

Université Blaise Pascal – Clermont II

École Doctorale SPI

Habilitation à Diriger des Recherches

Contributions à la commande et à la navigation par vision

Youcef MEZOUAR

Maître de Conférences à l'Université Blaise Pascal–Clermont II

Défendue le 19 Novembre 2009 devant le jury composé de :

Patrick Rives	Président
François Chaumette	Rapporteur externe
El Moustapha Mouaddid	Rapporteur externe
Philippe Martinet	Rapporteur interne
Seth Hutchinson	Examineur
Helder Araujo	Examineur

Table des matières

Avant propos	3
Chapitre 1 Activités scientifiques et administratives	7
1.1 Récapitulatif	8
1.2 Historique et contexte local	9
1.3 Responsabilités collectives et animation scientifique	9
1.4 Participation à des projets de recherche	10
1.5 Encadrements	11
1.6 Résumé de l'activité de recherche Post-thèse	13
1.7 Publications	19
Chapitre 2 Modèles pour les caméras grand angle	27
2.1 Modèle de projection Unifié	29
2.2 Reconstruction Euclidienne	33
2.2.1 Relations fondamentales	33
2.2.2 Estimation des homographies	35
2.3 Modèle unifié et caméra fish-eye	39
2.3.1 Modèle unifié versus modèle classique de distorsion radiale	41
Chapitre 3 Asservissement visuel	47
3.1 Introduction et préliminaires	47
3.2 Choix de l'inverse généralisé	52
3.2.1 Utilisation d'une Jacobienne étendue	53
3.2.2 ESM modifié	56
3.3 Asservissement visuel utilisant le modèle unifié	58
3.3.1 Asservissement visuel dans l'image	59

3.3.2	Asservissement visuel basé homographie	68
Chapitre 4 Navigation de robots mobiles par mémoire visuelle		75
4.1	Représentation de l'environnement par une mémoire sensorielle	79
4.1.1	Cartographie	79
4.1.2	Quelques définitions	81
4.1.3	Mémoire sensorielle	82
4.1.4	Système de navigation par mémoire sensorielle	83
4.2	Localisation dans une mémoire d'images grand angle	85
4.3	Suivi de chemin visuel	86
4.4	Exemple de résultats	87
Bilan et Perspectives		93
Bibliographie		99

Avant propos

L'attrait de l'Homme pour la technique et les nouvelles technologies n'a eu de cesse d'alimenter son imaginaire débordant. Aussi dans les années 60, à l'émergence de l'informatique, chercheurs et ingénieurs auguraient sans mal l'avènement des robots autonomes pourvus d'une intelligence artificielle à l'horizon 2000. De telles perspectives motivèrent les débats d'éthique quant à l'évolution des robots et de leur rapport à l'humain, agrémentés d'ouvrages de science fiction à caractère anticipatoire [Ass72]. Force est de constater que la communauté est en retard sur les prévisions. La très grande majorité de la population planétaire de robots officie dans un cadre industriel conditionné pour une productivité optimale, comme travailleurs automatiques à la chaîne. Ceux-ci répètent machinalement une séquence de mouvements déterminés, sans aucune capacité d'adaptation à leur environnement. Quoiqu'il en soit, ces systèmes robotiques n'ont pas aujourd'hui les moyens matériels nécessaires à cette adaptation car ils ne sont que très rarement équipés de capteurs extéroceptifs susceptibles de les informer sur leur environnement de travail. Les enjeux économiques de la robotique manufacturière sont clairs. Il s'agit, pour les industriels, d'optimiser la production en améliorant le rendement des chaînes à moindre coût. Le groupe Toyota Motors envisage d'ici dix ans l'usine entièrement robotisée, préférant miser à long terme sur la flexibilité, la performance et la robustesse de cette main d'œuvre robotique plutôt que sur des délocalisations d'activités à la recherche d'opérateurs certes peu coûteux pour leur qualification, mais dotés d'une certaine conscience propre et collective pouvant les amener un jour ou l'autre à la critique de leurs conditions de travail.

Le robot suscite donc des réactions paradoxales. Il est une machine, conçue et optimisée en vue d'une tâche identifiée, pour laquelle son rendement est supérieur à celui de l'ouvrier de chair et d'os. Il apparaît comme un rival contre lequel la bataille de l'emploi est perdue d'avance. Pourtant, outre les objectifs de productivité industrielle, le robot est conçu dans le but de se substituer à l'homme dans des tâches "machinales", pénibles, physiquement difficiles, pour lesquelles l'emploi d'une machine dédiée est fortement apprécié. Le robot peut rendre service. Doté d'une autonomie suffisante pour l'exécution de tâches somme-toute complexes, il force aujourd'hui les portes de la maison du particulier et entre dans sa vie, en lui apportant assistance et confort. Il est au centre des réflexions sur l'avenir des transports routiers et urbains. Les perspectives économiques sont énormes, et ce à moyen terme.

Les statistiques publiées par l'UNECE (United Nations Economic Commission for Europe) et l'IFR (International Federation of Robotics) confortent les dires du ministère de l'économie japonais, selon lequel l'industrie de la robotique couvrira à l'horizon 2025 un marché plus important que l'industrie de l'informatique [EGP04]. Elle devrait notamment profiter de la dynamique spectaculaire de l'évolution du matériel informatique qui devrait conduire le PC à des performances

équivalentes à celles du cerveau humain sur le même horizon temporel, selon Hans Moravec du Robotics Institute of Carnegie Mellon University [Mor98a, Mor98b].

Longtemps, seules des applications militaires et spatiales ont guidé la recherche et le développement de la robotique autonome, notamment aux États-Unis, où la Nasa et la Darpa ont su profiter de ces lignes directrices fortement financées pour asseoir leur savoir-faire et celui de leurs partenaires universitaires et industriels. Dans un registre plus civil, par le biais de grands programmes d'investissement nationaux, le Japon s'est affirmé mondialement à la pointe de la mécatronique, et compte sur l'essor à moyen et long terme de la robotique domestique et de service pour porter le développement de son industrie de la robotique. L'implication de grands industriels (dont les grands noms Sony, Honda, Fujitsu), la densité des laboratoires concernés sur le territoire et l'intérêt d'une population technophile forment pour cela un cadre idéal. Conscientes des enjeux économiques sous-jacents, les nations européennes investissent dans des projets nationaux et tentent de partager leurs connaissances via des réseaux européens d'excellence (citons parmi ceux-ci EURON et CLAWAR).

En France le feu département STIC (Sciences et Technologie de l'Information et de la Communication) du CNRS a affiché la robotique autonome comme l'une de ses priorités. Le programme interdisciplinaire "Robotique et entités artificielles" (ROBEA) a ainsi vu le jour en mai 2001, programme auquel s'est associé l'INRIA à partir de 2002. Le groupe GRAVIR du LASMEA s'est largement investi dans ce programme, en prenant part entre 2001 et 2005 à cinq projets couvrant le large spectre de ses activités scientifiques, de la perception à la commande de systèmes robotiques complexes. L'équipe ROSACE de GRAVIR a notamment profité de ce contexte pour donner un cadre applicatif porteur à ses activités de recherches. Historiquement, c'est dans le cadre d'un de ces projets (OMNIBOT) qu'ont débutées en 2002 mes activités de recherche dans le domaine de la commande par vision omnidirectionnelle. Le groupe GRAVIR a également intégré dans la même année un consortium à l'origine du projet WACIF, labélisé par le RNTL (2002-2004)¹. Les points clefs confiés au LASMEA dans le cadre de ce projet sont à l'origine de mes travaux autour de la navigation de robots mobiles par mémoire sensorielle. Je me suis investi par la suite, dans plusieurs projets nationaux, européens et internationaux dans la continuité de ces travaux dont le dénominateur commun est la commande référencée capteurs.

Comme cela a été évoqué précédemment, la grande majorité des systèmes robotiques, dépourvus de capteurs extéroceptifs, dispose d'une autonomie réduite au minimum. Les travaux décrits dans ce manuscrit s'inscrivent dans le domaine de la robotique et visent à doter ces systèmes d'un degré d'autonomie nécessaire à la réalisation d'une tâche donnée. L'objectif de mes travaux peut donc être résumé comme suit :

Proposer des stratégies de perception et de commande permettant à des robots de réaliser une tâche donnée de façon autonome, c'est-à-dire uniquement à partir d'informations fournies en temps réel par des capteurs.

¹Le RNTL (Réseau National des Technologies Logicielles) est une structure initiée en 1999 par les ministères français de la Recherche et de l'Industrie, pérennisé en 2005 dans le cadre de l'ANR (Agence Nationale de la Recherche). WACIF fut labélisé projet exploratoire en 2002 pour une durée de 24 mois.

Nous avons ainsi abordé cette vaste problématique avec comme fil conducteur la commande référencée capteurs. Celle-ci consiste à contrôler les mouvements d'un système dynamique à partir d'un ensemble d'informations extraites des images acquises par un ou plusieurs capteurs extéroceptifs. Même si nous avons considéré le cas des capteurs d'effort, nous nous sommes plus particulièrement intéressés aux capteurs visuels grand-angles. En outre, nous nous sommes intéressés aux tâches de positionnement de robots manipulateurs et aux tâches de navigation pour les robots mobiles.

Ce mémoire donne le contexte dans lequel ces travaux se sont développés et synthétise les résultats obtenus. Il est structuré en cinq chapitres abordant respectivement les sujets suivants :

- le premier chapitre décrira l'ensemble de mes activités professionnelles,
- le second chapitre sera consacré à la modélisation des caméras grands angles,
- dans le troisième chapitre, je présenterai quelques un de nos travaux en asservissement visuel,
- le quatrième chapitre sera consacré à la description de nos travaux ayant trait à la navigation de robots mobiles,
- finalement, le cinquième chapitre dresse un bilan de nos contributions dans les domaines de la commande et la navigation par vision et présente les perspectives qui en découlent.

Chapitre 1

Activités scientifiques et administratives

Ce chapitre synthétise mes activités scientifiques et administratives. On y trouve tout d'abord un récapitulatif de mes contributions et activités. Il détaille ensuite mes responsabilités collectives, mes activités d'animation scientifique et d'encadrement ainsi que mon implication dans des projets nationaux et internationaux. Il se termine par un résumé de mes travaux de recherche post-thèse, accompagné de la liste de mes publications.

1.1 Récapitulatif

Je donne ci-dessous un récapitulatif de mes contributions et activités.

Publications

La classification utilisée pour les journaux et conférences est celle du GDR robotique (pour les journaux : Rang A correspond à des revues d'excellent niveau, Rang B à des revues bon niveau ; pour les conférences : Rang A correspond à des conférences très sélectives et O aux autres conférences)

	Total	thèse	après thèse
Revue internationale	10	rang A : 2 rang B : 1	rang A : 6 rang B : 1
Revue nationale	1	0	1
Chapitres de livre	4	0	4
Congrès internationaux	45	rang A : 6 O : 1	rang A : 27 O : 11
Ateliers internationaux	13	1	12
Congrès nationaux	3	0	3
Totaux	76	11	65

Encadrements

Le tableau suivant récapitule les encadrements d'étudiants auxquels j'ai participé sur la période 2002-2009 (à partir du niveau M2). Le tableau est divisé en trois niveaux. Pour chaque niveau, le nombre d'étudiants encadrés est regroupé par taux d'encadrement.

	Plus ou 50%	entre 30% et 50%	Total
Niveau thèse	3	3	6
Niveau Master 2	7	0	7
Post-doctorant	3	0	3

Projets de recherche

Je participe ou j'ai participé à 6 projets internationaux (dont 2 projets européens), 7 projets nationaux et à 2 prestations industrielles.

Enseignements

J'effectue la majorité de mes enseignements dans le département de physique de l'UFR ST (Sciences et Technologies) à l'Université Blaise Pascal. J'ai mis en place des travaux pratiques en automatique, vision par ordinateur et robotique. J'ai régulièrement encadré des projets d'étudiants et de fin d'études. Je participe également au suivi des stages en entreprise, et aux jurys.

Responsabilités pédagogiques

- Responsable de la filière Mécatronique de l'IUP GSI (Clermont II) depuis 2004.
- Responsable du Master 2 Mécatronique de l'IUP GSI depuis 2004 et responsable de la Licence de l'IUP GSI de 2002 à 2004.

Autres éléments

- Co-responsable de l'équipe ROSACE du Lasmea depuis 2009 et responsable de l'activité de recherche VISIR de ROSACE depuis 2002.
- Membre des comités de pilotage et de gouvernance, co-responsable de l'axe recherche appliquée de l'UMT'08 Mécanisation-robotisation dans les filières viandes et produits carnés.
- Co-responsable de l'activité de recherche « structure sérielle redondante » du projet M2I (Machines et Mécanismes Innovants) de la fédération de recherche TIMS (Technologies de l'Information, de la Mobilité et de Sécurité).
- Participation à diverses commissions au niveau du LASMEA et des Universités Clermontoises (vice président de la commission de spécialistes 61ème section de l'UBP de 2006 à 2008).
- Membre du comité d'organisation des JNRR'03 et SSIR'08.
- Titulaire de la PEDR depuis 2006.

1.2 Historique et contexte local

Après un séjour Post-Doctoral dans le laboratoire de Robotique de l'Université de Columbia à New-York, j'ai été recruté le 1er Septembre 2002 en tant que Maître de conférences à l'Université Blaise Pascal. Depuis cette date, j'enseigne dans le département de physique de l'UFR Sciences et Technologie, en particulier à l'IUP Génie des Systèmes Industriels (GSI).

Mes activités de recherche, quant à elles, se déroulent au sein du LASMEA (Laboratoire des Sciences et Matériaux pour l'Electronique et d'Automatique). Le LASMEA, dirigé actuellement par Michel Dhome, est une Unité Mixte de Recherche (UMR) Université Blaise Pascal/CNRS du secteur des sciences pour l'ingénieur. Les travaux de recherche qui y sont menés relèvent des sciences et technologies de l'information et portent sur les matériaux pour l'électronique ou l'optoélectronique, l'électromagnétisme, la vision artificielle, la robotique. Le laboratoire est composé de deux groupes : MATELEC (MATériaux pour l'ELEctronique) et GRAVIR (Groupe d'Auto-matique et VIsion pour la Robotique), c'est dans ce dernier groupe que s'inscrivent mes activités. Ce groupe est organisé en thèmes. A ce jour, trois thèmes sont affichés : PERSYST (PERception SYSTems), ComSee (Computer that See) et ROSACE (RObtoticS and Autonomous Complex systEms). Mes activités de recherche sont en particulier rattachées au thème ROSACE, dont je suis co-responsable depuis 2009. Ce thème est organisé autour de trois activités scientifiques :

- VISIR VISual ServoIng for Robots,
- AGV Autonomous Guided vehicle,
- MICMAC Modeling, identification and control of complex machines.

Je suis responsable de l'activité scientifique VISIR et certains de mes travaux participent au développement des activités AGV et MICMAC. Trois actions principales sont menées dans VISIR :

- l'asservissement visuel par vision omnidirectionnelle,
- la navigation topologique de robots mobiles,
- la commande référencée multi-capteurs.

Mes travaux s'inscrivent également dans deux grands projets portés par la fédération de recherche TIMS (Technologies de l'Information, de la Mobilité et de Sécurité) : V2I (Véhicules et Infrastructures Intelligents) et M2I (Machines et Mécanismes Innovants).

Enfin, je participe au développement de l'UMT'08 Mécanisation-robotisation dans les filières viandes et produits carnés (créée en 2008) dans laquelle, je suis co-responsable de l'axe recherche appliquée et membre des comités de pilotage et de gouvernance.

1.3 Responsabilités collectives et animation scientifique

J'ai assumé ou assume les responsabilités suivantes :

- Co-responsable de l'équipe de recherche ROSACE (depuis 2009)
- Responsable de l'activité de recherche VISIR (depuis 2002)
- Co-responsable de l'axe recherche appliquée de l'UMT08 Mécanisation-robotisation dans les filières viandes et produits carnés (depuis 2009)
- Co-responsable de l'activité de recherche « structure sérielle redondante » du projet M2I (Machines et Mécanismes Innovants) de la fédération de recherche TIMS (Technologies de l'Information, de la Mobilité et de Sécurité)

- Membre des comités de pilotage et de gouvernance de l'UMT'08 Mécanisation-robotisation dans les filières viandes et produits carnés
- Responsable de la filière Mécatronique de l'IUP GSI (Université Blaise Pascal, Clermont-Fd) depuis 2004.
- Responsable du MASTER 2 Mécatronique de l'IUP GSI (Université Blaise Pascal, Clermont-Fd) depuis 2004
- Responsable de la Licence de l'IUP GSI (Université Blaise Pascal, Clermont-Fd) jusqu'en 2004
- Membre élu de la commission de spécialistes 61 de l'Université Blaise Pascal-Clermont-Fd de 2004 à 2008
- Vice-président de la commission de spécialistes 61 de l'Université Blaise Pascal-Clermont-Fd de 2006 à 2008.

J'ai par ailleurs été membre des comités d'organisation de SSIR 2008 (Summer School on Image and Robotics 2008) et du congrès national JNRR 2003 (Journée Nationale de Recherche en Robotique 2003).

Concernant les activités de relecture d'articles, j'ai apporté ma contribution aux journaux suivants :

TRO	IEEE Transactions on Robotics
ITS	IEEE Transactions on Intelligent Transportation System
IJRR	International Journal of Robotic Research
AR	Autonomous Robots
RAS	Robotic and Autonomous Systems
IJCV	International Journal of Computer Vision
CVIU	Computer Vision and Image Understanding
IJO	International Journal of Optomechatronics
TS	Traitement du Signal

J'ai par ailleurs été relecteur pour différents congrès (cette liste n'est pas exhaustive) : IEEE Int. Conf. on Robotics and Automation (ICRA), IEEE Int. Conf. on Intelligent Robots and Systems (IROS), IEEE Int. Conf. on Decision and Control (CDC), European Control Conference (ECC), IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR).

J'ai exercé en tant qu'expert pour les Appels à Projet ANR TECSAN et CAROTTE.

1.4 Participation à des projets de recherche

Je me suis investi et m'investit encore dans différents projets nationaux :

- **R-DISCOVER** (2009-2013). Réseaux de robots mobiles : couverture décentralisée de l'espace basé vision omnidirectionnelle. Projet ANR-CONTINT : Programme de recherche et d'innovation dans les transports terrestres.
- **SRDViand** (2007-2010). Développement de systèmes mécanisés-robotisés pour le désosage et la découpe des animaux de boucherie (2007-2010). Projet financé par le FCE (Fond de Compétitivité des Entreprises) et la Région Auvergne. Je suis le responsable administra-

- tif et scientifique de ce projet pour le LASMEA (budget de 500Keuros pour le LASMEA).
- **CITYVIP** (2008-2010). Déplacement sûr de véhicules individuels adaptés à l’environnement urbain. Projet ANR-PREDIT : Programme de recherche et d’innovation dans les transports terrestres.
 - **MOBIVIP** (2003-2006). Véhicules individuels publics pour la mobilité en centre ville. Projet financé par le PREDIT 3. Programme de recherche et d’innovation dans les transports terrestres.
 - **BODEGA** (2003-2005). Navigation autonome et sûre en environnement urbain . Projet financé par ROBEA : programme national robotique et entités artificielles du CNRS.
 - **WACIF** (2002-2004). Etude et développement d’un assistant robotique destiné au grand public réalisant des tâches de téléprésence et de télésurveillance et s’intégrant dans un contexte de maison communicante. Projet financé par le RNTL (Réseau National des Technologies Logicielles).
 - **OMNIBOT** (2002-2004). De la vision omnidirectionnelle à la commande de robots mobiles. Projet financé par ROBEA : programme national robotique et entités artificielles du CNRS.

Par ailleurs, j’ai participé ou je participe aux projets internationaux suivants :

- **NEXT** (2005-2009). Next generation production system. European 6th Framework Programme project.
- **SAFEMOVE** (2004-2005). Systèmes de transport intelligents et navigation autonome. Programme Franco - Coréen Science and Technology Amicable Research Programme (STAR).
- **FACT** (2005-2007). French asian cyber transportation. ICT-ASIA.
- **CITYHOME** (2008-2010). Cyber transportation in cities environment and mobile service robotic in home. ICT-ASIA.
- **μ -drone** (2005-2009). Navigation autonome de mini-drone. European 6th Framework Programme project. Nous intervenons dans ce projet en tant que sous traitant du CEA.
- **PICASSO**. Robotic Intelligence Laboratory Université de Jaume I - LASMEA (2001-2003) Asservissement visuel stéréo et Couplage vision/force.

J’ai également participé à deux prestations pour le compte du CEA (Centre d’Etude Atomique). La première a concerné la commande par vision d’un robot manipulateur mobile non instrumenté en vue d’intervention dans les centrale nucléaires (2003-2004). La seconde a porté sur la Cartographie 3d par vision pour drones.

1.5 Encadrements

Cette section reprend la liste de mes encadrements sur la période 2002-2009.

Doctorants

Je passe en revue ci-dessous les doctorants que j’ai encadré ou que j’encadre encore à l’heure actuelle. Ils sont tous inscrits à l’Université Blaise Pascal de Clermont-Ferrand. J’indique pour chaque doctorant l’année de démarrage et la date de soutenance quand elle a eu lieu, le pourcentage d’encadrement assuré et les noms des co-encadrants. Je mentionne également le sujet de thèse.

- **Saleha Belakelal** (Nov. 2008 - soutenance prévue en 2011) co-encadrement (60%) avec Mohamed Tadjine du Laboratoire d'automatique de l'Ecole National Polytechnique d'Alger et Nicolas Andreff (Directeur de thèse). Sujet : *Couplage vision-force pour les robots parallèles*. Les premiers travaux de Saleha ont porté sur une nouvelle loi de commande vision/effort pour des tâches d'insertion.
- **Franck Stephan** (Nov. 2007 - soutenance prévue en 2010) co-encadrement (70%) avec Philippe Martinet (Directeur de thèse). Sujet : *Saisie et manipulation d'objets déformables*. Franck s'attaque à la problématique peu explorée de la saisie et de la manipulation d'objets déformables. Il a débuté ces travaux en étudiant le cas des objets articulés.
- **Jonathan Courbon** (Nov. 2006 - soutenance prévue en 2009) co-encadrement (60%) avec Philippe Martinet (Directeur de thèse). Sujet : *Navigation de robots mobiles par mémoire sensorielle*. Jonathan travaille sur une méthodologie complète et générique pour la navigation de robots mobiles à roues et volants en milieu intérieur et extérieur à partir d'informations acquises par une caméra à point centrale unique (caméra conventionnelle, fish-eye et catadioptrique).
- **Redwan Dahmouche** (Nov. 2006 - soutenance prévue en 2009) co-encadrement (30%) avec Nicolas Andreff et Philippe Martinet (Directeur de thèse). Sujet : *Asservissement visuel dynamique*. Redwan travaille sur la modélisation des capteurs de vision à acquisition séquentielle. Il a proposé des méthodes originales de calcul de pose et vitesse qu'il exploite dans le contexte des asservissements visuels rapides.
- **Hicham Hadj-Abdelkader** (Nov. 2003 - Déc. 2006) co-encadrement (40%) avec Nicolas Andreff et Philippe Martinet (Directeur de thèse). Sujet : *Asservissement visuel en vision omnidirectionnelle*. Hicham a travaillé sur les modélisations géométrique et cinématique des informations acquises par une caméra catadioptrique pour la commande de robots manipulateurs et mobiles à roues.
- **Guillaume Blanc** (Nov. 2002) co-encadrement (40%) avec Thierry Chateau et Philippe Martinet (Directeur de thèse). Sujet : *Navigation de robots mobiles par mémoire visuelle*. Guillaume a participé à nos premiers développements de navigation topologique pour des robots mobiles à partir d'images acquises par une caméra conventionnelle en milieu intérieur. Il n'a pas souhaité mener à terme son travail de thèse pour créer sa société exoTIC Systems qui conçoit et développe des systèmes électroniques et informatiques embarqués.

Postdoctorants

J'ai également eu le plaisir d'encadrer trois stages post-doctoraux :

- **Michela Farenza** (1 an, 2007-2008) co-encadrement (50%) avec Adrien Bartoli. Sujet : *Navigation d'un drone par vision artificielle*. Michela a travaillé sur la conception d'un système de navigation d'un drone par vision artificielle robuste. Ceci inclut la détection et la gestion des singularités et une régularisation adaptative de la trajectoire.
- **Omar Tahri** (1 an, 2007-2008) encadrement (100%). Sujet : *Vision omnidirectionnelle pour les robots parallèles*. Omar a travaillé sur le suivi et la commande par vision omnidirectionnelle de la plateforme de Gough-Stewart. Il est actuellement associé au Laboratoire de vision et robotique de l'Université de Coimbra. Nous continuons à collaborer sur la thématique de l'asservissement visuel.

- **Gabriel Recatala Ballester** (1 an, 2004-2005) encadrement (100%). Sujet : *Stratégie de saisie pour la manipulation d'objets a priori inconnus*. Gabriel a travaillé sur un modèle de synthèse de point de saisie et d'exécution de tâche de saisie. Ce modèle est fortement inspiré par les travaux en neurophysiologie humaine et s'appuie sur une reconstruction 3D sélective de l'objet à saisir.

Stages niveau Master 2

Je me contente ici de lister mes encadrements à 100% de stagiaires de niveau Master 2. J'ai en outre encadré depuis 2002, 14 étudiants de niveau Licence.

- **Jonathan Courbon** (5 mois en 2006). *Sujet : Navigation autonome de robots mobiles dans une mémoire visuelle*. Stage du Master 2 Composants et Systèmes pour le Traitement de l'Information option VIsion pour la RObotique de l'Université Blaise-Pascal.
- **Jérôme roblet** (5 mois en 2005). *Sujet : Evitement d'obstacles pour robots mobiles*. Stage du Master 2 Mécatronique de l'Université Blaise-Pascal, Clermont-Ferrand.
- **Romain Riol** (5 mois en 2005). *Sujet : Accrochage virtuel de cycab par télémètre laser*. Stage du Master 2 Mécatronique de l'Université Blaise-Pascal, Clermont-Ferrand.
- **Mamadou Mbengue** (5 mois en 2004). *Sujet : Etude et simulation d'un capteur omni-directionnel muni d'une rétine log-polaire*. Stage du Master 2 GEII de l'Université Joseph Fourier, Grenoble.
- **Hicham Hadj-Abdelkader** (5 mois en 2003). *Sujet : Asservissement visuel avec des caméras catadioptriques*. Stage du Master 2 Composants et Systèmes pour le Traitement de l'Information option VIsion pour la RObotique de l'Université Blaise-Pascal, Clermont-Ferrand.
- **Mario Sanchez Prats** (5 mois en 2003). *Sujet : Couplage vision/force*. Stage de 3ème année ingénieur de l'IFMA/Universitat Jaume I de Castellon.
- **Sandra Quiroga Arias** (10 mois en 2002). *Sujet : Contribution à la commande de bras manipulateurs par couplage vision/force*. Stage de 3ème année ingénieur de Polytech'Clermont/Escuela Technica Superior de Ingenieros Industriales Valadolid.

1.6 Résumé de l'activité de recherche Post-thèse

L'asservissement visuel et plus généralement la commande référencée capteurs est la source principale d'inspiration de mes travaux. J'ai débuté dans cette thématique lors de ma thèse consacrée à la planification de trajectoires pour l'asservissement visuel dont les résultats sont présentés dans [Mez01]. J'ai ensuite abordé des sujets variés connexes à cette thématique dont les cibles applicatives vont de la micromanipulation de protéines à la navigation de mini-drones. Dans les chapitres suivants, je décrirai plus précisément certaines de mes contributions dans les domaines de l'asservissement visuel et de la navigation de robots mobiles. La suite de cette partie, quant à elle donne une vue d'ensemble de mes travaux post-thèse ainsi que quelques repères bibliographiques pour une lecture plus détaillée. L'ensemble de mes publications est listé dans la section suivante.

Micromanipulation de protéines par vision. Après ma thèse, au cours d'un post-doctorat

à l'Université de Columbia à New-York, j'ai étudié l'asservissement visuel à l'échelle du micron. Nous avons alors proposé un système complet pour le positionnement d'un outil par rapport à une cellule cristallisée. Le système proposé, composé d'un microscope observant le manipulateur et la cellule à manipuler, a permis une précision de positionnement inférieure à la centaine de micron. Les articles suivants donnent les détails de ces travaux :

Microrobotic crystal mounting using computer vision.

A. Georgiev, P. K. Allen, and Y. Mezouar.

WMB'03–*Workshop on Microrobotics for Biomanipulation* at IROS'03, Las Vegas, USA, October 2003.

Visual servoed micropositioning for protein manipulation tasks.

Y. Mezouar and P. K. Allen.

IROS'02–*IEEE/RSJ International Conference on Intelligent Robots and System*, Lausanne, Switzerland, October 2002.

Asservissement par vision omnidirectionnelle. Ce thème de recherche a été motivé par l'intérêt croissant de la communauté pour les capteurs omnidirectionnels dans des applications robotiques (localisation, navigation, commande d'une flotte de robots etc ...). En effet, alors que les caméras classiques souffrent d'un champ de vision limité, ce type de capteur permet de saisir l'ensemble de l'environnement de travail d'un robot en une seule prise de vue. Plus précisément, l'objectif de ces recherches est de proposer des stratégies de commande des systèmes robotiques en utilisant comme entrées d'une boucle d'asservissement visuel des primitives relatives à la projection dans le plan image de caméra panoramique d'une scène 3D. Afin de proposer de telles stratégies, il faut à la fois analyser les fonctions de projection permettant le passage du monde 3D au plan image de la caméra omnidirectionnelle et les relations cinématiques liant les mouvements de la caméra et les mouvements des primitives dans l'image.

Dans un premier temps, des techniques d'asservissement visuel classiques ont été abordées. Les primitives type points et droites ont été traitées pour les caméras catadioptriques à centre de projection unique. Des résultats théoriques sur la stabilité de l'asservissement visuel 2D points ont été obtenus. Nous avons également proposé l'utilisation de nouvelles primitives basées moments dans l'image afin d'améliorer les propriétés de découplage de la commande. Ensuite, des algorithmes de reconstruction euclidienne à partir des primitives points et droites ont été proposés. Ceux-ci ont permis de développer des schémas d'asservissement visuel originaux exhibant des propriétés de découplage intéressantes. Ces méthodes ont été développées en utilisant des modèles suffisamment génériques pour englober une classe entière de capteurs de vision. Plus précisément, elles permettent de traiter de manière indifférente les caméras catadioptriques centrale, sphérique et perspective ainsi qu'une grande partie des capteurs de vision avec optique fish-eye. Il en découle une généralité importante des résultats obtenus aussi bien dans le contexte de la reconstruction euclidienne que dans celui de l'asservissement visuel. Enfin, les résultats théoriques ont été validés par de nombreuses expérimentations à la fois sur un bras robotisé à 6 degrés de liberté, sur un robot parallèle (plateforme de Gough-Stewart) et sur des robots mobiles.

Nous décrivons de manière synthétique ces contributions dans les chapitres 2 et 3. Les articles suivants en donnent les détails.

Points based visual servoing with central cameras

H. Hadj Abdelkader, Y. Mezouar and P. Martinet

In Visual Servoing via Advanced Numerical Methods, Edited by : Graziano Chesi and Koichi Hashimoto, Publisher : Springer, to appear in 2009.

Visual servoing and pose estimation with cameras obeying the unified model

O. Tahri, Y. Mezouar, F. Chaumette and H. Araujo

In Visual Servoing via Advanced Numerical Methods, Edited by : Graziano Chesi and Koichi Hashimoto, Publisher : Springer, to appear in 2009.

Omnidirectional visual-servo of a gough-stewart platform.

O. Tahri, Y. Mezouar, N. Andreff, and P. Martinet.

TRO-*IEEE Transactions on Robotics*, 25(1) :178–183, February 2009.

Catadioptric visual servoing from 3d straight lines.

H. Hadj-Abdelkader, Y. Mezouar, P. Martinet, and F. Chaumette.

TRO- *IEEE Transactions on Robotics*, 24(3) :652–665, June 2008.

A generic fisheye camera model for robotic applications.

J. Courbon, Y. Mezouar, L. Eck, and P. Martinet.

IROS'07- *IEEE/RSJ International Conference on Intelligent Robots and Systems*, San-diego, California, October 2007.

Omnidirectional visual servoing from polar lines.

H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet.

ICRA'06-*IEEE International Conference on Robotics and Automation*, Orlando, Florida, USA, May 2006.

Robustness of central catadioptric image-based visual servoing to uncertainties on 3D parameters.

Y. Mezouar and E. Malis

IROS'04 - *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, October 2004.

Inverse généralisé de la matrice d'interaction. Comme nous le verrons dans le Chapitre 3, l'inverse généralisé de la matrice d'interaction joue un rôle primordial dans la construction de lois

de commande et conditionne le comportement en boucle fermée du système à réguler. Plusieurs choix sont possibles, parmi les plus classiques on retrouve les inverses généralisés reflétant les non linéarités du première ordre de la relation entre l'espace des observations et l'espace de commande (le choix le plus courant est alors l'inverse de Moore-Penrose). Nous avons proposé deux alternatives à ce choix permettant d'obtenir un inverse généralisé reflétant des non-linéarités d'ordre 2. Ce point sera abordé dans le chapitre 3 et les détails peuvent être consultés dans les articles suivants :

On the efficient second order minimization and image-based visual servoing.

O. Tahri and Y. Mezouar.

ICRA'08-IEEE International Conference on Robotics and Automation, Pasadena, California, USA, May 2008.

A hessian approach to visual servoing.

J. T. Lapreste and Y. Mezouar.

IROS'04-IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, October 2004.

Couplage vision/force et manipulation. Un de nos objectifs à moyen terme est la réalisation de tâches de saisie et de manipulation d'objets a priori inconnus et possiblement déformables. Les techniques d'asservissement visuel que nous développons contribuent largement à cet objectif. Nous nous sommes également penchés de manière plus spécifique sur le couplage vision/effort et l'estimation de points de saisie. Combiner adéquatement des informations visuelles et d'efforts dans une boucle de commande est une tâche très importante lorsque l'outil porté par le robot doit rentrer en contact avec l'environnement. Afin de répondre à ce problème, nous avons proposé une nouvelle stratégie pour le couplage vision/force. Dans le schéma proposé, la boucle de commande par vision est placée sous une boucle externe de commande en effort. Le couplage est réalisé dans l'espace image, la commande en effort modifiant les trajectoires des primitives visuelles. Nous avons montré que notre méthode permettait la convergence dans certaines configurations difficiles où les méthodes classiques échouaient. Cette stratégie a également été adaptée avec succès au cas des robots parallèles. En outre, nous avons mis en place une méthodologie pour la saisie d'objets complexes. Dans ce contexte, nous avons proposé un modèle pour la synthèse de point de saisie et l'exécution de la tâche. Ce modèle est fortement inspiré par les travaux en neurophysiologie humaine et s'appuie sur une reconstruction 3D sélective de l'objet à saisir. On trouvera les détails de ces méthodes dans les articles suivants :

Biologically-inspired 3d grasp synthesis based on visual exploration.

G. Recatalá, E. Chinellato, A. P. Del Pobil, Y. Mezouar, and P. Martinet.

AR-In Autonomous Robots, 25(1-2) :59-70, 2008.

Force/position control of parallel robot using exteroceptive pose measurements

S. Bellakehal, N. Andreff, Y. Mezouar, F. Pacot, and M. Tadjine.

Parallel'08–International workshop on Fundamental Issues and Future Research Directions for Parallel Mechanisms and Manipulators, Montpellier, France, October 2008.

External hybrid vision/force control.

Y. Mezouar, M. Prats, and P. Martinet.

ICAR'07–13th International Conference on Advanced Robotics, Jeju, Korea, August 2007.

Asservissement visuel basé sur l'acquisition séquentielle des primitives. L'acquisition séquentielle des primitives visuelles a un double intérêt en asservissement visuel. Le premier est que ce type d'acquisition permet d'augmenter de manière significative la fréquence d'échantillonnage de la boucle de commande en réduisant la quantité de données à acquérir et à transmettre. Le second avantage est associé au modèle de projection qui dépend alors à la fois de la pose et de la vitesse de l'objet observé. En exploitant ces propriétés, nous avons proposé de nouvelles méthodes d'estimation de pose et de vitesse et un schéma de commande original dont le vecteur de commande est composé des torseurs cinématique et dynamique. Cette loi de commande permet d'obtenir des performances très satisfaisantes de la commande pour des suivis de trajectoires à vitesse élevée. Les détails de ces méthodes sont données dans :

Visual servoing based on sequential region of interest acquisition.

R. Dahmouche, N. Andreff, Y. Mezouar, and P. Martinet.

IROS'09–IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, October 2009.

High-speed pose and velocity measurement from vision.

R. Dahmouche, O. Ait-Aider, N. Andreff, and Y. Mezouar.

ICRA'08–IEEE International Conference on Robotics and Automation, Pasadena, California, USA, May 2008.

Navigation topologique par mémoire sensorielle. La quasi-totalité des stratégies de navigation autonome se base sur l'analyse des signaux capteurs afin de construire une carte globale ou locale de l'environnement. Cette carte est en général référencée dans le monde 3D et est plus ou moins précise selon les types de capteurs utilisés et l'effort algorithmique associé. La localisation du robot est alors réalisée dans la carte et ses mouvements sont générés en extrayant la consigne de la carte et en la comparant à la localisation courante. L'objectif de ces recherches est de proposer une alternative à ces méthodes afin d'améliorer la précision des mouvements du robot tout en limitant la complexité algorithmique. Pour cela, il a été proposé de construire une carte qui ne soit pas référencée dans le monde 3D mais plutôt directement dans l'espace des capteurs à travers une mémoire sensorielle organisée de manière à représenter la topologie

de l'environnement. Le principal intérêt de cette nouvelle approche est la conservation d'une image fidèle de l'environnement en omettant la reconstruction 3D d'un monde virtuel (soumis à propagation d'erreurs). Dans ce cadre, la localisation du robot s'effectue directement en comparant des informations capteurs et ses mouvements sont générés en extrayant la consigne de la mémoire sensorielle. Des commandes référencées capteurs peuvent alors être mises en œuvre pour contrôler précisément les mouvements du robot. Un logiciel spécifique, dénommé SOVIN (SOftware for VIvisual Navigation), a été élaboré afin d'aborder des sites de navigation de très grande taille (échelle d'un campus ou d'une ville par exemple). Cette approche de navigation topologique a été validée pour différents types de robots mobiles terrestres (Intérieur : Pioneer 2DX et 3AT, Extérieur : ROBUCAB) ainsi que sur un robot mobile volant : le drone X4 Flyer du CEA. D'autre part, l'approche proposée est également valide pour une classe importante de caméras (caméra classique, caméra catadioptrique centrale, caméra avec optique fish-eye). Les dernières démonstrations sur le campus des Cégeux ont montré l'efficacité de telles approches pour des tâches de navigation sur des distances importantes (une expérimentation de navigation autonome d'un véhicule urbain électrique sur 1.7km a été réalisée).

Autonomous navigation of vehicles from a visual memory based on the use of generic camera model.

J. Courbon, Y. Mezouar, and P. Martinet.

ITS-IEEE Transaction on Intelligent Transportation System, to appear in 2009.

Visual navigation of a quadrotor aerial vehicle

J. Courbon, Y. Mezouar, N. Guenard, P. Martinet.

IROS'09-IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, October, 2009.

Indoor navigation of a non-holonomic mobile robot using a visual memory.

J. Courbon, Y. Mezouar, and P. Martinet.

AR-In Autonomous Robots, 25(3) :253-266, 2008.

Navigation of urban vehicle : An efficient visual memory management for large scale environments.

J. Courbon, Y. Mezouar, L. Lequievre, and L. Eck.

IROS'08-IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, September 2008.

Efficient hierarchical localization method in an omnidirectional images memory.

J. Courbon, Y. Mezouar, L. Eck, and P. Martinet.

ICRA'08-IEEE International Conference on Robotics and Automation, Pasadena, California, USA, May 2008.

Indoor navigation of a wheeled mobile robot along visual routes.

G. Blanc, Y. Mezouar, and P. Martinet.

ICRA'05–IEEE International Conference on Robotics and Automation, Barcelona, Spain, April 2005.

Construction de carte métrique par vision monoculaire. Il est clair que l'ajout de connaissances a priori dans les méthodes séquentielles de reconstruction 3D par vision peut améliorer la qualité des résultats. Dans cette optique, nous avons proposé de coupler au processus de reconstruction une fonction de pénalité contraignant la caméra à suivre une trajectoire continue. En général, le poids d'une telle fonction est fixé par des heuristiques et des tests. Nous avons proposé de l'estimer en utilisant une méthode de validation croisée. Dans sa forme de base, cette méthode est très coûteuse. Afin d'en réduire le coût, nous en avons proposé une version largement modifiée ("Gauss-Newton Cross Validation").

Efficient camera smoothing in sequential structure-from-motion using approximate cross-validation.

M. Farenzena, A. Bartoli, and Y. Mezouar.

ECCV'2008–European Conference on Computer Vision, Marseille, France, October 2008.

Automatically smoothing camera pose using cross validation for sequential vision-based 3d mapping.

M. Farenzena, A. Bartoli, and Y. Mezouar.

IROS'08–IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, September 2008.

1.7 Publications

Articles dans des revues internationales à comité de lecture

- [1] J. Courbon, Y. Mezouar, and P. Martinet. Autonomous navigation of vehicles from a visual memory based on the use of generic camera model. *IEEE Transactions on Intelligent Transportation System*, 2009.
- [2] E. Malis, Y. Mezouar, and P. Rives. Robustness of image-based visual servoing with a calibrated camera in the presence of uncertainties in the 3d structure. *IEEE Transactions on Robotics*, To appear 2009.
- [3] Y. Mezouar. Optimal camera trajectory under visibility constraint in visual servoing. *Advanced Robotics*, 23(12-13) :534–549, October 2009.
- [4] O. Tahri, Y. Mezouar, N. Andreff, and P. Martinet. Omnidirectional visual-servo of a gough-stewart platform. *IEEE Transactions on Robotics*, 25(1) :178–183, February 2009.

- [5] J. Courbon, Y. Mezouar, and P. Martinet. Indoor navigation of a non-holonomic mobile robot using a visual memory. *Autonomous Robots*, 25(3) :253–266, 2008.
- [6] G. Recatalá, E. Chinellato, A. P. Del Pobil, Y. Mezouar, and P. Martinet. Biologically-inspired 3d grasp synthesis based on visual exploration. *Autonomous Robots*, 25(1-2) :59–70, 2008.
- [7] H. Hadj-Abdelkader, Y. Mezouar, P. Martinet, and F. Chaumette. Catadioptric visual servoing from 3d straight lines. *IEEE Transactions on Robotics*, 24(3) :652–665, June 2008.
- [8] Y. Mezouar and F. Chaumette. Optimal camera trajectory with image-based control. *International Journal of Robotics Research, IJRR*, 22(10) :781–804, October 2003.
- [9] Y. Mezouar and F. Chaumette. Avoiding self-occlusions and preserving visibility by path planning in the image. *Robotics and Autonomous Systems*, 41(2) :77–87, November 2002.
- [10] Y. Mezouar and F. Chaumette. Path planning for robust image-based control. *IEEE Transactions on Robotics and Automation*, 18(4) :534–549, August 2002.

Articles dans des revues nationales à comité de lecture

- [1] H. Hadj Abdelkader, Y. Mezouar, P. Martinet, and F. Chaumette. Asservissement visuel en vision omnidirectionnelle à partir de droites. *Traitement du Signal*, 22(5) :462–482, September 2005.

Chapitres de livre

- [1] H. Hadj Abdelkader, Y. Mezouar, and P. Martinet. *Points based visual servoing with central cameras*, chapter 17. Edited by Graziano Chesi and Koichi Hashimoto : Visual Servoing via Advanced Numerical Methods. Springer, 2009.
- [2] O. Tahri, Y. Mezouar, F. Chaumette, and H. Araujo. *Visual Servoing and Pose Estimation with Cameras obeying the Unified Model*, chapter 18. Edited by Graziano Chesi and Koichi Hashimoto : Visual Servoing via Advanced Numerical Methods. Springer, 2009.
- [3] Y. Mezouar. *A variational approach to trajectory planning in visual servoing*, chapter 8. Edited by Graziano Chesi and Koichi Hashimoto : Visual Servoing via Advanced Numerical Methods. Springer, 2009.
- [4] Y. Mezouar, H. Hadj Abdelkader, and P. Martinet. *Omnidirectional Vision-Based Control From Homography*, chapter 21, pages 387–408. Vision Systems Applications, Edited by Goro Obinata and Ashish Dutta, ISBN 978-3-902613-01-1. I-Tech Education and Publishing, Vienna, Austria, June 2007.

Articles dans des congrès internationaux avec actes et comité de lecture

- [1] J. Courbon, Y. Mezouar, N. Guenard, and P. Martinet. Visual navigation of a quadrotor aerial vehicle. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'09*, St. Louis, MO, USA, 11–15 October 2009.
- [2] R. Dahmouche, N. Andreff, Y. Mezouar, and P. Martinet. Visual servoing based on sequential region of interest acquisition. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'09*, St. Louis, MO, USA, 11–15 October 2009.
- [3] H. H. Abdelkader, Y. Mezouar, and P. Martinet. Decoupled visual servoing based on the spherical projection of points. In *IEEE International Conference on Robotics and Automation, ICRA'09*, Kobe, Japan, April 2009.
- [4] O. Tahri, Y. Mezouar, F. Chaumette, and P. Corke. Generic decoupled image-based visual servoing for cameras obeying the unified projection model. In *IEEE International Conference on Robotics and Automation, ICRA'09*, Kobe, Japan, April 2009.
- [5] J. Courbon, Y. Mezouar, L. Lequievre, and L. Eck. Efficient visual memory based navigation of indoor robot with a wide-field of view camera. In *International Conference on Control, Automation, Robotics and Vision, ICAR-CV'2008*, pages 268–273, Hanoi, Vietnam, December 2008.
- [6] M. Farenzena, A. Bartoli, and Y. Mezouar. Efficient camera smoothing in sequential structure-from-motion using approximate cross-validation. In *European Conference on Computer Vision, ECCV'2008*, Marseille, France, October 2008.
- [7] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet. Efficient hierarchical localization method in an omnidirectional images memory. In *IEEE International Conference on Robotics and Automation, ICRA'08*, pages 13–18, Pasadena, California, USA, 19–23 May 2008.
- [8] J. Courbon, Y. Mezouar, L. Lequievre, and L. Eck. Navigation of urban vehicle : An efficient visual memory management for large scale environments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'08*, pages 1817–1822, Nice, France, 22–26 September 2008.
- [9] M. Farenzena, A. Bartoli, and Y. Mezouar. Automatically smoothing camera pose using cross validation for sequential vision-based 3d mapping. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'08*, pages 3616–3621, Nice, France, 22–26 September 2008.
- [10] R. Dahmouche, O. Ait-Aider, N. Andreff, and Y. Mezouar. High-speed pose and velocity measurement from vision. In *IEEE International Conference on Robotics and Automation, ICRA'08*, pages 107–112, Pasadena, California, USA, 19–23 May 2008.
- [11] O. Tahri, F. Chaumette, and Y. Mezouar. New decoupled visual servoing scheme based on invariants from projection onto a sphere. In *IEEE International Conference on Robotics and Automation, ICRA'08*, pages 3238–3243, Pasadena, California, USA, 19–23 May 2008.

- [12] O. Tahri and Y. Mezouar. On the efficient second order minimization and image-based visual servoing. In *IEEE International Conference on Robotics and Automation, ICRA'08*, pages 3213–3218, Pasadena, California, USA, 19–23 May 2008.
- [13] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet. A generic fisheye camera model for robotic applications. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'07*, pages 1683–1688, San-Diego, California, USA, 29 October–2 November 2007.
- [14] O. Tahri, Y. Mezouar, N. Andreff, and P. Martinet. Omnidirectional visual-servo of a gough-stewart platform. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'07*, pages 1326–1331, San-Diego, California, USA, 29 October–2 November 2007.
- [15] H. Hadj-Abdelkader, Y. Mezouar, and P. Martinet. Decoupled visual servoing from a set of points imaged by an omnidirectional camera. In *IEEE International Conference on Robotics and Automation, ICRA'07*, pages 1697–1702, Roma, Italy, 10–14 April 2007.
- [16] Y. Mezouar, M. Prats, and P. Martinet. External hybrid vision/force control. In *13th International Conference on Advanced Robotics, ICAR'07*, pages 170–175, Jeju, Korea, 21.–24 August 2007.
- [17] E. Chinellato, G. Recatala, A. P. del Pobil, Y. Mezouar, and P. Martinet. 3d grasp synthesis based on object exploration. In *IEEE International Conference on Robotics and Biomimetics, ROBIO'06*, pages 1065–1070, 17–20 December 2006.
- [18] H. Hadj-Abdelkader, Y. Mezouar, and P. Martinet. Path planning for image based control with omnidirectional cameras. In *45th IEEE Conference on Decision and Control, CDC'06*, pages 1764–1769, San-Diego, California, USA, 13–15 December 2006.
- [19] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Decoupled homography-based visual servoing with omnidirectional cameras. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'06*, pages 2332–2337, Beijing, China, 9–14 October 2006.
- [20] T. Dallej, N. Andreff, Y. Mezouar, and P. Martinet. 3d pose visual servoing relieves parallel robot control from joint sensing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'2006*, pages 4291–4296, Beijing, China, 9–14 October 2006.
- [21] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Homography based visual servoing with omnidirectional cameras : the cases of points and lines. In *IEEE International Conference on Methods and Models in Automation and Robotics, MMAR'06*, Miedzyzdroje, Poland, 28–31 August 2006.
- [22] G. Blanc, Y. Mezouar, and P. Martinet. A path planning strategy for obstacle avoidance. In *3rd International Conference on Informatics in Control, Automation and Robotics, ICINCO'06*, Setubal, Portugal, 1–5 August 2006.

-
- [23] O. Ait Ader, G. Blanc, Y. Mezouar, and P. Martinet. Pattern tracking and visual servoing for indoor mobile environment mapping and autonomous navigation. In *3rd International Conference on Informatics in Control, Automation and Robotics, ICINCO'06*, Setubal, Portugal, 1–5 August 2006.
- [24] G. Recatala, E. Chinellato, A. P. del Pobil, Y. Mezouar, and P. Martinet. 3d grasp synthesis based on a visual cortex model. In *IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechanics, BioRob'06*, pages 164–169, Pisa, Italy, 20–22 February 2006.
- [25] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Omnidirectional visual servoing from polar lines. In *IEEE International Conference on Robotics and Automation, ICRA'06*, pages 2385–2390, Orlando, Florida, USA, 15–19 May 2006.
- [26] G. Blanc, J. Roblet, Y. Mezouar, and P. Martinet. Control assistance of a mobile robot navigating into an encumbered environment. In *15th International Symposium on Measurement and Control in Robotics, ISMCR'05*, 18–22 November 2005.
- [27] G. Blanc, Y. Mezouar, and P. Martinet. Indoor navigation of a wheeled mobile robot along visual routes. In *IEEE International Conference on Robotics and Automation, ICRA'2005*, pages 3354–3359, Barcelona, Spain, 18–22 April 2005.
- [28] H. H. Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Image-based control of mobile robot with central catadioptric cameras. In *IEEE International Conference on Robotics and Automation, ICRA'2005*, pages 3522–3527, Barcelona, Spain, 18–22 April 2005.
- [29] H. H. Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. 2 1/2 d visual servoing with central catadioptric cameras. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'2005*, pages 3572–3577, Edmonton, Canada, 2–6 August 2005.
- [30] G. Blanc, T. Chateau, Y. Mezouar, Ait-Ader O., Martinet P., Eck L., Moreau V., and A. Nadim. Implementation of a vision based navigation framework on a house mobile robot prototype. In *International Conference on Advances in Intelligent Systems - Theory and Applications*, Kirchberg, Luxembourg, November 2004.
- [31] Y. Mezouar and E. Malis. Robustness of central catadioptric image-based visual servoing to uncertainties on 3d parameters. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'04*, volume 2, pages 1389–1394, Sendai, Japan, 28 September–2 October 2004.
- [32] J. T. Lapreste and Y. Mezouar. A hessian approach to visual servoing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'04*, volume 1, pages 998–1003, Sendai, Japan, 28 September–2 October 2004.
- [33] Y. Mezouar, H. H. Abdelkader, P. Martinet, and F. Chaumette. Central catadioptric visual servoing from 3d straight lines. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'04*, volume 1, pages 343–348, Sendai, Japan, 28 September–2 October 2004.

- [34] H. H. Abdelkader, Y. Mezouar, and P. Martinet. Paracatadioptric visual servoing from lines. In *35th International Symposium on Robotics, ISR'04*, Paris, France, 23–26 March 2004.
- [35] G. Blanc, O. Ait Ader, Y. Mezouar, T. Chateau, and P. Martinet. Indoor navigation of mobile robot : An image based approach. In *35th International Symposium on Robotics, ISR'04*, Paris, France, 23–26 March 2004.
- [36] G. Blanc, Y. Mezouar, and P. Martinet. An image driven vision based control framework for wheeled mobile robots. In *International Symposium on Robotics and Automation, ISRA'04*, pages 507–512, Queretaro, Mexico, 25–27 August 2004.
- [37] G. Blanc, O. Ait Aider, Y. Mezouar, T. Chateau, and P. Martinet. Autonomous image based navigation in indoor environment. In *fifth IFAC Symposium on Intelligent Autonomous Vehicles, IAV'04*, Instituto Superior Técnico, Lisbon, Portugal, 5–7 July 2004.
- [38] Y. Mezouar and P. K. Allen. Visual servoed micropositioning for protein manipulation tasks. In *IEEE/RSJ International Conference on Intelligent Robots and System, IROS'02*, volume 2, pages 1766–1771, Lausanne, Switzerland, 30 September–5 October 2002.
- [39] Y. Mezouar, A. Remazeilles, P. Gros, and F. Chaumette. Images interpolation for image-based control under large displacement. In *Proc. IEEE International Conference on Robotics and Automation, ICRA '02*, volume 4, pages 3787–3794, Washington, DC, USA, 11–15 May 2002.
- [40] Y. Mezouar and F. Chaumette. Model-free optimal trajectories in the image space : application to robot vision control. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'01*, volume 1, pages I-1155–I-1162, Kauai, Hawaii, USA, 2001.
- [41] Y. Mezouar and F. Chaumette. Model-free optimal trajectories in the image space. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'01*, volume 1, pages 25–30, Maui, Hawaii, USA, 29 October–3 November 2001.
- [42] Y. Mezouar and F. Chaumette. Visual servoing by path planning. In *European Conference on Control, ECC'01*, pages 2904–2909, Porto, Portugal, September 2001.
- [43] Y. Mezouar and F. Chaumette. Avoiding self-occlusions and preserving visibility by path planning in the image. In *9th International Symposium on Intelligent Robotic Systems, SIRS'01*, Toulouse, France, July 2001.
- [44] Y. Mezouar and F. Chaumette. Design and tracking of desirable trajectories in the image space by integrating mechanical and visibility constraints. In *IEEE International Conference on Robotics and Automation, ICRA'01*, volume 1, pages 731–736, Seoul, Korea, 2001.
- [45] Y. Mezouar and F. Chaumette. Path planning in image space for robust visual servoing. In *IEEE International Conference on Robotics and Automation, ICRA'00*, volume 3, pages 2759–2764, San-Francisco, USA, 24–28 April 2000.

Articles dans des workshops internationaux

- [1] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet. A generic framework for topological navigation of urban vehicles. In *Workshop on Safe navigation in open and dynamic environments Application to autonomous vehicles at IEEE International Conference on Robotics and Automation, ICRA'09*, 11–15 October 2009.
- [2] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet. Visual navigation of indoor/outdoor non-holonomic robot with wide field of view cameras. In *Workshop on Omnidirectional Robot Vision*, Venice, Italy, November 2008.
- [3] S. Bellakehal, N. Andreff, Y. Mezouar, F. Pacot, and M. Tadjine. Force/position control of parallel robot using exteroceptive pose measurements. In *Second International workshop on Fundamental Issues and Future Research Directions for Parallel Mechanisms and Manipulators, Parallel'08*, Montpellier, France, October 2008.
- [4] J. Courbon, G. Blanc, and P. Martinet. Navigation of a non-holonomic mobile robot with a memory of omnidirectional image. In *Workshop on Planning, Perception and Navigation for Intelligent Vehicles at IEEE International Conference on Robotics and Automation, ICRA'07*, Roma, Italy, 10–14 April 2007.
- [5] P. Martinet, Y. Mezouar, and H. Hadj Abdelkader. Omnidirectional visual servoing. In *International Workshop on Vision and Visual Servoing, IWVV07*, Sungkyunkwan University, Suwon, Korea, 14 May 2007.
- [6] H. Hadj Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Central catadioptric vision for robot control. In *2nd French-Korean Workshop on Dependable Robotic Navigation, SAFEMOVE'05*, Suwon, Korea, 27–28 October 2005.
- [7] G. Blanc, Y. Mezouar, and P. Martinet. Navigating mobile robots using the visual memory concept. In *1st French-Korean Workshop on Dependable Robotic Navigation, SAFEMOVE'04*, Suwon, Korea, 4–5 October 2004.
- [8] Y. Mezouar, H. Hadj Abdelkader, P. Martinet, and F. Chaumette. Visual servoing from 3d straight lines with central catadioptric cameras. In *Fifth Workshop on Omnidirectional Vision, Omnivis'2004*, Prague, Czech Republic, May 2004.
- [9] A. Georgiev, P. K. Allen, and Y. Mezouar. Microrobotic crystal mounting using computer vision. In *Microrobotics for Biomanipulation Workshop, IROS'03, Las Vegas, NV*, 2003.
- [10] Y. Mezouar and P. K. Allen. Fast manipulation of proteins by visual servoing. In *NorthEast Structural Genomic Consortium, Crystallography process Workshop, Columbia University, New York*, March 2002.
- [11] Y. Mezouar and F. Chaumette. Path planning for visual servoing. In *2nd Summer Systems and Control Workshop, Sidney, Australie*, February 2002.

Articles dans des congrès nationaux

- [1] L. Lequievre, J. Courbon, Y. Mezouar, and P. Martinet. Sovin, a software platform for visual navigation. In *3rd National Conference on Control Architectures of Robots, CAR'08*, Bourges, France, 29–30 September 2008.
- [2] H. Hadj Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Utilisation des droites polaires en asservissement visuel omnidirectionnel. In *Congrès sur la Reconnaissance des Formes et Intelligence Artificielle, RFIA'2006*, January 2006.
- [3] H. Hadj Abdelkader, Y. Mezouar, and P. Martinet. Utilisation des droites en asservissement visuel omnidirectionnel. In *Congrès sur la Reconnaissance des Formes et Intelligence Artificielle, RFIA'2004*, January 2004.

Chapitre 2

Modèles pour les caméras grand angle

Ce chapitre aborde la modélisation des systèmes de vision grand angle. Il couvre, dans un premier temps, la présentation du modèle de projection utilisé dans la suite du document. Il présente ensuite les méthodes mises en œuvre pour la reconstruction Euclidienne à partir de la projection d'un ensemble de points ou de droites dans une image panoramique. Il discute finalement l'utilisation du modèle "sphérique" (couramment utilisé pour les caméras catadioptriques à point central unique) pour les caméras fish-eye.

Les systèmes de vision omnidirectionnelle ont suscité un intérêt particulier dans la communauté roboticienne ces dix dernières années. Les premières applications robotiques pour ce type d'imageur remontent au début des années 90. Yagie *et al.* dans [YK90, YK94] proposent ainsi d'embarquer un capteur combinant un miroir conique à une caméra conventionnelle sur un robot mobile. Un peu plus tard Pegard *et al.* en 1995 utilisent ce même type de capteur dans [PM96] pour la navigation et la localisation 3D d'un robot mobile. En ce qui me concerne, c'est en 2002 que j'ai fait mes premiers pas dans le monde de l'omnidirectionnelle par le biais du projet ROBEA OMNIBOT. Je m'intéressais alors à l'utilisation des capteurs catadioptriques dans le contexte de l'asservissement visuel. A l'époque peu de chercheurs s'intéresse à cette problématique, le seul travail relevant de cette discipline étant à mettre à l'actif de Baretto *et al.* en 2002 [BMH02]. Dans cet article, l'asservissement visuel 2D point est étendu au cas où les observations sont réalisées par une classe particulière de caméras omnidirectionnelles (les caméras catadioptriques à point central unique sur lesquelles nous reviendrons). Nous présentons ici quelques éléments de modélisation qui nous seront fort utiles dans les deux chapitres suivants ayant trait à l'asservissement visuel et à la navigation autonome de robots mobiles.

Bien souvent, la géométrie des systèmes de vision omnidirectionnelle ne peut pas être décrite par un modèle de projection simple [BK00]. Toutefois, lorsque le centre de projection du capteur (point central) est supposé unique (c'est à dire que chaque pixel sur le plan image mesure la luminance du rayon passant par le point central dans une direction particulière et connue) alors

un modèle de projection relativement simple est disponible [GD00]. Ces capteurs sont d'un intérêt particulier car après étalonnage, leurs propriétés géométriques sont similaires à celles des caméras modélisées par une projection perspective. Cela implique que des algorithmes conçus pour les caméras classiques pourront assez aisément être transposés au cas plus général des caméras centrales.

Dès 1637 René Descartes, dans son célèbre Discours de la Méthode, présente le concept de système optique à point central unique [DS37]. Il démontre alors que les surfaces réfractives et réfléchives coniques permettent de focaliser les rayons lumineux en un point unique si elles sont illuminées à partir d'une source dont l'emplacement est soigneusement choisi. Ce concept est reformulé bien plus tard par Feynman *et al.* en 1963 [FLS63] et Hecht *et al.* en 1974 [HZ74]. Baker et Nayar le revisitent en 1998 et posent les principes fondateurs de la vision omnidirectionnelle moderne [Nay97, BN98, BN99]. Ils déduisent des propriétés géométriques des systèmes de vision centraux, une liste exhaustive des capteurs à points centraux combinant des miroirs de révolution avec des imageurs conventionnels (caméra catadioptrique). Une de leurs conclusions est qu'un système catadioptrique central peut être obtenu en combinant un miroir plan, sphérique, conique, hyperbolique ou elliptique avec une caméra perspective et un miroir parabolique avec une caméra orthographique. Cependant, seuls les miroirs hyperboliques et paraboliques sont utilisés en pratique pour concevoir les caméras catadioptriques à point central unique [PSH01]. En effet, le miroir plan n'apporte aucun accroissement du champ visuel. Les miroirs coniques et sphériques, quant à eux, sont des solutions dégénérées pour un système catadioptrique central et le miroir elliptique n'est pas utilisé en pratique car le champ de vision du capteur catadioptrique associé est inférieur à une demi-sphère.

Les associations miroir hyperbolique-caméra perspective et miroir parabolique-caméra orthographique combinent donc deux caractéristiques importantes : un centre de projection unique et un champ de vision important. Cependant, d'un point de vue applicatif les capteurs catadioptriques ont des inconvénients majeurs : 1) une zone importante au centre de l'image acquise est inexploitable, 2) la présence du miroir rend le capteur encombrant et fragile. Afin d'accroître le champ de vision d'une caméra, deux autres solutions sont envisageables. La première est basée sur la génération d'une mosaïque à partir d'une série d'images acquise par une caméra en mouvement (le plus généralement pivotant autour d'un axe perpendiculaire à son axe optique) ou par plusieurs caméras conventionnelles [FAB⁺00]. Cette solution a l'avantage de fournir des images panoramiques avec une très grande résolution et l'inconvénient d'être complexe à mettre en œuvre (étalonnage, synchronisation). La seconde solution (qui nous intéressera dans la suite) consiste à utiliser des lentilles spécifiques à très courte focale (lentille fish-eye). Cette solution est très attractive d'un point de vue applicatif car elle permet une observation grand angle simple à mettre en œuvre, sans zone morte dans l'image, peu encombrante et bien moins fragile qu'un système catadioptrique.

Geyer *et al.* dans [GD00] ont proposé un modèle unifié pour les capteurs catadioptriques centraux. Dans la suite de ce manuscrit, nous exploiterons largement ce modèle désormais classique. Nous l'exploiterons dans son contexte original (capteurs catadioptriques centraux) mais également pour les caméras grand angle (fish-eye). C'est pourquoi dans ce chapitre, nous discutons sur un plan théorique et pratique sa validité dans le cadre de nos applications robotiques. Nous abordons également les méthodes mises en œuvre pour la reconstruction Euclidienne à partir de la projection d'un ensemble de points ou de droites dans une image panoramique. Dans ce

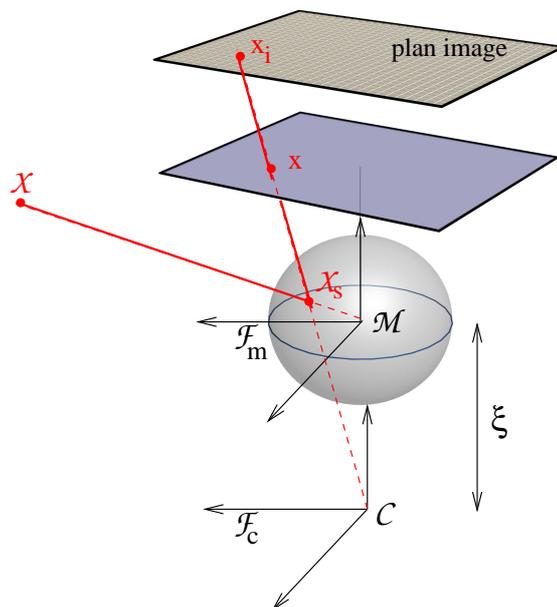


FIG. 2.1 – Modèle de projection unifié

contexte et à notre connaissance, seule l'estimation de la matrice essentielle avait été abordée dans la littérature. Nous avons de notre côté proposé des méthodes d'estimation de la transformation homographique entre deux images d'un objet plan. Notre motivation était alors d'étendre les techniques d'asservissement visuel hybride à l'ensemble des caméras à centre de projection unique (ce point sera abordé dans le chapitre 3). En effet, ces techniques exploitent largement cette transformation afin d'éviter les singularités de la matrice essentielle en particulier à la convergence. Ces méthodes nous seront donc bien sûr utiles dans le cadre de l'asservissement visuel mais également dans le contexte de la navigation autonome.

La première partie de ce chapitre couvre la présentation du modèle de projection utilisé dans la suite du document. La seconde partie présente les méthodes mises en œuvre pour la reconstruction Euclidienne à partir de la projection d'un ensemble de points ou de droites dans une image panoramique. Enfin la dernière partie discute l'utilisation du modèle "sphérique" (couramment utilisé pour les caméras catadioptriques à point central unique) pour les caméras fish-eye.

2.1 Modèle de projection Unifié

Le modèle standard pour une caméra perspective (modèle pinhole) transforme tout point \mathcal{X} de la scène, en l'intersection de la droite joignant \mathcal{X} au centre optique avec le plan image normalisé de coordonnées homogènes \mathbf{x}_p :

$$\mathbf{x}_p = \begin{bmatrix} \frac{X}{Z} & \frac{Y}{Z} & 1 \end{bmatrix}^T \propto \mathbf{P}\mathbf{X} \quad (2.1)$$

$\mathbf{X} = [X Y Z]^\top$ représente les coordonnées de \mathcal{X} dans \mathcal{F}_c et $\mathbf{x}_p = [x_p y_p 1]^\top$ représente les coordonnées homogènes de sa projection dans le plan image normalisé. $\mathbf{P} \in \mathbb{R}^{3 \times 4}$ est la matrice de projection, $\mathbf{P} = [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]$. Le point dans le plan image réel est alors obtenu après une transformation homographique \mathbf{K}_p :

$$\mathbf{p} = \mathbf{K}_p \mathbf{x}_p \quad (2.2)$$

$\mathbf{K}_p \in \mathbb{R}^{3 \times 3}$ contient les paramètres intrinsèques de la caméra. Notons f la focale de la caméra et θ l'angle d'incidence du rayon $\mathcal{X}\mathcal{C}$. La distance r_p entre le point image de coordonnées \mathbf{x}_p et le point central (projection du centre optique dans l'image) est donnée par :

$$r_p(\theta) = f \tan \theta = f \sqrt{x_p^2 + y_p^2} \quad (2.3)$$

Le modèle simple de projection donné par (2.1) ou (2.3) est largement utilisé pour des caméras sans distorsion. Pour les capteurs avec distorsion, comme les caméras catadioptrique ou fish-eye, d'autres modèles sont disponibles dans la littérature. Nous allons plus particulièrement nous intéresser au modèle unifié de projection proposé dans [GD00]. Il permet une description simple de l'ensemble des capteurs catadioptriques centraux. Ce modèle repose sur deux projections successives, une première projection centrale sur une sphère virtuelle (induisant une distorsion de l'image) suivie d'une projection perspective sur le plan image.

Considérons une sphère virtuelle unitaire centrée en \mathcal{M} et associons lui le repère \mathcal{F}_m comme indiqué sur la Figure 2.1. On peut maintenant supposer qu'une caméra perspective placée en l'origine du repère caméra \mathcal{F}_c observe le monde 3D à travers cette sphère unitaire. Nous supposons dans la suite que \mathcal{F}_m et \mathcal{F}_c sont liés par une translation de $-\xi$ le long de l'axe Z de \mathcal{F}_m . Soit \mathcal{X} un point de la scène 3D de coordonnées $\mathbf{X} = [X Y Z]^\top$ dans \mathcal{F}_m . Le point \mathcal{X} est projeté dans le plan en un point de coordonnées homogènes $\mathbf{x}_i = [x_i y_i 1]^\top$. La formation du point \mathbf{x}_i peut être décomposée en trois étapes :

- **Premièrement**, \mathcal{X} est projeté sur la surface de la sphère unité :

$$\mathbf{X}_s = \frac{1}{\rho} [X \quad Y \quad Z]^\top \quad (2.4)$$

où $\rho = \|\mathbf{X}\| = \sqrt{X^2 + Y^2 + Z^2}$.

- **Ensuite**, le point \mathbf{X}_s appartenant à la sphère unité est projeté perspectivement sur le plan image normalisé $Z = 1 - \xi$ en un point de coordonnées homogènes :

$$\mathbf{x} = \mathbf{f}(\mathbf{X}) = \left[\frac{X}{Z + \xi\rho} \quad \frac{Y}{Z + \xi\rho} \quad 1 \right]^\top \quad (2.5)$$

- **Finalement**, le point dans l'image réelle \mathbf{x}_i est obtenu après une transformation homographique \mathbf{K} :

$$\mathbf{x}_i = \mathbf{K} \mathbf{x}$$

$\mathbf{K} \in \mathbb{R}^{3 \times 3}$ contient les paramètres intrinsèques de la caméra couplés aux paramètres intrinsèques du miroir. La transformation \mathbf{K} et le paramètre ξ peuvent être obtenus en utilisant une procédure

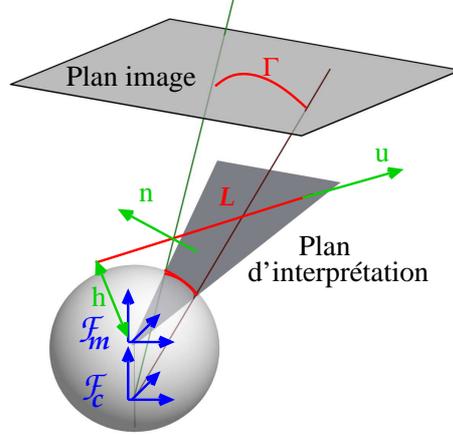


FIG. 2.2 – Modèle de projection unifié pour une droite

d'étalonnage comme celle proposée dans [BA02a] par exemple. De manière évidente, le modèle pinhole peut être obtenu en posant $\xi = 0$ dans l'équation (2.5).

Supposons maintenant le capteur étalonné. Alors le transformé d'un point appartenant au plan image en un point de la sphère unitaire peut être obtenu en inversant les deuxième et troisième étapes du processus de formation de l'image. Le point \mathbf{x} dans le plan image normalisé est obtenu simplement :

$$\mathbf{x} = [x \ y \ 1]^\top = \mathbf{K}^{-1}\mathbf{x}_i \quad (2.6)$$

Le point correspondant sur la sphère unité est ensuite obtenu en inversant la fonction de projection (2.5) :

$$\mathbf{X}_s = \mathbf{f}^{-1}(\mathbf{x}) = \lambda \left[x \ y \ 1 - \frac{\xi}{\lambda} \right]^\top \quad (2.7)$$

où $\lambda = \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1}$.

Supposons maintenant que la caméra observe une droite de la scène 3D. Dans un espace euclidien, celle-ci peut être définie par un couple de vecteurs et un scalaire (coordonnées

de Plücker binormalisées) :

$$(\mathcal{L}) : \begin{bmatrix} \mathbf{u} \\ \mathbf{h} \\ h \end{bmatrix} \quad (2.8)$$

Géométriquement, le vecteur \mathbf{u} est un vecteur unitaire directeur de \mathcal{L} , le vecteur \mathbf{h} est un vecteur unitaire normal au plan d'interprétation passant par l'origine du repère de définition et contenant la droite \mathcal{L} (voir Figure 2.2) et h est la profondeur de la droite \mathcal{L} (distance de la droite au centre du repère associé au miroir \mathcal{F}_m). Il est bien connu que la projection perspective de la droite \mathcal{L} sur le plan image est la droite 2D de coordonnées \mathbf{l}_p avec :

$$\mathbf{l}_p = \mathbf{K}_p^{-\top} \mathbf{h} = \mathbf{K}_p^{-\top} \begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix} \quad (2.9)$$

De manière plus générale, la transformée d'une droite par le modèle unifié est une conique dont la forme matricielle est $\mathbf{\Omega}_i = \mathbf{K}^{-\top} \mathbf{\Omega} \mathbf{K}^{-1}$ avec [BA05] :

$$\mathbf{\Omega} \propto \begin{pmatrix} h_x^2 - \xi^2(1 - h_y^2) & h_x h_y(1 - \xi^2) & h_x h_z \\ h_x h_y(1 - \xi^2) & h_y^2 - \xi^2(1 - h_x^2) & h_y h_z \\ h_x h_z & h_y h_z & h_z^2 \end{pmatrix} \quad (2.10)$$

Un point $\mathbf{x}_i = [x_i \ y_i \ 1]^\top$ appartient donc à cette conique s'il vérifie l'équation quadratique suivante :

$$A_0 x_i^2 + A_1 y_i^2 + 2A_2 x_i y_i + 2A_3 x_i + 2A_4 y_i + A_5 = 0 \quad (2.11)$$

où les coefficients A_i sont fonctions de \mathbf{h} et du paramètre ξ . L'équation (2.11) est définie à un facteur d'échelle près. Cette ambiguïté peut cependant être levée en normalisant cette équation par un des coefficients A_i ou encore une combinaison linéaire, non linéaire ou mixte entre les coefficients de l'équation quadratique. Nous utiliserons dans le chapitre 2, les coefficients de (2.11) normalisés par A_5 (notés $B_i \ i \in 0 \dots 4$).

En outre, l'image d'une droite 3D (une fois normalisée) est définie par cinq coefficients alors qu'elle ne possède que deux degrés de liberté. Par multiplication à gauche de la forme matricielle $\mathbf{\Omega}_i$ de la conique Γ par un point \mathbf{A} appartenant au plan de définition de la conique, on obtient une droite appelée **droite polaire** de \mathbf{A} par rapport à Γ . Soit $\mathbf{l}_i \propto \mathbf{\Omega}_i \mathbf{O}_i$ la droite polaire associée au point principal de l'image ($\mathbf{O}_i = [u_0 \ v_0 \ 1]^\top$) par rapport à la conique $\mathbf{\Omega}_i$. Les coordonnées de cette droite polaire sont données par

$$\mathbf{l}_i \propto \mathbf{K}^{-\top} \mathbf{h} \quad (2.12)$$

Ce résultat est intéressant car le vecteur normal au plan d'interprétation apparaît explicitement dans l'équation de projection. Rappelons que l'image catadioptrique d'une droite 3D est complètement définie par son plan d'interprétation. Dans le cas d'une caméra conventionnelle, $\mathbf{l}_p \propto \mathbf{K}_p^{-\top} \mathbf{h}$ n'est rien d'autre que l'image d'une droite 3D appartenant au plan d'interprétation défini par \mathbf{h} . La droite polaire \mathbf{l}_i associée au point principal de l'image par rapport à la conique est donc la projection perspective de \mathcal{L} dans l'image d'une caméra virtuelle avec comme

paramètre intrinsèque \mathbf{K} . Ce résultat est intéressant puisqu'il nous permet de représenter la projection d'une droite 3D dans l'image d'une caméra répondant au modèle unifié par une simple droite (polaire) au lieu d'une conique. Connaissant uniquement le point principal \mathbf{O}_i , il est donc possible d'utiliser le modèle linéaire (modèle sténopé) pour la projection des droites 3D au lieu du modèle unifié non-linéaire.

2.2 Reconstruction Euclidienne

La contrainte épipolaire est une propriété de la projection centrale jouant un rôle primordial dans la reconstruction euclidienne d'un déplacement d'une caméra entre deux prises de vue. Cette contrainte géométrique peut être appliquée aux capteurs catadioptriques centraux. Svoboda et Padjla sont les premiers à avoir étudié la géométrie épipolaire pour les capteurs catadioptriques centraux à miroirs hyperbolique [SPH98a] et parabolique [PSH01]. Une formulation générale de cette géométrie pour tout capteur à projection centrale est présentée dans [SP02a]. En particulier, il a été noté que les capteurs à point central unique calibrés admettent une géométrie épipolaire similaire à celles des caméras conventionnelles [SPH98a]. Ce résultat est important car les algorithmes d'estimation des géométries projective et euclidienne conçus pour les caméras conventionnelles, peuvent être assez directement utilisés pour des capteurs catadioptriques centraux. Comme nous l'avons évoqué précédemment, seule l'estimation de la matrice essentielle a été abordée dans la littérature. De notre côté, nous avons proposé des méthodes d'estimation de la transformation homographique entre deux images d'un ensemble de points ou de droites afin d'alimenter les schémas de commande hybride qui seront présentés dans le Chapitre 3.

2.2.1 Relations fondamentales

Considérons maintenant deux positions du capteur décrites par \mathcal{F}_m et \mathcal{F}_m^* (voir figure 2.3). Ces deux positions sont liées par une rotation \mathbf{R} et une translation \mathbf{t} . Un point \mathcal{X} de la scène 3D, avec comme coordonnées $\mathbf{X} = [X \ Y \ Z]^\top$ et $\mathbf{X}^* = [X^* \ Y^* \ Z^*]^\top$ dans \mathcal{F}_m et \mathcal{F}_m^* , se projette sur la sphère unité pour les deux positions de la caméra en des points de coordonnées $\mathbf{X}_s = \rho^{-1}\mathbf{X}$ et $\mathbf{X}_s^* = \rho^{*-1}\mathbf{X}^*$. Définissons le plan épipolaire par le point de la scène \mathcal{X} et les centres de la sphère unitaire. Clairement, les projections de \mathcal{X} sur la sphère unité dans les deux positions appartiennent également au plan épipolaire. Cette contrainte de co-planarité se traduit par la relation :

$$\mathbf{X}_s \mathbf{R} (\mathbf{t} \times \mathbf{X}_s^{*\top}) = \mathbf{X}_s \mathbf{E} \mathbf{X}_s^{*\top} = 0 \quad (2.13)$$

où $\mathbf{E} = \mathbf{R}[\mathbf{t}]_\times$ représente la matrice essentielle. De nombreux algorithmes permettent d'estimer la matrice essentielle à partir d'un ensemble de correspondances entre les images. Quand cela sera nécessaire nous utiliserons l'algorithme proposé dans [Nis04].

Définissons maintenant un plan Π de la scène par ses coordonnées $\pi^* = [\mathbf{n}^{*\top} \ d^*]^\top$ dans le repère \mathcal{F}_m et notons $d(\mathcal{X}, \Pi)$ la distance du point \mathcal{X} au plan Π . On vérifie alors aisément la relation suivante :

$$\rho \mathbf{X}_s = \rho^* \mathbf{H} \mathbf{X}_s^* + \alpha \mathbf{t} \quad (2.14)$$

Avec $\mathbf{H} = \mathbf{R} + \frac{\mathbf{t}}{d^*} \mathbf{n}^{*\top}$ et $\alpha = -\frac{d(\mathcal{X}, \pi)}{d^*}$. Si le point de la scène 3D \mathcal{X} appartient au plan de

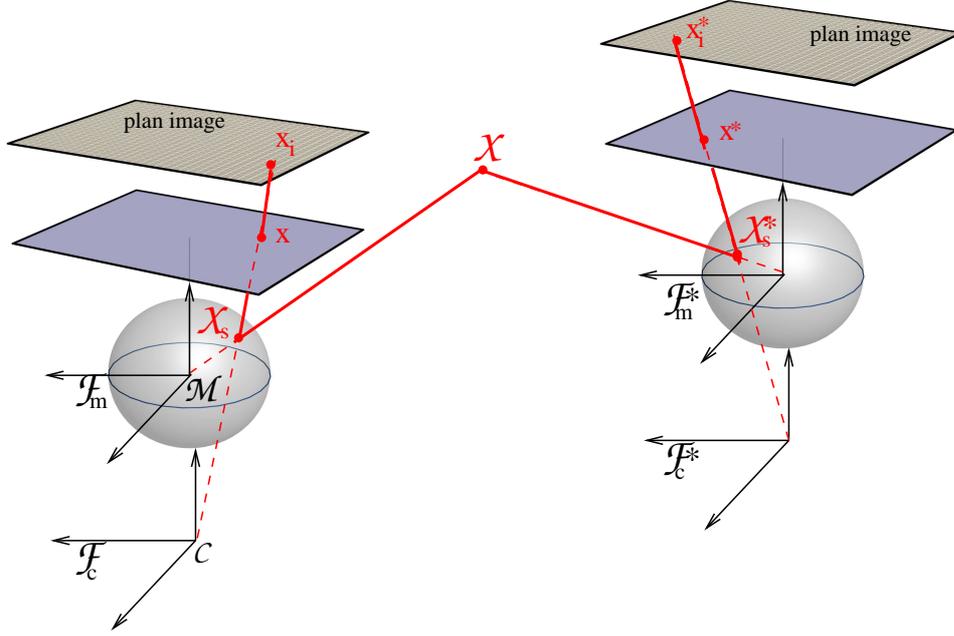


FIG. 2.3 – Géométrie de deux vues.

référence Π (*i.e* $\alpha = 0$) alors on a :

$$\mathbf{X}_s \propto \mathbf{H}\mathbf{X}_s^* \quad (2.15)$$

\mathbf{H} est une matrice 3×3 non singulière appelée matrice d'homographie euclidienne relative au plan de référence Π . En effet, elle réalise une transformation linéaire de \mathcal{S}^2 dans \mathcal{S}^2 des points images relatifs à la projection catadioptrique des points de la scène appartenant au plan de référence, sur la sphère unitaire. Lorsqu'une caméra perspective est considérée, on obtient une relation homographique identique mais réalisant une transformation de \mathcal{P}^2 dans \mathcal{P}^2 des points images relatifs à la projection perspective des points de la scène appartenant au plan de référence.

Supposons maintenant que la caméra observe une droite \mathcal{L} . Ses projections dans les deux images sont données par Ω_i et Ω_i^* . Soit \mathbf{l}_i et \mathbf{l}_i^* les droites polaires associées respectivement aux points principaux \mathbf{O}_i dans l'image courante \mathcal{I} et \mathbf{O}_i^* dans l'image désirée \mathcal{I}^* par rapport aux coniques Ω_i et Ω_i^* . En rappelant que ces droites polaires peuvent être interprétées comme les images de la droite \mathcal{L} obtenues avec une caméra perspective virtuelle, on peut directement utiliser un résultat connu pour les caméras pinholes :

$$\mathbf{l}_i \propto \mathbf{G}^{-\top} \mathbf{l}_i^* \quad (2.16)$$

où

$$\mathbf{G} = \mathbf{K}\mathbf{H}\mathbf{K}^{-1} = \mathbf{K}\left(\mathbf{R} + \frac{\mathbf{t}}{d^*}\mathbf{n}^{*\top}\right)\mathbf{K}^{-1} \quad (2.17)$$

est la matrice de colinéation (matrice d'homographie exprimée en pixel).

2.2.2 Estimation des homographies

Estimation à partir de points. Si l'on dispose de points coplanaires, une méthode purement linéaire peut être utilisée pour estimer la matrice d'homographie \mathbf{H} à un facteur d'échelle près. En écrivant (2.15) sous la forme d'une équation homogène linéaire :

$$\mathbf{X}_s \times \mathbf{H}\mathbf{X}_s^* = 0 \quad (2.18)$$

On obtient alors un système de $3n$ équations linéaires avec $n + 8$ inconnues. On peut donc résoudre ce système si l'on dispose d'au moins 4 points coplanaires. Si seuls 3 points appartenant au plan de référence sont disponibles, alors au moins 5 points supplémentaires sont nécessaires pour estimer la matrice d'homographie \mathbf{H} en utilisant par exemple l'algorithme linéaire proposé par Malis *et al.* dans [MC00a]. A partir de la matrice d'homographie $\hat{\mathbf{H}} = k\mathbf{H}$ estimée, il est possible de retrouver \mathbf{R} , \mathbf{t}_{d^*} et \mathbf{n}^* en utilisant les algorithmes proposés dans [FL88] ou [ZDFL95]. Dans le cas général, il existe deux solutions différentes. Cependant, cette indétermination peut être éliminée si l'on dispose d'informations additionnelles (par exemple le vecteur normal \mathbf{n}^* au plan de référence) lorsque la cible est plane, ou en considérant un second plan de référence et en choisissant la solution commune lorsque la cible n'est pas plane [FL88]. En plus du mouvement partiel de la caméra (catadioptrique ou perspective) entre les deux prises de vue, la matrice d'homographie \mathbf{H} peut nous informer sur la structure de la scène, notamment sur le rapport σ entre les deux normes ρ et ρ^* du point \mathcal{X} :

$$\sigma = \frac{\rho}{\rho^*} = \det(\mathbf{H}) \frac{\mathbf{n}^{*\top} \mathbf{X}_s^*}{\mathbf{n}^{*\top} \mathbf{R}^\top \mathbf{X}_s} \quad (2.19)$$

Naturellement, les paramètres présentés ci-dessus peuvent être reconstruits à partir de la matrice essentielle estimée. Néanmoins, dans le cadre de l'asservissement visuel nous préférons utiliser la matrice d'homographie, la matrice essentielle étant singulière à la convergence. Nous donnons dans le paragraphe suivant un exemple de résultat de reconstruction partielle tiré du document synthétisant les travaux de thèse d'Hicham Hadj-Abdelkader [HA06b]. Nous renvoyons le lecteur à ce même document pour des résultats plus complets.

Exemple de résultats de reconstruction partielle à partir de points. Le capteur considéré dans ces simulations est une caméra para-catadioptrique (miroir parabolique combiné avec une caméra orthographique). Cela implique que le paramètre ξ du modèle de projection est égal à 1. Les paramètres exacts (dans la matrice \mathbf{K}) du capteur para-catadioptrique sont $\alpha_u = \alpha_v = 161$, $\alpha_{uv} = 0$, et les coordonnées du point principal dans le repère image valent 300 pixels sur l'axe des abscisses et 269 pixels sur l'axe des ordonnées. Dans le cas d'une expérimentation réelle, les paramètres d'étalonnage peuvent être entachés d'erreurs. Pour se rapprocher de ces conditions, nous avons utilisé des paramètres d'étalonnage erronés ($\hat{\mathbf{K}}$) en ajoutant $\pm 10\%$ aux valeurs des focales réelles et ± 2 pixels aux coordonnées exactes du point principal. En outre, pour le paramètre ξ du miroir parabolique, nous avons utilisé $\hat{\xi} = 0.98$ au lieu de 1. Nous avons considéré une cible plane composée de 14 points pour l'estimation de la matrice d'homographie à partir de l'équation homogène linéaire (2.18). Un bruit aléatoire à distribution uniforme d'une variance de 1 pixel a été ajouté aux coordonnées des points dans l'image. Le déplacement considéré est composé d'une

translation $\mathbf{t} = [35 \ 28 \ 29]^\top$ cm et d'une rotation $\mathbf{u}_\theta = [-22 \ 54 \ -165]^\top$ degrés. Nous avons répété le processus d'estimation de la matrice d'homographie (et donc de calcul de déplacement) 100 fois. Après l'estimation de la rotation $\widehat{\mathbf{R}}$ et de la translation $\widehat{\mathbf{t}}$ à un facteur d'échelle près à partir de la matrice d'homographie estimée $\widehat{\mathbf{H}}$, nous calculons les erreurs suivantes :

- l'erreur de rotation : l'angle α_r extrait de la rotation $\mathbf{R}\widehat{\mathbf{R}}^\top$.
- l'erreur de translation : l'angle α_t formé par les vecteurs de translation réelle $\mathbf{t}/\|\mathbf{t}\|$ et estimée $\widehat{\mathbf{t}}/\|\widehat{\mathbf{t}}\|$.

La Figure 2.4 donne l'erreur médiane, l'écart médian et l'erreur maximale de α_r et de α_t en fonction du nombre de points. De manière assez classique, on observe que les erreurs dues aux bruits de mesure et aux erreurs d'étalonnage diminuent lorsque le nombre de points augmente. A partir de 12 points, les erreurs maximales en rotation et sur la direction de la translation sont inférieures à 10 degrés. Les erreurs médianes sont inférieures à 4 degrés et les écarts médians sont d'environ 2 degrés. Dans le chapitre 3, nous verrons à travers des expérimentations réelles que ces résultats sont acceptables dans le contexte de l'asservissement visuel.

Estimation à partir de droites. L'approche classique pour estimer la matrice de colinéation \mathbf{G} est basée sur une estimation linéaire de ses éléments à partir de l'équation (2.16) valable pour chaque couple de droites polaires $(\mathbf{l}_{i_k}, \mathbf{l}_{i_k}^*)$. Si l'on dispose de n couples, il est donc possible d'obtenir un système linéaire de la forme :

$$\mathbf{C}_g \mathbf{g} = 0 \quad (2.20)$$

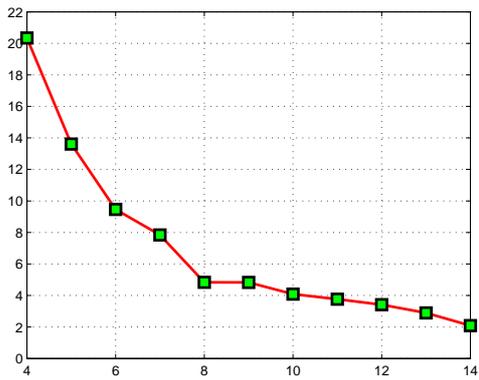
où \mathbf{C}_g est une matrice de dimension $3n \times 9$ contenant les coordonnées de droites polaires courantes et désirées et \mathbf{g} est un vecteur contenant les 9 coefficients de \mathbf{G} . Comme \mathbf{G} est définie à un facteur d'échelle près, un minimum de 4 couples de coordonnées de droites polaires correspondants aux 4 droites $3D$ coplanaires, sont nécessaires pour résoudre (2.20). Le problème d'estimation linéaire peut alors être écrit en présence de bruits de mesure comme suit :

$$\begin{aligned} \min_{\mathbf{g}} \|\mathbf{C}_g \mathbf{g}\| \\ \|\mathbf{g}\| = 1 \end{aligned} \quad (2.21)$$

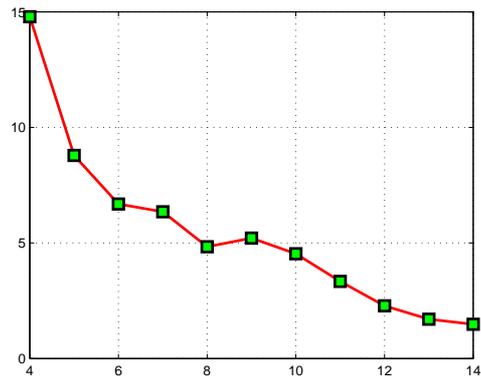
Cependant, cette minimisation au sens des moindres carrés est algébrique et n'a aucune signification physique. Afin de construire une fonction de coût significative physiquement, il est possible d'extraire deux points images $\mathbf{x}_{i_k}^j$ avec $j = \{1, 2\}$ pour chaque droite polaire \mathbf{l}_{i_k} , et de les utiliser ensuite pour estimer la fonction de coût suivante :

$$\mathcal{C}_1^2 = \sum_{k=1}^n \sum_{j=1}^2 \mathbf{x}_{i_k}^{j\top} \mathbf{G}^{-\top} \mathbf{l}_{i_k}^* \quad (2.22)$$

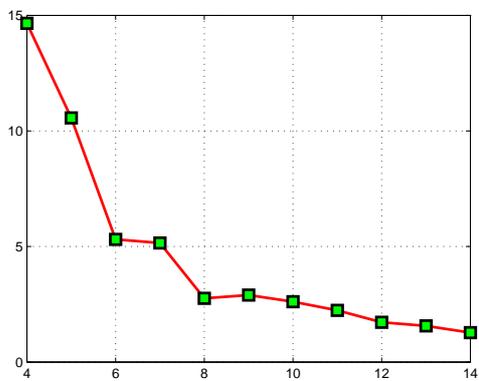
Évidemment, cette erreur est une somme de termes algébriques. Elle est toutefois basée sur une erreur avec une signification physique pourvu que les points $\mathbf{x}_{i_k}^j$ soient dans l'image. Au moins deux points sont nécessaires pour contraindre la droite correspondante. Ces deux points peuvent être choisis comme l'intersection de la droite polaire associée \mathbf{l}_{i_k} avec le cercle limitant l'image catadioptrique. Cependant, cette solution dégénère si la droite polaire se trouve proche de la tangente ou à l'extérieur du cercle limitant l'image. Afin d'éviter ces cas dégénérés, il est possible



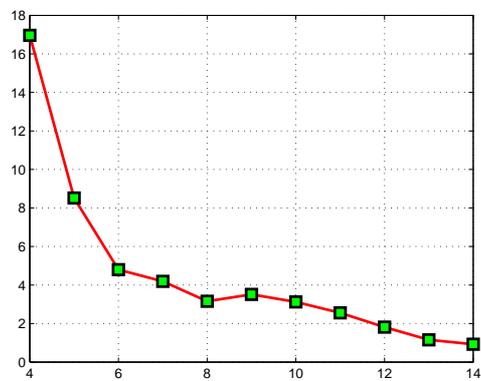
(a) erreur médiane sur α_r



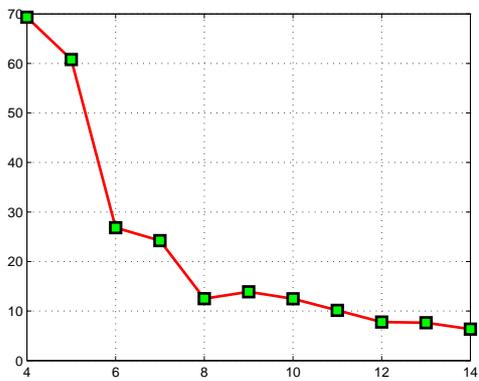
(d) erreur médiane sur α_t



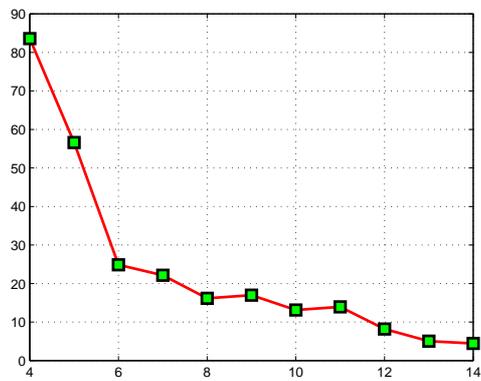
(b) écart médian sur α_r



(e) écart médian sur α_t



(c) erreur maximale sur α_r



(f) erreur maximale sur α_t

FIG. 2.4 – Erreur de reconstruction en fonction du nombre de points considérés.

de choisir les deux points comme l'intersection des deux tangentes au cercle limitant l'image catadioptrique et perpendiculaires à la droite polaire \mathbf{l}_k . Afin d'accroître encore la robustesse du calcul de \mathbf{G} par rapport aux bruits de mesure, il est souhaitable de minimiser une fonction de coût qui ait une interprétation géométrique. Cette erreur géométrique peut être obtenue en mesurant, par exemple, les distances entre chaque point $\mathbf{x}_{i_k}^j$ de la droite polaire \mathbf{l}_k et la droite polaire correspondante $\mathbf{G}^{-\top} \mathbf{l}_k^*$:

$$\mathcal{C}_2^2 = \sum_{k=1}^n \sum_{j=1}^2 d^2(\mathbf{x}_{i_k}^j, \mathbf{G}^{-\top} \mathbf{l}_k^*) \quad (2.23)$$

où la fonction $d()$ est la distance euclidienne. La fonction de coût est alors non linéaire en les éléments de la matrice de colinéation à estimer. Par ailleurs, elle est symétrique si les points $\mathbf{x}_{i_k}^{*j}$ appartenant aux droites polaires \mathbf{l}_k^* sont également utilisés :

$$\mathcal{C}_3^2 = \sum_{k=1}^n \left(\sum_{j=1}^2 d^2(\mathbf{x}_{i_k}^j, \mathbf{G}^{-\top} \mathbf{l}_k^*) + \sum_{j=1}^2 d^2(\mathbf{x}_{i_k}^{*j}, \mathbf{G}^{\top} \mathbf{l}_k) \right) \quad (2.24)$$

Dans le cadre d'une commande référencée vision, nous nous limiterons à la première solution linéaire afin de garantir un temps de calcul acceptable dans une boucle d'asservissement. En outre, les résultats obtenus au cours des travaux de thèse d'Hicham Hadj-Abdelkader ont montrés que l'utilisation d'une erreur algébrique était amplement suffisante dans notre contexte [HA06b]. Nous donnons un peu plus loin un exemple de résultats tirés de ces travaux de thèse. Notons, avant cela, que le calcul de la matrice de colinéation nécessite uniquement les coordonnées du point principal dans l'image exprimées en pixel comme paramètre intrinsèque (afin d'obtenir les droites polaires). Celles-ci peuvent aisément être estimées par exemple en utilisant les projections dans l'image de trois droites 3D. Le point principal est alors défini par le point d'intersection des droites liant les points d'intersections des coniques dans l'image associées aux trois droites 3D [BA02a]. En outre, la connaissance des autres paramètres intrinsèques de la caméra permet l'estimation de la matrice d'homographie (dans l'image normalisée) via la relation $\mathbf{H} = \mathbf{K}^{-1} \mathbf{G} \mathbf{K}$. Comme pour les points, à partir de la matrice d'homographie estimée, il est alors possible de retrouver le déplacement du capteur ainsi que certaines grandeurs 3D comme le rapport entre les profondeurs h et h^* :

$$r_h = \frac{h}{h^*} = \det(\mathbf{H}) \frac{\|\mathbf{K}^{\top} \mathbf{l}_i\|}{\|\mathbf{K}^{\top} \mathbf{l}_i^*\|} \frac{\|\mathbf{n}^* \times \mathbf{K}^{\top} \mathbf{l}_i^*\|}{\|\mathbf{R} \mathbf{n}^* \times \mathbf{K}^{\top} \mathbf{l}_i\|} \quad (2.25)$$

Exemples de résultats de reconstructions partielles à partir de droites. Nous considérons maintenant un ensemble de 11 droites 3D appartenant à une cible plane. Nous estimons les droites polaires associées au point principal en utilisant les coniques dans l'image du capteur para-catadioptrique. Deux méthodes d'estimation de la matrice de colinéation $\widehat{\mathbf{G}}$ (la matrice d'homographie $\widehat{\mathbf{H}}$ après la transformation plan-à-plan par la matrice d'étalonnage) sont utilisées dans ces simulations. La première méthode est linéaire et basée sur une minimisation aux moindres carrés d'une erreur algébrique (2.21). Tandis que la seconde méthode est basée sur la minimisation non linéaire d'une erreur géométrique (2.23) initialisée en utilisant la première

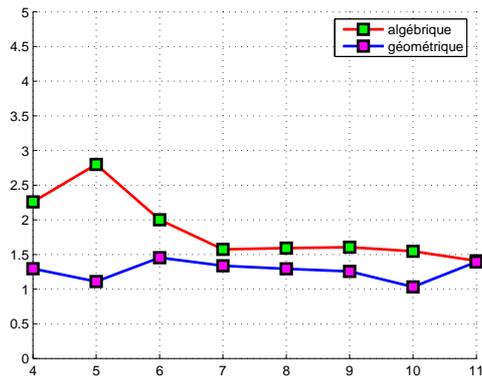
méthode. Notons que l'image catadioptrique d'une droite 3D lorsque le capteur est une caméra para-catadioptrique et que $\alpha_u = \alpha_v$, est un cercle. Le bruit de mesure dans l'image est alors introduit de la manière suivante :

- le cercle Ω_i issu de la projection para-catadioptrique d'une droite 3D sur l'image catadioptrique est échantillonné avec un pas que nous avons fixé à 10 pixels, sur un angle choisi égal à $\pi/4$ radians par rapport au centre du cercle Ω_i ;
- un bruit aléatoire à distribution uniforme de variance 1 pixel est ajouté à chaque point obtenu après l'échantillonnage. Le bruit est ajouté suivant la direction de la normale à la tangente du cercle au point considéré ;
- un lissage 3D robuste sur les points bruités en utilisant les paramètres d'étalonnage erronés $\hat{\mathbf{K}}$ et le paramètre $\hat{\xi}$ est réalisé. Le résultat de celui-ci est un cercle $\hat{\Omega}_i$;
- le point principal est obtenu en utilisant la projection dans l'image de trois droites 3D [BA02a] ;
- la droite polaire associée au point principal par rapport au cercle $\hat{\Omega}_i$ est calculée en utilisant le point principal estimé.

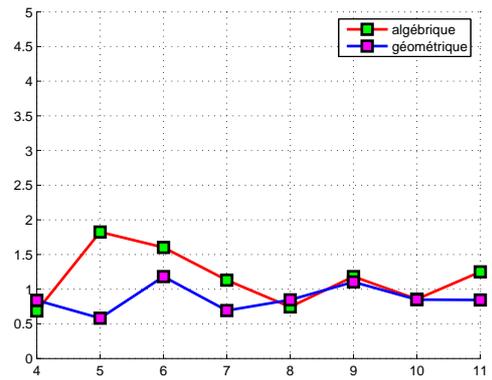
Quant aux erreurs sur les paramètres intrinsèques, elles sont simulées de la même manière que dans l'exemple de résultats précédent relatif à un ensemble de points. Les résultats présentés par la Figure 2.5 ont été obtenus pour un déplacement composé d'une translation $\mathbf{t} = [26 \ 30 \ 12]^\top$ cm et d'une rotation $\mathbf{u}_\theta = [16 \ -27 \ 25]^\top$ degrés. Ces résultats montrent la robustesse vis-à-vis des erreurs considérées. Les erreurs maximales en rotation et sur la direction de la translation sont inférieures à 5 degrés. Nous constatons que les erreurs médianes en rotation et sur la direction de la translation obtenues en utilisant la méthode géométrique sont plus faibles que lorsque la méthode linéaire est utilisée. La différence entre les deux méthodes devient négligeable à partir de 6 droites. Nous constatons également un saut des erreurs vers une valeur plus importante avec les deux méthodes en passant de 10 à 11 droites alors que l'accroissement du nombre de droites utilisées pour le calcul de l'homographie, devrait réduire encore les erreurs dues aux bruits d'étalonnage et de mesure. Nous supposons que cela est dû à une instabilité numérique. Comme le montrent les figures, les résultats obtenus par la méthode linéaire sont très proches de ceux obtenus avec les méthodes non linéaires. Par conséquent, nous utiliserons dans la suite plutôt la méthode linéaire, plus adaptée à des traitements à la cadence vidéo.

2.3 Modèle unifié et caméra fish-eye

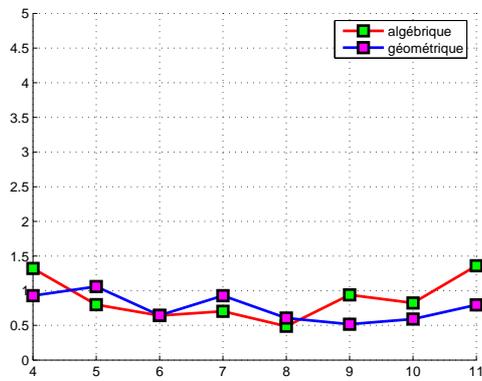
Le modèle précédent a été développé pour les capteurs catadioptriques centraux (même si nous avons vu qu'il intègre assez naturellement les caméras pinholes). Comme nous l'avons déjà mentionné la présence du miroir rend le capteur encombrant, fragile et rend une partie importante du centre de l'image inexploitable. L'utilisation d'une caméra fish-eye permet de résoudre en partie ces difficultés. Pour obtenir un tel système optique, une lentille optique divergente avant et une lentille convergente arrière sont combinées. Une caméra munie d'un objectif fish-eye n'est pas à point central unique. Cependant, la grande majorité des modèles font l'hypothèse d'un centre de projection unique avec un centre de distorsion confondu avec le centre de l'image. Les distorsions sont généralement décomposées en plusieurs composantes. Assez classiquement, nous ne nous intéresserons ici qu'aux distorsions radiales. Nous supposerons donc que les autres



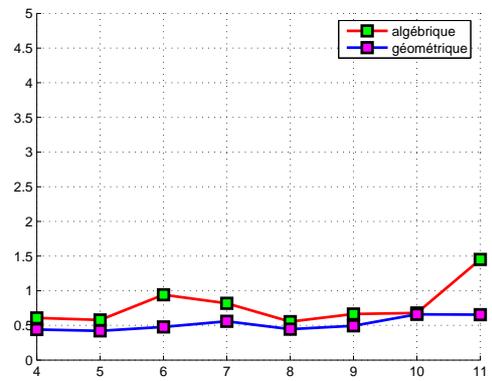
(a) erreur médiane sur α_r



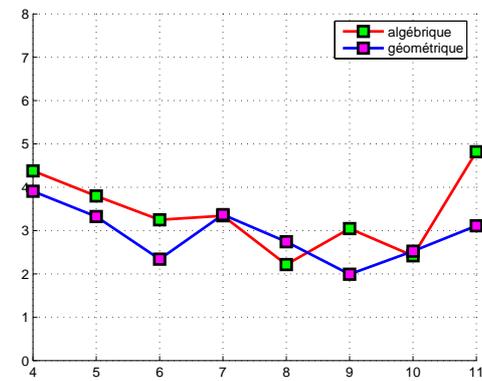
(d) erreur médiane sur α_t



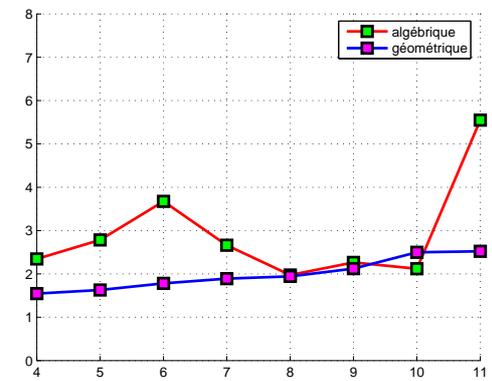
(b) écart médian sur α_r



(e) écart médian sur α_t



(c) erreur maximale sur α_r



(f) erreur maximale sur α_t

FIG. 2.5 – Erreur de reconstruction pour un déplacement générique en fonction du nombre de droites considérées.

composantes (en particulier la composante tangentielle) sont négligeables. Dans le cadre de nos applications (asservissement visuel et navigation autonome), nous verrons que cette hypothèse est tout à fait raisonnable. De nombreux modèles sont utilisés pour les caméras avec distorsions radiales. On peut cependant les regrouper en trois catégories. Le premier groupe de modèles est basé sur la modélisation pinhole. Dans ce cas, les points 3D sont dans un premier temps perspectivement projetés sur le plan image. On applique ensuite à ces points une transformation modélisant les distorsions radiales. Cette transformation peut être par exemple une fonction polynomiale [HZ00a] ou rationnelle [Fit01], logarithmique [BL95] ou de type cosinus hyperbolique [PK02, KKK04]. La seconde catégorie de modèles fait intervenir une transformation entre la direction du rayon d'incidence et la distance entre le centre de l'image et le point projeté dans l'image. Cette transformation peut également être de type polynomial [XT97] ou faire intervenir des fonctions sinusoïdales comme dans [Fle94, SF95]. La dernière catégorie est basée sur le modèle unifié de projection centrale [GD00]. Elle s'est développée à partir du constat que les images obtenues avec des caméras à optiques fish-eye et catadioptriques ont de fortes similitudes [YH04]. Par exemple, une droite de l'espace se projette en une conique dans le plan image d'un capteur catadioptrique à centre de projection unique [BA02a] et dans le plan image de certaines caméras à optique fish-eye [SJA99]. Cette catégorie de modèles inclut par exemple les propositions décrites dans [MR07, Fit01, YH04, Bar06]. Dans la suite, nous allons voir que le modèle unifié peut en fait être interprété comme un modèle appartenant à la première ou à la seconde des catégories mentionnées précédemment.

2.3.1 Modèle unifié versus modèle classique de distorsion radiale

Les distorsions radiales dépendent uniquement de l'angle d'incidence du rayon lumineux. Comme nous l'avons précédemment évoqué, les modèles de projection les intégrant sont nombreux mais ils respectent tous les deux contraintes fondamentales suivantes (les indices f et p seront utilisés dans la suite pour désigner respectivement des grandeurs se rapportant à des caméras avec ou sans distorsion) :

Contrainte 1 *Le rayon arrivant le long de l'axe optique n'est pas déformé : $r_f(0) = 0$.*

Contrainte 2 *Le rayon r_f croît de façon monotone quand le point s'éloigne du centre de l'image.*

Pour les caméras avec distorsion, la distance entre le centre de l'image et le point image est différente de celle obtenue dans le cas sténopé :

$$r_f(\theta) \neq r_p(\theta) = f \tan \theta$$

Deux grandes familles de modèles permettent la description des caméras avec distorsion. La première repose sur le modèle sténopé. Le rayon r_f est obtenu en ajoutant des distorsions radiales au rayon r_p (transformation T_1) :

$$r_p \xrightarrow{T_1} r_f$$

La seconde famille de modèles est basée sur la définition d'une transformation T_2 entre l'angle d'incidence θ et la distance r_f :

$$\theta \xrightarrow{T_2} r_f$$

$$\begin{aligned}
 r_f^1(r_p) &= r_p L(r_p, n) \\
 r_f^2(r_p) &= \frac{r_p}{L(r_p, n)} \\
 r_f^3(r_p) &= \frac{r_p}{1 + k_1 r_p^2} \\
 r_f^4(r_p) &= r_p \frac{L_1(r_p, n_1)}{L_2(r_p, n_2)} \\
 r_f^5(r_p) &= s \log(1 + \lambda r_p) \\
 r_f^6(r_p) &= \frac{1}{\omega} \arctan \left(2r_p \tan \frac{\omega}{2} \right)
 \end{aligned}$$

TAB. 2.1 – Quelques exemples de transformation de type T_1 : dans [SA96, PWH97, HZ00a, AF01, EMF03], les distorsions radiales sont décrites par un modèle polynomial $r_f^1(r_p)$ où $L(r_p, n) = 1 + k_1 r_p^2 + k_2 r_p^4 + \dots + k_n r_p^{2n}$ (où $k_i, i = 1, 2, \dots, n$ sont les paramètres du modèle). Le modèle fractionnel $r_f^2(r_p)$ a été suggéré dans [Fit01] et a été utilisé avec un seul paramètre (fonction $r_f^3(r_p)$) dans [Fit01]. Le modèle rationnel $r_f^4(r_p)$ a été proposé dans [LH05] et met en jeu $n_1 + n_2$ paramètres. La transformation logarithmique $r_f^5(r_p)$ avec comme paramètres un facteur d'échelle s et un gain λ contrôlant l'amplitude des distorsions a été proposée dans [BL95]. Le modèle FOV (Field of View) $r_f^6(r_p)$ où ω représente le champ de vue de la caméra a, quant à lui, été proposé dans [DF01].

Les Tableaux 2.1 et 2.2 donnent des exemples de transformation de type T_1 et T_2 .

En introduisant, les coordonnées \mathbf{x}_p (Eq. 2.1) dans le modèle de projection (2.5) puis en utilisant la relation (2.3), il est assez aisé de reformuler le modèle unifié sous la forme d'une transformation de type T_1 [CMEM07] :

$$r_f = r_f(r_p) = \frac{r_p}{1 + \xi \sqrt{\frac{r_p^2}{f^2} + 1}} \quad (2.26)$$

ou alors de type T_2 [CMEM07] :

$$r_f = r_f(\theta) = \frac{f \tan \theta}{1 + \xi \sqrt{\tan^2 \theta + 1}} \quad (2.27)$$

Dans les deux cas, on vérifie facilement que les contraintes 1 et 2 sont respectées. On remarque également que de manière classique, le modèle unifié présenté sous la forme T_1 n'est pas valide pour $\theta = \pi/2$ et :

$$\lim_{\substack{\theta \rightarrow \pi/2 \\ \theta < \pi/2}} r(\theta) = \frac{f}{\xi}$$

$$\begin{aligned}
r_f^1(\theta) &= f\theta \\
r_f^2(\theta) &= 2f \tan\left(\frac{\theta}{2}\right) \\
r_f^3(\theta) &= f \sin \theta \\
r_f^4(\theta) &= f \sin\left(\frac{\theta}{2}\right) \\
r_f^5(\theta) &= f(k_1\theta + k_2\theta^3 + \dots + k_n\theta^{2(n-1)+1}) \\
r_f^6(\theta) &= \alpha \sin(\beta\theta) \\
r_f^7(\theta) &= a \tan(\theta/b) + c \sin(\theta/d)
\end{aligned}$$

TAB. 2.2 – Quelques exemples de transformation de type T_2 : la transformation f-theta ou projection equi-angulaire $r_f^1(\theta)$ proposée dans [Kin89], permet seulement de modéliser des caméras ayant de faibles distorsions. La projection stéréographique $r_f^2(\theta)$ proposée dans [Fle94, SF95] permet de préserver les symétries locales. Les transformations sinusoïdales $r_f^3(\theta)$ et $r_f^4(\theta)$ sont décrites dans [Ray94] et [Smi92]. Des fonctions polynomiales du type $r_f^5(\theta)$ (utilisées par exemple dans [XT97, KB04, Sch05, SMS06]) permettent d’améliorer la précision en augmentant le nombre de paramètres. Le modèle $r_f^6(\theta)$ avec deux paramètres (α et β) a été proposé dans [KB00]. Enfin le modèle $r_f^7(\theta)$ proposé dans [BP02] combine une projection stéréographique à deux paramètres (a, b) avec une projection de l’angle equisolide à deux paramètres (c, d)

Les relations (2.26) et (2.27) d’une interprétation assez simple au regard du modèle unifié, permettent donc une description conventionnelle du phénomène de distorsion radiale. L’intérêt majeur est ici d’exploiter un même modèle pour les caméras conventionnelles, catadioptriques ou fish-eye. Par exemple, cela nous permettra d’utiliser directement les méthodes de reconstruction euclidienne présentées dans la partie 2.2, dans les contextes de l’asservissement visuel et de la navigation autonome pour les différents types de caméra.

Bien sûr, les relations (2.26) et (2.27) ne permettent pas de conclure directement sur la capacité du modèle unifié à modéliser efficacement les caméras fish-eye (par ailleurs, ceci est vrai pour l’ensemble des modèles proposés dans la littérature). De manière expérimentale, nous pouvons cependant nous faire une idée sur ce point. Dans cette optique, nous avons présenté dans [CMEM07] un ensemble de résultats sur des données simulées et réelles montrant que le modèle unifié était tout à fait suffisant pour nos applications. A titre d’exemple, la Figure 2.6 présente un résultat d’estimation des paramètres de la projection de droites dans l’image d’une caméra fish-eye en utilisant le modèle unifié et deux points cliqués à la souris sur chaque droite (c’est le minimum). On constate (ici visuellement) que le résultat est correct. Un autre exemple de résultat, plus quantitatif celui-là est donné par la Figure 2.7. L’objectif est ici d’évaluer les performances du modèle unifié pour une caméra fish-eye dans une application de reconstruction

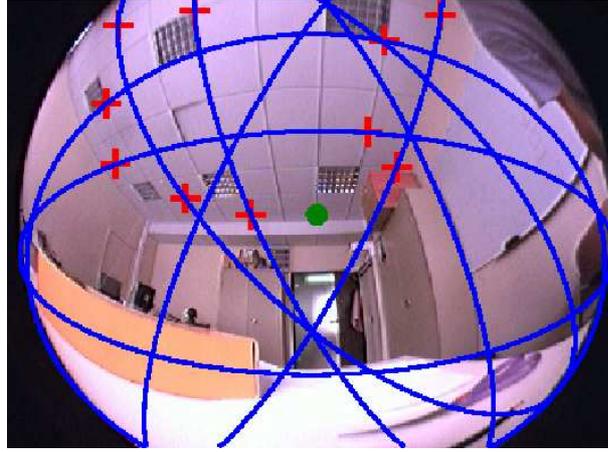


FIG. 2.6 – Estimation des paramètres de la projection de droites dans l’image d’une caméra fish-eye en utilisant le modèle unifié et deux points cliqués à la souris sur chaque droite :

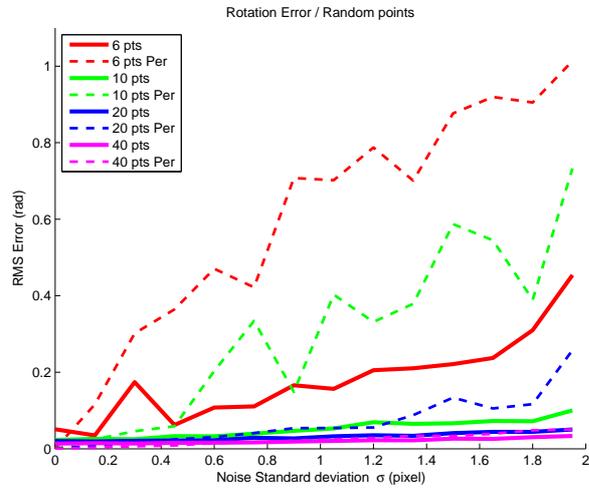
euclidienne à partir d’images synthétiques. La caméra est placée dans un environnement de travail virtuel et observe un plan π contenant N points placés de manière aléatoire. Les images synthétiques de taille 640×480 sont générées en projetant ces points dans le plan image en utilisant le modèle de projection $r_f^1(r_p)$ avec $r_f^1(r_p) = r_p(1 - 0.238.r_p^2 + 0.040.r_p^4 + 0.002.r_p^6 - 0.001.r_p^8)$ (voir Tab. 2.1). Ces paramètres correspondent à une caméra réelle (caméra Marlin F131B avec un objectif fish-eye ORIFL190-3) par ailleurs également calibrée pour le modèle unifié. Un bruit bidimensionnel Gaussien de moyenne nulle et d’écart-type σ est ensuite ajouté à chaque point dans l’image. Pour chaque expérience, deux positions de la caméra sont générées aléatoirement avec un déplacement composé d’une rotation \mathbf{R} et d’une translation \mathbf{t} . La matrice d’homographie \mathbf{H}_π relative au plan π est alors estimée en utilisant le modèle unifié. Après l’estimation de la rotation $\hat{\mathbf{R}}$ et de la translation $\hat{\mathbf{t}}$ à un facteur d’échelle extrait de la matrice d’homographie estimée $\hat{\mathbf{H}}_\pi$, nous calculons une nouvelle fois les erreurs suivantes :

- l’erreur de rotation : l’angle α_r extrait de la rotation $\mathbf{R}\hat{\mathbf{R}}^\top$.
- l’erreur de translation : l’angle α_t formé par les vecteurs de translation réelle $\mathbf{t}/\|\mathbf{t}\|$ et estimée $\hat{\mathbf{t}}/\|\hat{\mathbf{t}}\|$.

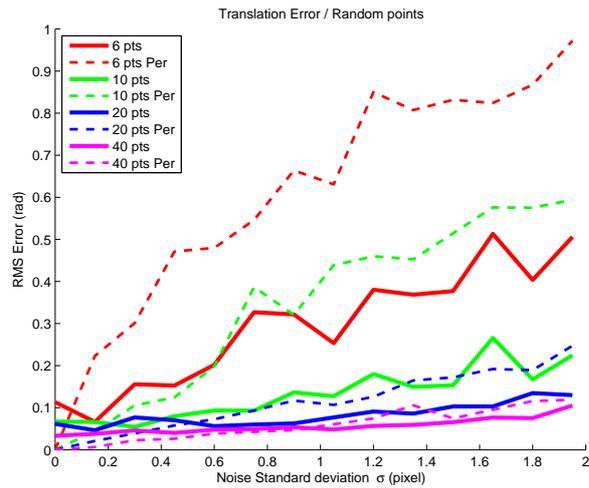
Afin de comparer avec le cas bien connu d’une estimation avec une caméra perspective, ces grandeurs sont également calculées avec le modèle de projection pinhole (utilisé à la fois pour générer les images synthétiques et estimer la matrice d’homographie). Les erreurs RMS (Root Mean Square) obtenues avec 200 réalisations de chaque expérience sont représentées sur la figure 2.7 pour un nombre de points et un écart-type σ variants.

On observe de manière classique que l’erreur RMS augmente quand l’écart-type du bruit augmente et qu’elle diminue avec le nombre de points. On remarque également que même quand l’écart type du bruit est nul les résultats obtenus avec le modèle unifié ne sont pas parfaits. En effet, dans le cas de la caméra fish-eye les images ont été générées avec un modèle polynomial de distorsion alors que la matrice d’homographie a été estimée en utilisant le modèle unifié (dans le cas d’une caméra conventionnelle les résultats sont bien sûr parfaits puisque ces deux opé-

rations sont réalisées en utilisant le modèle sténopé). On notera également que pour un faible nombre de points, dès que l'écart-type du bruit augmente, les meilleurs résultats sont obtenus avec la caméra fish-eye et le modèle unifié. Quoi qu'il en soit, les résultats de reconstruction euclidienne pour une caméra fish-eye sous le modèle de projection unifié sont toujours similaires (voir plus satisfaisants) que les résultats pour une caméra perspective sous le modèle pinhole. Comme évoqué précédemment, des résultats complémentaires validant le modèle unifié pour les caméras fish-eye sont donnés dans [CMEM07]. En outre, dans les chapitres 3 et 4 nous utiliserons à de nombreuses reprises des caméras fish-eye dans le contexte de l'asservissement visuel et de la navigation autonome de robots mobiles.



(a)



(b)

FIG. 2.7 – Reconstruction euclidienne avec des données synthétiques : erreurs RMS en radian (a) translation α_t et (b) rotation α_r en utilisant 6, 10, 20 et 40 points en correspondances en fonction de l'écart-type $\sigma = 0 \dots 2$ pixel(s). Les lignes continues donnent les résultats obtenus avec le modèle unifié alors que les lignes en pointillées donnent les résultats obtenus avec le modèle sténopé.

Chapitre 3

Asservissement visuel

Ce chapitre est consacré à nos contributions dans le domaine de l'asservissement visuel. Il aborde tout d'abord la question du choix de l'inverse généralisé de la matrice d'interaction lorsque celle-ci n'est pas inversible. Il présente à cette occasion deux alternatives à l'inverse de Moore-Penrose. Le second point abordé concerne l'asservissement visuel lorsque la caméra obéit au modèle de projection unifié décrit dans le chapitre précédent. Il présente en particulier des résultats de stabilité de l'asservissement visuel 2D basé point en présence de paramètres 3D erronés ainsi que des schémas de commande 2D et 2D1/2 dans divers contextes applicatifs (robots mobiles, manipulateurs séries ou parallèles).

3.1 Introduction et préliminaires

On utilise communément le terme d'asservissement visuel pour désigner une classe particulière de la commande référencée capteurs, dans laquelle le capteur utilisé est composé d'un ou plusieurs capteurs visuels. Celui-ci est le plus souvent une caméra, mais peut également être tout type de dispositif imageur (comme par exemple un dispositif à ultrason dans [MKC09]). En ce qui concerne nos travaux, l'imageur utilisé est toujours une caméra conventionnelle ou omnidirectionnelle. La commande résultante du traitement des données acquises par le capteur est en général intégrée dans un schéma de commande en boucle fermée. Les premiers travaux intégrant la vision artificielle dans un tel schéma remontent aux années 70. On notera en particulier, les travaux décrits dans [BR73, SI73] pour la réalisation de tâches d'assemblage. Ce n'est cependant qu'à partir des années 80 que la recherche dans ce domaine est devenue très active. Ces recherches sont depuis constamment alimentées d'une part par des aspects pratiques comme l'évolution rapide des puissances de calcul, les nouvelles technologies et concepts de capteurs, ou encore l'évolution des applications robotiques et d'autre part, par des problèmes théoriques

transversaux de fond comme les analyses de robustesse, de stabilité et de couplage des lois de commande. Dans la suite, nous décrivons assez brièvement les concepts à la base des techniques d'asservissement visuel. Le lecteur intéressé par plus de détails pourra se référer aux tutoriaux [HHC96, CH06, CH07].

Notons tout d'abord que la caméra peut être déportée ou embarquée. Si la caméra est déportée alors elle n'a pas de liaison mécanique avec le robot asservi et est positionnée afin d'observer l'organe terminal du robot [FCM00], la configuration articulaire du robot [RH00] ou à la fois le robot et les objets d'intérêt situés dans l'espace du travail [HDE98]. Elle peut être dans ce cas fixe ou actionnée [FCM00], [MCSP01]. Si la caméra est embarquée alors elle est rigidement liée au robot (par exemple, fixée à l'organe terminal dans le cas d'un robot manipulateur) afin d'observer l'espace de travail du robot. Pour des raisons pratiques, certaines des expérimentations décrites dans ce document ont été réalisées avec une caméra déportée. Cependant, dans ces applications le passage embarqué-déporté ne pose pas de problème particulier. *La suite de nos développements concernera donc la configuration embarquée.*

Les schémas de commande par vision peuvent être classifiés, de manière assez standard, en utilisant deux critères : 1) l'utilisation ou non d'une boucle interne au niveau des actionneurs du robot, 2) l'espace de commande du robot. Lorsque le système ne possède pas de boucle interne, le système de vision fournit directement une estimation de l'état du robot aux articulations, se substituant à ses contrôleurs. La stabilisation est uniquement réalisée en utilisant le retour sur le système de vision [SW83]. Dans ce cas de figure, le système doit fournir une estimation de l'état à une cadence élevée. En raison de l'influence importante des perturbations sur l'estimation temps réel de l'état [BHS94], ce type de schéma a été peu utilisé. Par exemple, les travaux de Hashimoto dans [HEK96] portent sur la mise en œuvre d'une technique de commande optimale de type LQ. Weiss exploite cette structure pour contrôler un robot plan à trois degrés de liberté dont les articulations sont observées par une caméra dans [WSN87]. On retrouvera le même type de structure dans [WWR93] pour un robot à deux axes. Des travaux dans ce sens ont également porté sur la commande de robot à six degrés de liberté [GdM00]. Lorsque le schéma de commande utilisé est hiérarchique (utilisation d'une boucle interne), le système de vision fournit une consigne au contrôleur du robot qui calcule les déplacements à réaliser dans l'espace articulaire par l'intermédiaire d'une boucle fermée interne qui stabilise le système. En utilisant une telle structure, la cadence du processus peut être amenée au temps réel vidéo (généralement de 40 ms). Le processus de vision et le système robotique sont alors découplés, ce qui permet une plus grande portabilité et simplicité des processus de commande par vision. On pourra se reporter à [CG96] pour une étude sur les performances dynamiques de cette structure. *Les schémas de commande proposés dans ce chapitre supposent une telle structure.*

Le second point caractérisant les schémas de commande concerne le type de mesures utilisé pour définir le signal d'erreur à réguler. Celui-ci peut-être 2D (AV2D : l'erreur est définie dans l'image), 3D (AV3D : l'erreur est définie dans l'espace de travail du robot) ou hybride (AVH : l'erreur combine des informations 2D et 3D). Dans toutes ces configurations, l'asservissement visuel consiste à élaborer une loi de commande afin qu'un ensemble de k mesures décrites par le vecteur \mathbf{s} atteignent leurs valeurs désirées \mathbf{s}^* en contrôlant n (où $n \leq 6$) degrés de liberté du robot. Le vecteur \mathbf{s}^* , contenant les k mesures à la position désirée du robot, peut être constant comme il peut être fonction du temps, on parle alors du suivi de la trajectoire $\mathbf{s}^*(t)$. *Dans ce document,*

on considérera que la consigne \mathbf{s}^* ne varie pas dans le temps. La loi de commande se ramène donc à réguler à zéro un vecteur d'erreur fonction de \mathbf{s} et \mathbf{s}^* adéquatement choisi dans l'espace de mesure. Dans ce contexte, de nombreux schémas de commande par vision ont été proposés. Certains d'entre eux intègrent la dynamique du robot [RB97, KH91, Kel96, CWH97, TS00] mais la majorité concernent des commandes cinématiques en considérant le robot comme un intégrateur parfait. On trouve par exemple la commande par mode glissant appliquée au suivi de trajectoire cartésienne de la caméra [ZMP00], la commande par retour d'état non linéaire [MG99, TS00], la commande par placement de pôles [PKK91], basées sur un contrôleur GPC [GdM00, Gan99], voire robuste H_∞ [Kha99] ou par retour d'état continu non stationnaire dans le cas des robots mobiles non holonomes [TRS98]. Les commandes de type proportionnel sont les plus couramment utilisées [MGK96, Cha90, HHC96, CH06, CH07]. Elles permettent d'assurer une décroissance exponentielle d'une erreur $\mathbf{e}(t) = \mathbf{s} - \mathbf{s}^*$ en imposant :

$$\dot{\mathbf{e}} = -\lambda \mathbf{e}$$

La loi de commande est alors obtenue en notant que la dérivée temporelle de la mesure $\dot{\mathbf{s}}$ est liée au mouvement relatif entre la caméra et la scène observée par la relation :

$$\dot{\mathbf{s}} = \mathbf{L}_s \boldsymbol{\tau} \quad (3.1)$$

et en combinant les deux dernières relations :

$$\boldsymbol{\tau} = -\lambda \mathbf{C} \mathbf{e} \quad (3.2)$$

où

- $\boldsymbol{\tau}$ est le torseur cinématique de la caméra exprimé dans le repère caméra considéré dans la suite comme variable de commande.
- \mathbf{L}_s est la matrice d'interaction de dimension $k \times n$ ($n \leq 6$), associée à \mathbf{s} et fonction de la position relative entre l'objet observé et la caméra. \mathbf{C} de dimension $n \times k$ est l'inverse de \mathbf{L}_s quand $n = k$ et si \mathbf{L}_s n'est pas singulière ou un inverse généralisé de \mathbf{L}_s si $n < k$. Nous reviendrons, sur le choix de l'inverse généralisé un peu plus loin dans ce document. Notons à titre d'exemple qu'un des choix possibles est l'utilisation de l'inverse de Moore-Penrose $\mathbf{C} = \mathbf{L}_s^+$ de \mathbf{L}_s . Dans ce cas la loi de commande (3.2) revient à minimiser la fonction de coût $g_s = \frac{1}{2} \|\mathbf{s} - \mathbf{s}^*\|$ en utilisant l'algorithme de Gauss-Newton.

En pratique, seules des estimations $\hat{\mathbf{C}}$ de \mathbf{C} sont disponibles. Le vecteur de commande est alors obtenu en utilisant la relation suivante :

$$\boldsymbol{\tau} = -\lambda \hat{\mathbf{C}} \mathbf{e} \quad (3.3)$$

En combinant (3.3) et (3.1), on obtient le comportement en boucle fermée :

$$\dot{\mathbf{e}} = -\lambda \mathbf{Q} \mathbf{e} \quad (3.4)$$

avec $\mathbf{Q} = \mathbf{L}_s \hat{\mathbf{C}}$. Il est alors assez aisé de montrer la condition suffisante de stabilité asymptotique globale du point d'équilibre du système différentiel (3.4) :

$$\mathbf{Q} > 0 \quad (3.5)$$

Remarque 1. Si la dimension k de \mathbf{s} est strictement supérieure au nombre de degrés de liberté n alors la condition (3.5) ne peut pas être vérifiée. En effet, on a dans ce cas $\dim[\text{Ker}(\mathbf{Q})] > 0$. Si \mathbf{Q} est semi-définie positive, alors la loi de commande (3.2) garantit l'annulation de la fonction de tâche $\mathbf{e}_t = \hat{\mathbf{C}}\mathbf{e}$ sans garantir l'annulation de l'erreur \mathbf{e} . L'ensemble des configurations telles que $(\mathbf{s} - \mathbf{s}^*) \in \text{Ker} \hat{\mathbf{C}}$ correspondent à des minima locaux. Dans ce cas, il est cependant possible d'exhiber une condition nécessaire et suffisante de stabilité asymptotique locale du point d'équilibre \mathbf{e}^* :

$$\mathbf{Q}_t = \hat{\mathbf{C}}\mathbf{L}_s > 0 \quad (3.6)$$

Cette dernière condition peut-être vérifiée quand la dimension k de \mathbf{s} est strictement supérieure au nombre de degrés de liberté n à commander si $\hat{\mathbf{C}}$ et \mathbf{L}_s sont de rang plein et si $\hat{\mathbf{C}}$ est bien choisie.

Remarque 2. La réalisation de la tâche $\mathbf{e}_t = \hat{\mathbf{C}}\mathbf{e}$ ne nécessite pas toujours l'utilisation de tous les degrés de liberté du robot. Afin de mieux utiliser les fonctionnalités de ce dernier, le formalisme de la fonction de tâche permet de gérer aisément les degrés de liberté non contraints par la tâche principale en définissant, par exemple une seconde tâche qui est pré-multipliée par l'opérateur de projection orthogonale sur le noyau de la jacobienne de la fonction de tâche \mathbf{e} afin que la seconde tâche (tâche secondaire) n'affecte pas la première. La tâche secondaire peut être utilisée pour la navigation autour d'un objet [Ber99], pour l'évitement des singularités du robot [MCR96], des butées articulaires [MCR96, CM00], des obstacles [Cad99] ou des obstacles et des occultations [MH98]. Dans [FCM00], une tâche secondaire a été utilisée pour contrôler la translation d'une caméra montée sur un robot manipulateur observé par une seconde caméra.

Revenons maintenant un peu plus en détails sur les trois types de schéma mentionnés plus haut.

AV3D. Dans ce type de schéma d'asservissement visuel, \mathbf{s} est composé uniquement d'informations 3D obtenues à partir d'une ou plusieurs images de la scène. Dans cette approche, le choix de 6 informations permettant de contrôler 6 degrés de liberté est assez intuitif. En outre, il est assez simple de choisir ces informations de manière à obtenir des matrices d'interaction avec de bonnes propriétés de découplage des translations et des rotations. En contrepartie, aucun contrôle dans l'image ne garantit que l'objet observé reste dans le champ de vision de la caméra. Le vecteur de mesure peut par exemple être choisi comme une paramétrisation de l'attitude de la caméra par rapport à l'objet [MG99, TMCG02, WHB96, DDLR97]. Dans ce cas, le processus d'estimation 3D nécessite la connaissance de la géométrie de la scène (par exemple un modèle CAO de l'objet observé). Une autre possibilité consiste à construire le vecteur d'observation \mathbf{s} comme une représentation de la position et de l'orientation de la caméra entre ses configurations courante et désirée. Dans ce cas, le modèle de l'objet cible n'est plus nécessaire, \mathbf{s} pouvant être obtenu en utilisant les techniques d'estimation partielle du mouvement [MC02a].

De manière très classique, on peut utiliser les coordonnées 3D d'un point (par exemple le centre du repère caméra courant exprimé dans le repère caméra cible) afin de contrôler les mouvements de translation du robot et le vecteur angle/axe représentant la rotation à effectuer pour amener la caméra de son orientation courante à l'orientation désirée afin de contrôler les mouvements de rotations. Dans ce cas, la matrice d'interaction est bloc triangulaire et inversible

sauf pour une rotation de 2π [MC02a]. En dehors de cette singularité, la matrice \mathbf{C} peut alors être choisie comme l'inverse de la matrice d'interaction.

AV2D. L'idée est ici de réaliser un asservissement dont la grandeur à réguler est issue de l'image. Le vecteur de mesure \mathbf{s} est alors uniquement composé d'informations extraites et exprimées directement dans l'espace image. Dans ce cas, aucune connaissance *a priori* du modèle de l'objet n'est nécessaire pour construire les vecteurs d'observation \mathbf{s} et de consigne \mathbf{s}^* . Différents types d'informations visuelles peuvent être utilisés. Le choix de ceux-ci est limité par deux conditions. La première porte sur l'existence d'algorithmes permettant de les extraire dans l'image. L'autre condition est qu'il soit possible d'estimer l'interaction associée à ces primitives. F. Chauvette a proposé une méthode générale pour exprimer de manière analytique l'interaction associée à des informations visuelles définies à partir de primitives géométriques paramétrables (points, droites, ellipses, etc.) [Cha90]. D'autres types de primitives visuelles comme la décomposition en série de Fourier de la signature polaire du contour de l'objet [CC00a] ou encore le signal de luminance [CM09] ont été étudiés. Un premier élément clé de l'AV2D est alors la modélisation de l'interaction entre le mouvement relatif de la caméra par rapport à la scène et la variation du vecteur de mesures. Nous reviendrons un peu plus loin sur ce point critique. Cependant notons d'ors et déjà que, même si l'erreur est complètement définie dans l'image, la modélisation de l'interaction fait intervenir des informations 3D. Un autre élément clé de l'AV2D (fortement lié au précédent) concerne le choix des informations visuelles 2D. En effet, la commande dans l'espace image garantit un comportement satisfaisant des primitives visuelles sans garantir le comportement spatial de la caméra. Si par exemple \mathbf{s} est construit à partir des coordonnées de trois points dans l'image (c'est bien sûr le nombre minimal pour contrôler six degrés de liberté) : 1) il existe quatre configurations spatiales de la caméra aboutissant à la même observation (laquelle choisir ?), 2) la matrice d'interaction s'avère singulière pour des configurations particulières de la caméra par rapport aux trois points (comment traverser ces singularités ?), 3) la trajectoire de la caméra peut s'avérer sous-optimale voire surprenante. Les singularités 1) et 2) peuvent être évitées en utilisant plus de points que nécessaire ($k > n$). Dans ce cas \mathbf{C} est prise comme un inverse généralisé de la matrice d'interaction. Se posent alors les questions suivantes : l'inverse de Moore-Penrose est-il le meilleur choix ? comment ce choix influe-t-il sur le comportement en boucle fermée, sur le découplage du système (et donc sur les trajectoires spatiales de la caméra) ? Dans la section 3.2, nous apporterons quelques éléments de réponses à ces questions.

Afin d'améliorer le comportement du schéma de commande, un autre point de vue consiste non pas à jouer sur le choix de l'inverse généralisé mais sur le type de primitives visuelles. L'idéal étant de définir six mesures conduisant à une matrice d'interaction 6×6 diagonale et constante. Bien sûr, dans ce cas la loi de commande (3.2) aboutira à un découplage total des six degrés de liberté à commander avec un point d'équilibre globalement asymptotiquement stable. On pourra par exemple se reporter à [CH01, MCC02, HM02, IO02, Tah04, TC05, TFC08, TCM08, TMCC09] pour des travaux ayant cette ligne conductrice. Nous reviendrons un peu plus en détails sur ce dernier point. En effet, dans la continuité des travaux de thèse présentés dans [Tah04], nous avons proposé l'utilisation de nouvelles primitives obtenues à partir des moments sur la sphère unité intervenant dans le modèle de projection décrit dans le chapitre 2 afin de découpler les mouvements de translation de ceux en rotation tout en diminuant l'influence des paramètres 3D

sur les variations de la matrice d'interaction.

AVH. Dans un asservissement visuel hybride, le vecteur de mesure \mathbf{s} est composé à la fois d'informations tridimensionnelles et de primitives extraites directement de l'image. L'asservissement visuel hybride peut alors être vu comme un compromis entre les asservissements visuels 2D et 3D. En effet, il permet de contrôler à la fois la trajectoire de la caméra dans son espace de travail et la trajectoire de certaines primitives dans l'espace image. Le premier schéma d'asservissement visuel hybride a été proposé par Malis *et al* dans [Mal98]. Il utilise une reconstruction partielle de la pose sans connaissance *a priori* du modèle 3D de l'objet par l'intermédiaire de l'estimation de la matrice d'homographie relative à un plan de référence. La fonction de tâche, dans ce cas, est composée de deux parties. La première partie contient les coordonnées images d'un point et une grandeur associée au ratio des profondeurs de celui-ci pour les positions courante et désirée de la caméra. La seconde composante représente la rotation (axe/angle) entre les positions courante et désirée de la caméra. La matrice d'interaction associée à cette fonction de tâche a une forme simple et est partiellement découplée, ce qui permet d'aboutir à des résultats formels sur la stabilité en présence d'erreurs de modélisation. Une approche similaire à celle décrite dans [Mal98] a été proposée dans [MLS⁺00]. Le choix des primitives géométriques 2D utilisées dans la définition de la fonction de tâche assure dans ce cas la visibilité de l'objet. Cependant, la forme analytique de l'interaction associée n'a plus les propriétés intéressantes obtenues dans [Mal98] et l'analyse de stabilité est alors beaucoup plus difficile. Citons finalement les travaux de Cervera *et al.* décrit dans [CPBM03]. Ils proposent de combiner les coordonnées dans l'image de points à leurs profondeurs. Les trajectoires 3D de la caméra sont dans ce cas satisfaisantes et comparables à celles obtenues avec un asservissement visuel 3D. De notre côté, nous avons proposé, dans un premier temps, l'extension des techniques hybrides basées sur l'observation d'un ensemble de points par une caméra conventionnelle au cas où cette observation est faite par une caméra obéissant au modèle unifié de projection centrale. Nous avons ensuite considéré l'utilisation d'un ensemble de droites dans ce type de schéma d'asservissement.

Comme nous l'avons évoqué précédemment, les recherches dans le domaine de l'asservissement visuel se nourrissent de considérations à la fois théoriques et pratiques. Nos contributions, synthétisées dans la suite du document ont trait à ces deux aspects. Le premier point que nous allons évoquer concerne le choix de l'inverse généralisé \mathbf{C} lorsque la matrice d'interaction n'est pas inversible. Nous présentons à cette occasion deux alternatives à l'inverse de Moore-Penrose. Le second point concerne nos contributions à l'asservissement visuel lorsque la caméra obéit au modèle de projection unifié.

3.2 Choix de l'inverse généralisé

Rappelons tout d'abord que la loi de commande (3.3) est obtenue via l'inversion de la relation (3.1). La matrice $\hat{\mathbf{C}}$ (estimation de l'inverse généralisé de la matrice d'interaction) joue alors un rôle primordial dans la construction de la loi de commande et conditionne le comportement en boucle fermée du système à réguler à travers la relation (3.4). Plusieurs choix sont possibles. Parmi les plus classiques, on retrouve les inverses généralisés reflétant les non linéarités du premier

ordre de la relation (3.1) [CH06, CH07] :

- $\hat{\mathbf{C}} = \hat{\mathbf{L}}_{\mathbf{s}}^+$ où $\hat{\mathbf{L}}_{\mathbf{s}}^+$ est la pseudo-inverse de Moore-Penrose d'une approximation de la matrice d'interaction.
- $\mathbf{C} = \hat{\mathbf{L}}_{\mathbf{s}^*}^+ = \hat{\mathbf{L}}_{\mathbf{s}}^+|_{\mathbf{s}=\mathbf{s}^*}$ où la matrice $\hat{\mathbf{L}}_{\mathbf{s}}^+$ est calculée pour la position désirée de la caméra.

On peut également utiliser une combinaison de ces deux matrices $\mathbf{C} = \beta \hat{\mathbf{L}}_{\mathbf{s}}^+ + (1 - \beta) \hat{\mathbf{L}}_{\mathbf{s}^*}^+$ où β est un scalaire tel que $0 \leq \beta \leq 1$ [MC08]. Dans le cas où $\beta = 1/2$, on obtient $\mathbf{C} = (\hat{\mathbf{L}}_{\mathbf{s}}^+ + \hat{\mathbf{L}}_{\mathbf{s}^*}^+)/2$ utilisée dans la méthode ESM [Mal04]. Dans ce cas certaines non-linéarités d'ordre 2 de (3.1) sont prises en compte. Nous reviendrons un peu plus en détails sur ce dernier choix dans la suite.

Remarque. Dans tous les cas cités précédemment, la condition de stabilité asymptotique locale du point d'équilibre du système différentiel (3.4) s'écrit :

$$\mathbf{Q}_t = (\hat{\mathbf{L}}_{\mathbf{s}}^+ \mathbf{L}_{\mathbf{s}})|_{\mathbf{s}=\mathbf{s}^*} > 0$$

Notons finalement que l'inverse généralisé peut également être obtenu de manière purement numérique en utilisant une phase d'apprentissage. Une première possibilité consiste à estimer numériquement la matrice d'interaction puis à l'inverser [HA94]. Une seconde possibilité consiste à estimer directement l'inverse généralisé de la matrice d'interaction [LJDC04]. Dans ce dernier cas, l'inverse généralisé obtenu peut alors refléter des non-linéarités d'ordre élevé de (3.1). En contre partie, la phase d'apprentissage peut s'avérer délicate à mettre en place et l'analyse de stabilité est impossible.

Dans [LM04] et [TM08], nous avons proposé deux autres possibilités permettant d'obtenir un inverse généralisé reflétant des non-linéarités d'ordre 2.

3.2.1 Utilisation d'une Jacobienne étendue

Supposons que nous soyons dans la situation classique et assez générale où un vecteur de paramètres $\boldsymbol{\theta} = (\theta_i)_{1 \leq i \leq n}$ explique le vecteur de mesures $\mathbf{x} = (x_j)_{1 \leq j \leq m}$ par une relation du type

$$\mathbf{x} = f(\boldsymbol{\theta}) = (f_j(\boldsymbol{\theta}))_{1 \leq j \leq m}. \quad (3.7)$$

Le problème inverse est d'estimer à partir des observations \mathbf{x} , les paramètres $\boldsymbol{\theta}$. f est généralement non-linéaire et non-inversible. Le problème d'inversion de f peut alors se poser comme la recherche d'une solution locale

$$\Delta \boldsymbol{\theta} = \mathbf{C} \Delta \mathbf{x},$$

pour $\Delta \mathbf{x}$ au voisinage de zéro et pour \mathbf{C} à estimer. Pour cela, de manière traditionnelle, on peut utiliser le développement de Taylor au premier ordre de \mathbf{x} au voisinage de $\boldsymbol{\theta}$,

$$f(\boldsymbol{\theta} + \Delta \boldsymbol{\theta}) = f(\boldsymbol{\theta}) + \mathbf{J}_f(\boldsymbol{\theta}) \Delta \boldsymbol{\theta} + \|\Delta \boldsymbol{\theta}\| \epsilon(\Delta \boldsymbol{\theta})$$

$\mathbf{J}_f(\boldsymbol{\theta})$ étant la matrice jacobienne de f évaluée en $\boldsymbol{\theta}$. En utilisant la pseudo-inverse de Moore-Penrose i.e. $\mathbf{C} = \mathbf{J}_f^+$, cela conduit à

$$\Delta\boldsymbol{\theta} \approx \mathbf{J}_f^+(\boldsymbol{\theta})(f(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) - f(\boldsymbol{\theta})) = \mathbf{J}_f^+(\boldsymbol{\theta})\Delta\mathbf{x}$$

Cette relation exprime les variations locales du vecteur de paramètres par une relation linéaire avec les variations locales du vecteur de mesures. Considérons maintenant le problème de minimisation $\min_{\boldsymbol{\theta}} g_f(\boldsymbol{\theta})$ avec

$$g_f(\boldsymbol{\theta}) = \frac{1}{2} \|f(\boldsymbol{\theta}) - f(\boldsymbol{\theta}^*)\|^2$$

Celui-ci peut être résolu au sens des moindres carrés de manière itérative en appliquant des déplacements $\Delta_k\boldsymbol{\theta}$ au vecteur de paramètres à partir d'une valeur initiale de celui-ci (méthode de Gauss-Newton) :

$$\Delta_k\boldsymbol{\theta} = -\lambda\mathbf{J}_f^+|_{\mathbf{x}=\mathbf{x}_k}(\mathbf{x}_k - \mathbf{x}^*) \quad (3.8)$$

Analogie avec la commande (3.2). Prenons $\mathbf{s} = f(\boldsymbol{\theta})$ comme la fonction de projection donnant \mathbf{s} connaissant $\boldsymbol{\theta}$ un vecteur de paramètres décrivant l'état de la caméra. On peut formuler le problème d'asservissement visuel comme la minimisation de la fonction g_s en utilisant la relation (3.8). En supposant que $\dot{\boldsymbol{\theta}} = \boldsymbol{\Gamma}\boldsymbol{\tau}$, on obtient la commande : $\boldsymbol{\tau} = -\lambda\mathbf{L}_s^+(\mathbf{s} - \text{sestimer}^*)$ (avec $\mathbf{L}_s = \mathbf{J}_s\boldsymbol{\Gamma}$). Cette dernière relation est équivalente à la loi de commande (3.2) avec $\mathbf{C} = \mathbf{L}_s^+$. En pratique \mathbf{L}_s^+ n'est pas estimée exactement et on prendra $\hat{\mathbf{C}} = \hat{\mathbf{L}}_s^+$ ou encore $\hat{\mathbf{C}} = \hat{\mathbf{L}}_{s^*}^+$.

Revenons maintenant au cadre général. Si les dérivées secondes de la fonction f par rapport à ses paramètres $\boldsymbol{\theta}$ sont disponibles, il est alors possible de considérer un développement de Taylor au second ordre de \mathbf{x} au voisinage de $\boldsymbol{\theta}$:

$$f(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \approx f(\boldsymbol{\theta}) + \mathbf{J}_f(\boldsymbol{\theta})\Delta\boldsymbol{\theta} + \frac{1}{2}\Delta\boldsymbol{\theta}^\top \mathbf{H}_f(\boldsymbol{\theta})\Delta\boldsymbol{\theta} + \|\Delta\boldsymbol{\theta}\|^2\epsilon(\Delta\boldsymbol{\theta}) \quad (3.9)$$

Maintenant, définissons une fonction \tilde{f} de $N = n + (n(n+1))/2$ variables ν_k : les paramètres θ_i et les paramètres supplémentaires $\theta_{il} = \theta_i\theta_l$, avec $1 \leq i \leq l \leq n$. Cette nouvelle fonction \tilde{f} est linéaire vis-à-vis de ν_k dans \mathbb{R}^N , et est par ailleurs identique à f dans la variété M_θ de dimension n générée par θ_i et $\theta_i\theta_l$. Il est intéressant de noter qu'alors le Jacobien $\mathbf{J}_{\tilde{f}}$ de \tilde{f} peut être exprimé à partir des composantes du Jacobien et du Hessien de f . Le développement de Taylor au premier ordre de la fonction \tilde{f} s'écrit :

$$\tilde{f}(\boldsymbol{\nu} + \delta\boldsymbol{\nu}) \approx \tilde{f}(\boldsymbol{\nu}) + \mathbf{J}_{\tilde{f}}(\boldsymbol{\nu})\Delta\boldsymbol{\nu} + \|\Delta\boldsymbol{\nu}\|\epsilon(\Delta\boldsymbol{\nu}),$$

Il est alors aisé de vérifier que le reste du premier ordre est en fait du second ordre quand la fonction est restreinte à la variété M_θ . En prenant les n premières lignes de la pseudo-inverse de Moore-Penrose de la Jacobienne $\mathbf{J}_{\tilde{f}}$ (notée $\tilde{\mathbf{H}}_f^+$ dans la suite), on peut alors estimer les variations des n paramètres θ_i en fonction des mesures en tenant compte des non-linéarités d'ordre 2 de la relation 3.7 :

$$\Delta\boldsymbol{\theta} = \tilde{\mathbf{H}}_f^+ \Delta\mathbf{x}$$

Un exemple illustratif simple. Afin d'illustrer cette méthode, considérons une fonction $f(\theta_1, \theta_2)$ de \mathbb{R}^2 dans \mathbb{R}^p , $f = (f_i)_{1 \leq i \leq p}$ définie par :

$$x_i = f_i(\theta_1, \theta_2) = a_{i1}\theta_1 + a_{i2}\theta_2 + a_{i3}\theta_1^2 + a_{i4}\theta_1\theta_2 + a_{i5}\theta_2^2.$$

En $(0, 0)$, la matrice Jacobienne de f , \mathbf{J}_f s'écrit :

$$\mathbf{J}_f(\theta_1, \theta_2) = ((a_{ij})_{1 \leq i \leq p, 1 \leq j \leq 2})$$

Définissons maintenant \tilde{f} comme décrit précédemment,

$$x_i = \tilde{f}_i(\theta_1, \theta_2, \dots, \theta_5) = a_{i1}\theta_1 + a_{i2}\theta_2 + a_{i3}\theta_3 + a_{i4}\theta_4 + a_{i5}\theta_5$$

\tilde{f} est linéaire en les 5 paramètres $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ et est identique à f dans la variété de \mathbb{R}^5 générée par $(\theta_1, \theta_2, \theta_1^2, \theta_1\theta_2, \theta_2^2)$. $\mathbf{J}_{\tilde{f}}$ est alors une matrice constante :

$$\mathbf{J}_{\tilde{f}}(\theta_1, \theta_2, \dots, \theta_5) = ((a_{ij})_{1 \leq i \leq p, 1 \leq j \leq 5})$$

Si on suppose que l'on s'intéresse à une inversion au voisinage de l'origine, on peut par exemple procéder des deux manières suivantes :

1. calcul de \mathbf{J}_f et \mathbf{J}_f^+ ,
2. calcul de $\mathbf{J}_{\tilde{f}}$ et extraction de $\tilde{\mathbf{H}}_f^+$ à partir de $\mathbf{J}_{\tilde{f}}^+$

Le Tableau 3.1 montre la précision des différentes approximations. Les coefficients de \tilde{f} sont choisis aléatoirement dans $[0, 1]$, les écarts types de la norme $\|\mathbf{C}\Delta\mathbf{x} - \Delta\boldsymbol{\theta}\|$ pour 1000 valeurs aléatoires de $\boldsymbol{\theta}$ sont présentés (la matrice \mathbf{C} étant prise égale à \mathbf{J}_f^+ ou $\tilde{\mathbf{H}}_f^+$). Pour les deux premières lignes du tableau, les valeurs de $\boldsymbol{\theta}$ sont choisies dans l'intervalle $[-0.05, 0.05]$ alors que pour les deux dernières, elles sont prises dans l'intervalle $[-5, 5]$.

number of values	1	2	3	4	5	6
\mathbf{J}_f^+	1e-2	2e-3	1e-3	1e-3	9e-4	1e-3
$\tilde{\mathbf{H}}_f^+$	1e-2	1e-2	1e-2	6e-3	1e-16	5e-17
$\mathbf{J}_{\tilde{f}}^+$	1e+1	4e+1	2e+1	1e+1	1e+1	1e+1
$\tilde{\mathbf{H}}_f^+$	3e+0	3e+0	4e+0	4e+0	1e-14	1e-14

TAB. 3.1 – Précisions des inversions locales.

Pour obtenir une bonne estimation de $\tilde{\mathbf{H}}_f^+$, il faut au minimum 5 points (puisque 3 paramètres supplémentaires ont été introduits pour définir \tilde{f}). Le résultat est alors parfait (aux erreurs d'arrondis près), la fonction de départ étant quadratique. On peut également noter que dans ce cas le domaine de validité de l'approximation \mathbf{J}_f^+ est un faible voisinage autour de l'origine alors que c'est l'espace de définition complet de \tilde{f} pour $\tilde{\mathbf{H}}_f^+$ car c'est dans ce cas une matrice constante.

Cette méthode d'approximation des variations des paramètres peut être utilisée assez directement dans le cadre de l'asservissement visuel. Pour cela, on choisit $\tilde{\mathbf{H}}_s$ comme inverse généralisé

C apparaissant dans la loi de commande (3.2). Dans [LM04], nous avons utilisé cette dernière loi de commande pour un asservissement visuel 2D basé sur les coordonnées de points dans l'image. Nous avons alors observé que cette méthode permet d'améliorer le découplage des mouvements contrôlés. En particulier, le phénomène bien connu de couplage entre les mouvements de translation le long de l'axe optique et de rotation autour de ce même axe est complètement annulé. La Figure 3.1 montre un résultat de simulation pour une rotation pure de 179deg autour de l'axe optique confirmant cette dernière remarque.

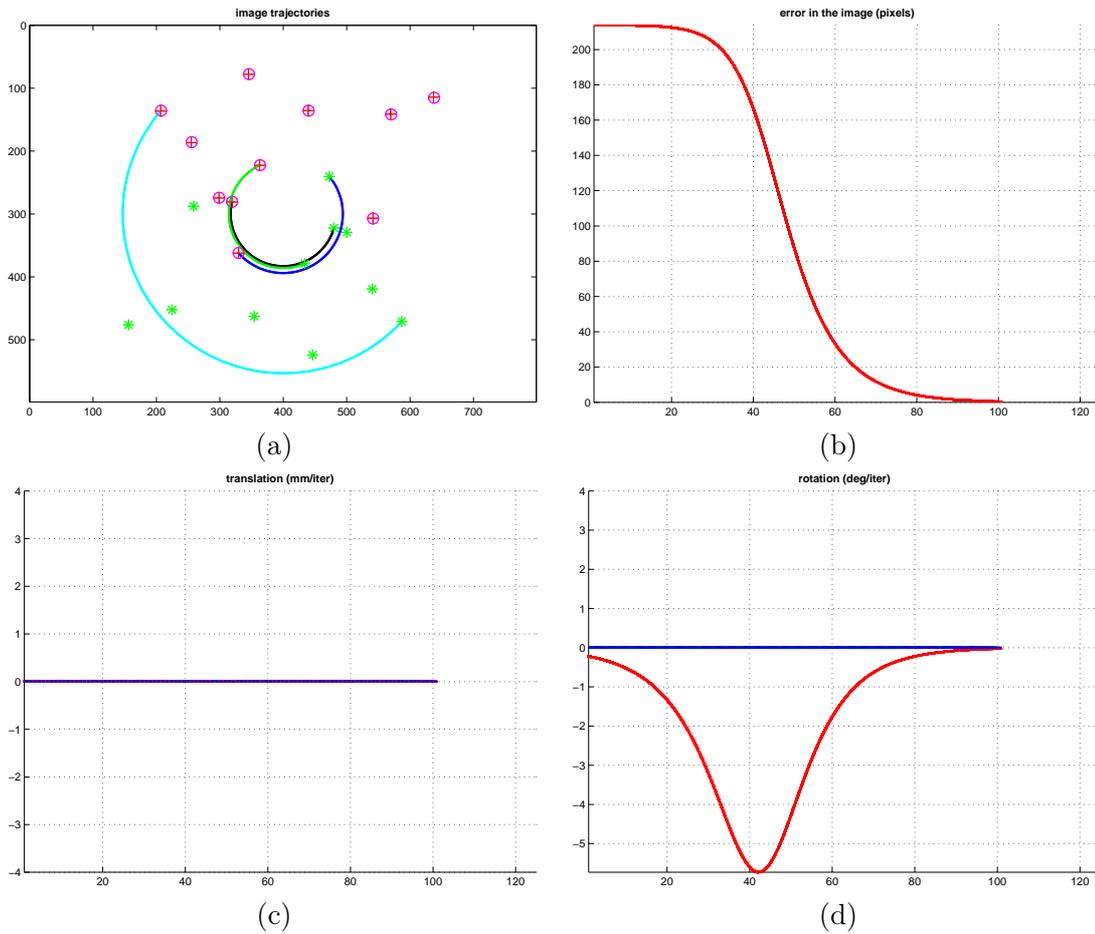


FIG. 3.1 – Simulation d'un asservissement visuel 2D-points en utilisant la matrice Jacobienne étendue : (a) trajectoire des points dans l'image, (b) norme de l'erreur, (c) vitesse de translation (mm/s), (d) vitesse de rotation (deg/s). On observe aucun mouvement le long de l'axe optique.

3.2.2 ESM modifié

Il a été démontré dans [Mal04] que l'utilisation d'une combinaison de la valeur moyenne des matrices d'interaction calculées à l'équilibre et à chaque itération pour définir l'inverse généralisé

de la matrice d'interaction améliore le comportement spatial de la caméra lorsque des coordonnées de points dans l'image sont utilisées. En particulier, le couplage de la rotation autour de l'axe optique de la caméra et de la translation le long du même axe peut être annulé. Dans [TM08], nous avons proposé une modification de cette approche que nous résumons ici. Pour cela, supposons que $f(\boldsymbol{\theta}^*)$ et $\mathbf{J}_f(\boldsymbol{\theta})^*$ soient disponibles, ce qui est généralement le cas dans le contexte de l'asservissement visuel. Considérons alors le développement de Taylor au second ordre de f au voisinage de $\boldsymbol{\theta}^*$:

$$f(\boldsymbol{\theta}^* + \Delta\boldsymbol{\theta}) \approx f(\boldsymbol{\theta}^*) + \mathbf{J}_f(\boldsymbol{\theta}^*)\Delta\boldsymbol{\theta} + \frac{1}{2}\Delta\boldsymbol{\theta}^\top \mathbf{H}_f(\boldsymbol{\theta}^*)\Delta\boldsymbol{\theta} + \|\Delta\boldsymbol{\theta}\|^2 \epsilon(\Delta\boldsymbol{\theta}) \quad (3.10)$$

En combinant les développements de Taylor au second ordre de f autour de $\boldsymbol{\theta}$ et $\boldsymbol{\theta}^*$ (Equations (3.9) et (3.10)), il est montré dans [Mal04] que les variations du vecteur de paramètre $\boldsymbol{\theta}$ peuvent être estimées en utilisant les inverses généralisés \mathbf{C}_1 et \mathbf{C}_2 reflétant les non-linéarités du second ordre de la relation (3.7) :

$$\mathbf{C}_1 = \frac{1}{2}(\mathbf{J}_f(\boldsymbol{\theta}) + \mathbf{J}_f(\boldsymbol{\theta}^*))^+ \quad (3.11)$$

$$\mathbf{C}_2 = 2(\mathbf{J}_f(\boldsymbol{\theta})^+ + \mathbf{J}_f(\boldsymbol{\theta}^*)^+)^+ \quad (3.12)$$

Dans [TM08], nous avons pointé quelques limitations de cette approche. En particulier, nous avons souligné que l'utilisation des inverses généralisés \mathbf{C}_1 et \mathbf{C}_2 dans le contexte de l'asservissement visuel ne conduit pas à une approximation du second ordre pour l'inversion de la relation (3.1). L'explication basique est que les développements de Taylor (3.9) et (3.10) introduisent des variations des paramètres $\boldsymbol{\theta}$ qu'il faut mettre en relation avec le vecteur de commande : le torseur cinématique exprimé et appliqué dans un repère associé à la caméra dans sa configuration courante. Pour cela, la combinaison des variations des paramètres obtenue via les Jacobiennes $\mathbf{J}_f(\boldsymbol{\theta})$ et $\mathbf{J}_f(\boldsymbol{\theta}^*)$ doit être traduite en une combinaison du torseur cinématique via les matrices d'interaction \mathbf{L}_s et \mathbf{L}_s^* . Ce qui amène à la définition suivante des inverses généralisés :

$$\mathbf{C}_{1mod} = \frac{1}{2}(\mathbf{L}_s^+ + \mathbf{T}\mathbf{L}_s^+)^+ \quad (3.13)$$

$$\mathbf{C}_{2mod} = 2(\mathbf{L}_s + \mathbf{L}_s^* \mathbf{T}^{-1})^+ \quad (3.14)$$

où \mathbf{T} permet le passage d'un torseur cinématique exprimé et appliqué dans le repère caméra désiré à un torseur cinématique exprimé et appliqué dans le repère caméra courant. On remarque alors que lorsque les deux positions de la caméra sont proches *i.e* $\mathbf{T} \approx \mathbf{I}_{6 \times 6}$ alors on retrouve $\mathbf{C}_{1mod} = \frac{1}{2}(\mathbf{L}_s^+ + \mathbf{L}_s^+)^+$ et $\mathbf{C}_{2mod} = 2(\mathbf{L}_s + \mathbf{L}_s^*)^+$. D'autre part, la condition de stabilité asymptotique locale s'écrit une fois encore $\mathbf{Q}_t = (\widehat{\mathbf{L}}_s^+ \mathbf{L}_s)|_{s=s^*} > 0$.

Résultat. A titre d'exemple, nous allons comparer le comportement de la loi de commande (3.3) pour trois choix de l'inverse généralisé :

- (cas A) la pseudo-inverse de Moore-Penrose de la matrice d'interaction,
- (cas B) la moyenne des pseudo-inverses des matrices d'interaction proposée dans [Mal04]
- (cas C) l'inverse généralisé (3.13).

Pour cela, nous utilisons une cible composée de quatre points et les primitives visuelles décrites dans [TCM08]. Le déplacement entre les positions initiale et finale est :

$$\boldsymbol{\theta}_{\mathbf{u}} = \begin{pmatrix} 0 & 10.57 & -79.29 \end{pmatrix} dg \quad (3.15)$$

$$\mathbf{t} = \begin{pmatrix} -0.4 & -0.1 & -0.6 \end{pmatrix} \text{mètre} \quad (3.16)$$

Les résultats sont donnés par la Figure 3.2. On observe que pour le choix A, malgré les couplages sur les translations, le comportement de l'erreur (Figure 3.2.(b)) et des rotations (Figure 3.2.(a)) est satisfaisant. On remarque également que l'utilisation de B n'améliore pas les résultats obtenus avec A. Au contraire, des couplages supplémentaires apparaissent sur les vitesses (Figures 3.2.(e), 3.2.(f)). Par contre, l'utilisation de C conduit à des résultats très satisfaisants (Figures 3.2.(g), 3.2.(h) et 3.2.(i)). Dans ce cas les couplages observées pour les deux autres choix sont quasiment éliminés.

Nous avons vu que pour un vecteur de mesure \mathbf{s} donné le comportement de la loi de commande (3.3) dépendait fortement du choix de l'inverse généralisé \mathbf{C} . Ce comportement est également sensiblement affecté par le choix de \mathbf{s} . Dans la suite, nous nous consacrons à cette problématique lorsque le modèle de projection de la caméra est le modèle de projection unifié décrit dans le premier chapitre (se résumant par la relation (2.5)). En outre, la loi de commande (3.3) sera utilisée avec \mathbf{C} prise égale à l'inverse (quand celle-ci existe) ou à la pseudo-inverse de Moore-Penrose de la matrice d'interaction.

3.3 Asservissement visuel utilisant le modèle unifié

Une grande majorité des travaux dans le domaine de la commande par vision omnidirectionnelle traite de la commande et de la navigation de robots mobiles. On pourra se référer à [CH00a, SDF⁺01, GWSV00, MAPP05, PPC04, VSS03] pour quelques exemples. Dans ce contexte, nous avons présenté de notre côté dans [AMAM05b, HAMMC08] des solutions pour le suivi de droites 3D lorsque le robot mobile est non-holonyme. En outre, nous avons également proposé des schémas d'asservissement visuel découplant les mouvements de translation des mouvements de rotation en considérant une tâche unique [HAMAM06a, HAMM07]. Ces méthodes peuvent aisément être appliquées aux cas des robots mobiles holonomes.

Assez peu de travaux traitent d'asservissement visuel omnidirectionnel pour des robots à six degrés de liberté. Dans ce contexte, Barreto *et al.* dans [BMH02] ont étendu le schéma d'asservissement visuel 2D à partir de points observés par une caméra conventionnelle à un capteur catadioptrique. Benhimane *et al.* ont proposé un schéma d'asservissement visuel basé homographie dans [BM06b]. Plus récemment Tatsambon *et al.* ont présenté dans [TFC08] une loi de commande asymptotiquement stable en exploitant les propriétés spécifiques de la projection sphérique d'une sphère. Nous décrivons dans la suite nos contributions à l'asservissement visuel avec une caméra répondant au modèle unifié de projection. Ces contributions auraient pu être organisées de diverses manières : en fonction de la catégorie de robot commandé (à roue, série, parallèle), en fonction du type de primitives visuelles utilisé : points, droites et moments ou enfin en fonction de l'espace de commande (2D, 3D, 2D1/2). C'est cette dernière organisation qui a été choisie.

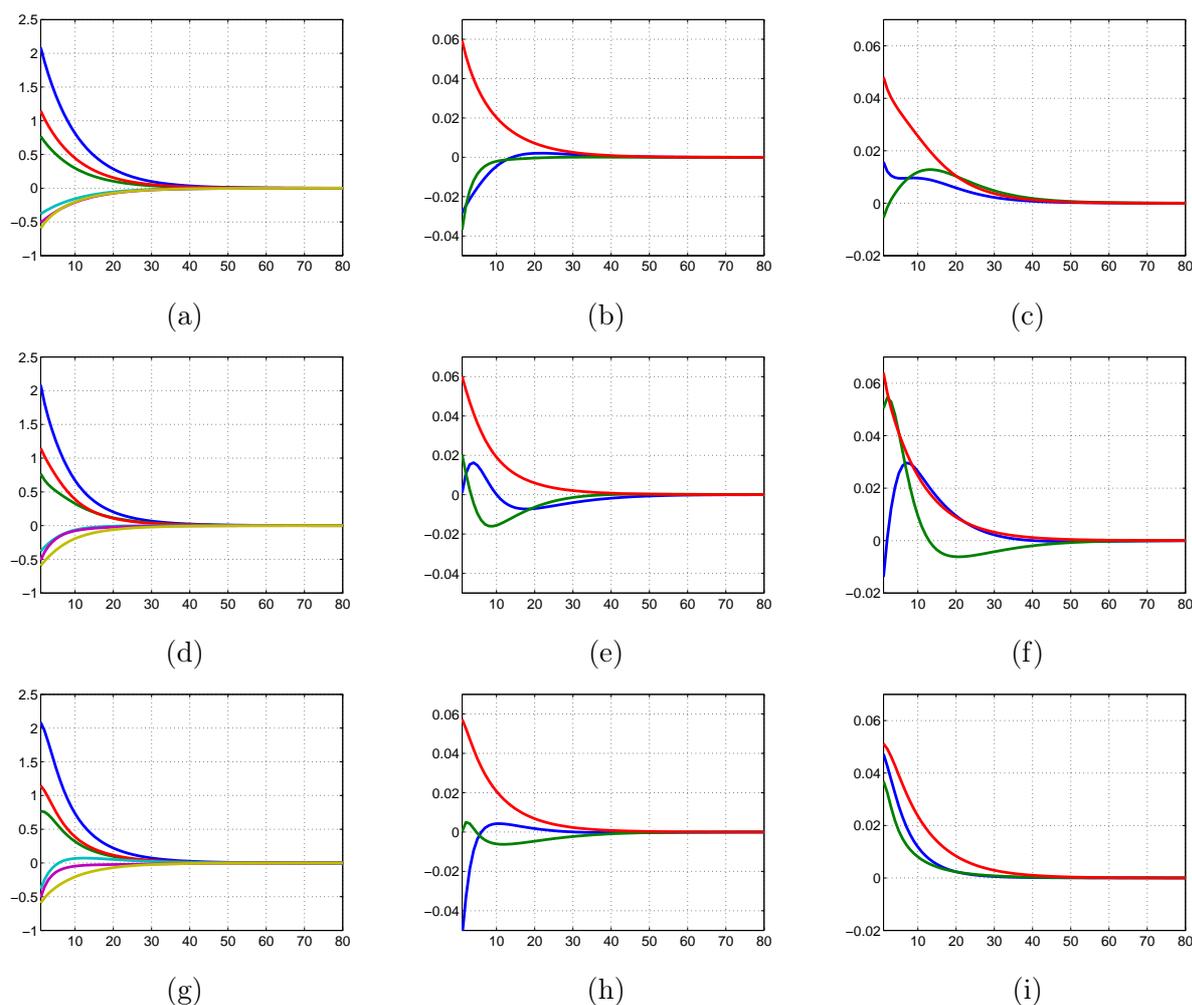


FIG. 3.2 – Positionnement d’une caméra par rapport à quatre points : résultat dans le cas A ((a) erreurs, (b) vitesses de rotation rad/s, (c) vitesses de translation (m/s)), résultats dans le cas B ((d) erreurs, (e) vitesses de rotation (rad/s), (f) vitesses de translation (m/s)), résultats dans le cas C ((g) erreurs, (h) vitesses de rotation (rad/s), (i) vitesses de translation (m/s))

3.3.1 Asservissement visuel dans l’image

Cette partie résume donc nos contributions à l’asservissement visuel 2D (AV2D). Le premier paragraphe synthétise l’étude menée dans [MMR09] concernant la stabilité asymptotique locale de l’AV2D basé points lorsque l’estimation des profondeurs est erronée. Le second paragraphe, quant à lui, résume nos contributions dans le contexte de l’AV2D basé droites. Nous présentons dans ce cadre des résultats sur la commande de robots à roue non-holonomes, manipulateurs sériel et parallèle. Ces contributions sont détaillées dans [HA06b], [AMM09b], [HAMMC08], [TMAM09]. Enfin, le dernier paragraphe de cette section présente notre approche AV2D basé moment plus précisément décrite dans [TMCA09], [TMCC09].

Des points sont extraits de l'image

Un choix assez classique de mesure est le vecteur de coordonnées $\mathbf{s}_i = [x_i \ y_i]^\top$ de la projection dans l'image d'un point 3D \mathcal{X}_i ayant pour coordonnées $\mathbf{X}_i = (X_i, Y_i, Z_i)$ dans \mathcal{F}_m . La matrice d'interaction relative à cette primitive $\mathbf{L}_{\mathbf{s}_i} = [\mathbf{A}_i \ \mathbf{B}_i]$ est donnée par [BMH02] :

$$\mathbf{A}_i = \begin{bmatrix} -\frac{1+x_i^2(1-\xi(\sigma_i+\xi))+y_i^2}{\rho_i(\sigma_i+\xi)} & \frac{\xi x_i y_i}{\rho_i} & \frac{x_i \sigma_i}{\rho_i} \\ \frac{\xi x_i y_i}{\rho_i} & -\frac{1+x_i^2+y_i^2(1-\xi(\sigma_i+\xi))}{\rho_i(\sigma_i+\xi)} & \frac{y_i \sigma_i}{\rho_i} \end{bmatrix} \quad (3.17)$$

$$\mathbf{B}_i = \begin{bmatrix} x_i y_i & -\frac{(1+x_i^2)\sigma_i - \xi y_i^2}{\sigma_i + \xi} & y_i \\ \frac{(1+y_i^2)\sigma_i - \xi x_i^2}{\sigma_i + \xi} & -x_i y_i & -x_i \end{bmatrix} \quad (3.18)$$

où $\sigma_i = \sqrt{1 + (1 - \xi^2)(x_i^2 + y_i^2)}$. On peut noter qu'en posant $\xi = 0$, alors on retrouve le résultat obtenu en utilisant le modèle pinhole. Il est assez aisé de reformuler la partie gauche de la matrice d'interaction sous la forme :

$$\mathbf{A}_i = \frac{1}{Z_i} \mathbf{G}_i$$

où \mathbf{G}_i ne contient pas d'information 3D. En utilisant cette formulation seule la profondeur Z_i est une inconnue de la matrice d'interaction. La matrice d'interaction relative aux coordonnées dans l'image de n points $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n)$ s'écrit $\mathbf{L}(\mathbf{z}, \mathbf{s}) = (\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n)$ et peut être décomposée en deux sous-matrices : $\mathbf{L}(\mathbf{z}, \mathbf{s}) = [\mathbf{A}(\mathbf{z}, \mathbf{s}) \ \mathbf{B}(\mathbf{s})]$, avec $\mathbf{A}(\mathbf{z}, \mathbf{s}) = \mathbf{D}(\mathbf{z}) \ \mathbf{G}(\mathbf{s})$ où $\mathbf{G}(\mathbf{s}) = (\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_n)$ et $\mathbf{D}(\mathbf{z})$ est une matrice diagonale contenant les profondeurs des points. A partir de cette formulation de la matrice d'interaction, nous avons étudié la stabilité asymptotique locale du point d'équilibre de la loi de commande (3.2) pour $\mathbf{C} = \beta \mathbf{L}_{\mathbf{s}} + (1 - \beta) \mathbf{L}_{\mathbf{s}^*}$ en présence d'erreurs sur les profondeurs \mathbf{z} . Dans ce cas, on rappelle que la condition de stabilité asymptotique locale est donnée par la relation (3.6). On peut également noter que cette étude s'étend aux cas où $\mathbf{C} = \mathbf{D} \mathbf{L}_{\mathbf{s}} + \mathbf{E} \mathbf{L}_{\mathbf{s}^*}$ si \mathbf{D} et \mathbf{E} sont deux matrices, qui évaluées au point d'équilibre, sont égales à l'identité (dans ce cas la condition de stabilité reste (3.6)). Cela nous permet d'étendre les résultats présentés dans [MMR09] à l'ESM modifié présenté dans le paragraphe 3.2.2. En particulier, une condition nécessaire et suffisante et une condition suffisante de stabilité asymptotique locale du point d'équilibre ont été proposées. Un exemple de résultats est donné par les Figures 3.3 et 3.4 pour une cible planaire de normale $\mathbf{n}(\theta, \phi) = (\cos(\theta) \sin(\phi), \sin(\theta) \sin(\phi), \cos(\phi))$. Dans ce cas, les erreurs sur les profondeurs sont équivalentes à des erreurs sur les paramètres θ et ϕ . Ces figures montrent d'une part, que les zones de stabilité pour les différents capteurs centraux sont similaires et d'autre part, qu'une augmentation du nombre de points n'améliore que légèrement les résultats. Notons également qu'un choix assez intuitif lorsque l'on n'a aucune idée sur la position du plan cible *i.e* $\mathbf{n}(\theta, \phi) = (0, 0, 1)$ conduit à une commande instable dans le cas où 3 points sont utilisés.

Nous avons également constaté le résultat surprenant décrit par la conjecture 1 et démontré le résultat décrit par le théorème 1 :

Conjecture 1. Si une caméra centrale observe une cible symétrique et centrée (c.à.d pour

chaque point $\mathbf{m} = (x, y, 1)$, le point $\mathbf{m}' = (-x, -y, 1)$ appartient à la cible) parallèle au plan image (i.e. $\mathbf{n} = (0, 0, 1)$), la loi de commande (3.2) avec $\mathbf{C} = \beta\mathbf{L}_s + (1 - \beta)\mathbf{L}_{s^*}$ a un point d'équilibre asymptotiquement stable quelles que soient les erreurs commises sur les profondeurs.

Théorème 1. Si une caméra perspective observe une cible centrée composée de quatre points formant un carré (i.e. l'image de référence consiste en les quatre points $\mathbf{m}_1 = (-\alpha, -\alpha, 1)$, $\mathbf{m}_2 = (\alpha, -\alpha, 1)$, $\mathbf{m}_3 = (\alpha, \alpha, 1)$, $\mathbf{m}_4 = (-\alpha, \alpha, 1)$) alors la loi de commande (3.2) avec $\mathbf{C} = \beta\mathbf{L}_s + (1 - \beta)\mathbf{L}_{s^*}$ a un point d'équilibre asymptotiquement stable quelles que soient les erreurs commises sur les profondeurs.

La Figure 3.5 illustre la conjecture 1. La cible est composée de six points symétriques par rapport au point principal et est parallèle au plan image (voir Figure 3.5(a)). Elle est observée par une caméra combinant un imageur pinhole et un miroir hyperbolique. Les coordonnées exprimées en pixel de ces six points sont $\{(320, 418), (180, 81), (180, 418), (320, 81), (106, 250), (394, 250)\}$. Dans ce cas, la loi de commande (3.2) est stabilisante quelles que soient les erreurs sur les profondeurs (se reporter à la Figure 3.5(b)). La cible est ensuite légèrement déplacée afin qu'elle n'apparaisse plus comme symétrique dans l'image (voir Figure 3.5(c)-(d)). Les nouvelles coordonnées des points dans l'image sont $\{(322, 422), (183, 90), (181, 414), (320, 80), (114, 250), (401, 250)\}$ et comme on peut le constater sur la Figure 3.5(d), le domaine de stabilité reste très important. L'image de référence est maintenant modifiée de manière plus importante (les coordonnées des points dans l'image sont alors $\{(324, 427), (190, 103), (183, 409), (317, 88), (125, 250), (394, 250)\}$, (voir Figure 3.5(e)-(f)). Le domaine d'instabilité s'accroît alors de manière plus importante. Ce résultat est très intéressant car il peut permettre de choisir adéquatement la cible et sa position lors de la phase d'acquisition de la référence afin d'assurer la stabilité asymptotique du point d'équilibre. Plus précisément, quand cela est possible, on choisira une cible symétrique et une configuration désirée telle que l'axe optique de la caméra soit perpendiculaire au plan image et passe par le barycentre de la cible.

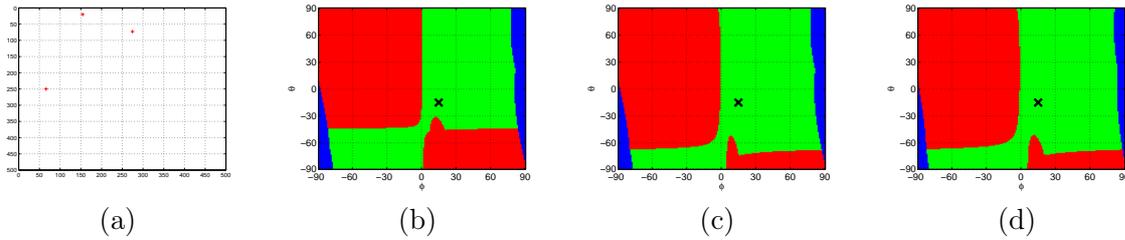


FIG. 3.3 – (a) Configuration des points dans l'image ; Domaine de stabilité : (b) caméra conventionnelle, (c) caméra avec miroir parabolique, (d) caméra avec miroir hyperbolique. Si les paramètres sont choisis dans la zone verte alors le point d'équilibre est asymptotiquement stable. Si les paramètres sont choisis dans la zone rouge alors le point d'équilibre n'est pas asymptotiquement stable. Si les paramètres sont choisis dans la zone bleue alors un des points a une profondeur négative.

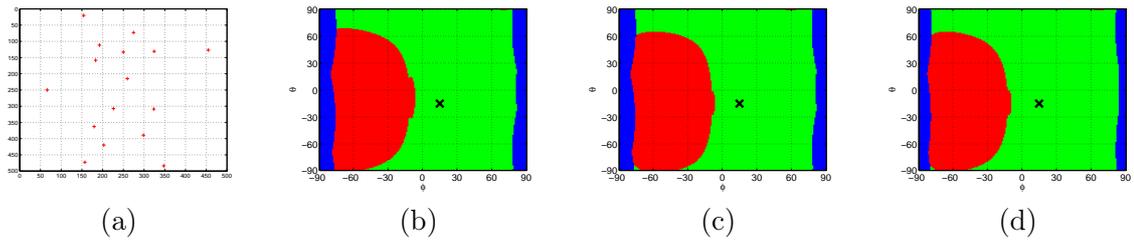


FIG. 3.4 – (a) Configuration des points dans l'image ; Domaine de stabilité : (b) caméra conventionnelle, (c) caméra avec miroir parabolique, (d) caméra avec miroir hyperbolique. En comparant avec la Figure 3.3, on observe que la surface de la zone de stabilité varie peu avec l'augmentation du nombre de points.

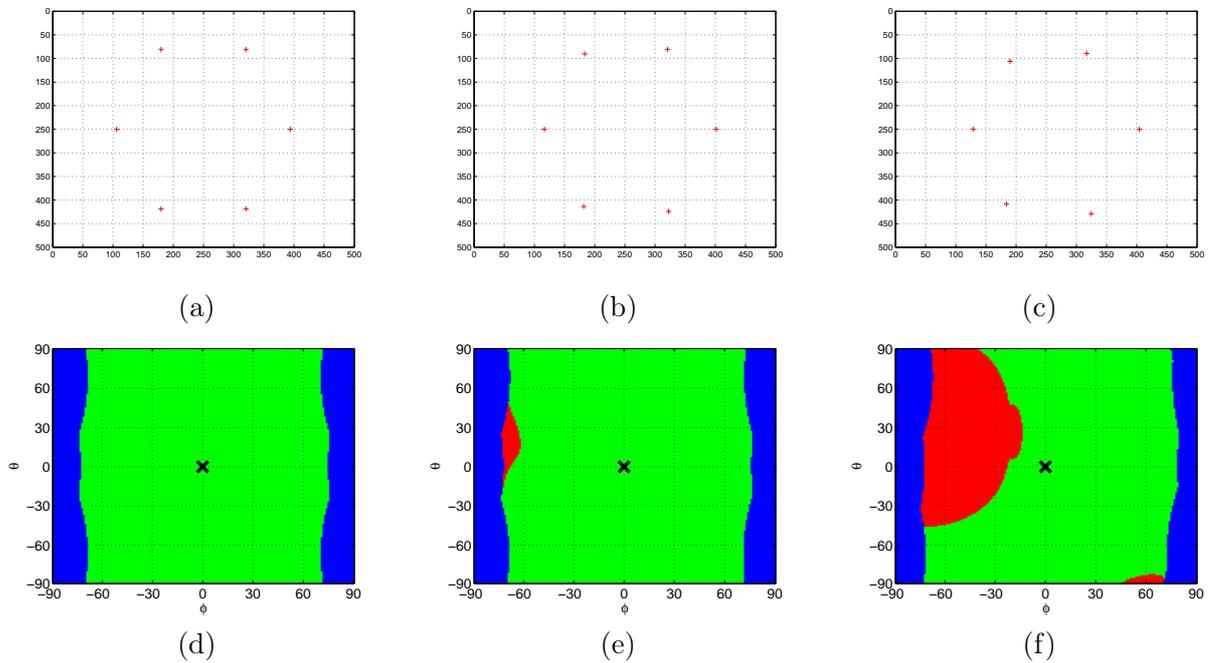


FIG. 3.5 – (a)-(b)-(c) Configuration des points dans l'image ; (d)-(e)-(f) les régions de stabilité associées. On observe que la surface de la zone de stabilité (en vert) diminue avec la symétrie de la cible.

Maintenant des coniques sont extraites de l'image

Comme évoqué précédemment, dans le cas des primitives droites nous avons étudié la commande de robots sériel, parallèle et à roues non-holonome. Bien que les lois de commande proposées dans ces trois cas de figures exploitent des concepts différents, les vecteurs d'observations sont toujours obtenus à partir des coefficients normalisés B_k des équations des coniques corres-

pendant à la projection de droites 3D dans l'image (voir Equation 2.11).

Le robot est sériel. Considérons une conique Ω_i issue de la projection d'une droite 3D \mathcal{L} via le modèle de projection unifié. Comme nous l'avons présenté dans le chapitre 2, la conique Ω_i représentant la droite 3D \mathcal{L} dans l'image peut être transformée en une droite 2D \mathbf{l}_i dans l'image d'une caméra perspective virtuelle (droite polaire associée au point principal par rapport à Ω_i). Il est alors possible d'utiliser **tout schéma d'asservissement visuel** adapté aux primitives de type droite. On pourra alors se référer à [CM00].

Il est également possible d'utiliser directement deux (parmi quatre) des coefficients normalisés B_k obtenus à partir de l'équation de la conique dans l'image réelle puisque la projection d'une droite dans l'image ne possède que deux degrés de liberté. En combinant au moins trois droites on peut alors utiliser la loi de commande (3.2) avec $\mathbf{C} = \mathbf{L}_s^+$. En pratique, il est préférable d'utiliser plus de trois droites afin d'éviter les configurations singulières. La forme analytique de la matrice d'interaction correspondante est donnée dans [HAMMC08]. Un exemple de résultat obtenu avec une caméra catadioptrique en situation déportée est donné par la Figure 3.6. La librairie VISP [MSC05] a été utilisée pour les traitements algébriques et d'images. De manière assez classique, on observe un couplage assez important entre les divers degrés de liberté contrôlés. Nous verrons par la suite que ce point peut être amélioré avec un asservissement visuel basé homographie.

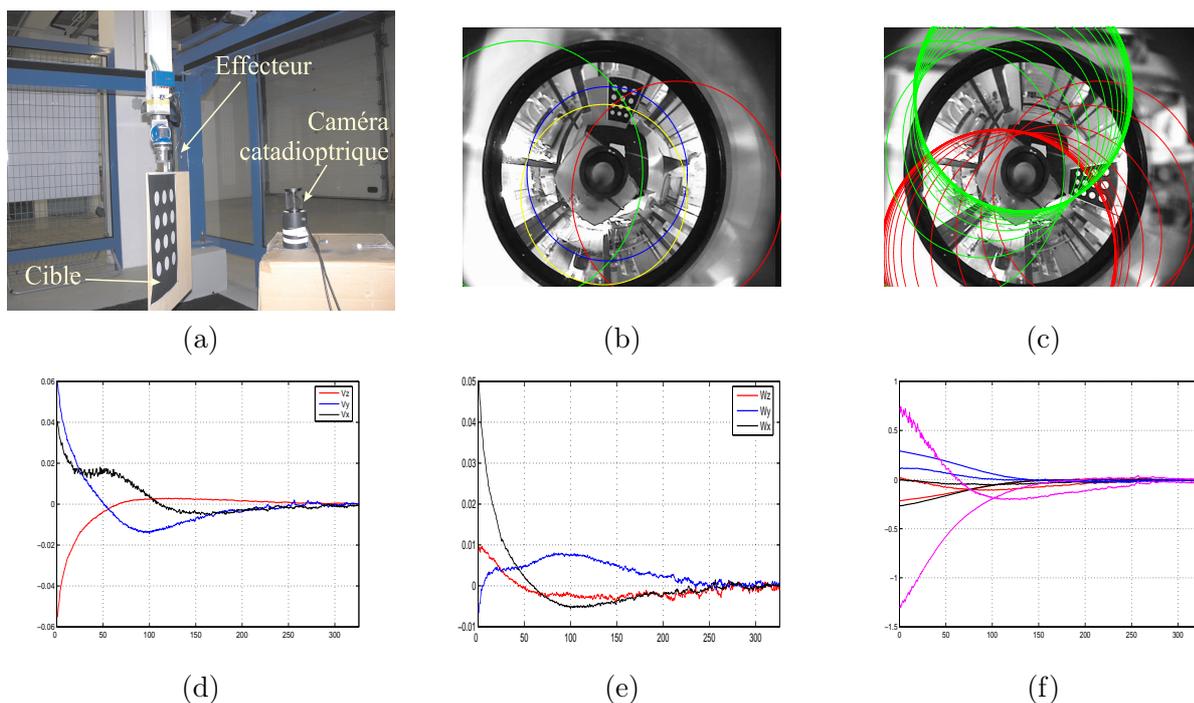


FIG. 3.6 – (a) Plateforme expérimentale, (b) image initiale, (c) image de référence et trajectoires de 2 coniques parmi les 4 utilisées pour la commande, (d) vitesses de translation [m/s], (e) vitesses de rotation [rad/s] et (f) erreurs dans l'image $\mathbf{s} - \mathbf{s}^*$ versus nombre d'itérations.

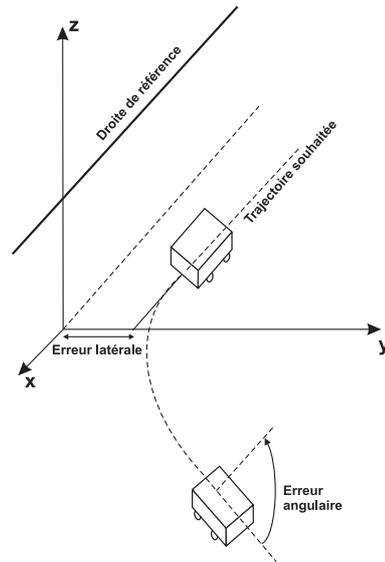


FIG. 3.7 – Tâche à réaliser : en fin de tâche l’axe X du repère de commande est parallèle et à une distance constante de la droite 3D de référence

Le robot est à roues non-holonome. Dans [HAMMC08], nous avons proposé, comme pour un robot sériel, l’utilisation des coefficients normalisés B_k issus de la projection d’une droite 3D de référence afin de réguler la position et l’orientation d’un robot mobile à roues non-holonome. La tâche que nous avons considérée consiste à amener l’axe X du repère de commande parallèle et à une distance constante de la droite 3D de référence (se référer à la Figure 3.7). Pour cela, nous avons construit une loi de commande en utilisant une modélisation par système chaîné [Sam95]. De manière classique, cette modélisation fait intervenir les erreurs latérale et angulaire du robot par rapport à la droite 3D de référence. Nous avons montré dans [HAMMC08], qu’il était possible de reformuler la loi commande de manière à exprimer ses objectifs et ses paramètres directement dans l’image. Par conséquent, le calcul de la loi de commande proposée requiert uniquement l’estimation des coefficients normalisés B_k issus de la projection de la droite 3D de référence. Une caractéristique importante de cette loi de commande est que ses performances sont indépendantes de la vitesse longitudinale du robot. Celle-ci peut alors être choisie constante ou variable par exemple pour assurer la sécurité du véhicule. Un exemple de résultat expérimental est donné par les Figures 3.8 et 3.9. On observe bien sûr que les erreurs latérale et angulaire sont bien régulées à zéro.

Le robot est parallèle. L’utilisation d’une caméra omnidirectionnelle est très avantageuse dans le cas de la commande d’une plateforme parallèle. Pour s’en convaincre, considérons le problème de la commande de la plateforme de Gough-Stewart (Figure 3.11). A cette fin, [ADM07, AM06] propose des schémas d’asservissement visuel 2D et 3D en observant directement les jambes de la plateforme avec une caméra perspective classique. Malheureusement, il est alors difficile de positionner correctement la caméra afin d’observer simultanément toutes les jambes. Dans

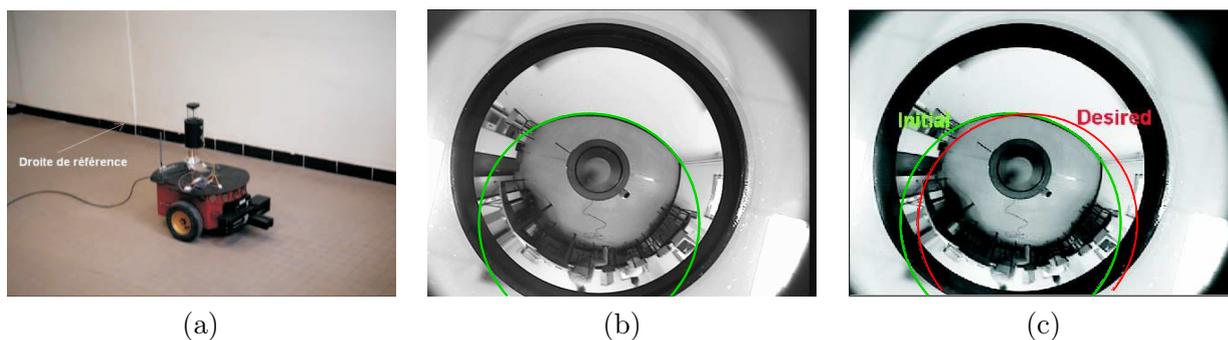


FIG. 3.8 – (a) Robot Pioneer 3 équipé avec une caméra omnidirectionnelle, (b) droite de référence projetée dans l'image initiale, (c) droite de référence dans ses configurations initiale (en vert) et finale (en rouge) projetée dans l'image désirée

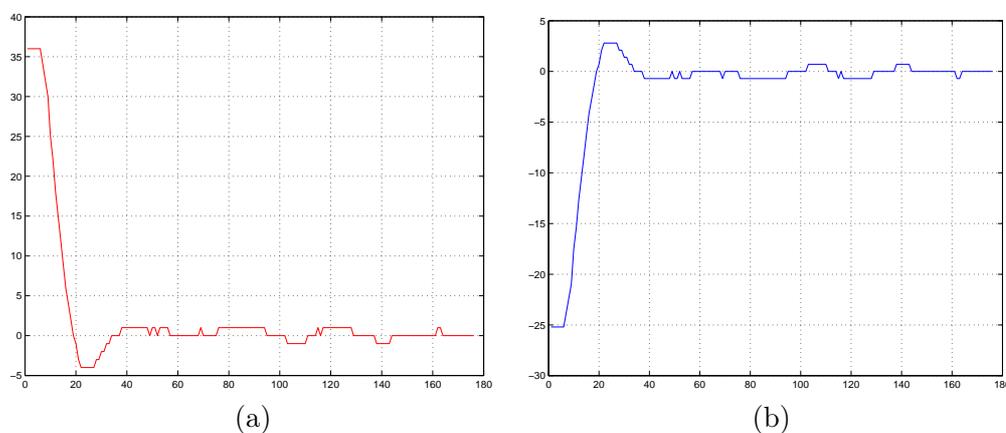


FIG. 3.9 – Erreurs (a) latérale [cm], (b) angulaire [deg]

[ADM07, AM06], la caméra a été positionnée face à la plateforme (voir Figure 3.11.(a)). Dans ce cas, l'extraction des primitives relatives aux jambes arrières peut s'avérer délicate en particulier en raison des auto-occultations (voir la Figure 3.11.(b)). Ceci constitue un inconvénient majeur car la loi de commande par vision suppose que toutes les jambes soient observables pendant l'intégralité de la tâche de régulation. Une première solution à ce problème consiste à observer la plateforme avec plusieurs caméras. Dans ce cas, les données fournies par chacune des caméras doivent être synchronisées et le système multi-caméras doit être calibré. Une solution plus simple et que nous avons proposée dans [TMAM09] consiste à positionner une caméra omnidirectionnelle au centre de la plateforme (voir Figure 3.11.(c)). De cette manière toutes les jambes peuvent être observées simultanément sans occultation possible (voir Figure 3.11.(d)). Si l'on suppose que les jambes de la plateforme sont cylindriques alors leur observation peut se faire à travers la projection des limbes dans l'image. En effet, à chacune des jambes, on peut associer deux droites 3D parallèles (les limbes) se projetant dans l'image en deux coniques (se référer

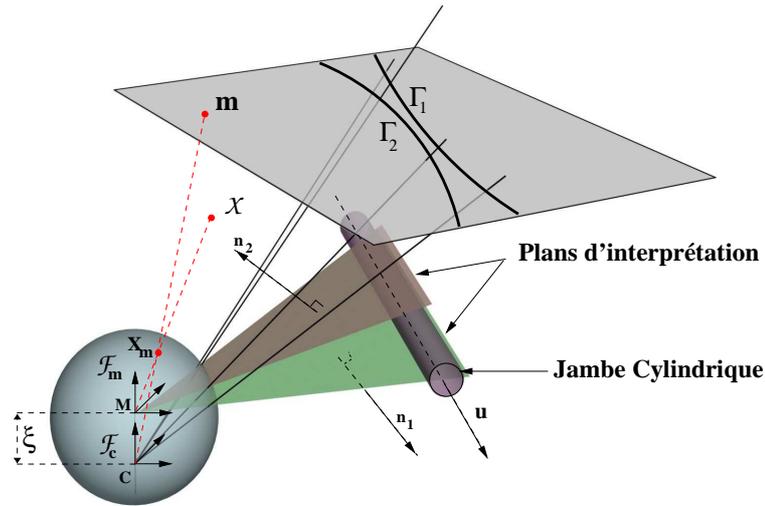


FIG. 3.10 – Une jambe de la plateforme de Gough-Stewart observée par une caméra répondant au modèle unifié : chaque jambe cylindrique se projette en deux coniques dans l'image.

à la Figure 3.10). A partir de l'observation de ces coniques (et donc une nouvelle fois en utilisant les coefficients normalisés B_k relatifs à la projection des limbes), il est alors possible de construire des vecteurs d'observation représentant l'orientation des jambes où encore les plans d'interprétation des limbes. Les commandes associées peuvent alors être construites de manière similaire à celles proposées dans [ADM07]. Un exemple de résultats est donné par les Figures 3.12 et 3.13. Dans cet exemple, on donne l'évolution des erreurs du vecteur d'observation dans deux cas : dans le premier cas l'observation est constituée de l'orientation des jambes, dans le second cas elle est constituée des coordonnées des plans d'interprétation des limbes. Comme nous l'avons montré dans [TMAM09], d'une part ces observations sont mesurables directement dans l'image omnidirectionnelle et d'autre part les meilleurs résultats sont obtenus en régulant les plans d'interprétation des limbes plutôt que les directions des jambes. Les résultats expérimentaux ont également confirmé que l'utilisation d'une caméra omnidirectionnelle (plutôt qu'une caméra conventionnelle) permettait d'améliorer de manière sensible le comportement global du système.

Asservissement visuel 2D à partir de moments

Le choix des informations visuelles à utiliser dans les schémas de commande 2D reste une question largement ouverte dans le domaine de l'AV2D. Comme nous l'avons déjà mentionné, la complexité de la matrice d'interaction et donc le comportement en boucle fermée du système à réguler dépend très intimement de ce choix. Si la matrice d'interaction est trop complexe (fortement non-linéaire et couplée) alors l'analyse *a priori* des propriétés de la commande devient très difficile et le comportement du système est souvent imprévisible et inadéquat pour des déplacements importants. En continuité avec les travaux décrits dans [Tah04] et afin de réduire le

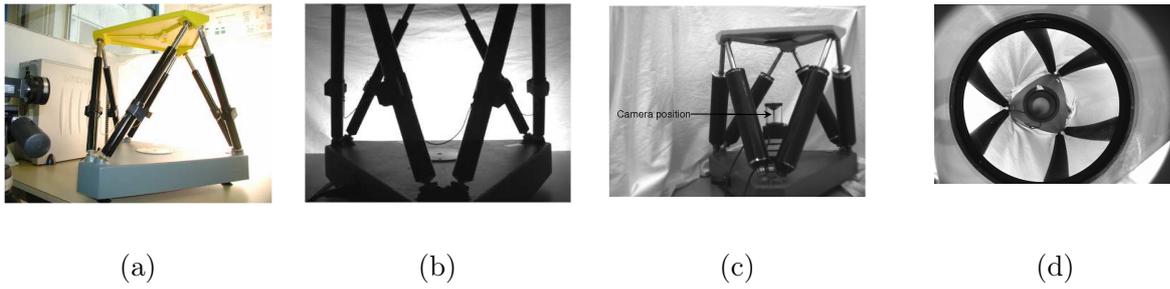


FIG. 3.11 – La plateforme de Gough-Stewart observée par une caméra conventionnelle : (a) vue externe, (b) image acquise par la caméra. La plateforme de Gough-Stewart observée par une caméra omnidirectionnelle : (c) vue externe (d) image acquise par la caméra



FIG. 3.12 – Exemple de résultat de commande par vision de la plateforme de Gough-Stewart : (a) et (b) configuration et image initiales, (c) et (d) configuration et image désirées.

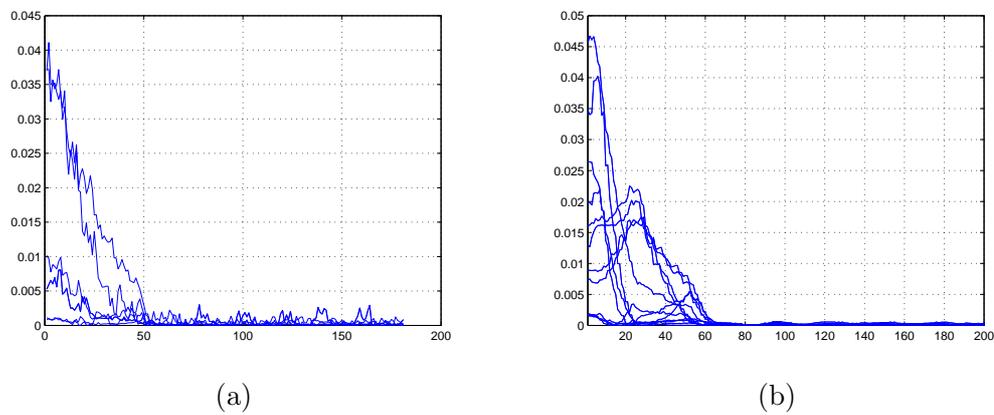


FIG. 3.13 – (a) Résultats obtenus avec l'observation de l'orientation des jambes (erreurs sur le vecteur d'observation) ; (b) Résultats obtenus avec l'observation des plans d'interprétation des limbes (erreurs sur le vecteur d'observation).

couplage observé entre les différentes composantes de la matrice d'interaction, nous avons proposé l'utilisation de nouvelles primitives visuelles basées sur les moments de la projection sur la sphère unité (intervenant dans le modèle unifié) d'un ensemble de points ou de contours fermés [TCM08, TMCC09, TMCA09]. Alors que les travaux précédents sur les moments concernaient uniquement des objets plans observés par une caméra classique, nos résultats s'appliquent également aux objets non-plans observés par une caméra obéissant au modèle unifié décrit dans le chapitre 2. De manière plus précise, afin de contrôler les mouvements de translation indépendamment des mouvements de rotation, nous avons proposé l'utilisation de deux nouvelles primitives invariantes aux rotations :

- $S_t = \frac{1}{\sqrt{S}}$, où S représente la surface de la projection de la cible sur la sphère unité,
- $P_t = \frac{1}{\sqrt{P}}$ où P est un polynôme de degré deux dont les paramètres sont des moments d'ordre deux de la projection de la cible sur la sphère unité.

Outre les propriétés d'invariance vis à vis des rotations, on notera également que les matrices d'interaction associées à ces primitives ont le bon goût de peu varier avec la profondeur de la cible.

Exemples de résultats. Le résultat présenté sur les Figures 3.14 et 3.15 a été obtenu en utilisant notre plateforme cartésienne équipée d'une caméra munie d'une optique fisheye. Quatre invariants S_t calculés pour quatre triangles définis à partir de la combinaison de points de la cible sont utilisés afin de contrôler les mouvements de translation. Les configurations initiale et désirée de la caméra sont liées par une rotation pure de $80dg$ autour de l'axe optique. Les images initiale et désirée correspondantes sont données par la Figure 3.14. Comme attendu, on observe que les vitesses de translation sont quasiment nulles (voir Figure 3.15).

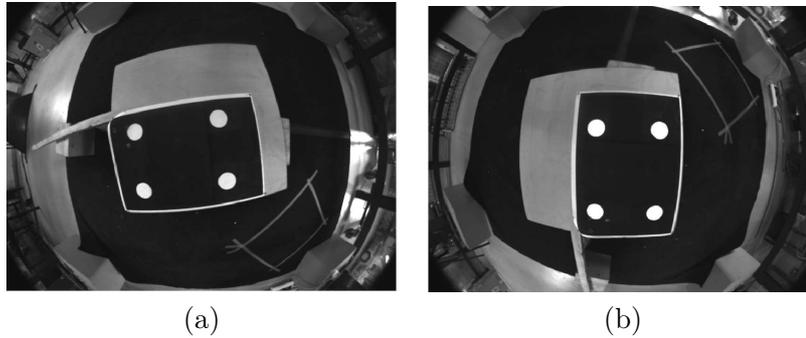


FIG. 3.14 – (a) Image désirée, (b) image initiale

3.3.2 Asservissement visuel basé homographie

Comme évoqué précédemment, les propriétés de la commande sont directement liées à l'interaction \mathbf{L}_s entre les mesures \mathbf{s} et les mouvements de la caméra. En particulier, le couplage entre les différents degrés de liberté dépend de la forme de la matrice d'interaction. Par exemple, le lien entre les variations des coordonnées d'un point dans l'image et le déplacement de la caméra

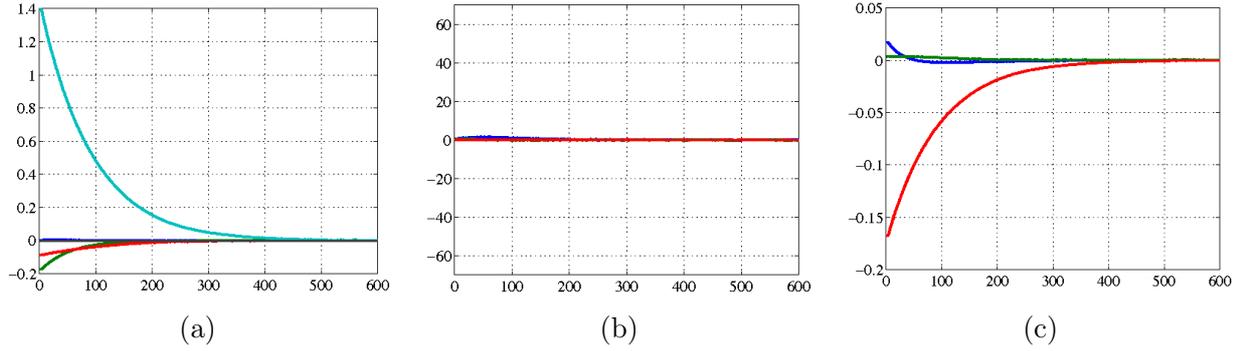


FIG. 3.15 – Résultats obtenus pour une rotation pure de $80dg$ autour de l'axe optique : (a) Erreurs, (b) vitesse de translation (mm/s), (c) vitesse de rotation (rad/s).

est fortement non linéaire et présente de forts couplages entre les différents degrés de liberté. Cela explique largement les difficultés pour réaliser des déplacements importants avec ce type de primitive. De nombreuses approches ont été proposées afin de résoudre ces difficultés. Nous nous intéressons dans cette section plus particulièrement aux approches basées sur le calcul du déplacement de la caméra à partir de la transformation homographique liant deux images d'une cible plane. Nous avons étendu ces approches connues pour les situations où l'on dispose d'un ensemble de points observé par une caméra perspective aux cas plus généraux des caméras centrales observant des points ou des droites. Dans la suite, nous définissons le vecteur de primitives visuelles de la manière suivante :

$$\mathbf{s} = \begin{bmatrix} \tilde{\mathbf{s}} & \theta \mathbf{u}^\top \end{bmatrix}^\top$$

Le vecteur $\tilde{\mathbf{s}}$, de dimension 3, varie lorsque la caméra a un mouvement de translation. Comme nous le verrons, il peut être choisi dépendant ou non des mouvements de rotation de la caméra. Le vecteur $\theta \mathbf{u}$, représente les déplacements en rotation de la caméra et est, quant à lui, indépendant des mouvements de translation. Par conséquent, la matrice d'interaction \mathbf{L}_s de dimension 6×6 relative au vecteur \mathbf{s} s'écrit :

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_{\tilde{\mathbf{s}}_v} & \mathbf{L}_{\tilde{\mathbf{s}}_\omega} \\ \mathbf{0}_3 & \mathbf{L}_\omega \end{pmatrix}$$

On peut noter que dans le cas où $\tilde{\mathbf{s}}$ est indépendant des mouvements de rotation de la caméra alors $\mathbf{L}_{\tilde{\mathbf{s}}_\omega} = \mathbf{0}_{3 \times 3}$ et \mathbf{L}_s est une matrice bloc-diagonale. D'autre part, on trouvera la forme de \mathbf{L}_ω dans [Mal98].

On suppose dans la suite que l'on peut estimer une transformation homographique liant les images courante et désirée d'un ensemble de points ou de droites (se reporter au chapitre 2).

Dans un premier temps la caméra observe un ensemble de points.

Par analogie avec le schéma 2D1/2 proposé dans [MCB99], un premier choix possible pour $\tilde{\mathbf{s}}$ (choix **A**) est :

$$\tilde{\mathbf{s}} = \begin{bmatrix} \tilde{\mathbf{s}}_1^\top & \tilde{\mathbf{s}}_2 \end{bmatrix}^\top$$

$\tilde{\mathbf{s}}_1 = [x \ y]^\top$ représente ici les coordonnées d'un point dans le plan image et $\tilde{\mathbf{s}}_2 = \log(\rho/\rho^*)$ où ρ représente la norme du vecteur composé des coordonnées d'un point 3D exprimées dans \mathcal{F}_m (on rappelle que $\tilde{\mathbf{s}}_2 = \log(\rho/\rho^*)$ peut être calculé en utilisant la relation (2.19)). Les deux premières lignes de la matrice d'interaction relative à $\tilde{\mathbf{s}}$ sont données par les matrices (3.17) et (3.18) et la dernière ligne par :

$$\mathbf{L}_{\tilde{\mathbf{s}}_2} = \frac{1}{\sigma\rho^*} \begin{pmatrix} -\Phi x & -\Phi y & \Phi \frac{\xi^2(x^2+y^2)-1}{1+\gamma\xi} & 0 & 0 & 0 \end{pmatrix} \quad (3.19)$$

avec $\Phi = \frac{1+\gamma\xi}{\gamma+\xi(x^2+y^2)}$. Les variations de la norme du point 3D sont indépendantes des mouvements de rotation de la caméra (ce que l'on observe également dans l'Equation 3.19). Afin d'exploiter, cette propriété nous avons proposé d'utiliser la norme de 3 points pour contrôler les mouvements de translation (choix **B**) :

$$\tilde{\mathbf{s}} = [\log(\rho_1/\rho_1^*) \quad \log(\rho_2/\rho_2^*) \quad \log(\rho_3/\rho_3^*)]^\top$$

Dans ce cas $\mathbf{L}_{\tilde{\mathbf{s}}_w} = \mathbf{0}_{3 \times 1}$ et :

$$\mathbf{L}_{\tilde{\mathbf{s}}_v} = \begin{pmatrix} \frac{\Phi_1}{\sigma_1\rho_1^*} & 0 & 0 \\ 0 & \frac{\Phi_2}{\sigma_2\rho_2^*} & 0 \\ 0 & 0 & \frac{\Phi_3}{\sigma_3\rho_3^*} \end{pmatrix} \begin{pmatrix} -x_1 & -y_1 & \frac{\xi^2(x_1^2+y_1^2)-1}{1+\gamma_1\xi} \\ -x_2 & -y_2 & \frac{\xi^2(x_2^2+y_2^2)-1}{1+\gamma_2\xi} \\ -x_3 & -y_3 & \frac{\xi^2(x_3^2+y_3^2)-1}{1+\gamma_3\xi} \end{pmatrix}$$

La matrice d'interaction est alors bloc-diagonale, ce qui implique que les mouvements de translation sont découplés des mouvements de rotation. Enfin, il est également possible de retrouver les bonnes propriétés obtenues avec le point 3D ou les primitives proposées dans [MC02a] en utilisant (choix **C**) :

$$\tilde{\mathbf{s}} = \frac{\rho}{\rho^*} \mathbf{X}_s = \sigma \mathbf{X}_s$$

où \mathbf{X}_s est donné en fonction des points dans l'image par la relation (3.5). Dans ce cas la matrice d'interaction s'écrit :

$$\mathbf{L}_{\tilde{\mathbf{s}}} = \begin{pmatrix} -\frac{1}{\rho^*} \mathbf{I}_3 & [\tilde{\mathbf{s}}]_\times \end{pmatrix}$$

Pour les deux derniers choix (**B** et **C**), la seule inconnue dans $\mathbf{L}_{\tilde{\mathbf{s}}}$ est le paramètre ρ^* qui apparaît comme un gain appliqué aux vitesses de translation. Il suffit alors de fixer ρ^* à une valeur strictement positive pour assurer la stabilité globale de la loi de commande.

Exemples de résultats. Les lois de commande présentées ci-dessus ont été validées par de nombreuses expérimentations sur notre plateforme cartésienne équipée d'une caméra munie d'une optique fisheye [AMM09b] ou observée par une caméra catadioptrique [HA06b]. Le résultat présenté ici a été obtenu en utilisant une caméra fisheye montée sur la plateforme cartésienne. Dans ce jeu d'expérimentation, les configurations initiale et désirée de la caméra sont liées par une rotation pure autour de l'axe optique. Les trois lois de commandes présentées ci-dessus sont tout d'abord testées pour une rotation très importante (140 degrés). Dans ce cas, seule la loi de commande utilisant le ratio de la norme de trois points (choix **B**) permet la convergence.

Les deux autres lois de commande (**A** et **C**) amènent le robot en limite articulaire en raison du couplage entre les mouvements de rotation et de translation. Pour celles-ci, nous avons alors réduit le mouvement de rotation à 90 degrés. Les résultats sont présentés Figure 3.16. On peut y noter le découplage total entre les mouvements de rotation et de translation pour la loi de commande **B**.

La caméra observe maintenant un ensemble de droites.

Afin de concevoir un schéma d'asservissement visuel hybride, il est possible d'utiliser les informations issues des droites polaires. En effet, comme nous l'avons vu dans le chapitre 2, il est possible d'estimer une transformation homographique entre deux images de quatre droites \mathcal{L}_k , $k = 1, 2, 3, 4$ appartenant à un plan (π). Pour cela, on peut utiliser les droites polaires de coordonnées dans l'image \mathbf{l}_k , $k = 1, 2, 3, 4$ associées au plan principal par rapport aux Γ_k , $k = 1, 2, 3, 4$ (projections des droites dans l'image). La décomposition de l'homographie relative au plan (π) permet alors l'estimation des vecteurs $\theta\mathbf{u}$ et $\tilde{\mathbf{s}} = [\log(r_1) \log(r_2) \log(r_3)]^\top$ où $r_k = h_k/h_k^*$ $k = 1, 2, 3$ où h_k représente la profondeur de la droite \mathcal{L}_k dans le repère \mathcal{F}_m (voir Equation 2.25). La partie relative aux mouvements de rotation de la matrice d'interaction pour ce vecteur d'observation est nulle (les profondeurs étant des grandeurs invariantes aux rotations). La matrice d'interaction globale est donc une nouvelle fois bloc-diagonale avec :

$$\mathbf{L}_{\tilde{\mathbf{s}}_v} = \begin{pmatrix} \|\mathbf{K}^\top \mathbf{l}_{i1}\| r_{h1} h_1^* & 0 & 0 \\ 0 & \|\mathbf{K}^\top \mathbf{l}_{i2}\| r_{h2} h_2^* & 0 \\ 0 & 0 & \|\mathbf{K}^\top \mathbf{l}_{i3}\| r_{h3} h_3^* \end{pmatrix}^{-1} \begin{pmatrix} (\mathbf{u}_1 \times \mathbf{K}^\top \mathbf{l}_{i1})^\top \\ (\mathbf{u}_2 \times \mathbf{K}^\top \mathbf{l}_{i2})^\top \\ (\mathbf{u}_3 \times \mathbf{K}^\top \mathbf{l}_{i3})^\top \end{pmatrix} \quad (3.20)$$

où \mathbf{u}_k est le vecteur directeur associé à la droite \mathcal{L}_k et exprimé dans \mathcal{F}_m . La matrice d'interaction $\mathbf{L}_{\tilde{\mathbf{s}}_v}$, contrôlant les mouvements de translation n'est singulière que si le centre du miroir du capteur catadioptrique se trouve sur le plan de référence (π). Cette configuration singulière n'est jamais observée en pratique. La matrice d'interaction $\mathbf{L}_{\tilde{\mathbf{s}}_v}$ est donc inversible durant toute la tâche d'asservissement. D'autre part, l'expression précédente laisse apparaître les vecteurs directeurs \mathbf{u}_k de chaque droite \mathcal{L}_k . Cependant, ceux-ci peuvent être exprimés en fonction des coordonnées des droites polaires dans les positions courantes et désirées. la seule inconnue est alors le paramètre h^* qui apparaît une nouvelle fois comme un gain appliqué aux vitesses de translation. Il suffit alors de fixer h^* à une valeur strictement positive pour assurer la stabilité globale de la loi de commande.

Exemples de résultats. Un exemple de résultats obtenus avec cette dernière loi de commande est donné par la figure 3.17. Le déplacement à effectuer dans cette expérimentation est composé d'une translation $\mathbf{t} = [-10 \ -80 \ 60]^\top$ cm et d'une rotation $\mathbf{u}_\theta = [0 \ 0 \ 100]^\top$ degrés. Les profondeurs h_k^* ont été fixées approximativement (notons que la convergence est peu sensible à ce choix). Comme nous le constatons, la composante de rotation autour de l'axe Z est plus importante. On observe également que les deux autres composantes apparaissent faiblement. Cela peut-être expliqué par l'étalonnage très grossier du passage caméra-effecteur. Cependant, un découplage intéressant est tout de même assuré.

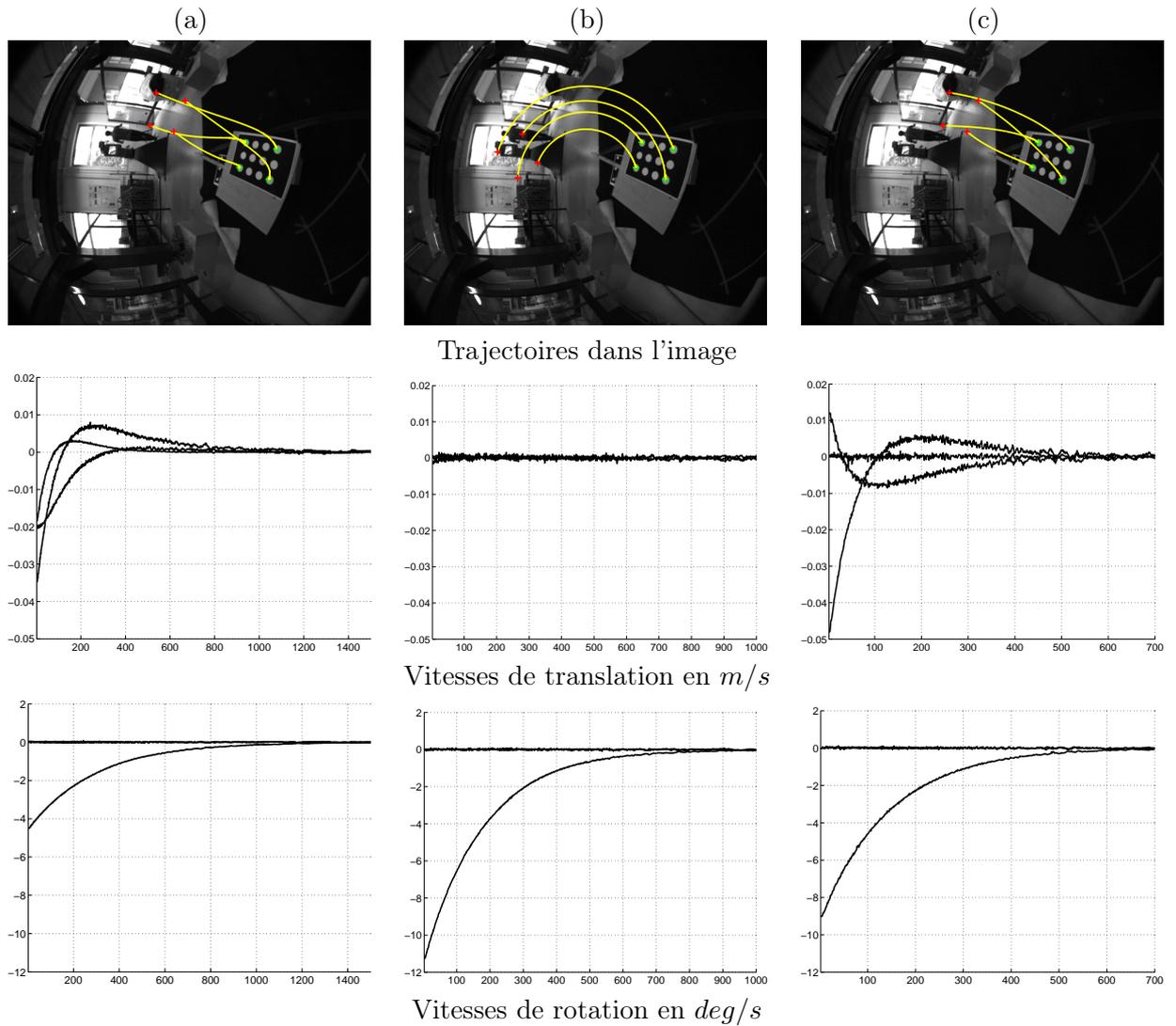


FIG. 3.16 – Comportement des asservissements visuels hybrides pour une rotation pure autour de l'axe optique : (a) choix **A** et rotation de 90 degrés ; (b) choix **B** et rotation de 140 degrés ; et (c) choix **C** et rotation de 90 degrés.

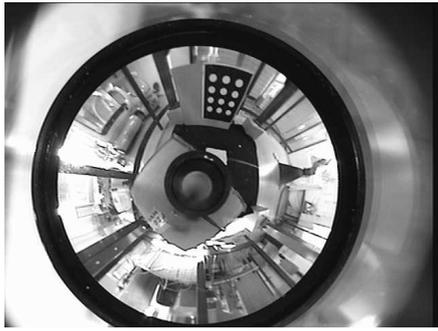


image initiale

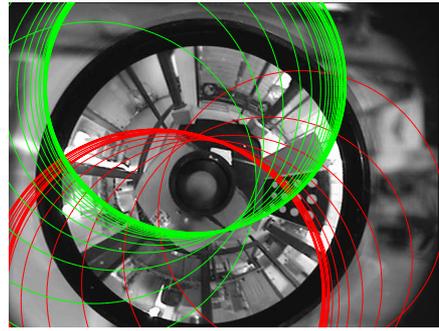
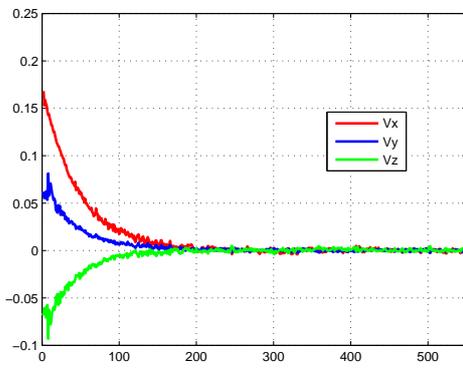
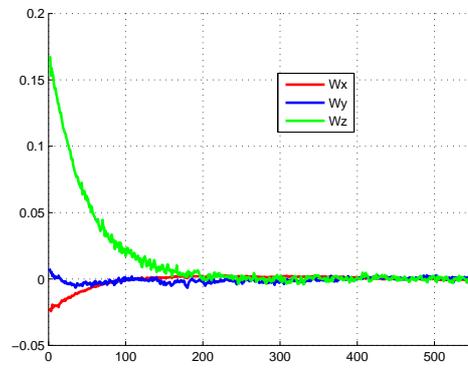


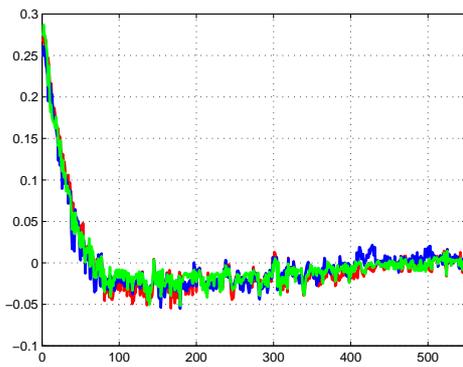
image désirée et trajectoire



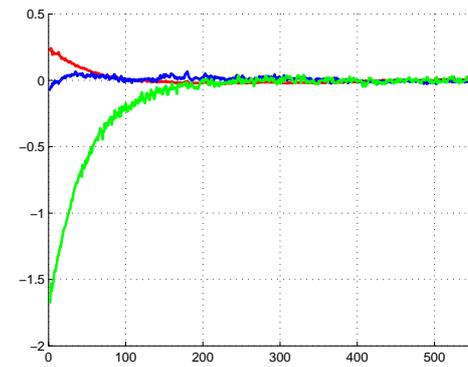
vitesses de translation (m/s)



vitesses de rotation (rad/s)



$\log(h_k/h_k^*)$



erreur d'orientation u_θ

FIG. 3.17 – AV2D1/2 basé droites.

Chapitre 4

Navigation de robots mobiles par mémoire visuelle

Ce chapitre est consacré à nos contributions dans la thématique “navigation de robots mobiles par vision”. Après une partie introductive, il présente la représentation de l’environnement utilisée pour localiser le robot et générer ses mouvements. Il décrit également les méthodes de localisation et de commande basées image proposées et présente des résultats obtenus en utilisant le logiciel SOVIN (Software for VISual Navigation) développé afin d’aborder les sites de navigation de grande taille (échelle d’un campus ou d’une ville par exemple) pour différents types de robots mobiles terrestres (Intérieur : Pioneer 2DX, et 3AT, Extérieur : ROBUCAB) ainsi que pour un robot mobile volant : le drone X4 Flyer du CEA.

Contrairement aux robots manipulateurs, les robots mobiles sont amenés à effectuer des déplacements importants. Apparus dans les années 70, peu après les robots manipulateurs, ceux-ci peuvent être à roues, aériens, sous-marins, marcheurs ou bipèdes. Jusqu’aux années 2000, les efforts de recherche ont essentiellement porté sur les robots à roues plus simples du point de vue mécanique et se déplaçant dans un espace limité à deux dimensions. Aujourd’hui les recherches sur les autres types de robots mobiles sont en plein essor en raison des progrès considérables dans les domaines de l’informatique, de la mécatronique et de l’expérience acquise sur les robots mobiles à roues.

Nos travaux dans ce domaine portent sur la navigation autonome du robot en site a priori inconnu. Cette thématique est très vaste. Elle adresse des problématiques complexes qui suscitent chacune indépendamment l’intérêt de nombreuses équipes de recherche, mais qu’il est difficile de découpler pour bâtir un système autonome complet. Ces problématiques sont classiquement répertoriées selon le triplet *perception - décision - action*, qui synthétise les capacités dont doit disposer le robot pour accomplir une tâche de navigation. Mon objectif scientifiques dans ce cadre est la définition en vue de l’implémentation sur démonstrateur d’une approche complète de navigation allant de l’apprentissage de site à la commande automatique de l’agent mobile.

Fondamentalement, ce problème n'est pas nouveau. Il préoccupe la communauté scientifique roboticienne depuis les balbutiements de la discipline. Jusqu'à présent, les efforts d'investigation du domaine ont surtout porté sur les aspects algorithmiques. L'avènement de systèmes de perception performants, notamment basés sur l'utilisation de caméras numériques et de télémètres lasers, a incité la communauté à se focaliser sur leur exploitation. Il résulte de ceci de nombreux travaux très avancés proposant des modélisations géométriques denses et précises de l'environnement de navigation des robots mobiles embarquant de tels capteurs. Ces travaux permettent d'envisager une localisation précise du robot, ainsi que la planification de ses actions et leur exécution automatique, en se référant à ce modèle numérique de son environnement. Malheureusement, malgré les progrès spectaculaires des matériels informatiques en terme de performance, les résultats de ces travaux ont du mal à s'extirper des laboratoires et des centres de recherche, tant leur mise en œuvre pratique demande des ressources matérielles difficilement intégrables dans un matériel dédié. Notre objectif a par conséquent été de bâtir une approche certes complète, mais suffisamment souple et légère d'un point de vue algorithmique pour être intégrée sur une électronique dédiée destinée à équiper un robot autonome. Bien entendu, le robot étant en mesure de communiquer avec tout équipement informatique relié au même réseau, il peut profiter d'une puissance de calcul déportée. Néanmoins, afin que l'agent robotique soit une entité à part entière et non le capteur mobile d'un système informatique central, nous avons souhaité qu'il soit le plus autonome possible.

La complexité des tâches de navigation pour un robot mobile résulte en particulier de l'étendue spatiale de l'environnement de travail du robot et de la nature dynamique de celui-ci. Cependant, cet environnement de navigation contient généralement des éléments fixes permettant de le caractériser et dont l'observation peut s'avérer très utile à la localisation. Dans le cas d'espaces urbains ou domestiques, sur lesquels nous nous focalisons, l'environnement de travail statique du robot a été façonné par l'homme, lequel lui a donné une structuration forte, qu'il modélise habituellement de manière topographique et géométrique, mais dont il possède aussi un modèle interne, issu de processus complexes de mémorisation de données allothétiques et idiothétiques acquises au fur et à mesure de ces expériences de navigation. La mise en correspondance de sa perception d'une scène réelle observée et des modèles issus de ses expériences passées lui permet de se localiser relativement à la scène observée et de planifier ses déplacements.

Le développement d'une stratégie de navigation nécessite l'utilisation d'une représentation de l'environnement. Cette représentation doit, bien entendu, permettre au robot d'alimenter tous les processus nécessaires à sa navigation, mais peut également s'avérer un élément clef, un vecteur de communication privilégié, entre le robot et l'utilisateur au service duquel il navigue. L'utilisation de capteurs extéroceptifs pouvant conduire à une estimation de la géométrie des éléments structurants de la scène dans laquelle le robot est immergé, a naturellement conduit à la construction de représentations de l'environnement de navigation du robot dont l'intérêt est, non seulement, de répondre aux besoins du système de navigation, mais également de proposer une modélisation de la scène aisément interprétable par un opérateur humain. Une telle représentation peut alors être projetée au centre du dialogue entre le robot et son utilisateur. L'utilisateur peut notamment spécifier un objectif de navigation au robot en désignant directement une situation à atteindre dans la représentation. Le robot est alors à même de proposer une visualisation de sa localisation dans cette même représentation.

Au delà de cet aspect fonctionnel non négligeable, la construction d'une représentation métrique de l'environnement s'avère souvent être un point clef du système de navigation car les bases des travaux de planification de mouvements et de commande des robots mobiles reposent sur l'hypothèse forte que le robot dispose d'un modèle géométrique de son espace navigable, dans lequel sa localisation absolue est connue en permanence, de manière exacte et non bruitée. Malgré l'abondance des travaux pour en assurer la validité, ces hypothèses restent difficilement admissibles en pratique. Le succès des approches référencées capteurs (cela tombe bien puisque c'est notre cheval de bataille) pour le contrôle des mouvements des robots manipulateurs, s'affranchissant d'une localisation géométrique absolue du robot, encourage la transcription des méthodes formalisées dans ce cadre à la navigation des robots mobiles. Comme nous l'avons vu dans le chapitre 3, les approches référencées capteurs consistent à spécifier la tâche robotique dans l'espace d'observation du système de perception. Les informations sensorielles acquises continuellement par le robot alimentent directement une loi de commande destinée à stabiliser ces informations sensorielles sur une consigne de même nature. Les approches référencées capteurs doivent leur popularité à leur robustesse intrinsèque. Spécifiée à l'aide d'informations sensorielles, la tâche robotique ne nécessite pas de modèle complexe de la scène ni de processus d'estimation d'une pose absolue du robot, coûteuse et incertaine, pour être accomplie. Cependant, l'utilisation d'approches référencées capteurs en robotique mobile se heurte à deux problèmes majeurs :

- D'une part, le robot est sujet à de grands déplacements, ce qui induit que les informations sensorielles de consigne ne peuvent pas forcément être mises en correspondance avec les informations sensorielles courantes.
- D'autre part, les robots mobiles se caractérisent en grande majorité par des contraintes de déplacement fortes (non-holonomie, sous actionnement, . . .) affectant leur cinématique. La nature non linéaire de ces contraintes cinématiques impose une remise en cause des approches de commande communément utilisées en robotique manipulatrice.

Au premier de ces deux problèmes, les solutions reposent sur la conception d'une représentation de l'environnement adéquate pour la navigation, de manière à ce que le robot puisse disposer d'une description de la tâche à accomplir sous forme d'un ensemble d'objectifs à atteindre consécutivement, spécifiés dans l'espace d'observation de son système de perception extéroceptif. Le deuxième problème est bien souvent contourné. Par exemple, en dotant les capteurs extéroceptifs embarqués de degrés de liberté par rapport à une base mobile non-holonome, la cinématique non-holonome de celle-ci est intégrée dans une modélisation conférant aux capteurs une cinématique holonome. La tâche de contrôle des déplacements du robot consistant à réguler les informations capteurs sur une référence, la non-holonomie de la base mobile n'affecte alors pas la tâche. En contrepartie, dans cette approche la pose de la base mobile n'est pas explicitement contrôlée.

L'approche de navigation présentée dans ce chapitre a pour objectif de proposer une solution viable en réponse aux problématiques énumérées précédemment. Elle contribue aux approches de navigation se référant à une carte topologique de l'environnement en définissant une représentation originale de l'environnement appelée "mémoire sensorielle". Ce modèle de l'environnement de navigation du robot est fondé sur la mémorisation d'informations extéroceptives issues de capteurs embarqués, lorsque le robot découvre son espace de travail sous l'égide d'un guide humain. Les problématiques scientifiques sous-jacentes relèvent classiquement de la perception de l'environnement pour la localisation, de la planification des déplacements du robot et de leur

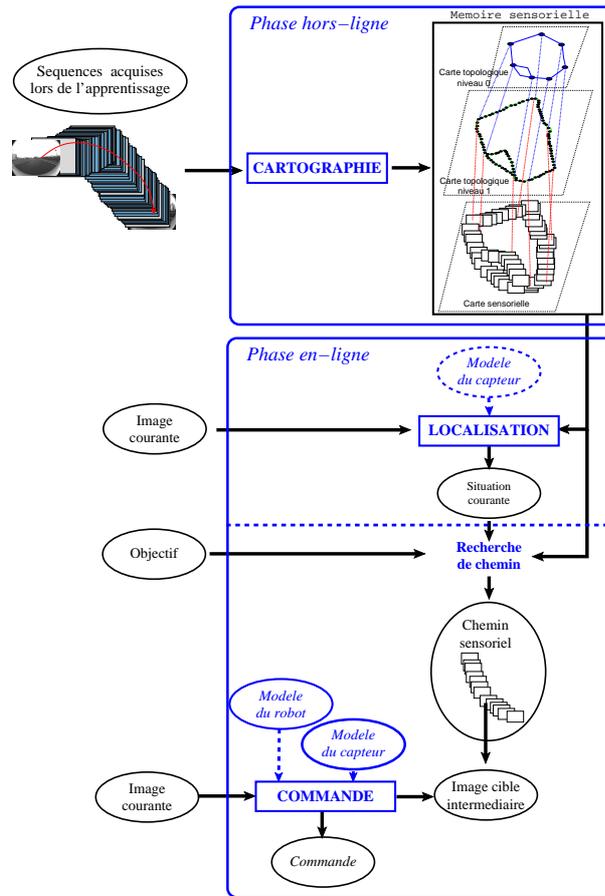


FIG. 4.1 – Overview of the proposed framework.

exécution. Les champs disciplinaires relatifs sont vastes. Nous nous sommes par conséquent focalisés sur la définition d’une approche cohérente dans son formalisme pour nous concentrer plus particulièrement sur le guidage du robot dans sa mémoire sensorielle. Nous avons adressé plus précisément le cas d’une mémoire visuelle, cas particulier d’une mémoire sensorielle constituée exclusivement d’images issues d’une caméra embarquée. Une vue générale de notre stratégie de navigation est présentée Figure 4.1. Elle peut-être divisée en trois étapes :

1. **Construction de la mémoire visuelle :** pendant un déplacement téléopéré, une séquence d’images est acquise par une caméra montée sur le robot. Afin de réduire la complexité de la séquence d’images, seules des images “clés” formant une “trajectoire visuelle” sont stockées. L’ensemble des chemins visuels forme notre carte de l’environnement (mémoire sensorielle). Dans la section 4.1, ce point sera plus précisément évoqué.
2. **Localisation initiale :** avant le déplacement autonome, une localisation du système robotique est réalisée. Celle-ci consiste à trouver l’image de la mémoire visuelle la plus “ressemblante” à l’image courante. Nous reviendrons sur ce point dans la section 4.2.
3. **Navigation autonome :** étant donnée une image cible contenue dans la mémoire senso-

rielle, la mission de navigation est définie par la concaténation d'un ensemble de trajectoires visuelles (chemin visuel) amenant le robot de sa configuration initiale à la configuration cible. Le robot est ensuite contrôlé le long du chemin visuel en utilisant des lois de commande basées vision adaptées à la nature du robot. Nos stratégies de commande seront évoquées dans la section 4.3.

La section suivante a pour objectif de présenter le concept de mémoire sensorielle. Notre approche se basant sur l'utilisation d'une représentation globale de l'environnement de navigation, identifiable par le terme générique de "carte", cette section se focalise par conséquent sur ce type de représentation. Nous définissons ensuite la structure de la mémoire sensorielle et détaillons les éléments clefs qui régissent son évolution et son exploitation.

4.1 Représentation de l'environnement par une mémoire sensorielle

Une stratégie globale de navigation consiste à planifier délibérément les déplacements du robot dans son environnement de navigation, en se référant à une représentation globale de la structure spatiale de celui-ci. Cette représentation prend généralement le nom de carte. Une carte doit par conséquent supporter trois fonctionnalités essentielles :

- La cartographie : la carte offre une représentation de l'espace navigable du robot. Cette représentation fige une mémorisation de la structure spatiale de l'environnement. Elle résulte d'une phase d'apprentissage, ou cartographie, collectant et intégrant des informations perceptuelles caractérisant l'état de l'environnement à l'instant d'observation. Dans le cadre de la robotique mobile autonome, ces informations perceptuelles sont généralement issues du système de perception embarqué.
- La localisation : à partir de l'observation de son état dans son environnement, le robot doit pouvoir déterminer son attitude dans la carte qu'il exploite.
- La planification : étant donnée sa localisation courante et désirée dans la carte, la planification permet au robot de prévoir les déplacements à effectuer pour atteindre son objectif.

Ces trois fonctionnalités sont très clairement interdépendantes. L'existence d'une carte, résultant d'une cartographie, et d'un processus de localisation dans cette carte sont deux prérequis sans lesquels la planification n'a pas de sens.

4.1.1 Cartographie

La cartographie peut être considérée indépendamment des autres processus. La carte peut résulter d'une modélisation préalable de l'environnement de navigation, que l'on confie au robot comme support de navigation. La phase de localisation doit alors permettre au robot de déterminer sa position dans ce modèle connu, au moyen des observations dont il dispose grâce à son système de perception. Il est fréquent de disposer de telles cartes quand le robot est destiné à naviguer en environnement très structuré. En environnement urbain, elle peut, par exemple, résulter de relevés topographiques. De la même manière, un plan d'architecte peut déboucher sur une carte modélisant la structure spatiale d'un bâtiment. Dans ce cadre, la carte modélise la

structure spatiale temporellement invariante d'un environnement de navigation. Le succès d'une localisation repose alors sur la capacité du processus dédié à distinguer les éléments de cette structure parmi les informations issues du système de perception du robot dans son environnement local. Le modèle exploité doit, de plus, s'avérer le plus fidèle possible à la réalité de l'environnement de navigation.

La rigidité de cette contrainte conduit à considérer la cartographie comme un processus intégrant les informations sensorielles acquises par le robot lorsque celui-ci explore ou navigue dans son environnement. Cela conduit clairement à un accroissement de l'autonomie du système robotique mais les phases de cartographie et de localisation deviennent alors fortement interdépendantes. D'une part, l'existence même d'une carte est indispensable à la localisation. D'autre part, l'incrémentement d'une carte incomplète repose sur la localisation préalable du robot afin d'ajouter convenablement à la carte en construction les nouvelles informations sensorielles acquises. Il s'agit par conséquent de traiter simultanément de localisation et de cartographie, comme d'un seul et unique problème, identifié dans la littérature, très dense à ce sujet, par l'acronyme SLAM (Simultaneous Localization And Mapping) proposé par Leonard et Durrant-Whyte [LDW91].

Une carte est un support de localisation absolue. L'état des éléments qui la constituent peut être décrit dans un référentiel unique. Les cartes se déclinent en deux grandes classes principales, les cartes métriques d'une part et les cartes topologiques d'autre part, offrant une structure de mémorisation des informations perceptuelles sur laquelle une stratégie de navigation, qu'elle soit topologique ou métrique, peut s'appuyer [Fil01, Thr02, FM03a, FM03b].

Une carte métrique positionne dans un référentiel géométrique absolu les éléments d'une scène pouvant constituer un obstacle pour le robot. Les travaux relevant de cartographie en milieu structuré, notamment en environnement intérieur, modélisent généralement l'environnement de navigation sur une base euclidienne à deux dimensions. Ce parti pris s'avère peu limitatif dans un tel contexte, où l'hypothèse d'évolution du robot dans un monde plan horizontal peut être acceptée en pratique.

La construction d'une telle carte repose généralement sur l'exploitation conjointe des deux sources d'informations complémentaires que sont les capteurs proprioceptifs et extéroceptifs du robot. Les capteurs proprioceptifs rapportent une mesure caractérisant un état interne du robot, indépendant du monde dans lequel il est immergé. En robotique mobile, le plus répandu de ces capteurs est l'odométrie, qui offre une mesure des vitesses de roues du robot. Son exploitation la plus courante consiste à déduire de ces mesures un déplacement du robot dans un plan. Elle offre par conséquent un moyen simple de localiser le robot dans son plan d'évolution relativement à un référentiel absolu.

Les obstacles de l'environnement de navigation sont intégrés à la carte par le biais de modèles métriques associés aux capteurs extéroceptifs du robot. Ces modèles réalisent la conversion des données brutes de ces capteurs dans l'espace euclidien support de la carte. Les capteurs télémétriques réalisent par définition cette conversion. Leur utilisation est par conséquent très répandue, sous forme de sonars et de lidars notamment. Les caméras qui, dans ce contexte, sont fréquemment employées en paire stéréoscopique, sont également très prisées mais nécessitent des efforts algorithmiques importants pour délivrer une image de profondeur de la scène observée. Ces capteurs permettent d'extraire de l'image de l'environnement local qu'ils fournissent un en-

semble de primitives géométriques le caractérisant. Ces primitives se présentent essentiellement sous forme de points, dont l'ensemble est généralement segmenté en primitives d'ordre supérieur (segments, coins, polygones), afin d'obtenir une représentation plus continue de la forme des obstacles détectés. Chaque obstacle peut alors être décrit par les coordonnées dans le plan des points constituant son contour détecté.

Une autre approche de représentation de l'espace navigable par le robot consiste à discrétiser le plan d'évolution de celui-ci sous forme d'une grille, nommée grille d'occupation, dont chaque cellule se voit associée une probabilité de présence d'obstacle. Cette probabilité résulte de la prise en compte d'une incertitude des mesures capteurs pour lesquelles la cellule concernée est visible.

Une carte topologique décrit un environnement de navigation par un ensemble de lieux caractéristiques de celui-ci. L'accessibilité d'un lieu depuis un autre est symbolisée par un lien topologique. La carte ne modélise donc pas l'environnement en terme d'espace libre, mais mémorise l'ensemble des lieux atteignables par le robot et l'ensemble des chemins permettant de se déplacer d'un lieu vers un autre. Un lieu se matérialise par un ensemble de données extéroceptives. Reconnaître un lieu, c'est mettre en correspondance les données extéroceptives perçues avec celles des autres lieux de la carte. Détecter un nouveau lieu, c'est percevoir un ensemble de données extéroceptives qui appelle l'intérêt. L'objet mathématique support d'une telle représentation est un graphe, dont les nœuds sont les lieux de la carte, et dont les arêtes symbolisent l'accessibilité entre deux lieux. Nous rappelons tout d'abord les quelques définitions relatives aux graphes qui nous seront utiles pour décrire ensuite la représentation de l'environnement choisie.

4.1.2 Quelques définitions

Un *graphe* $G = (\mathbf{N}, \mathbf{A})$ est un ensemble fini non nul de nœuds (ou sommets) $\mathbf{N} = \{N_1, N_2, \dots, N_{n_n}\}$ associé à un ensemble (possiblement vide) d'arcs (ou arêtes) $\mathbf{A} = \{a_1, a_2, \dots, a_{n_a}\}$ où chaque arc peut être défini par deux sommets $a_r = a_{s,t} = \{N_s, N_t\}$.

G est un *graphe orienté* ou *digraphe* si tous les éléments de \mathbf{A} sont des arêtes orientées ($a_{s,t} \neq a_{t,s}$).

Un *pseudo-graphe* est un graphe dans lequel au moins deux nœuds sont liés par au moins deux arcs. Un *graphe simple* est un graphe sans boucle dans lequel deux nœuds sont liés par au plus un arc.

Une *chaîne* Γ d'un graphe G est une suite de sommets adjacents. Un *chemin* d'un graphe orienté G est une suite ordonnée de sommets reliés par des arcs (identique à une chaîne dans un graphe orienté).

Un *graphe connexe* est un graphe dans lequel toute paire de sommets distincts est reliée par une chaîne.

Un *graphe pondéré* est un graphe G suppléé d'une fonction $\gamma : \mathbf{A} \rightarrow \mathfrak{R}^+$ qui associe un poids aux arêtes.

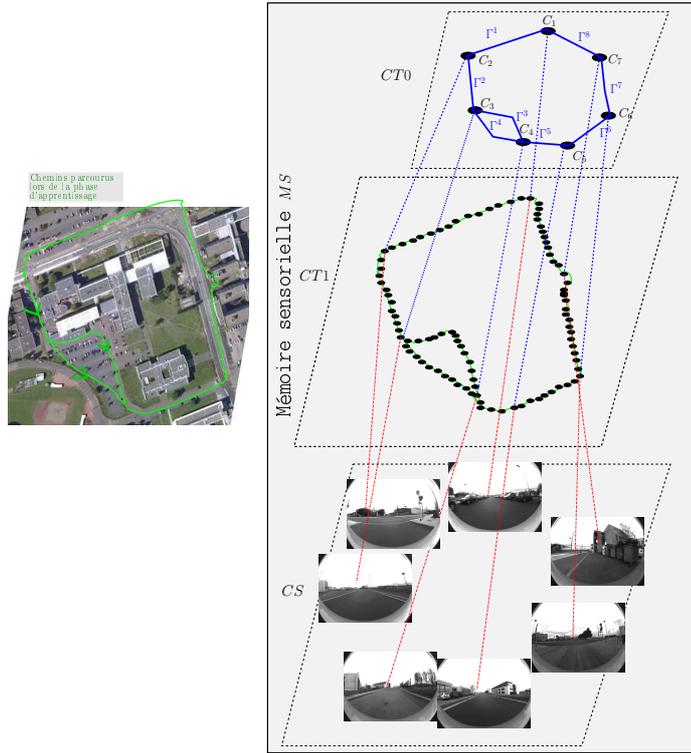


FIG. 4.2 – Mémoire visuelle d'un robot urbain à roues.

4.1.3 Mémoire sensorielle

Nous proposons de représenter l'environnement par une mémoire sensorielle (MS) composée de trois cartes ($MS = \{CS, CT1, CT0\}$) : une carte sensorielle (CS) et une carte topologique à deux niveaux ($CT1$ et $CT0$). Un exemple de mémoire sensorielle dans le cadre de la navigation d'un véhicule urbain à roues muni d'une caméra grand-angle est représenté Figure 4.2. Sur cette figure, seules les images correspondant aux nœuds extrêmes des chaînes simples sont représentées. La carte sensorielle est particulièrement adaptée au formalisme de la commande référencée capteurs et à l'interaction entre le robot et l'utilisateur alors que la carte topologique permet d'introduire des notions de navigabilité et de commandabilité dans MS . Plus précisément :

- La carte sensorielle est définie par un graphe $CS = (N_{CS}, \emptyset)$ avec $N_{CS} = \{\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_{n_N}\}$ où \mathcal{I}_j est la j -ème images clés.
- La carte topologique $CT1$ est définie par un graphe pondéré $CT1 = (N_{CT1}, A_{CT1})$ et la carte $CT0$ par le pseudo-graphe pondéré $CT0 = (N_{CT0}, A_{CT0})$.

Quelques détails concernant $CT1$.

Les nœuds du premier niveau $N_{CT1} = \{N_1, N_2, \dots, N_{n_N}\}$ représentent implicitement les positions du robot lors des prises de vue. A chaque image de la carte sensorielle on fait correspondre

un nœud unique de $CT1$:

$$\begin{cases} CS & \mapsto CT1 \\ \mathcal{I}_i & \rightarrow N_i \quad , \forall i \in \{1, 2, \dots, n_N\} \end{cases}$$

En outre, la présence d'un arc entre deux nœuds de $CT1$ assure la navigabilité et la commandabilité entre les positions correspondantes. Les hypothèses suivantes formalisent ces contraintes :

Hypothèse 1 Soient les repères de commande \mathcal{F}_i et \mathcal{F}_j associés respectivement au robot pour les nœuds N_i et N_j liés par une arête orientée $a = (N_i, N_j)$. Alors, il existe un chemin admissible (Υ) allant de \mathcal{F}_i à \mathcal{F}_j pour le robot dont on connaît les contraintes de déplacement.

Hypothèse 2 Les images \mathcal{I}_i et \mathcal{I}_j correspondant aux nœuds N_i et N_j , liés par une arête orientée $a = (N_i, N_j)$, sont telles qu'il est possible d'estimer l'état alimentant la loi de commande tout au long du chemin (Υ) à parcourir entre \mathcal{F}_i et \mathcal{F}_j .

Le graphe $CT1$ peut également être muni d'une fonction de pondération γ :

$$\begin{cases} A_{CT1} & \mapsto Re^+ \\ a_{i,j} & \rightarrow \gamma(a_{i,j}) \end{cases}$$

Les notions de navigabilité et de commandabilité étant implicitement contenues dans A_{CT1} , il n'est pas nécessaire de les introduire via la fonction de pondération. On peut alors envisager une pondération unitaire ou des fonctions de pondération :

- dépendant des données capteurs (par exemple de la qualité des appariements entre les images \mathcal{I}_i et \mathcal{I}_j),
- dépendant de la commande (par exemple de la qualité de l'estimation d'état à partir des données contenues dans N_i et N_j).

Quelques détails concernant $CT0$.

Afin d'accélérer les traitements lors des phases de planification et de navigation autonome, les nœuds de $CT1$ sont regroupés en séquences (*i.e* en *chemins* ou en *chaînes* lorsque les arêtes de $CT1$ sont non orientées) dans un second niveau topologique $CT0$. De manière plus précise, l'ensemble des nœuds $N_{CT0} = \{C_1, C_2, \dots, C_{n_C}\}$ de $CT0$ est un sous ensemble des nœuds N_{CT1} de $CT1$ représentant des "carrefours". Les n_Γ arcs $A_{CT0} = \{\Gamma^1, \Gamma^2, \dots, \Gamma^{n_\Gamma}\}$ de $CT0$ sont des chemins simples pondérés obtenus par concaténation des arcs de $CT1$ joignant deux nœuds de $CT0$. Par exemple, les images acquises dans un couloir en milieu intérieur ou sur une route entre deux carrefours seront regroupées. En outre, la fonction de pondération de $CT0$ est directement obtenue à partir de celle de $CT1$:

$$\begin{cases} A_{CT0} & \mapsto Re^+ \\ \Gamma^p & \rightarrow \gamma(\Gamma^p) = \sum_{j=1}^{n^p-1} \gamma(a_j) \end{cases}$$

4.1.4 Système de navigation par mémoire sensorielle

Nous rappelons que le processus de navigation proposé peut être décomposé en trois phases : (1) construction de la mémoire visuelle, (2) localisation initiale, (3) navigation autonome.

Construction de la mémoire sensorielle

L'étape de construction de MS se déroule hors ligne lors d'une phase d'apprentissage. Le robot est téléopéré dans son environnement de travail afin d'acquérir une séquence d'images $S = \{\mathcal{I}[i] \mid i \in \{1, 2, \dots, n_S\}\}$. Les trois cartes composant MS sont ensuite construites en trois phases :

1. Un ensemble d'images clés $\{\mathcal{I}[k] \mid k \in \{1, 2, \dots, n_S\}\}$ est sélectionné.
2. Pour chaque image clé $\mathcal{I}[k]$:
 - $\mathcal{I}[k]$ est ajoutée à CS ,
 - un nœud N_k est défini dans $CT1$,
 - si $\mathcal{I}[k]$ n'est pas la première image de CS alors une arête liant N_k au nœud correspondant à l'image précédemment ajoutée est créée dans $CT1$.
3. Le graphe $CT1$ est décomposé en chaînes simples dans $CT0$.

Localisation initiale

Avant le déplacement autonome, une localisation du système robotique est réalisée dans CS . Celle-ci consiste à trouver l'image de la mémoire visuelle la plus "ressemblante" à l'image courante. Dans cette phase le robot est immobile, une localisation par inférence directe doit donc être employée. Les données courantes sont comparées avec celles contenues dans CS via la définition d'une distance d entre deux images. Soient \mathcal{I}_c l'image courante acquise par le capteur extéroceptif et CS la carte sensorielle. La localisation courante du robot par rapport à sa mémoire sensorielle est le nœud de la carte sensorielle \mathcal{I}_l tel que :

$$\mathcal{I}_l = \arg \min_{\mathcal{I}_i \in \mathbf{N}_{CS}} d(\mathcal{I}_c, \mathcal{I}_i)$$

Nous récapitulons, dans la section 4.2, notre approche de localisation lorsque les images sont acquises par une caméra omnidirectionnelle (les détails étant disponibles dans [CMEM08]).

Navigation autonome

Extraction d'un chemin dans la carte sensorielle. On suppose maintenant que la localisation initiale $\mathcal{I}_{i1} = \mathcal{I}_l$ et l'objectif de navigation $\mathcal{I}_{in} = \mathcal{I}_e$ sont connus dans CS . Il est alors possible de construire une chaîne Γ de la mémoire liant \mathcal{I}_{i1} à \mathcal{I}_{in} :

$$\Gamma = \{\mathcal{I}_{ij} \mid j \in \{1, 2, \dots, n\}\}$$

L'extraction du chemin à suivre emploie les cartes topologiques $CT1$ et $CT0$. Connaissant \mathcal{I}_{i1} et \mathcal{I}_{in} , il est aisé de retrouver les nœuds correspondants N_l et N_e dans $CT1$ ainsi que les chaînes Γ^l et Γ^e correspondantes dans $CT0$. La construction de Γ peut alors être conduite via des algorithmes de recherche du plus court chemin bien connus en théorie des graphes tels que les algorithmes de Dijkstra, Bellman–Ford ou Floyd–Warshall.

Suivi de chemin. La carte MS et donc le chemin Γ ne sont pas référencés de manière absolue dans le monde 3D mais plutôt dans l'espace des capteurs. Les mouvements du robot peuvent alors être générés en exploitant les concepts de la commande référencée capteurs. Nous décrirons plus en détail les cas des robots mobiles à roues non-holonomes dans la section 4.3.

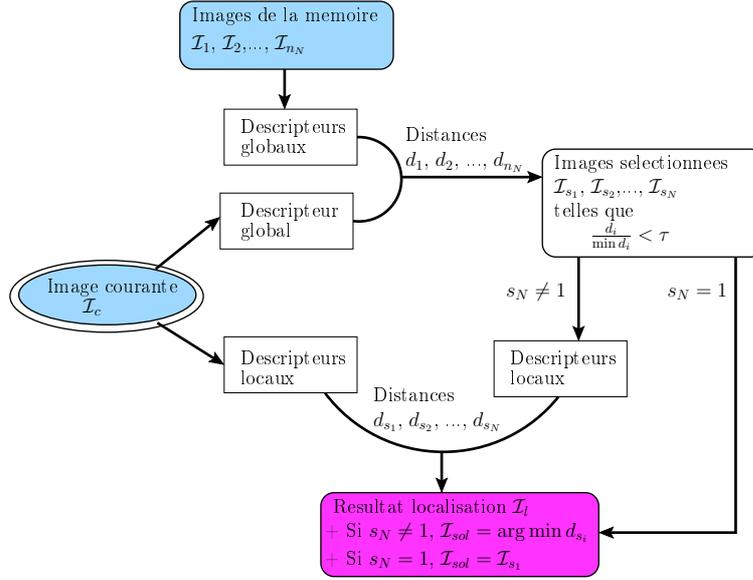


FIG. 4.3 – Approche hiérarchique de localisation.

4.2 Localisation dans une mémoire d'images grand angle

Dans cette section, nous résumons notre stratégie de localisation initiale. Les détails sont présentés dans [CMEM08]. La méthode proposée est hiérarchique dans le sens où elle combine un descripteur global de l'image et des descripteurs locaux sur le voisinage de points de Harris. Cette approche est résumée Fig. 4.3.

Une image en niveau de gris \mathcal{S} peut être vue comme une surface en trois dimensions où la troisième coordonnée correspond au niveau de gris :

$$\mathcal{S} : \begin{cases} [0, 1, \dots, N] \times [0, 1, \dots, M] & \mapsto [0, 255] \\ (u, v) & \rightarrow \mathcal{S}(u, v) \end{cases}$$

La surface $\mathcal{S}(u, v)$ peut ensuite être localement approximée par la surface $f(s, t)$, $s \in [0; 1]$, $t \in [0; 1]$. A cette fin, nous utilisons une fonction d'interpolation cubique appliquée à des points situés sur un maillage triangulaire régulier dans l'image. Ce type de maillage, obtenu en utilisant le générateur de maillage proposé dans [PS04], permet d'obtenir des points approximativement équidistants un à un. Le descripteur global Z est alors défini comme l'altitude des points de contrôle de la surface interpolée. En outre, afin d'augmenter la robustesse face aux changements d'illumination, l'histogramme de l'image est égalisé avant de réaliser l'interpolation. Les descripteurs globaux Z_c et Z_i (relatifs aux images courante \mathcal{I}_c et mémorisée \mathcal{I}_i) sont comparés via la distance $L1$ ($d_i = d(\mathcal{I}_c, \mathcal{I}_i)$). Les images de la mémoire candidates à la localisation sont alors telles que $\frac{d_i}{d_{min}} \leq \tau$ où τ est un seuil fixé par l'utilisateur. Le résultat obtenu est ensuite affiné en utilisant les scores d'appariement (corrélation croisée normalisée (ZNCC)) des points de Harris détectés dans l'image mémorisée et dans les images candidates.

Nous renvoyons le lecteur à [CMEM08] pour la validation expérimentale de cette approche.

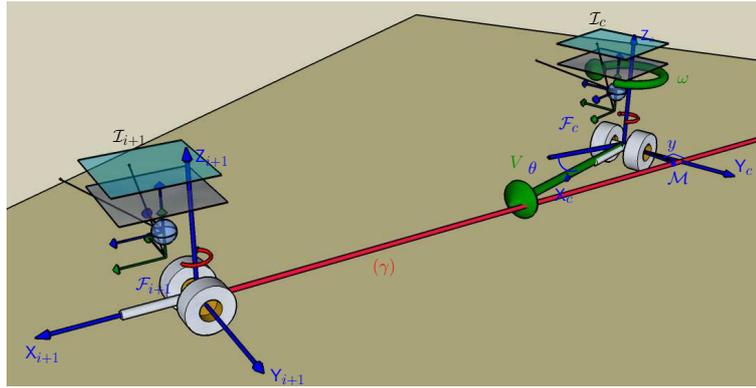


FIG. 4.4 – Suivi de chemin le long d’une droite.

Les résultats présentés dans cet article ont montrés que celle-ci constituait un très bon compromis entre coût calculatoire et précision.

4.3 Suivi de chemin visuel

Comme nous l’avons déjà évoqué, une grande majorité des robots mobiles se caractérise par des contraintes fortes sur leurs déplacements. Une approche possible pour contourner cette difficulté consiste à doter les capteurs extéroceptifs embarqués de degrés de liberté supplémentaires afin que ceux-ci ne soient pas affectés par les contraintes cinématiques de la base mobile (se référer à [PGR91, PGR95, Cad99, FC07] pour des robots mobiles à roues et à [NDC⁺07] pour un drone quadrirotor). De notre côté, nous n’avons pas souhaité équiper le capteur d’un système mécanique complémentaire. Nous avons donc considéré que les contraintes de mouvement de la caméra étaient identiques à celles de la base mobile. Dans ce contexte, nous avons proposé des solutions pour le suivi de chemin visuel par des robots mobiles à roues non-holonomes de type char et bicyclette ainsi que pour les drones sous-actionnés de type quadrirotor. Dans la suite, nous donnons quelques détails concernant les robots mobiles à roues non-holonomes et nous renvoyons le lecteur à [CMGM09] pour plus de détails concernant notre stratégie de commande du quadrirotor.

Nous supposons que le véhicule roule sans glisser sur une surface localement plane et que les paramètres extrinsèques du capteur sont connus. On note ici \mathcal{I}_c l’image courante, \mathcal{I}_{i+1} l’image cible sur le chemin visuel Γ préalablement extrait de la mémoire sensorielle, $\mathcal{F}_c = (O_c, X_c, Y_c, Z_c)$ un repère associé au robot lors de la prise de vue \mathcal{I}_c et $\mathcal{F}_{i+1} = (O_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1})$ un repère associé au robot lors de la prise de vue \mathcal{I}_{i+1} (se référer à la Figure 4.4). L’objectif de commande est que la caméra rejoigne successivement les positions correspondantes aux prises de vue définissant le chemin visuel Γ . Pour cela, notre stratégie consiste à guider \mathcal{I}_c vers \mathcal{I}_{i+1} en régulant asymptotiquement l’axe Y_c sur la droite $(\gamma) = (O_{i+1}, X_{i+1})$ (voir Fig. 4.4). L’objectif de commande est alors atteint si l’axe d’avance courant X_c est régulé sur (γ) avant que l’origine de \mathcal{F}_c atteigne l’origine de \mathcal{F}_{i+1} . Celui-ci peut être réalisé par un véhicule à roues de type char en utilisant la loi de commande suivante [BMM05, CMM09] développée en utilisant le principe des

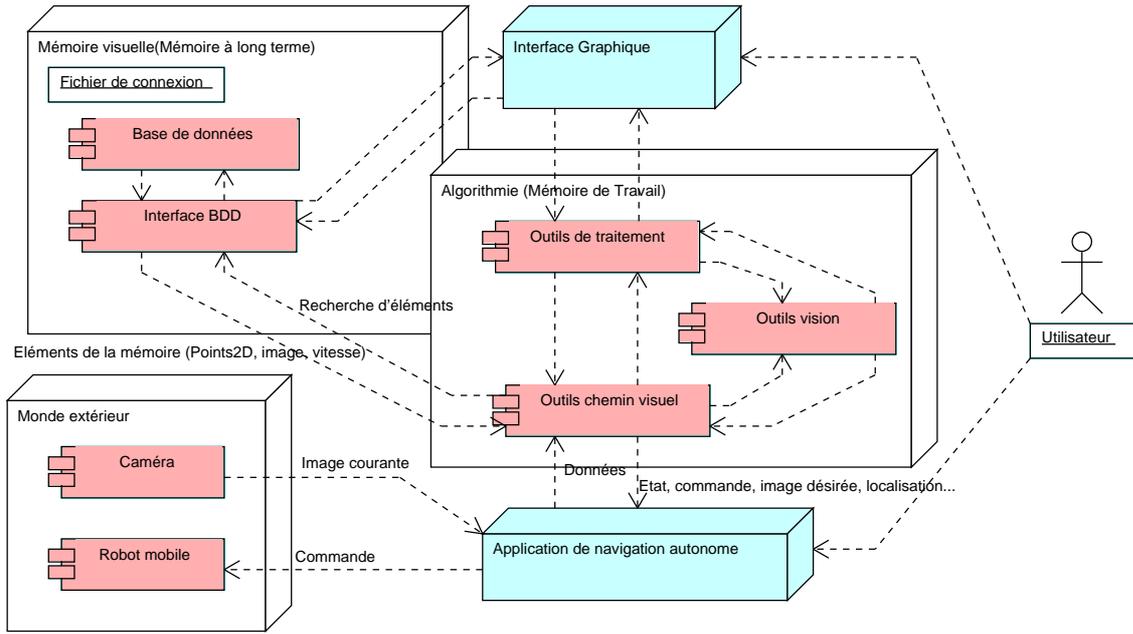


FIG. 4.5 – Implémentation du système complet de navigation par vision.

systèmes chaînés [Sam95] :

$$\omega(y, \theta) = -V \cos^3 \theta K_p y - |V \cos^3 \theta| K_d \tan \theta \quad (4.1)$$

ω et V représentent les vitesses longitudinale le long de l'axe X_c et angulaire autour de l'axe Z_c . Les variables d'entrée θ et y représentent les erreurs angulaire et latérale du robot par rapport à la trajectoire (γ). Elles peuvent être estimées via les méthodes de reconstruction partielle décrites dans le chapitre 2. De manière remarquable, les performances de cette loi de commande sont fixées par une distance d'établissement (*i.e* distance à parcourir avant d'atteindre le chemin à suivre) si la vitesse longitudinale est non-nulle [TBMM04]. K_p et K_d sont alors deux gains positifs qui permettent de fixer ces performances pour une distance d'établissement donnée. Les performances de la loi de commande (4.1) étant fixées par une distance (et non un temps) d'établissement, celles-ci ne dépendent pas de la vitesse longitudinale V . La vitesse longitudinale V peut alors être choisie constante ou variable afin d'assurer des objectifs secondaires (par exemple sécurité et confort des passagers). Notons finalement que, pour les robots mobiles à roues de type bicyclette, une loi de commande similaire peut être obtenue en utilisant les mêmes principes (se référer à [CMM09]).

4.4 Exemple de résultats

D'un point de vue expérimental, notre objectif est d'aborder des sites de navigation de très grande taille (échelle d'une ville). En théorie, la méthodologie décrite précédemment se prête bien à cela car elle ne nécessite la mémorisation que d'une quantité d'informations relativement

faible au regard des ressources actuellement disponibles sur PC. En pratique, cet objectif est cependant difficilement atteignable sans la mise en place d'un outil adapté à la gestion de la mémoire sensorielle. A cette fin, nous avons développé un logiciel nommé SOVIN (Software for Visual Navigation) et dont les principes sont détaillés dans [CMLE08a, LCMM08]. Ce système, représenté Figure 4.5 peut-être décomposé en trois modules. Le premier est dédié à la gestion de la mémoire sensorielle. Le second module intègre les principaux algorithmes de traitements mathématiques et d'images. Enfin, le troisième module réalise l'interface Homme-Machine. Notre méthodologie a été testée avec différents types de robots mobiles en exploitant ce logiciel :

- les robots terrestres d'intérieur de type char 2DX, 3AT conçus par Activ Media et Peeke conçu par Wanny Robotics,
- les robots terrestres d'extérieur de type bicyclette CYCAB et ROBUCAB conçus par Robosoft,
- le robot mobile volant de type quadrirotor X4 Flyer conçu par le CEA.

En outre, notre approche a été validée avec différents types de caméras (conventionnelle, catadioptrique et fisheye). Nous renvoyons le lecteur à [CMM08] pour des exemples de résultats obtenus avec les robots terrestres d'intérieur, à [CMM09] pour des résultats concernant les robots terrestres d'extérieur et à [CMGM09] pour des résultats obtenus avec le quadrirotor X4-Flyer. Nous présentons ici un exemple de résultats obtenus avec un RobuCab (voir Fig. 4.6). Le véhicule est équipé d'un DGPS différentiel cinématique temps réel RTK-DGPS Sagitta de Thales qui a une précision de 1 cm dans le plan horizontal et de 20 cm sur l'axe vertical. La fréquence d'acquisition du DGPS est de 10 Hz. La vitesse d'avance maximale est de 18 km/h. La distance d'établissement est fixée ici à $d_m=15$ m.

Le RobuCab est amené initialement près du chemin *Grand Trajet* quelques jours après la phase d'apprentissage. Il est alors localisé sur Γ^1 et un chemin visuel, contenant près de 1400 images clés et correspondant à un trajet de 1700 mètres, est manuellement spécifié. Le chemin à parcourir (voir Fig. 4.7) comprend des virages importants (par exemple près des lieux *STAPS*, *IFMA* et le virage en U au lieu *Tramway*) ainsi que des dénivelés (par exemple entre les lieux *ISIMA* et *IFMA* puis autour du lieu *IFMA*). La durée de la navigation autonome est de 26 minutes pour cette expérimentation. Les trajectoires suivies par le robot lors des phases d'apprentissage et autonome sont représentées Figure 4.8). Un zoom (Fig. 4.9) permet d'apprécier le comportement du véhicule lors du virage très important en U *Tramway*. On observe que le véhicule revient sur le chemin appris après s'en être écarté au plus fort du virage. Les erreurs latérale et angulaire sont présentées Figure 4.10. L'erreur latérale moyenne entre le chemin appris et le chemin désiré (données GPS) est de 22 cm avec un écart type de 30 cm. 115 appariements robustes ont été calculés en moyenne avec un temps moyen de calcul de 82 ms. Notons finalement, qu'à notre connaissance, ce trajet parcouru de manière autonome eu utilisant uniquement une caméra est le plus important reporté dans la littérature. Afin de comparer les chemins parcourus lors de la phase d'apprentissage et lors de la phase de navigation autonome, les données DGPS sont enregistrées. Bien entendu, lors de la phase de navigation, seul le capteur visuel est utilisé pour commander le véhicule.

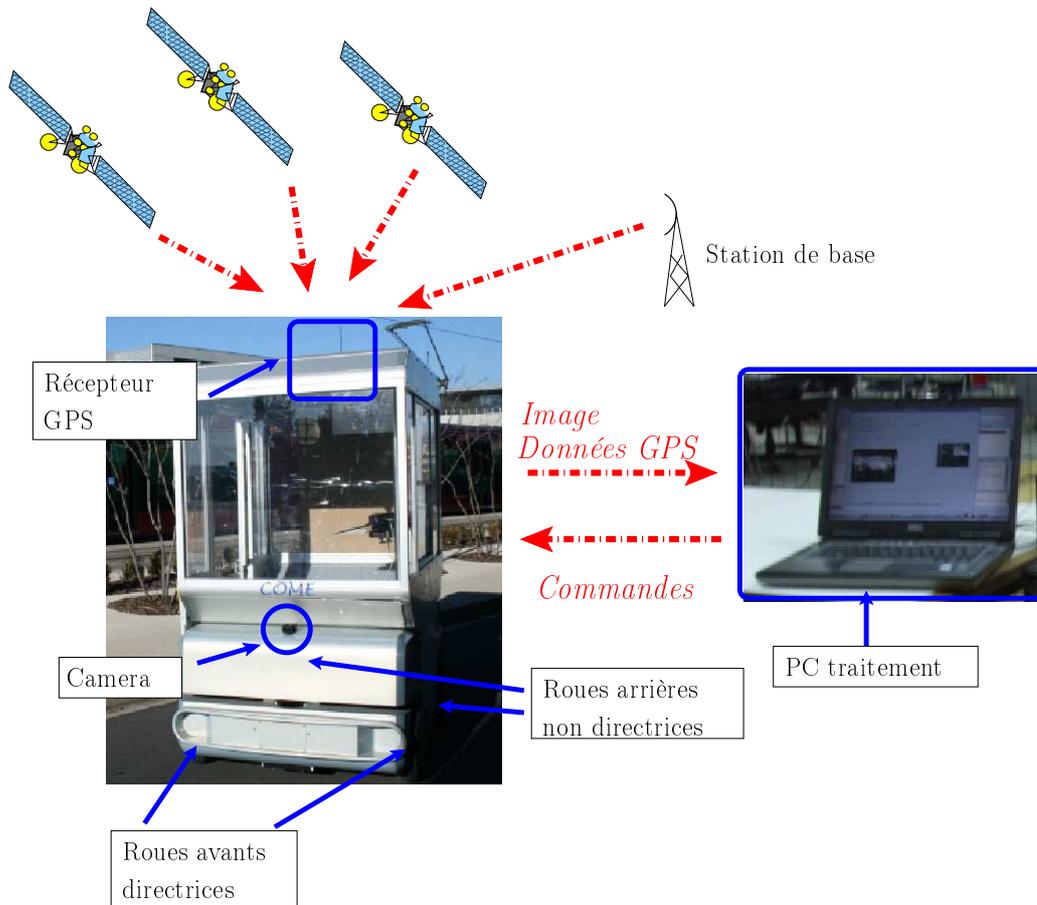


FIG. 4.6 – RobuCab

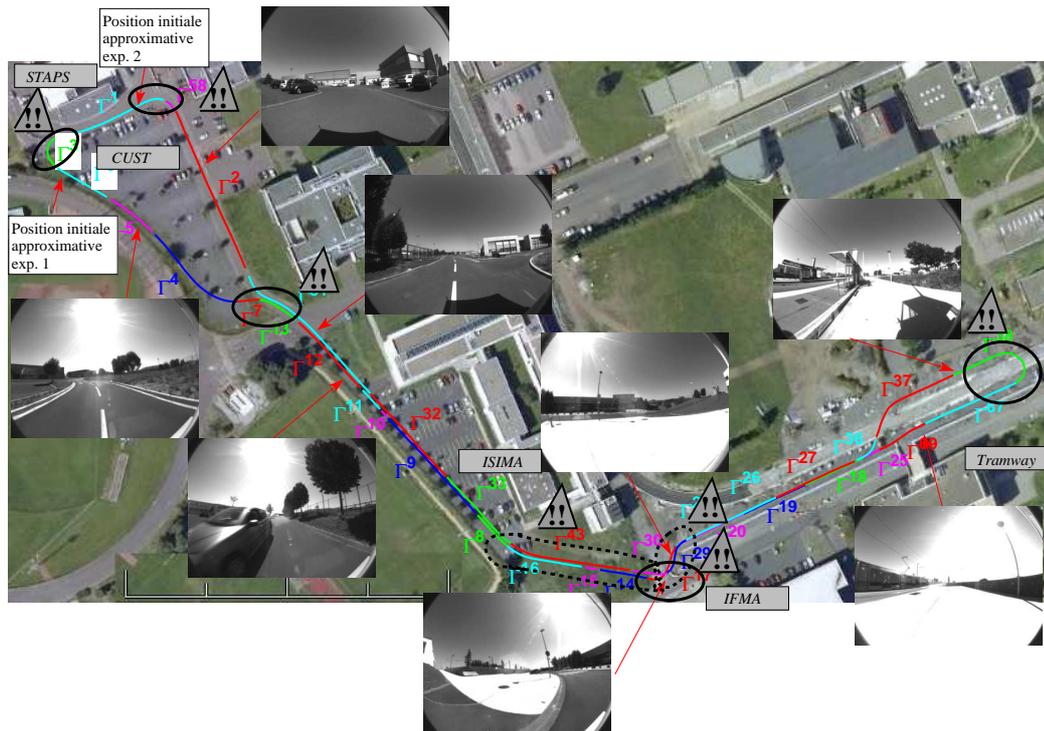


FIG. 4.7 – Chemins contenus dans *MS* pour l'environnement *Grand Trajet*. Ces chemins contiennent des virages importants (ellipses en traits pleins) et des dénivelés (ellipses en pointillés). Quelques images de la mémoire sont également représentées.

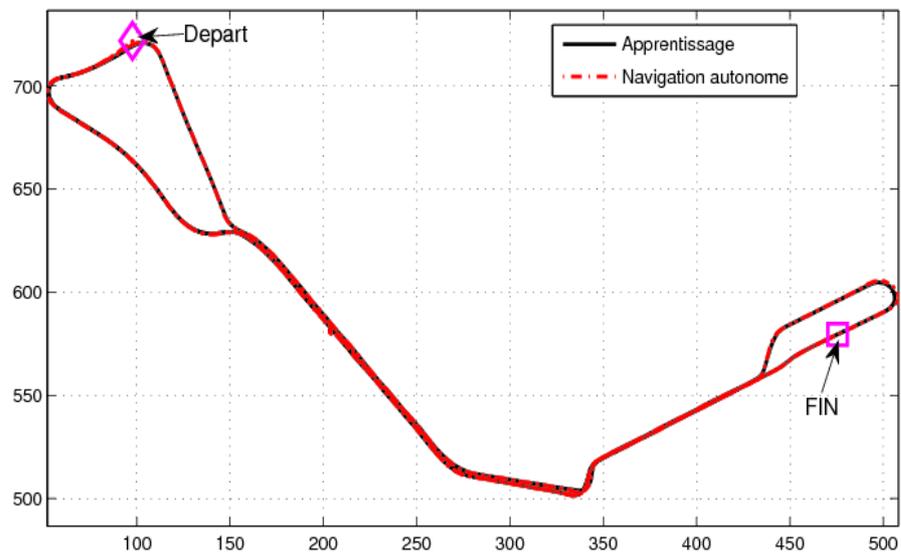


FIG. 4.8 – Suivi de *Grand Trajet* : Chemin appris versus chemin parcouru (en mètres).

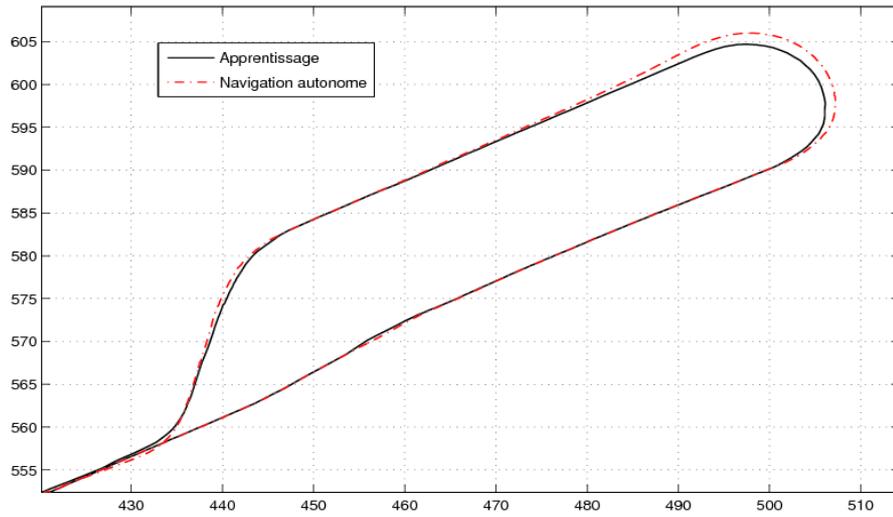
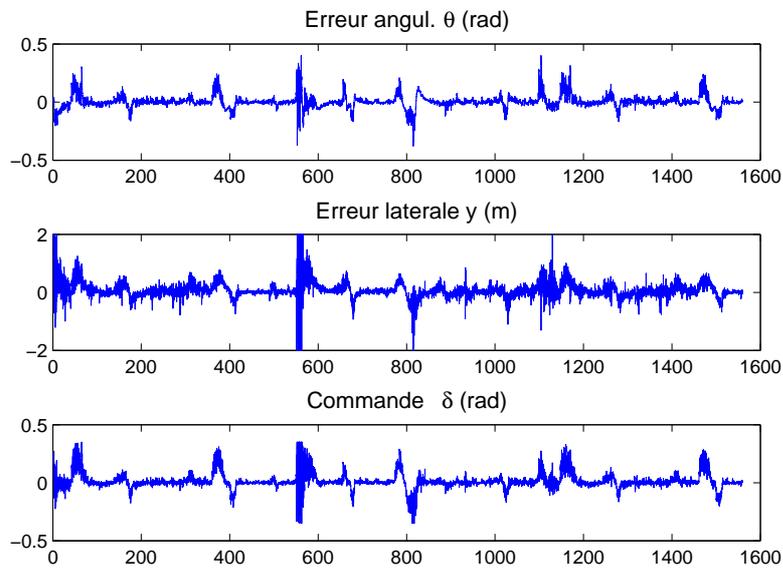
FIG. 4.9 – Zoom sur le virage en U *Tramway* (en mètres)

FIG. 4.10 – Erreurs angulaire et latérale, commande

Bilan et Perspectives

Ce chapitre récapitule nos contributions scientifiques dans les domaines de la commande et de la navigation référencées capteurs. Il présente également les perspectives qui en découlent.

Asservissement visuel

Les contributions que nous avons apportées dans la thématique asservissement visuel ont été en partie présentées dans le chapitre 3. Ces recherches ont été alimentées d'une part par des problèmes théoriques transversaux (analyses de la robustesse, de la stabilité et du couplage) et d'autre part par des aspects pratiques et expérimentaux (nouvelles technologies de capteurs, évolution des applications robotiques).

Le premier point que nous avons abordé concerne le choix de l'inverse généralisé de la matrice d'interaction. A cette occasion, nous avons développé deux alternatives à l'inverse de Moore-Penrose. Celles-ci (Jacobienne étendue et ESM modifié) conduisent à une approximation au second ordre des variations de déplacement de la caméra connaissant les variations des primitives visuelles. De ce fait, leur utilisation dans la commande a permis d'améliorer de manière sensible le comportement spatial de la caméra. On a par exemple observé une disparition complète du phénomène de retrait pour des rotations pures autour de l'axe optique pour l'asservissement visuel 2D avec des points. On notera également que les singularités relatives à l'ESM mises en évidence dans [MC09] ne sont pas partagées par nos deux propositions. Bien entendu, cela ne prouve pas l'absence de singularité.

Le second point que nous avons développé est relatif à l'asservissement visuel lorsque la caméra obéit au modèle de projection unifié (situation englobant les cas où une caméra conventionnelle est employée). Dans ce contexte, nous avons obtenu des résultats relatifs à la robustesse de l'AV2D lorsque la caméra observe un ensemble de points dont les profondeurs sont mal estimées. Il a été mis en évidence que même si la région de stabilité peut être assez importante, des erreurs faibles sur les paramètres 3D peuvent conduire à un système instable. Nous avons également vu que le cas le plus favorable correspond à une cible symétrique, centrée dans l'image et parallèle au plan image. Dans ce cas l'AV2D basé points assure la stabilité asymptotique locale du point d'équilibre quelles que soient les erreurs sur les paramètres 3D.

Toujours dans ce contexte (la caméra obéissant au modèle de projection unifié), nous avons

développé des schémas de commande 2D et 2D1/2. Les primitives type points et droites ont été traitées. Nous avons également proposé l'utilisation de nouvelles primitives basées moments dans l'image afin d'améliorer les propriétés de découplage de la commande. Les primitives droites ont été exploitées pour la commande de robots manipulateurs série et parallèle et pour la commande de robots à roues. En outre, les méthodes proposées permettent de traiter de manière indifférente les caméras catadioptriques centrale, sphérique et perspective ainsi qu'une grande partie des capteurs de vision avec optique fish-eye. Il en découle donc une généralité importante des résultats obtenus. Afin d'aboutir à cette généralité, nous avons montré comment la transformation homographique entre deux images d'un plan de référence, contenant le déplacement du capteur, pouvait être obtenue à partir de la projection d'un ensemble de points ou de droites via le modèle unifié. En particulier, nous avons vu que de la même manière que pour une caméra répondant au modèle pinhole, le problème d'estimation de la matrice d'homographie entre deux images pouvait être formulé linéairement. Dans le cas des points, ce résultat est obtenu en utilisant plutôt les points sur la sphère virtuelle apparaissant dans le modèle de projection que leurs correspondants dans l'image. Dans le cas des droites, ce résultat est obtenu en ayant recours aux droites polaires associées aux coniques issues de la projection des droites 3D correspondantes. Ces résultats peuvent être exploités pour résoudre les problèmes d'estimation du déplacement partiel du capteur entre deux prises de vue et de reconstruction de la structure de la scène :

- reconstruction euclidienne sous hypothèse d'étalonnage pour les points ;
- reconstruction projective sous hypothèse d'étalonnage faible pour les droites
- et, à partir de cette dernière et des paramètres d'étalonnage restants, reconstruction euclidienne pour les droites.

Nous avons également montré que le modèle unifié pouvait être interprété comme un modèle de projection avec distorsions radiales en le réécrivant sous une forme conventionnelle (modèle basé pinhole ou basé angle d'incidence). Cela nous a permis d'étendre de manière immédiate tous nos résultats aux caméras fish-eye.

Nos perspectives de recherche dans le domaine de l'asservissement visuel s'orientent dans plusieurs directions. La première, sur laquelle s'acharne les ferrus d'asservissement visuel depuis de nombreuses années (nous compris !), concerne le recherche "du" schéma de commande dans l'image permettant à la fois un découplage maximal et globalement stabilisant. Ce problème, nous tient d'autant plus à cœur qu'il semble à notre portée à l'aune de nos derniers résultats.

La seconde direction dans laquelle vont s'inscrire nos recherches a trait aux types de primitives utilisés dans nos schémas de commande. A ce jour, la majorité de nos résultats est basée sur l'observation de primitives géométriques (points, droites). Dans ce cas, le schéma d'asservissement visuel doit intégrer une mise en correspondance préalable des primitives visuelles (initialisation) et leur suivi durant toute la phase d'asservissement, ce qui en limite les potentialités. C'est pourquoi afin de réduire ces contraintes, il est possible d'utiliser des primitives plus globales comme nous l'avons fait avec des moments dans [TMCA09] ou comme dans [CC02] avec les coefficients de la transformée de Fourier ou encore comme dans [CM09] avec le signal de luminance. Cependant, ces approches nécessitent une segmentation de l'image, ou quand ce n'est pas le cas, conduisent toujours à de forts couplages. Nous souhaitons donc à moyen terme contribuer au développement de stratégies permettant l'élimination complète de ces phases de mise en correspondance et de

suivi tout en assurant de bonnes propriétés de découplage (et pourquoi pas de stabilité globale) de la commande. Nous pensons que la persistance importante des informations visuelles dans les images acquises par des caméras omnidirectionnelles doit pouvoir être exploitée dans ce sens.

Une troisième perspective naturelle pour nos travaux concerne l'étude de schémas de commande dynamique dans l'image. En effet, suite à des résultats très encourageants obtenus dans notre laboratoire concernant l'estimation de pose et de vitesse avec des images acquises par une caméra à acquisition séquentielle, nous avons débuté une étude sur des commandes dont le vecteur d'entrée est composé des torseurs cinématique et dynamique. Ces travaux que nous allons assurément poursuivre n'en sont aujourd'hui qu'à leurs balbutiements. Dans ce cadre, nous pensons en particulier allier les avantages de la vision omnidirectionnelle et de l'acquisition séquentielle. Pour cela nous envisageons très prochainement d'exploiter des caméras omnidirectionnelles à acquisition séquentielle dont le principe repose sur un ensemble d'imageurs linéaires en rotation autour d'un axe fixe. Cette technologie a l'avantage par rapport aux caméras séquentielles classiques de fournir une image omnidirectionnelle à résolution ajustable (en faisant varier la vitesse de rotation des imageurs linéaires). Les perspectives de recherche sont ici très importantes puisque une grande partie des domaines de la vision géométrique/cinématique et de l'asservissement visuel cinématique/dynamique pourra être réexaminée à l'aune de ce principe d'acquisition d'image.

A plus long terme, nous envisageons de poursuivre nos travaux dans deux autres directions :

1. A ce jour, les cas de figures étudiés concernaient tous la situation mono-caméra et mono-robot. Par la suite, nous souhaiterions nous pencher sur deux scénarios différents. Le premier concerne les situations où un système hétérogène de caméras, composé d'une caméra conventionnelle et d'une caméra omnidirectionnelle est utilisé. Dans cette configuration, la caméra omnidirectionnelle donne une vue globale mais avec une faible résolution alors que la caméra conventionnelle donne une vue locale mais avec une résolution importante. Le second scénario, au contraire, concerne les situations où une caméra omnidirectionnelle observe un système multi-robots (par exemple plusieurs bras manipulateurs). Bien sûr dans les deux cas de figures, les problèmes à résoudre sont nombreux aussi bien au niveau de la perception (pensons à la mise en correspondance de primitive dans des images acquises par des capteurs différents ou encore à l'extraction et au suivi de plusieurs bras dans les images omnidirectionnelles) qu'au niveau de la commande (pensons ici au couplage des observations au niveau de la commande ou à la définition des tâches de coopération).
2. Le second problème sur lequel nous souhaitons porter nos efforts à long terme est relatif à la réalisation de tâches de saisie et de manipulation d'objets a priori inconnus et possiblement déformables par un système multi-bras (par exemple les bras d'un robot humanoïde). Les techniques d'asservissement visuel que nous développons contribuent largement à cet objectif. A ce jour, nous nous sommes penchés de manière plus spécifique sur le couplage vision/effort et l'estimation de points de saisies. Les premiers résultats que nous avons obtenus dans ce domaine, bien que prometteurs, restent à approfondir dans deux directions : (1) la commande pour le couplage vision/force dans le cas multi-bras, (2) les méthodologies pour la saisie d'objets complexes (possiblement non-rigides) dans ce même cas.

Navigation autonome basée vision

Les contributions apportées dans la thématique “navigation par vision” ont été présentées dans le chapitre 4. De manière très synthétique, la stratégie de navigation adoptée se décline en quatre phases : (1) Lors de la première phase, le véhicule se déplace dans son espace de travail et acquiert un panel d’informations visuelles représentatives des lieux. Ces informations sont indexées et organisées dans une base de donnée. C’est la phase de constitution de la mémoire visuelle du robot. (2) Lors de la phase opérative, le robot se localise en comparant les données capteurs courantes à celles mémorisées. (3) Une fois le robot localisé, une suite ordonnée d’informations capteurs de la base est construite et décrit « l’itinéraire » à emprunter pour atteindre une position cible contenue dans la mémoire visuelle. (4) L’itinéraire déterminé par la suite d’informations capteurs est finalement parcouru en s’appuyant sur des méthodes de commande référencée capteurs.

Le principal intérêt de cette approche est la conservation d’une image fidèle de l’environnement en omettant la reconstruction 3D d’un monde virtuel soumis à propagation d’erreurs. Nous avons proposé une méthode efficace de localisation lorsque la mémoire visuelle est composée d’images acquises par une caméra grand angle et des stratégies de commande référencée capteurs adaptées à ce schéma de navigation. En outre, nous avons développé des lois de commande à la fois pour les robots mobiles de type char ou bicyclette et pour les drones de type quadrirotor, les commandes développées étant toujours valables pour l’ensemble des caméras centrales.

A l’heure actuelle, nous envisageons de poursuivre nos travaux dans cette thématique dans deux directions :

1. Dans un premier temps, nous chercherons à augmenter le degré d’autonomie du système robotique en supprimant l’intervention humaine lors de la phase d’apprentissage. Pour cela, une première difficulté réside dans la définition même de la tâche d’exploration. Cette première phase est primordiale pour définir correctement les objectifs d’exploration et donc les lois de commande associées. Une stratégie possible consiste à fixer les objectifs d’exploration en terme de couverture visuelle de l’espace. On pourra dans ce cas envisager deux tâches de couverture. La première, locale, consistant à maximiser le champ de perception du robot et la seconde, plus globale, consistant à biaiser l’exploration locale pour orienter le robot vers des sites d’intérêt particulier. Les stratégies d’exploration pourront également exploiter la “mémoire” du robot. Par exemple, on peut imaginer que la stratégie d’exploration d’un site puisse être orientée par l’expérience du système robotique sur un site similaire. A l’opposé des techniques classiques de type SLAM où la notion de navigation autonome est absente, la localisation, la cartographie et la navigation autonome devront être réalisées conjointement pour réaliser la tâche d’exploration. Un nouveau concept : le SLAN (pour Simultaneous Localisation And Navigation) est peut-être ici à inventer.
2. Un autre objectif, tout aussi important que le premier, sera de traiter les environnements dynamiques. Nous entendons par là les environnements dont le contenu a changé durablement (construction d’une école dans une ville) ou provisoirement (voitures stationnées ou roulantes, piétons, poules et renards ...). Le développement de stratégies d’exploration autonome, devrait apporter des solutions aux problèmes des changements durables dans

l'environnement de navigation. En ce qui concerne les changements provisoires de l'environnement, nous envisageons de les traiter comme un problème d'évitement d'obstacle. Dans l'approche adoptée, la phase opérative consiste en la réalisation d'une séquence d'objectifs définie dans l'espace capteur. Or, en pratique, il n'est pas nécessaire d'atteindre exactement et séquentiellement toutes les cibles. Dans cette optique, il serait probablement fructueux d'associer à notre modèle de représentation la notion d'espace libre. Celle-ci pourrait par exemple être utilisée pour replanifier (dans l'espace du capteur) les déplacements du robot afin d'éviter toute collision avec un obstacle. Dans ce contexte, il serait probablement intéressant de disposer d'un capteur capable d'estimer la vitesse des obstacles. Ceci constitue un point de convergence de nos perspectives. En effet, c'est précisément un des résultats attendus avec les caméras omnidirectionnelles à acquisition séquentielle.

Pour terminer, nous pensons que les résultats de ces différentes investigations nous mèneront naturellement à généraliser notre approche de navigation par mémoire sensorielle à la fois aux tâches de navigation multi-capteurs et multi-robots mais également à des tâches aussi complexes que la saisie et la manipulation d'objets inconnus par un système multi-bras. Nous sommes, en effet maintenant persuadés que la complexité de ces tâches pourra être significativement réduite si le robot dispose d'une mémoire "sensorielle" évolutive adéquatement organisée.

Bibliographie

- [ADM07] N. Andreff, T. Dallej, and P. Martinet. Image-based visual servoing of a Gough-Stewart parallel manipulator using leg observations. *International Journal of Robotics Research*, 26(7) :677–687, 2007.
- [AF01] M. Ahmed and A. Farag. Non-metric calibration of camera lens distortion. In *IEEE International Conference on Image Processing*, volume 2, pages 157–160, Thessaloniki Greece, October 7-10 2001.
- [AM06] N. Andreff and P. Martinet. Unifying kinematic modeling, identification and control of a Gough-Stewart parallel robot into a vision-based framework. *IEEE Transactions on Robotics*, 22(6) :1077–1086, 2006.
- [AMAM05b] H. H. Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Image-based control of mobile robot with central catadioptric cameras. In *IEEE International Conference on Robotics and Automation, ICRA'2005*, pages 3522–3527, Barcelona, Spain, 18–22 April 2005.
- [AMM09b] H. Hadj Abdelkader, Y. Mezouar, and P. Martinet. *Points based visual servoing with central cameras*, chapter 17. Edited by Graziano Chesi and Koichi Hashimoto : Visual Servoing via Advanced Numerical Methods. Springer, 2009.
- [Ass72] Isaac Assimov. *Le grand livre des robots*. Number 453. J'ai lu, j'ai lu edition, 1972.
- [BA02a] J. Barreto and H. Araujo. Geometric properties of central catadioptric line images. In *7th European Conference on Computer Vision, ECCV'02*, pages 237–251, Copenhagen, Denmark, May 2002.
- [BA05] João P. Barreto and Helder Araújo. Geometric properties of central catadioptric line images and their application in calibration. *IEEE Transactions Pattern Anal. Mach. Intell.*, 27(8) :1327–1333, 2005.
- [Bar06] J.P. Barreto. A unifying geometric representation for central projection systems. *Computer Vision and Image Understanding*, 103 :208–217, 2006. Special issue on omnidirectional vision and camera networks.
- [Ber99] F. Berry. *Contournement d'objet par asservissement visuel*. PhD thesis, Université Blaise Pascal–Clermont II, LASMEA-Clermont-ferrand, 1999.

- [BHS94] B. Bishop, S. Hutchinson, and M. Spong. On the performance of state estimation for visual servo systems. In *IEEE International Conference on Robotics and Automation*, pages 168–173, 1994.
- [BK00] R. Benosman and S. Kang. *Panoramic Vision*. Springer Verlag ISBN 0-387-95111-3, 2000.
- [BL95] A. Basu and S. Licardie. Alternative models for fish-eye lenses. *Pattern Recognition Letters*, 16(4) :433–441, 1995.
- [BM06b] S. Benhimane and E. Malis. A new approach to vision-based robot control with omni-directional cameras. In *IEEE International Conference on Robotics and Automation*, Orlando, USA, May 2006.
- [BMH02] J. P. Barreto, F. Martin, and R. Horaud. Visual servoing/tracking using central catadioptric images. In *International Symposium on Experimental Robotics*, pages 863–869, Bombay, India, July 2002.
- [BMM05] G. Blanc, Y. Mezouar, and P. Martinet. Indoor navigation of a wheeled mobile robot along visual routes. In *IEEE International Conference on Robotics and Automation, ICRA'2005*, pages 3354–3359, Barcelona, Spain, 18–22 April 2005.
- [BN98] S. Baker and S. K. Nayar. A theory of catadioptric image formation. In *International Conference on Computer Vision*, pages 35–42, Bombay, India, 1998.
- [BN99] S. Baker and S. K. Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2) :1–22, November 1999.
- [BP02] H. Bakstein and T. Pajdla. Panoramic mosaicing with a 180° field of view lens. In *IEEE Workshop on Omnidirectional Vision*, pages 60–67, Los Alamitos, US, 2002.
- [BR73] R. Bolles and R. Paul. The use of sensory feedback in a programmable assembly system. Technical report cs-396, Stanford Artificial Intelligence Laboratory, Computer Science Department, 1973.
- [Cad99] V. Cadenat. *Commande référencée multi-capteurs pour la navigation d'un robot mobile*. PhD thesis, Université Paul Sabatier de Toulouse, 1999.
- [CC00a] C. Collewet and F. Chaumette. A contour approach for image-based control of objects with complex shape. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'00*, volume 1, pages 751–756, Takamatsu, Japan, November 2000.
- [CC02] C. Collewet and F. Chaumette. Positioning a camera with respect to planar objects of unknown shape by coupling 2d visual servoing and 3d estimations. *IEEE Transactions on Robotics and Automation*, 18(3) :322–333, June 2002.

-
- [CG96] P. Corke and M.C. Good. Dynamic effects in visual closed-loop systems. *IEEE Transactions on Robotics and Automation*, 12(5) :671–683, October 1996.
- [CH00a] P. Chang and M. Herbert. First results in omnidirectional visual servoing. *Advanced Robotics*, 14(3) :205–220, 2000.
- [CH01] P. Corke and S. Hutchinson. A new partitioned approach to image-based visual servo control. *IEEE Transactions on Robotics and Automation*, 5(5) :507–515, August 2001.
- [CH06] F. Chaumette and S. Hutchinson. Visual servo control, part i : Basic approaches. *IEEE Robotics and Automation Magazine*, 13(4) :82–90, December 2006.
- [CH07] F. Chaumette and S. Hutchinson. Visual servo control, part ii : Advanced approaches. *IEEE Robotics and Automation Magazine*, 14(1) :109–118, March 2007.
- [Cha90] F. Chaumette. *La relation Vision-Commande : théorie et application à des tâches robotiques*. PhD thesis, Université de Rennes I, France, 1990.
- [CM00] F. Chaumette and E. Marchand. A new redundancy-based iterative scheme for avoiding joint limits : Application to visual servoing. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1720–1725, San Francisco, California, April 2000.
- [CM09] C. Collewet and E. Marchand. Photometry-based visual servoing using light reflexion models. In *IEEE International Conference on Robotics and Automation, ICRA '09*, pages 701–706, Kobe, Japan, May 2009.
- [CMEM07] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet. A generic fisheye camera model for robotic applications. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'07*, pages 1683–1688, San-Diego, California, USA, 29 October–2 November 2007.
- [CMEM08] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet. Efficient hierarchical localization method in an omnidirectional images memory. In *IEEE International Conference on Robotics and Automation, ICRA '08*, pages 13–18, Pasadena, California, USA, 19–23 May 2008.
- [CMGM09] J. Courbon, Y. Mezouar, N. Guenard, and P. Martinet. Visual navigation of a quadrotor aerial vehicle. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'09*, St. Louis, MO, USA, 11–15 October 2009.
- [CMLE08a] J. Courbon, Y. Mezouar, L. Lequievre, and L. Eck. Efficient visual memory based navigation of indoor robot with a wide-field of view camera. In *International Conference on Control, Automation, Robotics and Vision, ICAR-CV'2008*, pages 268–273, Hanoi, Vietnam, December 2008.
- [CMM08] J. Courbon, Y. Mezouar, and P. Martinet. Indoor navigation of a non-holonomic mobile robot using a visual memory. *Autonomous Robots*, 25(3) :253–266, 2008.

- [CMM09] J. Courbon, Y. Mezouar, and P. Martinet. Autonomous navigation of vehicles from a visual memory based on the use of generic camera model. *IEEE Transactions on Intelligent Transportation System*, 2009.
- [CPBM03] E. Cervera, A.P. Del Pobil, F. Berry, and P. Martinet. Improving image-based visual servoing with three-dimensional features. *International Journal of Robotics Research, Special Issue on Visual Servoing*, 22(10-11) :821–839, 2003.
- [CWH97] D-M. Chuang, S-C. Wu, and M-K. Hor. Adaptative fuzzy visual servoing in robot control. In *IEEE International Conference on Robotics and Automation*, pages 811–816, Albuquerque, New Mexico, USA, April 1997.
- [DDL97] N. Daucher, M. Dhome, J. T. Lapresté, and G. Rives. Speed command of a robotic system by monocular pose estimate. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 55–62, Grenoble, France, September 1997.
- [DF01] F. Devernay and O. Faugeras. Straight lines have to be straight. *Machine Vision and Applications*, 13(1) :14–24, 2001.
- [DS37] R. Descartes and D. Smith. *The Geometry of René Descartes*. Dovers, originally published in Discours de la Méthode, New york, 1637.
- [EGP04] B. Espiau, M. Ghallab, and F. Pierrot. Robotique domestique au japon. Technical Report SMM04_097, Ambassade de France à Tokyo, ADIT, Septembre 2004.
- [EMF03] M.T. El-Melegy and A.A. Farag. Statistically robust approach to lens distortion calibration with model selection. In *Conference on Computer Vision and Pattern Recognition CVPR-03, Workshop on Intelligent Learning*, volume 8, pages 150–156, Madison, Wisconsin, USA, 2003.
- [FAB⁺00] C. Fermüller, Y. Aloimonos, P. Baker, R. Pless, J. Neumann, and B. Stuart. Multi-camera networks : Eyes from eyes. In *Proceedings of the IEEE Workshop on Omnidirectional Vision, Hilton Head Island, USA*, pages 11–18, June 2000.
- [FC07] D. Folio and V. Cadenat. A new controller to perform safe vision-based navigation tasks amidst possibly occluding obstacles. In *European Control Conference, ECC'07*, Kos, Greece, July 2007.
- [FCM00] G. Flandin, F. Chaumette, and E. Marchand. Eye-in-hand / eye-to-hand cooperation for visual servoing. In *IEEE International Conference on Robotics and Automation*, San Francisco, CA, Avril 2000.
- [Fil01] David Filliat. *Cartographie et estimation globale de la position pour un robot mobile autonome*. PhD thesis, Université Paris 6, 2001.
- [Fit01] A. Fitzgibbon. Simultaneous linear estimation of multiple-view geometry and lens distortion. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2001*, volume 1, pages I–125–I–132, Hawaii, USA, 2001.

-
- [FL88] O. Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. *International Journal of Pattern Recognition and Artificial Intelligence*, 2(3) :485–508, 1988.
- [Fle94] M.M. Fleck. Perspective projection : the wrong imaging model. Technical report, Computer Science. University of Iowa. USA, 1994. Technical report 95-01.
- [FLS63] R. P. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures on Physics*. 1963.
- [FM03a] D. Filliat and J.A. Meyer. Map-based navigation in mobile robots - I. a review of localisation strategies. *Journal of Cognitive Systems Research*, 4(4) :243–282, 2003.
- [FM03b] D. Filliat and J.A. Meyer. Map-based navigation in mobile robots - II. a review of map-learning and path-planning strategies. *Journal of Cognitive Systems Research*, 4(4) :283–317, 2003.
- [Gan99] J. Gangloff. *Asservissements visuels rapides d'un robot manipulateur à 6 degrés de liberté*. PhD thesis, Université Louis Pasteur, 1999.
- [GD00] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical implications. In *European Conference on Computer Vision*, volume 29, pages 159–179, Dublin, Ireland, May 2000.
- [GdM00] J. Gangloff and M. de Mathelin. High speed visual servoing of a 6 dof manipulator using multivariable predictive control. In *IEEE International Conference on Robotics and Automation*, San Francisco, Californie, april 2000.
- [GWSV00] J. Gaspar, N. Winters, and J. Santos-Victor. Vision-based navigation and environmental representations with an omnidirectional camera. *IEEE Transactions on Robotics and Automation*, 16(6) :890–898, December 2000.
- [HA94] K. Hosoda and M. Asada. Versatile visual servoing without knowledge of true jacobian. *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'94*, pages 186–193, September 1994.
- [HA06b] Hicham Hadj-Abdelkader. *Asservissement visuel en vision omnidirectionnelle*. PhD thesis, Université Blaise-Pascal, Clermont-Ferrand, Décembre 2006.
- [HAMAM06a] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. Decoupled homography-based visual servoing with omnidirectional cameras. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'06*, pages 2332–2337, Beijing, China, 9–14 October 2006.
- [HAMM07] H. Hadj-Abdelkader, Y. Mezouar, and P. Martinet. Decoupled visual servoing from a set of points imaged by an omnidirectional camera. In *IEEE International Conference on Robotics and Automation, ICRA '07*, pages 1697–1702, Roma, Italy, 10–14 April 2007.

- [HAMMC08] H. Hadj-Abdelkader, Y. Mezouar, P. Martinet, and F. Chaumette. Catadioptric visual servoing from 3d straight lines. *IEEE Transactions on Robotics*, 24(3) :652–665, June 2008.
- [HDE98] R. Horaud, F. Dornaika, and B. Espiau. Visually guided object grasping. *IEEE Transactions on Robotics and Automation*, 14(4) :525–532, August 1998.
- [HEK96] K. Hashimoto, T. Ebine, and H. Kimura. Visual servoing with hand-eye manipulator - optimal control approach. *IEEE Transactions on Robotics and Automation*, 12(5) :766–774, October 1996.
- [HHC96] S. Hutchinson, G.D. Hager, and P. Corke. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5) :651–670, October 1996.
- [HM02] T. Hamel and R. Mahony. Visual servoing of an under-actuated dynamic rigid body system : an image-based approach. *IEEE Transaction on Robotics and Automation*, 18(2) :187–198, April 2002.
- [HZ74] E. Hecht and A. Zajac. *Optics*. 1974.
- [HZ00a] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN : 0521623049, 2000.
- [IO02] M. Iwatsuki and N. Okiyama. A new formulation of visual servoing based on cylindrical coordinates system with shiftable origin. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'02*, pages 354–359, Lausanne, Switzerland, October 2002.
- [KB00] J. Kumler and M. Bauer. Fisheye lens designs and their relative performance. In *SPIE*, volume 4093, pages 360–369, 2000.
- [KB04] J. Kannala and S. Brandt. A generic camera calibration method for fish-eye lenses. In *International Conference on Pattern Recognition, ICPR 2004*, volume 1, pages 10–13, Cambridge, UK, 2004.
- [Kel96] R. Kelly. Robust asymptotically stable visual servoing of planar robots. *IEEE Transactions on Robotics and Automation*, 12(5) :755–766, October 1996.
- [KH91] A. J Koivo and N. Houshangi. Real-time vision feedback for servoing robotic manipulator with self-tuning controller. *IEEE Transactions On Systems, Man, and Cybernetics*, 21(1) :134–142, February 1991.
- [Kha99] D. Khadraoui. *La commande référencée vision pour le guidage automatique des véhicules*. PhD thesis, Université Blaise Pascal, Clermont Ferrand, 1999.
- [Kin89] R. Kingslake. *A History of the Photographic Lens*. Academic Press, San Diego, 1989. 334 p., ISBN 0-12-408640-3.

-
- [KKK04] G. Klancar, M. Kristan, and R. Karba. Wide-angle camera distortions and non-uniform illumination in mobile robot tracking. *Robots and Autonomous Systems*, 46(2) :125–133, 2004.
- [LCMM08] L. Lequievre, J. Courbon, Y. Mezouar, and P. Martinet. Sovin, a software platform for visual navigation. In *3rd National Conference on Control Architectures of Robots, CAR'08*, Bourges, France, 29–30 September 2008.
- [LDW91] J.J. Leonard and H.F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, 7(3) :376–382, June 1991.
- [LH05] H. Li and R. Hartley. An easy non-iterative method for correcting lens distortion from nine point correspondences. In *Workshop on Omnidirectional Vision, Camera Networks and Non-classical cameras, OMNIVIS'05*, Beijing, China, 2005.
- [LJDC04] J.-T. Lapresté, F. Jurie, M. Dhome, and F. Chaumette. An efficient method to compute the inverse jacobian matrix in visual servoing. In *IEEE International Conference on Robotics and Automation, ICRA'04*, volume 1, pages 727–732, New Orleans, Louisiana, April 2004.
- [LM04] J. T. Lapreste and Y. Mezouar. A hessian approach to visual servoing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'04*, volume 1, pages 998–1003, Sendai, Japan, 28 September–2 October 2004.
- [Mal98] E. Malis. *Contributions à la modélisation et à la commande en asservissement visuel*. Thèse de doctorat, Université de Rennes I, IRISA, November 1998.
- [Mal04] E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *IEEE International Conference on Robotics and Automation*, New Orleans, USA, April 2004.
- [MAPP05] G.L. Mariottini, E. Alunno, J. Piazzzi, and D. Prattichizzo. Epipole-based visual servoing with central catadioptric camera. In *IEEE International Conference on Robotics and Automation*, pages 3516–3521, Barcelona, Spain, 2005.
- [MC00a] E. Malis and 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, June 2000.
- [MC02a] E. Malis and F. Chaumette. Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods. *IEEE Transactions on Robotics and Automation*, 18(2) :176–186, April 2002.
- [MC08] M. Marey and F. Chaumette. Analysis of classical and new visual servoing control laws. In *IEEE International Conference on Robotics and Automation, ICRA'08*, pages 3244–3249, Pasadena, California, May 2008.

- [MC09] M. Marey and F. Chaumette. A tentative to reach a visual singular configuration using halley's method. In *IEEE International Conference on Robotics and Automation, ICRA'09*, Kobe, Japan, May 2009.
- [MCB99] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Transactions on Robotics and Automation*, 15(2) :238–250, April 1999.
- [MCC02] R. Mahony, P. Corke, and F. Chaumette. Choice of image features for depth-axis control in image-based visual servo control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'02*, volume 1, pages 390–395, Lausanne, Switzerland, October 2002.
- [MCR96] E. Marchand, F. Chaumette, and A. Rizzo. Using the task function approach to avoid robot joint limits and kinematic singularities in visual servoing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'96*, pages 1083–1090, Osaka, Japon, Novembre 1996.
- [MCSP01] E. Marchand, F. Chaumette, F. Spindler, and M. Perrier. Controlling the manipulator of an underwater rov using a coarse calibrated pan tilt camera. In *International Conference on Robotics and Automation*, pages 2773–2778, Séoul, Corée du Sud, Mai 2001.
- [Mez01] Y. Mezouar. *Planification de trajectoires pour l'asservissement visuel*. PhD thesis, Université de Rennes 1, Mention informatique, November 2001.
- [MG99] P. Martinet and J. Gallice. Position based visual servoing using a nonlinear approach. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 531–536, Kyongju, Korea, October 1999.
- [MGK96] P. Martinet, J. Gallice, and D. Khadraoui. Vision based control. In *Proceedings of the World Automation Congress, WAC'96, Robotic and Manufacturing Systems*, 1996.
- [MH98] E. Marchand and G.D. Hager. Dynamic sensor planning in visual servoing. volume 3, pages 1988–1993, Leuven, Belgium, May 1998.
- [MKC09] R. Mebarki, A. Krupa, and F. Chaumette. Modeling and 3d local estimation for in-plane and out-of-plane motion guidance by 2d ultrasound-based visual servoing. In *IEEE International Conference on Robotics and Automation, ICRA'09*, pages 319–325, Kobe, Japan, May 2009.
- [MLS⁺00] G. Morel, T. Leibzeit, J. Szewczyk, S. Boudet, and J. Pot. Explicit incorporation of 2d constraints in vision based control of robot manipulators. In P. Corke and J. Trevelyan, editors, *Experimental Robotics VI*, pages 99–108. Springer Verlag, 2000.
- [MMR09] E. Malis, Y. Mezouar, and P. Rives. Robustness of image-based visual servoing with a calibrated camera in the presence of uncertainties in the 3d structure. *IEEE Transactions on Robotics*, To appear 2009.

-
- [Mor98a] H. Moravec. *ROBOT : mere machine to transcendent mind*. Oxford University Press, November 1998.
- [Mor98b] H. Moravec. When will computer hardware match the human brain? *Journal of Transhumanism*, 1, March 1998.
- [MR07] C. Mei and P. Rives. Single view point omnidirectional camera calibration from planar grids. In *IEEE International Conference on Robotics and Automation, ICRA '07*, pages 3945–3950, Rome, Italy, April 2007.
- [MSC05] E. Marchand, F. Spindler, and F. Chaumette. Visp for visual servoing : a generic software platform with a wide class of robot control skills. *IEEE Robotics and Automation Magazine*, 12(4) :40–52, December 2005.
- [Nay97] S.K. Nayar. Catadioptric Omnidirectional Camera. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 482–488, San Juan, Puerto Rico, Jun 1997.
- [NDC⁺07] A.E. Neff, L. DongBin, V.K. Chitrakaran, D.M. Dawson, and T.C. Burg. Velocity control for a quad-rotor uav fly-by-camera interface. In *SoutheastCon*, pages 273–278, Richmond, USA, 2007.
- [Nis04] D. Nistér. An efficient solution to the five-point relative pose problem. *Transactions on Pattern Analysis and Machine Intelligence*, 26(6) :756–770, 2004.
- [PGR91] R. Pissard-Gibollet and P. Rives. Asservissement visuel appliqué à un robot mobile : état de l’art et modélisation cinématique. Technical report, INRIA, Sophia-Antipolis, 1991. RR-1577.
- [PGR95] R. Pissard-Gibollet and P. Rives. Applying visual servoing techniques to control a mobile hand-eye system. In *IEEE International Conference on Robotics and Automation, ICRA '95*, pages 166–171, Nagoya, Aichi, Japan, may 1995.
- [PK02] J. Pers and S. Kovacic. Nonparametric, model-based radial lens distortion correction using titled camera assumption. In *Computer Vision Winter Workshop 2002*, Bad Aussee, Austria, 2002.
- [PKK91] N. Papanikolopoulos, P.K. Khosla, and T. Kanade. Vision and control techniques for robotic visual tracking. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 857–864, Sacramento, Californie, April 1991.
- [PM96] C. Pégard and E. M. Mouaddib. A mobile robot using a panoramic view. In *IEEE International Conference on Robotics and Automation, ICRA '96*, pages 89–94, Minneapolis, Minnesota, U.S.A, Mai 1996.
- [PPC04] J. Piazzi, D. Prattichizzo, and N.J. Cowan. Auto-epipolar visual servoing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 363–368, Sendai, Japan, October 2004.

- [PS04] P.-O. Persson and G. Strang. A simple mesh generator in MATLAB. *SIAM Review*, 46(2) :329–345, 2004.
- [PSH01] T. Pajdla, T. Svoboda, and V. Hlacac. *Panoramic Vision : Sensors, Theory and Applications*, chapter Epipolar Geometry of Central Panoramic Catadioptric Cameras, pages 73–102. R. Benosman and S. B. Kang, 2001.
- [PWH97] T. Pajdla, T. Werner, and V. Hlavác. Correcting radial lens distortion without knowledge of 3-d structure. Technical report, Czech Technical University. Prague. Czech Republic, 1997. Technical Report K335-CMP-1997-138.
- [Ray94] S.F. Ray. *Applied Photographic Optics*. Focal Press, Oxford, 1994. Second edition. ISBN-0240514998.
- [RB97] P. Rives and J-J. Borrelly. Visual servoing techniques applied to an underwater vehicle. In *IEEE International Conference on Robotics and Automation*, pages 1851–1856, Albuquerque, New Mexico, USA, April 1997.
- [RH00] A. Ruf and R. Horaud. Vision-based guidance and control of robots in projective space. In Springer, editor, *6th European Conference on Computer Vision (ECCV'00)*, volume 2, pages 50–66, Dublin, Ireland, June 2000.
- [SA96] S. Shah and J.K. Aggarwal. Intrinsic parameter calibration procedure for a (high distortion) fish-eye lens camera with distortion model and accuracy estimation. *Pattern Recognition*, 29(11) :1775–1788, 1996.
- [Sam95] C. Samson. Control of chained system. application to path following and time-varying stabilization of mobile robot. *IEEE Transactions on Automatic Control*, 40(1), 1995.
- [Sch05] E. Schwalbe. Geometric modelling and calibration of fisheye lens camera systems. In *International Society for Photogrammetry and Remote Sensing (ISPRS), Panoramic Photogrammetry Workshop*, Berlin, Germany, 24-25 February 2005.
- [SDF⁺01] J. Spletzer, A. Das, R. Fierro, C. Taylor, V. Humar, and J. Ostrowski. Cooperative localization and control for multi-robot manipulation. In *Proceedings of the Conference on Intelligent Robots and Systems*, volume 2, pages 631–636, 2001.
- [SF95] D.E. Stevenson and M.M. Fleck. Nonparametric correction of distortion. Technical report, Computer Science. University of Iowa. USA, 1995. Technical report TR 95-07.
- [SI73] Y. Shirai and H. Inoue. Guiding a robot by visual feedback in assembling tasks. *Pattern Recognition*, 5 :99–108, 1973.
- [SJA99] P.W. Smith, K.B. Johnson, and M.A. Abidi. Efficient techniques for wide-angle stereo vision using surface projection models. In *IEEE Conference on Computer*

Vision and Pattern Recognition, volume 1, pages 113–118, Ft. Collins, CO, USA, 1999.

- [Smi92] W.J. Smith. *Modern Lens Design : A Resource Manual*. McGraw-Hill, New-York, 1992. ISBN : 0070591784.
- [SMS06] D. Scaramuzza, A. Martinelli, and R. Siegwart. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In *Fourth IEEE International Conference on Computer Vision Systems, ICVS 2006*, pages 45–52, New York, NY, USA, 2006.
- [SP02a] T. Svoboda and T. Pajdla. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 49(1) :23–37, 2002.
- [SPH98a] T. Svoboda, T. Pajdla, and V. Hlavác. Epipolar geometry of panoramic cameras. In *European Conference On computer Vision*, pages 218–231, Freiburg, Germany, 1998.
- [SW83] A.C Sanderson and L.E Weiss. *Adaptative visual servo control of robots - editor A. Pugh, Dobot Vision*. Bedford, UK, 1983.
- [Tah04] O. Tahri. *Application des moments à l’asservissement visuel et au calcul de pose*. PhD thesis, Université de Rennes 1, mention informatique, Rennes, March 2004.
- [TBMM04] B. Thuilot, J. Bom, F. Marmoiton, and P. Martinet. Accurate automatic guidance of an urban electric vehicle relying on a kinematic GPS sensor. In *5th IFAC Symposium on Intelligent Autonomous Vehicles, IAV’04*, Instituto Superior Técnico, Lisbonne, Portugal, 2004.
- [TC05] O. Tahri and F. Chaumette. Point-based and region-based image moments for visual servoing of planar objects. *IEEE Transactions on Robotics*, 21(6) :1116–1127, December 2005.
- [TCM08] O. Tahri, F. Chaumette, and Y. Mezouar. New decoupled visual servoing scheme based on invariants from projection onto a sphere. In *IEEE International Conference on Robotics and Automation, ICRA’08*, pages 3238–3243, Pasadena, California, USA, 19–23 May 2008.
- [TFC08] R. Tatsambon Fomena and F. Chaumette. Visual servoing from two special compounds of features using a spherical projection model. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS’08*, pages 3040–3045, Nice, France, September 2008.
- [Thr02] S. Thrun. Robotic mapping : A survey. In G. Lakemeyer and B. Nebel, editors, *Exploring Artificial Intelligence in the New Millenium*. Morgan Kaufmann, 2002.
- [TM08] O. Tahri and Y. Mezouar. On the efficient second order minimization and image-based visual servoing. In *IEEE International Conference on Robotics and Automation, ICRA’08*, pages 3213–3218, Pasadena, California, USA, 19–23 May 2008.

- [TMAM09] O. Tahri, Y. Mezouar, N. Andreff, and P. Martinet. Omnidirectional visual-servo of a gough-stewart platform. *IEEE Transactions on Robotics*, 25(1) :178–183, February 2009.
- [TMCA09] O. Tahri, Y. Mezouar, F. Chaumette, and H. Araujo. *Visual Servoing and Pose Estimation with Cameras obeying the Unified Model*, chapter 18. Edited by Graziano Chesi and Koichi Hashimoto : Visual Servoing via Advanced Numerical Methods. Springer, 2009.
- [TMCC09] O. Tahri, Y. Mezouar, F. Chaumette, and P. Corke. Generic decoupled image-based visual servoing for cameras obeying the unified projection model. In *IEEE International Conference on Robotics and Automation, ICRA'09*, pages 1116–1121, Kobe, Japan, April 2009.
- [TMCG02] B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice. Position-based visual servoing : keeping the object in the field of vision. In *IEEE International Conference on Robotics and Automation*, pages 1624–1629, Washington DC, USA, May 2002.
- [TRS98] D. Tsakiris, P. Rives, and C. Samson. Extending visual servoing techniques to nonholonomic mobile robots. In G. Hager D. Kriegman and A. Morse, editors, *The Confluence of Vision and Control*, volume 237 of *LNCIS*, pages 106–117. Springer Verlag, 1998.
- [TS00] S. Tarbouriech and P. Souères. Advanced control strategy for the visual servoing scheme. In *6th IFAC Symposium on Robotic Control, SYROCO'00*, Vienna, Austria, 2000.
- [VSS03] R. Vidal, O. Shakernia, and S. Sastry. Formation control of nonholonomic mobile robots with omnidirectional visual servoing and motion segmentation. In *IEEE International Conference on Robotics and Automation*, pages 584–589, Taipei, Taiwan, September 2003.
- [WHB96] W.J Wilson, C.C Williams Hulls, and G.S Bell. Relative end-effector control using cartesian position-based visual servoing. *IEEE Transactions on Robotics and Automation*, 12(5) :684–696, 1996.
- [WSN87] L.E. Weiss, A.C. Sanderson, and C.P. Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE Journal of Robotics and Automation*, 3(5) :404–417, October 1987.
- [WWR93] S. W Wijesoma, D. F. H Wolfe, and R. J Richards. Eye-to-hand coordination for vision-guided robot control applications. *International Journal of Robotics Research*, 12(1) :65–78, February 1993.
- [XT97] Y. Xiong and K. Turkowski. Creating image-based vr using a self-calibrating fisheye lens. In *Conference on Computer Vision and Pattern Recognition, CVPR '97*, pages 237–246, Puerto Rico, USA, June 1997.

-
- [YH04] Xianghua Ying and Zhanyi Hu. Catadioptric camera calibration using geometric invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10) :1260–1271, October 2004.
- [YK90] Y. Yagi and S Kawato. Panoramic scene analysis with conic projection. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'90*, volume 1, pages 181–187, Tsuchiura. Japan, 1990.
- [YK94] Y. Yagi and S. Kawato. Panoramic scene analysis with conic projection. *IEEE Transactions on Robotics and Automation*, 10 :11–22, 1994.
- [ZDFL95] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal*, 78 :87–119, October 1995.
- [ZMP00] P. Zanne, G. Morel, and F. Plestan. Robust vision based 3d trajectory tracking using sliding mode control. In *IEEE International Conference on Robotics and Automation*, pages 2088–2093, San Francisco, CA, April 2000.