avait été fait jusqu'alors. Il est désormais possible d'analyser des configurations singulières et d'obtenir un recalage robuste, automatique et rapide, exploitable par un robot médical d'assistance à l'insertion d'aiguille en radiologie interventionnelle (CT-Bot).

Etat de l'art

Kwoh *et al.* [KHJH88], ont réalisé dès 1988 une biopsie pour une tumeur au cerveau à l'aide d'un robot industriel (PUMA 800). Le recalage entre l'aiguille attachée à l'organe terminal du robot et le patient est réalisé en deux étapes et en utilisant un ensemble de marqueurs stéréotaxiques formant 3 "N" et fixé sur une base circulaire, elle-même fixée sur la table d'opération. Le patient est alors avancé et sa tête est fixée à la base. Des images sont acquises permettant de repérer la tumeur (réduite à un point géométrique) dans le repère stéréotaxique. Le patient est alors enlevé puis l'aiguille est orientée et alignée sur ce point cible. Ce procédé d'alignement est une procédure longue, qui nécessite plusieurs coupes scanner, mais qui permet de référencer les coordonnées de la cible dans le repère de base du robot. Ce dispositif repris par Masamune *et al.* [MFP⁺01] puis par Gabor Fichtinger [FDP⁺02] a permis d'effectuer des biopsies, des drainages d'abcès et des ablations.

Lemieux and Jagoe [LJ94] ont étudié la propagation de l'erreur de détection des projections des marqueurs sur la localisation du cadre stéréotaxique. A notre connaissance, ils furent les premiers à établir un lien entre la configuration relative des marqueurs et les instabilités numériques observées lors de la détermination du recalage.

Susil *et al.* [SAT99] ont proposé une technique analytique qui détermine la position et l'orientation de marqueurs rectilignes, à partir d'une seule coupe. Ainsi, à l'aide d'un ensemble de 9 barres (3 "N" disposés sur 3 plans orthogonaux), Susil a obtenu un recalage précis (< 1.45 mm) par la réalisation d'un algorithme dédié à cette configuration spécifique des marqueurs. La méthodologie proposée ne permet pas de distinguer le modèle des marqueurs de celui de l'imageur car l'ensemble est imbriqué dans l'estimation du recalage. Il n'est donc pas possible d'intégrer un second ensemble de marqueurs, pour un arrangement différent des barres sans modifier l'algorithme de recalage proposé. D'autre part, il n'est pas non plus possible d'exhiber les configurations singulières.

Plus récemment Dai *et al.* ont proposé une méthode de recalage stéréotaxique également dédiée à un arrangement spécial de 12 marqueurs disposés en "V". Ce nombre très élevé de

marqueurs ne favorise pas la réduction de l'encombrement.

Enfin, Lee *et al.* [LFC02] ont proposé plusieurs méthodes numériques (descente de gradient) pour effectuer un recalage rigide à partir d'une seule coupe à l'aide de 9 marqueurs rectilignes. Là encore, il n'y a pas de distinction entre la modélisation géométrique du motif et celui du scanner. De plus il n'est pas possible de détecter *a priori* les configurations particulières du motif qui pourraient amener à des singularités du jacobien calculé à chaque étape de cette solution itérative.

Pour l'ensemble de ces travaux, il n'est pas proposé de recalage automatique et les correspondances entre les droites des marqueurs et leurs images sont établies par un appariement manuel, un procédé long et fastidieux. La méthode que nous proposons contribue à résoudre ces problèmes. En distinguant dans la formulation du recalage stéréotaxique la modélisation de l'imageur de la modélisation des marqueurs, nous avons développé une méthode permettant de prendre en compte tout arrangement de marqueurs rectilignes, d'étalonner le scanner, d'exhiber les configurations problématiques, les méthodes existantes établissaient le recalage pour des configurations spécifiques. De plus, nous fournissons un critère simple permettant d'évaluer s'il y a une réelle intersection entre le plan image du scanner et un marqueur. Ce dernier point permet d'envisager des simulations réalistes. Finalement, le recalage automatique de n'importe quel ensemble de motifs rectilignes est possible avec un nombre moins élevé de marqueurs (4). L'automatisation du recalage basée sur une estimation robuste (RANSAC) permet d'incorporer des contraintes géométriques supplémentaires ce qui conduit à rejeter rapidement de faux appariements [MDMG03, Mau05].

3.6.3 Méthodologie proposée pour le recalage stéréotaxique

Cette partie décrit brièvement les relations géométriques qui relient des droites d'un marqueur objet à leurs projetés dans l'image scanner, ainsi que la démarche d'une solution linéaire pour l'estimation de pose. Le lecteur intéressé par les détails et l'ensemble des résultats pourra consulter [MDMG03, Mau05, DMBM06].

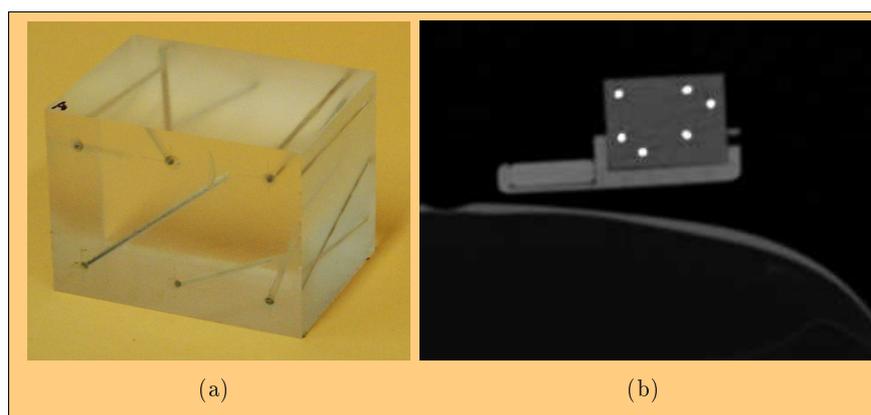


Fig. 3.43: (a) Un cube en plexiglass renferme, tel un objet rigide, un ensemble de 6 barres rectilignes métalliques utilisées comme marqueurs stéréotaxiques durant les expériences. (b) Une image scanner (amplifiée) quand le cube stéréotaxique est placé dans le "tunnel" du scanner, et positionné sur la base de la plate-forme robotique CT-Bot.

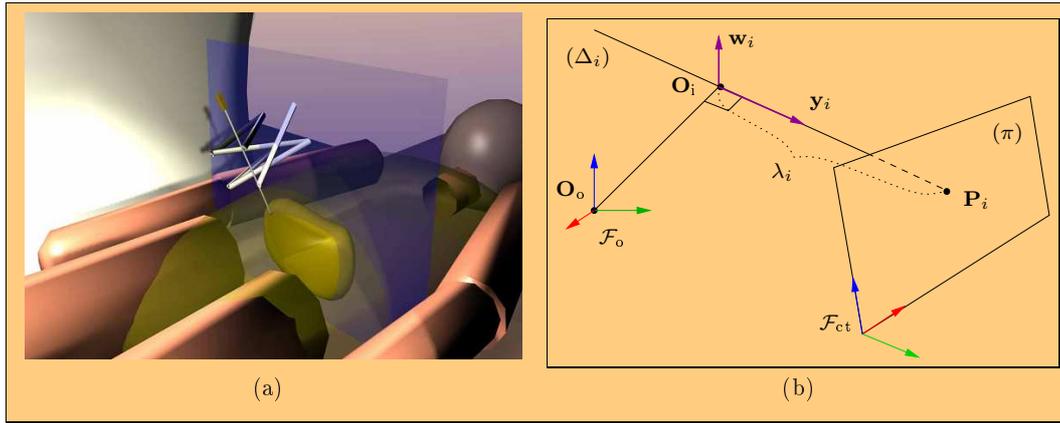


Fig. 3.44: (a) Recalage stéréotaxique avec marqueurs externes. Les intersections des marqueurs rectilignes avec le plan de coupe du scanner (bleu foncé) fournissent des points image qui sont employés pour retrouver la position et l'orientation des marqueurs par rapport au scanner à rayons X. (b) Une droite Δ_i de l'espace coupe le plan de coupe (π) au point P_i . Les vecteurs $(\mathbf{y}_i, \mathbf{w}_i)$ constituent la représentation plückerienne de la droite.

Modélisation géométrique

Généralement, les intersections entre les droites du marqueur et le plan de coupe doivent fournir autant de points image qu'il y a de droites (voir figure 3.43). Pour formuler ce problème géométrique, nous avons proposé un modèle affine composé de facteurs d'échelle, d'une transformation euclidienne et d'une projection orthogonale. Ce modèle permet de s'affranchir d'une analyse mathématique de la tomographie et nous permet de considérer cet imageur comme une boîte noire délivrant une image de l'objet radio-opaque.

Par la suite, nous définissons un repère géométrique \mathcal{F}_0 associé à l'objet, et un repère \mathcal{F}_{ct} attaché à l'image scanner. Un autre repère \mathcal{F}_I d'échelle différente est également associé à l'image scanner, dont l'unité de chaque vecteur est le pixel au lieu du millimètre (voir sur la figure 3.44). Un point dans l'espace tel que l'origine du repère \mathcal{F}_0 est noté \mathbf{O} et un vecteur entre deux points comme \mathbf{OO}_i . La projection d'un vecteur dans un repère \mathcal{F}_0 est noté ${}^0\mathbf{OO}_i$ et la transformation rigide entre \mathcal{F}_{ct} et \mathcal{F}_0 est écrit ${}^{ct}\mathbf{OO}_i = \mathbf{R} {}^0\mathbf{OO}_i + \mathbf{t}$, avec \mathbf{R} une matrice de rotation et \mathbf{t} un vecteur de position. Nous notons également par Δ_i la $i^{\text{ème}}$ droite du marqueur. L'intersection de cette droite avec le plan de coupe du scanner noté (π) est (généralement) un point $\mathbf{P}_i = \Delta_i \cap \pi$. Le modèle géométrique qui relie les coordonnées de ${}^{ct}\mathbf{OP}_i = (x_i, y_i, z_i)^T$ exprimées dans le repère du scanner \mathcal{F}_{ct} et les coordonnées du point image ${}^I\mathbf{Q}_i = (u_i, v_i)^T$ dans le repère image \mathcal{F}_I est :

$$\begin{aligned}
 {}^{ct}\mathbf{OP}_i = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} &= \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_{\text{proj. orthogonale}} \underbrace{\begin{bmatrix} s_x & g \\ 0 & s_y \end{bmatrix}}_{\text{paramètres intrinsèques}} \underbrace{\begin{bmatrix} u_i \\ v_i \end{bmatrix}}_{\text{position pixel}} \\
 &= {}^{ct}\Pi_\pi \mathbf{S}_{ct} {}^I\mathbf{Q}_i
 \end{aligned} \tag{3.54}$$

où la matrice (3×2) ${}^{ct}\Pi_\pi$ représente la projection orthogonale sur le plan de coupe (π) dans sa forme canonique. Les entrées non nulles de \mathbf{S}_{ct} sont les valeurs des paramètres intrinsèques du

scanner (facteurs d'échelle positifs s_x et s_y , et un paramètre g de distorsion ("de cisaillement") permettant de modéliser l'erreur due à l'inclinaison du statif du scanner le cas échéant ou de la torsion de la table pendant l'examen). En général, ce paramètre de distorsion est très petit par rapport aux autres paramètres et est généralement négligé [ZW96], bien qu'il puisse être identifié, car significatif, dans le cas des scanners IRM [ZW96, BZWF02]. Dans la suite, nous négligerons également cet effet. En considérant la transformation euclidienne entre le repère scanner \mathcal{F}_{ct} et le repère \mathcal{F}_0 des droites de l'objet, l'expression du vecteur \mathbf{OP}_i dans \mathcal{F}_{ct} est telle que ${}^{ct}\mathbf{OP}_i = \mathbf{R} {}^0\mathbf{OP}_i + \mathbf{t}$ et l'expression

$${}^0\mathbf{OP}_i = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 \end{bmatrix} \mathbf{S}_{ct} \begin{bmatrix} u_i \\ v_i \end{bmatrix} + \mathbf{t}' \quad (3.55)$$

représente la transformation de $\mathcal{F}_I \rightarrow \mathcal{F}_0$, où le vecteur \mathbf{r}_j est la $j^{\text{ème}}$ colonne de la matrice de rotation \mathbf{R}^T et le vecteur $\mathbf{t}' = -\mathbf{R}^T \mathbf{t}$. Le produit $(\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 \end{bmatrix} \mathbf{S}_{ct})$ est une matrice (3×2) , $\mathbf{L} = [\mathbf{l}_1 \ \mathbf{l}_2]$, qui vérifie les contraintes quadratiques provenant des composantes de la rotation :

$$\mathbf{l}_1^T \mathbf{l}_1 = s_x^2, \quad \mathbf{l}_2^T \mathbf{l}_2 = s_y^2, \quad \mathbf{l}_1^T \mathbf{l}_2 = 0. \quad (3.56)$$

La $i^{\text{ème}}$ droite Δ_i de l'objet peut également être représentée par une origine \mathbf{O}_i (3 ddl) et un vecteur directeur unitaire \mathbf{y}_i (2 ddl, voir figure 3.44). Cette représentation, bien que redondante (car seulement 4 ddl sont nécessaires pour représenter une droite dans l'espace 3-D euclidien) permet d'exprimer la position d'un point $\mathbf{P}_i = \Delta_i \cap \pi$, par :

$${}^0\mathbf{OP}_i = {}^0\mathbf{OO}_i + \lambda_i {}^0\mathbf{y}_i, \quad \lambda_i \in \mathbb{R} \quad (3.57)$$

où \mathbf{O}_i est la projection orthogonale de l'origine du repère \mathcal{F}_0 sur la droite Δ_i , qui vérifie aussi $\mathbf{O}_i \mathbf{P}_i \times \mathbf{y}_i = \mathbf{0}$. On peut alors déterminer \mathbf{L} , \mathbf{t} et les scalaires $\{\lambda_i\}$ par la relation

$${}^0\mathbf{OO}_i + \lambda_i {}^0\mathbf{y}_i = \mathbf{L} \begin{bmatrix} u_i \\ v_i \end{bmatrix} + \mathbf{t}'. \quad (3.58)$$

Pour s'affranchir des $\{\lambda_i\}$, nous introduisons la représentation en coordonnées plückeriennes des droites Δ_i . Avec cette représentation, la définition explicite de l'origine sur une droite n'apparaît plus. En définissant un nouveau vecteur \mathbf{w}_i tel que $\mathbf{w}_i = \mathbf{y}_i \times \mathbf{OP}_i$, les coordonnées plückeriennes $\mathcal{L}_i = (\mathbf{y}_i, \mathbf{w}_i)$ d'une droite Δ_i sont

$$\begin{aligned} \mathbf{y}_i &\equiv \mathbf{O}_i \mathbf{P}_i \\ \mathbf{w}_i &= \mathbf{y}_i \times \mathbf{OP}_i = \mathbf{y}_i \times \mathbf{OO}_i \Rightarrow {}^0\mathbf{w}_i^T {}^0\mathbf{OP}_i = 0 \end{aligned} \quad (3.59)$$

ou encore

$$\underbrace{\begin{pmatrix} [{}^0\mathbf{y}_i]_{\times} & -{}^0\mathbf{w}_i \\ {}^0\mathbf{w}_i^T & 0 \end{pmatrix}}_{\mathbb{L}_i^*} \begin{pmatrix} {}^0\mathbf{OP}_i \\ 1 \end{pmatrix} = \mathbf{0} \quad (3.60)$$

sous forme homogène, avec $[{}^0\mathbf{y}_i]_{\times}$ la matrice (3×3) antisymétrique (non-inversible) associée au vecteur ${}^0\mathbf{y}_i$, et exprimée dans le repère \mathcal{F}_0 . La matrice (4×4) \mathbb{L}_i^* est antisymétrique, de rang 2, et est appelée la matrice duale (de Plücker) pour la droite Δ_i . La représentation

plückerienne précédente d'une droite étant contenue dans \mathbb{L}_i^* , la dernière expression de ${}^0\mathbf{OP}_i$ peut s'écrire à l'aide de l'équation homogène :

$$\mathbb{L}_i^* \begin{bmatrix} \mathbf{L} & \mathbf{t}' \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ \beta_i \end{bmatrix} = 0. \quad (3.61)$$

La valeur du scalaire β_i dans l'équation (3.61) est utile dans le cas particulier où une droite du marqueur est parallèle au plan de coupe ($\lambda_i \rightarrow \infty$). Dans les autres cas, β_i peut être choisie librement, à 1 pour le cas homogène, et on peut alors employer la forme inhomogène des relations géométriques :

$$[{}^0\mathbf{y}_i]_{\times} \left(\mathbf{L} \begin{bmatrix} u_i \\ v_i \end{bmatrix} + \mathbf{t}' \right) = {}^0\mathbf{w}_i. \quad (3.62)$$

L'équation (3.62) est à la base des méthodes d'estimation de pose pour ce type d'imager. En pratique, les droites de l'objet étant matérialisées par des tiges de matériaux radio-opaques, celles-ci sont de longueurs finies et connues ($\lambda^{\min} \leq \lambda \leq \lambda^{\max}$). Il est alors pertinent de calculer les valeurs extrêmes pour une pose (\mathbf{R}, \mathbf{t}) donnée afin de vérifier si l'intersection avec le plan de coupe existe. En réalité, un certain nombre de points peuvent manquer dans l'image et, à l'inverse, il peut y avoir accidentellement des artefacts dus à la présence d'autres objets [LFC02]. Ainsi, on peut vérifier la visibilité d'une droite du marqueur et valider ou non l'existence d'un point image. L'expression du vecteur \mathbf{OO}_i dans l'équation (3.58) permet d'isoler le scalaire λ_i . En développant et en pré-multipliant par ${}^0\mathbf{y}_i^T$, on obtient une expression qui ne dépend pas du vecteur ${}^0\mathbf{w}_i$:

$$\lambda_i = {}^0\mathbf{y}_i^T \left(\mathbf{L} \begin{bmatrix} u_i \\ v_i \end{bmatrix} + \mathbf{t}' \right). \quad (3.63)$$

On peut alors envisager d'effectuer des simulations, car si $\lambda_i^{\min} \leq \lambda_i \leq \lambda_i^{\max}$, les coordonnées du spot ${}^I\mathbf{Q}_i$ sont calculées par :

$$\begin{bmatrix} u_j \\ v_j \end{bmatrix} = \mathbf{S}_{ct}^{-1} \left(\mathbf{F}_j^T \mathbf{F}_j \right)^{-1} \mathbf{F}_j^T \left({}^o\mathbf{w}_j - [{}^o\mathbf{y}_j]_{\times} \mathbf{t} \right), \quad (3.64)$$

avec $\mathbf{F}_j = [{}^o\mathbf{y}_j]_{\times} \mathbf{R} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$,

Estimation du recalage stéréotaxique avec une seule vue

Nous indiquons ici la démarche aboutissant à l'estimation quasi-linéaire du recalage, ce qui suppose que la mise en correspondance entre les droites et les points image ait été réalisée au préalable (la mise en correspondance automatique est décrite dans [Mau05]). Il s'agit de résoudre le système (3.62) pour les inconnues \mathbf{I}_1 , \mathbf{I}_2 et \mathbf{t}' . En développant, on peut écrire les

relations suivantes :

$$\begin{aligned}
 {}^0\mathbf{w}_i &= [{}^0\mathbf{y}_i]_{\times} \left(\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 \end{bmatrix} \mathbf{S}_{ct} \begin{bmatrix} u_i \\ v_i \end{bmatrix} + \mathbf{t}' \right) \\
 &= [(u_i, v_i, 1) \otimes [{}^0\mathbf{y}_i]_{\times}] \mathbf{S}_9 \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{t}' \end{bmatrix} \\
 &= [(u_i, v_i, 1) \otimes [{}^0\mathbf{y}_i]_{\times}] \boldsymbol{\xi}
 \end{aligned} \tag{3.65}$$

avec \otimes le produit de kronecker sous sa forme matricielle et où la matrice \mathbf{S}_9 est telle que

$$\mathbf{S}_9^{-1} = \begin{bmatrix} \mathbf{S}_{ct}^{\top} & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \otimes \mathbf{I}_{3 \times 3} \tag{3.66}$$

et où $\mathbf{I}_{3 \times 3}$ est l'identité. En supposant que n droites et n points sont en correspondances, nous pouvons envisager la résolution du vecteur $\boldsymbol{\xi} = (\mathbf{l}_1^{\top}, \mathbf{l}_2^{\top}, \mathbf{t}'^{\top})^{\top}$ comme un problème d'optimisation avec une contrainte d'égalité (quadratique). Nous fournissons également dans [Mau05] une approche numérique de type Newton-Raphson. Dans le cas d'un scanner calibré, ce recalage 2-D/3-D est réduit à un recalage rigide qui peut être résolu avec seulement 4 correspondances (2 solutions) en considérant le vecteur $\mathbf{x} = (\mathbf{r}_1^{\top}, \mathbf{r}_2^{\top}, \mathbf{t}'^{\top})^{\top}$, qui ne contient que les paramètres de pose.

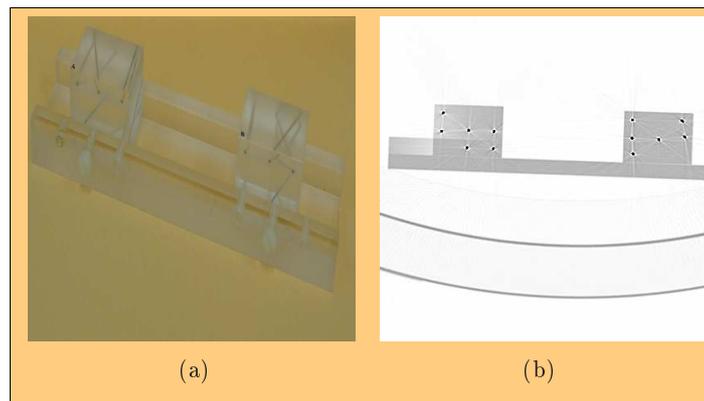


Fig. 3.45: (a) Les deux cubes utilisés pour les expériences, fixés sur un rail de guidage. (b) Image des deux cubes, extraite d'une séquence d'acquisition hélicoïdale.

Positionnement d'une aiguille

Dans cette expérience, le système robotique CT-Bot complet et étalonné (plate-forme, porte-aiguille et aiguille) a été sanglé sur un mannequin de façon que le plan de coupe du scanner traverse le marqueur stéréotaxique et que l'aiguille montée sur ce dispositif soit visible (voir la figure 3.47). La connaissance de la position de la pointe de l'aiguille n'est pas accessible précisément car celle-ci a été placée manuellement dans le porte-aiguille. Nous avons cherché à approcher cette pointe vers une cible correspondant à un point d'entrée à la surface de la peau du mannequin. La cible a été matérialisée par une seconde aiguille de biopsie d'un faible diamètre (1 mm), fixée à la peau du mannequin et placée perpendiculairement au plan de

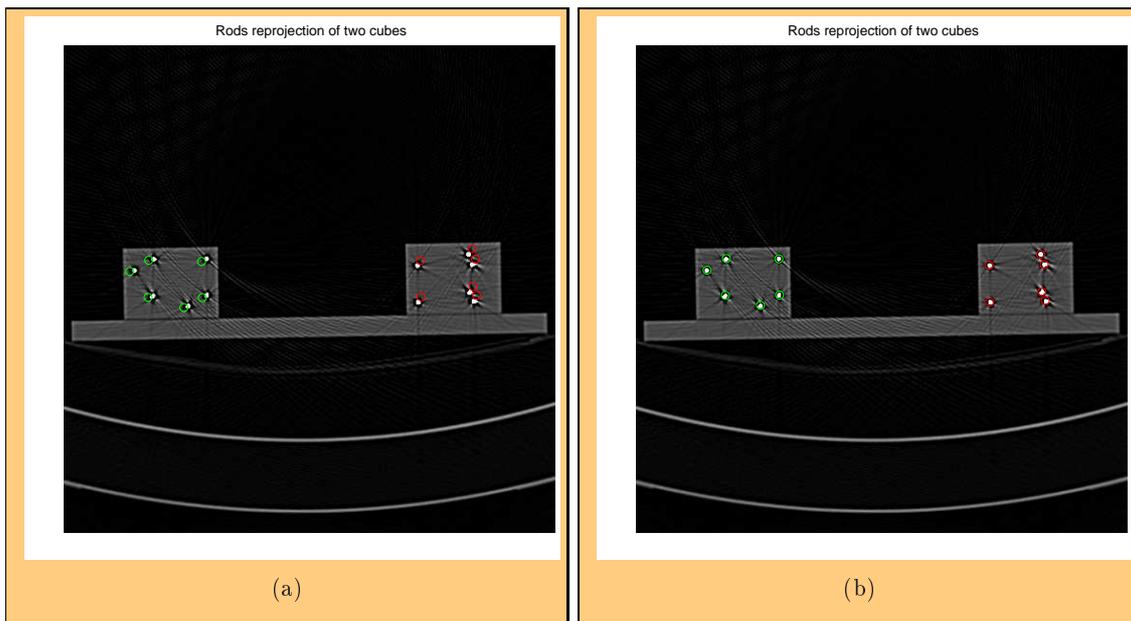


Fig. 3.46: Recalage indépendant (en vert et en rouge) des deux marqueurs cubiques de la figure 3.45. Les centres des cercles indiquent l'erreur de localisation dans l'image scanner, en considérant le recalage rigide estimé. (a) Estimation (de \mathbf{x}) par les moindres carrés (b) Estimation par l'approche numérique de Newton-Raphson.

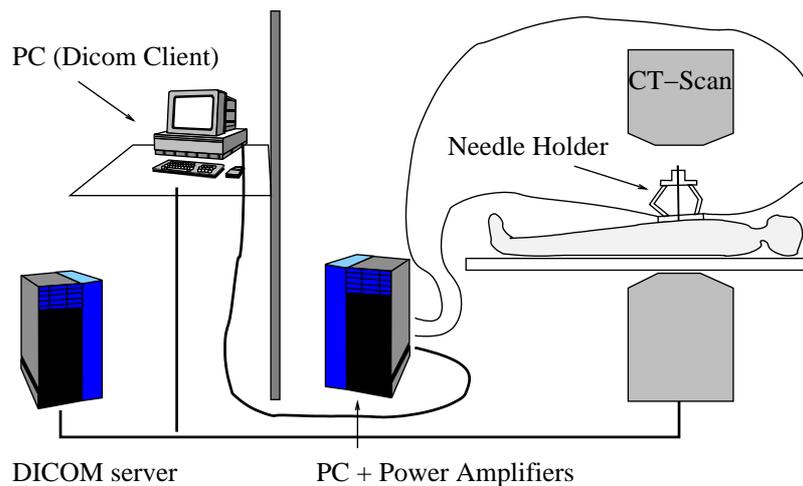


Fig. 3.47: Schéma global de l'installation pour la radiologie interventionnelle robotisée.

coupe (voir figure 3.49). Le positionnement est du type *look and move*, et la procédure est décomposée de la manière suivante :

1. mise en place du système dans sa configuration opérationnelle,
2. acquisition d'une image scanner (figure 3.48),
3. recalage rigide automatique entre le scanner et le robot (figure 3.50),
4. localisation dans l'image de la pointe de l'aiguille et de l'aiguille-cible,
5. calcul du mouvement relatif dans l'espace opérationnel (à orientation constante),
6. envoie de la commande au contrôleur du robot,
7. vérification visuelle par l'acquisition d'une nouvelle coupe scanner (figure 3.49).

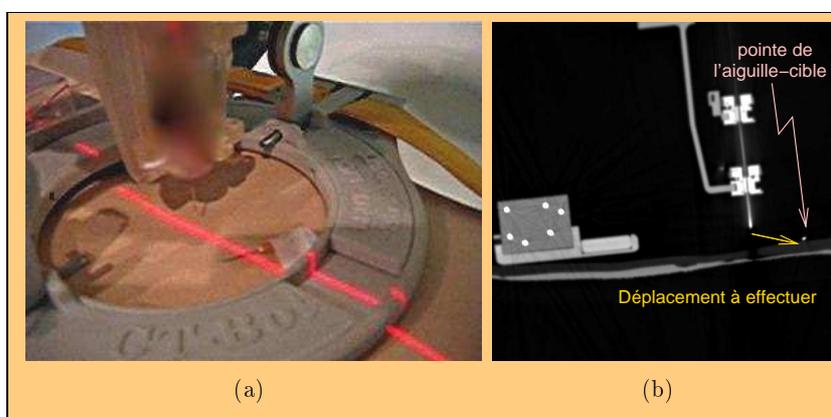


Fig. 3.48: Positionnement initial d'une aiguille de radiologie par recalage scanner à l'aide d'une seule vue. (a) Situation du manipulateur avant l'envoi de la consigne. (b) Image obtenue dans cette configuration.

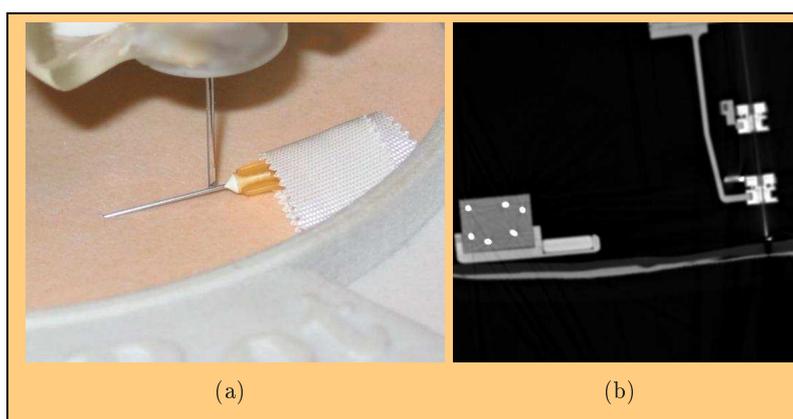


Fig. 3.49: Résultat du positionnement relatif. (a) La pointe de l'aiguille entre en contact avec la cible matérialisée par une aiguille. (b) Image obtenue dans cette configuration.

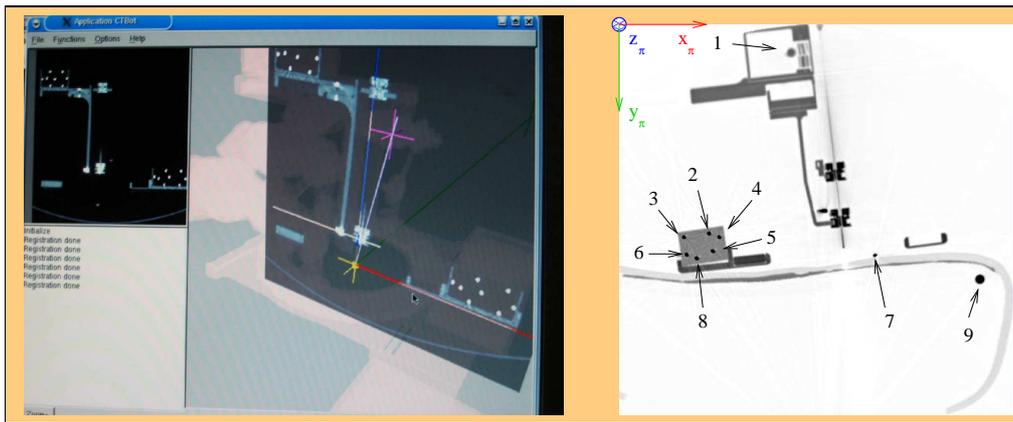


Fig. 3.50: (a) L'interface graphique développée pour la procédure de positionnement de l'aiguille. (b) Une image (format DICOM) incluant la détection de nombreux spots. Le recalage automatisé basé sur une approche de type RANSAC permet de rejeter les points aberrants et de résoudre les appariements droites/points.

Conclusion sur le recalage stéréotaxique

Nous avons fourni ici une méthodologie pour considérer le problème du recalage stéréotaxique à partir d'une coupe scanner comme un problème classique d'étalonnage de caméra et d'estimation de pose. Nous avons délibérément choisi une approche monoculaire afin d'envisager dans un futur proche l'analyse en temps réel des images scanner grâce au mode fluoroscopique, qui apparaît sur les récents dispositifs d'acquisition. La modélisation géométrique proposée pour formaliser le recalage a été validée par de nombreux résultats, qui sont répertoriés dans le mémoire de thèse de Benjamin Maurin [Mau05]. Il ressort de ce travail collaboratif important (qui par ailleurs décrit la modélisation, la conception, la réalisation et la mise en oeuvre du robot CT-Bot et de sa commande) que la précision du positionnement d'une aiguille de radiologie est d'environ 2 mm, le recalage stéréotaxique assurant une précision de position de l'ordre du mm et de 2 degrés pour l'orientation.

L'ensemble du volet correspondant au recalage stéréotaxique a permis la rédaction d'articles dans des congrès nationaux (Surgetica'05 [MDG⁺05], Orasis'05 [MDMG05] et RFIA'06 [DMBM06]), dans des conférences internationales, dont IEEE CVPR'03 [MDMG03] et MICCAI'04 [MGB⁺04]. Un article de revue a été soumis à IEEE Transactions on Medical Imaging et est à l'heure actuelle en révision.

Une application immédiate de ce travail concerne la compensation des petits mouvements du patient par la mise en oeuvre d'un suivi et d'un asservissement visuel sur les images peropératoires. Les bases de cette extension contiennent notamment une modélisation cinématique, à partir de laquelle nous montrons quels sont les arrangements relatifs des barres du marqueur stéréotaxique qui aboutissent à des configurations singulières de la matrice jacobienne assimilable à une matrice d'interaction dans ce cas (communication dans la conférence francophone RFIA en 2006 [DMBM06]).

D'un point de vue pratique, nous souhaiterions étendre ce travail en prenant en compte également l'image de l'aiguille avec le plan de coupe scanner. Cette information, bien que

partielle du point de vue du positionnement, permettrait de valider les corrections apportées par le mouvement des marqueurs qui se situent sur la plate-forme porte-aiguille, mais pas à l'extrémité de l'organe terminal.

Le passage au mode fluoroscopique afin de pouvoir corriger en temps réel la position de l'aiguille se heurte à des problèmes à la fois technique mais également à la sécurité des patients soumis à une très forte dose de rayonnement. Cependant, il faut garder de l'espoir, les améliorations apportées depuis plus de trente ans sur les images acquises par les scanners ont toujours été accompagnées d'une réduction des rayonnements émis.

3.7 RECONSTRUCTION 3-D PAR CODAGE DE LA LUMIÈRE STRUCTURÉE

Disposer de l'information tridimensionnelle en temps réel, relative à l'environnement dans lequel opère un chirurgien par l'emploi de la lumière structurée permet, par exemple dans le cadre d'interventions mini-invasives, de restituer visuellement la profondeur [HSN06] de la sène abdominale et, en s'appuyant sur la réalité augmentée, d'offrir au chirurgien un plus grand confort et une plus grande précision dans son geste. La connaissance en temps réel de la géométrie 3-D de la cavité abdominale pourrait aussi permettre de piloter des systèmes robotisés dans le but d'assister le chirurgien et d'aboutir à une plus grande sécurité pour le patient.

C'est ainsi que depuis deux ans et à travers les travaux de thèse de Chadi Albitar, nous nous intéressons à la conception de dispositifs actifs capables de fournir des reconstructions 3-D des structures internes de l'abdomen en temps réel, afin de pouvoir incorporer cette information dans les systèmes d'assistance à la chirurgie mini-invasive.

Dans le passé, des systèmes de capture 3-D à travers un endoscope ont été étudiés. Ainsi, les travaux assez anciens d'Okatani et Deguchi [OD97] ou plus récemment ceux de Stoyanov [DY04] ont contribué à acquérir une information 3-D par la réalisation de systèmes endoscopiques passifs. Comme travaux voisins au notre, on peut citer également ceux d'Hayashibe [HSN06], d'une part, nos travaux antérieurs lors de la thèse d'Alexandre Krupa [KGD⁺03], d'autre part, qui ont mis en œuvre des systèmes actifs de pointage laser pour extraire une information 3-D, respectivement à une fréquence de 6 images/s pour la reconstruction de surfaces à l'aide d'un galvano-scanner [HSN06] et en temps réel vidéo pour l'estimation de la distance de l'organe pointé par la projection d'un faisceau parallèle de 4 rayons laser.

Parmi toutes les approches habituellement mentionnées pour la reconstruction des surfaces et faisant appel à la lumière structurée, nous avons proposé d'utiliser une solution basée sur la projection d'un motif lumineux composé d'éléments géométriquement différents qui, une fois projetés sur la scène, sont imagés par un endoscope. Le principe d'un tel système est comparable à celui d'un système stéréoscopique traditionnel pour lequel on remplace une des caméras par un projecteur de motifs lumineux. Ainsi, pour déterminer la géométrie 3-D de la scène observée, on projette un motif de lumière sur la surface à reconstruire et on détermine par triangulation les coordonnées 3-D des points de la surface, matérialisés par la projection du motif. Ceci n'est évidemment possible que si l'on a pu mettre en correspondance les éléments du motif projeté et leurs positions respectives dans l'image. Il est donc nécessaire de pouvoir associer un code à chaque élément composant le motif, de manière à pouvoir les identifier sans ambiguïté dans l'image. Par conséquent, aussi bien la conception du motif que la stratégie de codage associée de ses éléments sont guidées par l'obtention rapide et robuste des appariements. Le problème lié au choix du motif utilisé et au codage qui lui est associé n'est pas nouveau [SBM98, CLZQ03], mais nous avons proposé une amélioration substantielle par rapport à ce qui existe par ailleurs, amélioration qui concerne le taux d'erreurs (données manquantes ou aberrantes) acceptable et la possibilité de retrouver des éléments manquants. En effet, dès que l'on s'intéresse à des scènes non contrôlées, comme cela est le cas pour des images acquises à l'intérieur de l'abdomen, on se trouve confronté aux problèmes classiques que peut poser la présence d'ombres, d'occultations ou de zones à courbure forte. De telles

situations peuvent produire des désordres ou des pertes d'information entre les composants du motif projeté et ainsi produire des erreurs lors de la phase de décodage. Pour résoudre ce problème, il est nécessaire d'utiliser des codages permettant de corriger ces erreurs [MOC⁺98]. En d'autres termes, il faut utiliser des codes autorisant une discrimination forte entre les mots. La distance de Hamming est un critère bien connu en théorie de l'information et sert à quantifier la différence entre deux séquences de symboles. Elle est utilisée ici pour caractériser la différence entre deux mots du code, correspondants à deux éléments distincts du motif.

L'approche retenue s'appuie sur le voisinage spatial et la théorie des matrices parfaites. Il s'agit d'une matrice de dimensions $n \times m$ où chaque élément de M prend sa valeur dans un alphabet de k symboles. M est appelée matrice parfaite car chaque sous-matrice de dimensions $r \times v$ n'apparaît qu'une seule fois dans M . Si de plus, M exclut la sous-matrice ne contenant que des valeurs nulles, M est appelée matrice pseudo-aléatoire ou *M-array*.

La robustesse liée à l'identification des éléments d'une matrice parfaite vient de sa capacité à identifier (décoder) les parties visibles grâce aux propriétés intrinsèques des *M-array*. Les travaux menés en ce sens (notamment [MYS88, PBT⁺92, GNY92, YG94, PSSM00]) se distinguent essentiellement par la taille des matrices proposées, l'utilisation ou non de la couleur et par la taille de l'alphabet utilisé.

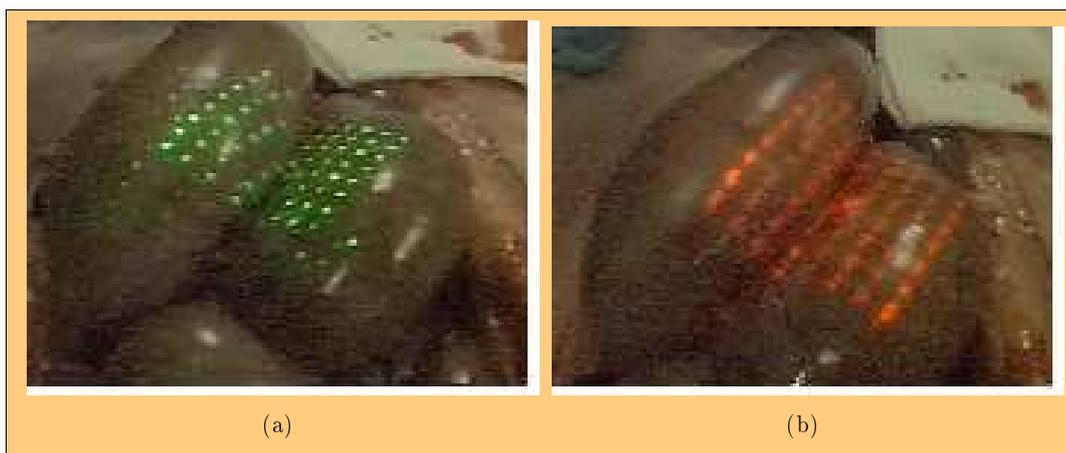


Fig. 3.51: Projection d'une matrice de points sur un organe de l'abdomen. (a) Avec une lumière verte. (b) Avec une lumière rouge.

Après avoir étudié l'interaction de la lumière sur les organes (voir la figure 3.51), en particulier les propriétés d'absorption et de diffusion, nous avons choisi une longueur d'onde d'émission correspondant à une couleur verte pour illuminer le motif. Le motif est donc monochromatique. D'autres travaux ont préféré associer des teintes de couleur aux symboles du motif, mais en visant d'autres applications que l'endoscopie [CLZ07]. Les symboles de notre motif sont associés à des primitives géométriques. Parmi les primitives considérées (voir la figure 3.52), nous avons introduit une primitive directionnelle afin de décrire l'orientation locale et contraindre la recherche des voisinages lors de la phase de décodage, phase qui a pour objet d'effectuer une rapide et efficace mise en place des appariements entre primitives.

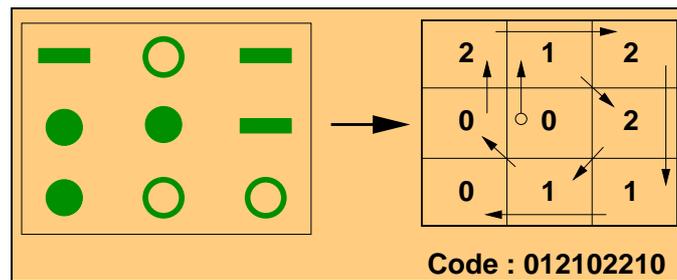


Fig. 3.52: Détermination du code de chaque sous-matrice 3×3 du motif.

Des premiers résultats ont été obtenus et un motif structurant, de très faible dimension, composé de 783 éléments (3 primitives différentes), a été réalisé grâce à la technologie de l'optique diffractive et fabriqué par micro-gravure.

Des problèmes, tels que le choix de la couleur, celui des primitives géométriques associées au code, la segmentation, le décodage et la reconstruction 3-D, ont tous été surmontés et les premiers résultats de segmentation et de reconstruction 3-D obtenus à partir d'acquisitions de surfaces à fortes variations de courbure et à partir d'organes de porcs viennent illustrer ce travail et permettent de valider qualitativement la méthode proposée. Par ailleurs, en choisissant la distance de Hamming comme critère directeur pour la conception du motif mais aussi comme critère de robustesse en présence de données aberrantes, nous fournissons une évaluation quantitative du taux maximal d'erreurs acceptable par l'ensemble du système de vision active.

Il est fort probable que le lecteur ait perçu que je ne souhaitais pas apporter d'autres illustrations sur ce sujet afin de laisser à Chadi le plaisir de présenter lui-même ses résultats de thèse de doctorat. De plus, nous avons récemment appris que ce travail pourra être présenté oralement lors de la conférence IEEE ICCV 2007 [AGD07c], ce qui nous permet d'envisager sereinement la troisième année de thèse. Néanmoins des résultats préliminaires ont été présentés lors des journées Orasis 2007 et peuvent être consultés [AGD07a].

Les perspectives à cours terme de ce travail concernent l'expérimentation *in vivo* par des tests de reconstructions 3-D métriques, durant une séquence. A plus longue échéance, nous mènerons une réflexion sur l'élaboration de la commande de robots par vision active codée (voir les perspectives du dernier chapitre) et la mise en place d'outils logiciels de réalité augmentée.

3.8 BILAN

Pour résumer l'ensemble de ces travaux, le problème de l'estimation de pose et plus généralement du recalage 2-D/3-D rigide a été abordé en empruntant plusieurs approches et différentes représentations, en fonction du contexte. Que ce soit à partir du filtre estimateur de Kalman avec des primitives géométriques non-linéaires, ou à partir des invariants projectifs avec des spots lasers et des marqueurs optiques, ou encore par une approche d'optimisation numérique basée sur les asservissements visuels virtuels et combinant des primitives géométriques différentes, de la modélisation des quadriques dégénérées pour la pose des instruments cylindriques ou finalement qu'il s'agisse du recalage stéréotaxique par une estimation robuste avec des primitives linéaires, le recalage 2-D/3-D rigide a été étudié tout au long de ces travaux de recherche, pour plusieurs types de dispositifs d'imagerie (caméra classique, caméra endoscopique, scanner à rayons X).

Le guidage 3-D a également été vu en réalisant uniquement un suivi qui pouvait être mis en oeuvre dans un environnement complexe (sous quelques hypothèses) sans passer par une étape de recalage. Ceci ne permet pas d'effectuer un positionnement d'instrument, mais d'offrir une assistance au chirurgien par le suivi automatique d'une zone d'intérêt à l'aide d'un endoscope articulé.

D'un point de vue pratique, il reste de nombreux aspects pour lesquels nous n'avons pas encore pu mener une investigation suffisante, certaines mises en oeuvre des techniques pré-citées étant très récentes. Il s'agit entre autre d'effectuer des tests *in vivo* pour le positionnement d'une aiguille circulaire lors de son passage à travers les tissus, dans le développement de l'assistance robotisée de la suture, ou encore de la détection directe de l'aiguille de radio-fréquence pour aboutir à un positionnement plus précis (pour certains degrés de liberté seulement) vis-à-vis du scanner (et pas uniquement au travers du recalage des marqueurs stéréotaxiques placés sur la plate-forme du robot). Il s'agit aussi de l'extension des techniques aux instruments non-marqués ou dotés d'un marquage adéquat (qui ne rend pas inopérantes les fonctionnalités propres de l'instrument de chirurgie). Bien que sur ce dernier point, nous avons proposé une technique de segmentation par l'emploi d'attributs de couleur, il reste à effectuer des tâches telles que le suivi et le positionnement de l'extrémité des instruments en vue de réaliser la saisie de l'aiguille ou d'opérer, sans l'aiguille, un contact direct avec les organes par exemple.

Un des objectifs à garder à l'esprit pour l'assistance au geste chirurgical est que le positionnement d'un instrument n'est pertinent que par rapport à l'organe sur lequel il doit interagir, et non par rapport à un capteur visuel. Or les organes présentent des surfaces qui ne semblent pas pouvoir être modélisées facilement par des méthodes exclusivement locales, la texture jouant un rôle important pour identifier et localiser les zones d'intérêt, sièges des interventions. Nous nous intéressons donc également à d'autres techniques existantes de positionnements relatifs telles que les asservissements élaborés à partir des approches du suivi visuel basées sur l'apparence au sens large afin de pouvoir conjuguer positionnement 3-D avec modèle géométrique ou CAO de l'objet d'intérêt d'une part (typiquement l'instrument de chirurgie) et asservissement visuel sans modèle géométrique de l'objet d'intérêt texturé (typiquement la surface des organes) d'autre part, comme une voie à explorer. Il est à noter que nous contribuons scientifiquement à l'extension de certaines de ces méthodes, notamment par la détection automatique des occultations [BDZM07].

Parallèlement, nous sommes également en train d'explorer une autre voie, en cherchant à modéliser les surfaces des organes à l'aide de la reconstruction 3-D par la lumière structurée et codée (thèse de Chadi Albitar en cours). Cette autre approche, outre le fait qu'elle peut également s'appliquer à la réalité augmentée, un domaine connexe au notre et dont des compétences existent au sein du consortium IRMC (équipes IGG du LSIIT et Virtuals de l'IRCAD), devrait permettre d'élaborer des lois de commande de robots pour endoscopes articulés et surmonter les problèmes rencontrés avec les techniques précédentes avec des zones peu texturées, et d'une manière générale à contenu informatif faible. Des premiers résultats de reconstruction 3-D en temps réel (< 50 ms) sur des surfaces présentant des discontinuités de courbure ont été obtenus cette année. Cette nouvelle action de recherche est soutenue par l'ULP qui nous a accordé un soutien financier en 2006 lors d'un appel à projets du conseil scientifique (BQR : Bonus Qualité Recherche). Je suis co-porteur de ce projet avec mon collègue Pierre Graebling (Prof. ULP/ENSPS), projet qui associe l'équipe AVR du LSIIT, le Laboratoire des Systèmes Photoniques de Strasbourg et l'IRCAD.

Chapitre 4

Perspectives de recherche

L'imagerie médicale est un moyen efficace d'assistance au clinicien pour aboutir à un diagnostic et au praticien pour guider le geste chirurgical. Les développements récents de la robotique, des technologies d'imagerie médicale et des techniques d'analyse pré-opératoires et per-opératoires ont permis de réaliser des progrès considérables dans l'assistance aux gestes médicaux et chirurgicaux. Dans ce domaine, les nombreux travaux sur le recalage rigide et non-rigide, ont apporté une contribution significative.

Ces progrès permettent d'envisager de nouvelles formes d'examens médicaux, plus précis et permettant un meilleur suivi du patient, et de nouvelles formes d'interventions chirurgicales, plus fortement assistées, plus sûres, mais aussi de prévoir une plus large planification de ces interventions.

4.1 PERSPECTIVES

Les travaux de recherche que nous avons effectués ont permis de proposer des modélisations mathématiques très détaillées du recalage 2-D/3-D pour plusieurs dispositifs d'imagerie, avec ou sans modèle de l'objet d'intérêt. Le positionnement précis des instruments par rapport aux organes en mouvement, l'interaction des instruments avec ces mêmes organes que nous avons abordés *in vitro*, requièrent la mesure simultanée de l'évolution des uns par rapport aux autres. Ainsi, le suivi sans modèle des régions d'intérêt va de pair avec le suivi avec modèle des instruments qui sont faciles à répertorier dans une base de connaissances. Nous avons montré que nous savions réaliser un guidage avec l'un ou l'autre. Finalement, les résultats que nous avons obtenus montrent la cohabitation de plusieurs disciplines : de la géométrie pour le recalage, aux traitements du signal causal ou non-causal pour les asservissements visuels et la segmentation des images. En reprenant certains des points mentionnés dans le bilan global présenté à la fin du chapitre précédent, nous proposons par la suite quelques perspectives qu'il serait souhaitable d'explorer à court ou moyen terme.

Multiplication et répartition des tâches visuelles

La navigation dans les environnements difficiles à appréhender tels que les environnements naturels ou l'intérieur du corps humain demandent de prendre en considération en permanence les modifications du milieu. La définition de tâches visuelles très diverses comme la détection des ombres, des spécularités, la recherche des zones texturées, la modélisation statistique de l'intensité de lumière, de la couleur, contribuera à améliorer le guidage, en participant, à des degrés différents et dans des étapes différentes, à la détection des objets d'intérêt. La gestion simultanée de plusieurs tâches visuelles, et particulièrement la distinction entre les tâches ne relevant, à un moment donné, que du suivi 2-D ou 3-D et celles devant être prises en compte, dans le même temps, dans l'élaboration d'un asservissement visuel, constitue une nouvelle étape dans l'appréhension des environnements non maîtrisés. En d'autres termes, le nombre de tâches de suivi à concevoir dans le dispositif de guidage doit être plus important qu'il ne l'est actuellement afin qu'un module *répartiteur* puisse élire, durant un laps de temps, le sous-ensemble des tâches qui doit être pris en considération pour l'élaboration d'une commande. La modification d'une trajectoire, due aux variations du milieu peut se traduire par une nouvelle définition des consignes successives (planification dynamique), comme cela se fait en robotique mobile [Kha86, Tso92, DCF⁺94, ZO02, SN04] mais également en asservissement visuel [Zan03, NZDM06].

Interactions avec les organes - modélisation

Spécifiquement pour la GMCAO, la compréhension des déformations des organes pris individuellement et dans leur ensemble, chacun d'entre eux ayant des propriétés mécaniques particulières (qui dépendent aussi de l'état de santé du patient), nous permettrait d'améliorer les performances d'interaction des instruments avec les organes [DY04]. En effet, comme le font souvent les chirurgiens, il est tout à fait envisageable de déplacer, à l'aide des instruments, tout un ensemble d'organes, pour parvenir ou maintenir visible le siège de l'opération, alors qu'il faut éviter généralement tout contact imprévu avec des éléments particuliers comme

les artères ou le système nerveux. La précision du positionnement n'est pas toujours le facteur le plus prépondérant dans l'élaboration d'une tâche en interaction avec les tissus. Comme nous nous en sommes aperçus avec l'aide à la suture, l'accès à toutes les informations 3-D pertinentes pour le bon déroulement de ce geste a été un frein au développement de ce travail. Le positionnement des instruments demande, certes, un degré de précision important à cause de l'effet de levier [NZDM06] en chirurgie laparoscopique par exemple et des grossissements mis en jeu dans les capteurs, mais il est surtout rendu difficile par la grande variabilité des images, celle-ci émanant du mouvement d'ensemble et des déformations des organes, et dont peu d'éléments nous permettent de quantifier ce qui est tolérable par les tissus de ce qui ne l'est pas. Il y a là, selon moi, à définir la navigation des instruments par rapport aux organes, non pas de manière systématique en évitant tout contact, mais en intégrant les interactions non dommageables comme des événements identifiés et intrinsèques au processus de guidage.

La vision active codée et ses applications en robotique

Si l'utilisation de scanners ou d'échographes est aujourd'hui courante, ces dispositifs ne constituent pas toujours les modalités d'imagerie les plus disponibles ni les plus appropriées. En effet, lorsqu'il s'agit de faire un diagnostic concernant l'état d'un organe ou d'une région de celui-ci en particulier, ou lorsqu'il faut pratiquer une intervention chirurgicale à l'intérieur de la cavité abdominale, il serait souhaitable de disposer d'un système d'imagerie que l'on puisse situer le plus près possible de la zone d'intérêt, qui restitue au médecin l'image la plus fidèle de la zone imagée, et qui puisse fournir des informations supplémentaires, telles que la géométrie 3-D de l'environnement par exemple. Un tel dispositif permettrait, d'une part, d'observer et de caractériser plus précisément la zone d'intérêt en évitant des couches intermédiaires qui affectent la qualité de la perception, et d'autre part, permettrait de localiser des instruments chirurgicaux, lors d'interventions, de manière à pouvoir contrôler ses déplacements à l'aide d'un système robotique. Il y aura alors lieu d'élaborer avec ce dispositif actif, des stratégies de commande par vision qui conservent les régions d'intérêt dans le champ de vision [CM99, MC02b, SM06].

Un système basé sur l'endoscopie active est non seulement capable de remplir ce rôle, mais il est de plus capable de fournir des informations supplémentaires non détectables par d'autres techniques (couleur, texture). Il peut être utilisé, soit pour établir un diagnostic (endoscopie diagnostique), soit pour traiter chirurgicalement une maladie ou un traumatisme (endoscopie opératoire). De par sa proximité avec l'organe à examiner, l'endoscopie apporte des informations complémentaires à l'imagerie à balayage, pouvant alors aboutir à un meilleur diagnostic (plus précis ou plus précoce) et offre aussi de nouveaux espoirs de suivi de traitement pour de nombreuses maladies, dont le cancer, car l'identification précise des lésions facilite le recours à la chirurgie (biopsie, ablation, hémostase) qui peut elle-même être assistée par un robot en fonction des informations fournies par l'endoscope. De plus, les systèmes endoscopiques sont moins coûteux et plus accessibles que les scanners. Je souhaite donc pouvoir contribuer durant les prochaines années à renforcer l'usage de l'endoscopie active dans l'assistance au diagnostic et à la thérapie. Cela inclut implicitement de s'attaquer aux limitations qui font que l'intégration de techniques basées sur l'endoscopie dans les systèmes d'assistance n'est pas encore (ou trop peu) systématique.

Pour ce faire, il est nécessaire de développer à la fois des instruments d'observation et des méthodes propres à l'endoscopie active. Des efforts doivent être portés autant sur le premier volet, celui de l'innovation technologique, que sur le second pour lequel la conception de nouvelles méthodes de détection [AGD07c], de recalage et de contrôle robotique des déplacements par retour visuel, spécifiques à cette modalité d'imagerie commencent à voir le jour [HSN06, PCCS06]. De nombreux problèmes restent cependant à résoudre. Ainsi, la modélisation des surfaces acquises, l'élaboration de commande 3-D dont les surfaces pourraient constituer les consignes, vont mettre en évidence des problèmes de convergence. Il n'est pas interdit de penser que les asservissements visuels de ce type pourraient être la source d'amélioration de la convergence des algorithmes de recalage 3-D/3-D comme l'ICP (*Iterative Closest Point*) ou ses variantes [NN96, JH03].

Au-delà de son application à l'endoscopie, la vision active et codée pourrait apporter des solutions intéressantes au délicat problème de l'interprétation d'image de grande variabilité en cherchant à contrôler la complexité algorithmique du processus d'appariement de primitives, par exemple.

Cela soulève aussi des problèmes de commande qui conditionnent le bon déroulement d'un guidage par endoscopie. Si les modèles rencontrés en endoscopie rigide sont maintenant bien identifiés, il y a de nombreux problèmes de commande à résoudre en endoscopie flexible.

De tels ingrédients devraient permettre d'améliorer la perception visuelle de l'intérieur du corps humain à l'avenir, en restituant la profondeur de l'environnement et par superposition d'indices visuels. Dans ce cas, la vision 3-D endoscopique présente un intérêt en termes de réalité augmentée. D'autre part, la connaissance 3-D en temps réel de l'intérieur de la cavité abdominale ou plus généralement de la surface d'un organe sera très utile pour accomplir des actes chirurgicaux robotisés. Actuellement, un robot déplaçant les instruments chirurgicaux ne dispose généralement que de peu d'informations pour reconstruire l'espace libre et ainsi planifier des trajectoires et les exécuter. Les retombées de ces travaux sont évidemment d'aboutir à un plus grand confort pour le praticien et le patient, à une sécurité accrue, à une plus grande précision dans l'exécution et à une plus grande flexibilité dans la planification des tâches robotiques et dans l'interface informatique assistant le geste chirurgical de demain.

4.2 CONCLUSION GÉNÉRALE

La lecture de ce rapport a, je l'espère, montré qu'autour du guidage par vision, un spectre assez large de problématiques pouvait être abordé. Bien que j'ai essayé de situer le recalage per-opératoire au coeur de nos travaux, les diverses contributions que j'ai développées ici dévoilent mon souci permanent de m'attaquer aux problèmes qui doivent être surmontés préalablement au recalage et au suivi visuel pour appréhender des environnements de moins en moins maîtrisés.

Après un chapitre introductif sur mon laboratoire d'accueil, certains de ses projets et les motivations qui m'ont guidé durant mes activités de recherche, le chapitre II a présenté les problématiques principales autour du guidage par vision, à savoir, selon moi, le recalage, les asservissements visuels, le suivi et plus succinctement la segmentation d'images en temps réel. Dans le chapitre III, une description, parfois brève, et une synthèse des travaux menés depuis 1994 (et surtout depuis 2000) ont été exposées avec un bilan empreint de perspectives encourageantes, dont certaines sont reprises dans cet ultime chapitre.

Issu de la communauté de l'automatique et de la robotique, j'ai, au fil des années, orienté de plus en plus mes travaux vers la vision par ordinateur, sans jamais avoir souhaité couper ce cordon ombilical. De par les compétences reconnues en sciences de l'image du LSIIIT, j'ai pu réaliser cela avec l'aide précieuse des doctorants et des stagiaires que j'ai eu la chance de co-encadrer et grâce aussi aux compétences de mes collègues de l'équipe AVR et de l'équipe MIV, associés à ces travaux. Par conséquent, il a été une évidence pour moi qu'il fallait maîtriser les techniques employées en traitement d'images dont l'application pouvait être envisagée en vision robotique, aider à leur diffusion, voire dans certains cas contribuer à concevoir et valider des algorithmes nouveaux. Dans ce panel, le cas des contours actifs est particulier et également révélateur, les temps de calculs importants et les problèmes encore nombreux de convergence ne contribuent pas à étendre leur champ d'applications. Dans l'état actuel des techniques permettant de les mettre en oeuvre, trop peu d'entre elles peuvent prétendre à pouvoir s'appliquer à une application robotique par vision dynamique.

Par opposition, on observe que de plus en plus de traitements faisant appel à la morphologie mathématique sont présents dans les modules de vision d'un système robotique. Nous avons d'ailleurs présenté ici une application de la squelettisation à la détection des aiguilles de suture en laparoscopie, d'autres traitements de la même famille ont également été employés dans nos travaux.

La vision robotique est une discipline qui a muri depuis que j'ai débuté ma thèse de doctorat en décembre 1990. Bien des concepts ont été validés depuis, à la fois en s'appuyant sur une théorie et par la multitude d'expériences qui attestent de leur pertinence et de leur applicabilité. C'est sans doute parce que la place qu'occupent les asservissements visuels dans mon esprit est privilégiée que je trouve que de nombreuses avancées notables ont été réalisées depuis presque deux décennies à leur égard, toujours est-il que nous possédons aujourd'hui de plus en plus d'éléments qui nous permettent de naviguer et d'interagir de plus en plus finement et sûrement dans des environnements complexes.

D'autres défis scientifiques sont alors à relever. En premier lieu, à travers l'innovation techno-

logique comme la conception de dispositifs d'imagerie endoscopique rigide et flexible que nous explorons au laboratoire. Sur le plan méthodologique, très peu de techniques existent à ce jour pour piloter un robot chirurgical à l'aide d'une mesure tridimensionnelle, précise, et rapidement disponible de l'environnement. De ce fait, malgré son caractère relativement local, l'imagerie par lumière structurée et codée ouvre une voie de recherche inexplorée en robotique médicale. Enfin, pour ce qui est de l'intérêt clinique, nos perspectives suggèrent une avancée notable dans les techniques de recalage sur des structures anatomiques non préparées, c'est-à-dire non équipées de capteurs. Les bénéfices de ce travail se situeront au niveau de l'imagerie médicale per-opératoire (recalage d'image 3-D à l'aide de projections de lumière structurée en temps réel) dans le but de fournir une assistance de plus en plus ciblée des gestes chirurgicaux.

BIBLIOGRAPHIE

- [AB94] R. Adams et L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6) :641–647, June 1994.
- [ABF99] G. Aubert et L. Blanc-Féraud. Some remarks on the equivalence between 2d and 3d classical snakes and geodesic active contours. *Int'l Journal of Computer Vision*, 34(1) :19–28, Septembre 1999.
- [ABL⁺04] L. Ascari, U. Bertocchi, C. Laschi, C. Stefanini, Antonina Starita, et P. Dario. A segmentation algorithm for a robotic micro-endoscope for exploration of the spinal cord. Dans *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 491–496, New Orleans, LA, April 2004.
- [AD02] M. Agrawal et L.S. Davis. Camera calibration using spheres : A semi-definite programming approach. Dans *IEEE Int'l Conf. on Computer Vision*, tome 2, pages 782–789, Nice, France, 2002.
- [AD03] A. Ansar et K. Daniilidis. Linear pose estimation from points or lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5) :578–589, Mai 2003.
- [ADA095] L. Albrecht, C. Doignon, G. Abba, et E. Ostertag. Multi-controller design for robot control with visual feedback and switching strategy. Dans *Proceedings of the European Control Conference*, pages 3480–3485, Rome, Italie, Septembre 1995.
- [AEH00] N. Andreff, B. Espiau, et R. Horaud. Visual servoing from lines. Dans *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2070–2075, San Francisco, USA, 2000.
- [AGD07a] C. Albitar, P. Graebing, et C. Doignon. Conception d'un motif monochrome pour le codage robuste de la lumière structurée. Dans *11ème journées des Jeunes Chercheurs en Vision par Ordinateur (Orasis'07)*, Obernai, Bas-Rhin, Juin 2007.
- [AGD07b] C. Albitar, P. Graebing, et C. Doignon. Design of a monochromatic pattern for a robust structured light coding. Dans *Proceedings of the IEEE Int'l Conf. on Image Processing (à paraître)*, San Antonio, USA, Septembre 2007.
- [AGD07c] C. Albitar, P. Graebing, et C. Doignon. Robust structured light coding for 3d reconstruction. Dans *IEEE Int'l Conf. on Computer Vision*, Rio de Janeiro, Brésil (accepté pour présentation orale), Octobre 2007.
- [Agi77] G. J. Agin. Servoing with visual feedback. Dans *7th Symposium on Industrial Robots*, pages 551–560, Tokyo, Octobre 1977.
- [Agi79] G. J. Agin. Real time control of a robot with a mobile camera. Dans *Technical Note 179, SRI International*, Février 1979.
- [Alb05] C. Albitar. Asservissement visuel sans modèle, 2005.
- [Alo90] Y. Aloimonos. Perspective approximations. *Image and Vision Computing*, 8(3) :177–192, Août 1990.

- [Alt94] T.D. Alter. 3-d pose from 3 points using weak-perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8) :802–808, Août 1994.
- [And99] N. Andreff. *Asservissement visuel à partir de droites et auto-étalonnage pince-caméra*. Thèse de doctorat, INP Grenoble et INRIA, France, 1999.
- [AQT02] M-A. Ameller, L. Quan, et B. Triggs. Le calcul de pose : de nouvelles méthodes matricielles. Dans *Conference AFRIF/AFIA Reconnaissance des Formes et Intelligence Artificielle*, Angers, Janvier 2002.
- [ASZ⁺02] P. Abolmaesumi, S.E. Salcudean, W.H. Zhu, M. Sirouspour, et S. DiMaio. Image-guided control of a robot for medical ultrasound. *IEEE Transactions on Robotics and Automation*, 18(1) :11–23, Février 2002.
- [Auj04] J.F. Aujol. *Contribution à l'analyse de textures en traitement d'images par méthodes variationnelles et équations aux dérivées partielles*. Thèse de doctorat, Université de Nice-Sophia Antipolis, 2004.
- [AWJ90] A.A. Amini, T.E. Weymouth, et R.C. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(9), Septembre 1990.
- [BAK91] H. Batatia, A. Ayache, et C. Krey. Reconnaissance d'objets polyédriques à partir d'une seule image utilisant des invariants projectifs. Dans *Conférence AFCET Reconnaissance des Formes et Intelligence Artificielle*, tome 1, pages 323–328, Novembre 1991.
- [Bal81] D.H. Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2) :111–122, 1981.
- [BB82] D. H. Ballard et C. M. Brown. *Computer vision*. Prentice Hall, Englewood Cliffs, NJ, 1982.
- [BC82] R.C. Bolles et R.A. Cain. Recognizing and locating partially visible objects, the local-feature-focus method. *Int'l Journal of Robotics Research*, 1(3) :57–82, 1982.
- [BCD⁺04] D. Burschka, J. J. Corso, M. Dewan, G.D. Hager, W. Lau, M. Li, H. Lin, P. Marayong, et N. Ramey. Navigating inner space : 3-d assistance for minimally invasive surgery. Dans *Workshop Advances in Robot Vision, in conjunction with the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 45–59, Sendai, Japan, Septembre 2004.
- [BDZM07] F. Bourger, C. Doignon, P. Zanne, et M. de Mathelin. A model-free vision-based robot control for minimally invasive surgery using esm tracking and pixels color selection. Dans *Proceedings of the IEEE Int'l Conf. on Robotics and Automation*, pages 3579–3584, Rome, Italie, Avril 2007.
- [BFM88] J.D. Boissonnat, B. Faverjon, et J.P. Merlet. *Techniques de la robotique : Perception et planification*. Editions Hermès (Traité des Nouvelles Technologies), 1988.
- [BH86] R.C. Bolles et R.P. Horaud. 3dpo : A three-dimensional part orientation system. *Int'l Journal of Robotics Research*, 5(3) :3–26, 1986.
- [BI98] A. Blake et M. Isard. *Active contours*. Springer-Verlag, Londres, 1998.
- [BM04] S. Baker et I. Matthews. Lucas-kanade 20 years on : A unifying framework. *Int'l journal of Computer Vision*, pages 221–255, Mars 2004.

- [BM06] S. Benhimane et E. Malis. Une approche directe pour l'asservissement visuel 2d. Dans *15^e congrès francophone AFRIF-AFIA sur la Reconnaissance des Formes et l'Intelligence Artificielle (RFIA)*, Tours, France, Janvier 2006.
- [Bou] J.-Y. Bouguet. Camera calibration matlab toolbox. *MRL - Intel Corp.*
- [Bou89] P. Bouthemy. A maximum likelihood framework for determining moving edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(5) :499–511, 1989.
- [BRO80] R. A. Brown, T. S. Roberts, et A. G. Osborne. Stereotactic frame and computer software for CT-directed neurosurgical localization. *Invest. Radiol.*, 15 :308–312, 1980.
- [Bro92a] C. Brown. Numerical evaluation of differential and semi-differential invariants. *Geometric Invariance in Computer Vision*, pages 215–227, 1992.
- [Bro92b] L.G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4) :325–376, 1992.
- [BS01] A. Bartoli et P. Sturm. The 3d line motion matrix and alignment of line reconstructions. Dans *Proceedings of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, pages 287–292, Hawaii, USA, 2001.
- [Bur04] D. Burschka. Vision-based slam with standard off-the-shelf monocular cameras. Dans *Workshop Advances in Robot Vision, in conjunction with the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 67–78, Sendai, Japan, Septembre 2004.
- [BZWF02] M. Breeuwer, W. Zylka, J. Wadley, et A. Falk. Detection and correction of geometric distortion in 3D CT/MR images. Dans *Proc. of Computer Assisted Radiology and Surgery*, pages 11–23, Paris, 2002.
- [Can86] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6) :679–698, Novembre 1986.
- [CAPL95] A. Casals, J. Amat, D. Prats, et E. Laporte. Vision guided robotic system for laparoscopic surgery. Dans *Proc. of the IFAC Int. Congress on Advanced Robotics*, pages 33–36, Barcelona, Spain, 1995.
- [CBBB95] P. Cinquin, E. Bainville, C. Barbe, et E. Bittar. Computer assisted medical interventions - passive and semi-active aids. *IEEE Engineering in Medicine and Biology*, 14(3) :254–263, 1995.
- [CC00] C. Collewet et F. Chaumette. A contour approach for image-based control of objects with complex shape. Dans *Proceedings of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 751–756, Takamatsu, Japon, Novembre 2000.
- [CD86] R. T. Chin et C. R. Dyer. Model-based recognition in robot vision. *ACM Computing Surveys*, 18(1) :67–108, 1986.
- [CD95] R. Coifman et D. Donoho. Translation invariant denoising. Rapport technique 475, Dept. of Statistics, Stanford University, Mai 1995.
- [CDF96] G. Cortelazzo, C. Doignon, et R. Frezza. High resolution visualisation of paintings and frescoes. Dans *Int'l Congress on Environment/Climate - UNESCO*, Rome, Italie, Mars 1996.

- [CH01] P. Corke et S. Hutchnison. A new partitioned approach to image-based visual servo control. *IEEE Trans. on Robotics and Automation*, 17(4) :507–515, 2001.
- [Cha90] F. Chaumette. *La relation vision-commande : théorie et application à des tâches robotiques*. Thèse de doctorat, Université de Rennes 1, IRISA, Juillet 1990.
- [Cha98] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. the confluence of vision and control. *LNCIS series*, 237 :66–78, 1998.
- [Che91] H. Chen. Pose determination from line-to-plane correspondences : Existence of solutions and closed-form solutions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6) :530–541, Juin 1991.
- [Che95] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(8) :790–799, Août 1995.
- [Chr98] S. Christy. *Localisation et modélisation tridimensionnelle par approximation successives du modèle perspectif de caméra*. Thèse de doctorat, INP Grenoble, GRAVIR-IMAG, INRIA, 1998.
- [CLS95] C.-H. Chen, J.-S. Lee, et Y.-N. Sun. Wavelet transformation for gray-level corner detection. *Pattern Recognition*, 28(6) :853–861, 1995.
- [CLZ07] S.Y. Chen, Y.F. Li, et J. Zhang. Real time structured light vision with the principle of unique color codes. Dans *IEEE Int'l Conf. on Robotics and Automation*, pages 429–434, Rome, Italie, Avril 2007.
- [CLZQ03] M. Chan, W. Lin, C. Zhou, et J.Y. Qu. Miniaturized three-dimensional endoscopic imaging system based on active stereovision. *Applied Optics*, 42(10), Avril 2003.
- [CM99] E. Cervera et Ph. Martinet. Visual servoing with indirect image control and a predictable camera trajectory. Dans *Proceedings of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, tome 1, pages 381–386, Kyongju, Korea, Octobre 1999.
- [CMC00] G. Chesi, E. Malis, et R. Cipolla. Automatic segmentation and matching of planar contours for visual servoing. Dans *IEEE Int'l Conf. on Robotics and Automation*, tome 3, pages 2753–2758, San Francisco, USA, Avril 2000.
- [CMCK93] E. Coste-Manière, P. Couvignou, et P. Khosla. Robotic contour following based on visual servoing. Dans *In IEEE/RSJ, Int'l Conf. on Intelligent Robots and Systems*, Yokohama, Japon, juillet 1993.
- [CO93] D. Cyganski et J.A. Orr. Application of tensor theory to object recognition and orientation determination. *IEEE Trans. on PAMI*, 7 :662–673, Novembre 1993.
- [Cor96] P.I. Corke. *Visual control of robots : high performance visual servoing*. Research Studies Press (John Wiley & sons inc.), 1996.
- [CP02] C. Hueand J-P. Le Cadre et P. Pérez. Sequential monte carlo methods for multiple target tracking and data fusion. *IEEE Trans. on Signal Processing*, 50(2) :309–325, Février 2002.
- [CPK93] P. Couvignou, N. Papanikolopoulos, et P. Khosla. On the use of snakes for 3d robotic visual tracking. Dans *In Proceedings of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, pages 750–751, 1993.

- [CR80] P. Coiffet et P. Rives. Reconnaissance par un robot de l'orientation d'objets tridimensionnels en vue de tâches de saisie automatique. *R.A.I.R.O Automatique/Systems Analysis and Control*, 14(1), 1980.
- [CR99] T.-J. Cham et J. Rehg. A multiple hypothesis approach to figure tracking. Dans *IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, Juin 1999.
- [CRE93] F. Chaumette, P. Rives, et B. Espiau. Classification and realization of the different vision-based tasks. Dans *World Scientific Series in Robotics and Automated Systems, Vol. 7*, pages 199–228. Visual Servoing, K. Hashimoto (ed.), Singapore, 1993.
- [CRM03] D. Comaniciu, V. Ramesh, et P. Meer. Kernel-based object tracking. *Transactions on Pattern Analysis and Machine Intelligence*, 25(5), Mai 2003.
- [DA99] C. Doignon et G. Abba. A practical multi-plane method for a low-cost camera calibration technique. Dans *Proceedings of the European Control Conference*, Karlsruhe, Allemagne, Septembre 1999.
- [DAO94] C. Doignon, G. Abba, et E. Ostertag. Recognition and localization of solid objects by a monocular vision system for robotic tasks. Dans *Proceedings of IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 2007–2014, Munich, Germany, Septembre 1994.
- [DAWO95] C. Doignon, G. Abba, N. Werghi, et E. Ostertag. Localization of objects by a monocular vision system. Dans *Proceedings of the fourth Int'l Conf. on Intelligent Autonomous Systems*, pages 513–520, IAS'4, Karlsruhe, Allemagne, Mars 1995.
- [DC00] T. Drummond et R. Cipolla. Application of lie algebras to visual servoing. *Int'l Journal of Computer Vision*, 37(1) :21–41, June 2000.
- [DCF+94] M. Devy, R. Chatila, P. Fillatreau, S. Lacroix, et F. Nashashibi. On autonomous navigation in a natural environment. Dans *Symposium on Intelligent robotic systems*, pages 213–220, 1994.
- [DCF96] C. Doignon, G. Cortelazzo, et R. Frezza. A robust camera calibration technique for high resolution image acquisition. Dans *European Conference on Multimedia Applications, Services and Techniques*, Louvain, Belgique, Mai 1996.
- [DD92] D. DeMenthon et L.S. Davis. Exact and approximate solutions of the perspective-three-point problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11) :1100–1105, Novembre 1992.
- [DD95] D. Dementhon et L.-S. Davis. Model-based object pose in 25 lines of code. *IJCV*, 15(1) :123–141, 1995.
- [DDDS02] P. David, D. DeMenthon, R. Duraiswami, et H. Samet. Softposit : Simultaneous pose and correspondence determination. Dans *European Conference on Computer Vision*, Copenhagen, Mai 2002.
- [DDK05] C. Delimal, C. Doignon, et D. Knittel. Localisation de fréquences de vibration d'une bande en mouvement à l'aide de la vision à lumière structurée. Dans *10ème journées des jeunes chercheurs en Vision par Ordinateur, (Orasis'05)*, Fournols, Puy-de Dome, Mai 2005.

- [DDL94] N. Daucher, M. Dhome, et J.-T. Lapresté. Camera calibration from spheres images. Dans *Proceedings of the third European Conference on Computer Vision*, tome 1, pages 449–454, Stockholm, Sweden, Mai 1994.
- [DF90] R. Deriche et O.D. Faugeras. Tracking line segments. Dans *Proceedings of the European Conference on Computer Vision*, pages 259–268, Antibes, France, 1990.
- [DF01] F. Devernay et O. Faugeras. Straight lines have to be straight : Automatic calibration and removal of distortion from scenes of structured environments. *IAPR Int'l Journal of Machine Vision and Application*, 13(1) :14–24, 2001.
- [DGM05] C. Doignon, P. Graebing, et M. de Mathelin. Real-time segmentation of surgical instruments inside the abdominal cavity using a joint hue saturation color feature. *Real-Time Imaging*, 11 :429–442, Novembre 2005.
- [DH73] R.O. Duda et P.E. Hart. *Pattern Classification and Scene Analysis*. Wiley idnterscience publication, New-York, 1973.
- [DK01] C. Doignon et D. Knittel. Non-circularity detection of a rolling winder by artificial vision. Dans *Proceedings of the fifth Int'l Conf. on Quality-Control by Artificial Vision*, Le Creusot, France, Mai 2001.
- [DK02] C. Doignon et D. Knittel. Detection of non-circularity and eccentricity of a rolling winder by artificial vision. *Eurasip Journal of Applied Signal Processing*, 2002(7) :714–727, Juillet 2002.
- [DK03] C. Doignon et D. Knittel. Detection and characterization of web vibrations by artificial vision. Dans *Proceedings of the seventh Int'l Conf. on Web Handling*, Stillwater, Oklahoma, USA, Juin 2003.
- [DK05] C. Doignon et D. Knittel. A structured light vision system for out-of-plane vibration frequencies location of a moving web. *IAPR Int'l Journal of Machine Vision and Applications*, 16(5) :289–297, Décembre 2005.
- [DLRR90] M. Dhome, J.-T. Lapreste, G. Rives, et M. Richetin. Spatial localization of modelled objects of revolution in monocular perspective vision. Dans *Proceedings of the European Conference on Computer Vision*, Antibes, France, 1990.
- [DM07] C. Doignon et M. de Mathelin. A degenerate conic-based method for a direct fitting and 3-d pose of cylinders with a single perspective view. Dans *Proceedings of the IEEE Int'l. Conf. on Robotics and Automation*, pages 4220–4225, Rome, Italie, Avril 2007.
- [DMBM06] C. Doignon, B. Maurin, B. Bayle, et M. de Mathelin. Vers un asservissement visuel en tomodensitométrie à rayons x. Dans *15^e congrès francophone AFRIF-AFIA RFA*, Tours, France, Janvier 2006.
- [DN95] J. Denzler et H. Niemann. Evaluating the performance of active contours models for real-time object tracking. Dans *Asian Conference on Computer Vision*, tome 2, pages 341–345, Singapour, 1995.
- [DN96] K. Deguchi et T. Noguchi. Visual servoing using eigenspace methods and dynamic calculation of interaction matrices. Dans *Proceedings of the Int'l Conf. on Pattern Recognition*, pages 302–306, 1996.

- [DNM04] C. Doignon, F. Nageotte, et M. de Mathelin. Detection of grey regions in color images : application to the segmentation of a surgical instrument in robotized laparoscopy. Dans *IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 3394–3399, Sendai, Japon, Octobre 2004.
- [DNM06a] C. Doignon, F. Nageotte, et M. de Mathelin. The role of insertion points in the detection and positioning of instruments in laparoscopy for robotic tasks. Dans *In Proceedings of the Int'l. Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pages 527–534 (Part I), Copenhagen, Danemark, Octobre 2006.
- [DNM06b] C. Doignon, F. Nageotte, et M. de Mathelin. Segmentation and guidance of multiple rigid objects for intraoperative endoscopic vision. Dans *Proceedings of the 2nd Int'l Workshop on Dynamical Vision, in conjunction with ECCV'2006*, Prague, République Tchèque, Mai 2006.
- [DNM07] C. Doignon, F. Nageotte, et M. de Mathelin. Segmentation and guidance of multiple rigid objects for intra-operative endoscopic vision. *Dynamical Vision*, LNCS 4358 :314–327, 2007.
- [Doi94] C. Doignon. *Etude et Réalisation d'un système de vision monoculaire pour la reconnaissance et la localisation d'objets en robotique*. Thèse de doctorat, Université Louis Pasteur (LSIIT), Janvier 1994.
- [Doi07] C. Doignon. An introduction to model-based pose estimation and 3d tracking techniques. *Scene Reconstruction, Pose Estimation and Tracking*, 2007.
- [DRLR89] M. Dhome, M. Richetin, J.-T. Lapreste, et G. Rives. Determination of the attitude of 3d objects from a single perspective view. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(12) :1265–1278, 1989.
- [DSG90] A. Djouadi, O. Snorrason, et F. Garber. The quality of training-sample estimates of the bhattacharyya coefficient. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12 :92–97, 1990.
- [DY04] D. Stoyanov and A. Darzi et G. Z. Yang. Dense 3d depth recovery for soft tissue deformation during robotically assisted laparoscopic surgery. Dans *In Proceedings of the Medical Image Computing and Computer-Assisted Intervention*, pages 41–48, St Malo, France, Septembre 2004.
- [ECR92] B. Espiau, F. Chaumette, et P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. Robotics and Automation*, 8(3) :313–326, Juin 1992.
- [Eji72] M. Ejiri. A prototype intelligent robot that assembles objects from plan drawings. *IEEE Transactions on Computers*, C-21 :161–170, Février 1972.
- [EMR02] F.-X. Espiau, E. Malis, et P. Rives. Robust features tracking for robotic applications : towards 2 1/2 d visual servoing with natural images. Dans *IEEE Int'l Conf. on Robotics and Automation*, tome 1, pages 574–579, Washington DC, USA, Mai 2002.
- [FB81] M. A. Fischler et R. C. Bolles. Random sample consensus : A paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM*, 24(6) :381–395, 1981.

- [FCH06] A. Foulonneau, P. Charbonnier, et F. Heitz. Affine-invariant geometric shape priors for region-based active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8) :1352–1357, 2006.
- [FDP⁺02] G. Fichtinger, T. L. DeWeese, A. Patriciu, A. Tanacs, D. Mazilu, J. H. Anderson, K. Masamune, R. H. Taylor, et D. Stoianovici. Robotically assisted prostate biopsy and therapy with intra-operative ct guidance. *Journal of Academic Radiology*, 9 :60–74, 2002.
- [FGL87] K. S. Fu, R. C. Gonzalez, et G. S. G. Lee. *Robotics : Control, Sensing, Vision, and Intelligence*. Mc Graw-Hill International Editions, 1987.
- [FH86] O.D. Faugeras et M. Hebert. The representation, recognition, and locating of 3d shapes from range data. *Int'l Journal of Robotics Research*, 5(3) :27–52, 1986.
- [FLM91] J. Feddema, C. S. G. Lee, et O. R. Mitchell. Weighted selection of image features for resolved rate visual feedback control. *IEEE Transactions on Robotics and Automation*, RA-7(1) :31–47, Février 1991.
- [FM99] J.T. Feddema et O.R. Mitchell. Vision-guided servoing with feature-based trajectory generation. *IEEE Trans. Robotics and Automation*, 5(5) :691–700, 99.
- [FMV93] M. Ferri, F. Mangili, et G. Viano. Projective pose estimation of linear and quadratic primitives in monocular computer vision. *Computer Vision, Graphics and Image Processing : Image understanding*, 58(1) :66–84, Juillet 1993.
- [FMZ⁺91] D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, et C. Rothwell. Invariant descriptors for 3-d object recognition and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10) :971–991, Octobre 1991.
- [FPB⁺71] J. A. Feldman, K. K. Pingle, T. O. Binford, G. Falk, A. Hay, R. Pau, R. F. Sproull, et J. M. Tenenbaum. The use of vision and manipulation to solve the puzzle. Dans *Proceedings of the Second Int'l Joint Conf. on Artificial Intelligence*, pages 359–364, London, Septembre 1971.
- [FRD00] J. Fayolle, L. Riou, et C. Ducottet. Robustness of a multiscale scheme of feature points detection. *Pattern Recognition*, 33 :1437–1453, 2000.
- [FS88] T.L. Faber et E.M. Stokely. Orientation of 3-d structures in medical images. *IEEE Trans. on PAMI*, 10(5) :626–633, Septembre 1988.
- [FS93] J. Flusser et T. Suk. Pattern recognition by affine moment invariants. *Pattern Recognition*, 16(1) :167–174, 1993.
- [FT87] O.D. Faugeras et G. Toscani. Camera calibration for 3d computer vision. Dans *Int'l Workshop on Machine Vision and Machine Intelligence*, Tokyo, 1987.
- [FTE⁺96] J. Funda, R.H. Taylor, B. Eldridge, S. Gomory, et K.G. Gruben. Constrained cartesian motion control for teleoperated surgical robots. *IEEE Trans. on Robotics and Automation*, 12(3), June 1996.
- [FTG01] V. Ferrari, T. Tuytelaars, et L. Vaan Gool. Real-time affine region tracking and coplanar grouping. Dans *Proceedings of the Int'l Conf. on Computer Vision and Pattern Recognition*, Hawaii, USA, 2001.
- [Gan84] S. Ganapathy. Decomposition of transformation matrices for robot vision. *Pattern Recognition Letter*, 2 :401–412, 1984.

- [Gan99] J. Gangloff. *Asservissements visuels rapides d'un robot manipulateur à 6 degrés de liberté*. Thèse de doctorat, Université Louis Pasteur, Janvier 1999.
- [Gan04] J. Gangloff. *Asservissements visuels et chirurgie robotisée*. Mémoire d'Habilitation à diriger des recherches, Université Louis Pasteur, Décembre 2004.
- [GB92] L.J. Van Gool et M.H. Brill. Semi-differential invariants for nonplanar curves. *Geometric Invariance in Computer Vision*, pages 293–309, 1992.
- [GCK⁺03] P. Grunert, K. Charalampi, M. Kassen, H. Boecher-Schwarz, R. Filippi, et P. Grunert Jr. Frame-based and frameless stereotaxy in the localization of cavernous angiomas. *NeuroSurgery Review*, 26(1) :53–61, Janvier 2003.
- [GD94] A. Gangi et J.-L. Dietermann. *Tomodensitométrie interventionnelle*. Editions Vigot, Paris, 1994.
- [GH91] E.L. Grimson et D.P. Huttenlocher. On the verification of hypothesized matches in model-based recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(12) :1201–1213, Décembre 1991.
- [GH95] C.A. Glasbey et G.W. Horgan. *Image Analysis for the Biological Sciences*. Vic Barnett editor - Wiley & Sons publishing, 1995. ISBN 0-471-93726-6.
- [GM02] J. Gangloff et M. de Mathelin. Visual servoing of a 6 dof manipulator for unknown 3d profile following. *IEEE Trans. on Robotics and Automation*, 18(4) :511–520, 2002.
- [GNY92] P. Griffin, L. Narasimhan, et S. Yee. Generation of uniquely encoded light patterns for range data acquisition. *Pattern Recognition*, 25(6) :609–616, 1992.
- [Gri90] E. Grimson. *Object Recognition by Computer : The Role of Geometric Constraint*. MIT Press, 1990.
- [GRL⁺98] S. Gold, A. Rangarajan, C.P. Lu, S. Pappu, et E. Mjolsness. New algorithms for 2d and 3d point matching : Pose estimation and correspondence. *Pattern Recognition*, 31 :1019–1031, 1998.
- [Gro98] P. Gros. *De l'appariement à l'indexation des images*. Thèse d'Habilitation à Diriger des Recherches de l'INP Grenoble, Décembre 1998.
- [GS99] V. Govindu et C. Shekhar. Alignment using distributions of local geometric properties. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(10) :1031–1043, Octobre 1999.
- [GSS93] N.J. Gordon, D.J. Salmond, et A. F. M. Smith. Novel approach to nonlinear / non-gaussian bayesian state estimation. Dans *IEE Proceedings, Part F, Radar, Sonar and Navigation*, 140, tome 2, pages 107–113, Avril 1993.
- [Har89] R.-M. Haralick. Monocular vision using inverse perspective projection geometry : Analytic relations. Dans *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 370–378, 1989.
- [Has04] K. Hashimoto. A hierarchical control architecture for high-speed visual servoing. Dans *Workshop Advances in Robot Vision, IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 32–44, Sendai, Japan, Septembre 2004.
- [HB98] G.D. Hager et Peter N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(10) :1025–1039, 1998.

- [HCLL89] R. Horaud, B. Conio, O. Leboulleux, et B. Lacolle. An analytic solution for the perspective 4-point problem. *Computer Vision, Graphics, and Image Processing*, 47 :33–44, 1989.
- [HCWS84] R.M. Haralick, Y.H. Chu, L.T. Watson, et L.G. Shapiro. Matching wire frame objects from their 2d perspective projection. *Pattern Recognition*, 17(6) :607–619, 1984.
- [HDH⁺02] J.S. Hong, T. Dohi, M. Hasizume, K. Konishi, et N. Hata. A motion adaptable needle placement instrument based on tumor specific ultrasonic image segmentation. Dans *Int'l Conf. on Medical Image Computing and Computer-Assisted Interventions*, Tokyo, Japon, Septembre 2002.
- [HH89] C. Hansen et T.C. Henderson. Cagd-based computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11) :1181–1193, Novembre 1989.
- [HHC96] S. Hutchinson, G. D. Hager, et P.I. Corke. A tutorial on visual servo control. *IEEE Trans. on Robotics and Automation*, 12(5) :651–670, Octobre 1996.
- [HLON91] R.M. Haralick, C. Lee, K. Ottenberg, et M. Nölle. Analysis and solutions of the three-point perspective pose estimation problem. Dans *IEEE Conf. Computer Vision and Pattern Recognition*, pages 592–598, Maui, Hawaii, USA, 1991.
- [HN01] M. Hayashibe et Y. Nakamura. Laser-pointing endoscope system for intraoperative geometric registration. Dans *Proceedings of the IEEE Int. Conf. on Robotics and Automation*, tome 2, pages 1543–1548, Seoul, Corée, Mai 2001.
- [Hor86] B. K. P. Horn. *Robot Vision*. MIT Press, McGraw-Hill, Mars 1986.
- [Hor87] R. Horaud. New methods for matching 3-d objects with single perspective views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(3) :401–412, Mai 1987.
- [Hou62] P.V.C. Hough. Method and means for recognizing complex patterns. *Brevet US 3 069 654*, 1962.
- [HP79] J. Hill et W. T. Park. Real time control of a robot with a mobile camera. Dans *Proceedings of the 9th ISIR*, pages 233–246, Washington D.C., Mars 1979.
- [HR93] D.P. Huttenlocher et W.J. Rucklidge. A multi-resolution technique for comparing images using the hausdorff distance. Dans *Proceedings of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, pages 705–726, 1993.
- [HS88] C.G. Harris et M. Stephens. A combined corner and edge detector. Dans *In 4th Alvey Vision Conference*, pages 147–151, 1988.
- [HS92] R. M. Haralick et L. G. Shapiro. *Computer and Robot Vision*, tome 1. Addison-Wesley Publishing, 1992.
- [HS02] A. Hanbury et J. Serra. A 3-d polar coordinate colour representation suitable for image analysis. Dans *Pattern Recognition and Image Processing Group*, tome Technical Report PRIP-TR-077, Université de Vienne, Autriche, December 2002.
- [Hsi03] J. Hsieh. *Computed Tomography : Principles, Design, Artifacts, and Recent Advances*. SPIE Press, 2003.
- [HSN06] M. Hayashibe, N. Suzuki, et Y. Nakamura. Laser-scan endoscope system for intraoperative geometry acquisition and surgical robot safety management. *Medical Image Analysis*, 10 :509–519, 2006.

- [Hu62] M. K. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8 :179–187, 1962.
- [HU90] D.P. Huttenlocher et Shimon Ullman. Recognizing solid objects by alignment with an image. *Int'l. Journal of Computer Vision*, 5(2) :195–212, 1990.
- [HYM95] H. Haneishi, Y. Yagihashi, et Y. Miyake. A new method for distortion correction of electronic endoscope images. *IEEE Trans. on Medical Imaging*, 14(3) :548–555, Septembre 1995.
- [HZ00] R. Hartley et A. Zisserman. *Multiple view geometry in computer vision*. Cambridge Univ. Press, 2000.
- [IB98] M.A. Isard et A. Blake. Condensation-conditional density propagation for visual tracking. *Int'l Journal of Computer Vision*, 29(1) :5–28, 1998.
- [IK88] Illingworth et Kittler. A survey of the hough transform. *Computer Vision, Graphics, and Image Processing*, 44 :87–116, 1988.
- [Ike87] K. Ikeuchi. Generating an interpretation tree from a cad model for 3-d object recognition in bin-picking tasks. *Int'l Journal of Computer Vision*, 1(2) :145–165, 1987.
- [IPV99] N. Ikonomakis, K. Plataniotis, et A. N. Venetsanopoulos. A region-based color image segmentation scheme. Dans *in Proceedings of Electrical Imaging*, tome 3653 of SPIE (San Jose, California), pages 1202–1209, January 1999.
- [JD02] F. Jurie et M. Dhome. Hyperplane approximation for template matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7) :996–1000, 2002.
- [JFEM03] A. Jepson, D. J. Fleet, et T. El-Maraghi. a robust on-line appearance models for vision tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10) :1296–1311, 2003.
- [JFW⁺97] C. R. Maurer Jr, J. M. Fitzpatrick, M. Y. Wang, R. J. Maciunas R. L. Galloway Jr, et G. S. Allen. Registration of head volume images using implantable fiducial markers. *IEEE Transactions on Medical Imaging*, 16(8) :447–462, Août 1997.
- [JH03] T. Jost et H. Hügli. A multi-resolution icp with heuristic closest point search for fast and robust 3d registration of range images. Dans *Proceedings of the Int'l. Conf. on 3-D Digital Imaging and Modeling*, Banff, Alberta, Canada, 2003.
- [JM92] J.M. Jolion et A. Montanvert. The adaptive pyramid : a framework for 2d image analysis. *CVGIP : Image Understanding*, 55(3) :339–348, 1992.
- [JMN77] R. Jain, D. Militzer, et H.H. Nagel. Separating non-stationary from stationary scene components in a sequence of real world tv-images. Dans *Proceedings of IJCAI*, 1977.
- [JU97] S.J. Julier et J.K. Uhlmann. A new extension of the kalman filter to nonlinear systems. Dans *Proceedings of the SPIE Aerosense Symposium*, Orlando, USA, Avril 1997.
- [JUDW95] S.J. Julier, J.K. Uhlmann, et H. Durrant-Whyte. A new approach for filtering nonlinear systems. Dans *Proceedings of the American Control Conference*, pages 1628–1632, 1995.

- [Jur99] F. Jurie. Solution of the simultaneous pose and correspondence problem using gaussian error model. *Computer Vision and Image Understanding*, 73(3) :357–373, Mars 1999.
- [Kal60] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82 :35–45, 1960.
- [KC03] D. Kragic et H.I. Christensen. A framework for visual servoing. Dans *Proceedings of the Int'l Conf. on Computer Vision Systems*, pages 345–354. Springer-Verlag, 2003.
- [KDGM02] A. Krupa, C. Doignon, J. Gangloff, et M. de Mathelin. Combined image-based and depth visual servoing applied to robotized laparoscopic surgery. Dans *Proceedings of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, tome 1, pages 323–329, Lauzanne, Suisse, Octobre 2002.
- [KG82] F. Kuhl et C. Giordani. Elliptic fourier features of a closed contour. *Computer Vision, Graphics and Image Processing*, 18 :236–258, 1982.
- [KGD⁺03] A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel, L. Soler, J. Leroy, et J. Marescaux. Autonomous retrieval and 3d positioning of surgical instruments in robotized laparoscopic surgery. *IEEE Trans. on Robotics and Automation*, 19(5) :842–853, Octobre 2003.
- [KH98] C. Kervrann et F. Heitz. A hierarchical markov modeling approach for the segmentation and tracking of deformable shapes. *Graphical Models and Image Processing*, 60(3) :173–195, Mai 1998.
- [Kha86] O. Khatib. Real time obstacle avoidance for manipulators and mobile robots. *Int'l. Journal of Robotics Research*, 5(1) :90–98, 1986.
- [KHJH88] Y. S. Kwok, J. Hou, E.A. Jonckheere, et S. Hayati. A robot with improved absolute positioning accuracy for CT guided stereotactic brain surgery. *IEEE Transactions on Biomedical Engineering*, 35(2) :153–160, Février 1988.
- [KJM⁺02] A. Krupa, J. Gangloff, M. de Mathelin, C. Doignon, G. Morel, L. Soler, J. Leroy, et J. Marescaux. Autonomous retrieval positioning of surgical tools in robotized laparoscopic surgery using visual servoing and lasers pointers. Dans *Proceedings of the Int'l Conf. on Robotics and Automation*, pages 3769–3774, Washington D.C. USA, Mai 2002.
- [KMD⁺01] A. Krupa, M. de Mathelin, C. Doignon, J. Gangloff, G. Morel, L. Soler, et J. Marescaux. Development of semi-autonomous control modes in laparoscopic surgery using visual servoing. Dans *Proceedings of the fourth Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*, tome LNCS 2208, pages 1306–1307, Utrecht, Pays-Bas, Octobre 2001.
- [KMD⁺02] A. Krupa, M. de Mathelin, C. Doignon, J. Gangloff, G. Morel, L. Soler, J. Leroy, et J. Marescaux. Automatic 3-d positioning of surgical instruments during robotized laparoscopic surgery using automatic visual feedback. Dans *Proceedings of the fourth Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*, tome LNCS 2488, pages 9–16, Tokyo, Japon, Septembre 2002.
- [Kru03] A. Krupa. *Commande par vision d'un robot de chirurgie laparoscopique*. Thèse de doctorat, INP Lorraine, Juillet 2003.

- [KWT87] M. Kass, A. Witkin, et D. Terzopoulos. Snakes : Active contours models. Dans *Proceedings of the Int'l Conf. on Computer Vision*, pages 259–269, Londres, 1987.
- [LB00] A. Leonardis et H. Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78(1) :99–118, 2000.
- [LD89] C.H. Lo et H.S. Don. 3-d moment forms : Their construction and application to object identification and positioning. *IEEE Trans. on PAMI*, 113(10) :1053–1064, Octobre 1989.
- [Lee83] J.S. Lee. Digital image smoothing and the sigma filter. *Computer Vision, Graphics, and Image Processing*, 24 :255–269, 1983.
- [Lef02] S. Lefevre. *Détection d'évènements dans une séquence vidéo*. Thèse de doctorat, Université François-Rabelais de Tours, Décembre 2002.
- [LF05] V. Lepetit et P. Fua. Monocular model-based 3d tracking of rigid objects : A survey. *Foundations and Trends in Computer Graphics and Vision*, 1(1) :1–89, Octobre 2005.
- [LFC02] S. Lee, G. Fichtinger, et G. S. Chirikjian. Numerical algorithms for spatial registration of line fiducials from cross-sectional images. *American Association of Physicists in Medicine*, 29(8) :1881–1891, Août 2002.
- [LH81] C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293 :61–62, Septembre 1981.
- [LHD88] S. Linnainmaa, D. Harwood, et L. S. Davis. Pose determination of a three-dimensional object using triangle pairs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5) :634–647, 1988.
- [LHF90] Y. Liu, T.S. Huang, et O.D. Faugeras. Determination of camera location from 2d to 3d line and point. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 12(1) :28–37, Janvier 1990.
- [LJ94] L. Lemieux et R. Jagoe. Effect of fiducial marker localization on stereotactic target coordinate calculation in CT slices and radiographs. *Physics in medicine and biology*, 39(11) :1915–1928, 1994.
- [LK81] B. Lucas et T. Kanade. An iterative image registration technique with an application to stereo vision. Dans *Int'l Joint Conference on Artificial Intelligence*, pages 674–679, 1981.
- [LLL05] A. Lehuger, P. Lechat, N. Laurent, et P. Pérez. Suivi de joueurs dans les séquences sportives à fort changements d'illumination. Dans *Actes de CORESA*, Novembre 2005.
- [Low80] D.G. Lowe. Solving for the parameters of object models from image descriptions. Dans *Proceedings of ARPA Image Understanding Workshop*, pages 121–127, College Park, MD, Avril 1980.
- [Low85] D.G. Lowe. *Perceptual Organization and Visual Recognition*. Boston, Mass : Kluwer Academic Publishers, 1985.
- [Low87] D.G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31(3) :355–395, Mars 1987.

- [Low04] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *Int'l Journal of Computer Vision*, 60(2) :91–110, 2004.
- [LSM⁺05] R. Lukac, B. Smolka, K. Martin, K.N. Plataniotis, et A.N. Venetsanopoulos. Vector filtering for color imaging. *IEEE Signal Processing Magazine, Special Issue on Color Image Processing*, 22(1) :74–86, 2005.
- [Luc01] L. Lucchese. Estimating affine transformations in the frequency domain. Dans *IEEE Int'l Conf. on Image Processing*, tome 2, pages 909–912, Thessalonique, Grèce, Octobre 2001.
- [LW88] Y. Lamdan et H.J. Wolfson. Geometric hashing : A general and efficient model-based recognition scheme. Dans *IEEE Int'l Conf. on Computer Vision*, pages 238–249, Tampa, FL, USA, Décembre 1988.
- [Mal89] S. G. Mallat. A theory for multiresolution signal decomposition. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 11(7) :674–693, Juillet 1989.
- [Mal98] E. Malis. *Contributions à la modélisation et à la commande en asservissement visuel*. Thèse de doctorat, Université de Rennes I, Novembre 1998.
- [Man06] N. Mansard. *Enchaînement de tâches robotiques*. Thèse de doctorat, Université de Rennes I, Décembre 2006.
- [Mar82] D. C. Marr. *Vision*. W. H. Freeman Editor, San Francisco, 1982.
- [Mar04] E. Marchand. *Contributions à la commande d'une caméra réelle ou virtuelle dans des mondes réels ou virtuels*. Mémoire d'Habilitation à diriger des recherches, Université de Rennes I, Novembre 2004.
- [Mau05] B. Maurin. *Conception et réalisation d'un robot d'insertion d'aiguille pour les procédures percutanées sous imageur scanner*. Thèse de doctorat, Université Louis Pasteur (LSIIT), Novembre 2005.
- [May92a] S.-J. Maybank. The cross-ratio and the j-invariant. *Geometric Invariance in Computer Vision*, pages 107–109, 1992.
- [May92b] S.-J. Maybank. The projection of two non-coplanar conics. *Geometric Invariance in Computer Vision*, pages 105–119, 1992.
- [MB94] F. Meyer et P. Bouthemy. Region-based tracking using affine motion models in long image sequences. *Computer Vision, Graphics and Image Processing : Image Understanding*, 60(2) :119–140, Septembre 1994.
- [MB04] E. Malis et S. Benhimane. Real-time image-based tracking of planes using efficient second-order minimization. Dans *IEEE/RSJ IROS*, Sendai, Japan, September 2004.
- [MC00] E. Malis et F. Chaumette. 2 1/2 d visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. *International Journal of Computer Vision*, 37(1) :79–97, 2000.
- [MC02a] E. Marchand et F. Chaumette. Virtual visual servoing : a framework for real-time augmented reality. Dans *Proceedings of the EUROGRAPHICS Conference*, tome 21 (3 of Computer Graphics Forum), pages 289–298, Saarbrück, Allemagne, Septembre 2002.

- [MC02b] Y. Mezouar et F. Chaumette. Path planning for robust image-based control. *IEEE Trans. on Robotics and Automation*, 18(4) :534–549, Août 2002.
- [MC04] E. Marchand et F. Chaumette. Features tracking for visual servoing purposes. Dans *Workshop Advances in Robot Vision, in conjunction with the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 10–20, Sendai, Japan, September 2004.
- [MDG⁺05] B. Maurin, C. Doignon, J. Gangloff, B. Bayle, M. de Mathelin, O. Piccin, et A. Gangi. Needle tip positioning on the abdomen with a novel stereotaxic robotic assistant. Dans *Proceedings of Surgetica'2005 : Computer-Aided Medical Interventions : tools and applications*, pages 157–164, Chambéry, France, Janvier 2005.
- [MDGD97] P. Martinet, N. Daucher, J. Gallice, et M. Dhome. Robot control using monocular pose estimation. Dans *Workshop on New Trends in Image-based Robot Servoing en conjonction avec IROS'97*, Grenoble, 1997.
- [MDK07] X. Maurice, C. Doignon, et D. Knittel. On-line vibrations characterization of a moving flexible web with a fast structured lighting stereovision system. Dans *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, Varsovie, Pologne, Mai 1-3 2007.
- [MDMG03] B. Maurin, C. Doignon, M. de Mathelin, et A. Gangi. Pose reconstruction from an uncalibrated computed tomography imaging device. Dans *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Madison, WN, USA, Juin 2003.
- [MDMG05] B. Maurin, C. Doignon, M. de Mathelin, et A. Gangi. Algorithmes d'estimation automatique de la pose pour la radiologie interventionnelle assistée par ordinateur. Dans *Actes des 10ème journées des Jeunes Chercheurs en Vision par Ordinateur (Orasis)*, Fournols, Puy-de Dome, Mai 2005.
- [Mey92] Y. Meyer. *Les Ondelettes : Algorithmes et Applications*. Editions Armand Colin, collection Acquis Avancés de l'Informatique, 1992.
- [MFP⁺01] K. Masamune, G. Fichtinger, A. Patriciu, R. C. Susil, R. H. Taylor, L. R. Kavoussi, J. H. Anderson, I. Sakuma, T. Dohi, et D. Stoianovici. System for robotically assisted percutaneous procedures with computed tomography guidance. *Computer-Aided Surgery*, 6 :370–383, 2001.
- [MGB⁺04] B. Maurin, J. Gangloff, B. Bayle, M. de Mathelin, O. Piccin, P. Zanne, C. Doignon, L. Soler, et A. Gangi. A parallel robotic system with force sensors for percutaneous procedures under ct-guidance. Dans *Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*, St Malo, France, Septembre 2004.
- [MGK96] P. Martinet, J. Gallice, et D. Khadraoui. Vision based control. Dans *Proceedings of the World Automation Congress*, Montpellier, Mai 1996.
- [MGM02] K. Miura, J. Gangloff, et M. de Mathelin. Robust and uncalibrated visual servoing without jacobian using a simplex method. Dans *Proceedings of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 311–316, 2002.
- [Mit94] A. Mitiche. *Computational Analysis of Visual Motion*. Plenum Press, New York, 1994.

- [MKM⁺95] K. Masamune, E. Kobayashiand, Y. Masutaniand, M. Suzuki, T. Dohi andH. Iseki, et K. Takakura. Development of a mri-compatible needle insertion manipulator for stereotactic neurosurgery. *Journal of Image Guided Surgery*, 1 :242–248, 1995.
- [MM86] F. Mokhtarian et A. Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1) :34–43, 1986.
- [MOC⁺98] A. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, et J. Nissanov. Structured light using pseudo-random codes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3) :322–327, 1998.
- [Mor93] L. Morin. *Quelques contributions des invariants projectifs à la vision par ordinateur*. Thèse de doctorat, Institut National Polytechnique de Grenoble (LIFIA), Janvier 1993.
- [MP76] D. C. Marr et T. Poggio. Cooperative computation of stereo disparity. *Science, New serie*, 194 :283–287, 1976.
- [MR96] R. Mukundan et K.R. Ramakrishnan. An iterative solution for object pose parameters using image moments. *Pattern Recognition Letters*, 7 :1279–1284, 1996.
- [MSKS03] Y. Ma, S. Soatto, J. Košeká, et S. Sastry. *Invitation to 3D Vision : From Images to Geometric Models*. Springer-Verlag, 1003.
- [MT00] J. MacLean et J. Tsotsos. Fast pattern recognition using gradient-descent search in an image pyramid. Dans *Proceedings of the Int'l Conf. on Pattern Recognition*, tome 2, pages 2873–2877, Barcelone, Espagne, 2000.
- [Muk92] R. Mukundan. Estimation of quaternion parameters from two dimensional image moment. *Graphical Model and Image Processing*, 54(4) :345–350, Juillet 1992.
- [MYS88] H. Morita, K. Yajima, et S. Sakata. Reconstruction of surfaces of 3-d objects by m-array pattern projection method. Dans *IEEE Int'l Conf. on Computer Vision*, pages 468–473, 1988.
- [Nag05] F. Nageotte. *Contributions à la Suture Assistée par Ordinateur en Chirurgie Mini-Invasive*. Thèse de doctorat, Université Louis Pasteur (LSIIT), Novembre 2005.
- [NF96] N. Navab et O.D. Faugeras. Monocular pose determination from lines : Critical sets and maximum number of solutions. Dans *Int'l Conf. on Computer Vision and Pattern Recognition*, pages 254–260, Juin 1996.
- [NHK⁺03] A. Nishikawa, T. Hosoi, K. Koara, D. Negoro, A. Hikita, S. Asano, H. Kakutani, F. Miyazaki, M. Sekimoto, M. Yasui, Y. Miyake, S. Takiguchi, et M. Monden. Face mouse : A novel human-machine interface for controlling the position of a laparoscope. *IEEE Trans on Robotics and Automation*, 19(5) :825–841, 2003.
- [Nil69] N. J. Nilsson. A mobile automaton : An application of artificial intelligence techniques. Dans *Proceedings of the first Int' Joint Conf. on Artificial Intelligence*, pages 509–520, Washington D.C., 1969.
- [NIN⁺06] A. Nishikawa, K. Ito, H. Nakagoe, K. Taniguchi, M. Sekimoto, S. Takiguchi, Y. Seki, M. Yasui, K. Okada, M. Monden, et F. Miyazaki. Automatic positioning of a laparoscope by preoperative workspace planning and intraoperative 3d instrument tracking. Dans *MICCAI Workshop on Medical Robotics : Systems and*

- Technology towards Open Architecture*, pages 82–91, Copenhagen, Danemark, Octobre 2006.
- [NMD⁺04] F. Nageotte, M. de Mathelin, C. Doignon, L. Soler, J. Ieroy, et J. Marescaux. Computer-aided suturing in laparoscopic surgery. Dans *Medical Robotics, Navigation and Visualization*, RheinAhrCampus Remagen, Germany, Mars 2004.
- [NN96] S.A. Nene et S.K. Nayar. Closest point search in high dimensions. Dans *Proceedings of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, pages 859–865, San Francisco, USA, 1996.
- [NNM96] S.K. Nayar, S.A. Nene, et H. Murase. Subspace methods for robot vision. *IEEE Trans. on Robotics and Automation*, 12(5) :750–758, Octobre 1996.
- [NZDM06] F. Nageotte, P. Zanne, C. Doignon, et M. de Mathelin. Visual servoing-based endoscopic path following for robot-assisted laparoscopic surgery. Dans *IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 2364–2369, Pékin, Chine, Octobre 2006.
- [NZMD05] F. Nageotte, P. Zanne, M. de Mathelin, et C. Doignon. A circular needle path planning method for suturing in laparoscopic surgery. Dans *Proceedings of the IEEE Int'l Conf. on Robotics and Automation*, pages 516–521, Barcelone, Espagne, Avril 2005.
- [OD97] T. Okatani et K. Deguchi. Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center. *Computer Vision and Image Understanding*, 66(2) :119–131, Mai 1997.
- [Oli94] J.C. Olivo. Automatic threshold selection using the wavelet transform. *CVGIP : Graphical Models and Image Processing*, 56(3) :205–218, Mai 1994.
- [Par00] N.K. Paragios. *Geodesic Active Regions and Level Set methods : Contributions and Applications in Artificial Vision*. Thèse de doctorat, Université de Nice Sophia-Antipolis, Janvier 2000.
- [Pav77] T. Pavlidis. Polygonal approximations by newton's method. *IEEE Transactions on Computer*, C-26(8) :800–807, 1977.
- [PBT⁺92] E.M. Petriu, T. Bieseman, N. Trif, W. McMath, et S. Yeung. Visual object recognition using pseudo-random grid encoding. Dans *Proceedings of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 617–624, Juillet 1992.
- [PCCS06] J. Pagès, Ch. Collewet, F. Chaumette, et J. Salvi. An approach to visual servoing based on coded light. Dans *IEEE Int'l Conf. on Robotics and Automation*, pages 4118–4123, 2006.
- [PD94] A. Pikaz et I. Dinstein. Using simple decomposition for smoothing and feature point detection of noisy digital curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8) :808–813, Août 1994.
- [PF77] E. Persoon et K. Fu. Shape discrimination using fourier descriptors. *IEEE Trans. Systems, Man and Cybernetics*, 7(3) :170–179, 1977.
- [PFR96] M. Pilu, A. Fitzgibbon, et R. Fisher. Ellipse-specific direct least-square fitting. Dans *IEEE Int'l Conf. on Image Processing*, Lausanne, Septembre 1996.
- [Pic91] J. Picard. Efficiency of the extended kalman filter. *SIAM Journal of Applied Mathematics*, 51(3) :843–885, 1991.

- [PKK93] N. P. Papanikolopoulos, P. K. Khosla, et T. Kanade. Visual tracking of a moving target by a camera mounted on a robot : A combination of control and vision. *IEEE Transactions on Robotics and Automation*, 9(1), Février 1993.
- [Por90] J. Porril. Fitting ellipses and predicting confidence envelopes using a bias corrected kalman filter. *Image and Vision Computing*, 8(1) :37–41, Février 1990.
- [Pre06] M. Pressigout. *Approches hybrides pour le suivi temps-réel d'objets complexes dans des séquences vidéos*. Thèse de doctorat, Université de Rennes 1, Décembre 2006.
- [PSM94] P. Perona, T. Shiota, et J. Malik. Anisotropic diffusion. Dans *Geometry-driven diffusion in Computer Vision*, pages 73–92. Kluwer Academic Publisher, 1994.
- [PSSM00] E.M. Petriu, Z. Sakr, H. Spoelder, et A. Moica. Object recognition using pseudo-random color encoded structured light. Dans *In Proceedings of the 17th IEEE Instrumentation and Measurement technology Conference*, tome 3, pages 1237–1241, 2000.
- [PV90] I. Pitas et A.N. Venetsanopoulos. *Nonlinear digital Filters, Principles and Applications*. Kluwer Academic Publishers, 1990.
- [QL99] L. Quan et Z. Lan. Linear n-point camera pose determination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8), Août 1999.
- [RAS93] O. Ravn, N.A. Andersen, et A.T. Sorensen. Auto-calibration in automation systems using vision. Dans *Proceedings of the Int'l. Symposium on Experimental Robotics*, pages 206–218, Kyoto, Japon, Octobre 1993.
- [RDDL94] S. Rémy, M. Dhome, N. Daucher, et J.T. Lapresté. Estimation de la distortion radiale d'un système optique. Dans *Actes du 9^{ème} congrès RFIA de l'AFCE*, Paris, Janvier 1994 1994.
- [RDL87] M. Richetin, M. Dhome, et J.T. Lapresté. Attitude spatiale des courbes vues en projection. Dans *Conference AFCE/INRIA Reconnaissance des Formes et Intelligence Artificielle*, pages 671–686, Antibes, France, Novembre 1987.
- [RDLR91] M. Richetin, M. Dhome, J.T. Lapresté, et G. Rives. Inverse perspective transform using zero-curvature contour points : Application to the localization of some generalized cylinders from a single view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2) :185–192, Février 1991.
- [RE87] P. Rives et B. Espiau. Estimation récursive de primitives 3d au moyen d'une caméra mobile. *Rapport de recherche INRIA n° 351*, 1987.
- [RL87] P. Rousseeuw et A. Leroy. *Robust Regression and Outlier Detection*. Wiley, NY, 1987.
- [RN67] C. A. Rosen et N. J. Nilsson. An intelligent automaton. Dans *IEEE International Convention Record*, Part 9, 1967.
- [Ros99] P.L. Rosin. Robust pose estimation. *IEEE Transactions on Systems, Man, and Cybernetics (PartB)*, 29(2), Avril 1999.
- [RRAC05] D.B. Russakoff, T. Rohlfing, J.R. Adler, et C.R.Maurer. Intensity-based 2d-3d spine image registration incorporating a single fiducial marker. *Journal of Academic Radiology*, 12(1) :37–50, 2005.

- [RW90] P. L. Rosin et G. W. West. Segmenting curves into elliptic arcs and straight lines. Dans *Proceedings of the 3rd ICCV*, pages 75–78, Osaka, Japon, 1990.
- [SA89a] Y. C. Shin et S. Ahmad. 3d location of circular and spherical features by monocular model-based vision. Dans *Proceedings of the IEEE International Conference on System, Man and Cybernetic*, pages 576–581, Boston, USA, November 1989.
- [SA89b] Y.-C. Shiu et S. Ahmad. 3d location of circular and spherical features by monocular model-based vision. Dans *Proceedings of the IEEE Int'l Conf. on Systems, Man and Cybernetics*, pages 576–581, Cambridge, MA, USA, Novembre 1989.
- [SAJ93] A. Schweikard, J.R. Adler, et J.C.Latombe. Motion planning in stereotaxic radio-surgery. *IEEE Transactions on Robotics and Automation*, 9(6) :764–774, 1993.
- [Sam82] P.D. Sampson. Fitting conic sections to very scattered data : an iterative refinement of the bookstein algorithm. *Computer Vision, Graphics, and Image Processing*, 18 :97–108, 1982.
- [SAT99] Robert C. Susil, James H. Anderson, et Russell H. Taylor. A single image registration method for CT guided interventions. Dans *Proceedings of the Second Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*, pages 798–808, Cambridge, UK, Septembre 1999.
- [SB97] S.M. Smith et J.M. Brady. Susan - a new approach to low level image processing. *Int. Journal of Computer Vision*, 23(1) :45–78, Mai 1997.
- [SB98] V. Sundareswaran et R. Behringer. Visual-servoing-based augmented reality. Dans *In Proceedings of the IEEE Int'l. Workshop on Augmented Reality*, San Francisco, USA, Novembre 1998.
- [SBE91] C. Samson, M. Le Borgne, et B. Espiau. Robot control : The task function approach. *Oxford University Press*, 1991.
- [SBM98] J. Salvi, J. Battle, et E. Mouaddib. A robust-coded pattern projection for dynamic 3d scene measurement. *Pattern Recognition Letters*, 19 :1055–1065, 1998.
- [SC04] D.C. Schuurman et D.W. Capson. Robust direct visual servo using network-synchronized cameras. *IEEE Transactions on Robotics and Automation*, 20(2) :319–334, Avril 2004.
- [Sco04] M.L.J. Scott. *Critical Values for the Test of Flatness of a Histogram Using the Bhattacharyya Measure*. Thèse de doctorat, University of Manchester, Angleterre, Novembre 2004.
- [Set96] J.A. Sethian. *Level Set Methods*. Cambridge University Press, 1996.
- [Set99] J.A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1999.
- [SF00] M. Sonka et J.M. Fitzpatrick, rédacteurs. *Handbook of medical imaging, volume 2 : Medical image processing and analysis*. SPIE Press, Bellingham, WA, 2000. ISBN 0-8194-3622-4.
- [Shi78] Y. Shirai. Recognition of real-world objects using edges cue. *Computer Vision Systems*, pages 353–362, 1978.
- [Shi87] Y. Shirai. *Three-dimensional computer vision*. Springer Series In Symbolic Computation, 1987.

- [SI73] Y. Shirai et H. Inoue. Guiding a robot by visual feedback in assembling tasks. *Pattern Recognition*, 5(2) :99–106, 1973.
- [SI98] S. Sclaroff et J. Isidoro. Active blobs. Dans *IEEE Int'l Conference on Computer Vision*, pages 1146–1153, 1998.
- [SLL06] P. Saeedi, P.D. Lawrence, et D.G. Lowe. Vision-based 3d trajectory tracking for unknown environments. *IEEE Trans. on Robotics*, 22(1) :119–136, Février 2006.
- [SM06] F. Schramm et G. Morel. Ensuring visibility in calibration-free path planning for image-based visual servoing. *IEEE Trans. on Robotics*, 22(4) :848–854, Août 2006.
- [SMB00] C. Schmid, R. Mohr, et C. Bauckhage. Evaluation of interest point detectors. *Int'l Journal of Computer Vision*, 37(2) :151–172, 2000.
- [SN04] R. Siegwart et I.R. Nourbakhsh. *Introduction to Autonomous Mobile Robots*. MIT Press, 2004.
- [SOD04] Y. Satoh, T. Okatani, et K. Deguchi. A color-based probabilistic tracking by using graphical models. Dans *IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 631–636, Sendai, Japon, Septembre 2004.
- [Spo87] M. W. Spong. Modeling and control of elastic joint robots. *Journal of Dynamic Systems, Measurement and control*, 109 :310–319, Septembre 1987.
- [SRTSB92] R. Safaee-Rad, I. Tchoukanov, K. C. Smith, et B. Benhabib. Three-dimensional location estimation of circular features for machine vision. *IEEE Transactions on Robotics and Automation*, 8(5) :624–640, Octobre 1992.
- [SSC90] R. Smith, M. Self, et P. Cheeseman. Estimating uncertain spatial relationship in robotics. *Autonomous Robot Vehicles*, pages 167–193, 1990.
- [ST94] J. Shi et C. Tomasi. Good features to track. Dans *IEEE Int'l Conf. on Computer Vision*, pages 593–600, Seattle, USA, Juin 1994.
- [ST96] P. Sturm et W. Triggs. A factorization based algorithm for multi-image projective structure and motion. Dans *In Proceedings of The European Conference on Computer Vision*, pages 709–720, 1996.
- [Sug88] K. Sugihara. Some location problems for robot navigation using a single camera. *Computer Vision, Graphics and Image Processing*, 42(1) :112–129, Avril 1988.
- [SW80] A. C. Sanderson et L. E. Weiss. Image-based visual servo control using relational graph error signals. Dans *Proceedings of the IEEE Int'l Conf. on Cybernetics and Society (SMC)*, pages 1074–1077, Cambridge, MA, Octobre 1980.
- [SW82] A. C. Sanderson et L. E. Weiss. Image-based visual servo control of robots. Dans *Proceedings of the 26th Annual SPIE Technical Symposium*, Août 1982.
- [SWA⁺98] D. Stoianovici, L.L. Whitcomb, J.H. Anderson, R.H. Taylor, et L.R. Kavoussi. A modular surgical robotic system for image guided percutaneous procedures. Dans *Int'l Conf. on Medical Image Computing and Computer-Assisted Interventions*, Cambridge, U.S.A., Octobre 1998.
- [Tah04] O. Tahri. *Utilisation des moments en asservissement visuel et calcul de pose*. Thèse de doctorat, Université de Rennes I, Mars 2004.

- [TAVA⁺04] E. Taillant, J.C. Avila-Vilchis, C. Allegrini, I. Bricault, et P. Cinquin. Ct and mr compatible light puncture robot : Architectural design and first experiments. Dans *Int'l. Conf. on Medical Image Computing and Computer-Assisted Intervention*, pages 145–152, St Malo, France, Octobre 2004.
- [TCB⁺05] J. Troccaz, P. Cinquin, P. Berkelman, A. Vilchis-Gonzales, et E. Doidard. Surgical robots at timc : where we are and where we go. *Springer Tracts in Advanced Robotics*, 2005.
- [TFE⁺95] R. Taylor, J. Fonda, B. Eldridge, S. Gomory, K. Gruben, D. LaRose, M. Talamini, L. Kavoussi, et J. Anderson. A telerobotic assistant for laparoscopic surgery. Dans *IEEE Engineering in medicine and biology*, pages 279–288, 1995.
- [TK92] C. Tomasi et T. Kanade. Shape and motion from image streams under orthography. *Int'l J. of Comp. Vis.*, 9(2) :137–154, 1992.
- [TM87] D.W. Thompson et J.L. Mundy. Three-dimensional model matching from an unconstrained viewpoint. Dans *Int'l Conf. on Robotics and Automation*, pages 208–220, Raleigh, NC, USA, 1987.
- [TMCG02] B. Thuilot, P. Martinet, L. Cordesses, et J. Gallice. Position-based visual servoing : keeping the object in the field of vision. Dans *Proceedings of the IEEE Int'l. Conf. on Robotics and Automation*, pages 1624–1629, Washington DC, USA, 2002.
- [Tri99] W. Triggs. Camera pose and calibration from 4 or 5 known 3d points. Dans *IEEE Int'l Conf. on Computer Vision*, pages 278–284, Corfou, Grèce, 1999.
- [Tsa87] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4) :323–344, 1987.
- [Tso92] J.K. Tsotsos. Motion understanding systems for mobile robots. *Active Perception and Robot Vision*, ed. by A. Sood and H. Wechsler, 83 :1–22, 1992.
- [TV98] E. Trucco et A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [TX02] H. Teramoto et G. Xu. Camera calibration by a single image of balls : From conics to the absolute conic. Dans *Asian Conf. on Computer Vision*, tome 2, pages 499–506, Melbourne, Australie, 2002.
- [VESS04] F. Voci, S. Eiho, N. Sugimoto, et H. Sekiguchi. Estimating the gradient threshold in the perona-malik equation. *IEEE Signal Processing Magazine*, 21(3) :39–46 :65, May 2004.
- [Vin01] M. Vincze. Robust tracking of ellipses at frame rate. *Pattern Recognition*, Février 2001.
- [VLF04] L. Vacchetti, V. Lepetit, et P. Fua. Stable real-time 3d tracking using online and offline information. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(10) :1385–1391, 2004.
- [VMO05] M.A. Vitrani, G. Morel, et T. Ortmaier. Automatic guidance of a surgical instrument with ultrasound based visual servoing. Dans *Int'l Conf. on Robotics and Automation*, pages 510–515, Barcelone, Espagne, Avril 2005.

- [VSGA05] M. Vincze, M. Schlemmer, P. Gemeiner, et M. Ayromlou. Vision for robotics : A tool for model-based object tracking. *IEEE Robotics and Autonomous Magazine*, 12(4) :53–64, Décembre 2005.
- [WADP95] C. Wren, A. Azarbayejani, T. Darrell, et A. Pentland. Pfunder : Real-time tracking of the human body. Dans *Photonics East, SPIE*, 1995.
- [WB91] D. M. Wuescher et K. L. Boyer. Robust contour decomposition using a constant curvature criterion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(1) :41–51, Janvier 1991.
- [WD95] N. Werghi et C. Doignon. Contour decomposition with wavelet transform and parametrisation of elliptic curves with an unbiased extended kalman filter. Dans *Proceedings of the Second Asian Conference on Computer Vision, ACCV'95*, tome 3, pages 186–190, Singapour, République de Singapour, Décembre 1995.
- [WDA96a] N. Werghi, C. Doignon, et G. Abba. Ellipses fitting and three-dimensional localisation of objects based on elliptic features. Dans *Proceedings of the IEEE Int'l Conf. on Image Processing*, tome 1, pages 57–60, Lausanne, Suisse, Septembre 1996.
- [WDA96b] N. Werghi, C. Doignon, et G. Abba. Pose estimation of object based on circular patterns in monocular computer vision. Dans *Proceedings of SPIE Optical Science, Engineering and Instrumentation Symposium*, Denver (Colorado) USA, Août 1996.
- [Wei84] L. Weiss. *Dynamic visual servo control of robots : an adaptative image-based approach*. Thèse de doctorat, Carnegie-Mellon University, 1984.
- [Wei93] I. Weiss. Noise-resistant invariants of curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9) :943–948, September 1993.
- [Wer96] N. Werghi. *Analyse et interprétation d'images pour la reconnaissance et la localisation d'objets en vision robotique*. Thèse de doctorat, Université Louis Pasteur, Décembre 1996.
- [WGG03] J. van de Weijer, T. Gevers, et J.M. Geusebroek. Color edge detection by photometric quasi-invariants. Dans *Proc. Int'l Conf. Computer Vision*, pages 1520–1526, Nice, 2003.
- [WM00] E.A. Wan et R. Van Der Merwe. The unscented kalman filter for nonlinear estimation. Dans *Proceedings of the Symposium on Adaptive Systems for Signal Processing, Communication and Control*, pages 153–158, 2000.
- [WMFM96] M. Y. Wang, C. R. Maurer, J. M. Fitzpatrick, et R. J. Maciunas. An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head. *IEEE Transactions on Biomedical Engineering*, 43(6) :627–637, 1996.
- [WS92] D.J. Williams et M. Shah. A fast algorithm for active contours and curvature estimation. *Computer Vision, Graphics and Image Processing : Image Understanding*, 55(1) :14–26, Janvier 1992.
- [WSN87] L. E. Weiss, A. C. Sanderson, et C. P. Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE Journal of Robotics and Automation*, RA-3(5) :404–417, Octobre 1987.

- [YG94] S. Yee et P. Griffin. Three-dimensional imaging system. *Optical Engineering*, 33(6) :2070–2075, 1994.
- [YHC92] A. Yuille, P. Hallinan, et D. Cohen. Feature extraction from faces using deformable templates. *Int'l Journal of Computer Vision*, 8(2) :99–111, 1992.
- [YR97] C. C. Yang et J. J. Rodriguez. Efficient luminance and saturation processing techniques for color images. *Journal of Visual Communication and Image Representation*, 3 :263–277, Septembre 1997.
- [Yua89] J.-C. Yuan. A general photogrammetric method for determining object position and orientation. *IEEE Transactions on Robotics and Automation*, 5(2) :129–142, Avril 1989.
- [Zan03] P. Zanne. *Contributions à l'asservissement visuel robuste*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, Juin 2003.
- [Zha93] Z. Zhang. Le problème de la mise en correspondance : l'état de l'art. Dans *Rapport de recherche INRIA 2146 (RobotVis)*, Décembre 1993.
- [Zha97] Z. Zhang. Parameter estimation techniques : a tutorial with application to conic fitting. *Image and vision computing*, 15(1) :59–76, 1997.
- [Zha99] Z. Zhang. A flexible new technique for camera calibration. Dans *Proceedings of the Int'l. Conf. on Computer Vision*, Corfou, Grèce, Septembre 1999.
- [ZO02] H. Zhang et J.P. Ostrowski. Visual motion planning for mobile robots. *IEEE Trans. on Robotics and Automation*, 18(2) :199–208, 2002.
- [ZP02] X. Zhang et S. Payandeh. Application of visual tracking for robot-assisted laparoscopic surgery. *Journal of Robotic systems*, 19(7) :315–328, 2002.
- [ZW96] W. Zylka et H.-A. Wischmann. On geometric distortions in CT images. Dans *Proceedings of the IEEE Int'l Conf. on Engineering in Medicine and Biology*, tome 3, pages 1120–1121, Amsterdam, Pays-Bas, Octobre 1996.

PUBLICATIONS CHOISIES