



Techniques d'automatique et de traitement du signal pour l'asservissement visuel et la perception auditive en robotique

Patrick Danès

► To cite this version:

Patrick Danès. Techniques d'automatique et de traitement du signal pour l'asservissement visuel et la perception auditive en robotique. Automatique / Robotique. Université Paul Sabatier - Toulouse III, 2010. <tel-00604522>

HAL Id: tel-00604522

<https://tel.archives-ouvertes.fr/tel-00604522>

Submitted on 29 Jun 2011

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

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

THÈSE

Présentée devant l'Université de Toulouse (Université Paul SABATIER)
en vue de l'obtention de l'Habilitation à Diriger des Recherches

par Patrick DANÈS
Maître de Conférence à l'Université Paul SABATIER
(École Doctorale Systèmes)

TECHNIQUES D'AUTOMATIQUE ET DE TRAITEMENT DU SIGNAL POUR L'ASSERVISSEMENT VISUEL ET LA PERCEPTION AUDITIVE EN ROBOTIQUE

Soutenue le 8 Décembre 2010, devant le jury composé de

M. Yannick DEVILLE	<i>Professeur Université Toulouse III, LATT</i>	Président
M. François CHAUMETTE	<i>Directeur de Recherche INRIA, IRISA Rennes</i>	Rapporteur
M. Michel DE MATHELIN	<i>Professeur ENSP Strasbourg, LSII/IRCAD</i>	Rapporteur
M. Philippe POIGNET	<i>Professeur Université Montpellier II, LIRMM</i>	Rapporteur
M. Germain GARCIA	<i>Professeur INSA Toulouse, LAAS-CNRS</i>	Examinateur
M. Bruno GAS	<i>Professeur Université Paris VI, ISIR</i>	Examinateur
M. Michel DEVY	<i>Directeur de Recherche CNRS, LAAS-CNRS</i>	Garant

Rapport LAAS N° 10858
Thèse préparée au LAAS-CNRS
7, avenue du Colonel Roche
31 077 Toulouse Cédex 4

Avant-Propos

Ce mémoire présente les travaux effectués au LAAS-CNRS depuis mon recrutement en tant que Maître de Conférence à l'Université Paul SABATIER de Toulouse. Le LAAS-CNRS constitue un cadre de travail exceptionnel, où la notion d'intérêt général se réalise au quotidien par l'existence de services communs disponibles et efficaces. Je remercie l'ensemble des directions successives, de même que les responsables successifs et membres de ces services, pour avoir permis que se pérennise cette singularité dans un paysage sans cesse plus incertain.

J'exprime ma sincère reconnaissance à MM. François CHAUMETTE, Directeur de Recherche INRIA à l'IRISA de Rennes, Michel DE MATHELIN, Professeur à l'École Nationale Supérieure de Physique de Strasbourg, et Philippe POIGNET, Professeur à l'Université de Montpellier II, pour m'avoir fait l'honneur de juger ce travail et d'en être les rapporteurs. Je suis également très honoré que MM. Germain GARCIA, Professeur à l'INSA de Toulouse, et Bruno GAS, Professeur à l'Université Pierre et Marie CURIE, aient accepté d'être examinateurs. Merci enfin à M. Yannick DEVILLE, Professeur à l'Université Paul SABATIER, de m'avoir fait l'honneur d'être examinateur et de présider le jury de soutenance.

Je remercie M. Michel DEVY, Directeur de Recherche au LAAS-CNRS et responsable bien-aimé de notre groupe de recherche pour la très bonne ambiance qu'il fait régner dans notre petite équipe, et pour s'être porté garant de ces travaux. Mes remerciements vont également aux responsables successifs du groupe Robotique et Intelligence Artificielle (RIA) puis des groupes qui en ont émané, pour leurs efforts continus dans l'animation de la recherche. Je sais particulièrement gré à M. Malik GHALLAB d'avoir permis ma réorientation scientifique dans RIA au tout début de ma carrière.

J'exprime ma reconnaissance et mon amitié à tous les doctorants, stagiaires et visiteurs que j'ai encadrés ou co-encadrés, à tous les collègues avec qui j'ai eu l'occasion de collaborer dans mon activité professionnelle, et plus largement à mes ami(e)s du LAAS et/ou de l'UPS qui se reconnaîtront aisément à la lecture de ces lignes. Je remercie tout particulièrement mes amis anciens membres du LAAS, qui m'ont apporté une aide inestimable : (par ordre alphabétique) Gérard AUTHIÉ (et Viviane), Françoise LE GALL-CHOLIN (et Sylvestre), ainsi que Mario PALUDETTO (et Françoise).

Enfin, merci à Marie-Laure, nos proches et amis pour m'avoir supporté dans tous les sens du terme ; à notre poussinette Emma pour le doux bonheur qu'elle nous apporte ; ainsi qu'à... un petit félin espiègle qui a partagé de longues heures de rédaction à mes côtés durant l'été.

À Emma et Marie-Laure,

Résumé

Les travaux présentés dans ce manuscrit d'Habilitation à Diriger des Recherches concernent essentiellement un ensemble de primitives du niveau fonctionnel de la Robotique, dont les fondements scientifiques sont ancrés dans l'Automatique et le Traitement du Signal. Bien plus qu'un domaine d'application de théories développées par ailleurs, la Robotique, par ses spécificités ou ses contraintes, « questionne » souvent l'état de l'art en Automatique et Signal, voire exige des extensions méthodologiques non triviales. C'est la richesse du dialogue entre ces trois disciplines qui a constitué le fil conducteur principal de nos recherches.

Un premier volet de nos travaux vise à développer un cadre générique pour l'analyse et la synthèse « multicritères» de commandes référencées vision, *i.e.* qui prennent en compte l'ensemble des contraintes (visibilité, saturations d'actionneurs, contraintes 3D pendant le déplacement, etc.). Ces problèmes sont ramenés à l'analyse en stabilité / la stabilisation de systèmes non linéaires certains rationnels sous contraintes rationnelles. Le support théorique est la théorie de Lyapunov et l'optimisation LMI. Par dualité, la localisation visuelle est abordée par des techniques de filtrage ensembliste robuste de systèmes rationnels.

Une deuxième contribution s'inscrit dans la thématique relativement récente de l'Audition en Robotique. Nous proposons des fonctions auditives bas-niveau pour la détection de sources sonores, leur localisation et leur extraction. Un capteur auditif intégré original conçu au LAAS-CNRS permet leur implémentation. Le support théorique est le traitement d'antenne (formation de voie, méthodes à haute résolution) et l'optimisation LMI. Plus récemment, nous avons posé les fondements d'une nouvelle approche pour la détection d'activité vocale, sur la base du filtrage adapté stochastique.

En parallèle à ces développements ciblés, un travail de fond dans les thématiques du filtrage stochastique et de la détection de ruptures s'est matérialisé par des collaborations scientifiques dans des domaines d'application ciblés : méthodes séquentielles de Monte Carlo et Quasi Monte Carlo pour le suivi visuel de personnes, de gestes, et la capture de mouvement par vision ; détection de ruptures et filtrage IMM entre modèles d'état d'ordres hétérogènes pour la surveillance de scènes dynamiques et la localisation ARGOS.

Certaines des thématiques exposées ci-dessus se « rejoignent » au sein des axes structurants de nos développements futurs. Un effort particulier concerne la perception active (*i.e.* exploitant la proprioception et le mouvement), déclinée dans le cadre de l'audition binaurale, ainsi que la fusion de données auditives embarquées et déportées dans des lieux intelligents.

Mots-clés : détection et estimation, asservissement visuel multicritère, théorie de LYAPUNOV, audition en robotique, traitement d'antenne, optimisation LMI.

Sommaire

Résumé

Introduction générale

1

A BILAN D'ACTIVITÉS

3

I CURRICULUM VITÆ

5

I.1	Fonctions Actuelles – Parcours – Vue globale des activités	7
I.2	Synthèse des Activités de Recherche	12
I.3	Synthèse des Activités d'Enseignement	28
I.4	Annexe : Résumé étendu de la Thèse de Doctorat	33

B ACTIVITÉS DE RECHERCHE (POST-THÈSE)

39

II Panorama des Activités de Recherche Post-Thèse

41

II.1	Résumé étendu des travaux en Détection et Estimation	43
II.2	Résumé condensé des travaux en Asservissement Visuel	54
II.3	Résumé condensé des travaux en Audition en Robotique	59

III Asservissement Visuel

65

III.1	Introduction	66
III.2	Contexte scientifique	66
III.3	Contribution à l'asservissement visuel multicritère	71
III.4	Approche de la localisation visuelle en tant que le dual de l'asservissement visuel	95
III.5	Conception et implémentation d'un environnement de simulation de commandes en robotique	98
III.6	Références	99

IV Audition en Robotique

105

IV.1	Introduction	106
IV.2	Contexte scientifique	106
IV.3	Contribution à l'audition en robotique	111
IV.4	Références	137

V Prospectives	141
V.1 Travaux en cours et prospectives associées à court terme	142
V.2 Axes structurants de notre recherche	145
V.3 Quelques références significatives	150
C SÉLECTION D'ARTICLES	153

Introduction générale

La perception que les chercheurs ont de la Robotique est très variable, y compris à l'intérieur de la thématique. Pour certains, la Robotique est une discipline à part entière, du fait qu'elle a exhibé des problèmes scientifiques propres et qu'elle a développé un corpus original pour leur résolution. Pour d'autres, il s'agit plutôt d'un champ d'application de disciplines connexes, implicitement considérées comme « plus fondamentales ». D'aucuns pensent que la recherche en Robotique est avant tout théorique et que l'expérimentation n'a qu'un rôle de validation, alors qu'*a contrario* la Robotique est aussi considérée comme une discipline expérimentale, au sens où seule l'expérimentation permet de révéler sa substance. Quels que soient les points de vue, il est toutefois communément admis que la mise en œuvre de la boucle perception-décision-action, centrale à la problématique du robot autonome et/ou en interaction avec les humains, nécessite un travail d'équipe au confluent de compétences multiples : Automatique, Traitement du Signal, Mathématiques Appliquées, Mécanique, Génie Logiciel, Informatique Temps Réel, Vision par Ordinateur, Intelligence Artificielle, etc.

Ce manuscrit d'Habilitation à Diriger des Recherches se situe à mi-chemin entre les positions exprimées ci-dessus. Les travaux que nous avons développés et que nous présentons ici concernent un ensemble de primitives du niveau fonctionnel de la Robotique, dont les fondements scientifiques ont été délibérément ancrés dans l'Automatique et le Traitement du Signal. La première étape des recherches correspondantes a par conséquent consisté à « s'imprégner » de techniques existantes, souvent développées sans préjuger d'un domaine d'application. Or, la simple application de ces techniques ne subvient pas aux besoins de la Robotique, pour de multiples raisons : la Robotique apporte systématiquement un lot de contraintes inédites, qui exige en retour des extensions théoriques non triviales ; la nécessité de construire un système et de le faire fonctionner suscite parfois des développements scientifiques nouveaux, à la frontière de la théorie et de l'implémentation ; enfin, le fait de se placer dans des environnements non contrôlés, variables et évolutifs rend souvent illusoire l'obtention d'une solution monolithique, auquel cas on recourt à une multiplicité d'algorithmes exécutés conditionnellement au contexte. Au-delà de telle ou telle contribution ponctuelle, c'est la richesse de ce dialogue entre Automatique, Signal et Robotique qui constitue le fil conducteur de nos recherches.

Les travaux que nous avons développés peuvent être partitionnés en trois volets :

- un volet « Détection et Estimation », qui a consisté en un travail de fond dans les thématiques du filtrage stochastique et de la détection de ruptures ; ce travail s'est matérialisé au sein de collaborations scientifiques avec des collègues spécialistes de domaines d'application ciblés ne relevant pas de notre compétence ;
- un volet « Asservissement Visuel », qui constitue une ligne de recherche propre dont le but principal est l'analyse et la synthèse *multicritères* de commandes référencées vi-

sion, *i.e.* qui prennent en compte l'ensemble des contraintes ; notre contribution à ce problème relativement peu abordé malgré le dynamisme de la communauté asservissement visuel s'appuie sur des développements récents en Automatique, qu'il a fallu adapter et compléter significativement pour répondre aux besoins de la Robotique ;

- un volet « Audition en Robotique », qui constitue également une ligne de recherche propre, dont le but est de participer à l'émergence de cette thématique très récente par la définition de primitives de localisation/extraction de sources sonores en temps réel, et par le développement d'un capteur audio intégré embarquable.

Ce document est structuré en deux parties. La Partie A présente un bilan d'activités. L'unique Chapitre I qui la constitue propose une synthèse des activités de recherche et des activités d'enseignement. Rédigé à la première personne à la manière d'un Curriculum Vitæ, il recense en particulier les encadrements, publications, participations à des projets et autres faits marquants de notre activité scientifique. Une annexe dédiée propose au lecteur un résumé étendu de la thèse de Doctorat, qui se situait hors du champ de la Robotique.

La Partie B constitue le mémoire de recherche proprement dit. Le Chapitre II propose d'abord un panorama succinct de nos recherches. Le but est d'en préciser le contexte, de résumer les principales contributions, et de les relier aux publications et encadrements. En raison de leur caractère plus individuel, nos travaux en « Asservissement Visuel » et « Audition en Robotique » sont ensuite présentés plus longuement dans les Chapitres III et IV en les résitant par rapport à l'état de l'art, en exposant les verrous scientifiques et en décrivant le principe des solutions apportées. Le Chapitre V conclut le document par les travaux en cours, les perspectives envisagées, ainsi que les axes qui structureront notre recherche à moyen terme.

Première partie

BILAN D'ACTIVITÉS

Chapitre I

CURRICULUM VITÆ

Sommaire

I.1 Fonctions Actuelles – Parcours – Vue globale des activités	7
I.2 Synthèse des Activités de Recherche	12
I.2.1 Travaux de Doctorat	12
I.2.1.1 Intitulé	12
I.2.1.2 Résumé condensé	12
I.2.1.3 Production scientifique associée	13
I.2.2 Bilan des thématiques de recherche postérieures au Doctorat	13
I.2.3 Encadrements	15
I.2.3.1 Post-Doctorant	15
I.2.3.2 Thèses de Doctorat	15
I.2.3.3 Stages de M2R/DEA	17
I.2.3.4 Encadrement de stages Ingénieur, M1 ou de Visiteurs	18
I.2.4 Publications	19
I.2.4.1 Contribution à un ouvrage collectif international	19
I.2.4.2 Journaux scientifiques	19
I.2.4.3 Conférences Internationales avec Actes et Comité de Lecture	20
I.2.4.4 Symposia Internationaux avec Actes et Comité de Lecture	22
I.2.4.5 Conférences Nationales avec Actes et Comité de Lecture	22
I.2.4.6 Manifestations avec Actes à Diffusion Limitée	23
I.2.5 Visibilité - Faits marquants - Animation scientifique	23
I.2.5.1 Rédaction de Rapport de Présoutenance de Thèse de Doctorat	23
I.2.5.2 Participation à un Jury de Soutenance de Thèse de Doctorat	24
I.2.5.3 Responsabilité Scientifique de Projets de Recherche	24
I.2.5.4 Contribution à des Projets de Recherche	25
I.2.5.5 Participation ponctuelle à des Projets de Recherche	26
I.2.5.6 Séminaires significatifs	26
I.2.5.7 Collaborations Scientifiques Suivies	27
I.2.5.8 Animation Scientifique	27
I.2.6 Divers	28
I.3 Synthèse des Activités d'Enseignement	28
I.3.1 Enseignements dispensés à l'Université Paul SABATIER	28
I.3.2 Investissement en Enseignements Pratiques	30

I.3.3	Responsabilité de la définition et de la mise en place d'enseignements théoriques	31
I.3.4	Responsabilités administratives - Tâches d'intérêt collectif	32
I.4	Annexe : Résumé étendu de la Thèse de Doctorat	33

I.1 Fonctions Actuelles – Parcours – Vue globale des activités

ÉTAT CIVIL

Patrick DANÈS - 42 ans (né le 06.08.1968) - Marié, un enfant - Nationalité française.

FONCTIONS ACTUELLES

Depuis le 01.09.1995,

- *Maître de Conférence en Automatique* (61^e section) dans l’UFR Physique-Chimie-Automatique de l’Université Paul SABATIER (UPS, Toulouse) sur le poste 6100 MCF 0443. Avancé au 7^o échelon le 01.08.2010.
- Chercheur dans le groupe *Robotique, Action, Perception (RAP)*, Département *Robotique et Intelligence Artificielle (RIA)* du Laboratoire d’Analyse et d’Architecture des Systèmes du CNRS (LAAS-CNRS, Toulouse).
- Titulaire de la *Prime d’Encadrement Doctoral et de Recherche* depuis 2008.

PARCOURS UNIVERSITAIRE

01.11.1991–20.01.1995 :

- *Doctorat en Automatique* sur le thème *Interfaçage Symbolique-Numérique dans la Simulation Qualitative des Systèmes Dynamiques*. Mention Très Honorable avec Félicitations du Jury. UPS & LAAS-CNRS.
→ Mots-clés : Raisonnement qualitatif, Interface symbolique-numérique, Commande optimale, Principe du maximum, Signaux qualitatifs, Simulation QSIM, Systèmes à paramètres qualitatifs.

1990–1991 :

- *DEA d’Automatique Informatique Industrielle et Traitement du Signal*, option Informatique Industrielle, UPS & LAAS-CNRS. Mention Bien.
- *Ingénierat en Intelligence Artificielle, Reconnaissance des Formes et Robotique*, équivalent au DESS du même nom, UPS. Mention Assez Bien.

1989–1990, 1988–1989, 1986–1988 :

- *Maîtrise EEA*, UPS. Mentions AB,B,AB,P. Salarié à mi-temps.
- *Licence EEA*, UPS. Mention B aux quatre unités de valeur.
- *DEUGS A*, UPS. Mention TB.

FONCTIONS CONTRACTUELLES D’ENSEIGNEMENT

01.09.1994–31.08.1995 :

- Attaché Temporaire en Enseignement et Recherche (quotité 50%) en Automatique et Traitement du Signal à l’UPS.

01.11.1991–31.08.1994 :

- Moniteur en Automatique, Informatique Industrielle et Traitement du Signal (UPS).

THÉMATIQUES DE RECHERCHE POSTÉRIEURES AU DOCTORAT

A. Détection et Estimation :

- Détection de ruptures pour la surveillance de scènes dynamiques.
 - ↪ Collaboration avec M. Thierry SENTENAC, LAAS/EMAC (1999–2002).
 - ↪ Mots-clés : filtrage de KALMAN linéaire, étendu, “unscented” ; détection statistique (principalement les tests GLRT et MLRT – Generalized/Marginalized Likelihood Ratio Test –, avec une focalisation sur le cas de modèles d'état d'ordres « hétérogènes »).
- Méthodes séquentielles de Monte Carlo pour le suivi visuel de personnes et de gestes.
 - ↪ Collaboration avec M. Frédéric LERASLE, LAAS/UPS (2002–2005).
 - ↪ Mots-clés : filtrage particulaire ; systèmes à sauts Markoviens.
- Méthodes séquentielles de Monte Carlo et de Quasi Monte Carlo pour la capture de mouvement par vision.
 - ↪ Collaboration avec M. Frédéric LERASLE, LAAS/UPS (2005–2008).
 - ↪ Mots-clés : filtrage particulaire ; méthodes séquentielles de Quasi Monte Carlo.
- Filtrage multimodèles pour la localisation ARGOS.
 - ↪ Collaboration avec MM Philippe GASPAR, Jean-Pierre MALARDÉ et François ROYER, Collecte Localisation Satellites, 31520 Ramonville (2010–).
 - ↪ Mots-clés : systèmes à sauts Markoviens ; filtres GPB, IMM (avec une focalisation sur le cas de modèles d'état d'ordres « hétérogènes »), etc.

B. Asservissement Visuel :

- Analyse et Synthèse multicritères de commandes référencées vision.
 - ↪ Activité de recherche propre (1999–2003, puis 2006–).
 - ↪ Collaboration avec M. Daniel F. COUTINHO, PUCRS, Porto Alegre, Brésil (2006–).
 - ↪ Mots-clés : théorie de LYAPUNOV ; optimisation LMI ; commande robuste.
- Dualité entre asservissement visuel et localisation visuelle.
 - ↪ Activité de recherche propre (*idem* que ci-dessus).
 - ↪ Mots-clés : optimisation LMI ; filtrage ensembliste robuste.
- Conception d'un environnement de simulation de commandes en robotique.

C. Audition en Robotique :

- Formation de voie pour l'extraction et la localisation de sources sonores.
 - ↪ Activité de recherche propre (2003–2006).
 - ↪ Mots-clés : antenne acoustique ; formation de voie large bande ; analyse modale ; optimisation convexe.
- Méthodes à haute résolution pour la localisation et la détection de sources sonores.
 - ↪ Activité de recherche propre (2006–2010).
 - ↪ Mots-clés : antenne acoustique ; extensions large bande de la méthode MUSIC ; détection basée sur des critères d'information.
- Détection de “patterns” audio.
 - ↪ Activité de recherche propre (2009–).
 - ↪ Mots-clés : filtrage adapté stochastique.
- Conception et implémentation d'un capteur auditif intégré pour la Robotique.
 - ↪ Activité de recherche propre (2003–).
 - ↪ Mots-clés : intégration matérielle et logicielle.

ENCADREMENTS

7 thèses de Doctorat (École Doctorale Systèmes, Toulouse) :

- Une thèse CIFRE à 25 % et deux thèses à 75 % en cours.
- Deux thèses à 50 %, une thèse à 75 % et une thèse à 100 % soutenues.

1 collaboration avec un visiteur Post-Doctorant (en cours).

10 stages M2R/DEA, dont 8 en co-encadrement.

15 stages M2, M1 ou visiteurs, dont 6 en co-encadrement.

PUBLICATIONS POSTÉRIEURES AU DOCTORAT

1 contribution à un ouvrage collectif international.

6 journaux scientifiques (+1 durant le Doctorat) :

- Int. Journal of Robust and Nonlinear Control, 2008.
- European Journal of Control, 2006.
- Int. Journal of Pattern Recognition and Artificial Intelligence, 2009.
- Machine Vision and Applications, 2010 (en ligne depuis 2008).
- Journal of the Robotics Society of Japan, 2010.
- Microwave and Optical Technology Letters, 2007.

23 conférences internationales avec actes et comité de lecture (+1 durant le Doctorat) :

- 5 IEEE Conf. on Decision and Control (CDC).
- 1 European Control Conference (ECC).
- 4 IEEE Int. Conf. on Image Processing (ICIP).
- 1 European Signal Processing Conference (EUSIPCO).
- 3 IEEE Int. Conf. on Robotics and Automation (ICRA), dont 1 session invitée.
- 8 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), dont 6 sessions invitées.
- 1 IEEE Int. Conf. on Advanced Robotics (ICAR).

3 symposia internationaux avec actes et comité de lecture :

- 1 IEEE RO-MAN, 1 IFAC ROCOND, 1 SIRS.

4 conférences nationales/francophones avec actes et comité de lecture (+1 durant le doctorat) :

- 3 RFIA, 1 GRETSI.

4 manifestations avec actes à diffusion limitée :

- 1 ORASIS, 1 IEEE HRI, 2 ECM²S.

VISIBILITÉ - FAITS MARQUANTS

Rapporteur d'une thèse de Doctorat, Australian National University, 2007.

Membre d'un jury de soutenance de thèse de Doctorat, LAAS-CNRS, 2006.

Responsabilité scientifique de projets :

- En cours : Porteur du projet ANR-JST BINAAHR (*Binaural Active Audition for Humanoid Robots*).

10 \I.1. FONCTIONS ACTUELLES – PARCOURS – VUE GLOBALE DES ACTIVITÉS

- En cours : Responsable français du projet STIC-AMSUD NCTVS (*Nonlinear Control Tools for Visual Servoing*).
- En cours : Porteur du projet AUDIO_HRP2 (*Capteur Audio Biomimétique et Fonctions Auditives pour le Robot Humanoïde HRP2*), financé par le LAAS-CNRS.

Participation significative à des projets :

- 2004-2010 : Participation au projet européen STREP COMMROB (Advanced Behaviour and High-level Multimodal Communication With and Among Robots) du FP6, et au projet intégré européen COGNIRON (The Cognitive Robot Companion) du FP6.
- 2002–2005 : Participation au pré-projet puis au projet EGOCENTRE du programme ROBEA du CNRS.

Séminaires significatifs :

- 5 séminaires invités : Université de Kyoto (2009) ; PUC Bogotá (2008 ×2) ; Université de Coimbra (2004) ; Workshop “Linear Matrix Inequalities in Control”, Toulouse (2004).
- 2 séminaires dans un GDR : GDR Robotique/ISIS (2006), GDR MACS (2005).
- 2 séminaires de pôle dans le cadre de la restructuration du LAAS-CNRS (2003).

Participation à l’organisation de conférences / sessions.

- Membre du comité local d’organisation de la conférence IFAC Triennal World Congress 2017, Toulouse, France.
- Co-organisateur des trois sessions invitées “Robot Audition” à la conférence IEEE/RSJ IROS’2010, Taipei, Taïwan.
- “Chair” de sessions à IEEE/RSJ IROS’2010 et IEEE CDC’2003

ENSEIGNEMENTS DISPENSÉS

- A. Introduction générale aux Systèmes. Cours : L1.
- B. Introduction à l’Automatique. TPs : L2.
- C. Analyse et Commande des systèmes linéaires invariants à temps continu : approche entrée-sortie (« fréquentielle »). Cours : L3. Correction d’annales : Agrégation. TDs : L3, L3Pro ou M1. TPs : L3, L3Pro ou M1.
- D. Analyse et Commande des systèmes linéaires invariants à temps discret : approche entrée-sortie. Cours/TD intégrés : L3. TDs : M1. TPs : M1.
- E. Analyse et Commande des systèmes linéaires invariants à temps continu ou discret dans l’espace d’état. Cours/TD intégrés : Agrégation. TDs : M1. TPs : M1 et Agrégation.
- F. Analyse et Commande des systèmes non linéaires : analyse de systèmes bouclés non linéaires par la méthode du premier harmonique ; analyse et commande dans le plan de phase. Cours/TD intégrés : Agrégations. TDs : L3. TPs : M1 et Agrégations.
- G. Analyse et Commande des systèmes à temps continu ou discret : approches entrée-sortie et techniques d’espace d’état. Introduction aux systèmes multidimensionnels. Cours/TDs intégrés : M2Pro. BEs : M2Pro.
- H. Identification paramétrique. Traitement numérique du signal. TPs : M1.
- I. Stabilité. Introduction à la théorie de LYAPUNOV et à la commande LQR. Introduction au régulateur RST. Cours/TDs intégrés : M1. TDs : M1. TPs : M1 et M2Pro.

- J. Introduction au Filtrage de KALMAN. Cours/TDs intégrés : M1 et M2Pro. TPs : M1 et M2Pro.
- K. Introduction à la navigation d'un robot mobile. TP : L1.
- L. Modélisation de robots manipulateurs. Cours : L3. TDs : L3.
- M. Introduction à la commande de robots manipulateurs. Cours/TDs intégrés : L3 et M1 International (en anglais). BEs : L3 et M1 International (en anglais).
- O. Travaux d'Étude et de Recherche d'Automatique. Nombreux thèmes : M1.
- P. Agrégations blanches, leçons et montages d'Automatique.
- Q. Formation au Projet Personnel et Professionnel.

RESPONSABILITÉS ADMINISTRATIVES ET PÉDAGOGIQUES

Responsabilité actuelle ou passée d'Unités d'Enseignements de M2, M1, L3, L2, L1. Participation à la définition et à la mise en place de contenus, à la rédaction de maquettes de diplômes, etc.

1999-2000 à 2003-2004 : Co-responsabilité puis responsabilité la salle de TP Automatique & Traitement du Signal de l'UPS.

Très fort investissement en Travaux Pratiques.

- Conception et déploiement de nombreux Travaux Pratiques et Travaux d'Étude et de Recherche.
- Pendant environ 10 ans, maintenance et évolution matérielle et logicielle de la salle de TPs Automatique & Traitement du Signal de l'UPS (MATLAB, "toolboxs" et matériel pour la commande par calculateur, etc.)
- Déploiement d'un réseau intranet, en collaboration avec un Ingénieur Réseau de l'UPS.

2002 : Président de la Commission d'Évaluation du DEUG Sciences et Technologie Industrielle, mention Sciences de la Matière.

De 1997-1998 à 2000-2001 : Centralisation des services d'enseignement de la 61^{ème} section UPS.
Participation annuelle au salon de l'enseignement supérieur INFOSUP.

TÂCHES D'INTÉRÊT COLLECTIF

Actuellement : Membre de l'équipe de direction de l'École Doctorale Systèmes de l'Université de Toulouse. Responsable (proposition, organisation, validation) de formations scientifiques et générales.

Actuellement : Référent auprès du Bureau d'Aide à l'Insertion Professionnelle de l'UPS pour certains Master 2.

De 2003-2004 à 2007-2008 : Membre élu suppléant de la CSE 61[°] section, UPS.

De 1999-2000 à 2003-2004 : Membre élu du Conseil de Département EEA de l'UPS et des commissions « DEUG » et « TPs » du Département. Membre élu du Conseil Pédagogique de l'UFR Physique-Chimie-Automatique de l'UPS.

De 1999-2000 à 2003-2004 : Membre élu titulaire de la CSE 61[°] section de l'UPS.

12 \I.2. SYNTHÈSE DES ACTIVITÉS DE RECHERCHE

De 2000-2001 à 2002-2003 : Membre nommé suppléant (2000-2001) puis titulaire des CSE mixtes 61°/63° sections et 27°/61° sections, INSA Toulouse.

De 1997-1998 à 1998-1999 : Membre de la commission de réforme du DEUG.

De 1997-1998 à 1998-1999 : Membre de la commission “DEUG” du Département EEA.

I.2 Synthèse des Activités de Recherche

Cette section propose un bilan condensé de mes activités de recherche. Du fait que mes travaux de Doctorat étaient très éloignés de mes thématiques actuelles, ils ne sont que brièvement référencés. Une description plus étendue figure dans l’annexe §I.4 en fin de chapitre. Ensuite, j’expose les grandes lignes des recherches que j’ai conduites consécutivement à ma réorientation scientifique – et à mon changement de groupe – en tant que Maître de Conférence. Celles-ci seront approfondies dans la deuxième partie du manuscrit.

Nos publications sont codifiées et recensées dans le §I.2.4, pages 19 et suivantes.

I.2.1 Travaux de Doctorat

I.2.1.1 Intitulé

Thèse de doctorat de l’Université Paul Sabatier, spécialité Automatique, sur le thème *Interface Symbolique-Numérique dans la Simulation Qualitative des Systèmes Dynamiques*.

Durée : 01.11.1991–20.01.1995. Lieu : LAAS-CNRS, opération « Automatique Qualitative et Supervision Experte » du groupe Décentralisation, Hiérarchisation, Parallélisme, puis, suite à une restructuration du laboratoire, groupe Automatique Symbolique. Directeur de thèse : M. Joseph AGUILAR-MARTIN, DR CNRS.

Rapporteurs : Mme Sylviane GENTIL (Prof. INP Grenoble), M. Laurent FOULLOY (Prof. Univ. Savoie, Annecy), M. Georges BITSORIS (Prof. Univ. Patras, Grèce).

Jury : Mme S. GENTIL, M. L. FOULLOY, M. G. BITSORIS, M. J. AGUILAR-MARTIN, Mme L. TRAVÉ-MASSUYÈS, M. M. COURDESSES.

Mention Très Honorable avec Félicitations du Jury.

I.2.1.2 Résumé condensé

Ce travail s’inscrivait dans le cadre de l’étude des systèmes dynamiques par l’utilisation conjointe des techniques de raisonnement qualitatif et des techniques classiques de l’Automatique. Sa problématique, dite de l’*interface symbolique-numérique*, concernait l’élaboration de méthodes permettant de déterminer la réponse d’un système représenté par son équation d’état à des entrées provenant d’une simulation qualitative, *i.e.* se présentant sous la forme de séquences d’intervalles indexées par rapport au temps.

Dans un premier temps, la représentation qualitative des signaux d’entrée a été complétée par des « enveloppes » délimitant un domaine compact de valeurs admissibles à chaque instant. Le problème a alors consisté en la détermination des enveloppes des réponses possibles du système, lorsque son vecteur d’état initial est lui-même décrit par des intervalles. Une reformulation en terme de problèmes de commande optimale à temps final fixé a été proposée, de même que les particularisations du principe du maximum de Pontriaguine nécessaires à leur traitement. Le cas des systèmes linéaires invariants ainsi que celui où la représentation

qualitative des entrées dégénère en des signaux déterministes ont fait l'objet d'une attention particulière.

Ces résultats ont ensuite été étendus à la détermination d'une approximation de la réponse d'un système à un signal obtenu par une simulation qualitative de type QSIM, où les instants qualitatifs sont eux-mêmes décrits par des intervalles temporels.

Enfin, l'approche a été complétée de façon à appréhender deux problèmes plus spécifiques intéressant la simulation qualitative : le cas où certains paramètres du système sont constants et définis par des intervalles a d'abord été traité ; ensuite, pour un système mono-entrée mono-sortie, une méthode permettant d'exploiter des intervalles de valeurs de la dérivée de l'entrée – lesquels sont parfois également fournis par une simulation de type QSIM – a été proposée en vue de déterminer plus précisément les évolutions possibles de la sortie.

Mots Clés : Raisonnement Qualitatif, Interface Symbolique-Numérique, Commande Optimale, Principe du Maximum, Signaux Qualitatifs, Signaux de type QSIM, Systèmes à Paramètres Qualitatifs.

Un résumé étendu est proposé dans une annexe en fin de ce chapitre (§I.4).

I.2.1.3 Production scientifique associée

Mes travaux de thèse ont donné lieu à une publication dans une revue internationale [JSI-1-THESE-1], une conférence internationale avec actes et comité de lecture [CIACL-01-THESE-1], à un colloque francophone avec actes et comité de lecture [CNACL-1-THESE-1], et à un exposé dans le groupe de recherche « Modélisation Qualitative et Diagnostic » du G.R. Automatique (1993). En raison de la réorientation scientifique consécutive à mon recrutement (§I.2.2), une partie conséquente demeure non publiée.

I.2.2 Bilan des thématiques de recherche postérieures au Doctorat

Lors de ma candidature au poste que j'occupe, j'ai clairement affiché auprès de mon ancienne équipe de recherche ma volonté d'entamer une réorientation scientifique. Suite à des discussions étendues avec le responsable et les membres du groupe Robotique et Intelligence Artificielle (RIA) du LAAS-CNRS, j'ai rejoint RIA pour y travailler sur des problématiques d'estimation et de commande.

À l'heure du « bilan d'étape » que constitue ce manuscrit, mes activités de recherche peuvent être partitionnées en les trois domaines suivants.

A. Détection et Estimation

Depuis mon intégration dans RIA, j'ai mené une activité continue dans ce thème. Il s'agit pour moi d'un travail de fond, principalement guidé par mon souhait d'étudier séquentiellement et en profondeur des sous-ensembles de la théorie de la détection et de l'estimation suscitant mon intérêt. Ce travail s'est systématiquement matérialisé au sein de collaborations scientifiques avec des collègues spécialistes de thématiques ou de domaines d'application connexes : d'abord à l'intérieur du groupe pour la *surveillance de scènes dynamiques*, le *suivi de personnes et de gestes dans des séquences video* et la *capture de mouvement par vision*, puis, actuellement, avec un laboratoire industriel pour le co-encadrement d'une thèse CIFRE en vue de développer une *nouvelle algorithmie de localisation pour le système ARGOS*. Le support théorique de ces activités est le *filtrage Bayésien* (filtrage de KALMAN linéaire, étendu, “unscented”, méthodes séquentielles

de Monte Carlo et Quasi Monte Carlo), la *détection statistique* (principalement les tests GLRT et MLRT – Generalized/Marginalized Likelihood Ratio Test –) et le *filtrage des systèmes à sauts Markoviens* (approches particulières, filtres GPB – Generalized Pseudo-Bayesian –, IMM – Interacting Multiple Model filter –, etc.). Les deux derniers points ont fait l'objet d'une focalisation sur le cas ouvert de modèles d'état d'ordres « hétérogènes ».

B. Asservissement Visuel

Je me suis inscrit dans cette thématique dès le début de ma réorientation scientifique. Le but principal est de développer des méthodologies et outils pour *l'analyse et la synthèse multicritères* de commandes référencées vision en Robotique, *i.e.* qui prennent en compte l'ensemble des contraintes (visibilité, saturation d'actionneurs, contraintes sur le déplacement 3D, etc.). Il s'agit d'un problème ouvert. Mes travaux sont fondés sur la *théorie de LYAPUNOV* et la *programmation semi-définie* (SDP = semi-definite programming) ou *optimisation LMI* (LMI = Linear Matrix Inequalities), et présentent de fortes connexions avec la *commande robuste*. Par dualité, la *localisation visuelle* est traitée par des techniques de *filtrage ensembliste robuste*. Sur le plan logiciel, j'ai conçu et co-développé – par l'encadrement de stagiaires – l'environnement de simulation MAVS (MAVS = “MAVS Ain’t Visual Servoing”) pour la *Simulation de Commandes Référencées Capteurs Extéroceptifs (Vision & Laser)*. MAVS est basé sur le logiciel de CACSD MATLAB et sur le moteur de rendu 3D GDHE développé par Matthieu HERRB au LAAS-CNRS (<http://www.laas.fr/~matthieu/gdhe>).

C. Audition en Robotique

En 2004, j'ai participé aux premiers travaux de RIA dans cette thématique relativement nouvelle, dont je suis actuellement le porteur. Mes contributions ont concerné le *développement d'un capteur auditif intégré*, l'*élaboration de primitives de détection, localisation et extraction (filtrage spatial) de sources sonores*, ainsi que *la détection et le rehaussement de “patterns” audio*. Sur le plan théorique, elles s'inscrivent dans le champ très vaste de l'*antennerie acoustique* (formation de voie, méthodes à haute résolution, détection basée sur des critères d'information), et, plus récemment, dans le contexte du *filtrage adapté stochastique*.

En résumé, la majorité de mes activités de recherche se situe à la frontière de l'Automatique, du Traitement du Signal, et de la Robotique. Je suis naturellement amené à étudier les potentialités offertes par des techniques de l'Automatique ou du Traitement du Signal pour la résolution de problèmes de Robotique. Cependant, je veille autant que possible à ce que mes travaux ne se limitent pas à une simple application de méthodes développées par ailleurs. En effet, la Robotique, par ses spécificités ou ses contraintes, « questionne » très souvent l'état de l'art en Automatique et en Traitement du Signal. Ma démarche consiste alors à tenter de proposer des extensions méthodologiques théoriquement étayées, qui correspondent aux besoins du problème considéré mais qui, par leur caractère générique, débordent éventuellement du cadre strict de la Robotique.

Sur un plan plus personnel, ce dialogue entre la Robotique – par la variété, la richesse et la complexité de ses problèmes – et les théories de l'Automatique et du Signal – dont j'apprécie au plus haut point la rigueur et l'esthétique – de même que la possibilité qu'il m'offre de contribuer modestement à l'une ou l'autre des disciplines, constitue une immense source de satisfaction dans mon activité de recherche. Celle-ci est complétée par une collaboration scientifique exaltante située hors de la Robotique, dont les domaines d'application recouvrent la

gestion et la protection des écosystèmes marins, la sécurité maritime, le suivi d'aide humanitaire, etc.

Les travaux effectués dans chacune de mes thématiques de recherche seront passés en revue dans le Chapitre II (§II.1,§II.2,§II.3). Ayant eu la quasi-totale « maîtrise d'ouvrage » et « maîtrise d'œuvre » de mes recherches en « Asservissement Visuel Multicritères » et en « Audition en Robotique », celles-ci seront ensuite détaillées dans des chapitres dédiés (Chapitres III et IV). En revanche, comme mentionné précédemment, mes activités en « Détection et Estimation » se sont inscrites dans différents contextes d'application qui ne forment pas nécessairement un ensemble cohérent et dont je ne prétends pas être spécialiste. Ainsi, bien que j'aie participé à la définition des objectifs et à l'obtention des résultats dans les encadrements et projets associés, l'essentiel des compétences relevant spécifiquement de ces contextes a été apporté par des collègues. C'est pourquoi il me semble opportun de ne proposer qu'un résumé étendu de ces activités, exclusivement focalisé sur la nature des problèmes de détection et d'estimation sous-jacents ainsi que sur les solutions apportées (§II.1). Ce travail de fond sera toutefois mis à profit dans des prospectives à plus long terme, lorsque sera examinée en profondeur la « jonction » de certaines thématiques actuellement développées indépendamment.

I.2.3 Encadrements

I.2.3.1 Post-Doctorant

Je collabore depuis mi-Mai 2010 avec un Post-Doctorant, que j'ai recruté sur le « projet LAAS » AUDIO_HRP2 [PROJ-R-1].

[POSTDOC-1] M. Alain SKAF, Docteur en Acoustique de l'Université de Valenciennes et du Hainaut Cambrésis. *Conception et Implémentation d'un Capteur Auditif Bioinspiré Reconfigurable pour le Robot Humanoïde HRP-2.* 15.05.2010–30.11.2011.

I.2.3.2 Thèses de Doctorat

Le pourcentage associé à chaque thèse en co-encadrement est approximé par un élément de l'ensemble {25 %, 50 %, 75 %}.

Je co-encadre actuellement les doctorants suivants, inscrits dans l'École Doctorale Systèmes.

[DOC-07] M. Rémy LOPEZ. *Développement d'une Nouvelle Algorithmie de Localisation adaptée à l'Ensemble des Mobiles Suivis par le Système ARGOS.*

Thèse CIFRE. Début : 01.01.2010. Co-encadrement (25 %) avec M. Philippe GASPAR, Directeur du Département Océanographie Spatiale (HdR 2009) de la société Collecte Localisation Satellites (CLS, www.cls.fr, 31520 Ramonville), en collaboration avec MM. Jean-Pierre MALARDÉ et François ROYER, également chercheurs à CLS.

[DOC-06] M. Julien BONNAL. *Développement d'un Capteur Sonore et de Fonctions Auditives pour l'Interaction Homme-Robot.*

Début : 01.11.2007. Co-encadrement (75 %) avec M. Marc RENAUD, Professeur à l'INSA de Toulouse. Soutenance prévue en Mars 2011.

16 \I.2. SYNTHÈSE DES ACTIVITÉS DE RECHERCHE

[DOC-05] M. Sylvain DUROLA. *Asservissement Multicapteurs Multicritères et Inégalités Matricielles.*

Début : 01.10.2007. Co-encadrement (75 %) avec M. Michel COURDESESSES, Professeur à l'UPS. Soutenance prévue en Juin 2011.

Dans le passé, j'ai obtenu l'autorisation d'encadrer ou de co-encadrer à titre individuel les Thèses de Doctorat suivantes.

[DOC-04] M. Mathias FONTMARTY. *Vision et Filtrage Particulaire pour le Suivi Tridimensionnel de Mouvement Humain. Applications à la Robotique.*

01.10.2005–02.12.2008. Co-encadrement (50 %) avec M. Frédéric LERASLE, Professeur en Vision par Ordinateur à l'UPS (MCf durant cette thèse).

Rapporteurs : Mme Marie-Odile BERGER (CR HdR INRIA, Réalité augmentée, modélisation 3D, fusion) et M. Patrick PÉREZ (DR INRIA, Suivi Visuel par filtrage particulaire). Jury : M. P. DALLE, Mme M.O. BERGER, M. P. PÉREZ, Mme J. BENOIS-PINEAU, M. M. DEVY.

Monsieur FONTMARTY occupait un poste d'ATER à l'INSA de Toulouse en 2009-2010, et a été récemment embauché par l'entreprise NOOMEO (www.noomeo.eu).

[DOC-03] M. Sylvain ARGENTIERI. *Conception d'un Capteur Sonore pour la Localisation de Source en Robotique Mobile.*

01.10.2003–08.12.2006. Co-encadrement (75 %) avec M. Philippe SOUÈRES, DR au LAAS-CNRS (CR HdR durant cette thèse).

Rapporteurs : MM Hiroshi OKUNO (Prof. U. Tokyo, "Robot Audition") et Frédéric BIMBOT (CR HdR INRIA, Traitement du Signal Audio). Jury : MM. P. PÉREZ, J.L. SCHWARTZ, Y. DEVILLE, M. DEVY.

Monsieur ARGENTIERI occupe actuellement un poste de Maître de Conférence en Signal-Robotique à l'Université Pierre et Marie Curie, et est rattaché à l'Institut des Systèmes Intelligents et de Robotique, Paris VI.

[DOC-02] M. Ludovic BRÈTHES. *Suivi Visuel par Filtrage Particulaire. Application à l'Interaction Homme-Robot.*

01.10.2002–13.12.2005. Co-encadrement (50 %) avec M. Frédéric LERASLE, Professeur à l'Université Paul Sabatier (MCf durant cette thèse).

Rapporteurs : MM Frédéric JURIE (CR HdR CNRS, Vision par Ordinateur) et Patrick PÉREZ (DR INRIA, Suivi Visuel par filtrage particulaire). Jury : MM R. CHATILA, F. JURIE, P. PÉREZ, P. DALLE.

Monsieur BRÈTHES est co-fondateur de l'entreprise NOOMEO (www.noomeo.eu), spécialisée dans la Vision 3D.

[DOC-01] M^{lle} Delphine BELLOT. *Contribution à l'Analyse et à la Synthèse de Schémas de Commande Référencée Vision.*

01.11.1999–19.12.2002. Encadrement à 100 %.

Rapporteurs : MM François CHAUMETTE (DR INRIA, Asservissement Visuel) et Laurent EL GHAOUI (Professeur U. Berkeley, Commande LMI). Jury : M. R. CHATILA, M. F. CHAUMETTE, M^{lle} S. TARBOURIECH, M. P. MARTINET, M. M. COURDESESSES.

Mademoiselle BELLOT occupe actuellement un poste de Maître de Conférence en Mécanique-Automatique-Robotique à l'Université Pierre et Marie Curie, et est rattachée à l'Institut des Systèmes Intelligents et de Robotique, Paris VI.

J'ai également participé officieusement à l'encadrement d'une partie du travail des thèses de Doctorat suivantes.

- [DOC-0b] M. Carlos DIAZ HERNANDEZ. *Estimation de Mouvement pour l'Asservissement Visuel.*
Début 2008–Mai 2009. Co-encadrement (33 %) avec M^{lle} Viviane CADENAT, Maître de Conférence à l'UPS et M. Thierry SENTENAC, Maître-Assistant à l'EMAC.
La thèse de Monsieur DIAZ HERNANDEZ a été interrompue à mi-parcours.
- [DOC-0a] M. Thierry SENTENAC. *Surveillance de Scènes Dynamiques avec une Caméra CCD dans le Proche Infrarouge : Application à la Détection Couplée de Feu et de Déplacement d'Objets.* Thèse préparée au LAAS-CNRS et à l'École des Mines d'Albi-Carmaux (EMAC), financée par un contrat impliquant l'entreprise LATÉCOÈRE. Présentée le 8.11.2002 à huis-clos, clause de confidentialité jusqu'à fin 2007.
Les résultats que j'ai encadrés sont consignés dans le Chapitre 4 du manuscrit (§1 : Caractéristique des phénomènes à détecter || §2 : Détection des phénomènes avec une même technologie || §3 : Mesure couplée des phénomènes || §4 : *Estimation de l'évolution des phénomènes* || §5 : Application à la surveillance de soutes d'avion).
M. SENTENAC est Maître-Assistant à l'EMAC.

I.2.3.3 Stages de M2R/DEA

J'ai encadré ou co-encadré les stages de recherche de M2R ou DEA suivants.

- [M2R-10] M. Anh Dung NGUYEN, 5^o année INSA Toulouse. *Analyse Théorique et Implémentation sur FPGA d'une Extension Large Bande de la Méthode MUSIC pour la Localisation de Sources Sonores en Robotique.* 02.2009–08.2009. Co-encadrement avec M. Jérôme MANHÈS, Ingénieur de Recherche au LAAS-CNRS.
- [M2R-09] M. Julien BONNAL, 3^o année ISEN Toulon et M2R à l'Université de Toulon. *Suivi Auditif et Visio-Auditif pour l'Interaction Homme-Robot.* 03.2007–09.2007. Co-encadrement avec M. Sylvain ARGENTIERI.
- [M2R-08] M^{lle} Hind MESTOURI, diplômée de l'Université de Fès. *Formation de Voie Large Bande Optimisée pour la Localisation de Source Sonore en Robotique Mobile.* 02.2006–06.2006. Co-encadrement avec M. Sylvain ARGENTIERI.
- [M2R-07] M. Thrasyvoulos Aris CHRONOPOULOS, diplômé en Génie Électrique à l'Université d'Athènes. *Application de Techniques de Commande LMI à l'Analyse et la Synthèse de Commandes Référencées Vision.* 03.2005–09.2005. Encadrement à 100 %.
- [M2R-06] M. Mathias FONTMARTY, 5^o année INSA Toulouse. *Suivi Spatial d'un Bras Humain à l'aide d'un Système Stéréoscopique.* 02.2005–07.2005. Co-encadrement avec M. Frédéric LERASLE.
- [M2R-05] M^{lle} Zhongjian HUA, 3^o année ENSEEIHT et année spéciale Automatique. *Estimation de Mouvements à partir d'Informations Visuelles.* 02.2003–09.2003. Co-encadrement avec M. Thierry SENTENAC.
- [M2R-04] M. Yann ROTROU, 5^o année INSA Toulouse. *Suivi d'Objets Déformables par Vision.* 02.2002–06.2002. Co-encadrement avec M. Frédéric LERASLE.

18 \I.2. SYNTHÈSE DES ACTIVITÉS DE RECHERCHE

- [M2R-03] M. Marc BELLETTI, diplômé de l’Université Paul Sabatier. *Étude d’une Approche LMI pour la Localisation d’un Robot Mobile à partir d’Informations Visuelles.* 02.2002–06.2002. Co-encadrement avec M^{lle} Delphine BELLOT.
- [M2R-02] M. Vincent RUGGIERI, 3[°] année ingénieur ENSEEIHT. *Détection et Estimation de Déplacements d’Objets.* 02.2000–07.2000. Co-encadrement avec M. Thierry SENTENAC.
- [M2R-01] M^{lle} Delphine BELLOT, 3[°] année ENSEEIHT et année spéciale Automatique. *Estimation de Mouvements à partir d’Informations Visuelles.* 04.1999–09.1999. Encadrement à 100 %.

I.2.3.4 Encadrement de stages Ingénieur, M1 ou de Visiteurs

- [STAGE-15] MM. T. MIZUMOTO et T. YOSHIDA, Doctorants à OKUNO’s Lab, Université de Kyoto. *Portage du logiciel Open Source HARK.* Séjour d’un mois, 11.2010–12.2010. Encadrement à 100 %.
- [STAGE-14] M. Jérémie GROSSMANN, M1 Systèmes Intelligents, UPS. *Implémentation d’un Système de Commande Générique sur un Robot 6 Axes.* Fin 02.2010–Fin 07.2010. Encadrement avec M^{lle} Viviane CADENAT.
- [STAGE-13] M^{lle} Lise MARTIN, MM. Alban PORTELLO et El Houssin EL BOUCHIKHI, 3[°] année ingénieur ENSEEIHT. *Localisation de Source en Robotique. Prototypage d’une Formation de Voie par des Techniques de Convolution Rapide sur Synplify DSP.* Projet Long, 11.2009–03.2010. Co-encadrement avec M. Jérôme MANHÈS.
- [STAGE-12] M. Damien DUFRESNE, 2[°] année SupÉlec. *Localisation de Source en Robotique. Étude de l’Antenne Semi-Circulaire.* Fin 06.2009–Fin 08.2009. Encadrement à 100 %.
- [STAGE-11] M^{lle} Estelle COMMENT, 2[°] année École Polytechnique. *Estimation de Mouvement pour l’Asservissement Visuel en Contexte Dynamique.* 04.2009–08.2009. Co-encadrement avec M^{lle} Viviane CADENAT et M. Thierry SENTENAC.
- [STAGE-10] M. Nicolas FERRERO, 2[°] année IUT A, UPS. *Intégration d’un Module « Formation de Voie » dans une Architecture VHDL.* 04.2009–07.2009. Co-encadrement avec M. Jérôme MANHÈS.
- [STAGE-09] M. Thierry BESSIÈRE, Master 1 UPS. *Implémentation de Primitives sur FPGA pour la Localisation de Source Sonore en Robotique Mobile.* 02.2008–09.2008. Co-encadrement avec M. Jérôme MANHÈS.
- [STAGE-08] M^{lle} Caroline P. CARVALHO, Ingénieur PUCRS, Porto Alegre, Brésil. *Développement d’un Environnement de Simulation de Commandes Référencées Capteurs Extéroceptifs.* 02.2008–07.2008. Encadrement à 100 %.
- [STAGE-07] M^{lle} Julie VAN DER HAM, 5[°] année INSA Toulouse. *Estimation Optimale de Trajectoires d’Animaux Marins.* 09.2006–03.2007. Co-encadrement avec M. Philippe GASPAR, Directeur du Département Océanographie Spatiale de l’entreprise Collecte Localisation Satellites (CLS, www.cls.fr, 31520 Ramonville).
- [STAGE-06] M. Gaël MARTINEZ, 2[°] année ingénieur ENSEEIHT. *Analyse Multicritère d’un Schéma Élémentaire de Commande Référencée Vision.* Fin 06.2006–Début 09.2006. Encadrement à 100 %.

- [STAGE-05] M. Abhishek SHARMA, visiteur de Institute of Techology / Banaras Hindu University, Varanasi, Inde. *Analyse d'un Schéma de Commande Référencée Vision Élémentaire par une Approche Bi-Quadratique.* 05.2002–07.2002. Encadrement à 100 % (“summer internship” de 3 mois, niveau Master 1).
- [STAGE-04] M. Sylvain OUDINET, 3^e année ingénieur ISEN Lille. *Conception et Implémentation d'un Environnement de Simulation de Commandes Référencées Vision.* 02.2001–08.2001. Encadrement à 100 %.
- [STAGE-03] M. Sébastien DESRUMAUX, 3^e année ingénieur ISEN Lille. *Conception et Implémentation d'un Environnement de Simulation de Commandes Référencées Vision.* 02.2000–08.2000. Encadrement à 100 %.
- [STAGE-02] M. Mohand Saïd DJOUADI, stagiaire visiteur de l'ENITA, Alger, Algérie. *Estimation de Mouvements 2D à partir d'Informations Visuelles pour des Tâches Robotiques.* 11.1998–02.1999. Encadrement à 100 %.
- [STAGE-01] M. Markus MÜLLER, Mastère de l'Université de Karlsruhe, Allemagne. *Conception et Réalisation d'une Commande Référencée Vision d'un Bras Manipulateur.* 04.1996–12.1996. Encadrement à 100 %.

I.2.4 Publications

Les publications sont référencées ci-dessous selon le format

[TYPE - NUMÉRO_GLOBAL - DOMAINE - SOUS_NUMÉRO]

où

- TYPE est l'acronyme de la catégorie de la publication : COCI = Contribution à un Ouvrage Collectif International, JSI = Journal Scientifique International, etc. ;
- NUMÉRO désigne le numéro de la publication dans sa catégorie ;
- DOMAINE correspond au domaine scientifique : AV = Asservissement Visuel ; DE = Détection-Estimation ; AR = Audition en Robotique ; THESE = Thèse ; MISC = Autre ;
- SOUS_NUMÉRO indice la publication dans sa catégorie et son domaine scientifique.

I.2.4.1 Contribution à un ouvrage collectif international

- [COCI-1-AV-1] P. DANÈS, D.F. COUTINHO, S. DUROLA. *Multicriteria Analysis of Visual Servos through Rational Systems, Biquadratic Lyapunov Functions, and LMIs.* Chapitre 10 de *Visual Servoing via Advanced Numerical Methods*, G. CHESI & K. HASHIMOTO Eds, pp. 169–188. Lecture Notes in Control and Information Sciences. Springer, Vol. 401/2010. DOI : 10.1007/978-1-84996-089-2_10.

I.2.4.2 Journaux scientifiques

- [JSI-7-AR-1] J. BONNAL, S. ARGENTIERI, P. DANÈS, J. MANHÈS, P. SOUÈRES, M. RENAUD. *The EAR Project.* Journal of the Robotics Society of Japan, special issue on Robot Audition, Vol.28, N.1, pp. 10–13, 2010.
- [JSI-6-DE-2] M. FONTMARTY, P. DANÈS, F. LERASLE. *Evaluations of Particle Filter based Human Motion Visual Trackers for Home Environment Surveillance.* Int. Journal of Pattern Recognition and Artificial Intelligence, Vol.23, N.7, pp. 1333–1368, 2009. DOI : 10.1142/S0218001409007661.

- [JSI-5-DE-1] L. BRÈTHES, F. LERASLE, P. DANÈS, M. FONTMARTY. *Particle Filtering Strategies for Data Fusion dedicated to Visual Tracking from a Mobile Robot*. Machine Vision and Applications, Vol.21, N.4, pp. 427–448, 2010 (en ligne depuis 2008). DOI : 10.1007/s00138-008-0174-7.
- [JSI-4-AV-2] D.F. COUTINHO, M. FU, A. TROFINO, P. DANÈS. *L₂-Gain Analysis and Control of Uncertain Nonlinear Systems with Bounded Disturbance Inputs*. Int. Journal of Robust and Nonlinear Control, Vol.18, pp. 88-110, 2008. DOI : 10.1002/rnc.1207.
- [JSI-3-MISC-1] B. ONILLON, P. DANÈS, B. BENAZET, O. LLOPIS. *An Optical Link for Microwave Clock Distribution using Optical Carrier Suppression and DC Drift Compensation*. Microwave and Optical Technology Letters, Vol.49, N.7, pp. 1634-1637, 2007.
- [JSI-2-AV-1] P. DANÈS, D. BELLOT. *Towards an LMI Approach to Multicriteria Visual Servoing in Robotics*. European Journal of Control, Special Issue on “Linear Matrix Inequalities in Control”, Vol.12, N.1, pp. 86–110, 2006. DOI : 10.3166/ejc.12.86-110. “Companion Report” étendu (62p.) sur <http://www.laas.fr/~danes/RESEARCH/VS/companionEJC2k5.pdf>.
- [JSI-1-THESE-1] P. DANÈS, J. AGUILAR-MARTIN. *The Symbolic-Numeric Interface : a “zosteric” approach*. Applied Artificial Intelligence Journal, Vol.9, N.5, pp. 451–478, 1995. DOI : 10.1080/08839519508945485.

I.2.4.3 Conférences Internationales avec Actes et Comité de Lecture

Remarque : sauf mention contraire explicite, les articles répertoriés ci-dessous ont fait l’objet d’une présentation orale. Les contributions à des sessions invitées ont été consécutives à des sollicitations des organisateurs, mais ont suivi le processus de review habituel.

- [CIACL-24-DE-9] R. LOPEZ, P. DANÈS, F. ROYER. *Extending the IMM Filter to Heterogeneous-Order State Space Models*. IEEE Conf. on Decision and Control (CDC’2010). Atlanta, GA, 2010.
- [CIACL-23-AR-8] J. BONNAL, P. DANÈS, M. RENAUD. *Detection of Acoustic Patterns by Stochastic Matched Filtering*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS’2010). Taipei, Taiwan, 2010. Session invitée “Robot Audition”.
- [CIACL-22-AR-7] P. DANÈS, J. BONNAL. *Information-Theoretic Detection of Broadband Sources in a Coherent Beamspace MUSIC Scheme*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS’2010). Taipei, Taiwan, 2010. Session invitée “Robot Audition”.
- [CIACL-21-AR-6] J. BONNAL, S. ARGENTIERI, P. DANÈS, J. MANHÈS. *Speaker Localization and Speech Extraction with the EAR Sensor*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS’2009), pp. 670–675. Saint-Louis, MO, 2009. Session invitée “Robot Audition”.
- [CIACL-20-DE-8] M. FONTMARTY, F. LERASLE, P. DANÈS. *Likelihood Tuning for Particle Filter in Visual Tracking*. IEEE Int. Conf. on Image Processing (ICIP’2009), pp. 4101–4104. Le Caire, Égypte, 2009. Accepté en tant que poster.

- [CIACL-19-DE-7] M. FONTMARTY, P. DANÈS, F. LERASLE. *Quasi Monte Carlo Partitioned Filtering for Visual Human Motion Capture*. IEEE Int. Conf. on Image Processing (ICIP'2009), pp. 2553–2556. Le Caire, Égypte, 2009. Accepté en tant que poster.
- [CIACL-18-AV-6] S. DUROLA, P. DANÈS, D.F.. COUTINHO, M. COURDESES. *Rational Systems and Matrix Inequalities to the Multicriteria Analysis of Visual Servos*. IEEE Int. Conf. on Robotics and Automation (ICRA'2009), pp. 1504–1509. Kobe, Japon, 2009. Session invitée “Visual servoing via advanced numerical methods”.
- [CIACL-17-AV-5] S. DUROLA, P. DANÈS, D.F. COUTINHO. *Set-membership Filtering for Uncertain Discrete-Time Rational Systems through Recursive Algebraic Representations and LMIs*. IEEE Conf. on Decision and Control (CDC'2008), pp. 684–689. Cancún, Mexique, 2008.
- [CIACL-16-DE-6] M. FONTMARTY, F. LERASLE, P. DANÈS. *Towards Real-Time Markerless Human Motion Capture from Ambiance Cameras using an Hybrid Particle Filter*. IEEE Int. Conf. on Image Processing (ICIP'2008), pp. 709–712, San Diego, CA, 2008.
- [CIACL-15-AR-5] S. ARGENTIERI, P. DANÈS. *Broadband Variations of the MUSIC High-Resolution Method for Sound Source Localization in Robotics*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'2007), pp. 2009–2014. San Diego, CA, 2007. Session invitée “Robot Audition”.
- [CIACL-14-DE-5] M. FONTMARTY, F. LERASLE, P. DANÈS. *Data Fusion within a Modified Annealed Particle Filter dedicated to Human Motion Capture*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'2007), pp. 3391–3396. San Diego, CA, 2007.
- [CIACL-13-DE-4] T. GERMA, F. LERASLE, P. DANÈS, L. BRÈTHES. *Visual Modalities dedicated to a Tour-Guide Robot*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'2007), pp. 3448–3453. San Diego, CA, 2007.
- [CIACL-12-AR-4] S. ARGENTIERI, P. DANÈS. *Convex Optimization and Modal Analysis for Beamforming in Robotics : Theoretical and Implementation Issues*. European Signal Processing Conference (EUSIPCO'2007), pp. 773–777. Poznán, Pologne, 2007.
- [CIACL-11-AV-4] D.F. COUTINHO, P. DANÈS. *Piecewise Lyapunov Functions to the Stability Analysis of Rational Systems subject to Multiple State Constraints*. IEEE Conf. on Decision and Control (CDC'2006), pp. 5801–5806. San Diego, CA, 2006.
- [CIACL-10-AR-3] S. ARGENTIERI, P. DANÈS, P. SOUÈRES. *Modal Analysis Based Beamforming for Nearfield or Farfield Speaker Localization in Robotics*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'2006), pp. 866–871. Pékin, Chine, 2006. Session invitée “Robot Audition”.
- [CIACL-09-AR-2] S. ARGENTIERI, P. DANÈS, P. SOUÈRES, P. LACROIX *An Experimental Testbed for Sound Source Localization with Mobile Robots using Optimized Wideband Beamformers*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'2005), pp. 909–914. Edmonton, Canada, 2005. Session invitée “Robot Audition”.

- [CIACL-08-AR-1] S. ARGENTIERI, P. DANÈS, P. SOUÈRES. *Prototyping Filter-Sum Beamformers for Sound-Source Localization in Mobile Robotics*. IEEE Int. Conf. on Robotics and Automation (ICRA'2005), pp. 3562–3567. Barcelone, Espagne, 2005.
- [CIACL-07-DE-3] L. BRÈTHES, F. LERASLE, P. DANÈS. *Data Fusion for Visual Tracking dedicated to Human-Robot Interaction*. IEEE Int. Conf. on Robotics and Automation (ICRA'2005), pp. 2087–2092. Barcelone, Espagne, 2005.
- [CIACL-06-AV-3] D. BELLOT, P. DANÈS. *An LMI solution to visual-based localization as the dual of visual servoing*. IEEE Conf. on Decision and Control (CDC'2003), pp. 5420–5425. Maui, HI, 2003.
- [CIACL-05-DE-2] P. MENEZES, L. BRETHÈS, F. LERASLE, P. DANÈS, J. DIAS. *Visual tracking of silhouettes for human-robot interaction*. IEEE Int. Conf. on Advanced Robotics (ICAR'2003), pp. 971–976. Coimbra, Portugal, 2003.
- [CIACL-04-AV-2] D. BELLOT, P. DANÈS. *Handling visual servoing schemes through rational systems and LMIs..* IEEE Conf. on Decision and Control (CDC'2001), pp. 3601–3606. Orlando, FL, 2001.
- [CIACL-03-DE-1] F. LERASLE, P. DANÈS. *Projected light beams tracking for efficient 3D reconstruction*. IEEE Int. Conf. on Image Processing (ICIP'2001), pp. 951–954. Thessalonique, Grèce, 2001.
- [CIACL-02-AV-1] D. BELLOT, P. DANÈS. *Towards an LMI approach to multicriteria visual servoing*. European Control Conference (ECC'2001), pp. 2930–2935. Porto, Portugal, 2001.
- [CIACL-01-THESE-1] P. DANÈS, L. TRAVÉ-MASSUYÈS, J. AGUILAR-MARTIN. *A Generic Method for Computing the Response of a Numerically-Known Dynamic System to Qualitative Inputs*. IEEE Conf. on Systems, Man and Cybernetics (SMC'1993), pp. 187–192, Le Touquet, France, 1993.

I.2.4.4 Symposia Internationaux avec Actes et Comité de Lecture

- [SIACL-3-DE-2] M. FONTMARTY, F. LERASLE, P. DANÈS. *Visual Human Motion Capture from a Mobile Robot*. IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN'2009), pp. 829–834. Toyama, Japon, 2009.
- [SIACL-2-AV-1] P. DANÈS, D.F. COUTINHO. *Multicriteria Analysis of Visual Servos via Biquadratic Lyapunov Functions*. IFAC Symp. on Robust Control Design (ROCOND'2006). Toulouse, France, 2006.
- [SIACL-1-DE-1] P. DANÈS, M.S. DJOUADI, D. BELLOT. *A 2-D point-wise motion estimation scheme for visual-based robotic tasks*. Int. Symp. on Intelligent Robotic Systems (SIRS'1999). Coimbra, Portugal, 1999.

I.2.4.5 Conférences Nationales avec Actes et Comité de Lecture

- [CNACL-5-DE-3] M. FONTMARTY, F. LERASLE, P. DANÈS. *Paramétrisation des Fonctions de Vraisemblance et Choix de Filtres Particulaires pour le Suivi Visuel*. Reconnaissance des Formes et Intelligence Artificielle (RFIA'2010). Caen, 2010.

- [CNACL-4-DE-2] M. FONTMARTY, F. LERASLE, P. DANÈS. *Une Stratégie Hybride de Filtrage Particulaire pour le Suivi de Mouvement Humain depuis un Robot Mobile.* Reconnaissance des Formes et Intelligence Artificielle (RFIA'2008). Amiens, 2008.
- [CNACL-3-DE-1] L. BRÈTHES, P. DANÈS, F. LERASLE. *Stratégies de Filtrage Particulaire pour le Suivi Visuel de Personnes : Description et Évaluation.* Reconnaissance des Formes et Intelligence Artificielle (RFIA'2006). Tours, 2006.
- [CNACL-2-AR-1] S. ARGENTIERI, P. DANÈS, P. SOUÈRES. *Conception et Implémentation d'un Système de Localisation de Source Sonore Large Bande par Formation de Voie pour des Applications Robotiques.* Colloque Francophone GRETSI sur le Traitement du Signal et des Images (GRETSI'2005). Louvain-la-Neuve, Belgique, 2005. Accepté en tant que poster.
- [CNACL-1-THESE-1] P. DANÈS, L. TRAVÉ-MASSUYÈS, J. AGUILAR-MARTIN. *Représentation Symbolique de la Réponse de Systèmes Linéaires à des Signaux Définis sur des Intervalles.* Colloque Francophone GRETSI sur le Traitement du Signal et des Images (GRETSI'1993). Juan les Pins, 1993. Accepté en tant que poster.

I.2.4.6 Manifestations avec Actes à Diffusion Limitée

- [MADL-4-DE-2] M. FONTMARTY, F. LERASLE, P. DANÈS, P. MENEZES. *Filtrage Particulaire pour la Capture de Mouvement dédiée à l'Interaction Homme-Robot.* Congrès francophone ORASIS'2007 des jeunes chercheurs en vision par ordinateur. Obernai, France, 2007.
- [MADL-3-DE-1] L. BRÈTHES, F. LERASLE, P. DANÈS. *Particle Filtering Strategies for Visual Tracking dedicated to H/R Interaction.* EURON Int. Workshop on Human-Robot Interaction (HRI'2006). Palerme, Italie, 2006.
- [MADL-2-AV-2] D. BELLOT, P. DANÈS. *An LMI approach to multicriteria visual servoing and its dual problem visual-based localization.* Workshop on Electronics, Control, Modelling, Measurement and Signals (ECM²S), Liberec, République Tchèque, 2003.
- [MADL-1-AV-1] D. BELLOT, P. DANÈS. *Analysis and synthesis of a multicriteria visual servoing scheme through robust linear control theory.* Workshop on Electronics, Control, Modelling, Measurement and Signals (ECM²S), Toulouse, France, 2001.

I.2.5 Visibilité - Faits marquants - Animation scientifique

I.2.5.1 Rédaction de Rapport de Présoutenance de Thèse de Doctorat

- [RAPPORT-1] *Rapporteur du manuscrit de thèse* de Michael Ian Yang WILLIAMS, intitulé *Fundamental Limits of Space-Time Processing*, Australian National University (ANU), 2007. Encadrant : Prof. Rodney KENNEDY (l'HdR n'est pas requise par l'ANU pour rapporter).

I.2.5.2 Participation à un Jury de Soutenance de Thèse de Doctorat

[JURY-1] À la suite de ma contribution au prototypage d'un compensateur de dérive d'un modulateur électro-optique avec le groupe Composants et Intégration des Systèmes Hyperfréquences pour Télécommunications du LAAS-CNRS, j'ai été membre du jury de soutenance de la Thèse de Doctorat de M. Bertrand ONILLON sur le thème *Liaisons optiques faible bruit pour la distribution de références de fréquences micro-ondes* (LAAS-CNRS, 17.10.2006) et co-auteur de l'article [JSI-3-MISC-1].

I.2.5.3 Responsabilité Scientifique de Projets de Recherche

[PROJ-R-3] Porteur du projet ANR-JST BINAAHR (*Binaural Active Audition for Humanoid Robots*), accepté en 2009 dans le cadre du programme blanc international France-Japon (JST = Japan Science and Technology Society). Partenaires : LAAS-CNRS, Toulouse ; Institut des Systèmes Intelligents et de Robotique (ISIR), Paris ; Laboratoire de Psychologie de la Perception (LPP), Paris ; Université de Kyoto (H.G. OKUNO's Lab, Prof. OKUNO étant coordinateur japonais) ; Tokyo Institute of Technology (K. NAKADAI) ; Kumamoto University (M. KUMON). Projet notifié le 01.03.2010, devant théoriquement se terminer le 28.02.2013.

Contenu global du projet :

- Conception et implémentation d'un capteur binaural.
- Définition et intégration de fonctions auditives actives bas-niveau : localisation active de sources sonores ; séparation géométrique de sources et détection d'activité vocale ; annulation du bruit propre au robot ; implémentation.
- Interaction auditive haut-niveau : reconnaissance de locuteur ; reconnaissance de parole ; interaction multi-partenaires.
- Approche « psychologie de la perception » de la perception auditive active, et développement d'une plateforme expérimentale.
- Intégration sur une tête binaurale robotisée et sur un robot humanoïde. Expérimentations et évaluations.

[PROJ-R-2] Responsable français du projet STIC-AMSUD (4^e appel) NCTVS (*Nonlinear Control Tools for Visual Servoing*). Partenaires : Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brésil (L.F. ALVES PEREIRA, porteur) ; Pontificia Universidad Católica de Chile (PUC Chile), Santiago, Chili (M.T. TORRES-TORRITI) ; Pontificia Universidad Católica do Rio Grande do Sul, Porto Alegre, Brésil (D.F. COUTINHO) ; LAAS-CNRS. 2010–2011.

Contenu global du projet :

- Analyse et Synthèse multicritères de commandes référencées vision. Évaluation de ces techniques. Propositions pour limiter le conservatisme.
- Traitement de la localisation visuelle par filtrage ensembliste en tant que le problème “dual”. Évaluation. Comparaison avec les méthodes classiques.
- Implémentation sur un robot. Transfert de connaissances.

[PROJ-R-1] Porteur du projet AUDIO_HRP2 (*Capteur Audio Biomimétique et Fonctions Auditives pour le Robot Humanoïde HRP2*), financé par le LAAS-CNRS, et impliquant des partenaires du LAAS-CNRS. 09.2008–08.2011.

Contenu global du projet :

- Développement de fonctions auditives sur la base d'un capteur arc-circulaire pour un robot humanoïde.
- Conception et implémentation d'un capteur audio bio-inspiré reconfigurable.

I.2.5.4 Contribution à des Projets de Recherche

[PROJ-C-3] De Mars 2007 à Juin 2010 : Participation au projet européen STREP COMMROB (Advanced Behaviour and High-level Multimodal Communication With and Among Robots) du FP6, cf. www.commrob.eu, impliquant 6 partenaires et coordonné par TUW, Vienne.

Objectif global du projet : déploiement de chariots de supermarché robotisés, embarquant des facultés avancées de navigation et de communication en interaction avec des humains et d'autres chariots.

Implication :

- Participation à la définition du projet.
- Workpackage *Human Motion Interpretation*, coordonné par le LAAS-CNRS, au titre de la perception auditive.
- Contribution à la rédaction de 6 livrables et aux supports visuels des revues de projet, participation à plusieurs réunions du consortium, etc.

[PROJ-C-2] 2004–2007 : Participation au projet intégré européen COGNIRON (The Cognitive Robot Companion) du FP6, cf. www.cogniron.org, impliquant 8 partenaires et coordonné par le LAAS-CNRS.

Objectif global du projet : conception et intégration d'un robot cognitif doté de facultés de perception, d'interprétation, de raisonnement et d'apprentissage en interaction avec l'homme.

Implication :

- Contribution à l'activité de recherche *Detection and Understanding of Human Activity*, coordonnée par l'Univ. Karlsruhe, au titre de : suivi de personnes et de gestes (2005-2006) ; suivi tridimensionnel (2006-2007) ; audition en robotique (2007).
- Contribution à la rédaction de 2 livrables et aux supports visuels des revues de projet, participation à plusieurs réunions du consortium, etc.
- Participation à la rédaction des supports visuels de l'école d'hiver CWSHRI "Robot Perception in HRI" liée au projet, Lausanne, 2008 (6 contributeurs).

[PROJ-C-1] 2002–2005 : Participation au pré-projet puis au projet *EGOCENTRE : Enchaînement dynamique de tâches référencées multi-capteurs pour l'exécution de mouvements complexes en robotique mobile* du programme de recherche ROBEA (Robotique et Entités Artificielles) du CNRS, impliquant le LAAS-CNRS (porteur P. SOUÈRES), le CerCo Toulouse, l'ENI Tarbes, l'IRISA Rennes.

Thématique : enchaînement de tâches multi-capteurs pour la réalisation de tâches robotiques de haut niveau ; étude des liens et des apports potentiels entre le fonctionnement et la modélisation du système perceptuel animal et humain ; modélisation et réalisation de tâches robotiques à base de capteurs extéroceptifs.

- Participation au titre de : asservissement visuel ; antenne acoustique.
- Contribution à la rédaction de 3 rapports d'avancement.

I.2.5.5 Participation ponctuelle à des Projets de Recherche

La participation aux projets suivants a été / est très ponctuelle (aucune contribution écrite).

- [PROJ-P-3] Depuis Mars 2008 : Participation ponctuelle au projet ANR AMORCES (Algorithmes et MODèles pour un Robot Collaboratif Éloquent et Social), impliquant 5 partenaires et dirigé par le LAAS-CNRS.
→ Workpackage *Cooperative perception/action loops*.
- [PROJ-P-2] 2005 : Participation ponctuelle au projet HR+ du programme de recherche RO-Botique et Entités Artificielles (ROBEA) du CNRS : Suivi de personnes et de gestes pour l'interaction homme-robot.
- [PROJ-P-1] 2003–2004 : Participation au projet CNRS-GRICES 14155 sur la thématique de l'interaction visuelle homme-robot, impliquant le LAAS-CNRS (porteur F. LERASLE) et l'Institute of Systems and Robotics (ISR), Coimbra, Portugal.
→ Participation dans le cadre de [DOC-02].
→ Séjour d'une semaine à l'ISR en 2004.

I.2.5.6 Séminaires significatifs

À la liste des séminaires ci-dessous s'ajoutent des présentations plus confidentielles, dans réunions de projets, ainsi que dans des ateliers du groupe RAP ou du département RIA.

- [SEM-09] Séminaire invité : *Robot Audition Research at LAAS-CNRS*. Université de Kyoto, Japon, 12.05.2009 (en marge de la conférence IEEE ICRA'2009).
- [SEM-08] Séminaire invité : *Rational Systems and Matrix Inequalities to Multicriteria Visual Servoing and Visual-Based Localization*. Pontificia Universidad Javeriana, Bogotá, Colombie, 09.05.2008.
- [SEM-07] Séminaire invité : *Sound Source Localization in Robotics : from Biomimetic to Array Processing Methods*. Pontificia Universidad Javeriana, Bogotá, Colombie, 13.05.2008.
- [SEM-06] Séminaire dans un GDR : *Systèmes Rationnels et Inégalités Matricielles pour l'Asservissement Visuel Multicritères et la Localisation Visuelle*. Journées du GdR Robotique et du GdR ISIS, ENST Paris, 20.06.2007. Version de [SEM-05] adaptée à un public ne connaissant pas les techniques LMI.
- [SEM-05] Séminaire dans un GDR : *Systèmes Rationnels et Inégalités Matricielles pour l'Asservissement Visuel Multicritères et la Localisation Visuelle*. Journées « Commande Robuste en Robotique » du groupe de travail MOSAR en commande robuste du GDR « Modélisation, Analyse et Conduite des Systèmes Dynamiques » (MACS), LIRMM Montpellier, 09.11.2006.
- [SEM-04] Séminaire invité : *Towards an LMI Approach to Multicriteria Visual Servoing in Robotics*. Mobile Robotics Laboratory de l'Institute of Systems and Robotics de l'Université de Coimbra, Portugal, 12.09.2004. Version de [SEM-03] adaptée à un public moins automaticien.
- [SEM-03] Séminaire invité : *Towards an LMI Approach to Multicriteria Visual Servoing in Robotics*. Workshop on Linear Matrix Inequalities in Control (LMI'04), LAAS-CNRS, 01-02.07.2004, sponsorisé par le projet MOCOROVA, le groupe de travail MOSAR en commande robuste, et le GDR

MACS. J'ai présenté les travaux développés avec M^{lle} BELLOT devant une assemblée prestigieuse, de nombreux scientifiques extrêmement renommés spécialistes des méthodes LMI en commande ayant donné une dimension internationale à cet atelier (V.R. BALAKRISHNAN, L. VANDENBERGHE, C.W. SCHERER, J.C. GEROMEL,... cf. <http://www.laas.fr/~arzelier/lmi04> ou <http://www.laas.fr/~henrion/lmi04>).

- [SEM-02] Séminaire de pôle (restucturation du LAAS-CNRS) : *Problématiques pour l'interaction visuelle homme-robot.* Premier atelier du pôle « RObotique et Systèmes Autonomes ». LAAS-CNRS, 2003. Présentation avec F. LERASLE.
- [SEM-01] Séminaire de pôle (restucturation du LAAS-CNRS) : *Problèmes d'estimation en robotique.* Premier atelier du pôle « MOdélisation et COnduite des SYstèmes ». LAAS-CNRS, 2003.

I.2.5.7 Collaborations Scientifiques Suivies

- [COLL-2] Depuis Septembre 2004 (suite à CDC'03) : Collaboration suivie (accueil pour 3 séjours d'une semaine fin 2004, puis en 2005 et 2007) avec M. Daniel F. COUTINHO, enseignant-chercheur en Automatique à PUCRS, Porto Alegre, Brésil.
→ Thématique : Commande et Filtrage pour l'asservissement visuel multicritères et la localisation visuelle.
- [COLL-1] Depuis 2004 : Collaboration avec le Dépt Océanographie du Labo. Collecte Localisation Satellite (Toulouse).
→ Thématique : Modélisation et estimation des mouvements migratoires des tortues-luth dans l'Atlantique Nord. Co-encadrement annuel de stagiaires Master 1 ou Master 2 (e.g. M^{lle} Julie VAN DER HAM [STAGE-07]).

I.2.5.8 Animation Scientifique

- [ANIM-6] Membre du comité local d'organisation de la conférence IFAC Triennal World Congress 2017, Toulouse, France. En charge, avec Gilles ROUX, professeur à l'IUT A de l'UPS, des "popularization events" dont le but est d'ouvrir le Congrès et de rendre l'Automatique accessible au grand public.
- [ANIM-5] Co-organisateur des trois sessions invitées "Robot Audition" à la conférence IEEE/RSJ IROS'2010, Taipei, Taïwan. "Chair" de l'une de ces sessions.
- [ANIM-4] Participation à la création du CD des actes du 6th IFAC Symposium on Intelligent Autonomous Vehicles (IAV'2007), Toulouse, 2007.
- [ANIM-3] "Chair" de la session "Visual Servocontrol" à la conf. IEEE CDC'2003, Maui, HI.
- [ANIM-2] Révision d'articles de journaux (IEEE Trans. on Automatic Control, IEEE Trans. on Robotics, IEEE Trans. on Mechatronics, European Journal of Control, EURASIP, Autonomous Robots, JESA, Speech Communication) et conférences (CDC, ACC, ICRA, IROS, MSC, ROCOND, IAV, RFIA, JDMACS,...)
- [ANIM-1] Évaluateur de deux "Projets LAAS", préalablement à leur financement sur fonds propres par le laboratoire.

I.2.6 Divers

- [DIV-2] Correspondant de l'Association Bernard Grégory (ABG) au LAAS-CNRS, pour l'insertion des Docteurs : aide à l'écriture de CV et envoi à l'ABG pour diffusion.
- [DIV-1] De 2005 à 2007 : Membre nommé du Conseil de Laboratoire du LAAS-CNRS.

I.3 Synthèse des Activités d'Enseignement

I.3.1 Enseignements dispensés à l'Université Paul Sabatier

La mention « définition du contenu » signifie que l'enseignement correspondant n'existe pas ou bien a été entièrement redéfini. La mention « contenu original » signifie que j'ai pris la suite d'un enseignement existant mais en ai proposé une présentation personnalisée. Les cas où mes interventions n'ont pas été accompagnées de la définition ou mise à jour de supports d'enseignement sont également précisés.

- A. Introduction générale aux Systèmes.
 - Cours (contenu original) en DEUG Technologie Industrielle (semestre 1).
- B. Introduction à l'Automatique.
 - TPs (définition du contenu, montage des manipulations en binôme) en DEUG Sciences de la Matière (semestre 4).
- C. Analyse et Commande des systèmes linéaires invariants à temps continu : approche entrée-sortie (« fréquentielle »).
 - Cours (contenu original) en : Licence 3 IUP Systèmes Intelligents ; Licence EEA Formation Continue.
 - Corrections d'annales de l'épreuve “Asservissement” (définition du contenu, montage en binôme) en Agrégation Mécanique.
 - TDs (définition ou participation à la définition des contenus, différents selon les filières) en : Maîtrise EEA ou IUP Systèmes Intelligents ; Maîtrise Technologie Mécanique ; Licence EEA ; MST Physique Habitat ; Licence Professionnelle Conception et Commande des Systèmes Électriques Embarqués.
 - TPs (contenus différents selon les filières ; définition du contenu de certains TPs et montage des manipulations) en : Maîtrise EEA ; Licence EEA ou Ingénierie Électrique ou IUP Systèmes Intelligents ; MST Physique habitat ; Licence Professionnelle Conception et Commande des Systèmes Électriques Embarqués.
- D. Analyse et Commande des systèmes linéaires invariants à temps discret : approche entrée-sortie.
 - Cours/TDs intégrés (définition du contenu) en Licence 3 IUP Architecture et Intégration des Systèmes Électroniques et Microélectroniques.
 - TDs (sans participation à leur définition) en Maîtrise EEA ou IUP Systèmes Intelligents.
 - TPs (participation à la définition du contenu et au montage des manipulations) en Maîtrise EEA ou IUP Systèmes Intelligents.
- E. Analyse et Commande des systèmes linéaires invariants à temps continu ou discret dans l'espace d'état.
 - Cours/TDs intégrés (contenu original) en Agrégation Génie Électrique.
 - TDs (sans participation à leur définition) en Maîtrise EEA.

- Définition de sujets de TP et Montage des manipulations correspondantes pour la Maîtrise EEA (sans encadrement d'étudiants).
- TP (définition du contenu) pour l'Agrégation Génie Électrique.

F. Analyse et Commande des systèmes non linéaires : analyse de systèmes bouclés non linéaires par la méthode du premier harmonique ; analyse et commande dans le plan de phase.

- Cours/TDs (contenu original) en Agrégations Génie Électrique et Physique Appliquée.
- TDs (« premier harmonique ») en Licence EEA.
- TP (définition du contenu de certains TP et montage des manipulations correspondantes) en : Maîtrise EEA ; Agrégations Génie Électrique et Physique Appliquée.

G. Analyse et Commande des systèmes à temps continu ou discret : approches entrée-sortie et techniques d'espace d'état. Introduction aux systèmes multidimensionnels.

- Cours/TDs intégrés (contenu original) en DESS/Master Pro Électronique de Puissance.
- BEs (hors « systèmes multidimensionnels ») en DESS/Master Pro Électronique de Puissance (définition du contenu des sujets et montage des manipulations ; réorganisation avec soutenance orale).

H. Identification paramétrique. Traitement numérique du signal.

- TP (sans participation à leur définition) en Maîtrise EEA format° initiale et continue.

I. Stabilité. Introduction à la théorie de LYAPUNOV et à la commande LQR. Introduction au régulateur RST.

- Cours/TDs intégrés (« stabilité, LQR », définition du contenu) en Master 1 Automatique et Informatique Temps Réel.
- Travaux Dirigés (« RST », définition du contenu en binôme) en Master 1 Automatique et Informatique Temps Réel.
- TP/BEs en Master 1 Automatique et Informatique Temps Réel et en DESS/Master Pro Électronique de Puissance (définition du contenu).

J. Introduction au Filtrage de KALMAN.

- Cours/TDs intégrés, ainsi que TP (définition du contenu, adapté à chaque public) en : Master 1 Information Signal Image Instrumentation ; M2 Pro IUP Ingénierie du Diagnostic, de l'Instrumentation et de la Mesure.

K. Introduction à la navigation d'un robot mobile.

- TP (définition du contenu, montage de la manipulation en binôme) en Licence 1.

L. Modélisation de robots manipulateurs.

- Cours et TDs (contenu original) en Licence 3 IUP Systèmes Intelligents.

M. Introduction à la commande de robots manipulateurs.

- Cours/TDs intégrés (définition du contenu) en : Licence 3 IUP Systèmes Intelligents ; M1 IUP Systèmes Intelligents International (enseignement en anglais).
- BE (définition du contenu) en : Licence 3 IUP Systèmes Intelligents ; M1 IUP Systèmes Intelligents International (enseignement en anglais).

30 \I.3. SYNTHÈSE DES ACTIVITÉS D'ENSEIGNEMENT

O. Travaux d'Étude et de Recherche d'Automatique.

- Maîtrise EEA et Master 1 Automatique et Informatique Temps Réel. Définition de nombreux sujets depuis la mise en place de ces enseignements en 2001, qui visent à développer une étude théorique puis de procéder à son implémentation effective sur un procédé. Encadrement d'une proportion importante des groupes d'étudiants en 2000-2001, 2002-2003, 2003-2004 (arrêt maladie en 2002), puis, grâce à des recrutements de collègues, plus d'une proportion plus raisonnable à partir de 2005.

P. Agrégations blanches, leçons et montages d'Automatique.

- Agrégation Génie Électrique.

Q. Formation au Projet Personnel et Professionnel.

- DEUG 1^{ère} année.

I.3.2 Investissement en Enseignements Pratiques

Je me suis très fortement investi dans la définition, le montage, et la maintenance de Travaux Pratiques. J'ai ainsi mis en place ou contribué à une grande proportion des TPs d'Automatique encore en vigueur aujourd'hui. Il me semble pertinent d'y consacrer une section dédiée, d'autant que les aspects « maintenance » (textes, logiciel, matériel) associés représentent un travail lourd, peu gratifiant, et peu reconnu.

Investissement dans la salle de TPs d'Automatique & Traitement du Signal de l'UPS.

- Maintenance du parc informatique (entre 20 et 25 machines) et du matériel associé de 1995 à 2005 environ, du fait qu'aucun personnel technique qualifié ne nous était affecté.
 - Installation et mise à jour de logiciels et matériels.
 - Prise en charge de la transition logicielle de CC-EXE et MATRIXx vers MATLAB-SIMULINK, et maintenance (1997).
- Responsable de l'organisation et des finances de la salle (cf. plus loin).
- Mise en place de systèmes génériques de commande par calculateur.
 - Encadrement de stagiaires (deux années consécutives) pour la définition d'une architecture logicielle commune à tous nos procédés.
 - Puis, lorsque les finances l'ont permis, acquisition de matériel et logiciel pour le prototypage rapide de commandes par calculateur.
 - Acquisition de cartes dSPACE et de la “toolbox” Real-Time Workshop, et maintenance.
 - Acquisition de cartes d'acquisition (RTI-815 Analog Devices & PCI-6024 National Instruments) et de la “toolbox” Real-Time Windows Target, et maintenance.
- Ré-agencement de la salle de TPs d'Automatique & Traitement du Signal (2003).
- Déploiement d'un réseau intranet, en collaboration avec un Ingénieur Réseau de l'UPS (2003) : installation, définition des modalités de partage SAMBA, etc.

Définition du contenu et montage de nombreuses manipulations de TPs, (avec parfois, sur la base d'un même procédé, des déclinaisons adaptées à divers niveaux de formations).

- Procédé électromécanique : automatique de base.
- Procédé électromécanique : analyse et commande à temps continu dans l'espace d'état.
- Procédé électromécanique : régulation RST.
- Procédé électromécanique : commande par divers types de relais.
- Procédé électromécanique : analyse d'un jeu mécanique.

- Étude simulée d'un asservissement de température.
- Asservissement de tension (fabrication de la maquette).
- Procédé hydraulique à trois bacs : analyse et commande dans l'espace d'état continu.
- Procédé « bille sur rail » : analyse et commande par calculateur dans l'espace d'état.
- Navigation d'un robot mobile Pekee (en binôme).
- Étude simulée de stratégies de commande d'un robot RP.
- Suivi de tortues-luth dans l'Atlantique Nord par filtrage de KALMAN.
- Analyse et Commande d'un procédé « sustentation magnétique ».
- etc.

Proposition de nombreux sujets de TERs (Travaux d'Étude et de Recherche) en Master 1 Automatique Industrielle et Informatique Temps Réel.

- Régulation RST.
- Commande LQR.
- Commande DLQR.
- Filtrage de KALMAN.
- Commande d'un modulateur électro-optique (avec l'IRSAMC-UPS).
- Implémentation d'un prédicteur de Smith pour la commande d'un procédé « soufflerie d'air chaud ».
- etc.

I.3.3 Responsabilité de la définition et de la mise en place d'enseignements théoriques

Participation régulière à la définition des maquettes d'enseignement de la plupart des diplômes dans lesquels j'interviens, à des jurys de recrutement d'étudiants, à des soutenances de stages, etc.

Depuis 2007-2008 : Responsabilité de l'Unité d'Enseignements « Commande et Observation de Systèmes par Calculateur » du Master 2 Pro Ingénierie du Diagnostic, de l'Instrumentation et de la Mesure. Animation de l'équipe pédagogique chargée de la définition du contenu (Commande événementielle, Régulation RST, Filtrage de KALMAN) et contribution à sa mise en place.

Depuis 2003-2004 : Responsabilité de l'Unité d'Enseignements « Automatique » du Master 2 Pro Électronique de Puissance. Définition du contenu des Cours et TPs.

Depuis 2003-2004 : Correspondant pour la définition du contenu et l'organisation des Travaux Pratiques « Navigation d'un Robot Mobile » en Licence 1 (semestre 1, de 350 à 850 étudiants selon les habilitations et les années).

Depuis 2002-2003 : Responsabilité de l'Unité d'Enseignements « Conception Avancée de Systèmes de Commande » en Maîtrise EEA Automatique, qui a évolué dans le cadre du LMD en Master 1 Automatique et Informatique Temps Réel. Il s'agit d'une Unité Optionnelle dont j'ai proposé le principe en 2001. Définition du contenu (stabilité, LYAPUNOV, régulateur LQR, commande RST) et mise en place avec l'équipe pédagogique.

Depuis 2002-2003 : Responsabilité de l'Unité d'Enseignements « Automatique » en 2^{nde} année de l'IUP Architecture et Intégration des Systèmes Électroniques et Microélectroniques, qui a évolué dans le cadre du LMD et de la fusion puis de la séparation avec l'IUP Instrumentation Capteurs Mesures. Une évolution de cette unité est actuellement mise en place dans le Master 1

32 \I.3. SYNTHÈSE DES ACTIVITÉS D'ENSEIGNEMENT

IUP Ingénierie du Diagnostic, de l'Instrumentation et de la Mesure. Définition du contenu et mise en place de l'Unité originelle en binôme. Depuis, animation de l'équipe pédagogique.

De 1999-2000 à 2003-2004 : Responsabilité de l'Unité « Automatique » de la Licence (devenu L3) EEA. De 45 à 90 étudiants selon les années.

2002-2003 : Rédaction de la maquette de l'Unité « Automatique » de la Licence Professionnelle Conception et Commande des Systèmes Électriques Embarqués. Mise en place des enseignements en binôme. Co-organisation des enseignements de l'Unité lors de son ouverture.

De 2000-2001 à 2002-2003 : Co-responsabilité de l'Unité optionnelle « Électrotechnique et Automatique » du DEUG (semestre 4). Définition et mise en place du contenu en binôme.

I.3.4 Responsabilités administratives - Tâches d'intérêt collectif

Depuis 2009-2010 : Membre de l'équipe de direction de l'École Doctorale Systèmes de l'Université de Toulouse.

- Responsable de la proposition, organisation et validation des formations scientifiques relatives à la spécialité « Systèmes Embarqués ».
- Mai 2010 : Mise en place du cours « Estimation et Détection pour les Systèmes Embarqués » dispensé par M. Igor NIKIFOROV, Prof. à l'Université de Troyes.
- Relais pour l'organisation et Responsable de la validation des formations générales relevant du PRES Université de Toulouse ou en lien avec l'Association Bernard Grégory.
- Suivi des Doctoriales, Coordinateur local du *Nouveau Chapitre de la Thèse*® en lien avec l'Association Bernard Grégory, etc.

Depuis 2009-2010 : Référent auprès du Bureau d'Aide à l'Insertion Professionnelle de l'UPS pour certains Master 2 de l'UFR Physique-Chimie-Automatique.

De 2003-2004 à 2007-2008 : Membre élu suppléant de la Commission de Spécialistes de la 61^{ème} section de l'UPS.

De 2000-2001 à 2002-2003 : Membre nommé suppléant (2000-2001) puis titulaire des Commissions de Spécialistes mixtes 61^{ème}/63^{ème} sections et 27^{ème}/61^{ème} sections de l'INSA Toulouse.

De 1999-2000 à 2003-2004 : Membre élu du Conseil de Département EEA de l'UPS. En tant que membre élu du bureau du Conseil, membre du Conseil Pédagogique de l'UFR Physique-Chimie-Automatique de l'UPS (entre 5000 et 6000 étudiants)

- Passage au LMD ; crédits pédagogiques ; projets TICE (correspondant TICE de l'UFR) ; etc.
- Membre de la commission des Travaux Pratiques.
- Membre de la commission « DEUG ».
 - Participation à la définition du parcours « Licence Sciences de l'Ingénieur » entre les départements EEA et Sciences Appliquées de l'UPS.

De 1999-2000 à 2003-2004 : Membre élu titulaire de la Commission de Spécialistes de la 61^{ème} section de l'UPS.

2002 : Président de la Commission d'Évaluation du DEUG Sciences et Technologie Industrielle, mention Sciences de la Matière.

1999-2000 à 2003-2004 : Co-responsabilité puis responsabilité de l'organisation et des finances de la salle de TP Automatique & Traitement du Signal de l'UPS.

- environ 10000 h*étudiants/an ; budget annuel ≈ 20kEUR/an.

De 1997-1998 à 1998-1999 : Membre de la commission de réforme du DEUG Sciences et Technologie Industrielle, mention Sciences de la Matière, mis en place en 1999-2000.

- Membre de la cellule de définition du programme de mathématiques.
- Membre de la cellule de définition de l'Unité Méthodologique Disciplinaire.

De 1997-1998 à 1998-1999 : Membre de la commission “DEUG” du Département EEA de l'UPS.

Participation annuelle au salon de l'enseignement supérieur INFOSUP à destination des lycéens, au titre de diverses formations (EEA, IUPs, etc.).

De 1997-1998 à 2000-2001 : Centralisation des services d'enseignement de la 61^{ème} section de l'UPS (\approx 30 permanents, ATERs et moniteurs à cette époque).

I.4 Annexe : Résumé étendu de la Thèse de Doctorat

Mon travail de thèse s'inscrivait dans le cadre du Raisonnement Qualitatif. Cette discipline, située au confluent de l'Automatique, de l'Intelligence Artificielle et des Mathématiques, a pour objectif de définir des concepts et des techniques permettant à l'Automaticien d'aborder l'étude de systèmes dynamiques pour lesquels les méthodes classiques s'avèrent inadaptées ou inefficaces. C'est le cas lorsque la connaissance dont il dispose est incomplète ou se présente sous une forme non numérique, et/ou lorsqu'il ne recherche qu'une description globale des comportements, en termes de zones de monotonies, points particuliers ou tendances. Le Raisonnement Qualitatif connaissait alors un essor significatif, comme en témoignait la parution d'un numéro spécial de la revue Artificial Intelligence [Bobrow 84]. Outre son apport en phase de conception des systèmes, les potentialités qu'il offrait pour l'automatisation des tâches du niveau de supervision telles que la surveillance, le diagnostic ou la prédiction avaient été clairement identifiées [Dubois 94]. En particulier, de nombreux travaux visaient à développer des schémas de supervision experte dits « de seconde génération », intégrant des connaissances « profondes » sur un processus sous la forme de modèles, au contraire des systèmes experts classiques – ou « de première génération » – dont la représentation interne des connaissances se limitait à une base de règles. À juste titre, le Raisonnement Qualitatif apparaissait incontournable dans cet effort de modélisation et de raisonnement, en permettant la construction et l'exploitation de modèles dont le niveau d'abstraction pouvait être adapté à la granularité requise.

La littérature proposait plusieurs algorithmes de simulation qualitative des systèmes dynamiques. La stratégie la plus aboutie était certainement l'algorithme QSIM (Qualitative Simulation) mis au point par Kuipers [Kuipers 86]. Sur la base de la description d'un système dynamique sous forme d'« équation différentielle qualitative », QSIM développait un arbre de comportements solutions. D'autres approches exploitaient les liens de causalité existant dans le système considéré [De Kleer 84][Forbus 84][Bousson 93]. Des extensions de ces schémas de simulation purement symboliques permettaient la génération de signaux sous la forme de séquences d'intervalles de \mathbb{R} indexés par des « instants qualitatifs ». Des intervalles d'appartenance de la dérivée première ou des dérivées d'ordres supérieurs de ces signaux étaient parfois également établis entre deux instants qualitatifs consécutifs. Selon que l'échelle temporelle était synchrone ou événementielle, les instants qualitatifs étaient soit parfaitement connus, soit associés à un intervalle numérique. La Figure I.1 illustre les propriétés des signaux délivrés par QSIM [Kuipers 86][Berleant 90] et CA-EN (CAusal ENgine) [Bousson 93][Bousson 94].

Malgré leur attractivité *a priori*, ces techniques de simulation conduisaient à des résultats décevants dès lors qu'elles étaient appliquées à un procédé un tant soit peu complexe. Des stratégies telles que QSIM garantissaient que les comportements réels étaient effectivement couverts par ceux simulés sur la base d'une description qualitative. Cependant, la nature locale des algorithmes de simulation qualitative et l'abstraction inhérente aux formalismes associés conduisait généralement à des comportements n'ayant aucune réalité physique (appelés "spurious behaviors" dans [Kuipers 86]). La présence de ces comportements invalides entraînait une explosion combinatoire de l'arbre de simulation, croissante avec la complexité du système étudié.

Ces problèmes étant difficilement solubles, il semblait donc judicieux de ne pas se limiter à des modélisations "tout-qualitatif", mais de faire cohabiter les techniques qualitatives avec les méthodes numériques courantes. Ainsi, l'utilisation conjointe d'information qualitative et quantitative avait été proposée préalablement à mon arrivée dans le langage de modélisation et de simulation ALCMEN (Automaticians Language for Causal Modelisation for Expert Knowledge) [Aguilar-Martin 91]. ALCMEN reposait sur une description par schéma-blocs, où les parties bien connues d'un système étaient traitées par les méthodes classiques de l'Automatique alors que les parties entachées d'imprécision étaient abordées par des techniques qualitatives. La réduction de la taille des sous-systèmes décrits qualitativement offrait la perspective de limiter l'explosion combinatoire de leurs arbres de comportements. En contrepartie, deux problèmes d'interfaçage apparaissaient. Il s'agissait d'une part d'étudier l'interface numérique-symbolique, *i.e.* la « transcription symbolique » d'un signal numérique placé en entrée d'un bloc qualitatif pour permettre sa simulation qualitative. En outre, il était nécessaire de développer une méthode générique d'obtention de la réponse d'un bloc quantitatif, *i.e.* d'un sous-système décrit par une représentation d'état dont les entrées, les conditions initiales, ou les paramètres étaient décrits de manière qualitative. C'est cette dernière tâche, également appelée *interface symbolique-numérique*, qui a fait l'objet de ma thèse (Figure I.2).

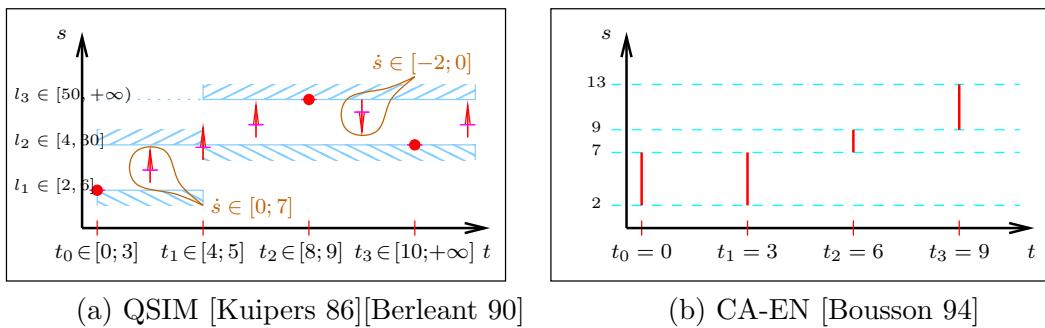


FIGURE I.1 – Représentations qualitatives de signaux. (a) Un signal de type QSIM [Kuipers 86] est décrit selon une échelle temporelle événementielle, les instants qualitatifs étant liés au passage du signal par des "landmarks" prédéfinis l_1, l_2, l_3 . Les landmarks et les instants qualitatifs, par nature symboliques, peuvent être encadrés par des intervalles numériques via [Berleant 90]. Dans l'intervalle temporel séparant deux instants qualitatifs successifs, le signal évolue nécessairement de manière monotone entre deux landmarks, et un affinement de l'intervalle des valeurs des dérivées est parfois disponible. (b) CA-EN [Bousson 93][Bousson 94] produit une suite d'intervalles de valeurs possibles du signal, indexée selon une échelle temporelle synchrone.

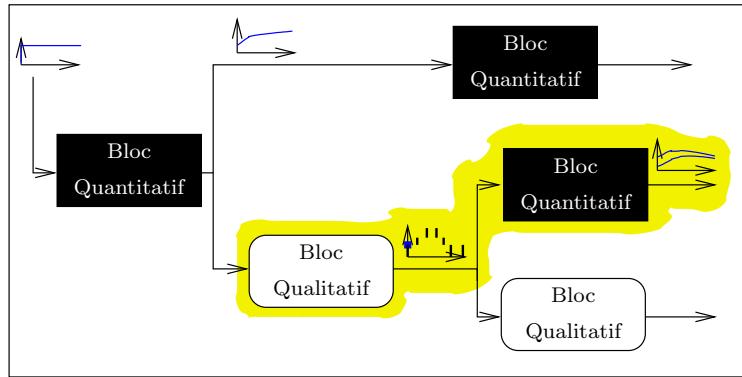
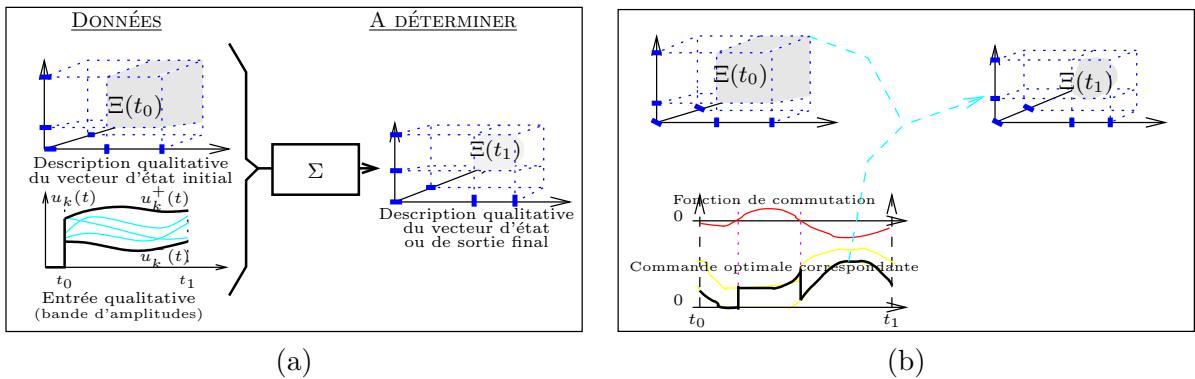


FIGURE I.2 – ALCMEN [Aguilar-Martin 91] et l’interface symbolique-numérique.

FIGURE I.3 – (a) Problème de la réponse d’un système dynamique à une bande d’amplitudes, le vecteur d’état initial étant lui-même défini par un paralléléotope $\Xi(t_0)$, produit cartésien d’intervalles. (b) Solution dans le cas linéaire invariant.

Le premier problème abordé a concerné la détermination d’enveloppes des trajectoires d’état et de sortie d’un système dynamique dont le vecteur d’état initial se situe dans un paralléléotope, et dont chaque entrée évolue dans une « bande d’amplitudes », *i.e.* entre deux signaux possiblement variants dans le temps (Figure I.3-(a)). J’ai reformulé ce problème en un ensemble de problèmes de commande optimale à temps final fixé : à chaque instant, il s’agit de déterminer les excursions minimale et maximale de chaque composante des vecteurs d’état et de sortie. Un ensemble de conditions nécessaires d’optimalité a été établi sur la base du principe du maximum de Pontriaguine (PMP) [Pontriaguine 74][Athans 66], qu’il a préalablement fallu adapter de façon à prendre en compte des critères dépendant explicitement du vecteur d’état initial, des contraintes inégalités sur ce vecteur, ainsi que des domaines de commandes admissibles non invariants[†]. Ceci a conduit à une méthode générique, théoriquement étayée, et permettant un énoncé rigoureux d’hypothèses suffisantes de bien posé sur le système et les enveloppes des bandes d’amplitudes d’entrée (*e.g.* sur la base des approches [Pontriaguine 74] ou [Leitmann 66] ou [Halkin 67]). Les problèmes de commande optimale dégénérés, permet-

[†]. Un problème simplifié, invariant par translation temporelle, a d’abord été étudié. Il s’agit du cas où le système est invariant et les enveloppes de l’entrée qualitative sont des échelons. La solution du cas général, sensiblement plus difficile à obtenir rigoureusement, a été obtenue dans un deuxième temps.

tant d'appréhender un modèle d'état dont le vecteur d'état initial vit dans un paralléotope mais dont le signal de commande est parfaitement défini, ont fait l'objet d'une attention particulière[†]. De même, le cas des systèmes linéaires a été détaillé en dégageant les aspects qui leur sont propres : hormis des cas pathologiques clairement identifiés, un extrémum (minimum ou maximum) d'une composante du vecteur d'état ou de sortie est atteint par application d'une commande bang-bang depuis l'un des sommets du paralléotope de conditions initiales admissibles (Figure I.3-(b)). En revanche, pour des systèmes non linéaires, des commandes optimales peuvent comporter des arcs singuliers, situés à l'intérieur des bandes d'amplitudes admissibles, sans que ce soit le signe d'une propriété pathologique. Enfin, il convient de mentionner que la détermination des enveloppes des trajectoires d'un système soumis à des bandes d'amplitudes a permis de conclure définitivement sur le bien posé de la détermination de la réponse d'un système à un signal défini simplement comme une séquence d'intervalles, tel que sur la Figure I.1-(b), en précisant dans quels cas une condition supplémentaire – enveloppe, limitation de spectre,... – doit être rajoutée sur les entrées.

Les développements ultérieurs ont concerné un système mono-entrée soumis à un signal $u(.)$ issu d'une simulation qualitative de type QSIM (Figure I.1-(a)). Dans un premier temps, seules les excursions minimale u^- et maximale u^+ de $u(.)$ entre deux instants qualitatifs consécutifs t_i et t_{i+1} étaient supposées données, conjointement à des intervalles encerclant les valeurs de t_i et t_{i+1} . La nature locale des conditions nécessaires fournies par le PMP a permis la détermination de l'excursion maximale des états ou des sorties du système sur l'intervalle relatif à t_{i+1} en réponse à la bande d'amplitudes $[u^-; u^+]$, et pour une condition initiale qualitative exprimée sur l'intervalle encadrant t_i . Il suffit en effet de « mettre en concurrence » des problèmes de commande optimale dont l'instant initial et/ou l'instant final est/sont fixe(s) et/ou localement non constraint(s), puis à déduire l'optimum optimorum par comparaison des solutions obtenues.

Ensuite, il s'est agit de déterminer en quoi la connaissance d'intervalles $[v^-; v^+]$ des valeurs admissibles de la dérivée $\dot{u}(.)$ sur $[t_i; t_{i+1}]$ permet de raffiner les évolutions possibles des sorties. Dans cas le plus souvent rencontré en pratique, où v^- et v^+ sont finis et de même signe, le problème a été reformulé en un ensemble de problèmes de commande optimale dans lesquels $\dot{u}(.)$ joue le rôle de la commande. Comme précédemment, le PMP offre un cadre théorique rigoureux quel que soit le système considéré, qui se décline à coup sûr en conditions constructives si celui-ci est linéaire invariant. Lorsque v^- et v^+ sont de signes contraires, les problèmes de commande optimale sont soumis à des contraintes sur les variables de phase. Bien que leur résolution soit considérablement plus complexe, des fondements théoriques sont disponibles dans la littérature [Pontriaguine 74, Chapitre 6][Leitmann 66, Chapitre 4].

Enfin, l'approche a été complétée de façon à traiter théoriquement le cas où certains paramètres du système considéré sont constants et définis par des intervalles.

En conclusion, la transposition du problème de l'interfaçage symbolique numérique dans le contexte de la commande optimale a permis son traitement générique et rigoureux. En effet, les preuves du PMP s'appuient sur des résultats directement issus de la théorie des équations différentielles ordinaires. Pour certains systèmes non linéaires, les conditions nécessaires d'optimalité peuvent s'avérer non constructives, *e.g.* lorsque les équations du vecteur d'état adjoint

[†]. Conceptuellement, ceux-ci se ramènent donc à des problèmes d'optimisation statique sur le vecteur d'état initial. La « subtilité » vient du fait que si le vecteur d'état initial est localement non constraint alors les conditions de transversalité sont des conditions nécessaires à la fois pour la maximisation et la minimisation du critère.

sont insolubles. Néanmoins, tous les problèmes intéressant l’interfaçage avec des signaux de type QSIM ont pu être déclinés. En outre, certains auteurs ont montré qu’il existe un lien entre la preuve du principe du maximum et l’ensemble des points atteignables d’un système augmenté à un instant donné, cette dernière notion étant assez proche de ce que nous cherchions à obtenir.

Mots Clés : Raisonnement Qualitatif, Interface Symbolique-Numérique, Commande Optimale, Principe du Maximum, Signaux Qualitatifs, Signaux de type QSIM, Systèmes à Paramètres Qualitatifs.

Quelques références significatives

- [Aguilar-Martin 91] J. Aguilar-Martin. *Knowledge-Based Real-Time Supervision of Dynamical Processes. Dynamic Knowledge Representation and Simulation*. Rapport technique, G.R. Automatique, LAAS-CNRS, Toulouse, 1991.
- [Athans 66] M. Athans & P.L. Falb. *Optimal control. an introduction to the theory and its applications*. Mc Graw-Hill, 1966.
- [Berleant 90] D. Berleant & B.J. Kuipers. *Combined Qualitative and Numerical Simulation with Q3*. In International Workshop on Qualitative Physics, pages 140–152, Lugano, Switzerland, 1990.
- [Bobrow 84] D.G. Bobrow. *Qualitative Reasoning about Physical Systems : An Introduction*. Artificial Intelligence, vol. 24, pages 1–5, 1984.
- [Bousson 93] K. Bousson & L. Travé-Massuyès. *Fuzzy Causal Simulation in Process Engineering*. In International Joint Conference on Artificial Intelligence (IJCAI'93), pages 1536–1541, Chambéry, France, August 1993.
- [Bousson 94] K. Bousson & L. Travé-Massuyès. *Putting More Numbers in the Qualitative Simulator CA-EN*. In International Conference on Intelligent Systems Engineering, pages 62–69, Harburg, Germany, September 1994.
- [De Kleer 84] J. De Kleer & J.S. Brown. *A Qualitative Physics Based on Confluences*. Artificial Intelligence, vol. 24, pages 7–83, 1984.
- [Dubois 94] D. Dubois & S. Gentil. *Intelligence Artificielle et Automatique*. Revue d’Intelligence Artificielle, vol. 8, no. 1, pages 7–27, 1994.
- [Forbus 84] K.D. Forbus. *Qualitative Process Theory*. Artificial Intelligence, vol. 24, pages 85–168, 1984.
- [Halkin 67] H. Halkin. *Mathematical Foundations of System Optimization*. In G. Leitmann, éditeur, Topics in Optimization, chapitre 6, pages 197–262. Academic Press, 1967.
- [Kuipers 86] B.J. Kuipers. *Qualitative Simulation*. Artificial Intelligence, vol. 29, pages 289–338, 1986.
- [Leitmann 66] G. Leitmann. *An introduction to optimal control*. Mc Graw-Hill, 1966.
- [Pontriaguine 74] L. Pontriaguine, V. Boltianski, R. Gamkrélidzé & E. Michtchenko. *Théorie mathématique des processus optimaux*. Editions Mir, Moscou, 1974.

Deuxième partie

ACTIVITÉS DE RECHERCHE (POST-THÈSE)

Chapitre II

Panorama des Activités de Recherche Post-Thèse

Sommaire

II.1 Résumé étendu des travaux en Détection et Estimation	43
II.1.1 Détection de ruptures entre modèles d'état hétérogènes pour la surveillance de scènes dynamiques	43
II.1.2 Méthodes séquentielles de Monte Carlo et de Quasi Monte Carlo pour le suivi visuel de personnes, de gestes, et la capture de mouvement par vision	44
II.1.2.1 Suivi visuel de personnes et de gestes	45
II.1.2.2 Capture de mouvement par vision	48
II.1.3 Filtrage multimodèles pour la localisation ARGOS	50
II.1.4 Bilan en terme de publications, encadrements et projets	52
II.1.5 Quelques références significatives	52
II.2 Résumé condensé des travaux en Asservissement Visuel	54
II.2.1 Premiers pas...	55
II.2.2 Asservissement visuel multicritère	55
II.2.2.1 Proposition d'un cadre générique : les systèmes non linéaires incertains « rationnels » soumis à des contraintes inégalités rationnelles	55
II.2.2.2 Analyse et Synthèse multicritères par des arguments de stabilité quadratique	56
II.2.2.3 Représentations algébro-différentielles et fonctions de LYAPUNOV plus évoluées pour l'analyse multicritère	57
II.2.3 Dualité entre asservissement visuel et localisation visuelle	58
II.2.4 Conception et implémentation d'un environnement de simulation de commandes en robotique	58
II.2.5 Bilan en terme de publications, encadrements et projets	58
II.2.6 Quelques références significatives	59
II.3 Résumé condensé des travaux en Audition en Robotique	59
II.3.1 Formation de voie pour l'extraction et la localisation de sources sonores large bande	60
II.3.1.1 Formation de voie par analyse modale et optimisation convexe	60
II.3.1.2 Implémentation pratique d'une formation de voie	61
II.3.2 Méthodes à haute résolution pour la localisation et la détection de sources sonores large bande	61
II.3.2.1 “Broadband Beamspace MUSIC”	61

II.3.2.2	Détection de sources	62
II.3.3	Détection de “patterns” audio	62
II.3.4	Le capteur EAR	63
II.3.5	Bilan en terme de publications, encadrements et projets	63
II.3.6	Quelques références significatives	64

Ce chapitre dresse un bilan global des activités de recherche développées consécutivement au Doctorat. Comme cela a été argumenté dans le chapitre précédent, un résumé étendu du volet « Détection et Estimation » est d’abord proposé. Celui-ci est suivi d’une description plus succincte des thématiques « Asservissement Visuel » et « Audition en Robotique », qui sera complétée dans les deux prochains chapitres.

Chaque section se termine par une bibliographie succincte. Celle-ci ne contient pas nos publications, codifiées et recensées dans le §I.2.4, pages 19 et suivantes. De même, les encadrements relatifs à certains travaux sont référencés au §I.2.3, pages 15 et suivantes.

II.1 Résumé étendu des travaux en Détection et Estimation

II.1.1 Détection de ruptures entre modèles d’état hétérogènes pour la surveillance de scènes dynamiques

Un premier problème a concerné la proposition d’algorithmes de trajectographie 2D en position et orientation pour un robot mobile sur la base de la localisation visuelle de marqueurs embarqués par des caméras d’ambiance. La capacité de prédiction de l’occupation de l’espace prend tout son sens lorsque plusieurs robots en évolution autonome partagent l’environnement. Un objectif était de définir une base pour l’estimation de phénomènes plus complexes, *e.g.* trajectographie 3D, exploitation d’attributs visuels plus naturels, contexte multi-capteurs et/ou multi-cibles, gestion d’occultations, etc. Il s’agissait de : définir une bibliothèque des modes de déplacements « canoniques » admissibles ; synthétiser un schéma d’estimation propre à chaque mode ; en ligne, détecter les transitions entre divers modes et estimer le déplacement du robot.

Notre modélisation d’état des divers déplacements canoniques différençiait la trajectoire géométrique du robot du mouvement sur celle-ci, et permettait une formulation assez élégante des problèmes de rupture dans la dynamique du mouvement [SIACL-1-DE-1][STAGE-02]. Afin de faciliter la convergence du schéma d’estimation dans un contexte fortement non linéaire, celui-ci était basé sur un ensemble de filtres de KALMAN étendus différant par leurs initialisations, et dont les durées de vie étaient liées à leur vraisemblance vis à vis des innovations. Ainsi, après quelques observations, seuls étaient actifs le ou les filtres capables de suivre convenablement le mouvement effectif du robot. Dans la veine de [Bensalah 96], nous avons mis en place un détecteur de ruptures entre déplacements d’un même mode sur la base de l’implémentation de WILLSKY et JONES du *Test du Rapport de Vraisemblances Généralisé* (GLRT = Generalized Likelihood Ratio Test) [Willsky 76][M2R-01]. Néanmoins, l’approche était limitée par le coût de l’implémentation d’un ensemble de filtres pour chaque mode de déplacement, et par l’impossibilité de détecter des transitions entre des modes distincts. Une difficulté particulière réside dans le fait que les vecteurs d’état des modèles de déplacement admettent *des dimensions et des sémantiques hétérogènes*.

Un volet de la thèse de doctorat de M. Thierry SENTENAC, Maître-Assistant à l’École des Mines d’Albi-Carmaux, concernait la détection de feux dans des soutes d’avion et l’estimation de leur stade d’avancement à partir d’informations délivrées par des caméras CCD [DOC-0a]. Les modèles de déplacements s’étendant aisément à la description de tels phénomènes, M. SENTENAC et moi avons poursuivi conjointement les travaux précédents. L’estimateur relatif à chaque type de dynamique/mouvement a été simplifié, et remplacé par un seul

filtre de KALMAN étendu dont l'initialisation est réalisée au moyen d'une forme « information », plus légère à mettre en œuvre. L'algorithme de détection de ruptures basé sur le GLRT a été supprimé. Plutôt que de tester l'hypothèse de référence – pas de saut – contre l'hypothèse alternative composite – existence d'une rupture – pour l'instant et la nature du saut les plus « favorables » pour que cette dernière l'emporte, une autre option consiste à comparer la vraisemblance de l'hypothèse de référence avec la vraisemblance maximale d'occurrence d'une rupture dans le passé récent en marginalisant, au sens du calcul des probabilités, par rapport à toutes les paramétrisations possibles du saut. L'exploitation du *Test du Rapport de Vraisemblances Marginalisé* (MLRT = Marginalized Likelihood Ratio Test) proposée par GUSTAFSSON [Gustafsson 96], constitue une solution à ce problème dans le cas d'un système linéaire avec saut sur le vecteur d'état. L'extension que nous avons suggérée pour la détection multi-hypothèses sur la base de modèles d'état hétérogènes effectue une marginalisation par rapport à toutes les hypothèses alternatives de dynamique/mouvement et leurs paramétrisations [M2R-02]. Dans la veine de l'implémentation à 2 filtres de GUSTAFSSON pour son problème original [Gustafsson 91], lorsque M déplacements canoniques mutuellement exclusifs de paramétrisations équiprobables sont en concurrence, nous avons synthétisé un algorithme de détection de ruptures à $M + 1$ filtres : pour chaque instant de saut candidat dans la fenêtre de détection, les termes constitutifs du rapport de vraisemblances marginalisé sont obtenus au moyen d'un filtre avant basé sur l'hypothèse de référence et de M filtres arrières basés sur la forme rétrograde des modèles de déplacement et ne faisant l'hypothèse d'aucune connaissance *a priori*[†]. En raison de sa relative complexité, ce schéma de détection n'est pas exécuté systématiquement, mais seulement si un mécanisme de surveillance de l'innovation – *residual monitoring* –, d'implémentation très légère, révèle que l'innovation réelle à tendance à ne plus vérifier ses propriétés théoriques dictées par le type de dynamique/mouvement en cours de validité.

Ces développements théoriques et leur validation expérimentale constituent le Chapitre 4 de la thèse de M. SENTENAC [DOC-0a], mais n'ont pas été publiés en raison de la clause de confidentialité avec l'industriel co-financeur. La bibliothèque des mouvements canoniques a été complétée dans [M2R-05] de façon à couvrir l'ensemble des déplacements sans rebroussement d'un robot de type voiture dont chacune des deux roues motrices admet une évolution indépendante à vitesse ou accélération constante [Fleury 95].

II.1.2 Méthodes séquentielles de Monte Carlo et de Quasi Monte Carlo pour le suivi visuel de personnes, de gestes, et la capture de mouvement par vision

La nécessité de telles fonctionnalités visuelles est particulièrement évidente dans le cadre de la problématique de l'Interaction Homme-Robot, centrale à l'activité Robotique du LAAS-CNRS. Le fait de permettre à un robot d'évoluer de manière autonome en présence de public exige en effet que celui-ci soit capable d'évaluer l'activité humaine dans son voisinage, et d'interagir avec d'éventuels tuteurs par la reconnaissance de gestes élémentaires ou de postures pour l'exécution de tâches décidées par eux.

Les capteurs visuels, par leur performances, leur faible coût et la richesse des informations qu'ils délivrent, sont incontournables pour répondre à ces défis. C'est ainsi que la thématique de la Vision par Ordinateur pour l'Interaction Homme-Robot a été initié au LAAS-CNRS

[†]. Ces filtres arrière sont donc implémentés d'abord sous forme information puis sous forme classique.

par M. Frédéric LERASLE, collègue spécialiste du domaine. Suite à une collaboration ponctuelle [CIACL-03-DE-1], nous avons co-encadré plusieurs travaux relatifs au suivi visuel de personnes, au suivi et à la segmentation de gestes élémentaires, ainsi qu'à la capture de mouvement par vision [Gavrila 99][Wu 99][Moeslund 01], sur la base de séquences vidéo typiques des environnements variables et évolutifs de la Robotique. Notre objectif était de proposer une conjonction de techniques de Vision par Ordinateur et d'Estimation permettant de développer des systèmes de suivi dotés d'une relative robustesse aux changements d'apparence du sujet, aux occultations consécutives à son déplacement, ainsi qu'aux conditions de prise de vue : encombrement de la scène, variations d'éclairage, caméra statique ou non, présence d'autres objets mobiles, etc. La limitation de la complexité algorithmique et le respect de contraintes temporelles fortes constituent des contraintes supplémentaires imposées par le contexte de la Robotique.

II.1.2.1 Suivi visuel de personnes et de gestes

Notre premier encadrement commun [M2R-04] a concerné la détection hors ligne de gestes dans une séquence d'images monoculaires monochromes au moyen de la superposition sur celles-ci d'une forme prototype 2D déformable – ou “template” – correspondant à une main. Les techniques sous-jacentes ont trait à la segmentation par le mouvement, aux tests d'hypothèses statistiques, et à l'estimation paramétrique par recuit simulé.

Suite à ce travail préliminaire, nous avons abordé les problèmes de suivi visuel d'objets gauches et déformables dans le cadre de l'estimation Bayésienne de systèmes dynamiques Markoviens [Chen 03]. À tout instant k , le vecteur d'état x_k capture la paramétrisation de l'objet à suivre : grandeurs continues caractéristiques de sa projection et/ou de sa situation et configuration tridimensionnelles, paramètres discrets exprimant des configurations-types ou modes de déplacement. La dynamique *a priori* $p(x_k|x_{k-1})$ peut donc soit être continue, soit inclure des sauts Markoviens. Le vecteur de mesure z_k symbolise l'ensemble des données visuelles disponibles à chaque instant. Contrairement à de nombreux problèmes classiques de filtrage, la densité d'observation $p(z_k|x_k)$ n'est généralement pas définie analytiquement, et son échantillonnage pour une réalisation donnée x_k du vecteur d'état n'est pas faisable. Néanmoins, pour une observation z_k donnée, il demeure possible de caractériser les vraisemblances de différentes hypothèses sur l'état, *i.e.* d'évaluer $p(z_k|\xi)$ pour divers ξ .

Le *filtrage particulaire* – ou *méthodes séquentielles de Monte Carlo* – [Arulampalam 02] [Doucet 00][Doucet 01] est un moteur particulièrement indiqué pour le suivi. Il permet l'estimation récursive de la loi *a posteriori* de tout système Markovien, linéaire ou non, soumis à des excitations aléatoires de nature quelconque. Il peut fonctionner en temps réel sur une CPU classique pour des espaces d'état de dimension raisonnable. Enfin, il constitue un cadre générique et théoriquement étayé pour l'intégration d'attributs visuels variés.

Il convient ici d'en rappeler les fondements afin de mieux cerner nos contributions. La loi *a posteriori* $p(x_k|z_{1:k})$ est récursivement approximée par une distribution ponctuelle $\hat{p}(x_k|z_{1:k})$ exprimant la sélection de *particules* – valeurs admissibles du vecteur d'état – selon les probabilités – ou *poids* – associées. Les particules sont échantillonées aléatoirement selon une *fonction d'importance*, puis sont pondérées de façon à garantir la cohérence de $\hat{p}(x_k|z_{1:k})$. Afin de limiter le *phénomène de dégénérescence*, selon lequel quelques particules concentrent l'essentiel des poids après quelques itérations, un *rééchantillonnage* est parfois introduit, qui consiste à sélectionner avec remise un nouveau jeu de particules dans le nuage pondéré courant. Les particules de poids élevés sont alors dupliquées au détriment de celles, faiblement

pondérées, qui disparaissent, tout en conservant une approximation ponctuelle cohérente. Ces rééchantillonnages ne doivent toutefois pas être invoqués systématiquement sous peine d'*appauvrir* le nuage, *i.e.* de perdre en diversité dans l'exploration de l'espace d'état consécutivement à une duplication trop importante d'un nombre réduit de particules.

Dès lors qu'une procédure de rééchantillonnage est sélectionnée, tout filtre particulaire se dérive de cet algorithme générique (appelé SIR = Sampling Importance Resampling) par le choix d'une fonction d'importance, choix qui conditionne son efficacité et la qualité de l'estimation [Arulampalam 02][Doucet 00]. On peut distinguer les stratégies « FID »[†], qui admettent pour fonction d'importance la dynamique *a priori*. Les particules ainsi échantillonées sont ensuite pondérées par leur vraisemblance déduite de la densité d'observation, ce qui confère à ces schémas une structure prédiction / mise à jour comparable à celle du filtre de KALMAN. La CONDENSATION (CONDitional DENSiTy propagATION) [Isard 98a] s'inscrit dans cette classe. Bien que son implémentation soit très simple, elle présente souvent des problèmes significatifs d'efficacité et de précision, du fait de la possible non-couverture de zones de l'espace d'état fortement vraisemblables vis à vis de la mesure, voire de la dégénérescence du nuage dans le cas de dynamiques peu informatives ou de densités d'observation très étroites. À l'inverse, les stratégies « FIM »[†] basent leur fonction d'importance sur une information issue de la mesure, cf. par exemple l'ICONDENSATION [Isard 98b]. Elles surpassent la classe des FID dès lors que la fonction de mesure est plus informative que la dynamique *a priori* ou incohérente avec elle. Cependant, elles présentent le risque de disperser inutilement le nuage en cas de fausses mesures. En outre, du fait qu'une particule échantillonnée sur la base de la mesure est affectée d'un historique arbitraire, rien n'empêche qu'elle soit incompatible avec sa particule prédecesseur, à l'instant précédent, du point de vue de la dynamique. Elle est alors pourvue d'un faible poids, y compris si elle se situe dans un mode de la loi *a posteriori*. Pour pallier ces difficultés, il est judicieux que la fonction d'importance combine la dynamique *a priori* et l'information contenue dans la mesure, ce qui donne lieu à des stratégies « FIDM »[†]. Des améliorations supplémentaires peuvent être obtenues en exploitant un rééchantillonnage à des niveaux divers du filtre : stratégie partitionnée [MacCormick 00a] [MacCormick 00b] ou hiérarchisée [Pérez 04], rééchantillonnage de l'historique [Torma 03], etc. Par ailleurs, certains filtres particulaires écrivent l'approximation $\hat{p}(x_k|z_{1:k})$ comme une mixture de Gaussiennes, *e.g.* [Kotecha 03a, Kotecha 03b][van der Merwe 03].

Il existe une stratégie FDIM dite « optimale », ou « adaptée », au sens où elle permet de limiter au mieux la dégénérescence de l'estimateur. Du fait que sa fonction d'importance et son équation de pondération des particules sont souvent difficiles à mettre en œuvre en pratique, on recourt à des approximations. L'une est le filtre « à particules auxiliaires » [Pitt 99], qui doit toutefois être écarté si la dynamique est très bruitée ou lorsque la mesure est très informative. Le filtre particulaire “unscented” [Van Der Merwe 01a] associe à chaque particule une fonction d'importance Gaussienne obtenue par application d'un pas du filtrage de KALMAN “unscented” [Julier 97]. Une autre stratégie qui imite le cas optimal, quoique relativement coûteuse et difficile à implémenter, est le filtre “unscented” avec poids auxiliaires [Andrieu 01].

Dans un premier temps, la thèse de M. Ludovic BRÈTHES [DOC-02] a concerné le suivi temps réel de personnes et de gestes dans des séquences monoculaires couleur. Pour le suivi de per-

[†]. Les sigles FID, FIM, et FIDM, exprimant le fait que la fonction d'importance est définie sur la base de la dynamique *a priori*, de la mesure, ou de leur combinaison, ont été introduits dans [DOC-02]. Ils sont utilisés par convenance mais ne correspondent pas à des dénominations standardisées dans la littérature.

sonnes, notre objectif a été de définir des associations de schémas de filtrage et d'attributs visuels adaptées à divers scenarii d'interaction visuelle homme-robot. Dans le but de limiter la complexité et respecter les contraintes temporelles tout en permettant une éventuelle interprétation des résultats, nous nous sommes limités à une analyse spatio-temporelle dans le plan image, ou « suivi 2D ». Dans le prolongement de [Pérez 04], nous avons proposé un éventail d'attributs pertinents en Robotique – contours, couleur, mouvement, visages, blobs peau, etc. – et les avons partitionnés en tant qu'intermittents ou persistants. Les premiers proviennent généralement de modules de détection et sont par nature fortement discriminants. Les seconds sont disponibles de manière systématique mais peuvent apporter une information ambiguë. Afin d'assurer l'initialisation spontanée du suivi ainsi que sa réinitialisation automatique consécutivement à une perte de la cible, il est opportun de définir la fonction d'importance comme une mixture intégrant la dynamique *a priori* de l'objet suivi, une éventuelle distribution uniforme sur l'image, et une loi issue de la détection d'attributs intermittents. Accessoirement, des attributs persistants peuvent aussi être exploités dans cette étape d'échantillonnage des particules. En outre, nous avons montré que les dérives dans le suivi peuvent être limitées en combinant/fusionnant des attributs persistants qui se complètent au sein d'une fonction de mesure unique impliquée dans la pondération des particules. L'adaptation en ligne des caractéristiques de certains attributs – *e.g.* la mise à jour de l'histogramme couleur de référence du sujet suivi, lequel est nécessairement déduit du résultat du filtrage – s'en trouve fiabilisée.

Un ensemble cohérent de fonctions d'importance et de fonctions de vraisemblance multi-attributs a ainsi été défini, puis évalué en terme de pouvoir discriminant, précision et temps de calcul. Alors que la littérature ne proposait des associations qu'avec un nombre restreint de stratégies d'estimation, nous avons mené une étude approfondie du filtrage particulaire et considéré un nombre conséquent de ses déclinaisons mono-cible [CNACL-3-DE-1][JSI-5-DE-1] : filtre générique SIR et ses instanciations CONDENSATION [Isard 98a] et ICONDENSATION [Isard 98b], filtre à échantillonnage hiérarchisé [Pérez 04], à particules auxiliaires [Pitt 99], à rééchantillonnage de l'historique – dont nous avons exhibé des connexions originales avec le filtre à particules auxiliaires – et sa version Rao-Blackwellisée (RBHSSIR = Rao-Blackwellized Subspace History Sampling SIR) [Torma 03]. Une évaluation de nombreux couples (attributs,filtres) en terme de précision, taux d'échec et temps de traitement a permis la définition de trois systèmes de suivi : un suivi proximal – robot à l'arrêt, à une distance d'interaction inférieure à 1 m – basé sur la CONDENSATION, qui échantillonne les particules selon une dynamique *a priori* de type « marche aléatoire » et les pondère au moyen de fonctions de vraisemblance combinant des attributs de forme et de mouvement ; un système de suivi à mi-distance – robot en mouvement, situé entre 1 m et 3 m d'un humain, *e.g.* lors d'une tâche de guidage – combinant couleur et forme dans un RBHSSIR ; un système de surveillance à longue distance – robot à l'arrêt, à une distance supérieure à 3 m – combinant couleur et mouvement dans un RBHS-SIR [CIACL-13-DE-4].

Le problème du suivi et de la segmentation de gestes s'inscrivent naturellement dans le contexte du *filtrage particulaire pour des systèmes à sauts Markoviens*. Les modes sujets à sauts expriment les différentes configurations que peut prendre l'apparence de la main dans le cas de gestes statiques, ainsi que les classes de déplacements admissibles dans le cas de gestes dynamiques. Nous avons procédé au suivi 2D de gestes fronto-parallèles symboliques. Parallèlement à l'élaboration de fonctions visuelles de détection et de mesure dédiées, nous avons considéré deux schémas d'estimation : la Mixed-State CONDENSATION [Isard 98a],

qui échantillonne le vecteur d'état hybride discret/continu selon sa dynamique *a priori*, ainsi que le filtre particulaire unscented avec poids auxiliaires pour les systèmes à sauts Markoviens [Andrieu 03], qui exploite astucieusement l'information image dès l'échantillonnage des composantes discrètes (modes) du vecteur d'état.

II.1.2.2 Capture de mouvement par vision

M. LERASLE et moi-même avons poursuivi notre collaboration scientifique par le co-encadrement du stage de DEA [M2R-06] puis de la thèse [DOC-04] de M. Mathias FONTMARTY. Le but était de mettre en œuvre un suivi 3D de postures sur la base d'images issues de capteurs stéréoscopiques ou multi-oculaires, par recalage en situation et en configuration d'un modèle tridimensionnel fruste – ou “template” – établi à partir de données biomécaniques et anthropomorphiques, *e.g.* un ensemble de cônes tronqués reliés par des liaisons rotoides. Le recalage peut être guidé par l'apparence du modèle dans les plans images des capteurs mais aussi, lorsque cela est possible, par une « distance 3D » entre le template et un nuage de points obtenus par reconstruction 3D locale de la scène.

Un problème simplifié a d'abord été considéré dans [M2R-06], consistant en le suivi tridimensionnel de mouvements élémentaires d'un bras humain à partir de son apparence dans les images délivrées par un capteur stéréoscopique embarqué sur le robot. Sur le plan de la vision, le recalage exploitait des zones circulaires extraites de l'image ainsi que des contours et leurs normales. Du point de vue estimation, nous avons implémenté une forme « racine carrée » [Van Der Merwe 01b] du filtre de KALMAN “unscented”, plus efficace et d'une meilleure stabilité numérique que l'algorithme standard [Julier 97].

La thèse [DOC-04] a concerné la capture du mouvement du corps tout entier ou de l'ensemble {buste,bras} dans deux contextes robotiques distincts. Dans le cadre de l'interaction homme-robot, les informations visuelles sont supposées provenir d'une tête stéréoscopique embarquée sur une plateforme mobile. L'environnement est non contrôlé, variable et dynamique. La position initiale et l'apparence du sujet ne sont pas bien définies. Les contraintes temporelles sont très fortes car l'algorithme est déroulé en temps réel par une CPU embarquée sur le robot, laquelle est également en charge de processus concurrents. Le second contexte est celui de la robotique ubiquiste, où des réseaux de capteurs sont intégrés à l'environnement. Un ensemble de trois caméras fixes déportées, calibrées et admettant des champs de vue complémentaires, surveille un environnement contrôlé. L'éclairage est constant et l'arrière-plan connu. L'évolution du sujet est limitée au champ de vue du système multi-oculaire. La capture de ses mouvements doit également s'effectuer en temps réel, mais un supplément de complexité par rapport au cas embarqué est acceptable. En revanche, la configuration des caméras rend plus difficile toute reconstruction tridimensionnelle de la scène, même partielle.

Dans les deux cas, l'extrême difficulté du problème exige de se limiter à une seule cible. Sur le plan des techniques d'estimation, la littérature montre que l'approximation Gaussienne sous-jacente aux extensions du filtrage de KALMAN est mise en échec en présence de butées ou de configurations singulières [Deutscher 99]. Malgré le succès des approches particulières dans de telles conditions, le problème n'en demeure pas pour autant résolu. En premier lieu, la dimension de l'espace d'état est très élevée, *e.g.* égale au nombre de degrés de liberté – 14 ou 22 pour nos modèles volumiques correspondant à l'ensemble {buste,bras} ou au corps tout entier – pour une dynamique *a priori* de type marche aléatoire. Bien que des études théoriques soignées des méthodes séquentielles de Monte Carlo [Crisan 02] aient affirmé que celles-ci admettent une vitesse de convergence indépendante de la dimension de

l'espace d'état – et par conséquent ne sont pas victimes de la « malédiction de la dimensionnalité » (“beat the curse of dimensionality”) – DAUM et HUANG [Daum 03a] montrent qu'en général, leur complexité varie exponentiellement avec l'ordre du système pour une erreur normalisée donnée[†]. Une difficulté supplémentaire tient à l'existence de multiples modes très fortement marqués dans la fonction de vraisemblance, à la perte d'information par projection, et à l'impossibilité de délimiter directement les zones fortement vraisemblables de l'espace d'état. Pour l'ensemble de ces raisons, trois directions principales ont été suivies dans la littérature pour la synthèse de schémas d'estimation dédiés à la capture de mouvement par vision. Il s'agit d'une part d'explorer en priorité les directions de l'espace d'état difficiles à estimer, *e.g.* en intégrant l'estimé de la covariance *a posteriori* calculé par le filtre à l'instant précédent (“covariance-based sampling”) dans la définition de sa fonction d'importance [Sminchisescu 02]. D'autre part, l'échantillonnage hiérarchisé des variables relatives au tronc préalablement aux variables de configuration des extrémités, sous-jacent aux filtres particulaires partitionnés [MacCormick 00b], réduit la complexité à une fonction linéaire du nombre de partitions. Enfin, les filtres à recuit simulé [Deutscher 00] exploitent une séquence de fonctions d'importance et de fonctions de vraisemblance de plus en plus focalisées, de façon à focaliser graduellement les particules vers les modes de la loi *a posteriori*. Compte tenu de ce constat, notre contribution s'est déclinée en quatre points principaux.

- Un ensemble de traitements visuels a été défini, sur lesquel repose le suivi. Celui-ci inclut l'extraction d'indices visuels de forme (silhouette, silhouette duale, contours), d'indices visuels de couleur (patches, blobs peau, zones de couleur uniforme), ainsi que la reconstruction éparse de blobs 3D relatifs à la tête et aux mains à partir de l'extraction de blobs peau dans les images. La richesse offerte par les primitives d'apparence (attributs 2D) et le caractère discriminant des primitives tridimensionnelles ont permis la définition d'un ensemble de fonctions relativement complet pour la caractérisation de la vraisemblance de tout couple (situation, configuration) du “template” 3D associé au corps tout entier ou à sa partie supérieure.
- Nous avons proposé plusieurs schémas d'estimation originaux pour la capture de mouvement par vision. Parmi ceux-ci, l'un consiste en une extension du filtre à recuit simulé permettant l'exploitation de la reconstruction tridimensionnelle éparse mentionnée ci-dessus au plus tôt, dès l'échantillonnage du premier niveau de particules [CIACL-14-DE-5]. Un effort important a ensuite concerné les *méthodes séquentielles de Quasi Monte Carlo* (QMC), préconisées dans [Daum 03b], qui substituent à l'échantillonnage (pseudo-)aléatoire conventionnel des séquences randomisées à faible discrépance. Les phénomènes de « trous et amas » (“gaps and clusters”), inhérents à l'échantillonnage aléatoire et omniprésents dans les espaces de hautes dimensions, sont alors éliminés. La convergence théorique de l'erreur par rapport au nombre de particules s'en trouve améliorée [Fang 94]. En outre, les résultats d'estimation admettent une variation de Monte Carlo plus faible, qui limite le taux d'échec du suivi. Ces méthodes

[†]. Selon ces auteurs, la complexité est indépendante de l'espace de recherche pour des problèmes « vaguement Gaussiens », sous réserve que les filtres soient « soigneusement implémentés ». L'argument intuitif est qu'une distribution Gaussienne requiert l'exploration d'une boule multi-dimensionnelle, alors qu'une distribution quelconque nécessite d'échantillonner un volume qui croît à la manière d'un hypercube. Pour des espaces d'état de dimension d faible, ces deux problèmes admettent une complexité comparable. La décroissance drastique du volume relatif de la boule unité de dimension d par rapport à l'hypercube $[-1; +1]^d$ au-delà de $d = 5$ est à l'origine de la différence de complexité dans des espaces de grande dimension. Les auteurs étayent cette intuition par la simulation de contre-exemples judicieux.

demeurent toutefois assez confidentielles. Leur application au suivi 2D est restreinte à [Philomin 00] (séquences QMC déterministes) et [Ormoneit 01] (séquences QMC randomisées). Bien que des bénéfices conséquents aient été reportés, aucune autre référence n'en a fait état dans la communauté Vision par Ordinateur, y compris à des fins de suivi tridimensionnel. Notre contribution principale sur ce volet du problème a été d'exploiter la version QMC de la CONDENSATION issue des travaux de GUO et WANG [Guo 06], et d'en proposer une version partitionnée [CIACL-19-DE-7].

- Un ensemble de métriques pour l'évaluation systématique de diverses stratégies de suivi tridimensionnel a ensuite été défini. Celui-ci intègre le taux d'échec du suivi, les racines des erreurs quadratiques moyennes ($RMSE = "Root Mean Square Error"$) entre les positions 3D réelles et estimées des articulations, les biais moyens sur ces positions, ainsi que l'homogénéité des estimations. Bien que souvent absent dans la littérature, ce dernier critère est tout à fait important car il permet de caractériser la dispersion des estimés – pour une même séquence stéréoscopique ou multi-oculaire en entrée – induite par la nature stochastique du filtre.
- Enfin, des règles de sélection des paramètres libres dans les définitions de fonctions de vraisemblances ont été établies [CIACL-20-DE-8]. Il s'agit d'un problème récurrent à toute stratégie de suivi visuel, paradoxalement très peu traité dans la littérature.

Une évaluation d'un ensemble relativement exhaustif de combinaisons d'algorithmes de filtrage et de primitives visuelles, relativement aux métriques ci-dessus, a été effectuée, prolongeant ainsi les travaux de [Balan 05]. La base d'images et la vérité terrain étaient d'abord simulées sous le logiciel MAVS (§II.2.4), puis obtenues en conditions réelles dans un environnement équipé d'un système de capture de mouvement commercial. Dans le contexte de la robotique ubiquiste –système trinoculaire–, notre algorithme de filtrage QMC partitionné constitue un moteur d'estimation intéressant pour le suivi tridimensionnel. Les performances sont supérieures du point de vue de la quasi-totalité des critères aux autres stratégies [JSI-6-DE-2], y compris vis à vis du filtre à recuit simulé, pour une fréquence d'exécution avoisinant 1 Hz, 50 % supérieure à ce dernier. La RMSE se situe autour de 10 cm en dépit de la grossièreté du template 3D, pour des taux d'échec, biais et dispersion somme toute très acceptables. L'exploitation de séquences randomisées à faible discrépance apporte une diminution significative du nombre de particules requises pour un niveau de performances donné. Le contexte stéréoscopique permet un fonctionnement à 8 Hz mais conduit à une RMSE plus élevée du fait de la moindre information sur la profondeur de la scène. On retrouve les mêmes tendances que précédemment pour des séquences simples ou pour une étude des performances selon les axes parallèles aux plans image du système stéréoscopique. Cependant, si la fonction de mesure renseigne peu sur la profondeur, alors le problème est mal posé. Des stratégies basiques telles que la CONDENSATION peuvent l'emporter en termes de précision ou de dispersion, confirmant certaines constatations de [Balan 05].

II.1.3 Filtrage multimodèles pour la localisation ARGOS

Le système ARGOS, inspiré en 1978 par le CNES, la NASA et la NOAA, permet la localisation par satellite de balises embarquées sur de nombreux mobiles : navires, animaux terrestres ou marins, oiseaux, bouées dérivantes, voiliers, etc. Le positionnement est basé sur la mesure de l'effet Doppler affectant les signaux émis par chaque balise [CLS 08]. Depuis 1986, c'est l'entreprise Collecte Localisation Satellites (CLS), filiale du CNES et de l'Ifremer, qui est

l'unique exploitant du système de géolocalisation. Son siège social est situé à Ramonville, à quelques encablures du LAAS-CNRS.

Une collaboration informelle a été mise en place voici plusieurs années avec M. Philippe GASPAR, alors Directeur du Département d'Océanographie Spatiale de CLS. Celle-ci a d'abord pris la forme d'encadrements d'étudiants de Master 1 puis de Master 2 [STAGE-07] sur le thème de la trajectographie d'animaux marins [Royer 08]. Il s'agissait de reconstruire des trajectoires de tortues-luth à partir de mesures ARGOS en prenant en compte les caractéristiques suivantes du problème : ces animaux admettent plusieurs modes de déplacements (migrations au long cours, mouvements erratiques dans des zones de nourriture, etc.) ; une classe de précision est associée à toute mesure ; le signal émis par la balise peut ne pas être perçu pendant plusieurs jours (*e.g.* si l'animal est en plongée profonde lors du passage d'un satellite) ; il existe des localisations erronées (dites « localisations images » dans le jargon ARGOS) ; etc.

Plus récemment, de février à juin 2009, M. Rémy LOPEZ a effectué son stage de fin d'études à CLS, sous la responsabilité scientifique de MM Jean-Pierre MALARDÉ, du groupe Études Systèmes et Ingénierie Radiofréquence, et François ROYER, de la Direction Océanique et Spatiale. L'objectif était de poser les bases d'une refonte de l'algorithme de localisation originel. Celui-ci estimait le triplet (position,vitesse,fréquence d'émission) de toute balise et lui associait une classe de localisation au moyen de techniques de moindres carrés non linéaires couplées à des calculs de vraisemblances exprimant la continuité de la fréquence de la balise, la plausibilité de sa vitesse, etc. La solution proposée durant le stage de M. LOPEZ [Lopez 08] est basée sur la forme « racine carrée » [Van Der Merwe 01b] du filtre de KALMAN “unscented” [Julier 97], couplée avec des modules qui évitent la divergence passagère du filtre et garantissent l'intégrité du processus de localisation : calculs de vraisemblances et tests, résolution d'ambiguïtés, récupération d'états erronés, gestion particulière de macropassages, etc. Du fait qu'elle conduit à des performances significativement supérieures à l'algorithme original [Lopez 10], elle le remplacera prochainement, cf. le Flash #19 sur www.argos-system.org.

Dans ce contexte, depuis début 2010 je co-encadre avec les trois collègues de CLS ci-dessus mentionnés la thèse de doctorat de M. LOPEZ [DOC-07], au sein d'une convention CIFRE entre CLS et le LAAS-CNRS. L'objectif est de poursuivre la révision en profondeur de l'algorithme ARGOS. Une fonctionnalité intéressant particulièrement les clients et utilisateurs d'ARGOS est la localisation multi-modèles, *e.g.* pour des animaux dont le comportement est sujet à de subtils changements, pour des navires marchands tantôt en transit tantôt en action de pêche, etc. C'est pourquoi nous nous sommes penchés en priorité sur la problématique du filtrage pour un système dynamique à sauts Markoviens dont les vecteurs état continus associés aux divers modes de fonctionnement admettent *des dimensions et sémantiques hétérogènes*. Il s'agit d'un problème ouvert [Li 05, p. 1268].

Une extension du filtre IMM (“Interacting Multiple Model”) de BAR SHALOM [Blom 88] [Bar-Shalom 93], que nous avons appelée IMMHO (“IMM Heterogeneous-Order”), a été acceptée pour présentation prochaine dans [CIACL-24-DE-9]. Tout comme l'IMM standard, celle-ci intègre un filtre de KALMAN adapté à chacun des modes de fonctionnement. Son originalité réside dans sa stratégie d'interaction, nécessaire à ce que chaque filtre puisse être réinitialisé conditionnellement au mode associé. Nous avons montré qu'une coopération séparée des modèles affectant chaque couple de composantes du vecteur d'état « global » – *i.e.* obtenu par réunion des vecteurs d'état continus relatifs aux divers modes de fonctionnement – permet une approximation de la loi *a posteriori* dans la veine de l'IMM standard. L'IMMHO constitue ainsi une solution rigoureuse, qui ne requiert aucunement que les modèles d'état re-

latifs à chaque mode fassent intervenir le même jeu de variables, mais qui au contraire permet à un mode donné de partager de l'information sur les variables continues qui lui sont associées, voire, par ce biais, d'influencer indirectement l'estimation de variables n'apparaissant que dans d'autres modes.

II.1.4 Bilan en terme de publications, encadrements et projets

En termes quantitatifs, la thématique détection-estimation s'est déclinée en

- 3 thèses de doctorat [DOC-07] [DOC-04] [DOC-02], dont l'une a récemment commencé ; une participation officieuse à l'encadrement de deux autres thèses [DOC-0b] [DOC-0a] ;
- 5 stages de M2R/DEA [M2R-06] [M2R-05] [M2R-04] [M2R-02] [M2R-01], 3 stages autres [STAGE-11] [STAGE-07] [STAGE-02] ;
- en raison du caractère central de la thématique de l'interaction homme-robot dans RIA, une participation significative au projet de recherche européen COGNIRON [PROJ-C-2], et beaucoup plus ponctuelle aux projets ou collaborations [PROJ-P-2] [PROJ-P-1] ;
- les publications suivantes en Vision par Ordinateur ou Robotique : 2 articles de journaux [JSI-6-DE-2] [JSI-5-DE-1], les contributions en conférences ou symposia internationaux [CIACL-20-DE-8] [CIACL-19-DE-7] [SIACL-3-DE-2] [CIACL-16-DE-6] [CIACL-14-DE-5] [CIACL-13-DE-4] [CIACL-07-DE-3] [CIACL-05-DE-2] [CIACL-03-DE-1] [SIACL-1-DE-1], ainsi que 3 participations au congrès franco-phone RFIA [CNACL-5-DE-3] [CNACL-4-DE-2] [CNACL-3-DE-1] ; une publication plus théorique en Automatique [CIACL-24-DE-9].

II.1.5 Quelques références significatives

- [Andrieu 01] C. Andrieu, M. Davy & A. Doucet. *Improved Auxiliary Particle Filtering : Applications to Time-Varying Spectral Analysis*. In IEEE Workshop on Statistical Signal Processing, pages 309–312, 2001.
- [Andrieu 03] C. Andrieu, M. Davy & A. Doucet. *Efficient Particle Filtering for Jump Markov Systems. Application to Time-Varying Autoregressions*. IEEE Trans. on Signal Processing, vol. 51, pages 1762–1770, 2003.
- [Arulampalam 02] S. Arulampalam, S. Maskell, N. Gordon & T. Clapp. *A Tutorial On Particle Filters For On-Line Non-Linear/Non-Gaussian Bayesian Tracking*. IEEE Trans. On Signal Processing, vol. 50, no. 2, pages 174–188, 2002.
- [Balan 05] A. Balan, L. Sigal & M. Black. *A Quantitative Evaluation of Video-Based 3D Person Tracking*. In IEEE Int. Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'05), pages 349–356, 2005.
- [Bar-Shalom 93] Y. Bar-Shalom & X.R. Li. Estimation and tracking : Principles, techniques and software. YBS editions, Norwood, 1993.
- [Bensalah 96] F. Bensalah. *Estimation du Mouvement par Vision Active*. PhD thesis, Université de Rennes I, IRISA, 1996.
- [Blom 88] H.A.P. Blom & Y. Bar-Shalom. *The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients*. IEEE Trans. on Automatic Control, vol. 33, no. 8, pages 780–783, August 1988.
- [Chen 03] Z. Chen. Bayesian filtering : from Kalman filters to particle filters, and beyond. disponible sur <http://soma.ece.mcmaster.ca/~zhechen/BayesFilter.htm>, 2003.
- [CLS 08] Collecte Localisation Satellites. *Argos User Manual*, 2008.
- [Crisan 02] D. Crisan & A. Doucet. *A survey of convergence results on particle filtering for practitioners*. IEEE Trans. Signal Processing, vol. 50, no. 3, pages 736–746, 2002.
- [Daum 03a] F. Daum & J. Huang. *Mysterious computational complexity of particle filters*. In SPIE Signal and Data Processing of Small Targets, volume 4728, 2003.

- [Daum 03b] F. Daum & J. Huang. *Nonlinear filtering with Quasi-Monte Carlo methods*. In SPIE Signal and Data Processing of Small Targets, volume 5204, 2003.
- [Deutscher 99] J. Deutscher, B. North, B. Basile & A. Blake. *Tracking through Singularities and Discontinuities by Random Sampling*. In IEEE International Conference on Computer Vision (ICCV'99), 1999.
- [Deutscher 00] J. Deutscher, A. Blake & I. Reid. *Articulated Body Motion Capture by Annealed Particle Filtering*. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR'00), pages 126–133, 2000.
- [Doucet 00] A. Doucet, S.J. Godsill & C. Andrieu. *On Sequential Monte Carlo Sampling Methods For Bayesian Filtering*. Statistics and Computing, vol. 10, no. 3, pages 197–208, 2000.
- [Doucet 01] A. Doucet, N. De Freitas & N.J. Gordon. Sequential Monte Carlo methods in practice. Series Statistics For Engineering and Information Science. Springer-Verlag, New York, 2001.
- [Fang 94] K.T. Fang, Y. Wang & P.M. Bentler. *Some Applications of Number-Theoretic Methods in Statistics*. Statistical Science, vol. 9, no. 3, pages 416–428, 1994.
- [Fleury 95] S. Fleury, P. Souères, J.-P. Laumond & R. Chatila. *Primitives for Smoothing Mobile Robots Trajectories*. IEEE Transactions on Robotics and Automation, vol. 11, no. 3, pages 441–448, 1995.
- [Gavrila 99] D.M. Gavrila. *The Visual Analysis of Human Movement : a Survey*. Computer Vision and Image Understanding, vol. 73, no. 1, pages 82–98, 1999.
- [Guo 06] D. Guo & X. Wang. *Quasi-Monte Carlo Filtering in Nonlinear Dynamic Systems*. IEEE transactions on signal processing, vol. 54, no. 6, pages 2087–2098, 2006.
- [Gustafsson 91] F. Gustafsson. *A Two-Filter Off-Line Solution to Optimal Detection*. Rapport technique, Linköping University, 1991.
- [Gustafsson 96] F. Gustafsson. *The Marginalized Likelihood Ratio Test for Detecting Abrupt Changes*. IEEE Trans. on Automatic Control, vol. 41, no. 1, pages 66–78, 1996.
- [Isard 98a] M. Isard & A. Blake. *CONDENSATION – Conditional Density Propagation For Visual Tracking*. Int. J. Comput. Vision, vol. 29, no. 1, pages 5–28, 1998.
- [Isard 98b] M. Isard & A. Blake. *ICONDENSATION : Unifying Low-Level and High-Level Tracking in a Stochastic Framework*. In European Conference On Computer Vision (ECCV'98), pages 893–908, 1998.
- [Isard 98c] M. Isard & A. Blake. *A Mixed-State Condensation Tracker With Automatic Model-Switching*. In IEEE Int. Conf. on Computer Vision (ICCV'98), page 107, 1998.
- [Julier 97] S. Julier & J. Uhlmann. *A New Extension of the Kalman Filter to Nonlinear Systems*. In International Symposium on Aerospace/Defense Sensing, Simulation and Controls, 1997.
- [Kotecha 03a] J.H. Kotecha & P.M. Djuric. *Gaussian particle filtering*. IEEE Trans. on Signal Processing, vol. 51, no. 10, pages 2592–2601, 2003.
- [Kotecha 03b] J.H. Kotecha & P.M. Djuric. *Gaussian sum particle filtering*. IEEE Trans. on Signal Processing, vol. 51, no. 10, pages 2602–2612, 2003.
- [Li 05] X.R. Li & V.P. Jilkov. *Survey of Maneuvering Target Tracking. Part V : Multiple-Model Methods*. IEEE Trans. on Aerospace and Electronic Systems, vol. 41, no. 4, pages 1255–1321, October 2005.
- [Lopez 08] R. Lopez. *Évolution Algorithmique de la Localisation des Balises ARGOS*. Rapport technique, CLS & INSA Toulouse, 2008.
- [Lopez 10] R. Lopez. *Analyse d'Erreurs pour la Nouvelle Chaîne de Localisation ARGOS*. Rapport technique, CLS, 2010.
- [MacCormick 00a] J. MacCormick. *Probabilistic Modelling and Stochastic Algorithms for Visual Localisation and Tracking*. PhD thesis, Department of Engineering Science, University of Oxford, 2000.
- [MacCormick 00b] J. MacCormick & M. Isard. *Partitioned Sampling, Articulated Objects, and Interface-Quality Hand Tracking*. In European Conference on Computer Vision (ECCV'2000), pages 3–19, 2000.

- [Moeslund 01] T. Moeslund & E. Granum. *A Survey of Computer Vision-Based Human Motion Capture*. Computer Vision and Image Understanding, vol. 81, no. 3, 2001.
- [Ormoneit 01] D. Ormoneit, C. Lemieux & D.J. Fleet. *Lattice Particle Filters*. In Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence (UAI'01), pages 395–402, 2001.
- [Philomin 00] V. Philomin, R. Duraiswami & L.S. Davis. *Quasi-Random Sampling for CONDENSATION*. In European Conference on Computer Vision (ECCV'00), pages 134–149, Dublin, Ireland, 2000.
- [Pitt 99] M.K. Pitt & N. Shephard. *Filtering via Simulation : Auxiliary Particle Filters*. Journal of the American Statistical Association, vol. 94, no. 446, 1999.
- [Pérez 04] P. Pérez, J. Vermaak & A. Blake. *Data Fusion For Visual Tracking With Particles*. Proc. IEEE, vol. 92, no. 3, pages 495–513, 2004.
- [Royer 08] F. Royer & M. Lutcavage. *Filtering and Interpreting Location Errors in Satellite Telemetry of Marine Animals*. Journal of Experimental Marine Biology and Ecology, vol. 359, no. 1, pages 1–10, 2008.
- [Sminchisescu 02] C. Sminchisescu. *Estimation Algorithms for Ambiguous Visual Models - Three Dimensional Human Modeling and Motion Reconstruction in Monocular Video Sequences*. PhD thesis, INRIA Grenoble, 2002.
- [Torma 03] P. Torma & C. Szepesvári. *Sequential Importance Sampling For Visual Tracking Reconsidered*. In AI and Statistics, pages 198–205, 2003.
- [Van Der Merwe 01a] R. Van Der Merwe, N. De Freitas, A. Doucet & E. Wan. *The Unscented Particle Filter*. In Advances in Neural Information Processing Systems 13, 2001.
- [Van Der Merwe 01b] R. Van Der Merwe & E. Wan. *The Square-Root Unscented Kalman Filter for State and Parameter Estimation*. In IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'01), 2001.
- [van der Merwe 03] R. van der Merwe & E. Wan. *Gaussian Mixture Sigma-Point Particle Filters for Sequential Probabilistic Inference in Dynamic State-Space Models*. In IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'2003), April 2003.
- [Willsky 76] A.S. Willsky & H.L. Jones. *A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems*. IEEE Trans. on Automatic Control, vol. 21, no. 1, pages 108–112, 1976.
- [Wu 99] Y. Wu & T.S. Huang. *Vision-based Gesture Recognition : a Review*. In International Workshop on Gesture-Based Communication, pages 103–105, 1999.

II.2 Résumé condensé des travaux en Asservissement Visuel

Du fait de l'évolution de la technologie des capteurs pour la Robotique, un nouvel axe de recherche en commande de robots est apparu dans les années 90. Il concerne l'exploitation des capteurs extéroceptifs dans des boucles de commande, ce qui offre de nouvelles perspectives pour la réalisation de tâches de positionnement ou de suivi. En particulier, une formalisation du problème de commande dans les espaces de ces capteurs a été proposée, qui permet de s'affranchir d'une étape de localisation absolue, coûteuse en ressources et source d'incertitudes. Cette nouvelle approche méthodologique est souvent désignée par le vocable « commande référencée capteurs » (“sensor-based control”). Dans cette veine, l'utilisation de caméras vidéos a conduit à envisager des « commandes référencées vision » – ou « asservissements visuels » [Hutchinson 96][Chaumette 06][Chaumette 07]. Les schémas dits « 3D », ou “situation-based”, expriment la consigne et la variable contrôlée en termes d'un couple (position, orientation) et nécessitent donc une étape de localisation visuelle. Les stratégies « 2D », ou “image-based”, élaborent la commande sur la base de primitives image – points, droites, aires,... – qu'il s'agit d'asservir à une configuration de référence [Espiau 92].

II.2.1 Premiers pas...

Dans le cadre du stage [STAGE-01], nous avons réalisé un asservissement visuel de la position de l'organe terminal d'un robot manipulateur ROBOSOFT GT 6A pour une orientation figée. Ce schéma, basé sur le centre de gravité et l'aire de la projection d'une cible lumineuse dans le plan image d'une caméra “eye-in-hand”, a été couplé avec un estimateur de mouvement élémentaire. Pour une cible en déplacement rectiligne ou circulaire uniforme avec une vitesse horizontale et verticale de l'ordre de 15cm.s^{-1} et 4cm.s^{-1} , respectivement, le traînage était suffisamment faible pour permettre la saisie. Une difficulté méthodologique majeure était liée au système d'exploitation Albatros accompagnant le robot, qui introduisait des retards variables et non maîtrisés dans la boucle et empêchait toute identification rigoureuse de la dynamique du procédé. Une modélisation astucieuse à temps discret de l'asservissement et une analyse rudimentaire au moyen de lieux d'Evans ont permis d'établir des limites de stabilité théoriques du même ordre que celles obtenues expérimentalement. Malgré de très bons résultats expérimentaux, nous avons jugé le contenu théorique trop limité et trop *ad hoc* pour présenter ce travail à la communauté scientifique.

II.2.2 Asservissement visuel multicritère

Cette activité, que j'ai initiée par l'encadrement de la thèse de M^{le} Delphine BELLOT [DOC-01], vise à développer des techniques d'*analyse et synthèse multicritères* de schémas de positionnement référencé vision, *i.e.* prenant en compte l'ensemble des contraintes de la tâche. En effet, au-delà du volet « convergence », tout asservissement visuel doit assurer la visibilité de la projection de la cible dans le plan image, garantir l'acceptabilité de la trajectoire 3D, tenir éventuellement compte des saturations d'actionneurs, gérer les singularités, etc. L'analyse d'une loi donnée peut donc consister à caractériser des régions dans lesquelles elle permet la réalisation de la tâche, à caractériser la vitesse de positionnement, à évaluer les gammes de valeurs de certaines quantités, ou à étudier la sensibilité à divers types d'incertitudes. Sur le plan de la synthèse, l'objectif est de garantir de telles propriétés, en contraignant si besoin certains signaux : transmis aux actionneurs, reçus par les capteurs, liés au déplacement, etc. Ces problèmes sont encore ouverts.

II.2.2.1 Proposition d'un cadre générique : les systèmes non linéaires incertains « rationnels » soumis à des contraintes inégalités rationnelles

Nous avons proposé une modélisation d'état originale, qui unifie les asservissements 3D et 2D “eye-in-hand” dans le contexte de la stabilité/stabilisation d'un système non linéaire sous contraintes. En contexte cinématique, le vecteur d'état de la boucle ouverte décrit la situation relative caméra-cible. Ceci implique de considérer l'asservissement 2D en tant que la stabilisation de cette situation sur la base des informations visuelles perçues, ce qui présente plusieurs avantages par rapport à la formulation classique consistant à asservir la configuration des informations visuelles à une valeur de référence.

Nous nous sommes ensuite restreints aux *systèmes non linéaires incertains « rationnels »*, *i.e.* admettant une représentation d'état semblable au cas linéaire mais dont les matrices sont rationnelles en le vecteur d'état et les incertitudes. Nous avons montré que leur analyse ou commande *sous des contraintes inégalités scalaires rationnelles* recouvre l'analyse multicritère de la majorité des commandes référencées vision existantes de même que la synthèse multicritère de nouvelles classes de contrôleurs. Bien que l'analyse et la commande des systèmes

rationnels incertains soient relativement peu abordées en Automatique, la littérature propose quelques pistes, principalement d'inspiration « linéariste » car issues de la commande robuste linéaire.

II.2.2.2 Analyse et Synthèse multicritères par des arguments de stabilité quadratique

Nos premières études se sont inscrites dans la veine des travaux d'EL GHAOUI [El Ghaoui 96] sur l'analyse et la commande de systèmes rationnels soumis à des contraintes linéaires. La boucle fermée référencée vision est « *linéarisée globalement* » [Boyd 94] – *i.e.* « *immergée* » – dans un système linéaire à paramètres variants incertains [LPV]. L'analyse et la synthèse sont alors ramenées à des problèmes de *commande robuste linéaire*, au sens où il suffit qu'une propriété soit satisfaite par [LPV] pour qu'elle le soit également par l'asservissement visuel considéré. Le conservatisme (pessimisme) concomitant est heuristiquement limité du fait que l'incertitude –de type LFT (Linear Fractional Transform = Transformation Fractionnelle Linéaire)– de [LPV] est fortement structurée.

L'approche repose sur la recherche d'une fonction de LYAPUNOV $V(\xi) = \xi'P\xi$ pour [LPV], où la variable de décision matricielle P est assujettie à un ensemble de conditions complémentaires. Celles-ci assurent que $V(\cdot)$ est également une fonction de LYAPUNOV pour la boucle fermée originelle sur un paralléléotope donné de l'espace d'état, et que l'intérieur de l'iso-coût $\mathcal{E} \triangleq \{\xi : V(\xi) \leq 1\}$ est un *bassin d'attraction multicritère*, *i.e.* un ensemble de conditions initiales pour lesquelles la tâche est réalisée dans le respect des contraintes. L'intégration des conditions complémentaires sur P est effectuée au moyen de la S-procédures [Boyd 94]. Le problème d'analyse ou de synthèse est alors résolu *via* un programme d'optimisation sous contraintes en P et/ou $Q = P^{-1}$. Si celles-ci prennent la forme d'*Inégalités Linéaires Matricielles* (LMIs) [Boyd 94], alors le programme est convexe et peut être résolu numériquement en temps polynômial et avec une précision arbitraire au moyen de « solveurs » dédiés [CIACL-02-AV-1].

Nous avons d'abord prolongé les travaux d'EL GHAOUI de façon à introduire des contraintes rationnelles [CIACL-04-AV-2]. L'analyse multicritère de schémas quelconques, de même que la synthèse multicritère d'asservissements 3D, conduisent à des LMIs. En revanche, des problèmes d'optimisation non convexes, donc plus difficiles, sont généralement obtenus pour la synthèse d'asservissements 2D.

Il s'est rapidement avéré que la symétrie par rapport à 0 et la convexité des bassins d'attraction ellipsoïdaux issus de fonctions de LYAPUNOV quadratiques ne permet de considérer que des cas peu pertinents, où les situations capteur-cible initiale et désirée sont très voisines. Deux contributions importantes ont alors été développées en vue de contourner cette limitation fondamentale [DOC-01] [JSI-2-AV-1]. Pour l'analyse multicritère, une méthodologie basée sur des programmes LMIs a été définie de façon à compléter incrémentalement un bassin de convergence donné en une approximation interne plus fine du bassin d'attraction multicritère maximal. Reposant sur des séquences de linéarisations globales de la boucle fermée sur des parallélétopes de petite taille, elle conduit à des résultats suffisamment peu pessimistes pour permettre l'élaboration de conclusions non triviales. Cependant, elle suppose un nombre réduit de degrés de liberté du robot, et est relativement lourde à mettre en œuvre. Une seconde méthode s'applique également à l'analyse et étend significativement les cas de succès de la synthèse. Elle consiste à augmenter l'équation de dynamique au moyen d'un état supplémentaire en évolution autonome, de telle sorte que le bassin d'attraction de ce

système augmenté soit une « tranche » d'ellipsoïde. La projection de ce bassin dans l'espace d'état initial n'étant pas symétrique par rapport à 0, la variété des problèmes qui peuvent être traités s'en trouve considérablement étendue. Les programmes d'optimisation obtenus sont soit convexes – à base de LMIs – soit non convexes, selon des degrés de liberté dans la caractérisation de l'état supplémentaire en évolution autonome.

D'autres prolongements auraient sans doute été possibles, mais par trop incrémentaux. En effet, la restriction à des fonctions de LYAPUNOV quadratiques et la nécessité d'assurer des considérations de bien posé sur des parallélotopes de l'espace d'état constituent deux limitations fondamentales de cette première approche. C'est pourquoi j'ai ensuite exploré les travaux plus récents de TROFINO [Trofino 00] qui laissaient présager des résultats moins conservatifs [STAGE-05]. Cependant, en raison de mes autres engagements en recherche [DOC-02][DOC-03] et du fait que cette thématique n'était pas prioritaire dans RIA, je l'ai temporairement mise en sommeil.

II.2.2.3 Représentations algébro-différentielles et fonctions de Lyapunov plus évoluées pour l'analyse multicritère

Après être entré en contact avec M. TROFINO et conséutivement à notre rencontre à IEEE CDC'2003, la collaboration [COLL-2] a été établie à partir de fin 2004 avec M. COUTINHO, ancienement doctorant de M. TROFINO sur la thématique de l'analyse en stabilité des systèmes rationnels par les fonctions de LYAPUNOV biquadratiques ou polyquadratiques. Leur approche [Trofino 02][Coutinho 04] repose sur une reformulation du modèle d'état du système considéré en une *représentation algébro-différentielle* (DAR = Differential-Algebraic Representation). Les contraintes que doivent satisfaire les paramètres libres de la fonction de LYAPUNOV afin qu'un iso-coût de celle-ci définisse un bassin d'attraction sont incorporées par application successive de la S-procédure et du lemme de Finsler [Boyd 94], conduisant ainsi à des conditions suffisantes LMIs.

Notre première contribution [SIACL-2-AV-1] a été de montrer que le formalisme DAR et la méthodologie d'analyse qui en est issue s'étendent de manière élégante et efficace à la détermination de bassins d'attraction multicritères en présence de contraintes rationnelles sur le vecteur d'état. Cette nouvelle approche de l'analyse multicritère des systèmes rationnels conduit encore à un programme d'optimisation LMI et s'avère significativement plus générique que nos premiers travaux. En effet, les iso-coûts des classes de fonctions de LYAPUNOV considérées ne présentent pas nécessairement les propriétés géométriques de symétrie et de convexité pénalisantes pour notre problème de Robotique. De plus, ces fonctions de LYAPUNOV doivent être définies sur des polytopes convexes n'ayant pas nécessairement la forme de parallélotopes, ce qui simplifie considérablement les considérations de bien posé. Enfin, les fonctions de LYAPUNOV peuvent dépendre du vecteur des paramètres incertains dès lors que la dérivée de ceux-ci peut être bornée a priori, ce qui limite encore davantage le pessimisme. En revanche, un dilemme fondamental est apparu entre la richesse de la classe des fonctions de LYAPUNOV considérées d'une part, et la taille et le conditionnement numérique des programmes d'optimisation obtenus d'autre part [M2R-07]. Du fait que la relative limitation des conclusions obtenues par une approche biquadratique et de la complexité calculatoire engendrée par des fonctions de LYAPUNOV d'ordres supérieurs, nous avons posé les bases d'une solution « à mi-chemin » par fonctions de LYAPUNOV biquadratiques par morceaux [CIACL-11-AV-4].

Cette approche a été complétée et évaluée plus extensivement dans [STAGE-06] et dans

la thèse de M. Sylvain DUROLA [COCI-1-AV-1][CIACL-18-AV-6].

II.2.3 Dualité entre asservissement visuel et localisation visuelle

Notre formulation mathématique de l'asservissement visuel est telle que le problème dual – au sens de la dualité observation-commande – consiste en l'estimation de la situation relative entre un capteur visuel et une cible à partir des informations qu'il délivre sur celle-ci. C'est à notre connaissance la première fois que cette relation de dualité a été exhibée. La dernière contribution de la thèse de M^{lle} Delphine BELLOT [DOC-01], dans la suite de [M2R-03], a concerné l'applicabilité à ce problème de localisation de techniques duales de celles utilisées en commande. Nous avons supposé d'éventuelles incertitudes dans les modèles de la cible et de la dynamique du robot, et considéré une description ensembliste des bruits de dynamique et de mesure. La « linéarisation globale » du modèle – à temps discret – de la boucle ouverte permet de se ramener à un problème de *filtrage linéaire robuste en contexte ensembliste*, au sens où une approximation extérieure – *e.g.* en terme d'ellipsoïde de confiance – de l'ensemble des vecteurs d'état du système à paramètres variants dans lequel est effectué l'immersion, contient la situation relative capteur-cible que l'on veut estimer. Des résultats tout à fait intéressants [CIACL-06-AV-3] ont été obtenus *via* une extension mineure des travaux de EL GHAOUI et CALAFIORE [El Ghaoui 01] concernant le filtrage ensembliste des systèmes linéaires à temps discret dépendant rationnellement de paramètres incertains variants.

Dans le cadre de la thèse de M. Sylvain DUROLA [DOC-05], nous avons défini une stratégie de filtrage ensembliste pour un système discret non linéaire incertain défini par une représentation algébro-récurrente (RAR = Recursive Algebraic Representation) [CIACL-17-AV-5]. Celle-ci peut être utilisée en lieu et place de l'approche d'EL GHAOUI et CALAFIORE en tant que solution LMI au problème de la localisation visuelle.

II.2.4 Conception et implémentation d'un environnement de simulation de commandes en robotique

Afin de valider des commandes référencées vision préalablement à leur expérimentation sur nos plate-formes, j'ai conçu et co-développé, par les encadrements [STAGE-08][STAGE-04][STAGE-03], l'environnement de simulation MAVS (MAVS = “MAVS Ain’t Visual Servoing”). MAVS consiste en l’interfaçage du logiciel de CACSD MATLAB-SIMULINK et du serveur d'affichage 3D GDHE (Graphical Display for Hilare Experiments, <http://www.laas.fr/~matthieu/gdhe>) développé par M. Matthieu HERRB au LAAS-CNRS autour de la bibliothèque graphique OpenGL. Il permet la définition d'une grande variété de tâches robotiques incluant des capteurs visuels et/ou proximétriques (lasers), l'intégration de primitives de traitement de l'information codées en C/C++, l'implémentation de lois de commande de complexité quelconque, la simulation de l'expérience considérée MATLAB-SIMULINK, la visualisation du résultat en 3D ainsi que dans les espaces des capteurs extéroceptifs. Ce logiciel, actuellement à usage interne dans RIA, sera « toiletté » puis rendu disponible publiquement.

II.2.5 Bilan en terme de publications, encadrements et projets

En termes quantitatifs, la thématique de l'asservissement multicritère s'est déclinée en – 2 thèses de doctorat [DOC-05] [DOC-01], dont l'une se termine ;

- 2 stages de M2R/DEA [M2R-07] [M2R-03], 2 stages autres [STAGE-06] [STAGE-05] ; à ceux-ci se rajoutent 3 stages dédiés au développement du logiciel MAVS [STAGE-08] [STAGE-04] [STAGE-03] ;
- l'obtention d'un financement pour le projet STIC-AMSUD NCTVS [PROJ-R-2], en cours ; une participation ponctuelle au projet EGOCENTRE du programme ROBEA [PROJ-C-1] ;
- les publications suivantes dans le domaine de l'Automatique : 1 article de journal [JSI-2-AV-1], 1 article de journal connexe [JSI-4-AV-2], 6 articles en conférences et symposia [CIACL-17-AV-5] [CIACL-11-AV-4] [SIACL-2-AV-1] [CIACL-06-AV-3] [CIACL-04-AV-2] [CIACL-02-AV-1] ; les publications suivantes dans le domaine de la Robotique : 1 chapitre d'ouvrage [COCI-1-AV-1] et 1 conférence [CIACL-18-AV-6], tous deux sur invitation de MM G. CHESI et K. HASHIMOTO ;
- 3 séminaires invités [SEM-03] [SEM-04] [SEM-08] (avec bien sûr une mention particulière pour l'atelier LMI'04) ; 2 séminaires [SEM-05] [SEM-06] dans les GdR Automatique et Robotique/ISIS, un séminaire dans le GdR ISIS durant la thèse de doctorat de M^{le} Delphine BELLOT ainsi que plusieurs séminaires internes lors des ateliers de pôle ou de groupe au LAAS-CNRS.

II.2.6 Quelques références significatives

- [Boyd 94] S. Boyd, L. El Ghaoui, E. Feron & V. Balakrishnan. *Linear matrix inequalities in system and control theory*. SIAM, 1994.
- [Chaumette 06] F. Chaumette & S. Hutchinson. *Visual Servo Control, Part I : Basic Approaches*. IEEE Robotics and Automation Magazine, vol. 13, no. 4, pages 82–90, 2006.
- [Chaumette 07] F. Chaumette & S. Hutchinson. *Visual Servo Control, Part II : Advanced Approaches*. IEEE Robotics and Automation Magazine, vol. 14, no. 1, pages 109–118, 2007.
- [Coutinho 04] D.F. Coutinho, A.S. Bazanella, A. Trofino & A.S. Silva. *Stability Analysis and Control of a Class of Differential-Algebraic Nonlinear Systems*. Int. Jour. of Robust and Nonlinear Control, vol. 14, pages 1301–1326, 2004.
- [El Ghaoui 96] L. El Ghaoui & G. Scorletti. *Control of Rational Systems using Linear-Fractional Representations and Linear Matrix Inequalities*. Automatica, vol. 32, no. 9, pages 1273–1284, 1996.
- [El Ghaoui 01] L. El Ghaoui & G. Calafio. *Robust Filtering for Discrete-Time Systems with Bounded Noise and Parametric Uncertainty*. IEEE Trans. on Automatic Control, vol. 46, no. 7, 2001.
- [Espiau 92] B. Espiau, F. Chaumette & P. Rives. *A New Approach to Visual Servoing in Robotics*. IEEE Trans. on Robotics and Automation, vol. 8, no. 3, pages 313–326, 1992.
- [Hutchinson 96] S. Hutchinson, G.D. Hager & P.I. Corke. *A Tutorial on Visual Servo Control*. IEEE Trans. on Robotics and Automation, vol. 12, no. 5, pages 651–670, 1996.
- [Trofino 00] A. Trofino. *Bi-Quadratic Stability for Nonlinear Systems*. In IFAC Symp. on Robust Control Design (ROCOND'00), Prague, 2000.
- [Trofino 02] A. Trofino. Local, Regional and Global Stability : An LMI Approach for Uncertain Nonlinear Systems. Tech. Report, DAS-CTC, Univ. Federal de Santa Catarina, Brésil, 2002.

II.3 Résumé condensé des travaux en Audition en Robotique

En 2003, dans le cadre du projet EGOCENTRE [PROJ-C-1], M. Philippe SOUÈRES m'a proposé de co-encadrer la thèse de M. Sylvain ARGENTIERI [DOC-03]. Le but était de développer une tête multi-capteurs combinant entre autres des modalités visuelles et auditives, pour réaliser l'enchaînement de tâches de haut-niveau. M. ARGENTIERI a souhaité se concentrer sur le volet « audition », avec comme objectif premier la localisation et le filtrage spatial de sources

sonores. Ce travail étant fortement ancré dans le Traitement du Signal, M. SOUÈRES et moi avons décidé que j'assurerais la majeure partie de l'encadrement.

Comparativement à la Vision par Ordinateur, l'Audition en Robotique (“Robot Audition”) est une thématique relativement récente [Okuno 04]. Les approches binaurales – *i.e.* sur la base d'une tête munie de deux pavillons – sont très intéressantes dans leur principe mais donnent encore des résultats assez mitigés. Nous avons donc opté pour l'exploitation d'un réseau – ou antenne – linéaire de microphones. Bien que la littérature en Acoustique ou Traitement d'Antenne comprenne de nombreuses méthodes de localisation ou de filtrage spatial, la Robotique apporte des contraintes inédites : embarquabilité et intégration du capteur – taille, puissance de traitement, consommation –, traitement temps réel – *e.g.* génération de primitives de localisation à une dizaine d'Hertz –, sources large bande – *e.g.* la voix humaine – émettant en champ proche ou lointain, bruits, réverbérations,...

II.3.1 Formation de voie pour l'extraction et la localisation de sources sonores large bande

La *formation de voie* (“beamforming”) consiste à émuler une antenne continue *via* l'insertion de filtres numériques – généralement RIF – en aval des microphones, dont on somme les sorties [Van Veen 88]. Elle peut servir de base au filtrage spatial de sources sonores mais aussi à leur localisation, par détermination de l'énergie acoustique incidente selon diverses directions de polarisation de l'antenne. La formation de voie large bande conventionnelle, qui est la plus utilisée en Robotique, applique un décalage temporel sur les signaux perçus de façon que la sommation soit constructive pour la direction de polarisation choisie. La largeur importante du lobe de directivité principal obtenu aux basses fréquences ne permet cependant pas une focalisation convenable de l'antenne sur une voix humaine. C'est pourquoi nos premiers travaux ont concerné la synthèse d'un diagramme de directivité quasi-invariant sur la bande [400Hz; 3kHz], qui capture une proportion importante de la puissance de la voix humaine tout en assurant son intelligibilité pour un traitement ultérieur.

II.3.1.1 Formation de voie par analyse modale et optimisation convexe

Nous avons d'abord proposé une extension de la formation de voie faible bande de WANG *et al.* [Wang 03] à notre problème sous l'hypothèse de champ lointain. Celle-ci consiste en la minimisation de l'erreur pire cas entre le diagramme de l'antenne et un diagramme de référence sur une grille en azimuth et en fréquences [CIACL-08-AR-1]. Elle donne lieu à un programme d'optimisation sur le cône du second ordre (SOCP = Second-Order Cone Program), convexe, dont la résolution numérique peut s'effectuer par des techniques comparables à celles référencées au §II.2.2 [Boyd 94][Luo 03]. Bien que cette formulation autorise la gestion d'exigences antagonistes – *e.g.* limitation du gain d'antenne maximal en réponse à un bruit blanc – elle conduit à un programme de grande dimension souvent mal conditionné numériquement, et dont le résultat dépend fortement de la grille sélectionnée.

Notre contribution majeure se situe au confluent de l'*analyse modale* de diagrammes d'antenne [Abhayapala 99] et de l'*optimisation convexe*. Le diagramme de référence et celui faisant l'objet de l'optimisation sont décomposés sur la base des harmoniques sphériques. On minimise alors l'erreur pire cas séparant les « coefficients modaux » obtenus [CIACL-10-AR-3]. Comme ces coefficients ne dépendent pas des variables spatiales du problème, cette dépendance étant reportée sur les fonctions de base – harmoniques sphériques – et sur des fonctions de normali-

sation, seule une grille en fréquence est nécessaire. Un SOCP est à nouveau obtenu, autorisant comme précédemment l'incorporation de contraintes. Sa taille réduite et son meilleur conditionnement rendent sa résolution aisée. En outre, la synthèse en champ lointain ou en champ proche peuvent être effectuées au sein d'un formalisme unique. Enfin, notre méthode permet également de mieux appréhender les distorsions subies par tout diagramme de directivité invariant en fréquence lorsque la distance antenne-source varie.

II.3.1.2 Implémentation pratique d'une formation de voie

L'optimisation d'un diagramme de directivité sur un sous-ensemble spatio-fréquentiel limité exige d'étudier ses propriétés hors de ce domaine ainsi que leur incidence sur la réalisation physique du capteur. Ceci est particulièrement important lorsqu'une explosion du gain de l'antenne se produit sur certaines bandes de fréquences. Nous avons donc étudié le *lien entre le diagramme de directivité théorique d'une antenne et la réponse de son implémentation physique* intégrant une chaîne d'acquisition complète [CIACL-09-AR-2]. Sur cette base, nous avons montré que le conditionnement du programme d'optimisation peut être significativement amélioré si celui-ci exploite les caractéristiques de la chaîne d'acquisition [CIACL-12-AR-4].

Le coût calculatoire de l'exécution en temps réel d'une formation de voie – 8 filtres d'ordre 100 environ – est acceptable. Cependant, la focalisation simultanée de l'antenne selon plusieurs directions d'écoute s'avère prohibitive. C'est pourquoi l'implémentation sur FPGA de notre formation de voie est effectuée selon des techniques de *convolution rapide* de type “overlap-and-save” [Borgerding 06][STAGE-13].

II.3.2 Méthodes à haute résolution pour la localisation et la détection de sources sonores large bande

Des cartes acoustiques de l'environnement peuvent être construites par formations de voie. Il suffit de polariser électroniquement l'antenne selon les directions d'écoute à balayer, puis à intégrer l'énergie acoustique en provenance de chacune d'elles sur une fenêtre temporelle glissante. Comme cela a été montré en simulation dans [CIACL-09-AR-2], une formation de voie invariante en fréquence conduit à une localisation très supérieure à la formation de voie conventionnelle. Cependant, les résultats demeurent intrinsèquement limités par la taille – ou « ouverture » – de l'antenne, et s'avèrent assez décevants sur des signaux réels en environnement modérément bruité. Nous avons par conséquent envisagé d'explorer les méthodes à haute résolution, qui admettent de meilleures propriétés théoriques [Krim 96][Gonen 99].

II.3.2.1 “Broadband Beamspace MUSIC”

La méthode MUSIC (MULTiple SIgnal Classification) [Schmidt 86] est probablement la *méthode à haute-résolution* la plus renommée. Elle permet la localisation de sources faible bande sur la base de la décomposition en valeurs propres généralisées du faisceau matriciel constitué des matrices de covariance des signaux bruts et des bruits perçus par les microphones. Un pseudo-spectre fonction des variables spatiales – azimuth et distance pour une antenne linéaire – est déterminé, dont les pics indiquent les positions des sources. Cette fonction est en fait l'inverse de la distance Euclidienne quadratique entre le “sous-espace signal” issu de la décomposition en valeurs propres généralisées et le vecteur d'antenne. La difficulté d'une extension large bande de MUSIC par son application sur un ensemble de “bins” fréquentiels réside dans le fait que les sous-espaces signal associés à chaque bin sont distincts – ou « non

alignés » –, ce qui interdit la définition d'un sous-espace signal global par moyennage. La seule méthode à haute résolution utilisée en robotique consistait par conséquent en le moyennage des pseudo-spectres calculés sur chaque bin [Asano 99]. Cette approche « naïve » conduit à des résultats satisfaisants, ce qui a fait d'elle la référence dans la communauté, mais possède un coût calculatoire exorbitant.

Sur la base des travaux de WANG et KAVEH [Wang 85] et de WARD et ABHAYAPALA [Ward 04], nous avons proposé une *extension large bande cohérente* de MUSIC dans des espaces de sorties de formations de voie (“*Broadband Beamspace MUSIC*”) [CIACL-15-AR-5][M2R-08]. Les matrices de covariance des signaux et des bruits sont pré- et post-multipliées par des coefficients de formations de voie spécifiques, dont la synthèse est immédiate *via* notre méthode présentée au §II.3.1.1 [CIACL-10-AR-3]. Leur sommation conduit à un faisceau de matrices de covariance « alignées », dont la décomposition en valeurs et vecteurs propres généralisés conduit à la définition d'un sous-espace signal et d'un pseudo-spectre uniques. Cette méthode est particulièrement bien adaptée à la Robotique : notre implémentation permet la localisation d'au plus 3 sources, et présente un coût calculatoire très inférieur –dans un facteur d'au moins 50– à l'approche « naïve ».

II.3.2.2 Détection de sources

Les caractéristiques spatio-temporelles du bruit ambiant étant supposées connues, détecter des sources consiste à estimer quel est leur nombre le plus vraisemblable au regard des signaux perçus par l'antenne de microphones. Notre approche de ce problème a été d'instancier des travaux théoriques développés dans la communauté Traitement d'Antenne, qui montrent que MUSIC faible bande et ses extensions cohérentes large bande offrent un cadre théoriquement étayé pour la détection de sources à partir d'un réseau de transducteurs [Wax 85][Wang 85]. Le détecteur [CIACL-22-AR-7] que nous avons récemment adjoint à notre schéma de localisation “*Broadband Beamspace MUSIC*” (cf. [CIACL-15-AR-5]) estime le nombre de sources comme celui qui confère à la matrice de covariance alignée des signaux la structure la plus vraisemblable vis à vis des données brutes perçus par les microphones. Le problème est transposé dans le contexte de l'*identification statistique*. La ligne directrice est la minimisation de la divergence de Kullback-Leibler, laquelle conduit à l'estimé du maximum de vraisemblance dans le cas d'un seul modèle. Cependant, les modèles en compétition admettent un nombre distinct de paramètres libres. On recherche alors, parmi les nombres de sources candidats, celui qui minimise le *critère du minimum d'information d'AKAIKE* (MAICE = Minimum AKAIKE Information Criterion Estimate) [Akaike 74]. Cette stratégie ne nécessite pas de réglage de seuil *a priori* et s'avère très légère sur le plan calculatoire.

II.3.3 Détection de “patterns” audio

Le vocable « détection » admet ici un acceptation distincte. Il s'agit de détecter et d'isoler un « motif » –ou “pattern” – audio dans un signal temporel, *e.g.* le signal résultant de la focalisation spatiale du capteur vers un azimuth donné. Nous avons considéré la détection de patterns audio stationnaires noyés dans du bruit indépendant stationnaire additif, sur la base de leurs statistiques respectives [CIACL-23-AR-8]. Le support théorique est le *filtrage adapté stochastique*, qui à notre connaissance n'a été développé et exploité que pour des applications d'océanographie [Cavassilas 91]. Disposant de l'autocovariance du pattern faisant l'objet de la détection et de l'autocovariance du bruit, une base de filtres est définie hors ligne de façon

à augmenter le rapport signal sur bruit. La séquence audio est filtrée par – *i.e.* projetée sur cette base en ligne. Le vecteur de composantes résultant se distribue différemment selon que la séquence contient le bruit seul ou bien le pattern bruité. Ces distributions étant connues théoriquement, un test de vraisemblance peut être mis en place pour la détection en temps réel. Une extension de cette stratégie pourrait concerner la *détection d'activité vocale* (VAD = Voice Activity Detection).

II.3.4 Le capteur EAR

Afin d'implémenter et d'évaluer l'ensemble de nos fonctions auditives, nous avons développé le *capteur auditif intégré EAR* (“Embedded Audition for Robotics”). Celui-ci est constitué d'une antenne linéaire de 8 microphones régulièrement espacés de $\frac{\lambda_{3\text{kHz}}}{2} = 5.66 \text{ cm}$, d'une unité de polarisation des microphones et de conditionnement des signaux, d'une chaîne d'acquisition 8 voies entièrement programmable, d'une unité de traitement FPGA et d'une liaison USB avec un hôte UNIX. Il permet l'acquisition de données audio brutes et le filtrage spatial à 15 kHz, ainsi que la génération de primitives de localisation à 15 Hz.

Un premier prototype a été conçu par M. ARGENTIERI en collaboration avec M. Jérôme MANHÈS. Son développement a été poursuivi après le départ de M. ARGENTIERI. Une évolution plus puissante est en cours de fabrication. Les deux versions partagent une architecture VHDL modulaire et évolutive pour l'intégration de fonctions auditives sur le FPGA, ainsi que des librairies C/C++ relatives à la communication bas-niveau et au prototypage rapide de primitives évoluées [CIACL-21-AR-6]. J'ai co-encadré avec M. MANHÈS plusieurs stages liés à ces développements expérimentaux [STAGE-13][M2R-10][STAGE-09][STAGE-10].

II.3.5 Bilan en terme de publications, encadrements et projets

En termes quantitatifs, la thématique de l'audition en robotique s'est déclinée en

- 2 thèses de doctorat [DOC-06] [DOC-03], dont l'une se termine ;
- 1 collaboration en cours avec un visiteur post-doctorant [POSTDOC-1] ;
- 3 stages de M2R/DEA [M2R-10] [M2R-09] [M2R-08], 1 stage autre sur un sujet théorique [STAGE-12], 3 stages à vocation expérimentale [STAGE-13] [STAGE-09] [STAGE-10] ;
- l'obtention d'un financement pour le projet ANR-JST BINAHR [PROJ-R-3], en collaboration avec les équipes des deux chercheurs les plus en pointe sur le plan international (Prof. H.G. OKUNO, Kyoto University et Dr. K. NAKADAI, Honda Research Institute & Tokyo Institute of Technology) ; l'obtention d'un financement pour le projet LAAS-CNRS AUDIO_HRP2 [PROJ-R-1] ; la participation aux projets européens COMMROB [PROJ-C-3] et COGNIRON [PROJ-C-2], ainsi qu'au projet EGOCENTRE du programme ROBEA [PROJ-C-1] ;
- le rapport d'une thèse de doctorat de l'Australian National University [RAPPORT-1] ;
- les publications suivantes dans le domaine de la Robotique : 1 article de journal [JSI-7-AR-1] sur invitation de K. NAKADAI, des contributions [CIACL-23-AR-8] [CIACL-22-AR-7] [CIACL-21-AR-6] [CIACL-15-AR-5] [CIACL-10-AR-3] [CIACL-09-AR-2] aux sessions invitées “Robot Audition” de la conférence IEEE/RSJ IROS en 2005, 2006, 2007, 2009 et 2010 ($\times 2$) ainsi que l'article de conférence [CIACL-08-AR-1] ; les publications suivantes dans le domaine du Traitement du Si-

- gnal : 1 article en conférence internationale [CIACL-12-AR-4] et 1 colloque francophone [CNACL-2-AR-1] ;
- 2 séminaires invités [SEM-09] [SEM-07] ([SEM-09] ayant eu lieu à OKUNO's lab) ;
 - l'encadrement de projets étudiants (TER puis stage du M1 IUP Systèmes Intelligents) avec l'entreprise NEXTER, concernant un capteur audio embarquable breveté par NEXTER dans les années 90.

II.3.6 Quelques références significatives

- [Abhayapala 99] P.T.D. Abhayapala. *Modal Analysis and Synthesis of Broadband Nearfield Beamforming Arrays*. PhD thesis, Australian National University, 1999.
- [Akaike 74] H. Akaike. *A New Look at the Statistical Model Identification*. IEEE Trans. on Automatic Control, vol. 19, no. 6, pages 716–723, 1974.
- [Asano 99] F. Asano, H. Asoh & T. Matsui. *Sound Source Localization and Signal Separation for Office Robot Jijo-2*. In IEEE/SICE/RSJ Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems, pages 243–248, Taipei, Taiwan, 1999.
- [Borgerding 06] M. Borgerding. *Turning Overlap-Save into a Multiband Mixing, Downsampling Filter Bank*. IEEE Signal Processing Magazine, vol. 158, 2006.
- [Boyd 94] S. Boyd, L. El Ghaoui, E. Feron & V. Balakrishnan. Linear matrix inequalities in system and control theory. SIAM, 1994.
- [Cavassilas 91] J.F. Cavassilas. *Stochastic Matched Filter*. In Int. Conf. on Sonar Signal Processing, pages 194–199. Proceedings of the Institute of Acoustics, 1991.
- [Gonen 99] E. Gonen & J.M. Mendel. *Subspace-Based Direction Finding Methods*. In Madisetti, V.K. and Williams D.B., éditeur, The Digital Signal Processing Handbook, chapitre 62. Chapman & Hall, CRCnetBASE, 1999.
- [Krim 96] H. Krim & M. Viberg. *Two Decades of Array Signal Processing Research : The Parametric Approach*. IEEE Signal Processing Magazine, vol. 13, no. 4, pages 67–94, 1996.
- [Luo 03] Z.Q. Luo. Mathematical programming, volume 97, chapitre Applications of Convex Optimization in Signal Processing and Digital Communication. Springer-Verlag GmbH, 2003.
- [Okuno 04] H.G. Okuno, T. Ogata, K. Komatani & K. Nakadai. *Computational Auditory Scene Analysis and its Application to Robot Audition*. In IEEE Int. Conf. on Informatics Research for Development of Knowledge Society Infrastructure (ICKS'04), pages 73–80, 2004.
- [Schmidt 86] R.O. Schmidt. *Multiple Emitter Location and Signal Parameter Estimation*. IEEE Trans. on Antennas Propagation, vol. 34, no. 3, pages 276–280, 1986.
- [Van Veen 88] B.D. Van Veen & K.M. Buckley. *Beamforming : A Versatile Approach to Spatial Filtering*. IEEE ASSP Magazine, vol. 5, no. 2, pages 4–24, 1988.
- [Wang 85] H. Wang & M. Kaveh. *Coherent Signal-Subspace Processing for the Detection and Estimation of Angles of Arrival of Multiple Wide-Band Sources*. IEEE Trans. on Acoustics, Speech and Signal Processing, vol. 33, no. 4, pages 823–831, 1985.
- [Wang 03] F. Wang, V. Balakrishnan, P.Y. Zhou, J.J. Chen, R. Yang & C. Franck. *Optimal Array Pattern Synthesis Using Semidefinite Programming*. IEEE Trans. on Signal Processing, vol. 51, pages 1172–1183, 2003.
- [Ward 04] D.B. Ward & T.D. Abhayapala. *Range and Bearing Estimation of Wideband Sources using an Orthogonal Beamspace Processing Structure*. In IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, pages 109–112, Montreal, Canada, 2004.
- [Wax 85] M. Wax & T. Kailath. *Detection of Signals by Information Theoretic Criteria*. IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 33, no. 2, pages 387–392, 1985.

Chapitre III

Asservissement Visuel

Sommaire

III.1 Introduction	66
III.2 Contexte scientifique	66
III.2.1 Stratégies d'asservissement visuel eye-in-hand	66
III.2.2 Positionnement de nos contributions	67
III.3 Contribution à l'asservissement visuel multicritère	71
III.3.1 Proposition d'un cadre générique : les systèmes non linéaires incertains « rationnels » soumis à des contraintes inégalités rationnelles	71
III.3.1.1 Vers une « formulation standard » unifiée de l'asservissement 3D ou 2D multicritère	71
III.3.1.2 Principe d'une solution par la théorie de LYAPUNOV	76
III.3.1.3 Inégalités Matricielles et Lemmes importants associés	78
III.3.2 Analyse et Synthèse multicritères par des arguments de stabilité quadratique	80
III.3.2.1 Linéarisation globale	80
III.3.2.2 Exploitation et prolongements mineurs de techniques existantes	80
III.3.2.3 Remarques importantes	84
III.3.2.4 Extensions exigées par le contexte Robotique	85
III.3.2.5 Illustration de l'approche quadratique et de ses extensions	87
III.3.3 Représentations algébro-différentielles et fonctions de LYAPUNOV plus évoluées pour l'analyse multicritère	89
III.3.3.1 Terminologie	91
III.3.3.2 Fonctions de LYAPUNOV candidates	92
III.3.3.3 Vers l'obtention d'un bassin d'attraction multicritère	92
III.3.3.4 Remarques importantes	93
III.3.3.5 Application à un cas d'étude	94
III.4 Approche de la localisation visuelle en tant que le dual de l'asservissement visuel	95
III.4.1 Introduction	95
III.4.2 Aperçu de la formalisation mathématique du problème	96
III.4.3 Illustration sur des cas d'étude	97
III.5 Conception et implémentation d'un environnement de simulation de commandes en robotique	98
III.6 Références	99

III.1 Introduction

La flexibilité offerte par l'introduction d'informations visuelles dans les commandes des robots a été reconnue très tôt, alors que les applications de la robotique étaient essentiellement industrielles. Ainsi, le guidage visuel de tâches de manipulation proposé dans [Shirai 73] permettait pour la première fois d'envisager des environnements de travail moins contrôlés, en acceptant une certaine latitude dans le positionnement des objets assemblés. La stratégie mise en place reposait sur deux processus : la situation[†] à atteindre par l'organe terminal du robot était définie à partir d'une estimation de pose ; un asservissement réalisait ensuite le déplacement effectif de l'organe terminal. La séquence de ces deux étapes était itérée afin d'améliorer la précision de la tâche. Dans ces travaux pionniers, l'information visuelle était acquise le robot à l'arrêt et exploitée à des fins de génération de consigne, alors que la commande était exclusivement rebouclée sur des données proprioceptives, et s'effectuait donc en boucle ouverte par rapport à la vision. Par conséquent, cette stratégie, plus tard appelée *static look-and-move*, s'avérait très sensible aux erreurs de modèles et de calibration.

L'amélioration continue de la technologie des capteurs visuels et leur coût raisonnable ont amené la communauté à envisager de véritables commandes en boucle fermée sur des informations visuelles, dites *commandes référencées vision*, ou *asservissements visuels*. L'introduction de caméras dans les boucles de commande soulève naturellement les problématiques non triviales de modélisation, d'analyse et de synthèse, mais offre la perspective de réaliser des tâches référencées vision robustes et performantes, et d'aborder des problèmes plus ambitieux tels la manipulation rapide d'objets, l'asservissement sur une cible mobile, etc.

III.2 Contexte scientifique

Dans ce manuscrit, nous considérons essentiellement le positionnement référencé vision d'un robot mobile omnidirectionnel ou de l'organe terminal d'un bras manipulateur par rapport à une cible immobile. Nous supposons qu'une seule caméra perspective est exploitée, et en configuration *eye-in-hand*, *i.e.* embarquée sur l'organe commandé. Outre le fait que la faible variation de l'image entre deux instants d'échantillonnage consécutifs simplifie le traitement de l'information visuelle, le problème peut dans un premier temps être abordé en tant que le positionnement de la caméra par rapport à la cible.

III.2.1 Stratégies d'asservissement visuel eye-in-hand

Quatre classes de schémas de commande référencée vision eye-in-hand ont été distingués, selon le type des données sur lesquelles s'effectue le rebouclage [Hutchinson 96]. Les *asservissements 3D* – ou “*position-based*”, ou *asservissements en situation* – définissent la variable contrôlée comme la situation relative caméra-cible, qu'il s'agit d'amener et de maintenir à une référence constante. Tout comme la stratégie naïve présentée plus haut, de tels schémas nécessitent une étape de localisation par reconstruction 3D à chaque instant d'échantillonnage. Une réelle rupture scientifique, dont les bases ont été posées par [Weiss 87], a engendré la famille des *asservissements 2D* – ou “*image-based*”. L'idée est de poser le problème d'asservissement directement dans l'image, en définissant le signal de commande comme une fonctionnelle de la configuration courante de la cible dans l'image et de sa configuration de

†. On utilisera indifféremment les vocables *situation* et *pose* pour les couples (position, orientation).

référence. Il est alors possible de s'affranchir de l'étape de localisation et d'envisager des cadences d'échantillonnage plus soutenues. Sous leur forme la plus élémentaire, ces schémas exploitent des *indices visuels*, tels que des projections de points liés à la cible, des droites, des cercles, etc. [Espiau 92]. Afin d'éviter certains inconvénients des deux stratégies de base tout en bénéficiant de leurs avantages respectifs, des stratégies *hybrides* – ou $2.5D$ – ont ensuite été développées [Malis 99]. Enfin, bien que s'inscrivant dans une veine assez différente, des *asservissements* $\frac{d^2D}{dt}$ [Crétual 01] visent à asservir le champ des vitesses de l'image. De même que les asservissements basés sur des primitives autres que géométriques, cette classe ne sera pas considérée dans la suite.

L'asservissement visuel se situe donc au confluent de la Vision par Ordinateur et de l'Automatique, et questionne en permanence ces disciplines du fait des exigences de la Robotique : sûreté de fonctionnement, faiblesse de la connaissance *a priori* en vue de l'extension des champs d'application, contraintes matérielles et temps réel, complexité des tâches envisagées, etc. La Vision par Ordinateur doit fournir des algorithmes pour : la détection, le suivi inter-images et l'extraction de structures d'intérêt ; la modélisation théorique du capteur, *i.e.* l'analyse du lien entre le monde 3D et l'information 2D perçue ; son calibrage ; éventuellement, la modélisation tridimensionnelle de l'objet, sa localisation visuelle, l'estimation de déplacements, etc.

Du point de vue de l'Automatique, le problème est également très riche. Si on accepte que la tâche référencée vision soit réalisée suffisamment lentement, les effets dynamiques dans le déplacement de la caméra peuvent être négligés. Le modèle des interactions en boucle ouverte est alors dit *cinématique*. Malgré la simplification effectuée, ses équations, qui rendent compte de la manière dont le vecteur des vitesses articulaires du robot ou le torseur cinématique de la caméra s'intègrent en des paramètres de situation caméra-cible et/ou en des vitesses de primitives image, sont fortement non linéaires. Lorsque l'objectif est la performance, on doit définir un modèle *dynamique* [Corke 96][Gangloff 03] permettant de capturer la dynamique des actionneurs et du robot, d'éventuelles flexibilités dans les liaisons, des retards dans l'aquisition et/ou le traitement de l'information visuelle, etc. Quel que soit le degré de prise en compte des dynamiques ou des retards dans le problème, les sources d'incertitudes paramétriques sont multiples. Elles peuvent affecter les paramètres de calibrage du capteur, l'information extraite de l'image, éventuellement le modèle de la cible et/ou du robot, etc.

En outre, l'asservissement visuel exige la satisfaction simultanée de nombreuses exigences. En effet, à la nécessaire propriété de stabilité asymptotique relative à la réalisation du positionnement se rajoutent plusieurs contraintes, y compris lorsqu'aucune garantie de performance n'est requise. Ainsi, le maintien de la cible dans le champ de vue de la caméra est essentiel, sous peine qu'aucun signal ne puisse être transmis au contrôleur en charge de l'élaboration de la commande. De même, la gestion des capacités des actionneurs doit faire l'objet d'une attention particulière. Y compris en contexte cinématique où l'on peut décider *a priori* d'accepter des saturations en ligne, les conséquences d'un tel choix doivent être soigneusement analysées. Enfin, les trajectoires 3D de la caméra et du robot doivent être acceptables. La difficulté de ce point est due à la forte non-linéarité de l'application unissant les paramètres de situation de la caméra et des informations 2D extraites de l'image.

III.2.2 Positionnement de nos contributions

Nos recherches ont pour but de proposer une approche de l'analyse et de la synthèse de schémas de positionnement référencé vision qui prenne en compte la convergence de la caméra mais aussi l'ensemble des contraintes. Nous nommerons cette problématique *analyse/synthèse*

multicritère, ou plus simplement *asservissement visuel multicritère*. Notre démarche est de définir une reformulation du problème qui permette de tirer partie de résultats récents d'automatique avancée pour l'analyse multicritère de nombreux schémas de la littérature, et pour la synthèse multicritère de nouvelles lois de commande.

Ce problème n'est pas nouveau, mais demeure relativement ouvert en raison de sa difficulté. Nous rappelons ci-après quelques schémas d'asservissement visuel classiques, certains résultats de stabilité importants, ainsi que des méthodes de prise en compte des contraintes. Pour une vision plus complète de l'asservissement visuel, le lecteur est invité à se référer aux tutoriaux [Chaumette 06][Chaumette 07].

A Les asservissements visuels 3D Dans les approches 3D du positionnement référencé vision, la variable contrôlée r et sa valeur de référence r^* décrivent les paramétrisations courante et souhaitée de la situation relative caméra-cible. Le modèle en boucle ouverte exprime en tout instant la relation entre r et le torseur cinématique T_c de la caméra, où $r, T_c \in \mathbb{R}^6$ pour une caméra à 6 degrés de liberté (ddl). Il s'écrit

$$\dot{r} = J_{3D}(r)T_c, \quad (\text{III.1})$$

où $J_{3D}(\cdot)$ désigne la *matrice d'interaction 3D*, ou, par abus de langage, “Jacobienne 3D”. Moyennant une paramétrisation de la rotation sans singularité dans la zone de travail, $J_{3D}(\cdot)$ est de rang plein. Souvent, on cherche à mettre en place une commande qui, sous les hypothèses idéales d'une localisation visuelle parfaite de r et d'un robot obéissant au modèle simplifié (III.1), conduit à une boucle fermée linéarisée et découplée. Considérant qu'en pratique, du fait d'inévitables erreurs de localisation, seule une approximation \hat{r} de r est disponible, cette commande s'écrit

$$T_c = -\lambda[J_{3D}(\hat{r})]^{-1}(\hat{r} - r^*), \quad \lambda > 0. \quad (\text{III.2})$$

Sous les hypothèses idéales ci-dessus mentionnées, la paramétrisation de la position relative caméra-cible converge donc vers sa valeur de référence selon une dynamique exponentielle décroissante. La rotation relative caméra-cible décrit une géodésique vers sa valeur de référence selon la même dynamique. On peut alors isoler un point lié à la cible dont la trajectoire est rectiligne dans le plan image et dans le repère mobile lié à la caméra, voire dans le repère lié au monde.

Hélas, deux considérations réduisent fondamentalement l'utilité de cette propriété théorique de stabilité globale. D'une part, il est très difficile d'assurer *a priori* la visibilité de la cible durant la durée de la tâche [Martinet 98]. D'autre part, du fait que l'estimation \hat{r} de la pose r peut être très sévèrement dégradée en présence de bruit d'extraction des indices visuels dans l'image ou d'incertitudes dans les paramètres intrinsèques de la caméra, la stabilité asymptotique globale est remise en cause. Comme solutions garantissant la contrainte de visibilité dans un schéma d'asservissement 3D, citons par exemple [Thuilot 02]. Plus récemment, [Zanne 04] exploite un contrôleur 3D globalement asymptotiquement stable et un planificateur de chemins visibles, et assure formellement la visibilité de la cible par la bornitude des erreurs de suivi en dépit d'incertitudes de calibration.

B Les asservissements 2D Les stratégies d'asservissement visuel 2D visent à réguler un vecteur d'indices visuels s à une référence constante s^* . Le modèle en boucle ouverte s'écrit

$$\dot{s} = J_{2D}(s, z)T_c. \quad (\text{III.3})$$

La matrice d'interaction (2D), ou, par abus de langage, « Jacobienne image », $J_{2D}(\cdot, \cdot)$ dépend de la configuration courante s des indices visuels, mais aussi de l'information de profondeur z perdue par projection perspective. Pour une caméra à 6ddl, et lorsque s caractérise les projections de points rigidement liés à la cible, ceux-ci sont généralement non alignés et en nombre $M \geq 4$. Ceci permet d'assurer qu'à une valeur de s correspond une unique situation relative caméra-cible, et que la matrice d'interaction $J_{2D}(\cdot, \cdot) \in \mathbb{R}^{2M \times 6}$ est de rang plein 6. La commande permettant d'approcher un comportement linéarisé et découplé en boucle fermée s'écrit

$$T_c = -\lambda \left[\widehat{J_{2D}(s, z)} \right]^+ (s - s^*), \quad \lambda > 0, \quad (\text{III.4})$$

où $(\cdot)^+$ est l'opérateur de pseudo-inversion matricielle et $\widehat{J_{2D}(s, z)}$ désigne une approximation de $J_{2D}(s, z)$. Celle-ci peut soit être fondée sur l'expression analytique de $J_{2D}(\cdot, \cdot)$ et faire intervenir une approximation plus ou moins grossière \hat{z} du vecteur de paramètres de profondeur z – *e.g.* en estimant z en ligne, ou en assimilant \hat{z} à la valeur z^* de z lorsque la tâche est réalisée –, soit être obtenue par une méthode *ad hoc*. Une condition suffisante immédiate de stabilité globale de la boucle fermée s'exprime comme la définie positivité de la fonction matricielle suivante, à valeurs dans $\mathbb{R}^{2M \times 2M}$:

$$J_{2D}(s, z) \left[\widehat{J_{2D}(s, z)} \right]^+ > \mathbb{O}. \quad (\text{III.5})$$

S'agissant d'une matrice carrée d'au moins 8 lignes/colonnes dont le rang n'excède pas 6, cette condition n'est jamais satisfaite. Pour conclure, on peut recourir au formalisme des fonctions de tâche [Samson 91][Espiau 92]. Une fonction de tâche e est une fonction de sortie de dimension égale au nombre de degrés de mobilité de la caméra, de la forme

$$e = C(s - s^*), \quad (\text{III.6})$$

où C désigne une matrice de combinaison ($C \in \mathbb{R}^{6 \times 2M}$ pour une caméra à 6ddl). Le but est d'amener cette sortie critique e à 0 et de l'y maintenir. Pour que le problème soit bien posé, il faut assurer un lien régulier entre l'espace de mobilité de la caméra et l'espace de la fonction de tâche, au sens où une petite variation autour de la trajectoire de référence de la caméra doit induire une petite variation de e autour de 0 et vice-versa. La propriété de ρ -admissibilité [Samson 91] permet de répondre à cette exigence. Souvent, le conditionnement du problème est tel que cette propriété se ramène à l'inversibilité du Jacobien de la tâche au voisinage de la trajectoire de référence. On montre ainsi [Chaumette 06] que pour la matrice de combinaison – possiblement variante dans le temps – $C = \left[\widehat{J_{2D}(s, z)} \right]^+$, la commande (III.4) conduit à la stabilité asymptotique locale de l'état d'équilibre $e^* = 0$ si

$$\left[\widehat{J_{2D}(s, z)} \right]^+ J_{2D}(s, z) > \mathbb{O}, \quad (\text{III.7})$$

et que $(s - s^*)$ ne peut pas appartenir au noyau de C dans un voisinage de la position de référence correspondant à la réalisation de la tâche. Néanmoins, rien n'empêche que ceci se produise en des positions s suffisamment éloignées de s^* . La caméra converge alors vers un *minimum local*, pour lequel la tâche n'est pas réalisée. Plus largement, pour de telles commandes de type « Jacobienne inverse », un échec est également obtenu lorsque la matrice d'interaction $J_{2D}(s, z)$ admet des valeurs singulières infinies ou lorsque $\widehat{J_{2D}(s, z)}$ est singulière.

C'est le cas du "Chaumette conundrum" [Chaumette 98], où l'obtention d'une trajectoire simple dans l'image conduit à un mouvement 3D inacceptable.

Sur la base de ces considérations générales, de nombreux travaux ont été développés afin d'analyse plus en profondeur les asservissements 2D ou de pallier certaines de leurs limitations. Des analyses théoriques de stabilité rapport à des erreurs de calibration à des incertitudes sur la profondeur de la cible ont été développées dans [Deng 02][Malis 03]. Une approche par commutation entre différents potentiels [Hashimoto 00] et l'utilisation de coordonnées 2D étendues [Schramm 04] permettent l'évitement de minima locaux. Plusieurs approches ont concerné la synthèse d'une stratégie assurant la visibilité de la cible, *e.g.* [Mezouar 03][Chesi 04][Cowan 02].

Le fait d'opter pour une commande assurant une dynamique découpée et linéaire en boucle fermée soulève le problème fondamental de la sélection des indices visuels. Le but peut être de mieux maîtriser les mouvements de rotation, à l'origine des limitations exposées plus haut. La sélection d'indices visuels invariants à certains types de mouvements permet de bonnes propriétés de linéarisation et de découplage en boucle fermée. Le choix d'indices visuels de types « moments », initialement proposé dans [Chaumette 04] s'inscrit dans cette veine, et a fait l'objet de multiples extensions, *e.g.* [Bourquardez 06].

L'objectif d'un meilleur contrôle de la rotation de la caméra est à la base des approches 2.5D, dont certaines ne nécessitent pas de modèle de la cible et admettent des propriétés de stabilité clairement établies [Malis 02].

Pour conclure ce bref panorama, il convient de préciser que de nombreuses autres techniques de commande ont été exploitées pour l'asservissement visuel : commandes LQ et LQG [Hashimoto 93][Papanikolopoulos 93], H_∞ à séquencement de gain [Sznaier 00], commande prédictive non linéaire [Sauvée 06][Allibert 06], LMI [Tarbouriech 00][Tarbouriech 05],...

C Organisation du chapitre Nous avons reformulé le problème de l'asservissement visuel multicritère dans un cadre scientifique original, significativement « parallèle » aux travaux existants. Celui-ci est décrit avec un certain niveau de détail dans le §III.3, dans le but de mieux argumenter les raisons de son choix et de souligner en quoi il questionne certaines théoriques d'automatique. Afin de limiter la technicité du manuscrit, nous n'entrions pas dans le détail des méthodes d'analyse et de synthèse que nous avons développées. L'accent est plutôt mis sur la manière dont les concepts généraux se déclinent en approches constructives. Des compléments techniques sont donnés seulement s'ils sont nécessaires à la compréhension du raisonnement ou s'ils sont le siège de propriétés importantes. Notre approche a ouvert de nouveaux questionnements, tel que le traitement du problème de la localisation visuelle en tant que le dual – au sens de la dualité observation-commande – de l'asservissement visuel, qui fait l'objet du §III.4. Enfin, nous présentons brièvement un environnement de simulation dans le §III.5.

Nos prospectives feront l'objet des §V.1.2 et §V.2.1.

Pour mémoire, les projets de recherche associés figurent dans le §II.3.5 page 63.

III.3 Contribution à l'asservissement visuel multicritère

Rappelons que la tâche « canonique » considérée consiste en le positionnement référencé vision d'une caméra perspective par rapport à une cible fixe munie d'indices visuels ponctuels en contexte cinématique. Un modèle possiblement incertain de ces divers éléments est supposé disponible *a priori*.

III.3.1 Proposition d'un cadre générique : les systèmes non linéaires incertains « rationnels » soumis à des contraintes inégalités rationnelles

III.3.1.1 Vers une « formulation standard » unifiée de l'asservissement 3D ou 2D multicritère

A Introduction Afin d'instaurer un « dialogue » entre l'asservissement visuel et l'automatique, nous proposons une *formulation standard* – dans la veine de [Zhou 95] – d'un problème de commande référencée vision étendue à la prise en compte éventuelle de contraintes instantanées. L'élément central est un modèle d'état du système commandé, de vecteur d'état x et de vecteur de commande u . Le vecteur de sortie y de ce système est l'entrée d'un contrôleur en boucle fermée sur u . Ce contrôleur est *a priori* dynamique, de vecteur d'état x_c , de sorte que le système bouclé – autonome – admet pour vecteur d'état $\tilde{x} \triangleq (x' \ x_c')$. Le premier but de la commande est de garantir la *stabilité du vecteur d'état d'équilibre* $\tilde{x}^* = 0$ de la boucle fermée. Un vecteur d'entrées irréductibles w et un vecteur de sorties régulées z permettent éventuellement l'optimisation d'objectifs de performance exprimés comme une fonction d'une norme de z ou d'une norme induite du transfert entre w et z . La satisfaction de contraintes instantanées s'exprime comme la *bornitude de variables additionnelles* ζ_j . Les incertitudes du système commandé constituent le vecteur χ . Pour simplifier, celles-ci sont possiblement variantes dans le temps mais sans mémoire, de sorte que toutes les variables du système asservi, y compris les variables additionnelles, sont des fonctions statiques de (\tilde{x}, χ) .

Il est donc nécessaire d'exhiber un modèle en boucle ouverte pour l'asservissement visuel qui satisfasse la définition axiomatique d'un vecteur d'état. Il nous a semblé naturel de poser $x \triangleq (t' \ r')$, où $t \in \mathbb{R}^3$ et $r \in \mathbb{R}^3$ paramétrisent respectivement la position et l'orientation relative entre le repère lié à la caméra à l'instant courant et la situation de référence qu'il doit atteindre, laquelle est rigidement liée à la cible. En effet, dans le contexte cinématique considéré où la cible est immobile et en l'absence d'incertitudes, toutes les variables du système en boucle ouverte sont des fonctions instantanées du vecteur x précédemment défini, qui est en outre de dimension minimale. Le vecteur de commande u est naturellement le torseur cinématique de la caméra. L'équation d'état qui unit u et \dot{x} s'obtient à partir des lois de la cinématique des solides rigides. Plusieurs choix de repères sont possibles pour l'expression des paramètres t et r . On requiert cependant que le jeu des trois paramètres de rotation qui constituent r admette une singularité de représentation – *i.e.* une indétermination – seulement pour des situations de la caméra exclues d'avance du fait de l'existence des contraintes. Enfin, le mouvement de la caméra est supposé omnidirectionnel, de sorte que les vecteurs d'état et de commande admettent la même dimension.

Pour un asservissement 3D, le signal y sur lequel est réalisé le rebouclage est simplement le vecteur x . Pour une stratégie 2D, y est défini comme la différence $y = s - s^*$ entre les valeurs courante s et souhaitée s^* de la paramétrisation des indices visuels dans l'image. L'équation de sortie du modèle en boucle ouverte, qui unit x et y , fait alors intervenir les équations de projection perspective du capteur, ses paramètres de calibrage ainsi que le modèle de la cible.

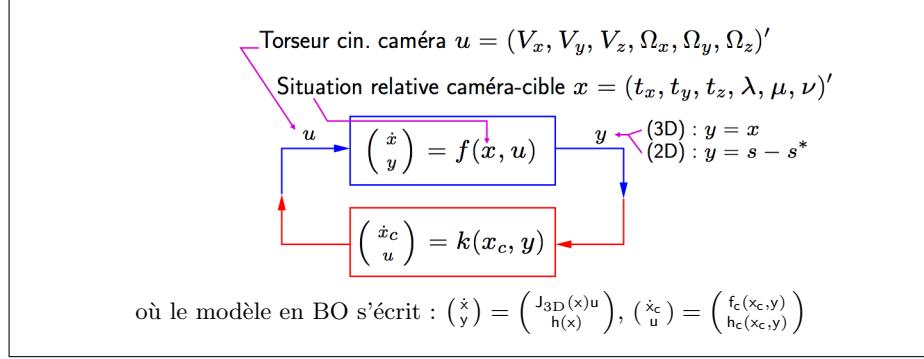


FIGURE III.1 – Unification des stratégies d'asservissement 3D et 2D.

Dans la suite du manuscrit, aucun objectif de performance ne sera spécifié. La Figure III.1 présente une unification des stratégies d'asservissement 3D et 2D en l'absence d'incertitudes. Le positionnement de la caméra par rapport à la cible est réalisé dès lors que le point d'équilibre $\tilde{x}^* = 0$ de la boucle fermée – autonome – est asymptotiquement stable[†]. La satisfaction des contraintes est équivalente à la bornitude de variables additionnelles scalaires $\zeta_j = \zeta_j(\tilde{x})$ par des intervalles prédéfinis $[\underline{\zeta}_j, \bar{\zeta}_j]$, $j = 1, \dots, J$. Sans perte de généralité, le vecteur de sortie y et les variables ζ_j , $j = 1, \dots, J$, sont définis de telle sorte que $y = 0$ et $\zeta_j = 0$ lorsque $x = 0$ et $\tilde{x} = \tilde{x}^* = 0$, respectivement.

B Équations du modèle dans le cas d'indices visuels ponctuels Nous proposons d'entrer brièvement dans le détail des équations du modèle en boucle ouverte dans le cas d'indice visuels ponctuels, afin de donner une idée de leur complexité. Pour cela, il convient de définir au préalable quelques notations, résumées sur la Figure III.2.

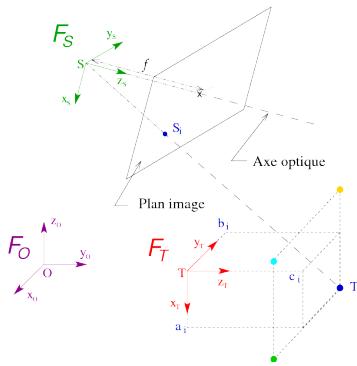


FIGURE III.2 – Notations.

s'écrivent $\bar{x}_i^* = f \frac{a_i}{c_i}$ et $\bar{y}_i^* = f \frac{b_i}{c_i}$.

Le vecteur de commande u est donc constitué de la superposition des vecteurs vitesses de translation $(v_x \ v_y \ v_z)'$ et de rotation instantanée $(\omega_x \ \omega_y \ \omega_z)'$ de F_S par rapport à F_O , exprimés

†. En toute rigueur, l'attractivité de $\tilde{x}^* = 0$ est suffisante à la convergence de la caméra. La stabilité asymptotique permet d'écartier des cas d'étude où la caméra effectue des mouvements d'aller-retour passant par l'infini, tels le *Chaumette conundrum*[Chaumette 98].

TABLE III.1 – Modèle d'état en boucle ouverte.

$$\begin{pmatrix} \dot{t}_x \\ \dot{t}_y \\ \dot{t}_z \\ \dot{\lambda} \\ \dot{\mu} \\ \dot{\nu} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 & -t_z & t_y \\ 0 & -1 & 0 & t_z & 0 & -t_x \\ 0 & 0 & -1 & -t_y & t_x & 0 \\ 0 & 0 & 0 & -1 & -\sin \lambda \tan \mu & \cos \lambda \tan \mu \\ 0 & 0 & 0 & 0 & -\cos \lambda & -\sin \lambda \\ 0 & 0 & 0 & 0 & \frac{\sin \lambda}{\cos \mu} & \frac{-\cos \lambda}{\cos \mu} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}, \quad \begin{aligned} \mathbf{y} &= \mathbf{x} && (\text{asservissement 3D}) \\ \mathbf{y} &= (\bar{x}_1 - \bar{x}_1^* \ \bar{y}_1 - \bar{y}_1^* \ \dots \ \bar{x}_M - \bar{x}_M^* \ \bar{y}_M - \bar{y}_M^*)' && (\text{asservissement 2D}) \end{aligned} \quad (III.8)$$

avec $\bar{x}_i - \bar{x}_i^* = \frac{f}{c_i z_i} (c_i x_i - a_i z_i)$, $\bar{y}_i - \bar{y}_i^* = \frac{f}{c_i z_i} (c_i y_i - b_i z_i)$
où

$$\begin{aligned} z_i &= t_z + a_i(-\cos \lambda \sin \mu \cos \nu + \sin \lambda \sin \nu) \\ &\quad + b_i(\cos \lambda \sin \mu \sin \nu + \sin \lambda \cos \nu) + c_i \cos \lambda \cos \mu \\ c_i x_i - a_i z_i &= c_i t_x - a_i t_z - a_i^2(-\cos \lambda \sin \mu \cos \nu + \sin \lambda \sin \nu) + c_i^2 \sin \mu - b_i c_i \cos \mu \sin \nu \\ &\quad - a_i b_i(\cos \lambda \sin \mu \sin \nu + \sin \lambda \cos \nu) + a_i c_i \cos \mu(\cos \nu - \cos \lambda) \\ c_i y_i - b_i z_i &= c_i t_y - b_i t_z - b_i^2(\cos \lambda \sin \mu \sin \nu + \sin \lambda \cos \nu) - c_i^2 \sin \lambda \cos \mu \\ &\quad - a_i b_i(-\cos \lambda \sin \mu \cos \nu + \sin \lambda \sin \nu) + a_i c_i(\sin \lambda \sin \mu \cos \nu + \cos \lambda \sin \nu) \\ &\quad + b_i c_i(\cos \lambda(\cos \nu - \cos \mu) - \sin \lambda \sin \mu \sin \nu). \end{aligned}$$

dans F_S . Les deux sous-vecteurs \mathbf{t} et \mathbf{r} de \mathbf{x} sont respectivement constitués des coordonnées $(\vec{ST})_{(F_S)} = (t_x \ t_y \ t_z)'$ du vecteur \vec{ST} dans F_S , et des angles de Bryant $(\lambda \ \mu \ \nu)'$ qui transforment $(\vec{x}_S, \vec{y}_S, \vec{z}_S)$ en $(\vec{x}_T, \vec{y}_T, \vec{z}_T)$. Ainsi, \mathbf{r} est monovalué si et seulement si $\mu \neq \pm \frac{\pi}{2}$ – *i.e.* ssi $\vec{z}_S \not\perp \vec{z}_T$ – [Renaud 96], ce qui peut être aisément garanti par l'adjonction d'une contrainte de visibilité au problème. Dans le cas d'un asservissement 2D, le vecteur de sortie est bien sûr défini comme $\mathbf{y} = \mathbf{s} - \mathbf{s}^* = (\bar{x}_1 - \bar{x}_1^* \ \bar{y}_1 - \bar{y}_1^* \ \dots \ \bar{x}_M - \bar{x}_M^* \ \bar{y}_M - \bar{y}_M^*)'$. On obtient alors la représentation d'état reportée Table III.1 [JSI-2-AV-1], où on reconnaît dans l'équation d'état une « Jacobienne 3D »† inversible pour tout $\mu \neq \pm \frac{\pi}{2}$ (modulo 2π).

C Potentialités du cadre théorique proposé Le cadre théorique précédent appelle quelques remarques importantes. En premier lieu, des incertitudes peuvent affecter les modèles d'état des interactions en boucle ouverte et/ou du contrôleur. Ensuite, il convient de noter que le but principal de la commande est la stabilité/stabilisation asymptotique du vecteur d'état du système bouclé quelles que soient les réalisations des incertitudes. Ceci implique la stabilité/stabilisation de la situation relative caméra-cible \mathbf{y} compris si un asservissement 2D est considéré. Cette position du problème permet de ne pas conclure hâtivement que le positionnement est effectivement réalisé alors que la caméra a convergé vers un minimum local. En fait, la convergence de \mathbf{s} vers \mathbf{s}^* est une conséquence de l'attractivité de l'état d'équilibre $\mathbf{x}^* = 0$.

Une potentialité intéressante de l'approche est la possibilité de considérer tout type de rebouclage référencé vision, qu'il soit 2D ou 3D, statique ou dynamique. Des commandes 2D « séquencées sur du 3D », *i.e.* admettant pour entrée \mathbf{y} un vecteur de primitives 2D mais dont les coefficients dépendent d'informations 3D, peuvent également être appréhendées. Il suffit pour cela de remplacer la représentation d'état du contrôleur de la Figure III.1 par

†. Qui, en toute rigueur, n'est pas une matrice Jacobienne car en général le vecteur vitesse de rotation instantanée n'est pas la dérivée d'un jeu de paramètres...

$$\dot{x}_c = f_c(x_c, x, y), u = h_c(x_c, x, y)^\dagger.$$

Enfin, quelques exemples simples illustrent la générnicité de la méthode de gestion des contraintes. Ainsi, il est possible de restreindre la projection d'indices visuels ponctuels aux limites du plan image de la caméra en définissant $\zeta_j = \bar{x}_j - \bar{x}_j^*$ ou $\zeta_j = \bar{y}_j - \bar{y}_j^*$, y compris pour des asservissements 3D. Des saturations d'actionneurs peuvent être exclues en bornant des composantes de u ou des normes de sous-vecteurs extraits de u^\ddagger . La garantie de contraintes 3D est permise y compris pour des asservissements 2D en bornant des distances $\zeta_j = d_{3D_j}$. Enfin, le fait de placer des limites sur la commande u ou sur les écarts $\zeta_j = \bar{x}_j - \bar{x}_j^*$ ou $\zeta_j = \bar{y}_j - \bar{y}_j^*$ permet d'éviter \S , lors de la mise en place d'une commande 2D « Jacobienne inverse séquencée » $u = -\lambda[J_{2D}(s, z)]^+(s - s^*)$ où $z = (z_1 \dots z_M)'$ et $J_{2D}(\cdot, \cdot)$ désigne la matrice d'interaction définie par $s = J_{2D}(s, z)u$ – le passage par des singularités de $J_{2D}(\cdot, \cdot)$ ou de $[J_{2D}(\cdot, \cdot)]^+$. Plus largement, l'introduction de variables additionnelles prévient les cas pathologiques où l'équation d'état de la boucle fermée est bien définie alors que le problème est mal posé du fait de valeurs infinies prises par des signaux internes à l'interconnexion.

D Restriction aux systèmes non linéaires rationnels Bien qu'un cadre théorique relativement générique ait été défini, le problème de l'analyse ou de la synthèse multicritère demeure entier. La difficulté majeure réside dans le caractère fortement non linéaire et incertain des équations du système en boucle ouverte, et dans les commandes elles-mêmes non linéaires souvent rencontrées dans la littérature. Afin de passer du concept à des résultats exploitables, nous avons tenté d'isoler une famille de systèmes non linéaires qui soit à la fois pertinente du point de vue du problème de robotique considéré, et pour laquelle il existe une littérature minimale en analyse et commande multicritère. Notre choix s'est porté sur la classe des systèmes non linéaires dits *rationnels*.

Définition III.1 (Système rationnel [El Ghaoui 96]) *Un système est dit rationnel s'il est défini par – les notations précédemment introduites étant conservées –*

$$\begin{pmatrix} \dot{x} \\ y \end{pmatrix} = \begin{pmatrix} A(x, \chi) & B(x, \chi) \\ C(x, \chi) & D(x, \chi) \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}, \quad (\text{III.9})$$

où $A(\cdot, \cdot)$, $B(\cdot, \cdot)$, $C(\cdot, \cdot)$, $D(\cdot, \cdot)$ désignent des fonctions matricielles rationnelles de (x, χ) bien définies – i.e. dont aucune composante n'admet de dénominateur nul – quels que soient $(x, \chi) \in \mathcal{X} \times \mathcal{X}_\chi$, où $\mathcal{X} \subset \mathbb{R}^{n_x}$ et $\mathcal{X}_\chi \subset \mathbb{R}^{n_\chi}$ sont des polytopes contenant l'origine 0.

Supposons que le modèle des interactions en boucle ouverte pour l'asservissement visuel s'écrit sous la forme (III.9). Son interconnexion avec un contrôleur rationnel

$$\begin{pmatrix} \dot{x}_c \\ u \end{pmatrix} = \begin{pmatrix} K_c(x, x_c) & K_{cy}(x, x_c) \\ K_u(x, x_c) & K_{uy}(x, x_c) \end{pmatrix} \begin{pmatrix} x_c \\ y \end{pmatrix}, \quad (\text{III.10})$$

†. Le contexte proposé permet de poser le problème d'analyse ou de synthèse multicritère de contrôleurs statiques ou dynamiques 2.5D généraux, en constituant y avec des informations 3D et 2D. Cette possibilité ne sera pas envisagée car la sous-famille des systèmes non linéaires rationnels introduite plus loin ne permet pas de recouvrir trivialement les asservissements 2.5D de la littérature, e.g. [Malis 99].

‡. En contexte cinématique, il est préférable d'*autoriser* – en maîtrisant leur effet – les saturations que de les *exclure* comme nous le proposons dans ce manuscrit.

§. Plus exactement, ceci permet le passage par des singularités, mais selon des directions *compatibles*.

où \mathbf{x}_c vit dans un polytope $\mathcal{X}_c \subset \mathbb{R}^{n_c}$ donné contenant 0, conduit à la boucle fermée rationnelle autonome

$$\dot{\tilde{\mathbf{x}}} = \tilde{\mathbf{A}}(\tilde{\mathbf{x}}, \chi)\tilde{\mathbf{x}}, \quad \tilde{\mathbf{x}} = (\mathbf{x}' \ \mathbf{x}_c)' \in \tilde{\mathcal{X}} \triangleq \mathcal{X} \times \mathcal{X}_c \subset \mathbb{R}^{\tilde{n}}, \quad \tilde{n} = n_{\mathbf{x}} + n_c, \quad \chi \in \mathcal{X}_{\chi} \subset \mathbb{R}^{n_{\chi}}, \quad (\text{III.11})$$

pour laquelle on suppose $\tilde{\mathbf{A}}(., .)$ bien définie sur $\tilde{\mathcal{X}} \times \mathcal{X}_{\chi}$, avec $0 \in \tilde{\mathcal{X}} \times \mathcal{X}_{\chi}$.

On considère en outre la bornitude de variables additionnelles

$$\forall j \in 1, \dots, J, \quad \forall \chi \in \mathcal{X}_{\chi}, \quad \zeta_j = Z_j'(\tilde{\mathbf{x}}, \chi)\tilde{\mathbf{x}} \in [\underline{\zeta}_j, \bar{\zeta}_j], \quad (\text{III.12})$$

où chaque $Z_j(., .)$ désigne une fonction rationnelle bien définie sur $\tilde{\mathcal{X}} \times \mathcal{X}_{\chi}$ à valeurs dans $\mathbb{R}^{\tilde{n}}$.

Les méthodes que nous exploitons pour l'analyse ou la commande d'un système rationnel incertain supposent à un moment donné son *immersion* – ou « plongée », ou *embedding* – dans une famille plus vaste, *e.g.* un système linéaire incertain dont il est supposé constituer une réalisation. De telles approches, dites *linéaristes*, s'apparentent donc à la commande robuste. L'obtention plus aisée de conclusions s'effectue naturellement au prix de *pessimisme*, ou « conservatisme ». La restriction aux systèmes rationnels limite toutefois la famille dans laquelle est effectuée l'immersion, et par conséquent le pessimisme associé. En contrepartie, le traitement simultané de toute une classe d'équivalence de systèmes incluant le système original occulte des propriétés structurelles propres à celui-ci.

Soulignons toutefois qu'à notre connaissance, la littérature ne propose aucune méthodologie complète pour l'analyse de stabilité ou la stabilisation d'un système non linéaire rationnel soumis à des contraintes elles-mêmes rationnelles (III.12).

E Pertinence des systèmes rationnels pour l'asservissement visuel Considérons le modèle en boucle ouverte (III.8). Comme les paramètres d'orientation apparaissent dans ses équations en tant qu'arguments de fonctions trigonométriques, il peut être aisément réécrit sous une forme rationnelle équivalente. Ainsi, sous l'hypothèse $\lambda \in [-\frac{\pi}{2}; +\frac{\pi}{2}]$, $\mu \in]-\frac{\pi}{2}; +\frac{\pi}{2}[$ et $\nu \in [-\pi; +\pi]$, le changement bijectif de vecteur d'état

$$\mathbf{x} \triangleq (t_x \ t_y \ t_z \ \tan(\frac{\lambda}{2}) \ \tan(\frac{\mu}{2}) \ \tan(\frac{\nu}{4}))' \quad (\text{III.13})$$

conduit à une boucle ouverte rationnelle semblable à (III.9), où $\mathbf{A}(., .) = \mathbb{O}$ et $\mathbf{D}(., .) = \mathbb{O}$.

Le rebouclage (III.10) décrit un contrôleur dynamique non linéaire séquencé, dont les paramètres sont mis à jour en ligne avec des expressions rationnelles de la situation relative caméra-cible \mathbf{x} et/ou du vecteur d'état du contrôleur \mathbf{x}_c . Cette stratégie peut se spécialiser en des commandes moins sophistiquées, telles des retours dynamiques linéaires 3D ou 2D, et recouvre de nombreuses stratégies de la littérature. Ainsi, les schémas classiques « Jacobienne inverse 3D » $\mathbf{u} = -\lambda \mathbf{J}_{3D}^{-1}(0)\mathbf{x}$ et $\mathbf{u} = -\lambda \mathbf{J}_{3D}^{-1}(\mathbf{x})\mathbf{x}$, avec $\mathbf{J}_{3D}(.)$ tel que $\dot{\mathbf{x}} = \mathbf{J}_{3D}(\mathbf{x})\mathbf{u}$, correspondent respectivement à un retour d'état statique linéaire $\mathbf{u} = \mathbf{K}\mathbf{x}$ et à un retour d'état statique non linéaire rationnel $\mathbf{u} = \mathbf{K}(\mathbf{x})\mathbf{x}$. De même, les contrôleurs « Jacobienne inverse 2D » $\mathbf{u} = -\lambda [\mathbf{J}_{2D}(\mathbf{s}^*, \mathbf{z}^*)]^+(\mathbf{s} - \mathbf{s}^*)$ et $\mathbf{u} = -\lambda [\mathbf{J}_{2D}(\mathbf{s}, \mathbf{z})]^+(\mathbf{s} - \mathbf{s}^*)$, avec $\mathbf{J}_{2D}(., .)$ la matrice d'interaction définie par $\dot{\mathbf{s}} = \mathbf{J}_{2D}(\mathbf{s}, \mathbf{z})\mathbf{u}$, peuvent être considérés. En effet, $\mathbf{J}_{2D}(., .)$ est aussi une fraction rationnelle de \mathbf{x} , de sorte que ces commandes 2D correspondent respectivement à un retour de sortie statique linéaire $\mathbf{u} = \mathbf{K}\mathbf{y}$ et à un retour de sortie statique non linéaire séquencé $\mathbf{u} = \mathbf{K}(\mathbf{x})\mathbf{y}$.

On montre également que lorsque le modèle en boucle ouverte (III.8) est transformé en (III.9) via le changement de variable (III.13), la bornitude des variables additionnelles

relatives aux diverses contraintes évoquées au §III.3.1.1–C se réécrit sous la forme rationnelle (III.12) [†].

Comme indiqué précédemment, les incertitudes doivent être sans mémoire et affecter le modèle de la boucle ouverte de manière rationnelle. Naturellement, on peut imaginer un modèle faisant intervenir tous les paramètres intrinsèques de la caméra avec leurs incertitudes associées, des incertitudes sur le modèle de la cible, etc. Il est tout aussi possible d'envisager l'étude de la fragilité d'un contrôleur vis à vis d'incertitudes du même type, en complétant de manière adéquate l'équation (III.10). Une extension à la prise en compte d'effets dynamiques dans le mouvement de la caméra paraît également concevable.

Pour conclure, évoquons quelques pistes lorsque le modèle des interactions en boucle ouverte et/ou du contrôleur référencé vision comporte des termes non linéaires qui ne sont pas immédiatement « rationnalisables ». Une première stratégie d'immersion consiste à « remplacer formellement » chacun de ces termes par un paramètre incertain variant, sans mémoire, et vivant dans un polytope. Cependant, afin de limiter le conservatisme des conclusions, il est préférable d'approximer chaque non-linéarité par une fraction rationnelle, puis plonger la différence ou le quotient des deux expressions dans de l'incertitude.

III.3.1.2 Principe d'une solution par la théorie de Lyapunov

La théorie de LYAPUNOV et l'optimisation LMI (*Linear Matrix Inequalities*, ou *Inégalités Matricielles Linéaires*) constituent le socle théorique de l'analyse de la boucle fermée (III.11) ou de la commande du système en boucle ouverte (III.9) par le contrôleur visuel (III.10), en présence des contraintes (III.12).

A Définition d'une fonction de Lyapunov L'analyse ou la synthèse reposent sur la définition d'une fonction $V(., .) : \tilde{\mathcal{X}} \times \mathcal{X}_\chi \rightarrow \mathbb{R}$ telle que l'ensemble $\tilde{\mathcal{E}} \triangleq \{\tilde{x} : V(\tilde{x}, \chi) \leq 1, \forall \chi \in \mathcal{X}_\chi\}$ soit un bassin d'attraction multicritère au sens suivant :

Définition III.2 (Bassin d'attraction multicritère) *Un ensemble $\tilde{\mathcal{E}}$ est un bassin d'attraction multicritère si toute trajectoire du système asservi (III.11) dont la condition initiale se situe dans $\tilde{\mathcal{E}}$ converge vers le vecteur d'état d'équilibre $\tilde{x}^* = 0$ et est entièrement contenue dans le sous-espace admissible défini par les contraintes (III.12).*

Les degrés de liberté de $V(., .)$ constituent une matrice P , et $V(., .) = V_{\mathsf{P}}(., .)$ est recherchée dans une classe spécifiée *a priori* de façon à satisfaire les trois propriétés suivantes :

1. $\tilde{x}^* = 0$ est localement asymptotiquement stable, si $V_{\mathsf{P}}(., .)$ est une fonction de LYAPUNOV sur $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$, i.e. $V_{\mathsf{P}}(., .)$ est continûment différentiable et vérifie

$$\forall (\tilde{x}, \chi) \in (\tilde{\mathcal{X}} \setminus \{0\}) \times \mathcal{X}_\chi, V_{\mathsf{P}}(\tilde{x}, \chi) > 0; \quad \forall \chi \in \mathcal{X}_\chi, V_{\mathsf{P}}(0, \chi) = 0; \quad (\text{III.14})$$

$$\forall (\tilde{x}, \chi) \in (\tilde{\mathcal{X}} \setminus \{0\}) \times \mathcal{X}_\chi, \dot{V}_{\mathsf{P}}(\tilde{x}, \chi) < 0; \quad \forall \chi \in \mathcal{X}_\chi, \dot{V}_{\mathsf{P}}(0, \chi) = 0. \quad (\text{III.15})$$

2. $\tilde{\mathcal{E}}$ est un bassin d'attraction de $\tilde{x}^* = 0$ pour le problème non contraint, si

$$\forall (\tilde{x}, \chi) : V_{\mathsf{P}}(\tilde{x}, \chi) \leq 1, \tilde{x} \in \tilde{\mathcal{X}}. \quad (\text{III.16})$$

[†]. Dans le cas – très courant – où le contrôleur n'exhibe pas de dépendance rationnelle en x_c , $\tilde{\mathbf{A}}(., .)$ et $\mathbf{Z}_j(., .)$, $j = 1, \dots, J$, sont rationnelles en (x, χ) , de sorte qu'on exige seulement qu'elles soient bien définies sur $\mathcal{X} \times \mathcal{X}_\chi$.

3. $\tilde{\mathcal{E}}$ est un bassin d'attraction multicritère de $\tilde{x}^* = 0$, i.e. la bornitude (III.12) des variables additionnelles est assurée, si

$$\forall j \in 1, \dots, J, \forall (\tilde{x}, \chi) : V_P(\tilde{x}, \chi) \leq 1, \underline{\zeta}_j \leq Z_j'(\tilde{x}, \chi)\tilde{x} \leq \bar{\zeta}_j. \quad (\text{III.17})$$

La Figure III.3 illustre les conditions ci-dessus lorsque $V(\cdot)$ et $Z_j(\cdot)$ ne dépendent pas de χ . Sur ce schéma, le sous-espace d'état admissible du point de vue des contraintes s'écrit $\tilde{\mathcal{A}} = \cap_{j=1}^J \tilde{\mathcal{A}}_j$, avec $\tilde{\mathcal{A}}_j \triangleq \{\tilde{x} : Z_j'(\tilde{x})\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]\}$. Quelques iso-coûts de la fonction de LYAPUNOV $V(\cdot)$ sont représentés, ainsi que le bassin d'attraction multicritère $\tilde{\mathcal{E}} \triangleq \{\tilde{x} : V(\tilde{x}) \leq 1\}$. Par construction, $\tilde{\mathcal{E}}$ est un ensemble invariant inclus à la fois dans $\tilde{\mathcal{X}}$ et $\tilde{\mathcal{A}}$.

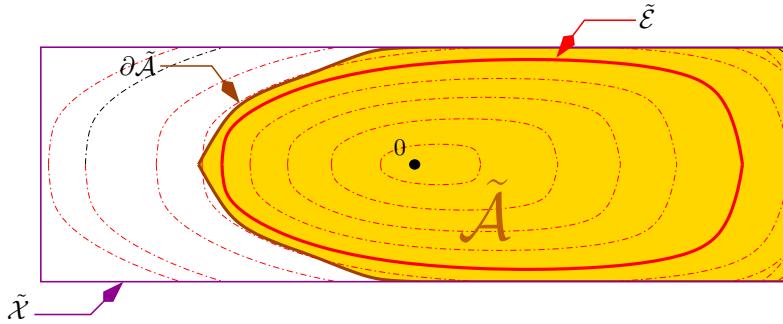


FIGURE III.3 – Principe d'une solution par LYAPUNOV, lorsque $V(\cdot)$ et $Z_j(\cdot)$ ne dépendent que de \tilde{x} . L'ensemble $\tilde{\mathcal{E}} \triangleq \{\tilde{x} : V(\tilde{x}) \leq 1\}$ est inclus dans $\tilde{\mathcal{X}}$ et dans le sous-espace admissible $\tilde{\mathcal{A}} = \cap_{j=1}^J \tilde{\mathcal{A}}_j$, avec $\tilde{\mathcal{A}}_j \triangleq \{\tilde{x} : Z_j'(\tilde{x})\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]\}$. Par construction, $\tilde{\mathcal{X}}$ est aussi invariant.

B De la faisabilité à l'optimisation d'un bassin d'attraction multicritère Les équations (III.14)–(III.17) définissent des conditions sur la variable de décision P pour l'obtention d'un ensemble $\tilde{\mathcal{E}}$. Celles-ci doivent être complétées en vue d'améliorer la pertinence de $\tilde{\mathcal{E}}$ vis-à-vis du problème considéré. À titre d'exemple, dans un but d'analyse du système bouclé pour des situations relatives caméra-cible appartenant à un domaine \mathcal{X}_0 donné, des contraintes supplémentaires doivent être placées sur P de telle sorte que $\tilde{\mathcal{X}}_0 \triangleq \mathcal{X}_0 \times \mathcal{X}_{c0} \in \tilde{\mathcal{E}}$, où \mathcal{X}_{c0} désigne l'ensemble des conditions initiales considérées du vecteur d'état du contrôleur. La recherche peut également être guidée par la minimisation d'un critère convexe en P relatif à la « taille » de $\tilde{\mathcal{E}}$ ou à la maximisation heuristique de l'étendue de $\tilde{\mathcal{E}}$ vers un ensemble de points $\{\lambda_h\} \subset \tilde{\mathcal{X}}$ sélectionné *a priori*. De même, il peut être intéressant de maximiser la taille d'un ellipsoïde inscrit dans $\tilde{\mathcal{E}}$, voire, dans un contexte de synthèse, de maximiser le taux de convergence de la boucle fermée.

L'objectif est donc de définir un programme de faisabilité ou d'optimisation en la variable P et d'autres variables de décision, qui possède des propriétés intéressantes, et dont la solution conduise à une fonction $V(\cdot, \cdot)$ et un ensemble $\tilde{\mathcal{E}}$ théoriquement satisfaisants. Ce programme peut ne constituer qu'une condition suffisante au problème original, auquel cas son obtention s'accompagne de pessimisme – ou *conservatisme* – qu'il s'agit de limiter.

Notons que la réunion de bassins d'attraction multicritères constitue un bassin d'attraction multicritère. Cette propriété élémentaire peut être mise à profit dans un contexte d'analyse où le but est d'établir un bassin d'attraction multicritère maximal. Ainsi, il est parfois opportun

de définir $\tilde{\mathcal{E}}$ comme la réunion $\tilde{\mathcal{E}} \triangleq \cup_{r \in \{1, \dots, R\}} (\tilde{\mathcal{E}})_r$ de R bassins $(\tilde{\mathcal{E}})_r$ calculés indépendamment pour des polytopes $(\tilde{\mathcal{X}} \times \mathcal{X}_\chi)_r$ et/ou des directions d'extension $(\{\lambda_h\})_r$.

Avant d'entrer dans le détail, un constat fondamental s'impose.

C Difficultés soulevées par le contexte robotique Les conditions que doit satisfaire la fonction de LYAPUNOV sous-tendant l'analyse ou la synthèse multicritère sont assez classiques. Cependant, le contexte de l'asservissement visuel en robotique soulève des difficultés non triviales, dont on verra qu'elles peuvent nécessiter des développements méthodologiques originaux y compris en automatique. En effet, les zones admissibles de l'espace d'état vis-à-vis des différentes contraintes ne sont généralement pas symétriques par rapport à $\tilde{x}^* = 0$, ni convexes, ni même connexes. Une conséquence fondamentale est que (III.17) risque de s'avérer extrêmement conservative pour certaines classes de fonctions de LYAPUNOV, *e.g.* si les iso-coûts définissent des ensembles symétriques par rapport à 0 et/ou convexes.

III.3.1.3 Inégalités Matricielles et Lemmes importants associés

A LMIs, BMIs et Problème de complémentarité sur le cône Cette section rappelle quelques rudiments de programmation semi-définie.

Définition III.3 (LMI [Boyd 94]) *Une contrainte $\mathcal{L}(x)$ sur une variable de décision vectorielle ou matricielle x est une Inégalité Matricielle Linéaire – ou Linear Matrix Inequality (LMI) – en x si elle s'écrit comme la (semi-)définie positivité ou négativité d'une combinaison matricielle affine en les composantes x_1, \dots, x_n de x , i.e. si elle s'écrit – les matrices $A_i = A'_i$ étant données –*

$$\mathcal{L}(x) : A_0 + x_1 A_1 + \dots + x_n A_n \leq \mathbb{O}. \quad (\text{III.18})$$

Ci-dessus, les valeurs de x sont supposées réelles et $'$ désigne l'opérateur de transposition.

Les contraintes LMI sont convexes. Par conséquent, la faisabilité d'un ensemble de LMIs de même que la minimisation d'un critère convexe sous des contraintes LMIs sont des programmes convexes. Des « solveurs » dédiés, *e.g.* les boîte à outils SeDuMi [Sturm 99] et SDPT3 [Toh 99] exploitables sous MATLAB grâce à YALMIP [Löfberg 04], permettent leur résolution numérique en temps polynomial et avec une précision arbitraire – en se limitant à certains types de critères dans le cas de l'optimisation. C'est pourquoi depuis de nombreuses années, de nombreux travaux visent à obtenir une « solution LMI » de problèmes d'Automatique non solubles analytiquement, *e.g.* en présence de contraintes contradictoires \dagger .

Nous ne rentrerons pas davantage dans le détail de l'optimisation LMI et de ses applications. Nous nous limiterons seulement à deux définitions supplémentaires et à leur caractérisation.

Définition III.4 (BMI) *Une Inégalité Matricielle Bilinéaire – ou Bilinear Matrix Inequality (BMI) – en la variable réelle x est de la forme*

$$\mathcal{B}(x) : A_0 + \sum_{i=1}^n x_i A_i + \sum_{i_1, i_2=1}^n x_{i_1} x_{i_2} B_{i_1 i_2} \leq \mathbb{O}, \quad (\text{III.19})$$

où $x = (x_1 \dots x_n)'$ et les matrices $A_i = A'_i$ et $B_{i_1 i_2} = B'_{i_1 i_2}$ sont données.

\dagger . Signalons ici que l'équipe de recherche en Commande Robuste du LAAS-CNRS (dirigée par J. BERNUSOU) a apporté des contributions fondatrices à la résolution numérique de problèmes d'Automatique difficiles par des programmes d'optimisation convexe, *e.g.* [Gerome91]

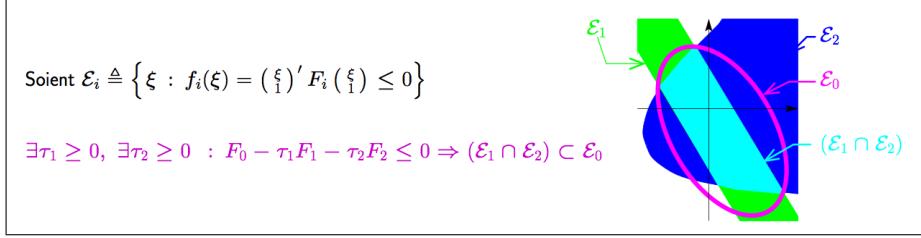


FIGURE III.4 – Illustration de la S-procédure.

Les BMIs n'étant généralement pas convexes, la faisabilité d'un ensemble de BMIs ou l'optimisation sous contraintes BMIs est de ce fait beaucoup plus complexe.

Définition III.5 (CCP [El Ghaoui 97]) *Le problème*

$$CCP(X, S, Z) : \mathcal{L}(X, S, Z), X > \mathbb{O}, S > \mathbb{O}, XS = \mathbb{I} \quad (\text{III.20})$$

en les variables de décision matricielles $X \in \mathbb{R}^{n \times n}$, $S \in \mathbb{R}^{n \times n}$ et Z , où \mathbb{I} désigne la matrice identité et $\mathcal{L}(\cdot, \cdot, \cdot)$ une LMI quelconque en ses arguments, est appelé Problème de Complémentarité sur le Cône – ou Cone Complementarity Problem (CCP).

Bien que non convexe en raison de la contrainte $XS = \mathbb{I}$, ce problème peut être résolu efficacement par l'algorithme [El Ghaoui 97] basé sur des séquences de LMIs.

B Un lemme important : la S-procédure La définition du bassin d'attraction multicritère implique des inclusions ensemblistes. Pour des ensembles définis par des fonctions quadratiques, le lemme ci-dessous, illustré Figure III.4, établit une jonction avec les LMIs.

Lemme III.1 (S-procédure [Boyd 94]) Soient $f_l(\xi) \triangleq \left(\begin{smallmatrix} \xi \\ 1 \end{smallmatrix} \right)' F_l \left(\begin{smallmatrix} \xi \\ 1 \end{smallmatrix} \right)$, $l = 0, \dots, L$, des fonctions quadratiques données, avec $F_l = F_l'$. Soient les ensembles $\mathcal{E}_l \triangleq \{\xi : f_l(\xi) \leq 0\}$. L'équation suivante constitue une condition suffisante à la relation ensembliste $\cap_{l=1}^L \mathcal{E}_l \subset \mathcal{E}_0$:

$$\begin{aligned} f_0(\xi) \leq 0, \forall \xi : f_1(\xi) \leq 0, \dots, f_L(\xi) \leq 0 \\ \uparrow \end{aligned} \quad (\text{III.21})$$

$$\exists \tau_1 \geq 0, \dots, \tau_L \geq 0 : \forall \xi, f_0(\xi) - \sum_{l=1}^L \tau_l f_l(\xi) \leq 0.$$

Si les composantes de ξ sont indépendantes, alors l'équation ci-dessus se réécrit en un programme de faisabilité LMI :

$$\begin{aligned} f_0(\xi) \leq 0, \forall \xi : f_1(\xi) \leq 0, \dots, f_L(\xi) \leq 0 \\ \uparrow \end{aligned} \quad (\text{III.22})$$

$$\exists \tau_1, \dots, \tau_L : \begin{cases} \tau_1 \geq 0, \dots, \tau_L \geq 0 \\ F_0 - \sum_{l=1}^L \tau_l F_l \leq 0. \end{cases}$$

D'autres lemmes liés aux LMIs sont utilisés dans le développement des résultats, tels la S-procédure pour des contraintes égalités et/ou des formes quadratiques, le cas où la S-procédure est non conservative, le lemme de Schur pour des inégalités strictes ou larges, et le lemme d'élimination [Boyd 94]. On donnera simplement le moment venu une condition suffisante à $\cap_{l=1}^L \mathcal{E}_l \subset \mathcal{E}_0$ moins conservatrice lorsque les composantes du vecteur ξ sont liées.

III.3.2 Analyse et Synthèse multicritères par des arguments de stabilité quadratique

Cette première approche a été développée dans le cadre de la thèse de Delphine BELLOT [DOC-01].

III.3.2.1 Linéarisation globale

Sous l'hypothèse $(\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi$, la boucle fermée visuelle non linéaire (III.11), dont nous rappelons l'équation

$$[\text{BFLN}] : \dot{\tilde{x}} = \tilde{A}(\tilde{x}, \chi)\tilde{x}, \quad (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi \subset \mathbb{R}^{\tilde{n}} \times \mathbb{R}^{n_\chi}, \quad 0 \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \quad (\text{III.23})$$

est « immérsee » dans le système linéaire incertain décrit par l'inclusion différentielle

$$[\text{LPV}] : \dot{\tilde{x}} = \tilde{A}(\delta)\tilde{x}, \quad \delta \triangleq ((\delta^{\tilde{x}})' (\delta^\chi)')' \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \quad (\text{III.24})$$

où le vecteur de paramètres incertains $\delta \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi$ est variant dans le temps et sans mémoire. En d'autres termes, chaque composante $\tilde{x}_1, \dots, \tilde{x}_{\tilde{n}}, \chi_1, \dots, \chi_{n_\chi}$ de \tilde{x} et χ intervenant – rationnellement – dans la matrice $\tilde{A}(.,.)$ du système non linéaire [BFLN] est remplacée formellement par un paramètre incertain $\delta_1 \triangleq \delta_1^{\tilde{x}}, \delta_2 \triangleq \delta_2^{\tilde{x}}, \dots, \delta_{\tilde{n}} \triangleq \delta_{\tilde{n}}^{\tilde{x}}, \delta_{\tilde{n}+1} \triangleq \delta_1^\chi, \dots, \delta_{\tilde{n}+n_\chi} \triangleq \delta_{n_\chi}^\chi$, vivant dans son intervalle de définition. Il s'en suit qu'*une trajectoire* $\tilde{x}(.)$ de [BFLN] est également trajectoire de [LPV] si et seulement si elle est entièrement contenue dans $\tilde{\mathcal{X}}$.

Ainsi, dans un contexte d'analyse (resp. de synthèse), on montre (resp. assure) qu'une propriété est vérifiée par la boucle fermée non linéaire originale [BFLN] en montrant (resp. assurant) qu'elle est satisfaite par chaque réalisation du système linéaire incertain englobant [LPV]. Cette technique, parfois dénommée *linéarisation globale* [Boyd 94], traite les non-linéarités et les incertitudes de manière unifiée. Elle induit naturellement du pessimisme, car de nombreuses trajectoires du système [LPV] n'ont pas de sens pour le problème considéré. Par conséquent, elle donne lieu à des conditions suffisantes d'analyse ou de synthèse multicritère, qui peuvent s'avérer trop pessimistes ou infaisables.

III.3.2.2 Exploitation et prolongements mineurs de techniques existantes

EL GHAOUI et SCORLETTI [El Ghaoui 96] proposent la définition d'un bassin d'attraction multicritère d'un système rationnel soumis à des contraintes linéaires sur le vecteur d'état, sur la base d'une fonction de LYAPUNOV quadratique $V(\tilde{x}, \chi) = V(\tilde{x}) = \tilde{x}' \mathbf{P} \tilde{x}$, $\mathbf{P} > \mathbb{O}$. Par hypothèse, $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$ est un paralléléotope contenant 0 et orienté selon les axes principaux de l'espace d'état. Nous avons dans un premier temps exploité puis prolongé ces travaux. Dans la présentation qui suit des aspects importants, nous commettons quelques abus de langage et sacrifices la rigueur mathématique afin de limiter la technicité et la longueur du manuscrit †.

A Écriture de [LPV] sous forme d'une LFT $\mathcal{F}_u(\Sigma, \Delta)$ Dans un premier temps, le système linéaire incertain [LPV] est réécrit sous la forme équivalente d'une *Transformation*

†. Le lecteur est invité à se référer aux publications jointes en annexe pour un traitement rigoureux.

Linéaire Fractionnaire (LFT = Linear Fractional Transform)

$$\dot{\tilde{x}} = \left(\tilde{A}^\sharp + \tilde{B}_p^\sharp \Delta (\mathbb{I} - \tilde{D}_{qp}^\sharp \Delta)^{-1} \tilde{C}_q^\sharp \right) \tilde{x} \quad (\text{III.25})$$

$$\text{avec } \Delta = \text{diag}(\delta_1 \mathbb{I}_{r_1}, \dots, \delta_{\tilde{n}} \mathbb{I}_{r_{\tilde{n}}}, \delta_{\tilde{n}+1} \mathbb{I}_{r_{(\tilde{n}+1)}}, \dots, \delta_{\tilde{n}+n_\chi} \mathbb{I}_{r_{(\tilde{n}+n_\chi)}}), \quad (\text{III.26})$$

moyennant un choix adéquat des matrices \tilde{A}^\sharp , \tilde{B}_p^\sharp , \tilde{D}_{qp}^\sharp , \tilde{C}_q^\sharp et du vecteur $r \triangleq (r_1 \dots r_{(\tilde{n}+n_\chi)})' \in \mathbb{R}^{(\tilde{n}+n_\chi)}$. Notons que la matrice Δ est diagonale et constituée de répétitions des composantes $\delta_1, \dots, \delta_{\tilde{n}+n_\chi}$ du vecteur incertain δ défini en (III.24). Si $\det(\mathbb{I} - \tilde{D}_{qp}^\sharp \Delta) \neq 0$ pour tout $\delta \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi$, alors la LFT (III.25)–(III.26) est dite *bien posée* sur $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$. Elle est alors équivalente à l’interconnexion $\mathcal{F}_u(\Sigma, \Delta)$ –elle-même bien posée– d’un système dynamique linéaire invariant Σ et d’un gain matriciel incertain égal à Δ , *i.e.*

$$\mathcal{F}_u(\Sigma, \Delta) : \begin{pmatrix} \dot{\tilde{x}} \\ \tilde{q}^\sharp \end{pmatrix} = \begin{pmatrix} \tilde{A}^\sharp & \tilde{B}_p^\sharp \\ \tilde{C}_q^\sharp & \tilde{D}_{qp}^\sharp \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{p}^\sharp \end{pmatrix}, \quad \tilde{p}^\sharp = \Delta \tilde{q}^\sharp. \quad (\text{III.27})$$

Une étape de *normalisation* permet la réécriture de (III.27) en †

$$\begin{pmatrix} \dot{\tilde{x}} \\ \tilde{q}^\sharp \end{pmatrix} = \begin{pmatrix} \tilde{A} & \tilde{B}_p \\ \tilde{C}_q & \tilde{D}_{qp} \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{p} \end{pmatrix}, \quad \tilde{p} = \Delta \tilde{q} \text{ où, dans (III.26), } \delta \in [-1; +1]^{\tilde{n}+n_\chi}. \quad (\text{III.28})$$

L’interconnexion (III.28) obtenue après normalisation est communément appelée *Inclusion Différentielle Linéaire Structurée et Bornée en Norme* (SNLDI = Structured Norm-Bounded Linear Differential Inclusion). Sous l’hypothèse de bien posé, elle est donc équivalente à [LPV]. Naturellement, toute trajectoire $\tilde{x}(.)$ de la boucle fermée originale [BFNL] entièrement contenue dans $\tilde{\mathcal{X}}$ est « couverte » par le faisceau de trajectoires de la SNLDI (III.28) émanant de sa condition initiale $\tilde{x}(0)$.

Dans un contexte d’analyse, les matrices intervenant dans la LFT (III.25)–(III.26), et donc dans la SNLDI (III.28), sont connues. Lorsque l’objectif est la synthèse, elles s’obtiennent à partir des matrices de la LFT/SNLDI issue de la linéarisation globale du modèle rationnel en boucle ouverte (III.9) et des paramètres, à déterminer, du contrôleur.

B Conditions pour que $V(\tilde{x}) = \tilde{x}' P \tilde{x}$ soit une fonction de Lyapunov pour $\mathcal{F}_u(\Sigma, \Delta)$
 La fonction $V(.)$ prouve la stabilité asymptotique globale de la SNLDI (III.28) englobant [BFNL] si elle satisfait $V(\tilde{x}) > 0$ et $\dot{V}(\tilde{x}) < 0$ pour tous $\tilde{x} \neq 0$ et toutes les réalisations de Δ . Les ensembles $\tilde{\mathcal{E}}_\gamma \triangleq \{\tilde{x} : V(\tilde{x}) \leq \gamma\}$, $\gamma > 0$, définissent alors des ellipsoïdes emboités centrés sur 0 tels que toute trajectoire de (III.28) émanant de $\tilde{\mathcal{E}}_\gamma$ demeure dans $\tilde{\mathcal{E}}_\gamma$ et converge vers 0. C’est en particulier le cas de $\tilde{\mathcal{E}} \triangleq \tilde{\mathcal{E}}_1 \triangleq \{\tilde{x} : V(\tilde{x}) \leq 1\}$. Accessoirement, un taux de décroissance des trajectoires de la SNLDI (III.28) ‡ au moins égal à α est assuré si on exige que $\dot{V}(\tilde{x}) < -2\alpha V(\tilde{x})$, auquel cas si $\tilde{x}(t_0) \in \tilde{\mathcal{E}}$, alors $\forall t, \tilde{x}(t) \in e^{-\alpha(t-t_0)} \tilde{\mathcal{E}}$.

Les considérations précédentes ne peuvent pas toutes être transformées en inégalités matricielles impliquant la matrice de LYAPUNOV P . Pour parvenir à une solution constructive, [El Ghaoui 96] exploite l’*approximation externe* suivante de la SNLDI (III.28) :

$$\begin{pmatrix} \dot{\tilde{x}} \\ \tilde{q}^\sharp \end{pmatrix} = \begin{pmatrix} \tilde{A} & \tilde{B}_p \\ \tilde{C}_q & \tilde{D}_{qp} \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{p} \end{pmatrix}, \quad \begin{pmatrix} \tilde{q} \\ \tilde{p} \end{pmatrix}' \begin{pmatrix} S & G \\ G' & -S \end{pmatrix} \begin{pmatrix} \tilde{q} \\ \tilde{p} \end{pmatrix} \geq 0, \quad (S, G) \in \mathcal{S}(r) \times \mathcal{G}(r). \quad (\text{III.29})$$

† . Les matrices \tilde{A} , \tilde{B}_p , \tilde{D}_{qp} , \tilde{C}_q dépendent bien sûr des bornes de $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$.

‡ . Le taux de décroissance [Boyd 94] est le plus grand réel β tel que $\lim_{t \rightarrow +\infty} \|\tilde{x}(t)\| = 0$.

Toute trajectoire de (III.28) est également trajectoire de (III.29) quelles que soient les valeurs admissibles des matrices constantes – ou *scalings* – S et G . Ces matrices vivent dans des ensembles $\mathcal{S}(r)$ et $\mathcal{G}(r)$ préalablement définis de façon à « capturer au mieux » la structure de Δ . Dans l’écriture des inégalités matricielles en P qui assurent que (III.29) est stable quadratiquement, des degrés de liberté sont rajoutés en considérant S et G comme des variables de décision, *i.e.* en autorisant l’approximation externe (III.29) la plus « favorable ». Les conditions suffisantes de stabilité quadratique de (III.29) décrites dans [El Ghaoui 96][Dussy 98a] assurent également le bien posé de (III.28).

C Conditions pour que $\tilde{\mathcal{E}} \triangleq \{\tilde{x} : V(\tilde{x}) \leq 1\}$ soit un bassin d’attraction multicritère pour [BFNL] Du fait que les trajectoires $\tilde{x}(.)$ de [BFNL] entièrement contenues dans $\tilde{\mathcal{X}}$ sont également trajectoires de [LPV], la fonction $V(\tilde{x}) = \tilde{x}'P\tilde{x}$ permettant d’établir la stabilité asymptotique – globale – de [LPV] est également une fonction de LYAPUNOV pour [BFNL] sur $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$, *i.e.* satisfait (III.14)–(III.15). L’incorporation via la S-procédure de la condition (III.16) assurant que $\tilde{\mathcal{E}}$ est un bassin d’attraction de [BFNL] en l’absence de contraintes, de même que l’inclusion – suggérée au §III.3.1.2-B – d’un ensemble de conditions initiales $\tilde{\mathcal{X}}_0$ dans $\tilde{\mathcal{E}}$, sont assez immédiates.

Il reste à contraindre P de façon que $V(.)$ satisfasse (III.17). Nous avons dans un premier temps considéré l’inclusion de l’ensemble $\tilde{\mathcal{E}} \triangleq \{\tilde{x} : V(\tilde{x}) \leq 1\}$ dans un sous-espace $\tilde{\mathcal{A}}_j \triangleq \{\tilde{x} : Z_j'(\tilde{x})\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]\}$ tel que la fonction rationnelle $Z_j(.,.)$ dans (III.17) ne dépend pas de l’incertitude χ . La littérature ne proposant pas de solution, des prolongements mineurs mais relativement originaux ont été nécessaires [CIACL-04-AV-2][†]. L’un consiste à établir une approximation interne *ad hoc* de $\tilde{\mathcal{A}}_j$ en tant que l’intersection $\cap_l \tilde{\mathcal{Q}}_{j,l}$ d’ensembles $\tilde{\mathcal{Q}}_{j,l}$ définis au moyen de fonctions quadratiques, de sorte que les relations $\tilde{\mathcal{E}} \subset \tilde{\mathcal{Q}}_{j,l}$ se prêtent directement à l’application de la S-procédure. Une seconde approche, plus élégante, repose sur la linéarisation globale de la contrainte définissant $\tilde{\mathcal{A}}_j$, *i.e.* sur l’exploitation des contraintes $Z_j'(\delta)\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]$, avec δ incertain vivant dans $\tilde{\mathcal{X}}$, afin d’assurer que $Z_j'(\tilde{x})\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]$. Plus rigoureusement, soit $\tilde{\mathcal{C}}_j \triangleq \{\tilde{x} : \forall \delta \in \tilde{\mathcal{X}}, Z_j'(\delta)\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]\}$. Du fait que $\tilde{\mathcal{B}}_j \triangleq \tilde{\mathcal{C}}_j \cap \tilde{\mathcal{X}} = \{\tilde{x} \in \tilde{\mathcal{X}} : \forall \delta \in \tilde{\mathcal{X}}, Z_j'(\delta)\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]\}$ est inclus dans $\tilde{\mathcal{A}}_j \cap \tilde{\mathcal{X}} = \{\tilde{x} \in \tilde{\mathcal{X}} : Z_j'(\tilde{x})\tilde{x} \in [\underline{\zeta}_j, \bar{\zeta}_j]\}$, et que $\tilde{\mathcal{E}} \subset \tilde{\mathcal{X}}$ est assuré par ailleurs, il suffit d’exiger[‡] $\tilde{\mathcal{E}} \subset \tilde{\mathcal{C}}_j$ pour assurer que $\tilde{\mathcal{E}} \subset \tilde{\mathcal{A}}_j$. Des conditions suffisantes s’obtiennent par la S-procédure. Si celles-ci s’avèrent trop conservatives ou infaisables, alors on doit alors recourir à la première approche, plus laborieuse.

Toute fonction vectorielle $Z_j(.,.)$ est donnée dans un contexte d’analyse, mais aussi lorsque le but est de satisfaire des contraintes image ou des contraintes 3D lors de la synthèse d’un contrôleur. La satisfaction de (III.17) se décline différemment dans un contexte de synthèse si $Z_j(.,.)$ dépend des paramètres, à déterminer, du contrôleur, *e.g.* lors de la gestion de saturations d’actionneurs alors que le contrôleur n’est pas entièrement spécifié.

[†]. L’évitement de saturations pour une commande 3D statique linéaire $u = Kx$ implique une fonction vectorielle $Z_j(\tilde{x})$ linéaire. La solution de ce cas étant assez immédiate [El Ghaoui 96][Dussy 98a], nous n’y ferons plus référence et nous focaliserons sur les fonctions $Z_j(.,.)$ rationnelles.

[‡]. Outre une contrainte de bien posé de l’inclusion statique $\zeta_j = Z_j'(\delta)\tilde{x}$ pour tout $\delta \in \tilde{\mathcal{X}}$, qui pose des problèmes aigus en asservissement visuel, cf. §III.3.2.3.

D Aperçu des programmes d'optimisation obtenus La Table III.2 schématise les inégalités matricielles que doit satisfaire la matrice de LYAPUNOV P . Supposons dans un premier temps que le taux de décroissance minimum α ne soit pas introduit dans l'étude, et considérons l'analyse multicritère. Des conditions suffisantes s'expriment alors en termes de LMIs en P et possiblement d'autres variables. On note que dans les LMIs assurant la décroissance de $V(.)$ le long des trajectoires de la boucle fermée [BFNL], apparaissent les “scalings” S, G impliqués dans l'approximation externe (III.29) de la SNLDI englobante (III.28). Le traitement de toute contrainte *via* la linéarisation globale de l'équation définissant l'ensemble admissible associé \tilde{A}_j conduit similairement à l'introduction de scalings S_j, G_j . La maximisation de la « taille » de $\tilde{\mathcal{E}}$ s'effectue au moyen du critère linéaire convexe $\text{trace}(P)$. Par conséquent, l'analyse multicritère d'asservissements visuels est temporairement considérée comme résolue du fait qu'elle se réduit à un programme de faisabilité/optimisation LMI.

Lors de la synthèse de contrôleurs 3D statiques linéaires $u = Kx$, les conditions s'expriment au moyen de LMIs en $Q = P^{-1}$ et d'autres variables, parmi lesquelles figurent des scalings T, H liés à l'approximation externe de (III.28) par (III.29)[†]. Le gain K du contrôleur s'obtient sur la base des matrices solutions du programme de faisabilité/optimisation. Une méthode de synthèse a également été développée pour des contrôleurs 3D statiques non linéaires rationnels tels que la minimalité de la LFT de la boucle fermée [BFNL] soit assurée. Le gain rationnel $K(x)$ de tels contrôleurs s'exprime comme une fonction de deux gains matriciels K_1, K_2 également déterminés à partir de matrices solutions d'un programme de faisabilité/optimisation LMI. Notons que la synthèse d'un contrôleur 3D assurant la convergence de la caméra et le respect des contraintes pour le plus vaste ensemble de conditions initiales peut être guidée par la maximisation du volume de $\tilde{\mathcal{E}}$, *via* la minimisation de $\log(\det(Q^{-1}))$.

Le fait de fixer le scalaire α à une valeur strictement positive permet, dans un contexte d'analyse, de se restreindre aux trajectoires de la boucle fermée présentant un taux de convergence minimal, et, dans un contexte de synthèse, de contraindre le temps de réponse. Le taux de décroissance α peut également être considéré comme une variable de décision dont on recherche le maximum. Dans ce cas, les inégalités matricielles permettant la satisfaction de $\dot{V}(\tilde{x}) < -2\alpha V(\tilde{x})$ sont non seulement linéaires en P (resp. en Q) pour α fixé, mais également linéaires en α pour P (resp. en Q) fixé. Le programme de faisabilité/optimisation obtenu revient en fait à minimiser la plus grande valeur propre généralisée d'un faisceau matriciel sous des contraintes LMI. Ce problème aux valeurs propres généralisées (GEVP = Generalized Eigenvalue Problem) est également soluble numériquement.

La synthèse d'une commande 2D linéaire statique $u = Ky$ ou d'une commande 2D dynamique – notée $u = \mathcal{K}y$ – conduit à des inégalités matricielles du même type. On remarque que celles qui assurent la décroissance de $\dot{V}(.)$ ou la non saturation des actionneurs sont bilinéaires non convexes. La détermination de l'optimum global du programme d'optimisation peut donc s'avérer hors de portée. Si on supprime certains degrés de liberté dans l'approximation externe (III.29) de la SNLDI (III.28), en l'occurrence si on fixe $G = \mathbb{O}$, alors la synthèse de la commande 2D se simplifie en deux étapes de complexité réduite. On résoud d'abord un problème de complémentarité sur le cône [El Ghaoui 97]. Les matrices solutions entrent ensuite en tant que constantes dans une LMI permettant de déterminer le gain du contrôleur. Cependant, ce dernier programme s'avère souvent infaisable en raison du pessimisme élevé de l'approximation externe (III.29) lorsqu'on fixe S ou G .

La synthèse d'un retour d'une commande 2D dynamique séquencée, dont les paramètres

[†]. Ceux-ci doivent demeurer libres, sous peine de dégrader significativement les résultats [CIACL-02-AV-1].

TABLE III.2 – Aperçu de quelques programmes d'optimisation (LMI/BMI/GEVP) pour l'analyse et la synthèse multicritères d'asservissements visuels.

Si possible, $\min(-\alpha) / \min \text{trace}(\mathbf{P}) / \min(\log(\det(\mathbf{Q}^{-1})))$, sous les contraintes Inégalités Matricielles

	$\tilde{\mathcal{E}} \triangleq \{\tilde{x} : V(\tilde{x}) \leq 1\}$ bassin d'attract° (III.14) & (III.15)/ avec $\mathcal{E} \subset \tilde{\mathcal{X}}$ (III.16)		$\tilde{\mathcal{E}} \subset \tilde{\mathcal{A}}_j$ (III.17)			$\tilde{x}_0 \in \tilde{\mathcal{E}}$	Reconstruction du contrôleur
Analyse	$\mathcal{L}_\alpha(\mathbf{P}, \mathbf{S}, \mathbf{G})$	$\mathcal{L}(\mathbf{P})$	$\mathcal{L}(\mathbf{P})$	$\mathcal{L}(\mathbf{P}, \mathbf{S}_j, \mathbf{G}_j)$	—	$\mathcal{L}(\mathbf{P})$	—
Synthèse $u = Kx$	$\mathcal{L}_\alpha(\mathbf{Q}, \mathbf{Y}, \mathbf{T}, \mathbf{H})$	$\mathcal{L}(\mathbf{Q})$	$\mathcal{L}(\mathbf{Q})$	$\mathcal{L}(\mathbf{Q}, \mathbf{T}_j, \mathbf{H}_j)$	$\mathcal{L}(\mathbf{Q}, \mathbf{Y})$	$\mathcal{L}(\mathbf{Q})$	$K = \mathbf{Y}\mathbf{Q}^{-1}$
Synthèse $u = K(x)x$	$\mathcal{L}_\alpha(\mathbf{Q}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{T})$	$\mathcal{L}(\mathbf{Q})$	$\mathcal{L}(\mathbf{Q})$	$\mathcal{L}(\mathbf{Q})$	$\mathcal{L}(\mathbf{Q}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{T})$	$\mathcal{L}(\mathbf{Q})$	$K_1 = \mathbf{Y}_1\mathbf{Q}^{-1}$ $K_2 = \mathbf{Y}_2\mathbf{Q}^{-1}$
Synthèse $u = Ky$ ou $u = \mathcal{K}y$	$\mathcal{B}(\mathbf{P}, \mathbf{K}, \mathbf{S}, \mathbf{G}, \alpha)$ ou bien $\mathcal{L}_\alpha(\mathbf{P}, \mathbf{Q}, \mathbf{S}, \mathbf{T})$ $\mathbf{P}\mathbf{Q}=\mathbb{I}$ $\mathbf{S}\mathbf{T}=\mathbb{I}$	$\mathcal{L}(\mathbf{P})$	$\mathcal{L}(\mathbf{P})$	$\mathcal{L}(\mathbf{P}, \mathbf{T}_j, \mathbf{H}_j)$ ou bien $\mathcal{L}(\mathbf{P}, \mathbf{S}_j, \mathbf{G}_j)$ & $\mathcal{L}(\mathbf{Q}, \mathbf{T}_j, \mathbf{H}_j)$	$\mathcal{B}(\mathbf{P}, \mathbf{K}, \mathbf{S}, \mathbf{G})$ avec reconstruction du contrôleur	$\mathcal{L}(\mathbf{P})$ ou bien $\mathcal{L}(\mathbf{P})$ $\mathcal{L}(\mathbf{Q})$	—
Synthèse $u = \mathcal{K}(\tilde{x})y$	$\mathcal{L}_\alpha(\mathbf{P}, \mathbf{Q}, \mathbf{Y}, \mathbf{S}, \mathbf{T})$ $\mathbf{S}\mathbf{T}=\mathbb{I}$	$\mathcal{L}(\mathbf{Q})$	$\mathcal{L}(\mathbf{Q})$	$\mathcal{L}(\mathbf{Q}, \mathbf{T}_j, \mathbf{S}_j)$	$\mathcal{L}(\mathbf{P}, \mathbf{Q}, \mathbf{Y})$	$\mathcal{L}(\mathbf{P})$	$\mathcal{L}(\mathbf{K})$

Ci-dessus,

- le traitement des contraintes $\tilde{\mathcal{E}} \subset \tilde{\mathcal{A}}_j$ s'effectue
 - [si $Z_j(\cdot)$ est parfaitement définie] (1) au moyen de l'approximation interne de $\tilde{\mathcal{A}}_j$ par l'intersection $\cap_l \tilde{\mathcal{Q}}_{j,l}$ d'ensembles $\tilde{\mathcal{Q}}_{j,l}$ définis par des fonctions quadratiques ; (2) par linéarisation globale de l'équation définissant $\tilde{\mathcal{A}}_j$ (valable aussi si $Z_j(\cdot, \cdot)$ dépend de l'incertitude) ;
 - [si $Z_j(\cdot, \cdot)$ dépend du correcteur en cours de synthèse] (3) par linéarisation globale de l'équation définissant $\tilde{\mathcal{A}}_j$ (valable aussi si $Z_j(\cdot, \cdot)$ dépend de l'incertitude) ;
- $\mathcal{L}(\cdot)$ et $\mathcal{B}(\cdot)$ désignent des inégalités matricielles linéaires et bilinéaires, respectivement ; $\mathcal{L}_\alpha(\mathbf{P}, \dots)$ désigne un problème aux valeurs propres généralisés (GEVP), *i.e.* $\mathcal{L}_\alpha(\mathbf{P}, \dots) = \mathcal{L}(\mathbf{P}, \alpha\mathbf{P}, \dots)$;
- K désigne un gain matriciel, la fonction matricielle $K(\cdot)$ désigne un gain fonction de son argument, l'opérateur \mathcal{K} désigne un système dynamique, l'opérateur $\mathcal{K}(\cdot)$ désigne un système dynamique dont les paramètres sont fonctions des quantités placées en argument.

sont affectés en ligne à une fonction rationnelle de x , conduit à des conclusions analogues.

III.3.2.3 Remarques importantes

Plusieurs remarques peuvent être formulées concernant une « approche quadratique » de l'asservissement visuel multicritère.

Lors de la phase de modélisation, l'écriture du système [LPV] (III.24) sous la forme LFT (III.25)–(III.26) n'est pas unique. Néanmoins, le nombre de répétitions r_n de tout paramètre δ_n , $n = 1, \dots, \tilde{n} + n_\chi$, est au moins égal à l'exposant maximal de celui-ci dans les monômes de $\tilde{\mathcal{A}}(\delta)$. La LFT (III.25)–(III.26) est dite *minimale* si la somme $\sum_{n=1}^{\tilde{n}+n_\chi} r_n$ des répétitions des $\{\delta_n\}_{n=1, \dots, \tilde{n}+n_\chi}$ dans Δ est minimale. Or, le conservatisme de l'approximation externe (III.29) de (III.28) est d'autant plus important que le nombre de répétitions des $\{\delta_n\}_{n=1, \dots, \tilde{n}+n_\chi}$ dans Δ est élevé. C'est pourquoi il est primordial d'obtenir une LFT minimale de [LPV]. La littérature propose quelques règles de construction et de réduction des LFTs [Zhou 95][Dussy 98a], mais il convient parfois de recourir à des méthodes *ad hoc*. Notons que ceci peut conditionner le choix des indices visuels.

Les prolongements mineurs proposés de [El Ghaoui 96][Dussy 98a] ne modifient pas les propriétés du bassin d'attraction multicritère sous-jacent à la méthode d'analyse ou de synthèse. S'agissant d'un ellipsoïde centré sur 0, la nécessité de l'inclure dans des ensembles

admissibles dissymétriques et/ou non convexes prévient parfois l'élaboration de conclusions sur des problèmes de Robotique triviaux, cf. §III.3.1.2-C et §III.3.2.5.

Enfin, une autre limitation drastique de la méthode provient de problèmes potentiels de bien posé concomitants à la linéarisation globale de l'équation d'une contrainte (§III.3.2.2-C). Ceux-ci sont assez subtils, mais imposent sur $\tilde{\mathcal{X}}$ une condition s'exprimant en termes géométriques simples, à savoir sa non-intersection avec une hypersurface \mathcal{O} de l'espace d'état fonction de l'équation des contraintes. L'expérience montre que cette variété a souvent la forme d'un « V ». Comme $\tilde{\mathcal{X}}$ est par hypothèse un paralléléotope orienté selon les axes principaux de l'espace d'état, son appartenance à un demi-espace délimité par \mathcal{O} empêche de considérer de nombreuses situations relatives caméra-cible initiales pourtant parfaitement pertinentes pour le problème d'asservissement visuel.

III.3.2.4 Extensions exigées par le contexte Robotique

Deux extensions significatives de la méthode précédente ont été proposées [JSI-2-AV-1]. Celles-ci demeurent ancrées dans les concepts de linéarisation globale et de stabilité quadratique, mais permettent de contourner tout ou partie des limitations mentionnées ci-dessus. Nous en expliquons les grandes lignes, puis caractérisons leur complexité. Notons que leur utilité peut largement déborder du contexte de l'asservissement visuel.

A Construction incrémentale d'un bassin d'attraction multicritère dans un contexte d'analyse Ce premier prolongement suppose la donnée *a priori* d'un bassin d'attraction multicritère $\tilde{\mathcal{E}}$ pour le problème considéré. Typiquement, $\tilde{\mathcal{E}}$ est la réunion $\cup_r(\tilde{\mathcal{E}})_r$ d'ellipsoïdes élémentaires optimisés indépendamment, telle que proposée au §III.3.1.2-B. La brique de base consiste à compléter $\tilde{\mathcal{E}}$ par un sous-ensemble d'un paralléléotope $\tilde{\Pi}$ avec lequel il présente une intersection non nulle. Cette opération élémentaire peut alors être répétée récursivement jusqu'à ce qu'elle conduise à un rajout insignifiant ou à un échec.

Outre les hypothèses triviales $0 \in \tilde{\mathcal{E}}$, $0 \notin \tilde{\Pi}$, $\tilde{\mathcal{E}} \cap \tilde{\Pi} \neq \emptyset$ et $\tilde{\mathcal{E}} \cap \tilde{\Pi} \neq \tilde{\mathcal{E}}$, on admet que la portion $(\partial\tilde{\mathcal{E}} \cap \tilde{\Pi})$ de la frontière $\partial\tilde{\mathcal{E}}$ de $\tilde{\mathcal{E}}$ située dans $\tilde{\Pi}$ est connexe (Figure III.5-gauche). Soit [LPV] le système incertain obtenu par linéarisation globale de [BFNL] sous l'hypothèse $\tilde{\mathcal{X}} = \tilde{\Pi}$. Il s'en suit que les trajectoires de [LPV] recouvrent les portions situées dans $\tilde{\Pi}$ des trajectoires de [BFNL]. Soit $V(\tilde{x}) = \tilde{x}'P\tilde{x}$ une fonction de LYAPUNOV pour [LPV]. Les portions de trajectoires de [BFNL] situées dans $\tilde{\Pi}$ intersectent alors au fil du temps des ellipsoïdes emboîtés correspondant à des valeurs décroissantes de $V(\cdot)$. Supposons que P soit tel que l'ellipsoïde $\tilde{\mathcal{E}}_1 \triangleq \{\tilde{x} : V(\tilde{x}) \leq 1\}$ ne franchisse la frontière $\partial\tilde{\Pi}$ de $\tilde{\Pi}$ qu'une fois, *i.e.* que $(\tilde{\mathcal{E}}_1 \cap \partial\tilde{\Pi})$ soit connexe. Dès lors, toute trajectoire de [BFNL] émanant de $(\tilde{\mathcal{E}}_1 \cap \tilde{\Pi})$ franchit $(\tilde{\mathcal{E}}_1 \cap \partial\tilde{\Pi})$. Si $(\tilde{\mathcal{E}}_1 \cap \partial\tilde{\Pi})$ est contenu dans $\tilde{\mathcal{E}}$, alors toute trajectoire de [LPV] émanant de $(\tilde{\mathcal{E}}_1 \cap \tilde{\Pi})$ converge vers 0. Il suffit d'exiger en outre que $(\tilde{\mathcal{E}}_1 \cap \tilde{\Pi})$ soit contenu dans la zone de l'espace d'état admissible vis à vis des contraintes, *i.e.* $(\tilde{\mathcal{E}}_1 \cap \tilde{\Pi}) \subset \tilde{\mathcal{A}}_j$, pour procéder à la réinitialisation de $\tilde{\mathcal{E}}$ selon $\tilde{\mathcal{E}} \leftarrow \tilde{\mathcal{E}} \cup (\tilde{\mathcal{E}}_1 \cap \tilde{\Pi})$.

L'ensemble des conditions précédentes sont transformées en des LMIs en P et d'autres variables *via* la S-procédures. Accessoirement, l'étendue de la portion rajoutée selon certaines directions peut être maximisée au moyen d'un critère adéquat.

Cette extension présente plusieurs avantages déterminants. D'une part, elle permet d'obtenir un bassin d'attraction multicritère dissymétrique et non convexe, pouvant mieux s'inscrire dans le sous-espace admissible de l'espace d'état. De plus, la petite taille de $\tilde{\Pi}$ limite le conservatisme de la linéarisation globale et réduit les problèmes de bien posé. Enfin, la complexité

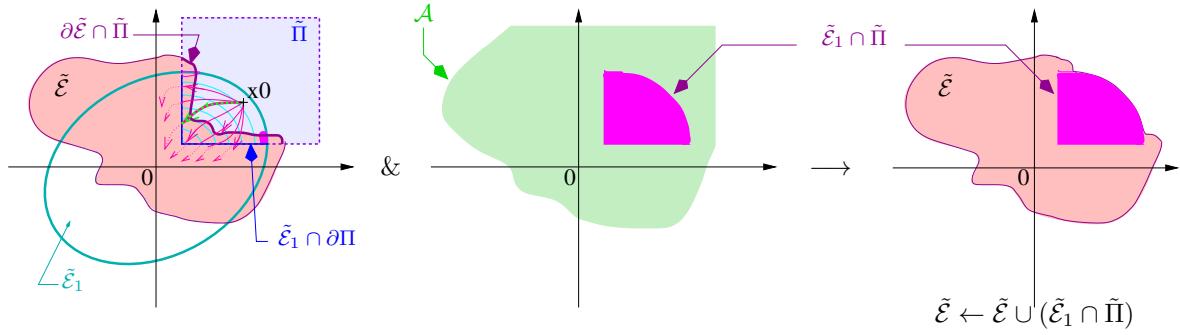


FIGURE III.5 – Extension de l'approche quadratique pour l'analyse multicritère.

demeure raisonnable du fait que des LMIs sont systématiquement obtenues. Inversement, l'implémentation de cette extension est relativement laborieuse, et difficile à imaginer pour des ordres supérieurs à 3.

B Obtention d'un bassin d'attraction multicritère asymétrique par introduction d'une variable fictive L'idée est ici d'appliquer la méthode « de base » du §III.3.2.2 au système $\overline{[BFNL]}$ défini comme l'augmentation de la boucle fermée $[BFNL]$ par une variable fictive \tilde{x}_f en évolution autonome selon l'équation

$$\dot{\tilde{x}}_f = -\alpha_f \tilde{x}_f, \quad \alpha_f > 0, \quad \tilde{x}_f(0) = \tilde{x}_f^0 > 0. \quad (\text{III.30})$$

Soit $\overline{[LPV]}$ le système incertain obtenu par linéarisation globale de $\overline{[BFNL]}$. La stabilité globale de $\overline{[LPV]}$ peut être étudiée au moyen d'une fonction de LYAPUNOV $V(\bar{x}) = \bar{x}' \bar{P} \bar{x}$, où $\bar{x} \triangleq (\bar{x}' \ \tilde{x}_f)'$. Du fait que $\tilde{x}_f(t)$ vit dans $[0; \tilde{x}_f^0]$ à tout instant t et converge vers 0 lorsque $t \rightarrow +\infty$, toute trajectoire $\bar{x}(.)$ de $\overline{[LPV]}$ émanant de la « tranche » $\overline{\mathcal{T}} \triangleq (\overline{\mathcal{E}} \cap \{\bar{x} \in \mathbb{R}^{(\bar{n}+1)} : \tilde{x}_f \in [0; \tilde{x}_f^0]\})$ de l'ellipsoïde $\overline{\mathcal{E}} \triangleq \{\bar{x} : V(\bar{x}) \leq 1\}$ demeure dans $\overline{\mathcal{T}}$ et converge vers l'origine de $\mathbb{R}^{(\bar{n}+1)}$. Par conséquent, toute trajectoire $\tilde{x}(.)$ de $\overline{[LPV]}$ débutant dans la projection $\tilde{\mathcal{P}} \triangleq \{\tilde{x} \in \mathbb{R}^{\bar{n}} : \exists \tilde{x}_f \in [0; \tilde{x}_f^0], \bar{x} \in \overline{\mathcal{T}}\}$ de $\overline{\mathcal{T}}$ sur $\mathbb{R}^{\bar{n}}$ demeure dans $\tilde{\mathcal{P}}$ et converge vers 0 (Figure III.6). Grâce à cette astuce, l'ensemble $\tilde{\mathcal{P}}$ peut parfaitement ne pas être symétrique par rapport à 0. Pour qu'il constitue un bassin d'attraction multicritère de la boucle fermée originale $[BFNL]$, il suffit d'assurer qu'il soit à la fois contenu dans $\tilde{\mathcal{X}}$ et dans les ensembles $\tilde{\mathcal{A}}_j$ admissibles du point de vue des contraintes.

Ces considérations se prêtent également à l'obtention de conditions suffisantes ayant la forme d'inégalités matricielles par la S-procédure. Dans un contexte d'analyse, il s'agit de LMIs ou de BMIs selon que les paramètres \tilde{x}_f^0 et α_f sont fixes ou constituent des variables de décision, respectivement. La méthode peut également être utilisée à des fins de synthèse. Cependant, contrairement aux résultats précédents reportés Figure III.2, la synthèse d'un asservissement 3D – *i.e.* d'un retour d'état – linéaire ou non linéaire est un problème de complémentarité sur le cône si \tilde{x}_f^0 et α_f sont fixes, et conduit à des BMIs sinon. La synthèse d'une commande 2D selon cette méthode conduit systématiquement à des BMIs.

Cette extension est plus générique que la précédente car elle s'applique aussi bien en analyse qu'en synthèse quel que soit l'ordre de l'asservissement, permet l'obtention de bassins d'attraction dissymétriques, et réduit également des problèmes de bien posé lors du traitement

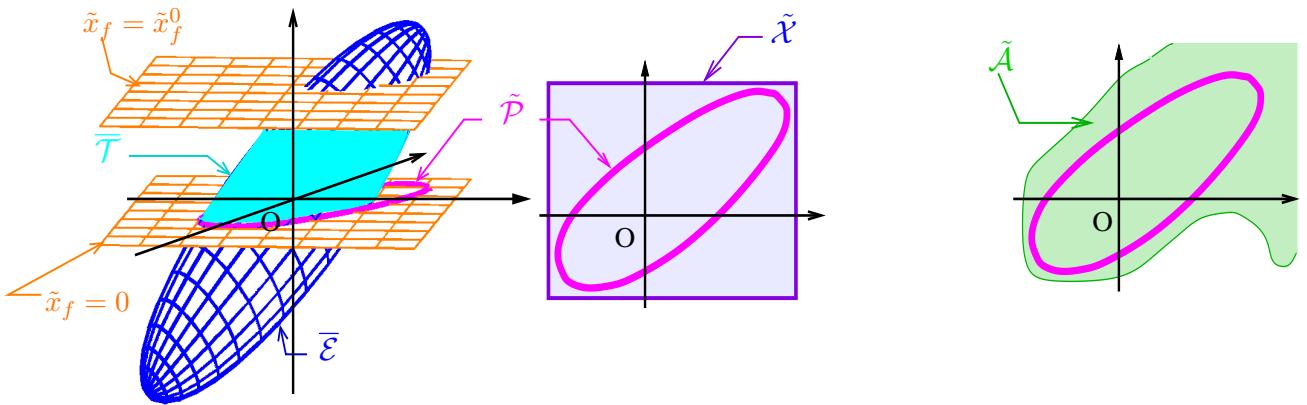


FIGURE III.6 – Bassin d'attraction multicritère obtenu par introduction d'une variable fictive pour l'analyse et la synthèse.

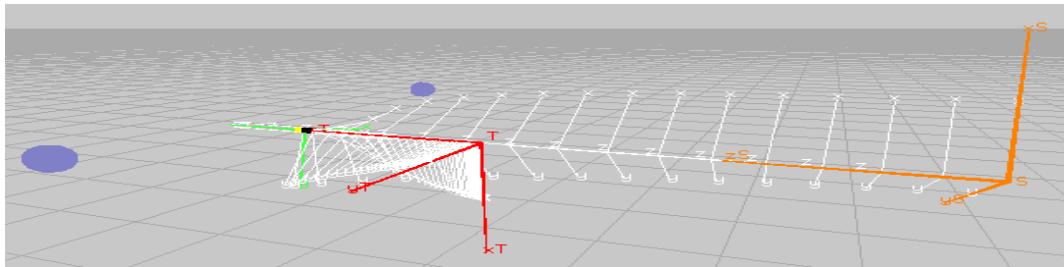


FIGURE III.7 – Asservissement 2D d'une caméra à 2ddl.

de contraintes. Ses inconvénients sont naturellement liés à sa complexité, et éventuellement à la convexité des bassins obtenus.

III.3.2.5 Illustration de l'approche quadratique et de ses extensions

Nous présentons deux cas d'étude.

A Analyse d'un asservissement 2D d'une caméra à deux degrés de liberté La Figure III.7 schématise une caméra pouvant se déplacer le long et autour de son axe optique. Celle-ci est commandée en boucle fermée sur la base de la projection dans son plan image d'une cible munie de 2 points, disposée orthogonalement à l'axe optique. Le contrôleur "image-based" est statique et admet pour équation $u = -\lambda[J_{2D}(s^*, z^*)]^\dagger(s - s^*)$, où λ est fixé à une valeur positive donnée et $J_{2D}(\cdot, \cdot)$ satisfait $\dot{s} = J_{2D}(s, z)u$. Le but est de déterminer un bassin d'attraction multicritère $\tilde{\mathcal{E}}$ de taille maximale en présence de contraintes de visibilité et de saturations d'actionneurs. Malgré son apparence simplicité, ce problème est ouvert.

La modélisation d'état vérifie donc $u = (v_z \omega_z)'$ et $\tilde{x} = x = (t_z N = \tan(\frac{\pi}{4}))'$. La Figure III.8 présente les bassins d'attraction multicritères obtenus par application de la méthode « de base ». L'espace d'état est muni de l'axe horizontal t_z et de l'axe vertical N . La situation finale désirée correspond au point $(0 0)'$. Les courbes situées à gauche délimitent les zones

admissibles vis à vis des contraintes. Le bassin présenté dans les trois tracés supérieurs est l'ellipsoïde de taille maximale obtenu par une application de la méthode « de base ». On note par exemple qu'il ne contient pas la condition – situation relative – initiale $\tilde{x}_0 = (2.5 \ 0)'$ du fait que $-\tilde{x}_0$ n'est pas admissible par rapport aux contraintes, alors qu'un raisonnement élémentaire démontre la convergence de la caméra depuis \tilde{x}_0 . Le bassin reporté sur les trois tracés inférieurs consiste en la réunion $\cup_r(\tilde{\mathcal{E}})_r$ de plusieurs ellipsoïdes optimisés séparément. Bien qu'il effleure certaines contraintes, son extension demeure très limitée.

La Figure III.9 illustre la réduction du conservatisme apportée par la méthode du §III.3.2.4-A. On note que $\cup_r(\tilde{\mathcal{E}})_r$ est considérablement augmenté. La méthode permet également d'établir une approximation interne du bassin d'attraction multicritère maximal pour la commande $u = -\lambda[J_{2D}(s^*, \hat{z}^*)]^+(s - s^*)$, où \hat{z}^* désigne une approximation de la profondeur z^* de la cible en la situation de référence. Elle conduit à des résultats intéressants, mais qui ne peuvent néanmoins pas prendre en compte la stationnarité de l'erreur commise lors de l'approximation.

B Synthèse d'un asservissement 3D d'une caméra à trois degrés de liberté On considère une caméra à 3 degrés de liberté, embarquée sur un robot mobile se déplaçant en translation et en rotation sur sol plan. Son torseur cinématique et la situation relative capteur-cible s'écrivent donc $u = (v_y \ v_z \ \omega_x)'$ et $x = (t_y \ t_z \ L=\tan \frac{\alpha}{2})'$. Celle-ci doit être positionnée face à une cible munie de 4 points coplanaires disposés sur un rectangle (Figure III.10). Dans la situation finale désirée de la caméra, l'axe optique est donc orthogonal au plan de la cible, et la projection de celle-ci est centrée dans le plan image. Des contraintes « raisonnables » sont placées sur le déplacement 3D – rester dans un couloir – et sur les actionneurs. L'objectif est de synthétiser un asservissement 3D de la caméra permettant la réalisation de la tâche avec le taux de convergence α maximal.

Pour que le problème soit soluble, il est nécessaire de relâcher de manière irréaliste la contrainte de visibilité. Les contraintes LMI du problème d'optimisation sont alors faisables seulement pour des situations caméra-cible initiales très voisines de 0 (Figure III.10-droite). La Figure III.11 caractérise le comportement obtenu. On note que l'ellipsoïde invariant sous-tendant la synthèse est effectivement inscrit dans les zones de l'espace d'état admissible du point de vue des contraintes, mais que leur dissymétrie limite considérablement la méthode.

L'extension §III.3.2.4-B appliquée à ce cas d'étude permet l'obtention d'une loi de com-

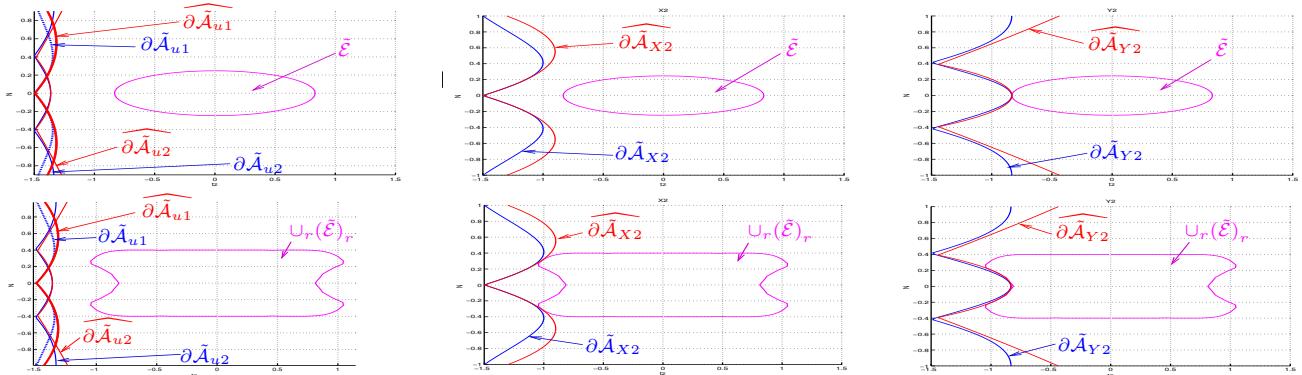


FIGURE III.8 – Analyse multicritère par la méthode « de base » du §III.3.2.2.

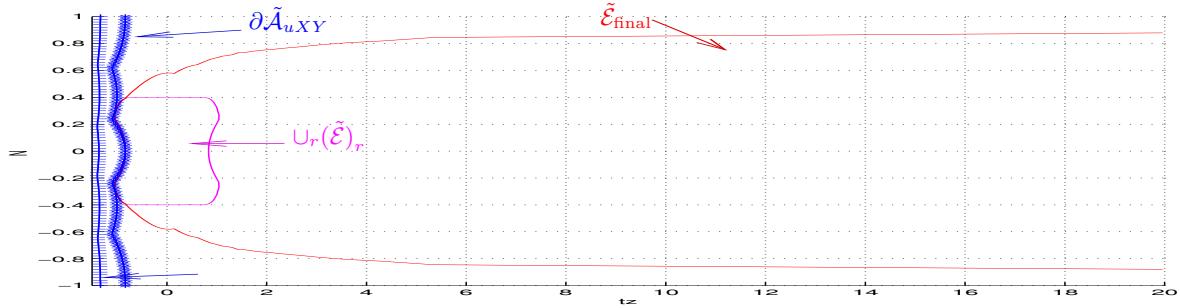


FIGURE III.9 – Analyse multicritère par la méthode du §III.3.2.4–A.

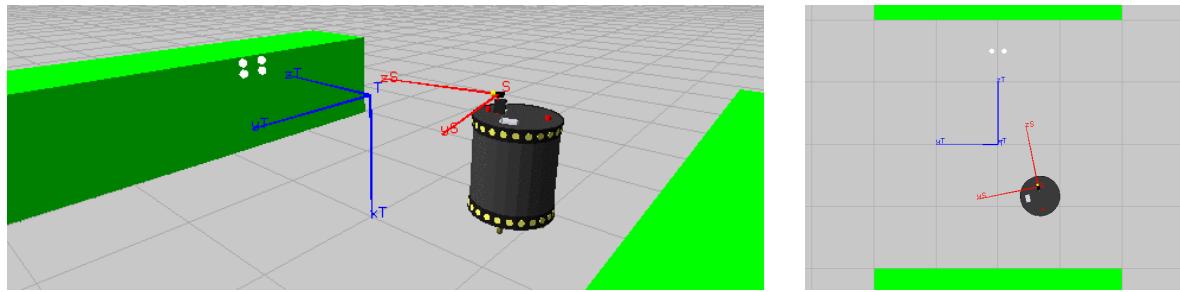


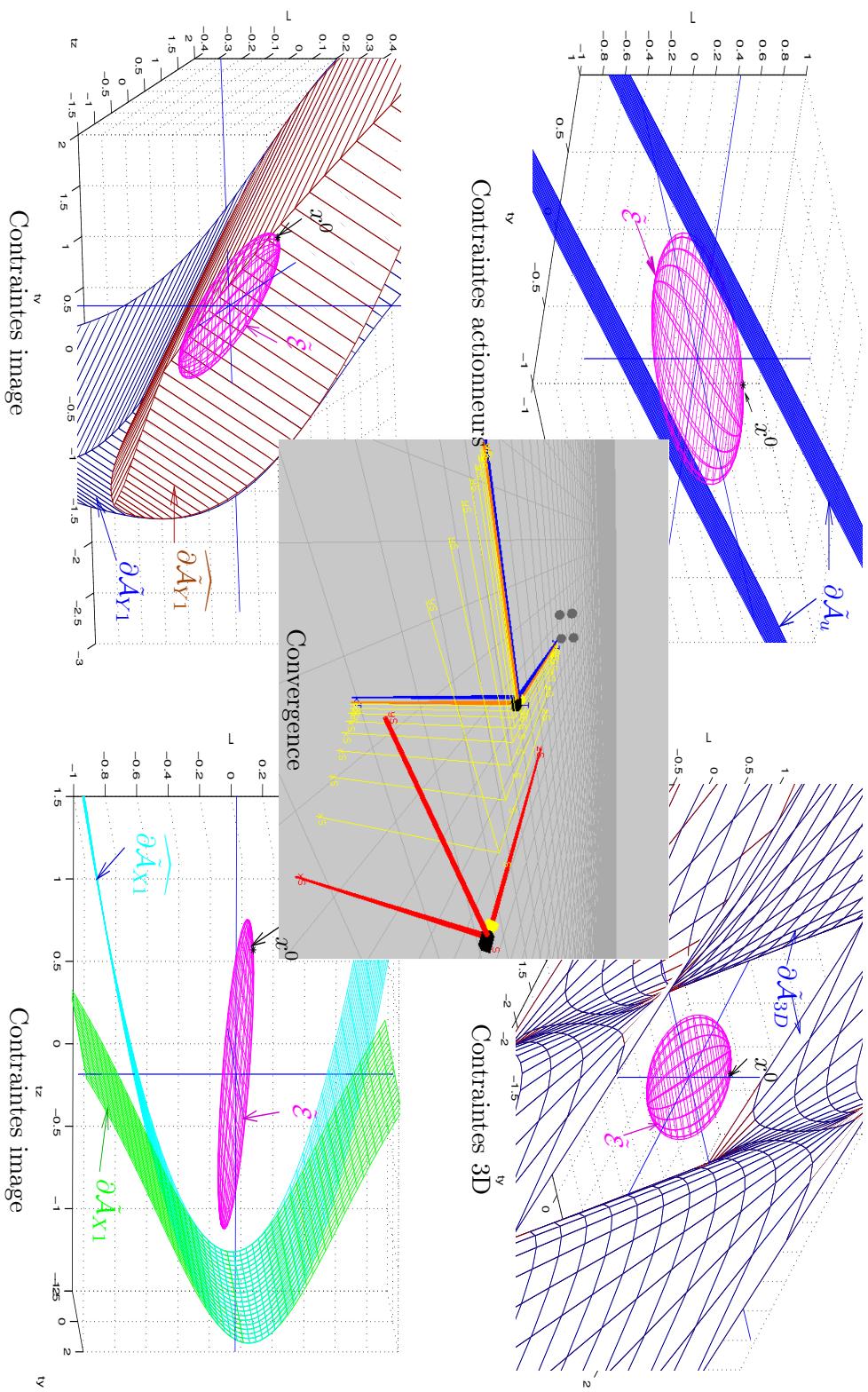
FIGURE III.10 – Asservissement 3D d'une caméra à 3ddl.

mande 3D statique linéaire pour une contrainte de visibilité réaliste, et pour une situation initiale de la caméra beaucoup plus éloignée de la cible (Figure III.12). Le temps de réponse est tout à fait crédible, et peut même être diminué en synthétisant une commande 3D statique rationnelle.

III.3.3 Représentations algébro-différentielles et fonctions de Lyapunov plus évoluées pour l'analyse multicritère

Outre leurs sources de pessimisme fondamentales – exploitation de la plongée de [BFNL] (eq. (III.23)) dans l'approximation externe (III.29) de la SNLDI (III.28), conservatisme éventuel de la S-procédures –, les travaux fondateurs sur lesquels s'appuie la méthode précédente présentent des limitations plus ponctuelles. Parmi celles-ci, nous avons cité la nécessité de définir des fonctions de LYAPUNOV sur des parallélétopes orientés selon la base canonique de l'espace d'état. Un autre écueil est la restriction à des fonctions de LYAPUNOV quadratiques indépendantes de l'incertitude, avec pour effet de bord l'impossibilité de prendre en compte une quelconque connaissance *a priori* sur la dynamique de celle-ci : existence de paramètres incertains constants, de dérivée bornée, etc.

Les travaux de TROFINO et COUTINHO [Trofino 02][Coutinho 04] pour l'analyse des systèmes rationnels incertains contournent ces deux dernières limitations. Dans leur version initiale, ils permettent la détermination d'un bassin d'attraction sur la base de fonctions de LYAPUNOV biquadratiques en le vecteur d'état et possiblement quadratiques en l'incertitude. Nos contributions, développées en collaboration avec Daniel F. COUTINHO puis dans le cadre de la thèse de Sylvain DUROLA [DOC-05], ont consisté à les étendre à la prise en compte de contraintes rationnelles en le vecteur d'état et les incertitudes [SIACL-2-AV-1] ainsi qu'à l'introduction de fonctions de LYAPUNOV biquadratiques par morceaux [CIACL-11-AV-4].



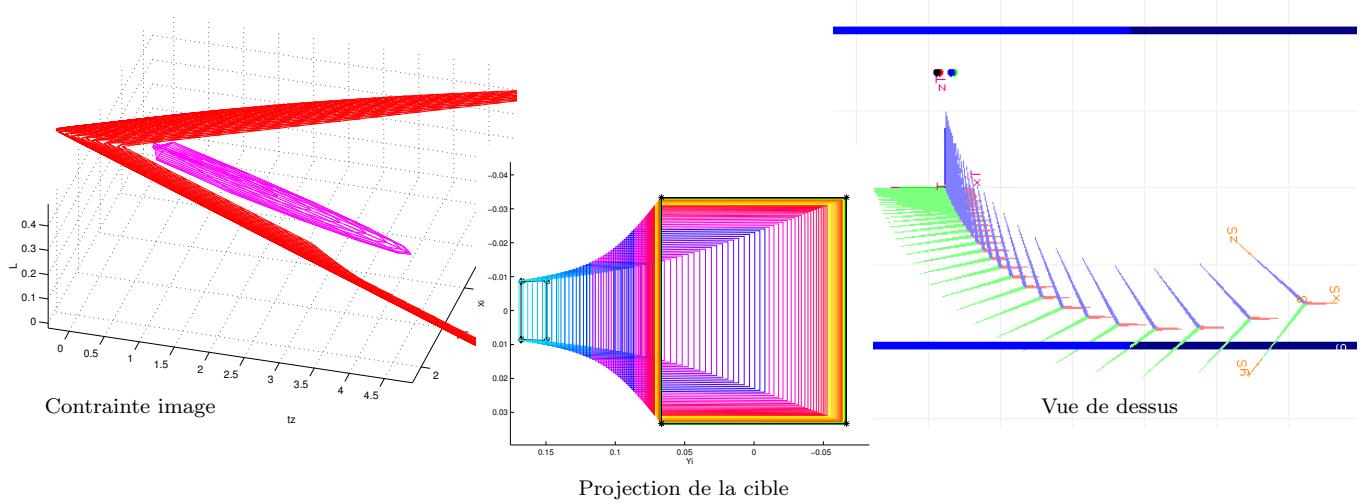


FIGURE III.12 – Synthèse multicritère par la méthode du §III.3.2.4-B.

Nous en décrivons les grandes lignes ci-dessous, ainsi que les conclusions obtenues pour notre problème de Robotique [CIACL-18-AV-6][COCI-1-AV-1]. Comme cela était le cas pour la méthode précédente, nous n’entrons pas dans les détails techniques. Néanmoins, nous présentons les idées principales dans un contexte sensiblement plus large, permettant l’extension de la méthode à des fonctions de LYAPUNOV plus complexes, telles des fonctions de LYAPUNOV polyquadratiques ou rationnelles, éventuellement par morceaux.

III.3.3.1 Terminologie

L’approche repose sur la représentation équivalente suivante de la boucle fermée référencée vision [BFNL]. Les polytopes $\tilde{\mathcal{X}}$ et \mathcal{X}_χ , donnés, sont supposés convexes et contiennent 0.

Définition III.6 (Représentation Algébro-Différentielle [Trofino 02][Coutinho 04])
La Représentation Algébro-Différentielle (DAR = Differential Algebraic Representation) d’un système non linéaire rationnel incertain est définie par

$$\begin{cases} \dot{\tilde{x}} = A_1 \tilde{x} + A_2 \pi \\ 0 = \Omega_1(\tilde{x}, \chi) \tilde{x} + \Omega_2(\tilde{x}, \chi) \pi, \end{cases} \quad (\text{III.31})$$

où $\pi = \pi(\tilde{x}, \chi) \in \mathbb{R}^{n_\pi}$ est une fonction vectorielle non linéaire de (\tilde{x}, χ) , A_1, A_2 sont des matrices constantes, et $\Omega_1(\cdot, \cdot), \Omega_2(\cdot, \cdot)$ sont des fonctions matricielles affines en leurs arguments. La représentation ci-dessus est dite bien posée si $\Omega_2(\cdot, \cdot)$ est de rang plein $\forall (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi$.

Une DAR est en fait d’un sous-ensemble des systèmes descripteurs non linéaires. De la même manière, les variables additionnelles $\zeta_j = Z_j'(\tilde{x}, \chi)\tilde{x}$ – cf. (III.12) – se réécrivent en [SIACL-2-AV-1]

$$\begin{cases} \zeta_j = K_{1j}' \tilde{x} + K_{2j}' \varrho \\ 0 = \Upsilon_{1j}(\tilde{x}, \chi) \tilde{x} + \Upsilon_{2j}(\tilde{x}, \chi) \varrho, \end{cases} \quad (\text{III.32})$$

où $\varrho = \varrho(\tilde{x}, \chi)$ est une fonction vectorielle non linéaire de (\tilde{x}, χ) , K_{1j}, K_{2j} désignent des vecteurs constants, et $\Upsilon_{1j}(\cdot, \cdot), \Upsilon_{2j}(\cdot, \cdot)$ sont des fonctions matricielles affines, avec $\Upsilon_{2j}(\cdot, \cdot)$ de rang plein sur $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$.

La méthodologie d'obtention des modèles (III.31) et (III.32) consiste à rassembler tous les termes non linéaires et/ou incertains de [BFNL] et ζ_j dans π et ϱ , respectivement, puis à relier ces vecteurs – en les augmentant si besoin – au vecteur d'état \tilde{x} par des équations algébriques admettant la structure requise. La boucle fermée [BFNL] admet une forme LFT – dans la veine de (III.25)–(III.26) pour [LPV] – ainsi qu'une DAR équivalente. Celles-ci s'écrivent

$$\begin{array}{c} [\text{BFNL}] \\ \Updownarrow \\ \left\{ \begin{array}{l} \dot{\tilde{x}} = \left(\tilde{A}^\sharp + \tilde{B}_p^\sharp \Delta (\mathbb{I} - \tilde{D}_{qp}^\sharp \Delta(\tilde{x}, \chi))^{-1} \tilde{C}_q^\sharp \right) \tilde{x} \\ \Delta(\tilde{x}) = \text{diag}(\tilde{x}_1 \mathbb{I}_{r_1}, \dots, \tilde{x}_{\tilde{n}} \mathbb{I}_{r_{\tilde{n}}}, \chi_1 \mathbb{I}_{r_{(\tilde{n}+1)}}, \dots, \chi_{n_\chi} \mathbb{I}_{r_{(\tilde{n}+n_\chi)}}) \end{array} \right. \end{array} \quad (\text{III.33})$$

$$\begin{array}{c} \Updownarrow \\ \left\{ \begin{array}{l} \dot{\tilde{x}} = A_1 \tilde{x} + A_2 \pi, \quad 0 = \Omega_1(\tilde{x}, \chi) \tilde{x} + \Omega_2(\tilde{x}, \chi) \pi \\ A_1 = \tilde{A}^\sharp, \quad A_2 = \tilde{B}_p^\sharp, \quad \Omega_1 = \Delta(\tilde{x}, \chi) \tilde{C}_q^\sharp, \quad \Omega_2 = \Delta(\tilde{x}, \chi) \tilde{D}_{qp}^\sharp - \mathbb{I}. \end{array} \right. \end{array} \quad (\text{III.34})$$

Tout comme une LFT, une DAR n'est pas unique. Toutefois, le pessimisme de l'approche présentée ci-après ne présente pas de lien évident avec la minimalité des vecteurs π et ϱ dans (III.31) et (III.32).

III.3.3.2 Fonctions de Lyapunov candidates

La classe des fonctions de LYAPUNOV candidates, définies sur $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$, s'écrit comme suit :

$$\forall (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \quad V(\tilde{x}, \chi) = \phi'_1(\tilde{x}, \chi) \mathsf{P} \phi_1(\tilde{x}, \chi), \quad \phi_1(\tilde{x}, \chi) \triangleq (\Theta'(\tilde{x}, \chi) \mathbb{I}_{\tilde{n}})' \tilde{x}, \quad (\text{III.35})$$

où la matrice symétrique P est une variable de décision, et $\Theta(., .)$ désigne une fonctions matricielle donnée rationnelle en ses arguments. Pour des raisons évidentes, la dérivée temporelle $\dot{\phi}_1(\tilde{x}, \chi)$ du vecteur $\phi_1(\tilde{x}, \chi)$ le long des trajectoires de [BFNL] est une fraction rationnelle en \tilde{x}, χ, π , de sorte qu'elle admet une DAR équivalente

$$\left\{ \begin{array}{l} \dot{\phi}_1 = \begin{pmatrix} C_1 \\ A_1 \end{pmatrix} \tilde{x} + \begin{pmatrix} \mathbb{O} \\ A_2 \end{pmatrix} \pi + \begin{pmatrix} C_2 \\ \mathbb{O} \end{pmatrix} \nu \\ 0 = \begin{pmatrix} \Lambda_1(\tilde{x}, \chi) \\ \Omega_1(\tilde{x}, \chi) \end{pmatrix} \tilde{x} + \begin{pmatrix} \mathbb{O} \\ \Omega_2(\tilde{x}, \chi) \end{pmatrix} \pi + \begin{pmatrix} \Lambda_2(\tilde{x}, \chi) \\ \mathbb{O} \end{pmatrix} \nu. \end{array} \right. \quad (\text{III.36})$$

III.3.3.3 Vers l'obtention d'un bassin d'attraction multicritère

Les notations précédemment introduites permettent de réécrire les conditions (III.14)–(III.15) de définition positivité de $V(., .)$ et de décroissance de sa dérivée le long des trajectoires de [BFNL] en deux équations de la forme

$$\forall (\tilde{x}, \chi) \in (\tilde{\mathcal{X}} \setminus \{0\}) \times \mathcal{X}_\chi, \quad \phi'_1(\tilde{x}, \chi) \mathsf{P} \phi_1(\tilde{x}, \chi) > 0 \quad (\text{III.37})$$

$$\forall (\tilde{x}, \chi) \in (\tilde{\mathcal{X}} \setminus \{0\}) \times \mathcal{X}_\chi, \quad \phi'_2(\tilde{x}, \chi) \mathcal{M}_2(\mathsf{P}) \phi_2(\tilde{x}, \chi) < 0, \quad (\text{III.38})$$

où $\phi_1(\tilde{x}, \chi) \triangleq (\Theta'(\tilde{x}, \chi) \mathbb{I}_{\tilde{n}})' \tilde{x}$, $\phi_2(\tilde{x}, \chi) \triangleq (\phi'_1(\tilde{x}, \chi) \pi'(\tilde{x}, \chi) \nu'(\tilde{x}, \chi))'$, et $\mathcal{M}_2(.)$ désigne une fonction matricielle affine en son argument.

La condition (III.16) (resp. (III.17)), nécessaire à la définition d'un bassin d'attraction multicritère, s'exprime comme une relation d'inclusion ensembliste entre des ensembles définis

comme une forme quadratique en $\phi_3(\tilde{x}, \chi) \triangleq (1 \phi'_1(\tilde{x}, \chi))'$ (resp. $\phi_4(\tilde{x}, \chi) \triangleq (1 \phi'_1(\tilde{x}, \chi) \varrho'(\tilde{x}, \chi))'$). L'application de la S-procédures conduit alors à des conditions suffisantes de la forme

$$\exists \tau_{3,k} \geq 0 : \forall (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \phi'_3(\tilde{x}, \chi) \mathcal{M}_3(\mathsf{P}, \tau_{3,k}) \phi_3(\tilde{x}, \chi) \leq 0, \quad (\text{III.39})$$

$$\exists \tau_{4,j} \geq 0 : \forall (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \phi'_4(\tilde{x}, \chi) \mathcal{M}_4(\mathsf{P}, \tau_{4,j}) \phi_4(\tilde{x}, \chi) \leq 0, \quad (\text{III.40})$$

où $\{\tau_{3,k}\}$ et $\{\tau_{4,j}\}$ désignent des multiplicateurs scalaires, et les fonctions matricielles $\mathcal{M}_3(., .)$ et $\mathcal{M}_4(., .)$ sont affines en leurs arguments.

La LMI suivante implique évidemment (III.37)–(III.38)–(III.39)–(III.17) :

$$\mathsf{P} > 0, \mathcal{M}_2(\mathsf{P}) < 0, \mathcal{M}_3(\mathsf{P}, \tau_{3,k}) \leq 0, \mathcal{M}_4(\mathsf{P}, \tau_{4,j}) \leq 0. \quad (\text{III.41})$$

Cependant, il s'agit d'une condition suffisante très conservative car elle ne prend pas en compte le lien existant entre les composantes de chaque vecteur $\phi_i(\tilde{x}, \chi)$, $i = 1, \dots, 4$. Or, il est possible d'exhiber des fonctions matricielles $\Psi_1(., .), \dots, \Psi_4(., .)$ affines en leur arguments telles que

$$\forall i \in \{1, \dots, 4\}, \forall (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \Psi_i(\tilde{x}, \chi) \phi_i(\tilde{x}, \chi) = 0. \quad (\text{III.42})$$

Ces fonctions matricielles sont par conséquent appelées *annihilateurs linéaires* des vecteurs correspondants. Leurs expressions font naturellement intervenir $\Omega_1(., .)$, $\Omega_2(., .)$, $\Upsilon_{1j}(., .)$, $\Upsilon_{2j}(., .)$, $\Lambda_1(., .)$, $\Lambda_2(., .)$, et peuvent être complétées par des lignes non redondantes établies *ad hoc*. Une application du lemme fondamental suivant conduit alors immédiatement à des conditions suffisantes moins pessimistes de (III.37)–(III.38)–(III.39)–(III.17).

Lemme III.2 ([Trofino 02][Coutinho 04]) Soient les vecteurs $\tilde{x} \in \tilde{\mathcal{X}}$ et $\chi \in \mathcal{X}_\chi$, où $\tilde{\mathcal{X}}, \mathcal{X}_\chi$ désignent des polytopes convexes donnés. Soit la fonction matricielle $\Sigma_{\tau_l, \mathsf{P}, \dots}(., .)$, symétrique, affine en ses arguments et en les variables de décision $\tau_l, \mathsf{P}, \dots$. On considère la contrainte suivante en une fonction vectorielle non linéaire donnée $\sigma(., .)$:

$$\forall (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \sigma'(\tilde{x}, \chi) \Sigma_{\tau_l, \mathsf{P}, \dots}(\tilde{x}, \chi) \sigma(\tilde{x}, \chi) \leq 0. \quad (\text{III.43})$$

Par convexité, si les LMIs $\Sigma_{\tau_l, \mathsf{P}, \dots}(\tilde{x}, \chi) \leq 0$ en $\tau_l, \mathsf{P}, \dots$ sont en vigueur en tous les sommets de $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$, alors elles le sont également sur $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$ et (III.43) est satisfaite. Cependant, s'il existe une fonction matricielle $\Lambda_\sigma(., .)$ affine en ses arguments telle que $\forall (\tilde{x}, \chi) \in \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \Lambda_\sigma(\tilde{x}, \chi) \sigma(\tilde{x}, \chi) = 0$ –i.e. s'il existe $\Lambda_\sigma(., .)$ annihilateur linéaire de $\sigma(., .)$ sur $\tilde{\mathcal{X}} \times \mathcal{X}_\chi$ – alors la condition suffisante suivante de (III.43) est moins conservative :

$$\begin{aligned} \exists \mathsf{L}, \text{ multiplicateur matriciel, tel qu'en tous sommets } (\tilde{x}, \chi) \text{ de } \tilde{\mathcal{X}} \times \mathcal{X}_\chi, \\ \text{la LMI suivante en les variables } \tau_l, \mathsf{P}, \dots, \mathsf{L} \text{ est satisfaite :} \end{aligned} \quad (\text{III.44})$$

$$\bar{\Sigma}_{\tau_l, \mathsf{P}, \dots, \mathsf{L}}(\tilde{x}, \chi) \triangleq \Sigma_{\tau_l, \mathsf{P}, \dots}(\tilde{x}, \chi) + \mathsf{L} \Lambda_\sigma(\tilde{x}, \chi) + \Lambda_\sigma'(\tilde{x}, \chi) \mathsf{L}' \leq 0.$$

III.3.3.4 Remarques importantes

En conclusion, l'analyse multicritère se réduit à un ensemble de conditions suffisantes LMIs dont la taille et le nombre de variables de décision augmente avec la complexité de $\Theta(., .)$. Une fonction $\Theta(., .)$ plus complexe peut conduire à des conclusions moins pessimistes, du fait de la richesse de la famille des fonctions de LYAPUNOV considérées, mais au risque d'un mauvais conditionnement numérique des programmes d'optimisation obtenus. Pour ces raisons, nous

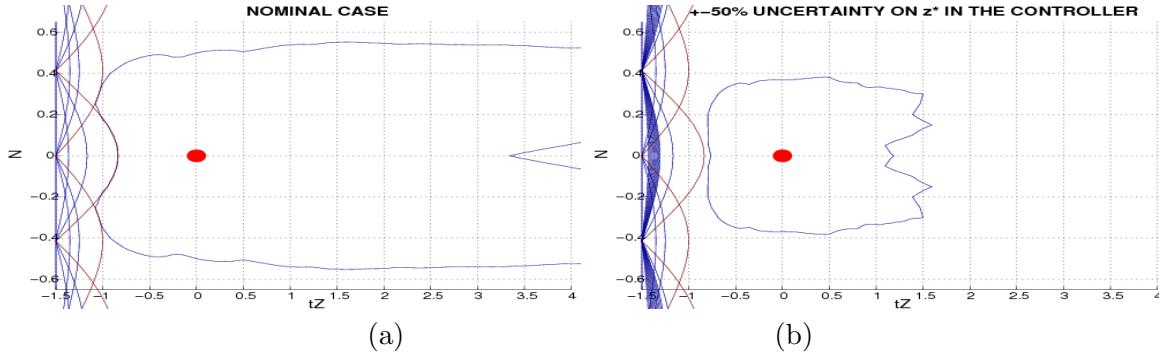


FIGURE III.13 – Analyse multicritère du cas d'étude §III.3.2.5-A sur la base de DARs et de fonctions de LYAPUNOV biquadratiques : (a) cas nominal ; (b) incertitude constante de $\pm 50\%$ sur la profondeur z^* exploitée dans le contrôleur 2D. $(0, 0)'$ désigne la situation capteur-cible désirée. Les contraintes sont matérialisées sur la gauche (et pour divers χ dans le cas (b)).

avons exclusivement considéré une fonction $\Theta(.,.)$ affine en ses arguments, de sorte que $V(.,.)$ est bi-quadratique. Néanmoins, un encadrement *a priori* de χ peut être mis à profit dans l'approche de façon à obtenir des contraintes LMI moins conservatives.

Comme alternative à la construction d'un bassin d'attraction multicritère en tant que la réunion de bassins calculés séparément et dont l'étendue est optimisée selon diverses directions de l'espace, nous avons également envisagé des fonctions de LYAPUNOV biquadratiques par morceaux. Le principe est assez classique. On partitionne $\tilde{\mathcal{X}}$ en régions polytopiques adjacentes $\tilde{\mathcal{X}}_1, \dots, \tilde{\mathcal{X}}_S$ contenant 0 et dont les frontières sont de dimension $(\tilde{n} - 1)$. La fonction de LYAPUNOV recherchée est définie sur $\tilde{\mathcal{X}}$ de telle sorte que sa restriction à chaque partition $\tilde{\mathcal{X}}_s$ soit une fonction biquadratique $V_s(.,.)$ de variable de décision associée P_s . Pour assurer la stabilité locale du point d'équilibre 0 de [BFNL], on exige de $V_s(.,.)$ qu'elle soit positive sur son domaine de définition, décroissante le long des trajectoires du système, et que sa valeur et sa dérivée le long des trajectoires du système soient continues aux frontières de $\tilde{\mathcal{X}}_s$ avec celles des fonctions biquadratiques adjacentes. Le bassin d'attraction $\tilde{\mathcal{E}} \triangleq \{\tilde{x} : V(\tilde{x}, \chi) \leq 1, \forall \chi \in \mathcal{X}_\chi\}$ s'écrit naturellement comme l'union $\tilde{\mathcal{E}} = \bigcup_{s \in \{1, \dots, S\}} \tilde{\mathcal{E}}_s$, où, pour chaque s , $\tilde{\mathcal{E}}_s \triangleq \{\tilde{x} \in \tilde{\mathcal{X}}_s : V_s(\tilde{x}, \chi) \leq 1, \forall \chi \in \mathcal{X}_\chi\}$. Il reste à exiger que chaque $\tilde{\mathcal{E}}_s$ soit inclus dans $\tilde{\mathcal{X}}$ et dans les zones de l'espace d'état admissibles vis à vis des contraintes pour garantir que $\tilde{\mathcal{E}}$ est un bassin d'attraction multicritère de [BFNL].

III.3.3.5 Application à un cas d'étude

Considérons à nouveau l'analyse de l'asservissement 2D d'une caméra à deux degrés de liberté présenté au §III.3.2.5-A. Dans un premier temps, $\tilde{\mathcal{E}}$ est défini comme la réunion $\cup_r (\tilde{\mathcal{E}})_r$ de bassins d'attraction multicritères optimisés séparément, sur la base de fonctions de LYAPUNOV $V(\tilde{x}) = (t_z^2, t_z N, N^2, t_z, N)' P(t_z^2, t_z N, N^2, t_z, N)$. Les résultats, présentés Figure III.13, montrent une amélioration par rapport à ceux établis Figure III.8-bas dans les mêmes conditions. Une approximation interne du bassin d'attraction multicritère maximal pour la commande $u = -\lambda[J_{2D}(s^*, \hat{z}^*)]^+(s - s^*)$, où $\hat{z}^* = z^*(1 + \chi)$ désigne une approximation de la profondeur z^* de la cible en la situation de référence est également reportée Figure III.13 pour $\chi \in [-0.5; +0.5]$. La fonction de LYAPUNOV $V(\tilde{x}, \chi) = (t_z^2, t_z N, N^2, t_z, N, \chi)' P(t_z^2, t_z N, N^2, t_z, N, \chi)$ dépend de l'incertitude, ce qui réduit

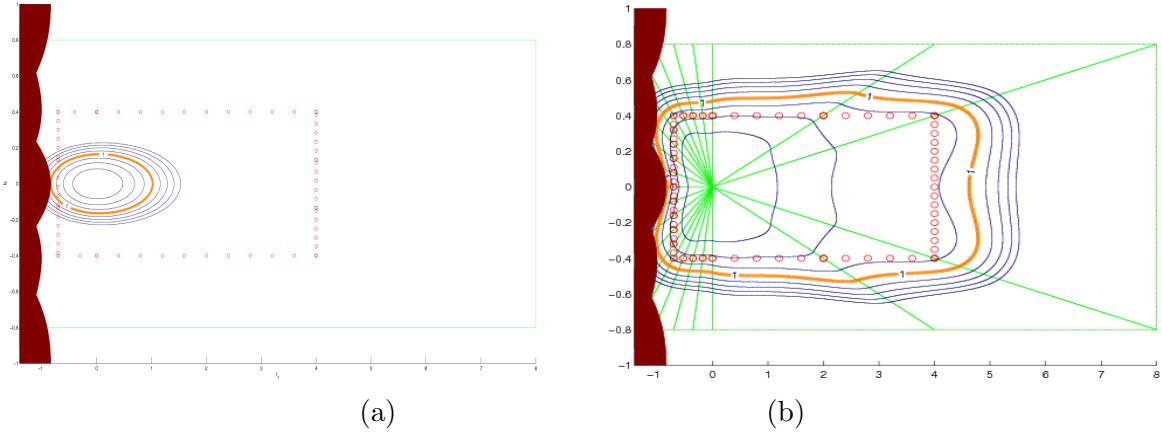


FIGURE III.14 – Bassin d'attraction multicritère, étendu vers les points matérialisés par des cercles, et obtenu par une fonction de LYAPUNOV (a) biquadratique, (b) biquadratique par morceaux.

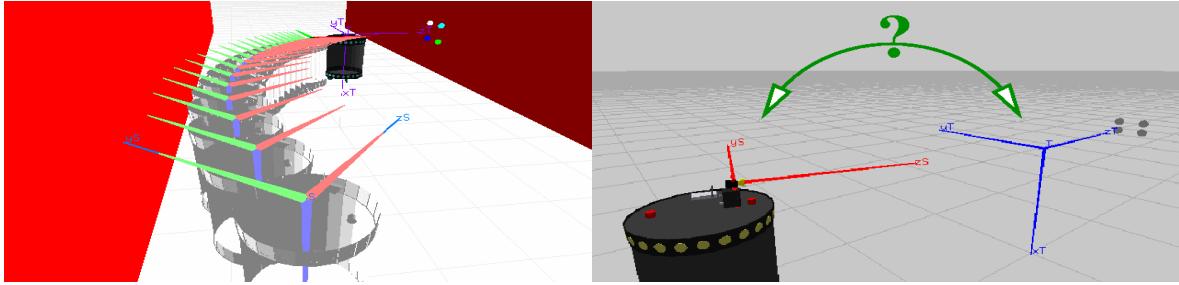


FIGURE III.15 – Dualité entre asservissement visuel (gauche) et localisation visuelle (droite).

le conservatisme de l'approximation, et la stationnarité de χ est explicitement prise en compte dans l'élaboration des conditions LMI.

Bien que les bassins élémentaires ($\tilde{\mathcal{E}}_r$) n'aient pas été optimisés dans la construction de $\tilde{\mathcal{E}}$ reportée Figure III.13, les résultats semblent encore limités. La Figure III.14 compare le bassin d'attraction multicritère obtenu dans le cas nominal précédent par des fonctions de LYAPUNOV biquadratiques et biquadratiques par morceaux, lorsque celui-ci est étendu simultanément selon plusieurs directions distinctes. L'exploitation de fonctions de LYAPUNOV définies par morceaux semble constituer une alternative intéressante car elle offre de nombreux degrés de liberté pour une complexité algorithmique raisonnable.

III.4 Approche de la localisation visuelle en tant que le dual de l'asservissement visuel

III.4.1 Introduction

Notre formulation du problème de commande référencée vision 2D consiste à stabiliser la situation relative d'une caméra mobile par rapport à une cible fixe, sur la base des informations qu'elle délivre sur celle-ci. À ce problème de stabilisation d'un vecteur d'état par retour de

96 \III.4. APPROCHE DE LA LOCALISATION VISUELLE EN TANT QUE LE DUAL DE L'ASSERVISSEMENT VISUEL

sortie correspond le problème dual – au sens de la dualité observation-commande – de sa reconstruction sur la base des valeurs passées et présentes des vecteurs de mesure et de commande. En d'autres termes, il s'agit de reconstruire la pose d'une cible fixe à partir des historiques des mesures visuelles délivrées par une caméra mobile et des torseurs cinématiques appliqués à cette caméra (Figure III.15).

Au-delà de son caractère « esthétique », cette propriété de dualité originale suggère d'approcher la localisation visuelle par des techniques duales de celles précédemment considérées pour l'asservissement visuel. Ainsi, le problème s'inscrit naturellement dans le cadre du filtrage des systèmes rationnels incertains.

III.4.2 Aperçu de la formalisation mathématique du problème

Nous nous sommes placés dans le contexte du filtrage ensembliste [CIACL-06-AV-3]. Disposant à l'instant k d'un ellipsoïde de confiance $\mathcal{E}(k)$ entourant la situation relative caméra-cible réelle – inconnue – $\mathbf{x}(k)$, et connaissant la mesure visuelle $\mathbf{y} = \mathbf{s} - \mathbf{s}^*$, on recherche l'ellipsoïde $\mathcal{E}(k+1)$ de taille minimale entourant les valeurs possibles $\mathbf{x}(k+1)$ à l'instant $k+1$.

Plus formellement, on suppose la donnée *a priori* d'un modèle rationnel à temps discret du système en boucle ouverte

$$[\text{BOz}] : \begin{cases} \mathbf{x}(k+1) = \mathbf{A}_k(\mathbf{x}(k), \chi(k))\mathbf{x}(k) + \mathbf{B}_k(\mathbf{x}(k), \chi(k))\mathbf{u}(k) + \mathbf{G}_k(\mathbf{x}(k), \chi(k))\mathbf{w}(k) \\ \mathbf{y}(k) = \mathbf{C}_k(\mathbf{x}(k), \chi(k))\tilde{\mathbf{x}}(k) + \mathbf{H}_k(\mathbf{x}(k), \chi(k))\mathbf{v}(k). \end{cases} \quad (\text{III.45})$$

L'équation d'état s'obtient par exemple par discréttisation de l'équation d'état de la boucle ouverte continue (III.9) sous l'hypothèse d'un blocage d'ordre zéro de la commande. Le bruit de dynamique additif $\mathbf{w}(k)$ se réalise dans un ellipsoïde de confiance de caractéristiques données. L'équation de mesure est la même que dans le cas continu. Outre d'éventuelles incertitudes rationnelles sur le modèle paramétrique de la cible ou sur les paramètres intrinsèques de la caméra, elle peut également comprendre un bruit de mesure additif $\mathbf{v}(k)$ caractérisé sous forme d'un ensemble elliptique. Un paralléléotope $\mathcal{X}_\chi(k)$ des valeurs admissibles de $\chi(k)$ est également donné.

Conceptuellement, le principe de la solution comporte deux étapes. Dans un premier temps, on isole les triplets (situation relatives caméra-cible $\mathbf{x}(k) \in \mathcal{E}(k)$; réalisation des incertitudes dans le modèle de mesure; réalisation du bruit de mesure) compatibles avec la mesure visuelle $\mathbf{y}(k)$ prélevée à l'instant k . Une prédiction des valeurs $\mathbf{x}(k+1)$ à l'instant suivant $k+1$ est ensuite effectuée depuis les $\mathbf{x}(k)$ ainsi isolés, pour toutes les réalisations admissibles des incertitudes et du bruit apparaissant dans l'équation d'état[†]. La recherche d'un ellipsoïde de taille minimale encerclant l'ensemble ainsi obtenu conduit à $\mathcal{E}(k+1)$.

Une extension mineure de l'algorithme de filtrage ensembliste des systèmes linéaires autonomes discrets à incertitude rationnelle proposé par EL GHAOUI et CALAFIORE [El Ghaoui 01] permet la mise en œuvre des idées précédentes. On détermine dans un premier temps le plus petit paralléléotope $\Xi(k)$ – orienté selon la base canonique de l'espace d'état – encerclant $\mathcal{E}(k)$, ce qui ne pose aucune difficulté. Ensuite, on linéarise globalement [BOz] sous l'hypothèse $\mathbf{x}(k) \in \Xi_k$, *i.e.* on remplace formellement dans les arguments des fonctions matricielles [BOz] le vecteur $\mathbf{x}(k)$ par un paramètre incertain $\delta^\chi(k)$ vivant dans Ξ_k . La forme LFT du système à paramètres variants $[\text{LPV}_k]$ ainsi obtenu se réécrit comme l'interconnexion d'un système

[†]. En fait, la méthode est sensiblement plus générique car elle permet que les équations d'état et de mesure partagent les mêmes incertitudes et/ou sources de bruit. Seules les réalisations de ces quantités partagées isolées lors de la première étape participent alors à la prédiction du vecteur d'état à l'instant $k+1$.

linéaire à temps discret avec un gain matriciel incertain, ou *Inclusion Récurrente Linéaire Structurée et Bornée en Norme* (SNLRI = Structured Norm-Bounded Linear Recursive Inclusion). L'immersion de cette SNLRI dans une approximation externe permet l'obtention d'une solution LMI au schéma de filtrage, par application de la S-procédure. Comme pour la commande, l'obtention d'une forme LFT minimale de $[LPV_k]$ est essentielle à la limitation du conservatisme des conclusions.

III.4.3 Illustration sur des cas d'étude

Considérons la localisation d'une cible constituée de 2 ou 3 points, disposée orthogonalement à l'axe optique d'une caméra immobile. La situation relative caméra-cible est représentée par le vecteur $\mathbf{x} = (t_z \ N \ \tan(\frac{\nu}{4}))'$, où, comme précédemment, t_z et N désignent respectivement la translation et la rotation relatives caméra-cible.

La Figure III.16 présente les résultats de localisation obtenus lorsque l'extraction des indices visuels est entâchée d'un bruit. La situation relative réelle est matérialisée par le point noté $\mathbf{x}(0)$, et le centre de l'ellipsoïde de confiance $\mathcal{E}(0)$ donné a priori est noté $\hat{\mathbf{x}}(0)$. À chaque instant k , la résolution du programme d'optimisation LMI conduit à l'ellipsoïde $\mathcal{E}(k)$ de centre $\hat{\mathbf{x}}(k)$. On reconnaît également la famille des parallélétopes $\Xi(\cdot)$ circonscrits aux $\mathcal{E}(\cdot)$, exploités pour la linéarisation globale.

La comparaison de la ligne supérieure et inférieure montre qu'il est absolument capital de disposer d'une LFT minimale de $[BOz]$. Dans le cas contraire, l'approximation externe de la SNLDI est très conservatrice, de sorte qu'à chaque instant un nombre élevé de valeurs de son vecteur d'état sont compatibles avec la mesure visuelle. Ce conservatisme augmente d'autant plus que le caractère non linéaire de $[BOz]$ est prononcé – *e.g.* lorsque le nombre

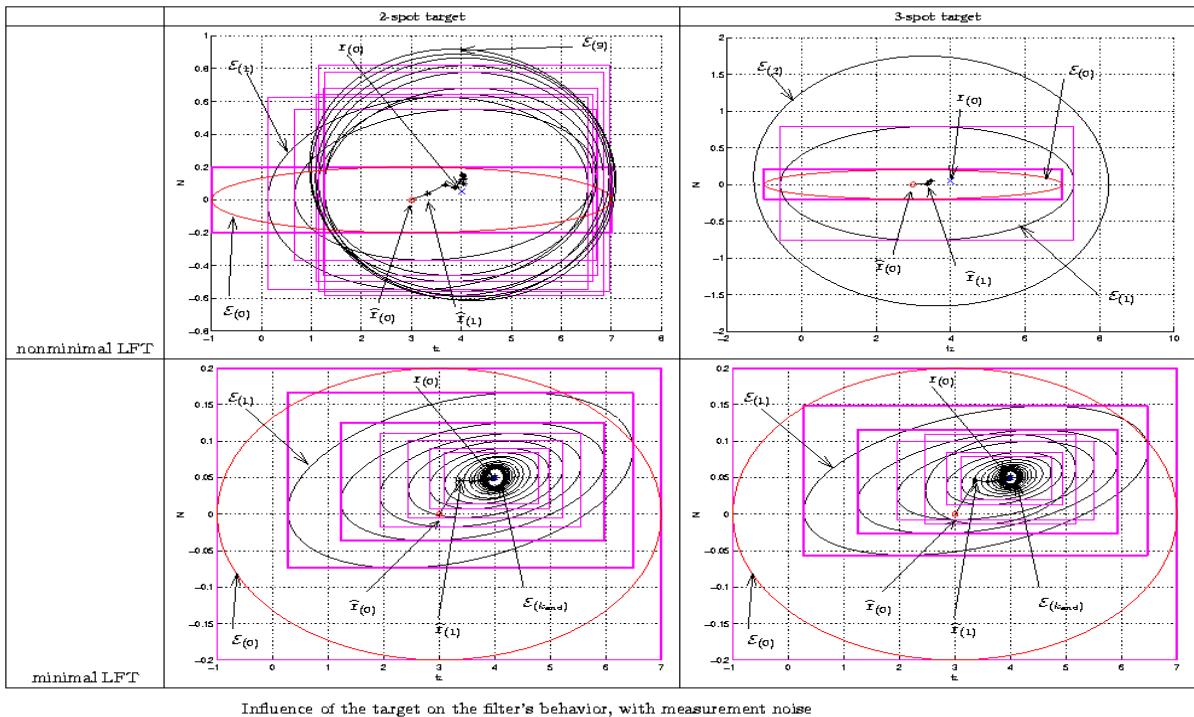


FIGURE III.16 – Localisation visuelle d'une cible par une caméra immobile.

98 \III.5. CONCEPTION ET IMPLÉMENTATION D'UN ENVIRONNEMENT DE SIMULATION DE COMMANDES EN ROBOTIQUE

d'indices visuels augmente – de sorte que les résultats sont pires pour une cible munie du nombre maximal de points.

En revanche, lorsqu'une LFT minimale de [BOz] est exploitée, l'observation des deux colonnes de la ligne inférieure fait apparaître deux propriétés extrêmement intéressantes. D'une part, les ellipsoïdes de confiance convergent vers un « régime permanent », d'autant plus regroupé autour de la vraie situation que la cible est informative. En outre, la vitesse de convergence vers ce régime permanent augmente avec le nombre de points.

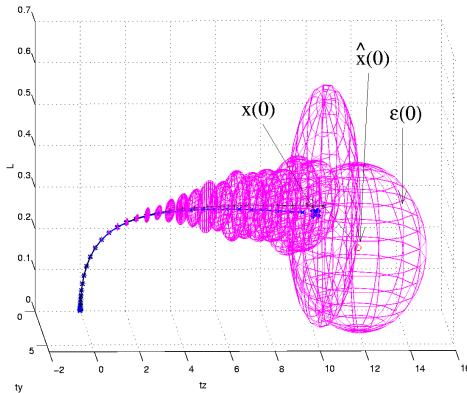


FIGURE III.17 – Localisation visuelle d'une cible par une caméra à 3 ddl.

quelques itérations de l'algorithme de façon à réutiliser la mesure $y(k)$ en supposant que la commande est nulle. Ainsi, $\mathcal{E}(k)$ peut être réduit en un ellipsoïde $\mathcal{E}'(k)$ – comme cela a été effectué dans le premier cas d'étude –, préalablement à la propagation de ce dernier en $\mathcal{E}(k+1)$ sur la base de $y(k)$, de $u(k)$ et du modèle de dynamique.

Par ailleurs, nous avons développé dans [CIACL-17-AV-5] une méthode alternative de filtrage ensembliste des systèmes rationnels incertains. Le principe est le même que celui, rappelé au §III.4.2, des travaux de EL GHAOUI et CALAFIORE, cependant sa mise en œuvre est différente quoique procédant en deux temps. [BOz] est d'abord réécrit sous la forme d'une Représentation Algébro-Récurrente (RAR = Recursive Algebraic Representation), puis, sur cette base, une condition LMI permettant la définition de l'ellipsoïde $\mathcal{E}(k+1)$ recherché est obtenue par application du Lemme III.2 page 93. L'application de ce résultat théorique à la localisation visuelle est en suspens.

III.5 Conception et implémentation d'un environnement de simulation de commandes en robotique

Depuis plusieurs années, Matthieu HERRB, collègue ingénieur de recherche au LAAS-CNRS, développe le logiciel GDHE (Graphical Display for Hilare Experiments, <http://www.laas.fr/~matthieu/gdhe>). Il s'agit d'un serveur d'affichage, construit autour de OpenGL, et acceptant des requêtes écrites en Tcl. Cet outil est très utilisé afin de visualiser des expérimentations robotiques, du fait qu'il encapsule des primitives de dessin dans des routines élémentaires, permettant ainsi de dessiner rapidement un objet articulé.

Une autre illustration de la méthode est proposée pour la localisation visuelle d'une cible munie de 4 points coplanaires non alignés par rapport à une caméra à 3 degrés de liberté telle que celle considérée au §III.3.2.5-B, lorsque celle-ci est asservie visuellement vers cette cible. L'évolution temporelle du vecteur d'état $x = (t_y \ t_z \ L=\tan \frac{\lambda}{2})'$ et de l'ellipsoïde de confiance associé sont reportés Figure III.17. On note que l'algorithme possède des propriétés de convergence intéressantes.

Pour conclure, il convient de mentionner deux extensions de la méthode. L'une, immédiate, s'applique dès lors que la caméra est en mouvement, *e.g.* dans le second cas d'étude présenté ci-dessus. Du fait que le problème est posé dans le contexte de la localisation ensembliste, il est possible d'insérer à chaque instant k

Nous avons procédé à l’interfaçage de MATLAB-SIMULINK avec GDHE dans le but de simuler une grande variété de tâches robotiques incluant des capteurs visuels et/ou proximétriques (lasers) en bénéficiant de la souplesse et de la performance de SIMULINK pour la simulation, ainsi que des capacités de GDHE/OpenGL pour le dessin, le rendu 3D, et la manipulation de transformations homogènes. Notre environnement MAVS (MAVS = “MAVS Ain’t Visual Servoing”) repose ainsi sur SIMULINK pour la modélisation des dynamiques des éléments de la scène, et fournit à l’utilisateur un ensemble de blocs lui permettant de définir leur géométrie, de les visualiser sous GDHE, voire de récupérer depuis GDHE – la communication étant bidirectionnelle – la matrice homogène exprimant leur situation. À titre d’exemple, il est possible de construire une chaîne articulée puis, au cours de la simulation, de modifier sa position depuis SIMULINK, de visualiser son évolution en 3D sous GDHE, et de récupérer la situation de son organe terminal calculée par OpenGL sans calcul explicite de son modèle géométrique. Tout élément peut embarquer une ou plusieurs caméras perspectives et/ou capteurs lasers, dont les caractéristiques sont paramétrables. Les possibilités d’interfaçage de SIMULINK permettent d’exploiter au fil de la simulation des routines de traitement de l’information extéroceptive codées en C/C++. Afin de garantir un résultat fiable, l’utilisateur doit seulement s’assurer qu’à chaque “minor time step” de la simulation SIMULINK, les opérations suivantes soient effectuées en séquence : 1/positionnement de tous les objets, 2/positionnement de tous les capteurs, 3/demande explicite de réaffichage du contenu de tous les observateurs (fenêtre de visu principale et capteurs extéroceptifs) sous GDHE, 4/exécution des routines de traitement des informations extéroceptives. Ceci exige de jongler avec les mécanismes d’affectation des priorités sous SIMULINK, ce qui peut entraîner quelques frustrations...

Enfin, quatre blocs génériques (`object2GDHE`, `configurableObject2GDHE`, `object2GDHEwithTrack`, `configurableObject2GDHEwithTrack`) permettent l’exploitation depuis SIMULINK de nouveaux objets graphiques définis dans un fichier Tcl. Cette possibilité est illustrée sur la Figure III.18.

Ce logiciel est actuellement utilisé de manière interne par quelques collègues de RIA. Un ultime « toilettage » est prévu avant sa diffusion en Open Source.

III.6 Références

- [Allibert 06] G. Allibert, E. Courtial & F. Chaumette. *Visual Servoing via Non-linear Predictive Control*. Chapitre 20 de *Visual Servoing via Advanced Numerical Methods*, G. CHESI & K. HASHIMOTO Eds, pp. 169–188. Lecture Notes in Control and Information Sciences. Springer, Vol. 401/2010
- [Bourquardez 06] O. Bourquardez, R.E. Mahony, T. Hamel & F. Chaumette. *Stability and Performance of Image Based Visual Servo Control using First Order Spherical Image Moments*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS’2006), pages 4304–4309, 2006.
- [Boyd 94] S. Boyd, L. El Ghaoui, E. Feron & V. Balakrishnan. Linear matrix inequalities in system and control theory. SIAM, 1994.
- [Chaumette 98] F. Chaumette. *Potential problems of stability and convergence in image-based and position-based visual servoing*. In The Confluence of Vision and Control, LNCIS series, volume 237, pages 66–78. Springer, 1998.

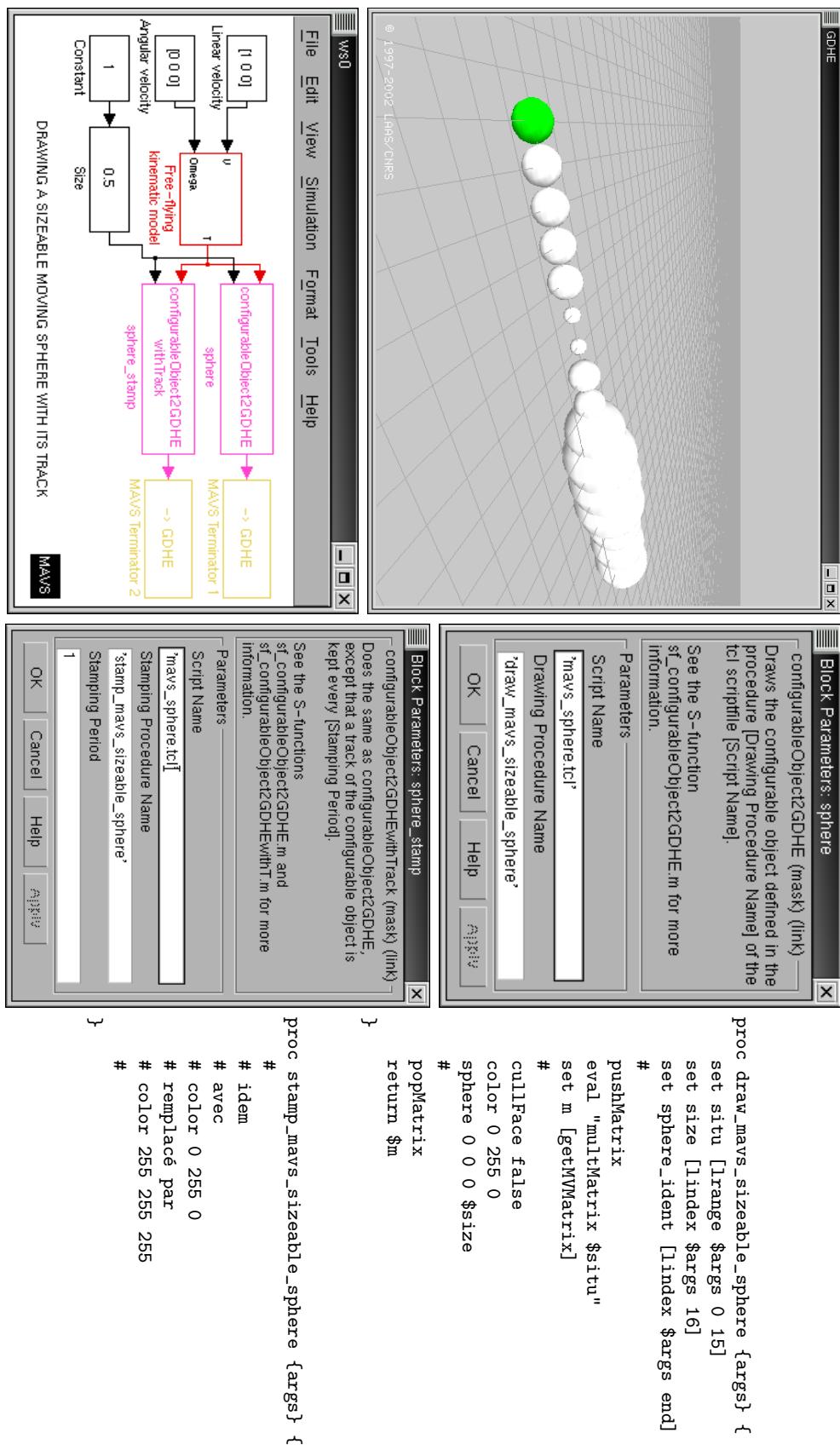


FIGURE III.18 – Aperçu des possibilités d'extension de MAVS : adjonction d'une sphère de taille modifiable depuis SIMULINK et de son « fantôme » (pour visualisation de son évolution spatio-temporelle à une période définie par l'utilisateur).

- [Chaumette 04] F. Chaumette. *Image Moments : A General and Useful Set of Features for Visual Servoing*. IEEE Trans. on Robotics, vol. 20, no. 4, pages 713–723, 2004.
- [Chaumette 06] F. Chaumette & S. Hutchinson. *Visual Servo Control, Part I : Basic Approaches*. IEEE Robotics and Automation Magazine, vol. 13, no. 4, pages 82–90, 2006.
- [Chaumette 07] F. Chaumette & S. Hutchinson. *Visual Servo Control, Part II : Advanced Approaches*. IEEE Robotics and Automation Magazine, vol. 14, no. 1, pages 109–118, 2007.
- [Chesi 04] G. Chesi, K. Hashimoto, D. Prattichizzo & A. Vicino. *Keeping features in the field of view in eye-in-hand visual servoing : a switching approach*. IEEE Trans. on Robotics, vol. 20, no. 5, pages 908–914, 2004.
- [Corke 96] P. Corke & M.C. Good. *Dynamic Effects in Visual Closed-Loop Systems*. IEEE Trans. on Robotics and Automation, vol. 12, no. 5, pages 671–683, 1996.
- [Coutinho 04] D.F. Coutinho, A.S. Bazanella, A. Trofino & A.S. Silva. *Stability Analysis and Control of a Class of Differential-Algebraic Nonlinear Systems*. Int. Jour. of Robust and Nonlinear Control, vol. 14, pages 1301–1326, 2004.
- [Cowan 02] N.J. Cowan, J.D. Weingarten & D.E. Koditschek. *Visual Servoing via Navigation Functions*. IEEE Trans. on Robotics and Automation, vol. 18, no. 4, pages 521–533, 2002.
- [Crétual 01] A. Crétual & F. Chaumette. *Visual Servoing Based on Image Motion*. Int. Journal of Robotics Research, vol. 20, no. 11, pages 857–877, 2001.
- [Deng 02] L. Deng, F. Janabi-Sharifi & W.J. Wilson. *Stability and Robustness of Visual Servoing Methods*. In IEEE Int. Conf. on Robotics and Automation (ICRA’02), pages 1605–1609, 2002.
- [Dussy 98a] S. Dussy. *Approche LMI de la commande robuste multicritère*. Thèse de doctorat, Université Paris-IX Dauphine, 1998.
- [Dussy 98b] S. Dussy & L. El Ghaoui. *Measurement-scheduled control for the RTAC problem : an LMI approach*. Int. Jour. of Robust and Non-linear Control, vol. 8, pages 377–400, 1998.
- [El Ghaoui 96] L. El Ghaoui & G. Scattolini. *Control of Rational Systems using Linear-Fractional Representations and Linear Matrix Inequalities*. Automatica, vol. 32, no. 9, pages 1273–1284, 1996.
- [El Ghaoui 97] L. El Ghaoui, F. Oustry & M. Ait-Rami. *A Cone Complementarity Linearization Algorithm for Static Output-Feedback and related Problems*. IEEE Trans. on Automatic Control, 1997.
- [El Ghaoui 01] L. El Ghaoui & G. Calafiore. *Robust Filtering for Discrete-Time Systems with Bounded Noise and Parametric Uncertainty*. IEEE Trans. on Automatic Control, vol. 46, no. 7, 2001.

- [Espiau 92] B. Espiau, F. Chaumette & P. Rives. *A New Approach to Visual Servoing in Robotics*. IEEE Trans. on Robotics and Automation, vol. 8, no. 3, pages 313–326, 1992.
- [Gangloff 03] J. Gangloff & M. de Mathelin. *High Speed Visual Servoing of a 6DOF Manipulator using Multivariable Predictive Control*. Advanced Robotics, vol. 17, no. 10, pages 993–1021, 2003.
- [Geromel91] J.C. Geromel, P.L.D. Peres & J. Bernussou. *On a Convex Parameter Space Method for Linear Control Design of Uncertain Systems*. SIAM Jour. on Control and Optimization, vol. 19, no. 2, pages 381–402, 1991.
- [Hashimoto 93] K. Hashimoto & H. Kimura. *LQ optimal and nonlinear approaches to visual servoing*. In K. Hashimoto, éditeur, *Visual Servoing*, volume 7 of *World Scientific Series in Robotics and Automated Systems*, pages 165–198. World Scientific Press, 1993.
- [Hashimoto 00] K. Hashimoto & T. Noritsugu. *Enlargement of Stable Region in Visual Servo*. In IEEE Conf. on Decision and Control, pages 3927–3932, 2000.
- [Hutchinson 96] S. Hutchinson, G.D. Hager & P.I. Corke. *A Tutorial on Visual Servo Control*. IEEE Trans. on Robotics and Automation, vol. 12, no. 5, pages 651–670, 1996.
- [Löfberg 04] J. Löfberg. *YALMIP : A Toolbox for Modeling and Optimization in MATLAB*. In IEEE CACSD Conf. Taipei, Taiwan, 2004.
- [Malis 99] E. Malis, F. Chaumette & S. Boudet. *2 $\frac{1}{2}$ 2D visual servoing*. IEEE Trans. on Robotics and Automation, vol. 15, no. 2, pages 234–246, 1999.
- [Malis 02] E. Malis & F. Chaumette. *Theoretical Improvements in the Stability Analysis of a New Class of Model-Free Visual Servoing Methods*. IEEE Trans. on Robotics and Automation, vol. 18, no. 2, pages 176–186, 2002.
- [Malis 03] E. Malis & P. Rives. *Robustness of Image-Based Visual Servoing with respect to Depth Distribution Error*. In IEEE Int. Conf. on Robotics and Automation (ICRA’03), pages 1056–1061, Taipei, Taiwan, 2003.
- [Martinet 98] P. Martinet. *Comparison of Visual servoing Techniques : Experimental Results*. In IAPR Workshop on Machine Vision Applications (MVA’98), pages 254–257, 1998.
- [Mezouar 03] Y. Mezouar & F. Chaumette. *Optimal Camera Trajectory with Image-Based Control*. Int. Jour. of Robotics Research, vol. 22, no. 10/11, 2003.
- [Papanikolopoulos 93] N.P. Papanikolopoulos, P.K. Khosla & T Kanade. *Visual Tracking of a Moving Target by a Camera Mounted on a Robot : A Combination of Control and Vision*. IEEE Trans. on Robotics and Automation, vol. 9, no. 1, pages 14–35, 1993.
- [Renaud 96] M. Renaud. *Comment définir l’orientation d’un corps ?* Rapport technique 96078, LAAS-CNRS, March 1996.
- [Samson 91] C. Samson, M. Le Borgne & B. Espiau. *Robot control : The task function approach*. Clarendon Press, Oxford, 1991.

- [Sauvée 06] M. Sauvée, P. Poignet, E. Dombre & E. Courtial. *Image Based Visual Servoing through Nonlinear Model Predictive Control*. In IEEE Conf. on Decision and Control (CDC'2006), pages 1776–1781, San Diego, CA, 2006.
- [Schramm 04] F. Schramm, G. Morel, A. Micaelli & A. Lottin. *Extended-2D visual servoing*. In IEEE Int. Conf. on Robotics and Automation (ICRA'04), pages 267–273, New Orleans, LA, 2004.
- [Scorletti 98] G. Scorletti & L. El Ghaoui. *Improved LMI Conditions for Gain Scheduling and Related Control Problems*. Int. Jour. of Robust and Nonlinear Control, vol. 8, pages 845–877, 1998.
- [Shirai 73] Y. Shirai & H. Inoue. *Guiding a Robot by Visual Feedback in Assembling Tasks*. Pattern Recognition, vol. 5, pages 99–108, 1973.
- [Sturm 99] J.F. Sturm. *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*. Optimization Methods and Software, vol. 11, pages 625–653, 1999.
- [Sznaier 00] M. Sznaier, B. Murphy & O.I. Camps. *An LPV Approach to Synthesizing Robust Active Vision Systems*. In IEEE Conf. on Decision and Control, pages 2545–2550, 2000.
- [Tarbouriech 00] S. Tarbouriech & P. Souères. *Advanced control strategy for the visual servoing scheme*. In IFAC Symp. on Robot Control (SYROCO'2000), pages 457–462, 2000.
- [Tarbouriech 05] S. Tarbouriech, P. Souères & B. Gao. *A Multicriteria Image-Based Controller based on a Mixed Polytopic and Norm-Bounded Representation of Uncertainties*. In IEEE Conf. on Decision and Control (CDC'2005), pages 5385–5390, 2005.
- [Thuilot 02] B. Thuilot, P. Martinet, L. Cordesses & J. Gallice. *Position based visual servoing : keeping the object in the field of vision*. In IEEE Int. Conf. on Robotics and Automation (ICRA'02), pages 1624–1629, 2002.
- [Toh 99] K.C. Toh, M.J. Todd & R.H. Tutuncu. *SDPT3 — A MATLAB Software Package for Semidefinite Programming*. Optimization Methods and Software, vol. 11, pages 545–581, 1999.
- [Trofino 00] A. Trofino. *Bi-Quadratic Stability for Nonlinear Systems*. In IFAC Symp. on Robust Control Design (ROCOND'00), Prague, 2000.
- [Trofino 02] A. Trofino. Local, Regional and Global Stability : An LMI Approach for Uncertain Nonlinear Systems. Tech. Report, DAS-CTC, Univ. Federal de Santa Catarina, Brésil, 2002.
- [Weiss 87] L. Weiss, A.C. Sanderson & C.P. Neuman. *Dynamic Sensor-Based Control of Robots with Visual Feedback*. IEEE Trans. on Robotics and Automation, vol. 3, no. 1, 1987.
- [Zanne 04] P. Zanne, G. Morel & F. Plestan. *Robust 3D Vision Based Control and Planning*. In IEEE Int. Conf. on Robotics and Automation (ICRA'04), pages 4423–4428, 2004.
- [Zhou 95] K. Zhou, J. Doyle & K. Glover. Robust and optimal control. Englewood Cliffs, 1995.

Chapitre IV

Audition en Robotique

Sommaire

IV.1 Introduction	106
IV.2 Contexte scientifique	106
IV.2.1 Analyse de Scènes Auditives Computationnelle et Audition en Robotique	106
IV.2.2 Positionnement de nos contributions	108
IV.2.2.1 Localisation, extraction et détection de sources sonores	108
IV.2.2.2 Ensemble matériel et logiciel pour l'audition en robotique	111
IV.2.2.3 Vers la détection d'activité vocale	111
IV.3 Contribution à l'audition en robotique	111
IV.3.1 Terminologie	111
IV.3.1.1 Généralités	111
IV.3.1.2 Antenne linéaire uniforme	113
IV.3.2 Techniques de formation de voie pour l'extraction et la localisation de sources sonores large bande	113
IV.3.2.1 Introduction à la formation de voie	113
IV.3.2.2 Formation de voie large bande « naïve » par optimisation convexe	115
IV.3.2.3 Analyse modale et optimisation convexe pour la formation de voie	116
IV.3.2.4 Implémentation pratique d'une formation de voie	120
IV.3.3 Méthode à haute résolution pour la localisation et la détection de sources sonores large bande	124
IV.3.3.1 Introduction	124
IV.3.3.2 Déclinaisons de la méthode MUSIC	125
IV.3.3.3 Une stratégie cohérente de localisation de sources adaptée à la Robotique	127
IV.3.3.4 Exploitation de MUSIC large bande cohérent pour la détection du nombre de sources	128
IV.3.3.5 Implémentation pratique du schéma de {détection,localisation}	130
IV.3.4 Vers la détection de “patterns” acoustiques	131
IV.3.5 Un capteur intégré embarquable basé sur un réseau de microphones pour l'audition en robotique	134
IV.4 Références	137

IV.1 Introduction

L’Audition en Robotique (“Robot Audition”) est une thématique relativement récente. Ainsi, ce n’est que depuis 2004 que des sessions dédiées sont organisées dans la conférence IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), véritable rendez-vous annuel de la communauté. Quasi-omnidirectionnelle et insensible aux conditions d’illumination, cette modalité perceptuelle constitue un complément idéal à la Vision, qui a fait l’objet de développements continus depuis plusieurs décennies. Son importance dans la thématique très actuelle de l’Interaction Homme-Robot est évidente, du fait des nombreuses perspectives offertes par un robot capable de détecter et localiser un locuteur se situant hors de son champ de vision, de le reconnaître, de le suivre, pour ensuite en interpréter les demandes et/ou dialoguer avec lui. Depuis Cog [Brooks 99], le nombre de robots embarquant un capteur auditif dédié a ainsi augmenté continûment. Les solutions *binaurales* initiales, exploitant deux transducteurs montés sur une tête – *e.g.* sur les prototypes SIG et SIG2 [Nakadai 00] –, sont souvent supplantes par des réseaux, ou *antennes*, cf. par exemple le robot Spartacus [Michaud 07].

IV.2 Contexte scientifique

Cette section propose une introduction à l’Audition en Robotique, en retracant l’apport historique de l’Analyse de Scènes Auditives Computationnelle à la discipline puis en listant ses spécificités. Nous ciblons alors les fonctions sur lesquelles ont porté nos contributions, et en dressons un bref état de l’art.

IV.2.1 Analyse de Scènes Auditives Computationnelle et Audition en Robotique

Notre système auditif nous offre la remarquable capacité de focaliser notre attention sur un locuteur d’intérêt en présence d’autres locuteurs, de bruit et de réverbérations, puis d’extraire et de comprendre son message. Connu sous le nom de *Cocktail Party Problem* (CPP), ce phénomène a suscité de nombreux travaux en physiologie, neurosciences et psychologie cognitive. Dans le but de conférer des capacités similaires à des machines, les premiers efforts ont principalement porté sur la reconnaissance automatique de parole (ASR = *Automatic Speech Recognition*). De multiples déclinaisons ont vu le jour, ancrées dans diverses techniques d’Intelligence Artificielle ou reposant sur des modèles statistiques de la parole. Comme leurs performances dans des environnements non contraints demeuraient très inférieures à celles du système auditif humain, un consensus est apparu sur la nécessité d’une compréhension plus profonde de celui-ci, non dans le but de le reproduire fidèlement dans tous ses aspects, mais plutôt d’en identifier les fondements computationnels importants [Haykin 05].

L’*Analyse de Scènes Auditives* (ASA) [Bregman 90] a posé les bases de la *spécification fonctionnelle* du système auditif humain la plus convaincante et la plus acceptée. Se positionnant d’emblée dans le contexte des mixtures de sons concurrents et simultanés, l’organisation proposée du traitement de l’information auditive explique l’émergence de sources subjectives (*flux* ou *streams*) relatives aux différentes sources physiques. Ces flux alimentent des activités haut-niveau de reconnaissance ou d’interprétation de scènes. Il est notable que les primitives les plus élémentaires mises en jeu dans les processus de ségrégation et de fusion intrinsèques à l’ASA peuvent dépendre de la fréquence, du timbre et de l’amplitude des sons mais aussi de leur localisation spatiale.

L'*Analyse de Scènes Auditives Computationnelle* (CASA = (*Computational Auditory Scene Analysis*) a par la suite tenté d'implémenter ou d'étudier sur des machines quelques principes de l'ASA. L'un de ses objectifs est donc d'extraire et de suivre automatiquement un ou plusieurs signaux d'intérêt dans un « environnement Cocktail Party » sur la base d'un modèle computationnel de la scène. Du fait que l'ASA postule, parfois à tort, l'existence de principes psychologiques communs (Gestalt) avec l'analyse de scènes visuelles, il se peut que leur exploitation par l'approche CASA ne soit pas en accord avec la réalité perceptive [De Cheveigné 02]. De même, tout comme les travaux de Marr en vision dans la lignée desquels elle se situe, on peut soupçonner l'approche CASA d'accorder une importance trop importante aux aspects « représentation ». Il n'empêche qu'elle a constitué un terrain fertile pour l'élaboration d'idées nouvelles [Rosenthal 97].

La thématique de l'*Audition en Robotique* (*Robot Audition*) a clairement bénéficié des développements de l'Analyse de Scènes Auditives Computationnelle. Ainsi, les travaux pionniers d'OKUNO [Okuno 04] ont identifié trois fonctions principales que doit comprendre tout système complet de traitement de l'information auditive par un robot. Il s'agit de

- la *localisation de sources*, qui intègre éventuellement leur suivi ;
- la *séparation de flux audio*, ou *extraction* de sources ;
- leur *reconnaissance*, qui inclut – mais n'est pas limitée à – l'ASR, et peut s'étendre à l'*interprétation* de scènes.

L'Analyse de Scènes Auditives Computationnelle suppose *a priori* un système auditif binaural, à l'instar du système auditif humain. Cependant, la catégorisation rappelée ci-dessus est communément admise aussi bien dans les approches de l'Audition en Robotique visant à imiter la géométrie ou l'algorithme du système auditif humain que dans les méthodes alternatives consistant en l'exploitation d'une antenne de microphones.

Les communautés Traitement du Signal et Acoustique traitent bien sûr depuis longtemps des problématiques de la localisation, extraction et reconnaissance. Néanmoins, l'objectif de doter un robot de capacités auditives soulève des problèmes inédits, souvent non pris en compte dans les méthodes existantes.

- La propriété d'*embarquabilité* d'un capteur sonore constraint son intégration sur un robot dans ses aspects géométrique et énergétique. Une solution binaurale dont la partie externe consiste en une tête munie de deux pavillons simplifie la géométrie et l'électronique embarquée. Inversement, une antenne de microphones est plus difficile à déployer mais peut impliquer une complexité algorithmique et des capacités de traitement de l'information moindres du fait de la redondance d'information qu'elle apporte.
- La Robotique exige l'exécution de nombreuses fonctions auditives en temps fortement constraint, voire en *temps réel*. Un traiteur du signal audio ou un acousticien souhaitant dresser une cartographie acoustique d'un environnement accepte de traiter l'information en différé pendant plusieurs secondes à plusieurs minutes. *A contrario*, l'exploitation de primitives de localisation de sources dans des tâches robotiques telles que la locomotion ou l'interaction (visio-)auditive nécessite leur élaboration à une dizaine d'Hertz.
- Une *contrainte fréquentielle* se rajoute, au sens où la plupart des signaux sonores utiles en Robotique recouvrent une bande de fréquences relativement large par rapport à sa fréquence centrale. Il en est ainsi de la voix humaine, dont nous considérerons que la bande de fréquences $F = [400\text{Hz}; 3\text{kHz}]$ suffit à capturer l'essentiel de la puissance tout en assurant l'intelligibilité du message. Les méthodes de la littérature dédiées au traitement de signaux monofréquentiels sont donc inadaptées, et leur extension au cas

“large bande” peut constituer un défi.

- Dans la même veine, la *distance des sources sonores au robot* peut engendrer des répercussions méthodologiques non triviales. Si les sources sont suffisamment éloignées, l’hypothèse de *champ lointain*, selon laquelle les fronts d’onde sont plans au voisinage du capteur, simplifie généralement la formulation du problème considéré et sa résolution. Cependant, il est parfois nécessaire de former l’hypothèse de *champ proche*, consistant à prendre explicitement en compte la sphéricité des fronts d’ondes, sous peine de conclusions erronées. Ces considérations sont particulièrement importantes dans un contexte d’interaction proximale homme-robot.
- La *variabilité* et l'*évolutivité* intrinsèques aux *environnements robotiques*, par nature non contrôlés, complexifie de manière déterminante toute exploitation de l’information sonore. Ceci pose bien sûr des problèmes de modélisation, mais exige aussi une adaptation ou une robustesse des traitements aux *réverbérations*, *bruits ambients* ou *sources parasites*.
- Quelques problèmes plus particuliers peuvent également être mentionnés. À titre d’exemple, l’analyse d’une scène auditive par un robot en mouvement nécessite le traitement des *bruits générés par son déplacement*. De même, dans une situation de *barge-in* où le robot est interrompu pendant qu’il émet intentionnellement des sons – *e.g.* lorsqu’un humain lui coupe la parole dans une situation d’interaction – il doit être capable de les supprimer en ligne de l’analyse de la scène. Enfin, on peut citer la fusion de l’audition avec la vision ou d’autres modalités sensorielles, l’exploitation du mouvement pour l’audition active, etc.

IV.2.2 Positionnement de nos contributions

Notre but a été de proposer un ensemble cohérent de fonctions auditives satisfaisant aux contraintes de la Robotique, à développer un capteur auditif intégré embarquable permettant leur implémentation, et à les expérimenter dans des environnements réels. En raison de nos compétences et de nos centres d’intérêt scientifiques, nous nous sommes limités aux niveaux bas de détection, localisation et extraction de sources. Nous ne sommes pas concernés par la reconnaissance de parole ou l’interprétation de scènes auditives.

Au commencement de notre activité, les approches binaurales donnaient lieu à des résultats assez mitigés, principalement à cause de leur forte sensibilité aux variations de l’environnement acoustique et de la difficulté à modéliser précisément les modifications induites par la présence de la tête. Ce constat – encore valide – ainsi que l’abondante littérature développée en Traitement d’Antenne nous ont orientés vers l’exploitation d’une antenne de microphones.

Nous proposons ci-dessous un bref état de l’art des sous-disciplines que nous avons abordées, puis expliquons notre ligne de recherche. Nos prospectives seront développées aux §V.1.3 et §V.2.2. Pour mémoire, les projets de recherche associés figurent dans le §II.3.5 page 63.

IV.2.2.1 Localisation, extraction et détection de sources sonores

A Méthodes utilisées en Robotique La localisation de sources a été le premier problème abordé par les roboticiens. Nous proposons une revue des approches bioinspirées et d’antenne dans [Argentieri 07]. Ces dernières possèdent une précision et une robustesse au bruit généralement supérieures, et se déclinent en trois classes.

Les méthodes de *localisation par corrélation* estiment les délais temporels séparant les signaux perçus par les microphones (TDOA = *Time Delay of Arrival*) sur la base de calculs d'intercorrélations, puis inversent le lien qui les unit avec les positions des sources pour un modèle de propagation donné. La première étape est la plus cruciale. Elle se décline en de nombreuses variantes de la *Generalized Cross-Correlation* (GCC) [Knapp 76], qui se place dans le domaine fréquentiel et propose l'introduction de pondérations. La solution probablement la plus aboutie est proposée dans [Valin 03] pour le calcul des TDOAs entre des microphones disposés en cube sur une plateforme Pioneer-2. Elle constitue une extension robuste au bruit de l'algorithme *Phase Transform* (PHAT) défini par [Omologo 97] de façon à n'exploiter que les phases des signaux. La seconde phase de l'algorithme s'appuie sur des techniques classiques. Les précisions affichées sont bonnes, de quelques degrés sous l'hypothèse de champ lointain. Toutefois, le traitement du champ proche semble mal conditionné et on observe une forte sensibilité aux caractéristiques du bruit.

De toutes les méthodes ancrées dans le Traitement du Signal, celles basées sur la *formation de voie (beamforming)* ont été les plus exploitées en Robotique, sans doute en raison de leur simplicité et de leur faible coût. Le but est de former le *diagramme de directivité (beampattern)* de l'antenne discrète *via* l'insertion de filtres en aval des microphones dont on somme les sorties. La *polarisation électronique* de l'antenne en direction d'une source d'intérêt permet alors d'en extraire le message, en dépit de sources ou bruits parasites. Une procédure de construction de cartes acoustiques se déduit immédiatement pour la localisation : il suffit de balayer l'ensemble des directions d'écoute « utiles » puis d'intégrer sur une fenêtre temporelle glissante l'*énergie acoustique incidente* selon chacune d'elles. Une approche récurrente en Robotique est la *formation de voie conventionnelle* sous l'hypothèse de champ lointain, qui décale temporellement les signaux perçus de façon que leur somme soit constructive pour la direction de polarisation sélectionnée. Du fait que cette méthode conduit à une mauvaise focalisation spatiale aux basses fréquences, son application [Mattos 04] à l'antenne de 8 microphones disposés sur la ceinture de la plateforme mobile EvBoy II suggère de filtrer les fréquences inférieures à 800 Hz. Dans la même veine, [Tamai 05] ne conserve que les fréquences supérieures à 1 kHz. Il s'agit toutefois de solutions par défaut, car les fréquences éliminées capturent une portion importante de la puissance de la voix humaine. En outre, certaines localisations sont erronées en raison de lobes secondaires importants dans les diagrammes de directivité obtenus. C'est pourquoi [Valin 04a] implémente le calcul d'énergie dans le domaine fréquentiel sur la base d'intercorrélations prébranchées « à la PHAT » [Omologo 97].

Enfin, une dernière alternative [Asano 99] s'inscrit dans la classe des méthodes « à haute résolution », appelées ainsi car leur résolution est théoriquement indépendante de l'ouverture de l'antenne. Elle consiste en l'extension large bande de la méthode MUSIC (MUltiple SIgnal Classification) [Schmidt 86]. Malgré son coût exorbitant et son incapacité à traiter les réverbérations, elle demeure l'alternative à [Valin 04a] la plus utilisée [Okuno 08][Ishi 09].

L'extraction de sources est reconnue comme devant soit s'adapter soit être robuste à des changements dynamiques de l'environnement, sous peine de mettre en péril l'exploitation des données qu'elle produit. À l'origine, elle reposait essentiellement sur des techniques de formation de voie, du fait de la préalable détermination des positions des sources. Bien que les phénomènes de diaphonie concomitants puissent être résolus par des méthodes de séparation aveugle, celles-ci n'exploitent pas d'information géométrique, nécessitent des aménagements en vue de traiter les corrélations de sources consécutives à des réverbérations, et sont sujettes à des ambiguïtés : permutations, ambiguïtés dues à un nombre de capteurs plus élevé que le

nombre de sources, etc. Une solution à mi-chemin a été mise au point dans [Valin 04b], sur la base d'une simplification de l'algorithme GSS (*Geometric Source Separation*) de [Parra 02] et d'un post-filtrage pour réhausser les résultats. La matrice des filtres permettant la séparation de mélanges convolutifs par GSS est déterminée au moyen d'un algorithme semblable à la formation de voie à variance minimale sous contrainte linéaire – *Linearly Constrained Minimum Variance (LCMV) beamformer* – à la différence près qu'elle minimise la somme des carrés des intercorrélations du vecteur des sorties séparées sous des contraintes de réponse en les directions des sources. L'implémentation est toutefois lourde, et les distorsions spectrales induites sur les sorties exigent des traitements annexes préalablement à toute exploitation pour l'ASR [Okuno 08].

Le problème de la détection du nombre de sources a bien sûr été identifié très tôt. Bien qu'il s'agisse de la fonction située au plus bas niveau, les approches proposées jusqu'ici effectuent cette estimation sur la base de résultats de localisation. Une première solution suggérée dans [Valin 04a] repose sur un post-traitement probabiliste appliqué à l'énergie acoustique calculée au moyen d'une formation de voie. Une stratégie plus récente estime les directions des sources et leur nombre à partir des TDOAs par le couplage d'une « inversion » par moindre carrés avec un algorithme K-means adaptatif permettant de ne conserver que les solutions les plus cohérentes. Enfin, [Ishi 09] définit une stratégie *ad hoc* basée sur l'extension MUSIC de [Asano 99]. Dans l'ensemble, ces méthodes sont limitées par un support théorique peu développé, par la nécessité d'une paramétrisation subjective, ou par leur complexité.

B Ligne de recherche Nous nous sommes d'abord proposés de définir de nouveaux algorithmes pour la localisation et le filtrage spatial de sources par *formation de voie*. Nous nous sommes placés dans le cadre des approches *large bande* et indépendantes des données (*data-independent beamforming*), qui consistent à approcher un gabarit de référence spatio-fréquentiel sans connaissance préalable des statistiques des sources et du bruit [Van Veen 88]. Il s'agissait en particulier d'évaluer si un diagramme d'antenne invariant en fréquences et doté d'une phase linéaire permet une localisation satisfaisante et un extraction suffisamment peu distordue. La nécessité, mentionnée dans [Van Veen 88], de limiter le *gain de bruit blanc* (*White Noise Gain*) lors de la synthèse nous a placés d'emblée dans le contexte d'optimisation, et de l'*optimisation semi-définie* en particulier. En effet, comme cela a été notre démarche pour nos travaux en commande (§III), l'obtention de programmes convexes solubles numériquement par des solveurs dédiés avec des propriétés satisfaisantes – temps polynômial, précision arbitraire – nous semble constituer une solution acceptable.

Notre second objectif a été de nous tourner vers des *méthodes à haute résolution large bande* alternatives à [Asano 99], dont le coût est prohibitif. Nous n'avons pas envisagé d'aborder la localisation de sources par d'autres méthodes spectrales [Krim 96] basées sur des formations de voie statistiquement optimales ou adaptatives pour deux raisons principales : la restriction imposée par la Robotique sur la taille de l'antenne limiterait à coup sûr les résolutions obtenues comparativement aux méthodes à haute résolution ; dans leurs versions de base, ces méthodes conduisent à de mauvais résultats en présence d'interférences corrélées avec les sources du fait de réverbérations.

Enfin, notre dernier but a été de définir une méthode de détection théoriquement étayée, de faible coût, et ne nécessitant qu'une paramétrisation subjective limitée.

IV.2.2.2 Ensemble matériel et logiciel pour l'audition en robotique

A Quelques réalisations significatives Il est admis que la diffusion de robots auditifs exige que les fonctions bas-niveau soient embarquées dans une solution matérielle dédiée embarquable et peu coûteuse. Ainsi, le portage du système de localisation, suivi et séparation de sources AUDIBLE vers une cible DSP est décrit dans [Brière 08]. Similairement, un système de séparation de sources sur un processeur reconfigurable dynamique est proposé par [Kuratori 05]. Sur le plan logiciel, la librairie Open Source HARK[†] (HRI-JP Audition for Robotics with Kyoto University) [Nakadai 08] intègre un ensemble conséquent de primitives pour l'Analyse de Scènes Auditives Computationnelle en Robotique.

B Ligne de recherche Nous avons conçu et développé un *ensemble matériel et logiciel* pour l'acquisition sonore de haute qualité et l'élaboration de primitives audio évoluées sur la base d'une antenne de microphones. Cet ensemble constitue le capteur intégré embarquable EAR (“Embedded Audition for Robotics”), et sert de base à l'ensemble de nos expérimentations. Les critères que nous avons retenus pour la conception du capteur EAR sont la simplicité et la puissance d'utilisation du point de vue de l'utilisateur, avec en particulier la possibilité de paramétrisation complète de l'acquisition – *e.g.* suite à un changement de transducteurs, à une redéfinition de la fréquence d'échantillonnage, etc. Le premier prototype ainsi que l'évolution en cours de fabrication intègrent une unité FPGA, permettant l'implémentation de traitements avec un parallélisme massif. Cette unité embarque une architecture modulaire, permettant à un concepteur d'incorporer facilement de nouveaux traitements *via* des “templates” pré-définis. À moyen terme, nous souhaitons diffuser EAR en Open Source et l'interfacer avec HARK.

IV.2.2.3 Vers la détection d'activité vocale

Enfin, nous avons récemment entamé un travail préliminaire en vue de proposer un schéma de *détection d'activité vocale* (VAD = Voice Activity Detection) *via* la détection de signaux voisés. Celui-ci concerne une nouvelle approche pour la *détection* – voire le *réhaussement* – de “patterns” acoustiques, sur la base du *filtrage adapté stochastique*. Son évaluation et sa mise au point constituent une proportion conséquente de notre activité actuelle.

IV.3 Contribution à l'audition en robotique

IV.3.1 Terminologie

Cette section recense les notations et formules de base sur lesquelles s'appuient nos travaux. Les démonstrations ne sont pas présentées, s'agissant de résultats très classiques en acoustique linéaire. Le lecteur peut se référer à l'ouvrage [Van Trees 02] ou aux thèses [Argentieri 06][Abhayapala 99].

IV.3.1.1 Généralités

Une antenne de N microphones omnidirectionnels échantillonne le champ de pression causé par $D < N$ sources ponctuelles. Celui-ci se propage librement dans un milieu homogène et

[†]. L'interjection “Hark!” signifie “Listen!” dans “The Tempest” de Shakespeare.

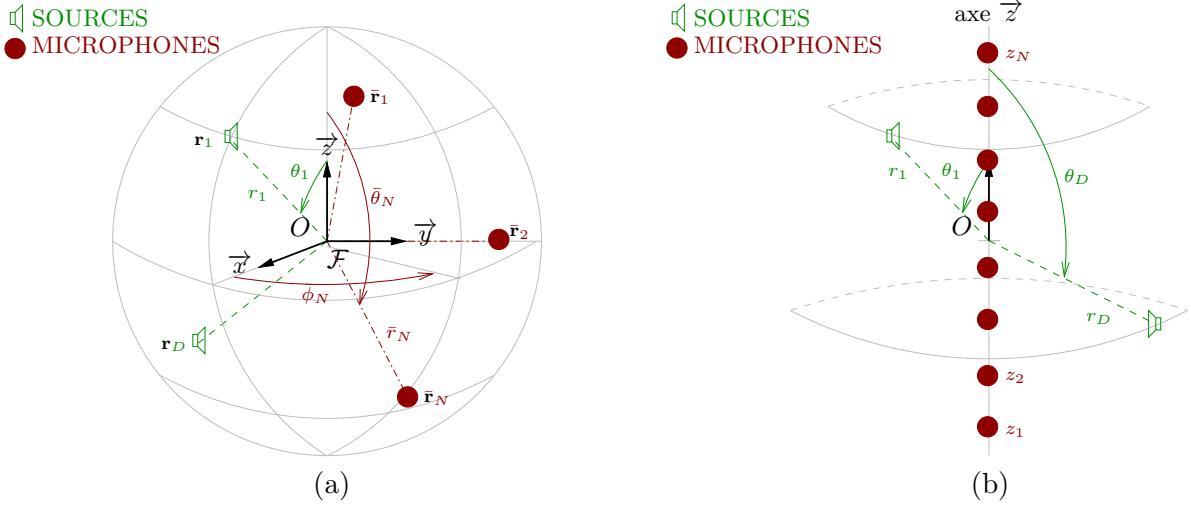


FIGURE IV.1 – Antenne de microphones. (a) Cas général. (b) Antenne linéaire uniforme.

isotrope à la vitesse c . Les positions des sources et des microphones sont respectivement repérées par les coordonnées sphériques $\mathbf{r}_d = (r_d, \theta_d, \phi_d)$, $d = 1, \dots, D$, et $\bar{\mathbf{r}}_n = (\bar{r}_n, \bar{\theta}_n, \bar{\phi}_n)$, $n = 1, \dots, N$, relativement à un repère $\mathcal{F} = (O, \vec{x}, \vec{y}, \vec{z})$ lié à l'antenne (Figure IV.1-(a)). En l'absence de bruit, le signal sonore causé en l'origine O de \mathcal{F} par la $d^{\text{ème}}$ source est noté $s_d(t)$. Les transformées de $\mathbf{s}(t) \triangleq (s_1(t), \dots, s_D(t))^T$ constituent $\mathbf{S}(k) = (S_1(k), \dots, S_D(k))^T \in \mathbb{C}^D$, où $k = \frac{2\pi f}{c}$ et f désignent les fréquences spatiales et temporelles.

Soit $\mathbf{V}(\mathbf{r}_d, k) \in \mathbb{C}^N$ le vecteur d'antenne relatif à la $d^{\text{ème}}$ source, i.e. la fonction de transfert entre le point O et les microphones lorsque seule la $d^{\text{ème}}$ source, placée en \mathbf{r}_d , est active en l'absence de bruit. Les transformées $\mathbf{N}(k) = (N_1(k), \dots, N_N(k))^T$ et $\mathbf{X}(k) = (X_1(k), \dots, X_N(k))^T$ du bruit $\mathbf{n}(t) \triangleq (n_1(t), \dots, n_N(t))^T$ et du signal $\mathbf{x}(t) \triangleq (x_1(t), \dots, x_N(t))^T$ en les N transducteurs vérifient

$$\begin{aligned} \mathbf{X}(k) &= \sum_{d=1}^D \mathbf{V}(\mathbf{r}_d, k) S_d(k) + \mathbf{N}(k), \\ &= \underline{\mathbf{V}}(\mathbf{r}_1, \dots, \mathbf{r}_D, k) \mathbf{S}(k) + \mathbf{N}(k), \text{ où } \underline{\mathbf{V}}(\mathbf{r}_1, \dots, \mathbf{r}_D, k) = (\mathbf{V}(\mathbf{r}_1, k) \mid \dots \mid \mathbf{V}(\mathbf{r}_D, k)). \end{aligned} \quad (\text{IV.1})$$

Lors de sa propagation à la vitesse c , l'onde sphérique issue de toute source ponctuelle placée en $\mathbf{r} = (r, \theta, \phi)$ admet une amplitude inversement proportionnelle à la distance parcourue. C'est pourquoi le vecteur d'antenne s'écrit

$$\mathbf{V}(\mathbf{r}, k) = \left(\frac{|r| e^{jk||\mathbf{r}|}}{||\bar{\mathbf{r}}_1 - \mathbf{r}||}, \dots, \frac{|r| e^{jk||\mathbf{r}|}}{||\bar{\mathbf{r}}_N - \mathbf{r}||} \right)^T. \quad (\text{IV.2})$$

La dimension physique L d'une antenne conditionne sa discrimination spatiale, au sens où elle est d'autant plus sensible à la position des sources qu'elle s'étend sur un nombre important de longueurs d'onde. En outre, en toute rigueur, une source de fréquence centrale f_0 et de largeur de bande Δf est dite *faible bande* si $\frac{L}{\lambda_0} \ll (\frac{\Delta f}{f_0})^{-1}$, avec $\lambda_0 = \frac{c}{f_0}$.

†. Un critère équivalent est que le TBWP (*Time BandWidth Product*) vérifie $\text{TWBP} = (\frac{L}{c})\Delta f \ll 1$.

IV.3.1.2 Antenne linéaire uniforme

L'essentiel de nos travaux suppose une *antenne linéaire uniforme* (ULA = Uniform Linear Array), dont les transducteurs sont espacés d'une distance l sur un segment de droite porté par le vecteur \vec{z} du repère \mathcal{F} et centré en O (Figure IV.1-(b)). En raison de la symétrie de révolution autour de l'axe \vec{z} , le comportement du réseau de microphones pour une source située en $\mathbf{r} = (r, \theta, \phi)$ n'est dicté que par la distance r et l'azimuth θ de celle-ci. Le vecteur d'antenne (IV.2) se particularise en

$$\mathbf{V}(r, \theta, k) = \left(r e^{jkr} \frac{e^{-jk\sqrt{r^2 + \bar{z}_1^2 - 2r\bar{z}_1 \cos\theta}}}{\sqrt{r^2 + \bar{z}_1^2 - 2r\bar{z}_1 \cos\theta}}, \dots, r e^{jkr} \frac{e^{-jk\sqrt{r^2 + \bar{z}_N^2 - 2r\bar{z}_N \cos\theta}}}{\sqrt{r^2 + \bar{z}_N^2 - 2r\bar{z}_N \cos\theta}} \right)^T, \quad (\text{IV.3})$$

où $\bar{z}_n = (n - \frac{N+1}{2})l$, $n = 1, \dots, N$, désigne l'abscisse du $n^{\text{ème}}$ microphone selon (O, \vec{z}) .

L'expression (IV.3) prend explicitement en compte la sphéricité des fronts d'onde, auquel cas la source est dite *en champ proche*. Si l'échantillonnage spatial réalisé par l'antenne peut être assimilé au cas où $r = +\infty$, *i.e.* si l'onde peut localement être considérée comme plane, alors la source est dite *en champ lointain*, et il vient

$$\mathbf{V}^\infty(\theta, k) \triangleq \lim_{r \rightarrow +\infty} \mathbf{V}(r, \theta, k) = (e^{jk\bar{z}_1 \cos\theta}, \dots, e^{jk\bar{z}_N \cos\theta})^T, \text{ où } \bar{z}_n = (n - \frac{N+1}{2})l. \quad (\text{IV.4})$$

Classiquement, on admet que (IV.3) se simplifie en (IV.4) dès lors que r est supérieur à la *distance de Rayleigh* $r_R = \frac{2L^2}{\lambda}$, fonction de la longueur λ de l'onde et de l'ouverture $L = Nl$ de l'antenne. Afin d'éviter que $x(t)$ prenne des valeurs identiques pour des sources d'azimuths différents, on exige que la distance inter-microphones l obéisse au *théorème de Shannon spatial*, selon lequel

$$l < \frac{\lambda}{2}. \quad (\text{IV.5})$$

Sauf mention contraire, les études simulées ou expérimentales présentées dans la suite supposent une antenne linéaire de $N = 8$ transducteurs régulièrement espacés de $l = \frac{\lambda_{3\text{kHz}}}{2} = 5.66\text{ cm}$. Cette configuration a été retenue pour nos prototypes de capteurs audios intégrés embarquables, cf. §IV.3.5. Elle assure une ouverture compatible avec la contrainte d'embarquabilité tout en respectant (IV.5) sur la bande $F = [400\text{Hz}; 3\text{kHz}]$. Il convient de noter qu'une source couvrant F doit être considérée comme large bande.

IV.3.2 Techniques de formation de voie pour l'extraction et la localisation de sources sonores large bande

IV.3.2.1 Introduction à la formation de voie

Comme indiqué au §IV.2.2-B, notre premier objectif a été de définir une méthode de formation de voie adaptée à la Robotique. Même s'il s'avèrera (§ IV.3.3) que d'autres stratégies sont mieux adaptées à la localisation, cette contribution n'en demeure pas moins essentielle, au sens où son exploitation déborde du contexte du filtrage spatial.

A Formulation mathématique Par définition, le signal $y(t)$ délivré par une formation de voie est la somme des sorties de filtres placés en aval des microphones. Désignons par $W_n(k)$

la réponse en fréquences du filtre associé au $n^{\text{ème}}$ microphone, $n = 1, \dots, N$, et par $\mathbf{W}(k)$ le vecteur $\mathbf{W}(k) = (W_1(k), \dots, W_N(k))^T \in \mathbb{C}^N$. La transformée de Fourier $Y(k)$ de $y(t)$ s'écrit

$$Y(k) = \sum_{n=1}^N W_n(k) X_n(k) = \mathbf{W}^T(k) \mathbf{X}(k) \quad (\text{IV.6})$$

soit, d'après (IV.1),

$$Y(k) = \sum_{d=1}^D D(\mathbf{r}_d, k) S_d(k) + \sum_{n=1}^N W_n(k) N_n(k), \quad (\text{IV.7})$$

où l'expression

$$D(\mathbf{r}, k) \triangleq \sum_{n=1}^N W_n(k) V_n(\mathbf{r}, k) = \mathbf{W}^T(k) \mathbf{V}(\mathbf{r}, k) \quad (\text{IV.8})$$

est appelée *réponse spatio-fréquentielle*, ou *diagramme de directivité* de l'antenne.

Dans la suite du manuscrit, la notation adoptée au §IV.3.1 demeure en vigueur. Ainsi, $\{\mathbf{V}(\mathbf{r}, k) = \mathbf{V}(r, \theta, \phi, k), D(\mathbf{r}, k) = D(r, \theta, \phi, k)\}$ désignent le vecteur et la réponse d'une antenne générale ; $\{\mathbf{V}(r, \theta, k), D(r, \theta, k)\}$ font l'hypothèse de microphones disposés selon l'axe (O, \vec{z}) ; $\{\mathbf{V}^\infty(\theta, k), D^\infty(\theta, k)\}$ supposent de plus les sources en champ lointain.

B Formation de voie faible bande Si on peut admettre que les sources émettent à une fréquence commune k_0 , alors le problème se ramène à la recherche d'un vecteur $\mathbf{W} = \mathbf{W}(k_0) \in \mathbb{C}^N$ conduisant à une réponse $D(\mathbf{r}, k_0)$ satisfaisante. La littérature propose un éventail de solutions. Dans le cas d'une antenne linéaire uniforme sous l'hypothèse de champ lointain, des méthodes de synthèse de filtres RIF peuvent être indifféremment exploitées pour la formation du diagramme de directivité en azimuth. En effet, le résultat de l'échantillonnage spatial, par des transducteurs séparés de la distance l , d'une onde plane monochromatique en fonction de son azimuth θ s'apparente à l'échantillonnage temporel à la période T_e d'un signal en fonction de sa fréquence f .

Le gain de bruit blanc s'écrit $G_n \triangleq W^T W$. Sa valeur doit être limitée [Van Veen 88].

C Formation de voie large bande sur la base de filtres RIF Dans le cas large bande, il est généralement opportun ne pas introduire des gains et déphasages invariants en fréquence. Si la formation de voie repose sur l'exploitation de filtres RIF, alors son diagramme de directivité admet une expression mathématique semblable à (IV.8). En effet, en notant w_{nq} le $q^{\text{ème}}$ coefficient du filtre RIF d'ordre Q associé au $n^{\text{ème}}$ transducteur, la réponse $W_n(k)$ de ce filtre s'écrit

$$W_n(k) = \sum_{q=0}^Q w_{nq} e^{-jkcqT_e}, \quad (\text{IV.9})$$

avec T_e la période d'échantillonnage temporel. Dès lors, y compris dans le cas d'un vecteur d'antenne général (IV.2), le diagramme de directivité devient †

$$D_{\text{LB}}(\mathbf{r}, k) = \sum_{n=1}^N \sum_{q=0}^Q w_{nq} e^{-jkcqT_e} V_n(\mathbf{r}, k) = \mathbf{W}_{\text{LB}}^T \mathbf{V}^{\text{LB}}(\mathbf{r}, k) \quad (\text{IV.10})$$

†. Le produit de Kronecker $\mathbf{X} \otimes \mathbf{Y}$ des matrices $\mathbf{X} \in \mathbb{C}^{n_X \times m_X}$ et $\mathbf{Y} \in \mathbb{C}^{n_Y \times m_Y}$ est la matrice de $\mathbb{C}^{(n_X n_Y) \times (m_X m_Y)}$ formée des $n_X m_X$ sous-matrices $Z_{n,m} = \mathbf{X}_{n,m} \mathbf{Y}$, $n = 1, \dots, n_X$, $m = 1, \dots, m_Y$.

où

$$\mathbf{W}_{\text{LB}} \triangleq (w_{10}, \dots, w_{1Q}, \dots, w_{N0}, \dots, w_{NQ})^T \quad (\text{IV.11})$$

et

$$\mathbf{V}^{\text{LB}}(\mathbf{r}, k) \triangleq \mathbf{V}(\mathbf{r}, k) \otimes \mathbf{V}^{\text{RIF}}(k) \text{ avec } \mathbf{V}^{\text{RIF}}(k) \triangleq (1, e^{-jkcT_e}, \dots, e^{-jkcQT_e})^T. \quad (\text{IV.12})$$

La bornitude de $G_{\text{LB}} \triangleq W_{\text{LB}}^T W_{\text{LB}}$ limite la sensibilité de la formation de voie au bruit.

Pour une antenne linéaire uniforme, la dépendance en \mathbf{r} se limite à une dépendance en (r, θ) . D'après (IV.4), la polarisation selon un azimuth θ_0 sous l'hypothèse de champ lointain par formation de voie conventionnelle serait obtenue pour $\forall k$, $(W_1(k), \dots, W_N(k))^T \equiv (\mathbf{V}^\infty(\theta_0, k))^*$, avec * l'opérateur de conjugaison.

IV.3.2.2 Formation de voie large bande « naïve » par optimisation convexe

Pour que notre alternative à la formation de voie large bande conventionnelle soit adaptée à la Robotique, elle doit approcher au mieux un diagramme de référence donné en dépit de la petite taille de l'antenne et de la bornitude *a priori* de son gain de bruit blanc. Comme indiqué précédemment, nous nous sommes placés dans le contexte de l'optimisation semi-définie. Un panorama des applications de cette discipline au Traitement du Signal figure dans [Balakrishnan 98][Luo 03]. Bien que la littérature propose un nombre significatif d'applications de la programmation semi-définie à la synthèse d'antennes, le problème demeure largement ouvert.

La formation de voie faible bande proposée dans [Wang 03] pour une antenne linéaire consiste à minimiser le module pire cas de l'erreur – complexe – entre son diagramme de directivité et un diagramme de référence sur une grille en azimuth. Les gains complexes associés aux microphones sont solutions d'un programme d'optimisation sur le cône du second ordre (SOCP = Second-Order Cone Program) [Boyd 94], convexe et soluble numériquement. L'approche autorise l'introduction de contraintes convexes en les coefficients recherchés, telles que – mais pas seulement – la bornitude du gain de bruit blanc. Une conception robuste vis à vis d'incertitudes en gain et/ou phase sur le vecteur d'antenne et/ou d'erreurs de quantification sur les coefficients peut être également obtenue. Une version large bande serait donc particulièrement bien adaptée à la Robotique.

Du fait que ce résultat théorique n'exige pas de structure particulière du vecteur d'antenne, nous en avons proposé une extension large bande « naïve » [CIACL-08-AR-1]. Étant donné un gabarit de référence $D_{\text{LB}}^{\text{REF}}(r_0, \theta, k)$ spécifié en azimuths et fréquences pour une distance r_0 – i.e. $\forall \theta, k$, $D_{\text{LB}}^{\text{REF}}(r_0, \theta, k) \equiv \bar{D}(\theta, k)$ donné – il suffit de résoudre, en reprenant les notations (IV.10)–(IV.11)–(IV.12) pour une antenne linéaire, le SOCP

$$\begin{aligned} & \min_{\epsilon \geq 0, \mathbf{W}_{\text{LB}} \in \mathbb{R}^{N(Q+1)}} \epsilon \\ & \text{sous } \begin{cases} \forall (\theta, k) \in \Theta \times K, |\mathbf{W}_{\text{LB}}^T \mathbf{V}^{\text{LB}}(r_0, \theta, k) - D_{\text{LB}}^{\text{REF}}(r_0, \theta, k)|^2 \leq \epsilon \\ \mathbf{W}_{\text{LB}}^T \mathbf{W}_{\text{LB}} \leq \delta, \end{cases} \end{aligned} \quad (\text{IV.13})$$

où la borne δ est définie *a priori* et $\Theta \times K$ désigne une grille du plan azimuth × fréquence.

Lorsque l'objectif est la construction de cartes d'énergie acoustique, le comportement en phase de l'antenne importe peu, de sorte qu'il est opportun d'envisager une formation de voie en gain seulement. Il suffit pour cela de remplacer le premier ensemble de

contraintes de (IV.13) par $\forall(\theta, k) \in \Theta \times K$, $\| |\mathbf{W}_{LB}^T \mathbf{V}^{LB}(r, \theta, k)| - |D_{LB}^{REF}(r, \theta, k)| \|_2^2 \leq \epsilon$. Le programme d'optimisation obtenu n'est pas convexe. Cependant, suivant la procédure proposée dans [Wang 03], il est équivalent à un programme en les variables de décision ϵ , \mathbf{W}_1 , \mathbf{W}_2 , multiconvexe en \mathbf{W}_1 , \mathbf{W}_2 – *i.e.* convexe en l'une de ces variables lorsque l'autre est fixée – et dont la somme des optima de \mathbf{W}_1 , \mathbf{W}_2 est égale à l'optimum recherché de \mathbf{W}_{LB} . Une initialisation judicieuse et des alternances d'optimisations sur ϵ et l'une des deux variables \mathbf{W}_1 , \mathbf{W}_2 , l'autre étant fixée, conduisent en théorie à un optimum local satisfaisant.

Malgré leur attractivité, ces deux méthodes sont d'une utilité pratique limitée. Un premier écueil réside dans la définition d'un maillage convenable, surtout selon l'axe des fréquences. En effet, un compromis satisfaisant doit être établi entre une finesse suffisante pour éviter un comportement médiocre entre les points de la grille, et la dimension du programme d'optimisation, qui influe directement sur son conditionnement – *e.g.* 289 variables de décision et 2650 contraintes pour des filtres RIF d'ordre $Q = 35$ et un diagramme de référence invariant en fréquences et à phase linéaire. Des spécifications trop ambitieuses sur le diagramme de référence se traduisent également par un mauvais conditionnement. Une autre difficulté tient à ce que le choix d'une faible valeur de δ dégrade significativement la réponse de l'antenne. Enfin, une synthèse en gain et phase (resp. en gain seulement) nécessite environ 30 mn (resp. plusieurs heures) de temps CPU.

IV.3.2.3 Analyse modale et optimisation convexe pour la formation de voie

Les conclusions précédentes nous ont incités à définir une nouvelle stratégie de formation de voie large bande par optimisation convexe, avec pour objectifs :

- une formulation du problème qui capture davantage l'*« essence »* d'un diagramme de directivité, par opposition à l'approche *« force brute »* proposée au §IV.3.2.2 ;
- la limitation de la complexité *via* la réduction de la grille sur laquelle sont évalués les écarts entre le diagramme courant et le gabarit de référence ;
- la possibilité de réaliser des synthèses en champ proche ou en champ lointain.

Il s'est avéré que l'analyse modale des diagrammes de directivité constitue un cadre théorique intéressant [Abhayapala 99]. Nous en rappelons brièvement les éléments essentiels ci-dessous.

A Décomposition modale d'un diagramme d'antenne Le champ de potentiel échantillonné par les microphones vérifie une équation d'onde linéaire homogène, faisant intervenir ses dérivées secondes par rapport aux variables d'espace – Laplacien – et au temps. Son amplitude complexe est solution d'une équation algébrique linéaire en les variables d'espace et la fréquence, dite *équation de Helmholtz*. Du fait qu'une formation de voie combine linéairement les signaux perçus par les microphones, sa réponse spatio-fréquentielle satisfait également l'équation de Helmholtz. Or, toute solution de cette équation dans le cas sphérique peut être décomposée sur la base des fonctions

$$Y_{mp}(\theta, \phi) \triangleq \sqrt{\frac{2m+1}{4\pi} \frac{(m-|p|)!}{(m+|p|)!}} P_{mp}(\cos \theta) e^{jp\phi}, \quad m \in \mathbb{N}, \quad p \in \{-m, \dots, m\}, \quad (\text{IV.14})$$

connues dans la littérature sous le nom d'*harmoniques sphériques*. Ci-dessus, $P_{mp}(.)$ désigne la fonction de Legendre associée de première espèce de degré m et d'ordre p .

Selon le produit scalaire habituellement défini sur les fonctions de (θ, ϕ) d'énergie finie, les fonctions de Legendre et les exponentielles sont orthogonales, de sorte que les harmoniques sphériques $\{Y_{mp}(\theta, \phi)\}_{m \in \mathbb{N}, p \in \{-m, \dots, +m\}}$ constituent une base orthonormale, *i.e.*, avec $*$ l'opérateur de conjugaison et $\delta_{(.,.)}$ le symbole de Kronecker,

$$\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} Y_{m_1 p_1}(\theta, \phi) Y_{m_2 p_2}^*(\theta, \phi) \sin \theta d\theta d\phi = \delta_{(m_1, m_2)} \delta_{(p_1, p_2)}. \quad (\text{IV.15})$$

Sur cette base, tout diagramme de directivité $D(\mathbf{r}, k)$ se décompose selon

$$D(\mathbf{r}, k) = D(r, \theta, \phi, k) = \sum_{m=0}^{\infty} \sum_{p=-m}^{+m} \alpha_{mp}(k) R_m(r, k) Y_{mp}(\theta, \phi), \quad (\text{IV.16})$$

avec

$$R_m(r, k) \triangleq r e^{jkr} h_m^{(2)}(kr) \quad (\text{IV.17})$$

et $h_m^{(2)}(.)$ la fonction sphérique de Hankel de seconde espèce d'ordre m . Dès lors, les *coefficients modaux* $\alpha_{mp}(k)$, qui vérifient

$$\alpha_{mp}(k) = \frac{1}{R_m(r, k)} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} D(r, \theta, \phi, k) Y_{mp}^*(\theta, \phi) \sin \theta d\theta d\phi, \quad (\text{IV.18})$$

caractérisent à eux seuls tout diagramme de directivité. Ces coefficients ne dépendent que de la fréquence. La dépendance de $D(\mathbf{r}, k)$ en les variables angulaires est capturée par les fonctions de base $\{Y_{mp}(\theta, \phi)\}_{m \in \mathbb{N}, p \in \{-m, \dots, +m\}}$, et l'influence de la distance r ne se répercute que sur la fonction de normalisation $R_m(r, k)$. Ainsi, la connaissance des coefficients modaux d'un diagramme de directivité pour une distance r_0 permet d'inférer le comportement de l'antenne pour toutes les valeurs de $r \neq r_0$. Champ proche et champ lointain sont unifiés grâce au passage à la limite

$$\lim_{r \rightarrow +\infty} R_m(r, k) = R_m(+\infty, k) = \frac{j^{m+1}}{k}. \quad (\text{IV.19})$$

Les équations (IV.16) et (IV.18) définissent une transformation orthogonale analogue à la série de Fourier. Elle possède les mêmes propriétés, telles que la linéarité et une relation de Parseval permettant de justifier le bien-fondé de l'approximation de la somme infinie (IV.16) par sa troncature aux $(M + 1)$ premiers modes

$$D(\mathbf{r}, k) \approx D_M(\mathbf{r}, k) \triangleq \sum_{m=0}^M \sum_{p=-m}^{+m} \alpha_{mp}(k) R_m(r, k) Y_{mp}(\theta, \phi). \quad (\text{IV.20})$$

On montre que l'approximation (IV.20) est d'autant meilleure que M croît, et que les modes d'ordres faibles sont les plus significatifs.

Pour une formation de voie de la forme $D(\mathbf{r}, k) = \mathbf{W}^T(k) \mathbf{V}(\mathbf{r}, k)$ réalisée sur la base de N microphones en $\bar{\mathbf{r}}_1 = (\bar{r}_1, \bar{\theta}_1, \bar{\phi}_1), \dots, \bar{\mathbf{r}}_N = (\bar{r}_N, \bar{\theta}_N, \bar{\phi}_N)$, il vient [Abhayapala 99]

$$\forall r = \|\mathbf{r}\| \gg \|\bar{\mathbf{r}}_n\|, \quad n = 1, \dots, N, \quad \alpha_{mp}(k) = -j4k\pi \sum_{n=1}^N W_n(k) j_m(k\bar{r}_n) Y_{mp}^*(\bar{\theta}_n, \bar{\phi}_n), \quad (\text{IV.21})$$

avec $j_m(.)$ la fonction de Bessel sphérique d'ordre m .

B Un nouvel algorithme de formation de voie Nous avons proposé dans [CIACL-10-AR-3] de synthétiser un diagramme de directivité par minimisation de l'erreur pire cas entre ses coefficients modaux et ceux d'un diagramme de référence. Plus précisément, on détermine d'abord – généralement de manière heuristique – l'indice M du mode pour lequel ces deux diagrammes peuvent être assimilés à leur troncature (IV.20). Cet indice est propre à la géométrie du réseau de microphones. Ensuite, on exprime au moyen de (IV.18) les coefficients modaux $\alpha_{00}^{\text{REF}}, \alpha_{1-1}^{\text{REF}}, \alpha_{10}^{\text{REF}}, \alpha_{11}^{\text{REF}}, \dots, \alpha_{MM}^{\text{REF}}$ – en nombre égal à $(M + 1)^2$ dans le cas général – du gabarit de référence. Afin de limiter d'éventuels phénomènes de Gibbs, celui-ci doit être suffisamment « lisse » ; une solution peut consister à ne retenir que les premiers termes de sa série de Fourier en (θ, ϕ) .

Si l'objectif est la synthèse faible bande à une fréquence k_0 , alors, d'après (IV.21), le vecteur de coefficients complexes $\mathbf{W} = \mathbf{W}(k_0) = (W_1(k_0), \dots, W_N(k_0))^T$ satisfait

$$\forall m \in \mathbb{N}, p \in \{-m, \dots, m\}, \alpha_{mp}(k_0) = \mathbf{W}^T \mathbf{V}_{mp}^{\text{MODAL}}(k_0), \quad (\text{IV.22})$$

$$\mathbf{V}_{mp}^{\text{MODAL}}(k) = -j4k\pi(j_m(k\bar{r}_1)Y_{mp}^*(\bar{\theta}_1, \bar{\phi}_1), \dots, j_m(k\bar{r}_N)Y_{mp}^*(\bar{\theta}_N, \bar{\phi}_N))^T \quad (\text{IV.23})$$

et s'obtient donc par résolution de

$$\begin{aligned} & \min_{\epsilon \geq 0, \mathbf{W} \triangleq \mathbf{W}(k_0) \in \mathbb{C}^N} \epsilon \\ & \text{sous } \begin{cases} \forall m \in \{0, \dots, M\}, \forall p \in \{-m, \dots, m\}, \\ |\mathbf{W}^T \mathbf{V}_{mp}^{\text{MODAL}}(k_0) - \alpha_{mp}^{\text{REF}}(k_0)|^2 \leq \epsilon \\ \mathbf{W}^T \mathbf{W} \leq \delta. \end{cases} \end{aligned} \quad (\text{IV.24})$$

Pour une synthèse large bande au moyen de filtres RIF dont les coefficients sont rassemblés dans le vecteur \mathbf{W}_{LB} – cf. (IV.9)–(IV.11) – le programme devient

$$\begin{aligned} & \min_{\epsilon \geq 0, \mathbf{W}_{\text{LB}} \in \mathbb{R}^{N(Q+1)}} \epsilon \\ & \text{sous } \begin{cases} \forall m \in \{0, \dots, M\}, \forall p \in \{-m, \dots, m\}, \forall k \in K, \\ \left| \mathbf{W}_{\text{LB}}^T (\mathbf{V}_{mp}^{\text{MODAL}}(k) \otimes \mathbf{V}^{\text{RIF}}(k)) - \alpha_{mp}^{\text{REF}}(k) \right|^2 \leq \epsilon \\ \mathbf{W}_{\text{LB}}^T \mathbf{W}_{\text{LB}} \leq \delta. \end{cases} \end{aligned} \quad (\text{IV.25})$$

Les développements précédents appellent plusieurs remarques fondamentales.

- En premier lieu, (IV.24) et (IV.25) sont encore des SOCPs. Aucune grille sur les variables angulaires (θ, ϕ) n'est requise. Le cas large bande ne nécessite qu'une grille en fréquences. Le nombre de contraintes est donc considérablement réduit, et le conditionnement s'en trouve amélioré.
- Comme cela a été déjà mentionné, l'approche modale unifie dans un même formalisme les réponses d'antenne pour différentes valeurs de r . Ainsi, soit $\bar{D}(\theta, \phi, k)$ un gabarit de référence en azimuts, élévations et fréquences, et soient $\bar{M}_{mp}(k) \triangleq \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \bar{D}(\theta, \phi, k) Y_{mp}^*(\theta, \phi) \sin \theta d\theta d\phi$ les coefficients de sa décomposition sur les harmoniques sphériques. Pour approcher ce gabarit à la distance r_0 , i.e. $\forall \theta, \phi, k, D^{\text{REF}}(r_0, \theta, \phi, k) \equiv \bar{D}(\theta, \phi, k)$, les coefficients modaux de référence $\{\alpha_{mp}^{\text{REF}[r_0]}(\cdot)\}_{m \in \mathbb{N}, p \in \{-m, \dots, +m\}}$ se déduisent de $\{\bar{M}_{mp}(\cdot)\}_{m \in \mathbb{N}, p \in \{-m, \dots, +m\}}$ par la relation $\forall m, p, k, \alpha_{mp}^{\text{REF}[r_0]}(k) = \frac{\bar{M}_{mp}(k)}{R_m(r_0, k)}$. Si l'objectif est la même synthèse en

r_1 – possiblement $r_1 = +\infty -$, i.e. $\forall \theta, \phi, k, D^{\text{REF}}(r_1, \theta, \phi, k) \equiv \bar{D}(\theta, \phi, k)$, alors les coefficients modaux de référence se déduisent des précédents par la formule $\forall m, p, k, \alpha_{mp}^{\text{REF}[r_1]}(k) = \alpha_{mp}^{\text{REF}[r_0]}(k) \frac{R_m(r_0, k)}{R_m(r_1, k)}$.

- De manière symétrique, si $\{\alpha_{mp}^{[r_0]}(.)\}_{m \in \mathbb{N}, p \in \{-m, \dots, +m\}}$ désignent les coefficients modaux de la réponse d'une antenne approchant au mieux un gabarit donné à la distance r_0 , alors cette même antenne admet en r_1 une réponse dont les coefficients modaux s'écrivent $\forall m, p, k, \alpha_{mp}^{[r_1]}(k) = \alpha_{mp}^{[r_0]}(k) \frac{R_m(r_0, k)}{R_m(r_1, k)}$. Ceci constitue un volet du phénomène de *réciprocité radiale* traité plus en détail dans [Kennedy 99].

C Cas d'une antenne linéaire Dans le cas d'une antenne linéaire constituée de N microphones d'abscisses $\bar{z}_1, \dots, \bar{z}_N$ selon l'axe (O, \vec{z}) , il vient

$$D(r, \theta, k) = \sum_{m=0}^{\infty} \alpha_m(k) R_m(r, k) Y_m(\theta), \quad (\text{IV.26})$$

où

$$Y_m(\theta) \triangleq Y_{m0}(\theta, \phi) = \sqrt{\frac{2m+1}{4\pi}} P_{m0}(\cos \theta), \quad (\text{IV.27})$$

$$\begin{aligned} \alpha_m(k) &\triangleq \alpha_{m0}(k) = \frac{2\pi}{R_m(r, k)} \int_{\theta=0}^{\pi} D(r, \theta, k) Y_m(\theta) \sin \theta \, d\theta \\ &= -2jk\sqrt{\pi(2m+1)} \sum_{n=1}^N W_n(k) j_m(k\bar{z}_n). \end{aligned} \quad (\text{IV.28})$$

Par ailleurs, nous avons montré dans [CIACL-10-AR-3] que si on se donne une fonction $\bar{D}(\theta, k)$ symétrique par rapport à une direction de polarisation θ_0 , alors la troncature $\bar{D}_L^F(\theta, k) = \sum_{l=0}^L a_l(k) \cos(l\theta)$ à l'ordre L de son développement en série de Fourier admet une projection nulle sur les harmoniques sphériques $Y_m(\theta)$ d'ordre $m > L$. Dès lors, si pour une distance r_0 donnée on définit le gabarit de référence comme $\forall \theta, k, D^{\text{REF}}(r_0, \theta, k) \equiv \bar{D}_L^F(\theta, k)$, alors les coefficients modaux $\alpha_m^{\text{REF}[r_0]}(k) = \frac{\sqrt{\pi(2m+1)}}{R_m(r_0, k)} \sum_{l=0}^L a_l(k) \int_{\theta=0}^{\pi} \cos(l\theta) P_{m0}(\cos \theta) \sin \theta \, d\theta, m \in \mathbb{N}$, sont tels que $\forall m > L, \alpha_m^{[r_0]}(k) \equiv 0$. Le SOCP (IV.24) pour la synthèse faible bande devient

$$\begin{aligned} \min_{\epsilon \geq 0, \mathbf{W} \triangleq \mathbf{W}(k_0) \in \mathbb{C}^N} \quad & \epsilon \\ \text{sous} \quad & \left\{ \begin{array}{l} \forall m \in \{0, \dots, M\}, M \geq L, \\ |\mathbf{W}^T \mathbf{V}_m^{\text{MODLIN}}(k_0) - \alpha_m^{\text{REF}}(k_0)|^2 \leq \epsilon \\ \mathbf{W}^T \mathbf{W} \leq \delta. \end{array} \right. \end{aligned} \quad (\text{IV.29})$$

avec

$$\mathbf{V}_m^{\text{MODLIN}}(k) = -2jk\sqrt{\pi(2m+1)}(j_m(k\bar{z}_1), \dots, j_m(k\bar{z}_N))^T. \quad (\text{IV.30})$$

Le nombre de contraintes est réduit de $(M+1)^2$ à $(M+1)$, avec M nécessairement supérieur

ou égal à L . La synthèse large bande (IV.25) se simplifie également en

$$\begin{aligned} & \min_{\epsilon \geq 0, \mathbf{W}_{LB} \in \mathbb{R}^{N(Q+1)}} \epsilon \\ & \text{sous } \left\{ \begin{array}{l} \forall m \in \{0, \dots, M\}, M \geq L, \forall k \in K, \\ \quad \left| \mathbf{W}_{LB}^T (\mathbf{V}_m^{\text{MODLIN}}(k) \otimes \mathbf{V}^{\text{RIF}}(k)) - \alpha_m^{\text{REF}}(k) \right|^2 \leq \epsilon \\ \mathbf{W}_{LB}^T \mathbf{W}_{LB} \leq \delta. \end{array} \right. \end{aligned} \quad (\text{IV.31})$$

Il convient de noter qu'une formation de voie invariante en fréquence synthétisée en champ lointain sur la base d'une antenne linéaire uniforme conduit à un diagramme de directivité très fortement distordu lorsque la distance aux sources diminue. Les possibilités offertes par notre algorithme ainsi qu'une comparaison des cartes de localisation théoriques obtenues par formation de voie conventionnelle sont illustrées Figure IV.2.

À l'usage, on note qu'une forte pénalisation de la borne supérieure δ de $\mathbf{W}_{LB}^T \mathbf{W}_{LB}$ dans (IV.31) conduit à un diagramme qui certes s'éloigne du gabarit de référence, mais dans des proportions significativement moindres que lors de la formation de voie « naïve » (IV.13). En contrepartie, la synthèse d'un diagramme en gain seulement semble hors de portée.

IV.3.2.4 Implémentation pratique d'une formation de voie

Cette section recense quelques réflexions suscitées par la nécessité d'intégrer et de faire fonctionner en conditions réelles nos algorithmes de formation de voie.

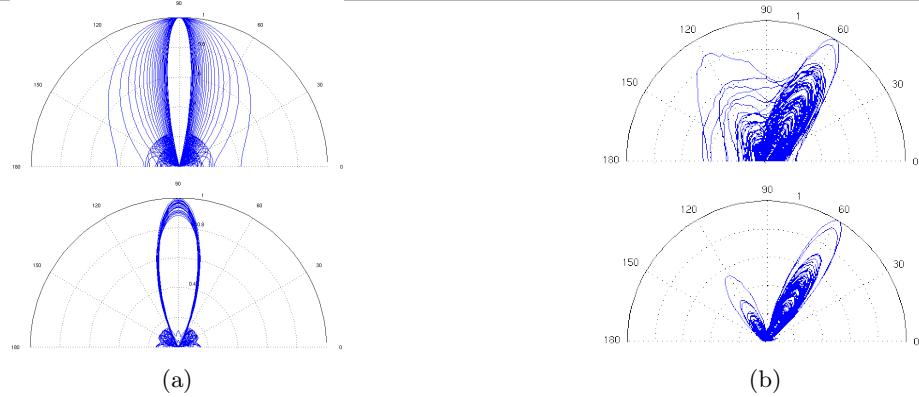
A Introduction de considérations pratiques dans la définition du problème de synthèse Nos algorithmes ne contraignent le diagramme de directivité qu'à l'intérieur d'un intervalle de fréquences. Or, on assiste parfois à une explosion du gain de l'antenne hors de ce domaine, principalement aux hautes fréquences. Nous avons donc étudié les implications pratiques de ce constat.

A.1 Diagramme de directivité théorique et filtrage spatial effectif Notre premier objectif a été de relier la réponse théorique d'une antenne et le comportement effectif de son implémentation [CIACL-09-AR-2], schématisée Figure IV.3. On utilise ponctuellement la fréquence temporelle f au lieu de sa contrepartie spatiale k , où, pour mémoire, $f = \frac{kc}{2\pi}$. $ACQ(f)$ désigne la fonction de transfert globale de la chaîne d'acquisition analogique ; elle contient donc au moins une coupure haute fréquence abrupte liée à la présence d'un filtre anti-repliement. En aval des N échantilleurs synchrones à la période T_e se trouvent des filtres numériques $NUM(f)$ – typiquement passe-bande – puis les filtres numériques $\{W_n(f)\}_{n=1,\dots,N}$ de la formation de voie[†].

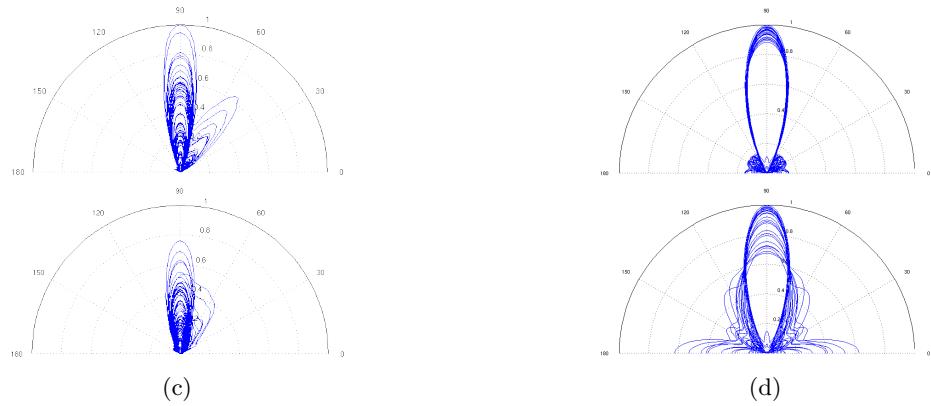
Soient $s(t)$ le signal présent en l'origine O du repère \mathcal{F} et $y[m]$ le signal numérique en sortie de formation de voie. On montre que les transformées $S(f)$ de $s(t)$ et $Y^{\dagger\dagger\dagger}(f)$ de $y^{\dagger\dagger\dagger}(t) = \sum_m y[m] \delta(t - mT_e)$, avec $\delta(\cdot)$ l'impulsion de Dirac, sont unies par

$$Y^{\dagger\dagger\dagger}(f) = \frac{1}{T_e} \sum_m Y^\sharp(f - \frac{m}{T_e}), \quad (\text{IV.32})$$

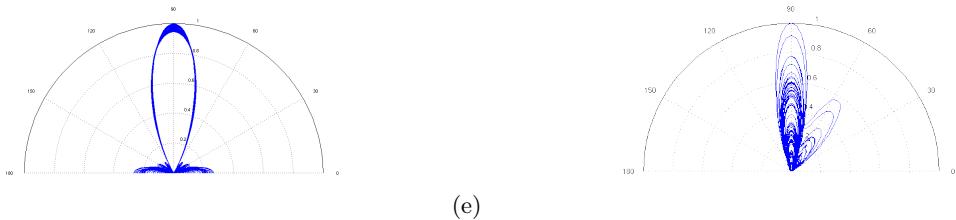
[†]. De par leur nature échantillonnée, $NUM(\cdot)$ et $W_n(\cdot)$, $n = 1, \dots, N$, dépendent de f via $e^{2j\pi f T_e}$.



(a) Diagramme de directivité en azimuth d'une formation de voie conventionnelle (haut) vs optimisée invariante en fréquences (bas) sous l'hypothèse de champ lointain, pour la polarisation $\theta_0 = 90^\circ$. Chaque courbe correspond à une fréquence. (b) Cartes d'énergie acoustique résultantes pour deux sources à l'infini et aux azimuths 60° et 120° . Chaque courbe est une carte calculée sur un "snapshot" temporel.



(c) Cartes d'énergie acoustique obtenues au moyen de formations de voie optimisées, synthétisées sous l'hypothèse de champ lointain. Les deux sources se trouvent à 50° et 90° , à la distance $r_0 = +\infty$ vs à $r_1 = 0.8$ m (bas). (d) La réponse d'antenne est effectivement invariante en fréquences pour $r_0 = +\infty$ (cf. haut pour $\theta_0 = 90^\circ$), mais se distord lorsque la source se rapproche en r_1 (bas) : les basses fréquences se trouvent en champ proche, ce qui explique la mauvaise localisation.



(e) Formation de voie optimisée invariante en fréquences synthétisée pour $r_1 = 0.8$ m (gauche). Carte d'énergie obtenue pour deux sources à 50° et 90° , et $r_1 = 0.8$ m (droite, à comparer avec (c)-bas).

FIGURE IV.2 – Étude simulée de cartes d'énergie acoustique obtenues par formations de voie conventionnelles et optimisées, pour l'antenne linéaire uniforme décrite au §IV.3.1.2 avec $M + 1 = L + 1 = 9$, $Q = 35$, $\delta = 1000$.

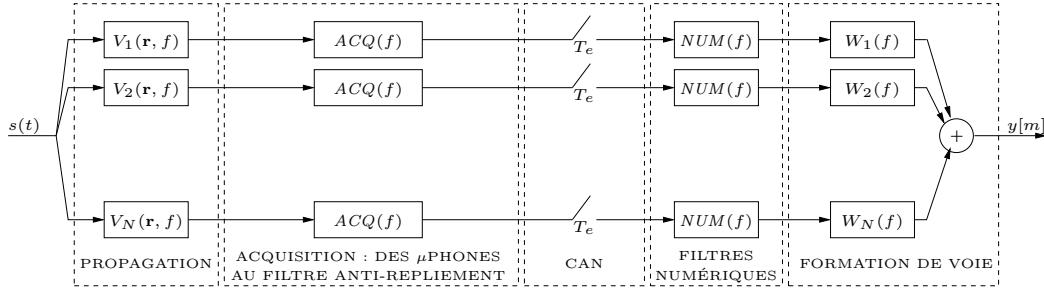


FIGURE IV.3 – Implémentation pratique d'une formation de voie.

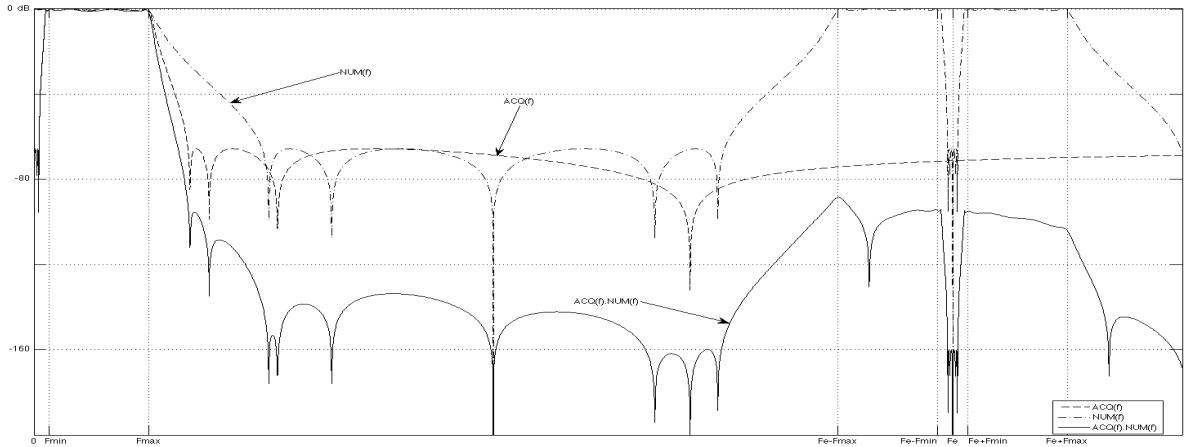


FIGURE IV.4 – Transferts impliqués dans l'implémentation pratique d'une formation de voie.

$$\text{avec } Y^\sharp(f) = D(\mathbf{r}, f)NUM(f)ACQ(f)S(f). \quad (\text{IV.33})$$

Même si les transferts $ACQ(f)$ et $NUM(f)$ sont égaux à 1, la relation fondamentale (IV.7) n'est pas en vigueur. En effet, le contenu fréquentiel $Y^{\sharp\sharp}(f)$ de la sortie de la formation de voie est obtenu par repliement et périodisation du contenu fréquentiel $Y^\sharp(f)$ du résultat du filtrage de $s(t)$ par la réponse d'antenne $D(\mathbf{r}, f)$. Il s'en suit que toute explosion numérique du diagramme d'antenne en hautes fréquences vient, par repliement, modifier les caractéristiques attendues pour les fréquences d'intérêt. À notre connaissance, cette conclusion importante ne figure pas dans la littérature.

A.2 Vers une révision du programme d'optimisation ? Comme indiqué précédemment, le diagramme de directivité n'est contraint à s'approcher du gabarit que sur la bande de fréquences spatiales K ; lorsque les sources sont des voix humaines, un intervalle de fréquences temporelles $F = [F_{MIN}; F_{MAX}]$ équivalent est typiquement $F = [400\text{Hz}; 3\text{kHz}]$. Or, $ACQ(f)$ et $NUM(f)$ permettent le filtrage de fréquences situées dans le complément \bar{F} de F , qui pourraient s'avérer problématiques lors du repliement et de la périodisation (IV.32) de $Y^\sharp(f)$. Le diagramme de gain de leur produit $ACQ(f)NUM(f)$ –en particulier sur $[F_e - F_{MAX}; F_e - F_{MIN}]$, avec $F_e = \frac{1}{T_e}$ la fréquence d'échantillonnage– renseigne sur l'explosion admissible du diagramme de directivité hors de la bande F (Figure IV.4).

La tentation est grande d'intégrer ces considérations pratiques à la formulation du problème

de synthèse [CIACL-12-AR-4]. Deux degrés de liberté supplémentaires demeurent alors exploitables. D'une part, l'augmentation de la fréquence d'échantillonnage F_e peut doter $ACQ(f)NUM(f)$ d'une meilleure réjection sur la bande $[F_e - F_{\text{MAX}}; F_e - F_{\text{MIN}}]$, et ainsi limiter les effets du repliement. D'autre part, il convient de placer sur $\mathbf{W}_{\text{LB}}^T \mathbf{W}_{\text{LB}}$ la plus grande borne δ qui cantonne dans les limites acceptables l'explosion du diagramme de directivité sur \bar{F} . Autrement dit, la borne δ ne doit pas être inutilement contraignante, sous peine d'éloigner le diagramme en cours de synthèse du gabarit sur la bande F . Bien que ce point de vue soit « peu académique », il constitue, avec la sensibilité de l'antenne au bruit, un guide pertinent pour la sélection de δ , pouvant conduire à un meilleur conditionnement du programme d'optimisation.

Une autre approche serait que la structure de l'antenne impose une réponse invariante en fréquence [Ward 95]. Hélas, cette alternative ne concerne pas les antennes linéaires uniformes.

B Formation de voie par techniques de convolution rapide L'expérience montre que pour une fréquence d'acquisition $F_e = 15 \text{ kHz}$, des filtres RIF d'ordre $Q = 35$ suffisent à synthétiser un diagramme de directivité invariant sur $F = [400\text{Hz}; 3\text{kHz}]$ par notre approche. L'insertion sur chaque voie d'acquisition d'un filtre RII d'ordre 8 ou RIF d'ordre 48 suffit à ne laisser passer que la bande F pour limiter les effets du repliement. Si on sélectionne cette dernière option afin de simplifier l'implémentation, alors $N = 8$ filtres RIF d'ordre 83 sont impliqués dans la polarisation de l'antenne pour chaque azimuth θ_0 . La polarisation selon $180^\circ - \theta_0$ s'effectue sans nouvelle synthèse, par simple permutation des filtres entre les $n^{\text{ème}}$ et $(N + 1 - n)^{\text{ème}}$ microphones, $n = 1, \dots, N$.

Ainsi, l'extraction en temps réel selon chaque direction d'écoute requiert $N = 8$ convolutions indépendantes par une séquence de $R = 84$ échantillons à la cadence de 15 kHz . Pour la génération de cartes d'énergie acoustique à environ 15 Hz , au moins $P = 50$ formations de voie synthétisées hors ligne doivent être exécutées sur la base des signaux acquis à 15 kHz , *e.g.* $P = 51$ pour des angles de polarisation $\{0^\circ, 5^\circ, 10^\circ, \dots, 45^\circ\} \cup \{50^\circ, 52^\circ, 54^\circ, \dots, 90^\circ\}$ et leurs supplémentaires. Accessoirement, leurs sorties doivent être mémorisées sur une fenêtre temporelle glissante de 1024 échantillons en vue de procéder au calcul des énergies incidentes.

Or, une convolution linéaire de deux séquences temporelles de longueur R admet une complexité en $O(R^2)$. Pour une solution « fréquentielle » consistant en le calcul du produit des transformées de Fourier rapides (FFT = Fast Fourier Transform) de ces séquences préalablement à sa transformation inverse (IFFT = Inverse Fast Fourier Transform), la complexité est en $O(R \log_2(R))$. Pour des ordres R suffisamment élevés, la solution fréquentielle est à coup sûr avantageuse. Toutefois, ces complexités théoriques doivent être modulées selon le matériel supportant les opérations arithmétiques : si la convolution dans le domaine temporel est effectuée sur un FPGA embarquant des coeurs de DSP, alors la solution fréquentielle peut s'avérer intéressante seulement pour des filtres d'ordre $(R - 1)$ supérieur à 50 [Borgerding 06]. Bien que notre capteur EAR (§IV.3.5) soit construit autour d'un tel FPGA, il est certain que la solution fréquentielle l'emporte. En effet, lorsque l'antenne est polarisée simultanément selon P directions, le signal perçu par chaque transducteur alimente P filtres. Ceci permet de « factoriser » le calcul de sa FFT dans les opérations, et conduit à un gain décisif en performances.

Il demeure que l'IFFT du produit des FFTs de deux séquences temporelles ne réalise pas leur convolution linéaire, mais leur convolution circulaire[†]. Si le filtre RIF est d'ordre

[†]. C'est pourquoi nous n'avons pas envisagé d'implémentation « fréquentielle » d'une formation de voie

($R - 1$), alors les valeurs de l'IFFT correspondant aux ($R - 1$) dates les plus anciennes sont erronées. Des techniques de *convolution rapide*, sur la base de séquences atomiques [FFT→produit→IFFT] ont été développées pour pallier ce problème. Parmi elles, l'approche *overlap-and-save* est particulièrement adaptée à nos besoins. Elle consiste à répartir chaque signal temporel à filtrer dans des blocs de taille $L \gg (R - 1)$ qui se chevauchent de $L_{ov} \geq (R - 1)$ échantillons. Les FFTs de ces blocs sont effectuées séquentiellement. Elles sont multipliées par les FFTs des réponses impulsionales des filtres RIF correspondants, lesquelles sont calculées hors ligne sur une fenêtre de taille L par bourrage de zéros. Ainsi, malgré l'élimination des ($R - 1$) premiers échantillons du résultat de la IFFT, l'ensemble du résultat de la convolution linéaire s'obtient sans erreur, bien qu'avec une latence de $L - L_{ov}$ échantillons. Dans notre implémentation, nous avons posé $L = 256$ et $L_{ov} = 128$. Ce choix offre un gain de performance et une latence acceptables, et met à disposition de l'utilisateur sans calcul supplémentaire des FFTs sur des blocs temporels contigus de taille 256 pour d'autres utilisations (§IV.3.3.5-A).

IV.3.3 Méthode à haute résolution pour la localisation et la détection de sources sonores large bande

IV.3.3.1 Introduction

La Figure IV.2 page 121 illustre le bénéfice apporté par la formation de voie invariante en fréquence pour la construction de cartes d'énergie acoustique. Toutefois, malgré son coût calculatoire maîtrisé et le fait qu'elle n'exige aucune hypothèse sur la nature du bruit ambiant, cette approche de la localisation comporte plusieurs inconvénients.

- Une antenne synthétisée par formation de voie conventionnelle en champ lointain admet une réponse semblable – et tout aussi peu satisfaisante – si la source est située en $r \neq +\infty$. Cependant, un diagramme de directivité invariant en fréquences est fortement distordu lorsque la distance à la source est très différente de celle supposée lors de son optimisation. Une localisation satisfaisante nécessite donc que les sources soient situées à une même distance connue d'avance.
- Si, en environnement réel, la puissance des poids des filtres n'est pas suffisamment limitée, alors la carte de localisation est en fait une « signature acoustique » de l'environnement, dépendant davantage du bruit ambiant que des sources. La borne δ doit alors être drastiquement réduite, au risque d'une résolution médiocre.
- Sur le plan théorique, la capacité de focalisation spatiale de l'antenne, et par conséquent la résolution de la localisation, sont fondamentalement limitées par son ouverture [Krim 96][Williams 07].

Parmi les alternatives « à haute résolution »[Gonen 99], nous avons étudié plus particulièrement la méthode MUSIC (MUltiple SIgnal Classification) [Schmidt 86], qui est relativement générique et a fait l'objet de nombreuses analyses et extensions. Le prix à payer pour le bénéfice prévisible est le coût calculatoire ainsi que la connaissance *a priori* les statistiques du bruit. Ce sont naturellement des « leviers » sur lesquels nous avons tenté de jouer.

large bande, *i.e.* par IFFT des sorties de formations de voie faible bande alimentées par les FFTs des N canaux d'entrée.

IV.3.3.2 Déclinaisons de la méthode MUSIC

MUSIC et ses déclinaisons présentées ci-après exigent la donnée de l'expression analytique du vecteur d'antenne, ou bien sa détermination expérimentale pour un ensemble discret suffisamment riche de positions et de fréquences. L'antenne ne doit pas être ambiguë, *i.e.* le rang de la matrice $\underline{\mathbf{V}}(\mathbf{r}_1, \dots, \mathbf{r}_D, k)$ doit être plein, et égal au nombre D – donné *a priori* – de sources, pour tous $\mathbf{r}_1 \neq \mathbf{r}_2 \neq \dots \neq \mathbf{r}_D$. Les processus $\mathbf{s}(t)$ et $\mathbf{n}(t)$ sont supposés mutuellement indépendants, centrés, stationnaires et ergodiques sur la fenêtre temporelle d'étude. La matrice de covariance $\underline{\mathbf{C}}_{\mathbf{s}}(k) \triangleq \mathbb{E}[\mathbf{S}(k)\mathbf{S}^H(k)]$ est inconnue. Le contenu fréquentiel et la distribution spatiale du bruit $\mathbf{n}(t)$ sont quelconques, et $\underline{\mathbf{C}}_{\mathbf{n}}(k) \triangleq \mathbb{E}[\mathbf{N}(k)\mathbf{N}^H(k)] = \sigma_{\mathbf{n}}^2 \underline{\mathbf{C}}_{\mathbf{n}}(k)$ est connue à la constante $\sigma_{\mathbf{n}}^2$ près. On note $\underline{\mathbf{C}}_{\mathbf{x}}(k) \triangleq \mathbb{E}[\mathbf{X}(k)\mathbf{X}^H(k)]$.

A L'algorithme faible bande La version originale [Schmidt 86] de MUSIC suppose que les D sources sont indépendantes et émettent à la fréquence k_0 . Leur matrice de covariance $\underline{\mathbf{C}}_{\mathbf{s}}(k_0)$ est donc de rang D . La matrice de covariance $\underline{\mathbf{C}}_{\mathbf{x}}(k_0)$ est de la forme

$$\underline{\mathbf{C}}_{\mathbf{x}}(k_0) = \underline{\mathbf{V}}(\mathbf{r}_1, \dots, \mathbf{r}_D, k_0) \underline{\mathbf{C}}_{\mathbf{s}} \underline{\mathbf{V}}^H(\mathbf{r}_1, \dots, \mathbf{r}_D, k_0) + \sigma_{\mathbf{n}}^2 \underline{\mathbf{C}}_{\mathbf{n}}(k_0). \quad (\text{IV.34})$$

La décomposition en valeurs propres généralisées (GEVD = Generalized EigenValue Decomposition) du faisceau matriciel $(\underline{\mathbf{C}}_{\mathbf{x}}(k_0), \underline{\mathbf{C}}_{\mathbf{n}}(k_0))$ s'écrit

$$\underline{\mathbf{C}}_{\mathbf{x}}(k_0) = \sum_{i=1}^N \lambda_i \underline{\mathbf{C}}_{\mathbf{n}}(k_0) \mathbf{U}_i \mathbf{U}_i^H \underline{\mathbf{C}}_{\mathbf{n}}^H(k_0), \quad (\text{IV.35})$$

où les valeurs propres généralisées satisfont $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_D > \lambda_{D+1} = \dots = \lambda_N = \sigma_{\mathbf{n}}^2$. Leurs vecteurs propres associés $\mathbf{U}_1, \dots, \mathbf{U}_N$ peuvent être sélectionnés de sorte que matrices $\underline{\mathbf{U}}_{\mathcal{S}} = (\mathbf{U}_1 | \dots | \mathbf{U}_D) \in \mathbb{C}^{N \times D}$ et $\underline{\mathbf{U}}_{\mathcal{N}} = (\mathbf{U}_{D+1} | \dots | \mathbf{U}_N) \in \mathbb{C}^{N \times (N-D)}$ vérifient $(\underline{\mathbf{U}}_{\mathcal{S}} | \underline{\mathbf{U}}_{\mathcal{N}})^H \underline{\mathbf{C}}_{\mathbf{n}}(k_0) (\underline{\mathbf{U}}_{\mathcal{S}} | \underline{\mathbf{U}}_{\mathcal{N}}) = \mathbb{I}_N$. Ainsi, le *sous-espace signal* \mathcal{S} généré par les colonnes de $\underline{\mathbf{U}}_{\mathcal{S}}$ est orthogonal, dans la métrique de $\underline{\mathbf{C}}_{\mathbf{n}}(k_0)$, à l'espace image \mathcal{N} de $\underline{\mathbf{U}}_{\mathcal{N}}$, appelé *sous-espace bruit*. Or, le vecteur d'antenne satisfait $\mathbf{V}^H(\mathbf{r}, k_0) \underline{\mathbf{U}}_{\mathcal{N}} = \mathbf{0}^T$ en les positions $\mathbf{r} \in \{\mathbf{r}_1, \dots, \mathbf{r}_D\}$ des sources. Par conséquent, le *pseudo-spectre*

$$h^{[k_0]}(\mathbf{r}) \triangleq \frac{1}{\mathbf{V}^H(\mathbf{r}, k_0) \underline{\Pi}_{\mathcal{N}} \mathbf{V}(\mathbf{r}, k_0)}, \quad (\text{IV.36})$$

avec $\underline{\Pi}_{\mathcal{N}} \triangleq \sum_{i=D+1}^N \underline{\mathbf{U}}_i \underline{\mathbf{U}}_i^H$ le *projecteur sur l'espace bruit* admet des pics infinis en les positions des sources. Lorsque le bruit est spatialement blanc, *i.e.* $\underline{\mathbf{C}}_{\mathbf{n}}(k_0) = \mathbb{I}_N$, $h^{[k_0]}(\mathbf{r})$ n'est autre que l'inverse de la distance Euclidienne quadratique entre le vecteur d'antenne $\mathbf{V}(\mathbf{r}, k_0)$ et le sous-espace signal \mathcal{S} .

En pratique, les covariances théoriques $\underline{\mathbf{C}}_{\mathbf{x}}(k_0)$ et $\underline{\mathbf{C}}_{\mathbf{n}}(k_0)$ ne sont pas connues. Seuls des estimés $\tilde{\underline{\mathbf{C}}}_{\mathbf{x}}(k_0)$, $\tilde{\underline{\mathbf{C}}}_{\mathbf{n}}(k_0)$ peuvent être calculés sur la base des séquences temporelles perçues. Un projecteur approché $\tilde{\underline{\Pi}}_{\mathcal{N}}$ est donc utilisé dans (IV.36) au lieu de $\underline{\Pi}_{\mathcal{N}}$. Les modes du pseudo-spectre pratique $\tilde{h}^{[k_0]}(\mathbf{r})$ sont alors finis, mais très fortement marqués. Si $\tilde{\underline{\mathbf{C}}}_{\mathbf{x}}(k_0)$ et $\tilde{\underline{\mathbf{C}}}_{\mathbf{n}}(k_0)$ sont asymptotiquement non biaisés, alors la localisation l'est également.

B Une extension large bande « naïve » Une extension immédiate de MUSIC au cas de sources large bande consiste à partitionner leur bande utile en B “bins” k_1, \dots, k_B , puis

à appliquer les résultats du §IV.3.3.2–A sur chacun d'eux. Une difficulté théorique majeure vient de ce que les sous-espaces signal \mathcal{S} et bruit \mathcal{N} varient selon les bins. De ce fait, on est amené à définir le pseudo-spectre large bande comme la moyenne –ou la somme– des pseudo-spectres faible bande relatifs à k_1, \dots, k_B , *i.e.*

$$h_{\text{LB NAIF}}(\mathbf{r}) \triangleq \sum_{b=1}^B h^{[k_b]}(\mathbf{r}). \quad (\text{IV.37})$$

Cette extension conduit à de bons résultats mais présente trois faiblesses majeures.

- Les matrices de covariance $\underline{\mathbf{C}}_s(k_b) = \mathbb{E}[\mathbf{S}(k_b)\mathbf{S}^H(k_b)]$ doivent être de rang plein pour tous $b \in \{1, \dots, B\}$. Ceci exclut tout traitement étayé de problèmes de Robotique où certaines sources sont cohérentes du fait de réverbérations.
- Le calcul de (IV.37) exige B GEVDs de faisceaux matriciels de $\mathbb{C}^{N \times N}$. Pour donner un ordre d'idée, si les $N = 8$ voies échantillonnées à $F_e = 15$ kHz sont décomposées par des FFTs calculées sur 256 points, alors la bande $F = [400\text{Hz}; 3\text{kHz}]$ comprend 45 bins. $B = 45$ GEVDs de faisceaux de $\mathbb{C}^{8 \times 8}$ sont donc requises pour chaque définition de la fonction pseudo-spectre.
- Enfin, aucune structure algébrique –faisceau matriciel, valeurs propres, sous-espace signal, sous-espace bruit– n'est exhibée, qui soit liée aux positions des sources.

C Une extension large bande cohérente Alors que l'équation (IV.37) constituait la base de la seule application de MUSIC à la Robotique [Asano 99], nous avons étudié les fondements des *extensions large bande cohérentes* de MUSIC, posés dans [Wang 85].

L'idée est de définir B matrices de focalisation $\underline{\mathbf{T}}(k_b) \in \mathbb{C}^{Q \times N}$, $b = 1, \dots, B$, de rang plein Q , où $D < Q \leq N$, et telles que pour une fréquence de référence k_0 ,

$$\forall \mathbf{r}, \underline{\mathbf{T}}(k_b)\mathbf{V}(\mathbf{r}, k_b) = \underline{\mathbf{T}}(k_0)\mathbf{V}(\mathbf{r}, k_0). \quad (\text{IV.38})$$

Soit $\mathbf{Z}(k) \triangleq \underline{\mathbf{T}}(k)\mathbf{X}(k)$, $\mathbf{Z}(k) \in \mathbb{C}^Q$. La somme des covariances de $\mathbf{Z}(k_b)$ sur les B bins constitue la matrice de covariance alignée de \mathbf{Z} , d'expression

$$\underline{\Gamma}_{\mathbf{z}} \triangleq \sum_{b=1}^B \underline{\mathbf{T}}(k_b)\underline{\mathbf{C}}_{\mathbf{x}}(k_b)\underline{\mathbf{T}}^H(k_b) = \underline{\mathbf{T}}(k_0)\mathbf{V}(\mathbf{r}, k_0)\underline{\Gamma}_{\mathbf{s}}\mathbf{V}^H(\mathbf{r}, k_0)\underline{\mathbf{T}}^H(k_0) + \sigma_{\mathbf{n}}^2\underline{\Gamma}_{\mathbf{n}}, \quad (\text{IV.39})$$

$$\text{avec } \underline{\Gamma}_{\mathbf{s}} \triangleq \sum_{b=1}^B \underline{\mathbf{C}}_{\mathbf{x}}(k_b) \text{ et } \underline{\Gamma}_{\mathbf{n}} \triangleq \sum_{b=1}^B \underline{\mathbf{T}}(k_b)\underline{\mathbf{C}}_{\mathbf{n}}(k_b)\underline{\mathbf{T}}^H(k_b). \quad (\text{IV.40})$$

$\underline{\Gamma}_{\mathbf{n}}$ est dite matrice de covariance alignée du bruit. La GEVD $\{\lambda_i, \mathbf{U}_i\}_{i=1,\dots,Q}$ de $(\underline{\Gamma}_{\mathbf{z}}, \underline{\Gamma}_{\mathbf{n}})$ vérifie $\lambda_1 \geq \dots \geq \lambda_D > \lambda_{D+1} = \dots = \lambda_Q = \sigma_{\mathbf{n}}^2$ et $\mathbf{V}^H(\mathbf{r}, k_0)\underline{\mathbf{T}}^H(k_0)\mathbf{U}_{D+1} = \dots = \mathbf{V}^H(\mathbf{r}, k_0)\underline{\mathbf{T}}^H(k_0)\mathbf{U}_Q = \mathbf{0}^T$. Elle permet la définition d'un sous-espace signal (resp. bruit) unique, combinant de manière cohérente les sous-espaces signal (resp. bruit) sur les divers bins. Il en résulte le pseudo-spectre suivant pour la localisation de $D < Q$ sources :

$$h_{\text{LB COHERENT}}(\mathbf{r}) \triangleq \frac{1}{\mathbf{V}^H(\mathbf{r}, k_0)\underline{\mathbf{T}}^H(k_0)\underline{\Pi}_{\mathcal{N}}\underline{\mathbf{T}}(k_0)\mathbf{V}(\mathbf{r}, k_0)}, \text{ avec } \underline{\Pi}_{\mathcal{N}} \triangleq \sum_{i=D+1}^Q \mathbf{U}_i \mathbf{U}_i^H. \quad (\text{IV.41})$$

En pratique, les matrices de covariance alignées théoriques sont remplacées par leurs estimées, et les commentaires du §IV.3.3.2–A demeurent valides. Il convient toutefois de noter que, contrairement à l'extension « naïve », $(\underline{\Gamma}_z, \underline{\Gamma}_n)$ possède une structure voisine du faisceau $(\underline{\mathbf{C}}_x(k_0), \underline{\mathbf{C}}_n(k_0))$ entrant en jeu dans l'algorithme faible bande. Le problème est bien posé si et seulement si $\underline{\Gamma}_s \triangleq \sum_{b=1}^B \underline{\mathbf{C}}_S(k_b)$ est de rang plein D . Cette condition est généralement vérifiée, y compris lorsque le rang de $\underline{\mathbf{C}}_S(k_1), \dots, \underline{\mathbf{C}}_S(k_B)$ n'est pas maximal du fait de multi-trajets. Enfin, l'élaboration du pseudo-spectre ne requiert, outre des sommes et produits matriciels, qu'une GEVD de matrices de $\mathbb{C}^{Q \times Q}$.

IV.3.3.3 Une stratégie cohérente de localisation de sources adaptée à la Robotique

Une extension large bande cohérente de MUSIC exige la définition des matrices de focalisation. Parmi les solutions proposées dans la littérature, nous nous sommes conformés à [Ward 04], qui présente des connexions intéressantes avec nos travaux précédents. La stratégie complète est résumée ci-dessous pour une antenne linéaire uniforme.

Pour chaque hypothèse de distance r , B matrices de focalisation $\underline{\mathbf{T}}(r, k_b) \in \mathbb{C}^{Q \times N}$, $b = 1, \dots, B$, sont définies sur la base des vecteurs de gains $\mathbf{W}_0(r, k_b), \dots, \mathbf{W}_{Q-1}(r, k_b)$ – à valeurs dans \mathbb{C}^N – de Q formations de voie faible bande à la fréquence k_b et à la distance r . Plus exactement, on pose $\underline{\mathbf{T}}(r, k_b) \triangleq \underline{\mathbf{W}}^H(r, k_b)$ avec $\underline{\mathbf{W}}^H(r, k_b) = (\mathbf{w}_0(r, k_b) | \dots | \mathbf{w}_{Q-1}(r, k_b))^H$. Du fait que le vecteur $\mathbf{Z}(r, k) = \underline{\mathbf{T}}(r, k)\mathbf{X}(k) = \underline{\mathbf{W}}^H(r, k)\mathbf{X}(k)$ appartient à l'espace des sorties de Q formations de voie, la méthode est qualifiée de *coherent beamspace MUSIC*.

Les auteurs suggèrent que les Q formations de voies admettent pour diagrammes de directivité les harmoniques sphériques d'ordres croissants, *i.e.*

$$\forall q \in \{0, \dots, Q-1\}, \forall b \in \{1, \dots, B\}, \forall r, \forall \theta,$$

$$D_q(r, \theta, k_b) = \underline{\mathbf{W}}_q^H(r, k_b)\mathbf{V}(\theta, k_b) = Y_q(\theta). \quad (\text{IV.42})$$

Ainsi, les formations de voie constituent une structure orthogonale, qui vérifie en tout r la propriété d'alignement $\forall k_b, \theta, \underline{\mathbf{T}}(r, k_b)\mathbf{V}(r, \theta, k_b) = \underline{\mathbf{T}}(r, k_0)\mathbf{V}(r, \theta, k_0)$. On montre en outre que si la matrice $\underline{\mathbf{W}}^\infty(k_b)$ regroupe l'ensemble des vecteurs de coefficients conduisant aux réponses en champ lointain recherchées – *i.e.* $\forall q, k_b, \theta, D_q(+\infty, \theta, k_b) = Y_q(\theta)$ – alors la matrice $\underline{\mathbf{W}}(r, k_b)$ des coefficients conduisant aux mêmes diagrammes à la distance r – *i.e.* $\forall q, k_b, \theta, D_q(r, \theta, k_b) = Y_q(\theta)$ – se déduit simplement par

$$\underline{\mathbf{W}}(r, k_b) = \underline{\mathbf{W}}^\infty(k_b)\underline{\mathbf{R}}_b^H(r), \quad (\text{IV.43})$$

$$\text{avec } \underline{\mathbf{R}}_b^H(r) = \text{diag}\left(j/(k_b R_0(r, k_b)), \dots, j^Q/(k_b R_{Q-1}(r, k_b))\right), \quad (\text{IV.44})$$

où les fonctions $R_q(., .)$ sont définies en (IV.17) page 117, et l'opérateur $\text{diag}(.)$ construit une matrice diagonale à partir de ses arguments.

Dans la veine de (IV.39)–(IV.40)–(IV.41), un faisceau matriciel $(\underline{\Gamma}_z(r), \underline{\Gamma}_n(r))$ est calculé pour chaque distance r , dont la décomposition en valeurs propres généralisées conduit à la définition d'un pseudo-spectre

$$h_{\text{LB COHERENT}}(r, \theta) = \frac{1}{\mathbf{V}^H(r, \theta, k_0)\underline{\mathbf{W}}(r, k_0)\underline{\Pi}_{\mathcal{N}}(r)\underline{\mathbf{W}}^H(r, k_0)\mathbf{V}(r, \theta, k_0)}. \quad (\text{IV.45})$$

Pour obtenir une méthode constructive complète, il reste à synthétiser les coefficients permettant de conférer à l'antenne des diagrammes de directivité en champ lointain approchant les harmoniques sphériques. Notre méthode de formation de voie large bande par analyse modale et optimisation convexe présentée au §IV.3.2.3 est extrêmement bien adaptée [CIACL-15-AR-5]. En effet, d'après (IV.26) page 119, les coefficients modaux $\{\alpha_m^{\text{REF}_q[+\infty]}(k_b)\}_{m \in \mathbb{N}}$, associés aux Q diagrammes de référence $D_q(+\infty, \theta, k_b) \equiv Y_q(\theta)$, $q = 0, \dots, Q - 1$, s'obtiennent par $\forall q \in \{0, \dots, Q - 1\}$, $\forall b \in \{1, \dots, B\}$, $\forall m \in \mathbb{N}$, $\alpha_m^{\text{REF}_q[+\infty]}(k_b) = \frac{1}{R_m(+\infty, k_b)} \delta_{(q,m)} = \frac{k_b}{j^{m+1}} \delta_{(q,m)}$.

Dans le cas d'une antenne linéaire, il est intéressant de signaler que les modes du pseudo-spectre $h_{\text{LB COHERENT}}(r, \theta)$ peuvent être isolés par la séquence deux optimisations unidimensionnelles, moins coûteuse qu'une exploration de l'espace (r, θ) . En effet, les azimuts des sources s'obtiennent en tant que solutions de $\theta^* = \arg \max_\theta h_{\text{LB COHERENT}}(+\infty, \theta)$, *i.e.* pour $r = +\infty$. Les distances associées vérifient alors $r^* = \arg \max_r h_{\text{LB COHERENT}}(r, \theta^*)$.

IV.3.3.4 Exploitation de MUSIC large bande cohérent pour la détection du nombre de sources

Notre contribution [CIACL-22-AR-7] a été de synthétiser des travaux théoriques développés en Traitement d'Antenne pour la détection du nombre de sources, et de les instancier en une méthode étayée et peu coûteuse pour la Robotique. Les extensions cohérentes large bande de MUSIC constituent un socle théorique particulièrement adapté [Williams 99][Wang 85].

Considérons à nouveau les notations afférentes au §IV.3.3.2-C. En théorie, le nombre de sources D peut être déduit du fait que le faisceau des matrices de covariances alignées $(\underline{\Gamma}_z, \underline{\Gamma}_n)$ comporte $(Q - D)$ valeurs propres minimales identiques. Cette propriété n'est presque sûrement jamais réalisée par le faisceau $(\tilde{\underline{\Gamma}}_z, \tilde{\underline{\Gamma}}_n)$ estimé en pratique. Le partitionnement des valeurs propres de $(\tilde{\underline{\Gamma}}_z, \tilde{\underline{\Gamma}}_n)$ au moyen d'un seuil subjectif n'est pas acceptable. De même, l'hypothèse $d \in \{0, \dots, Q - 1\}$ conduisant au pseudo-spectre dont le mode est le plus prononcé conduirait systématiquement à l'estimé $\hat{D} = (Q - 1)$.

En fait, l'estimée $\tilde{\underline{\Gamma}}_z$ de la covariance alignée de \mathbf{Z} s'obtient à partir des estimés $\tilde{\mathbf{C}}_x(k_b)$ de la covariance $\underline{\mathbf{C}}_x(k_b)$ sur les bins k_1, \dots, k_B . Chaque estimé $\tilde{\mathbf{C}}_x(k_b) = \frac{1}{J} \sum_{j=1}^J \mathbf{X}_{t_j}(k_b) \mathbf{X}_{t_j}^H(k_b)$ combine des réalisations indépendantes $\mathbf{X}_{t_1}(k_b), \dots, \mathbf{X}_{t_J}(k_b)$ de $\mathbf{X}(k_b)$ en J instants t_1, \dots, t_J .

Une position mathématique correcte du problème de détection consiste à considérer $\tilde{\underline{\Gamma}}_z$ comme la somme d'une matrice semi-définie positive de rang égal au nombre, inconnu, de sources, et du produit de $\tilde{\underline{\Gamma}}_n$ par un scalaire également inconnu [Wax 85]. En d'autres termes, $\tilde{\underline{\Gamma}}_z$ est vu comme un échantillon de la famille de matrices

$$\underline{\Gamma}_z^{(d)} = \underline{\Psi}^{(d)} + \sigma^2 \underline{\Gamma}_n, \quad (\text{IV.46})$$

où $\underline{\Psi}^{(d)} = (\underline{\Psi}^{(d)})^H \in \mathbb{C}^{Q \times Q}$ est semi-définie positive et de rang d , et où $d \in \{0, \dots, Q - 1\}$ et $\sigma \in \mathbb{R}$ sont inconnus. L'identification statistique des paramètres d et σ de (IV.46) qui expliquent le mieux $\tilde{\underline{\Gamma}}_z$ ne peut pas s'effectuer au moyen de leurs estimés du maximum de vraisemblance (MLE = Maximum Likelihood Estimator), car le nombre des paramètres libres des modèles (IV.46) varie selon les valeurs admissibles de d en compétition [Akaike 74]. Dans une telle situation, la recherche du minimum d'un estimé de la divergence de Kullback-Leibler, qui sous-tend l'obtention du MLE dans le cas d'un seul modèle, conduit à l'estimé du minimum du critère d'information d'AKAIKE (AIC = AKAIKE Information Criterion, MAICE = Minimum AKAIKE Information Criterion Estimate). Une alternative est l'estimé du critère de longueur

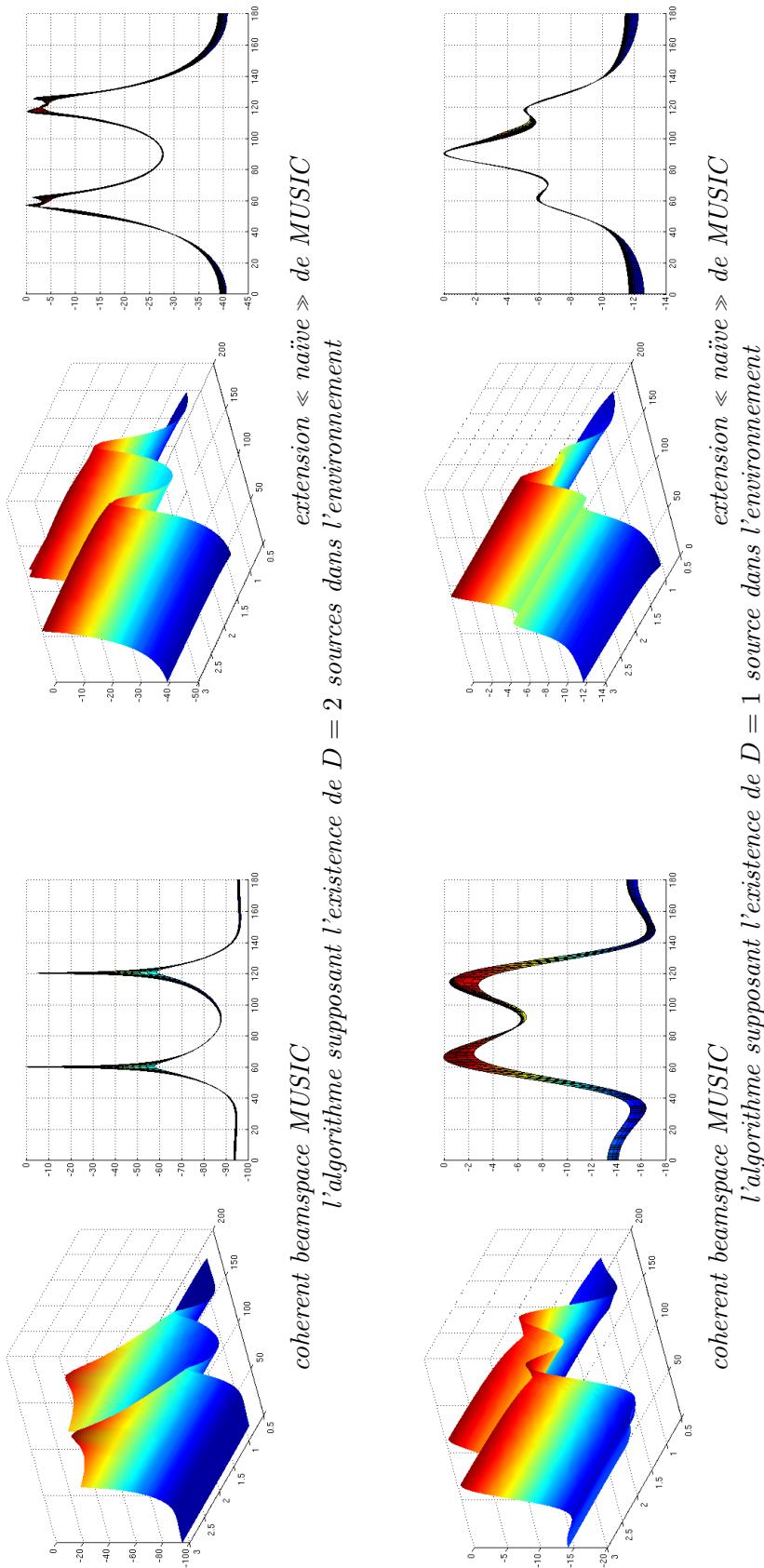


FIGURE IV.5 – Pseudo-spectres normalisés, en dB, obtenus en simulation pour une source S1 située en $(r_1, \theta_1) = (2\text{m}, 60^\circ)$ et une source S2, version retardée – donc cohérente – de S1, située en $(r_2, \theta_2) = (2\text{m}, 120^\circ)$. Le rapport signal sur bruit est de 10dB. Pour chaque type d’extension de MUSIC et chaque hypothèse de nombre de sources $D \in \{1, 2\}$ dans l’algorithme, le pseudo-spectre fonction de (r, θ) ainsi que sa coupe en azimut pour $r = r^*$ sont représentés.

de description minimale (MDL = Minimum Description Length), qui permet l'encodage des données observées par le code de longueur minimale [Wax 85].

Au final, l'estimé du nombre de sources se déduit des valeurs propres généralisées $l_1 \geq \dots \geq l_Q$ de $(\tilde{\Gamma}_z, \tilde{\Gamma}_n)$, par l'une des formules suivantes, où J est défini plus haut :

$$\hat{D} = \arg \min_d \text{AIC}(d), \text{ où } \text{AIC}(d) = -2J(Q-d) \ln \mathcal{R} + 2d(2Q-d), \quad (\text{IV.47})$$

$$\hat{D} = \arg \min_d \text{MDL}(d), \text{ où } \text{MDL}(d) = -J(Q-d) \ln \mathcal{R} + \frac{1}{2}d(2Q-d) \ln J, \quad (\text{IV.48})$$

$$\text{et } \mathcal{R} \triangleq \left(\frac{\sqrt[Q-d]{\prod_{i=d+1}^Q l_i}}{\frac{1}{Q-d} \sum_{i=d+1}^Q l_i} \right).$$

Cette méthode de détection est peu coûteuse et n'impose aucun réglage de seuil *a priori*. Sa mise en place sur la base de l'extension large bande « naïve » du §IV.3.3.2-B est impossible. La sélection du critère AIC ou MDL découle de la stratégie d'estimation utilisée. En théorie, (IV.48) détecte presque sûrement le nombre correct de sources lorsque $J \rightarrow +\infty$, alors (IV.47) tend à le surestimer [Xu 90].

IV.3.3.5 Implémentation pratique du schéma de {détection,localisation}

A Choix algorithmiques Notre implémentation en C/C++ du *coherent beamspace MUSIC* exploite $Q = 4$ formations de voie. Leur synthèse en champ lointain pour $B = 45$ bins au moyen de YALMIP [Löfberg 04] et SDPT3 [Toh 99] ne prend que quelques minutes sous MATLAB et ne pose aucun problème numérique. Une estimation fiable des covariances $\{\underline{\mathbf{C}}_n(k_b)\}_{b \in \{1, \dots, B\}}$ du bruit est nécessaire. Celle-ci est effectuée lorsqu'aucune source informative n'est active, au moyen du schéma d'estimation de $\{\underline{\mathbf{C}}_x(k_b)\}_{b \in \{1, \dots, B\}}$. Afin de satisfaire approximativement les hypothèses de stationnarité sous-jacentes à la méthode dans le cas de signaux voisins, lorsque la fréquence d'acquisition est de 15 kHz ce schéma requiert $J = 4$ FFTs sur des blocs contigus de taille $L = 256$.

B Résultats expérimentaux Nous proposons deux ensembles d'évaluations expérimentales de notre schéma de {détection,localisation}. Une évaluation qualitative est reportée Figure IV.6-haut en champ proche et en environnement silencieux. Un téléphone émet des sons d'animaux, entrecoupés de silence, depuis une table. On remarque que les silences sont correctement détectés. Des pseudo-spectres en azimuth et distance sont présentés. L'exploitation de la mire disposée sur la table et la calibration de celle-ci par rapport au capteur montreraient que les incertitudes de localisation sont tout à fait acceptables, n'excédant pas $\pm 5^\circ$ dans les pires cas.

Les Figures IV.6-bas et IV.7 concernent des évaluations quantitatives. Des pseudo-spectres calculés sous MATLAB à partir de données réelles – donc en tous points semblables à ceux calculés en temps réel par l'implémentation C/C++ – sont présentés, de même qu'un histogramme des azimuths conduisant à leurs valeurs maximales. On note que l'estimation de la distance n'est pas satisfaisante, sans doute à cause d'un alignement imparfait des sous-espaces signal et bruit sur les différents “bins” fréquentiels. Le biais et la dispersion de l'azimuth estimé peuvent être attribués à plusieurs phénomènes. En premier lieu, la calibration a été assez sommaire, et n'est donc absolument pas fiable ; nous prévoyons de ré-utiliser prochainement le système de capture de mouvement disponible au LAAS – assez lourd à mettre en œuvre pour

ce genre d'expérimentation – afin d'obtenir la vérité terrain. Il se peut également que la raison soit plus profonde, certains travaux théoriques mentionnant un biais de MUSIC lorsque les statistiques des signaux et du bruit sont estimées sur un horizon fini. Une conclusion définitive sur ce point demande une étude plus poussée.

IV.3.4 Vers la détection de “patterns” acoustiques

Il est opportun que les primitives acoustiques exploitées dans une stratégie de suivi (visio-)auditif de locuteur ne soient pas seulement géométriques ou énergétiques, mais qu'elles intègrent également une *détection d'activité vocale* (VAD = Voice Activity Detection) [Lehmann 07]. Considérant que les phonèmes voisés constituent un ensemble de « motifs » – ou, ci-après, “patterns” – audios localement stationnaires, nous avons développé les fondements d'une solution originale pour la détection d'un pattern stationnaire noyé dans du bruit stationnaire additif [CIACL-23-AR-8].

Il convient de préciser que ces développements sont en cours de finalisation, y compris concernant certains aspects fondamentaux.

Les fondements théoriques de l'approche sont formalisés dans le filtrage adapté stochastique (SMF = Stochastic Matched Filtering), proposé pour la première fois par [Cavassilas 91] pour des applications d'océanographie. Le pattern d'intérêt et le bruit additif sont supposés centrés, stationnaires, mutuellement indépendants et d'autocorrélations connues ou préalablement estimées. Pour une taille M de la fenêtre temporelle de détection, ils sont respectivement désignés par les vecteurs[†] $S \in \mathbb{R}^M$ et $N \in \mathbb{R}^M$.

Le calcul de la GEVD $\{\lambda_i, \Phi_i\}_{i \in \{1, \dots, M\}}$ des matrices d'autocorrélations[‡] de S et N permet la définition d'une séquence déterministe $\{\Psi_i\}_{i \in \{1, \dots, M\}}$ vérifiant les deux propriétés suivantes : $\{\Psi_i\}_{i \in \{1, \dots, M\}}$ et $\{\Phi_i\}_{i \in \{1, \dots, M\}}$ sont bi-orthogonales, *i.e.* $\forall i, j, \Psi_i^T \Phi_j = \delta_{(i,j)}$; $Z = S + N$ s'écrit $Z = \sum_{i=1}^M z_i \Psi_i$, où la séquence aléatoire $\{z_i\}_{i \in \{1, \dots, M\}}$ définie par $\forall i, z_i = Z^T \Phi_i$, est centrée et décorrélée. Supposons que $\lambda_1 \geq \dots \geq \lambda_Q > 1 \geq \lambda_{Q+1} \geq \dots \lambda_M$. On montre que la troncature $Z_\Phi \triangleq \sum_{i=1}^Q z_i \Psi_i$ est une forme réhaussée de Z , qui permet d'augmenter le rapport signal sur bruit (SNR = Signal to Noise Ratio) moyen.

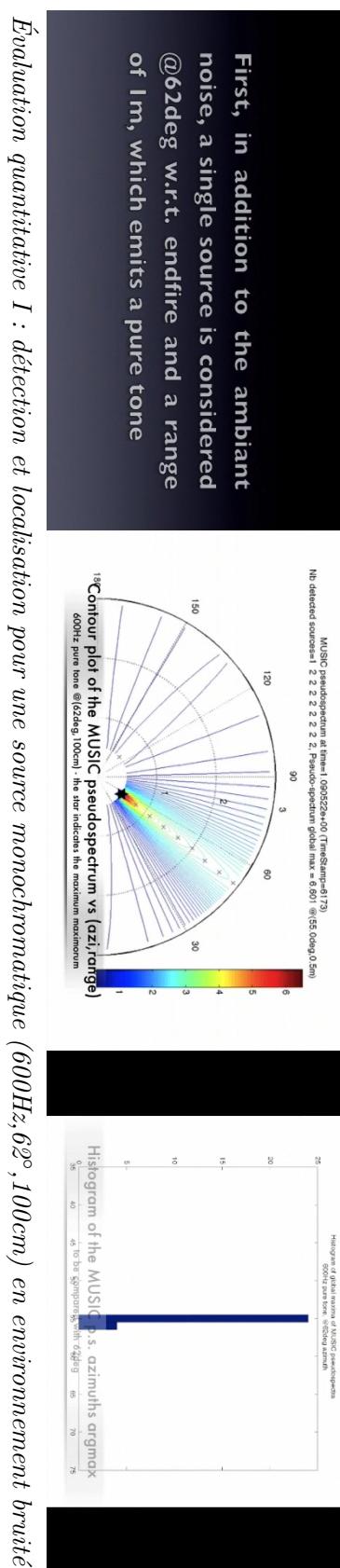
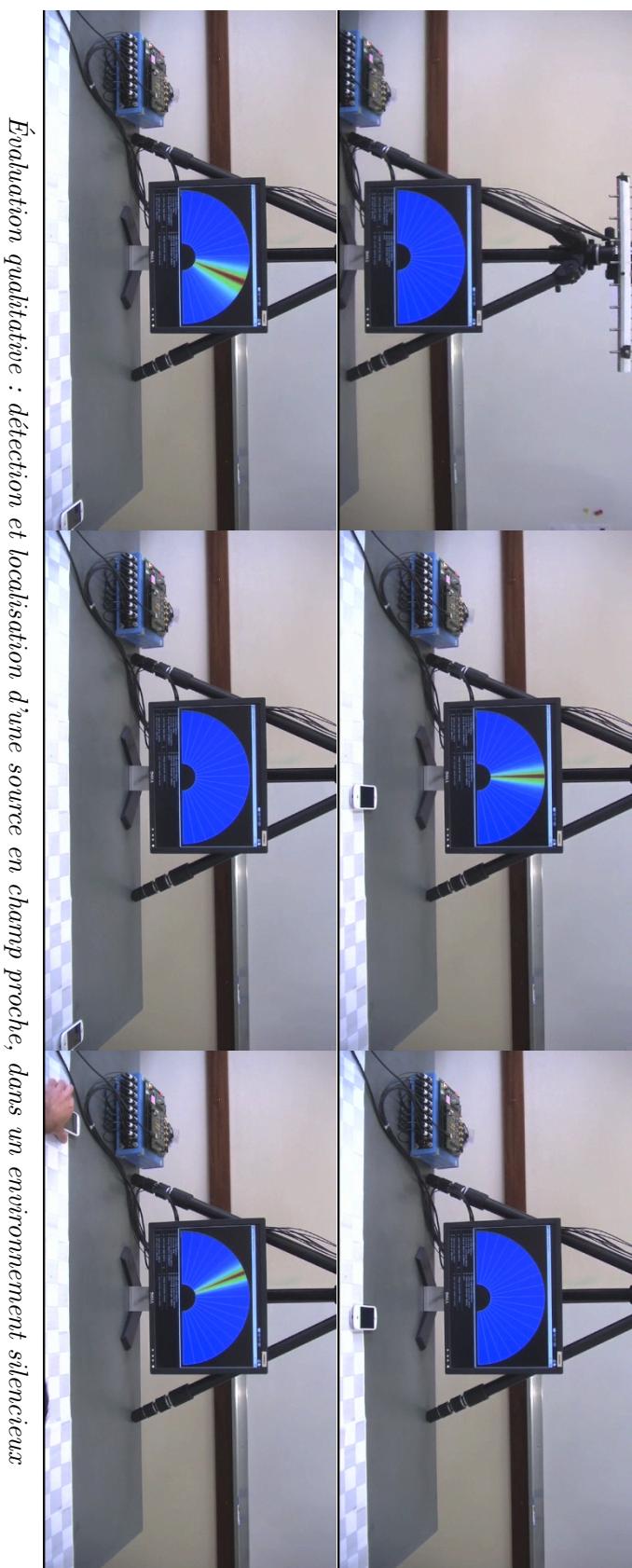
Le vecteur $z \triangleq (z_1, \dots, z_Q)^T$, où $\forall i, z_i = Z^T \Phi_i$, peut par conséquent être vu comme le résultat du filtrage de $Z = S + N$ par les Q filtres RIF dont les réponses impulsionales constituent les vecteurs Φ_1, \dots, Φ_Q . Similairement, $n \triangleq (n_1, \dots, n_Q)^T$ regroupe les sorties $n_i = N^T \Phi_i$ de ces Q filtres RIF lorsque Z est constitué du bruit N seul.

La connaissance *a priori* des statistiques sur S et N ainsi que le calcul de $\{\lambda_i\}_{i \in \{1, \dots, M\}}$ conduisent aux statistiques théoriques des vecteurs z et n . Si on suppose tous les signaux Gaussiens, un test de rapport de vraisemblances permet de détecter, à partir du filtrage d'un signal Z par les filtres $\{\Phi_i\}_{i \in \{1, \dots, M\}}$, l'hypothèse la plus vraisemblable parmi $[H_0 : Z \text{ est du bruit seul}]$ et $[H_1 : Z \text{ est le signal d'intérêt corrompu par du bruit}]$.

La complexité de cet algorithme est très réduite. En effet, la GEVD, le calcul de la « base du SMF » $\{\Phi_i\}_{i \in \{1, \dots, M\}}$, et la sélection du seuil du test du rapport de vraisemblances sont effectués hors ligne. La partie en ligne se limite à Q convolutions linéaires de deux séquences de \mathbb{R}^M et à la détection proprement dite.

[†]. Les notations sont propres à cette section et n'ont aucun rapport avec celles utilisées précédemment.

[‡]. Il s'agit de matrices de Toeplitz dont les diagonales sont constituées des valeurs de la fonction d'autocorrélation des signaux.



Évaluation quantitative I : détection et localisation pour une source monochromatique (600Hz, 62°, 10cm) en environnement bruité

FIGURE IV.6 – Évaluations du schéma de {détection,localisation} de sources sonores large bande.

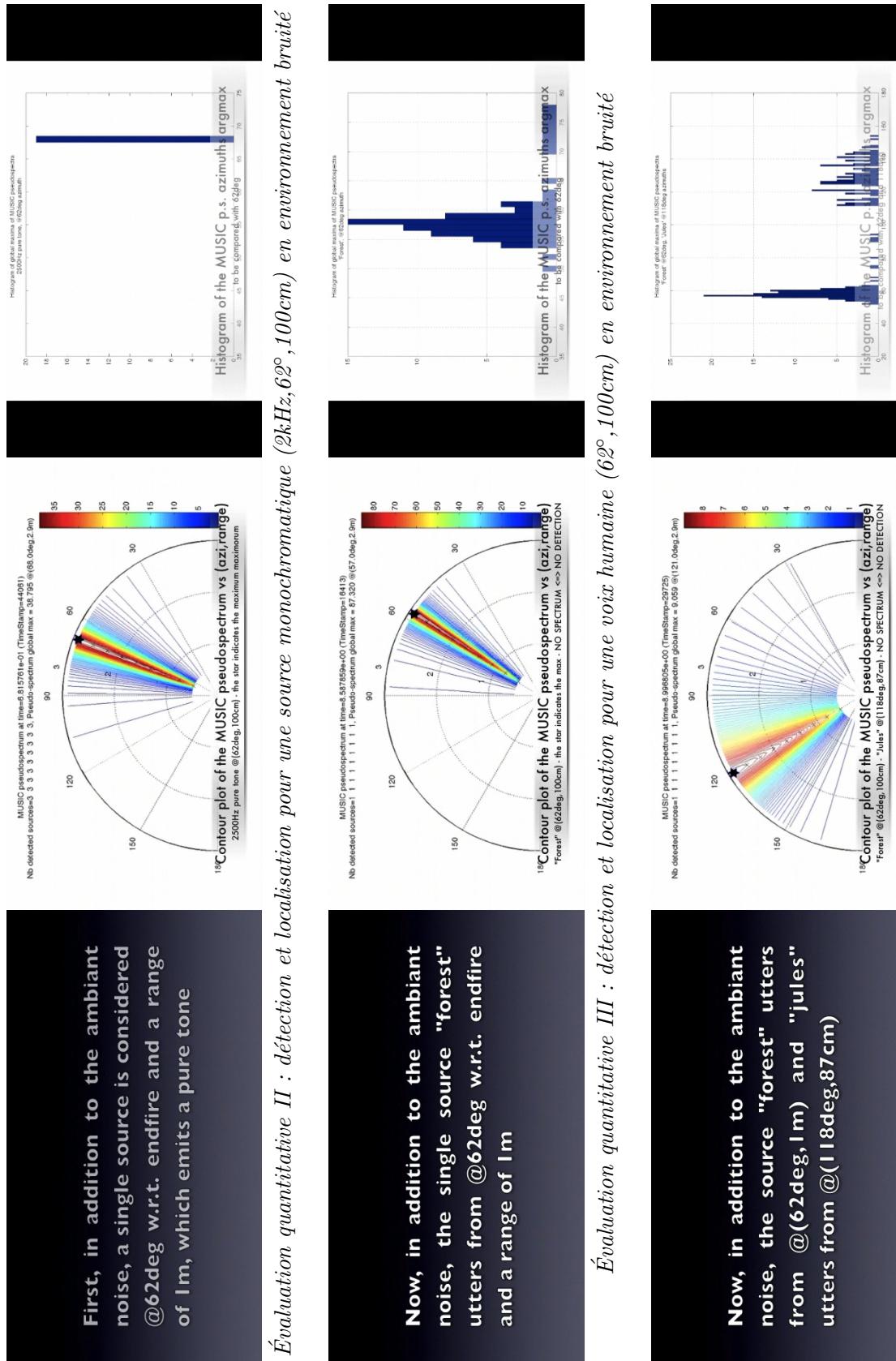


FIGURE IV.7 – Évaluations du schéma de {détection,localisation} de sources sonores large bande.

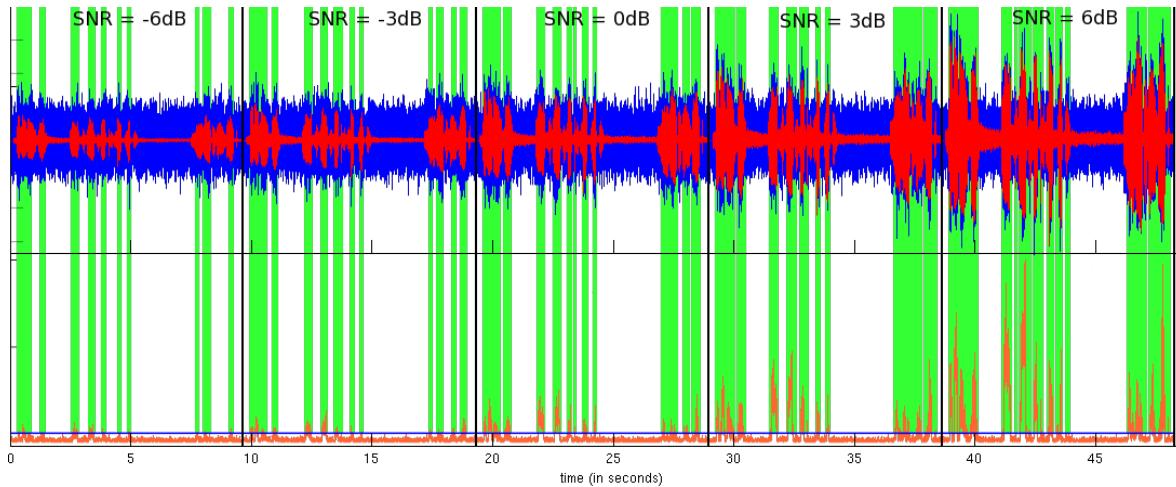


FIGURE IV.8 – Détection d'un pattern audio par le SMF. Les signaux Z et S sont respectivement tracés en bleu et rouge. Les zones vertes matérialisent la détection du pattern, lorsque $\Lambda(z)$ (orange) est supérieur au seuil ξ (bleu).

La Figure IV.8 présente nos premiers résultats expérimentaux relatifs à la détection d'un phonème voisé. Pour la fréquence d'échantillonnage $F_e = 15\text{ kHz}$, la dimension des signaux est fixée à $M = 100$. Un SMF comportant $Q = 37$ filtres RIF conduit à une détection satisfaisante pour un SNR supérieur à -6 dB . Toutefois, les performances chutent pour des SNR variables.

IV.3.5 Un capteur intégré embarquable basé sur un réseau de microphones pour l'audition en robotique

Un capteur audio pour la Robotique doit être embarquable et intégrer des traitements évolués avec un temps d'exécution garanti. Aucune solution commerciale ne nous ayant satisfait en terme de fiabilité et flexibilité pour l'acquisition multivoies synchrone et l'élaboration de primitives évoluées en temps réel, Sylvain ARGENTIERI [DOC-03] et Jérôme MANHÈS, ingénieur de recherche au LAAS-CNRS, ont conçu et réalisé le capteur auditif intégré EAR (“Embedded Audition for Robotics”, Figure IV.9-(a)). Julien BONNAL [DOC-06] a ensuite participé au développement des librairies associées.

Le cœur du capteur EAR est un kit de développement AVNET basé sur le FPGA Xilinx Virtex-4sx35, qui offre un parallélisme massif pour une consommation limitée. Sur cette carte sont enfichées deux cartes mezzanines développées au LAAS-CNRS pour l'acquisition synchrone de 8 voies. La communication avec un hôte UNIX – contrôle du capteur et récupération des données – s'effectue par USB1.1 au moyen d'un protocole dédié. La fréquence d'échantillonnage, les fréquences de coupure des filtres anti-repliement, et les gains des différents canaux sont entièrement programmables depuis l'hôte *via* le FPGA.

Le champ acoustique est échantillonné au moyen d'une antenne linéaire de $N = 8$ microphones $\frac{1}{4}''$ appariés en phase et régulièrement espacés de $l = \frac{\lambda_{3\text{kHz}}}{2} = 5.66\text{ cm}$ (cf. §IV.3.1.2). La polarisation des microphones et la mise en forme des signaux sont effectués par un conditionneur ICP® pouvant fonctionner sur batterie. EAR fournit également des signaux calibrés programmables pour le déclenchement précis de capteurs visuels, ainsi qu'une conversion numérique-analogique pour la récupération de toute donnée audio brute ou élaborée

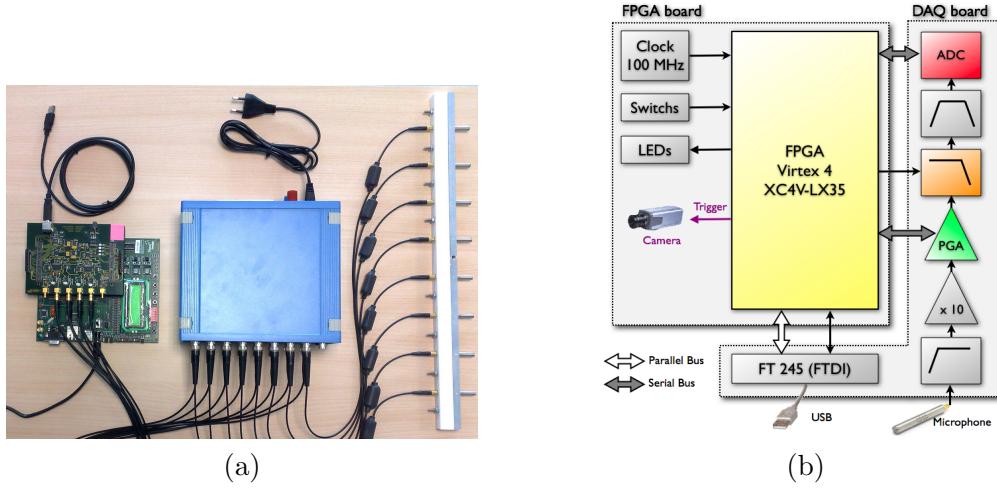


FIGURE IV.9 – Le capteur EAR. (a) Vue d'ensemble. De gauche à droite : unité de traitement FPGA et cartes d'acquisition mezzanines, conditionneur, antenne de microphones. (b) Schéma fonctionnel. L'acquisition (DAQ) de chaque voie est constituée d'un filtre passe-haut du 2^e ordre ($f_c = 60\text{ Hz}$), d'un amplificateur analogique dont le gain est réglable manuellement sur un continuum, d'un amplificateur de gain programmable (PGA), d'un filtre anti-repliement à capacités commutées (f_c programmable dans [2kHz; 50kHz]) suivi d'un étage passe-bande pour l'élimination du bruit d'horloge, puis d'un convertisseur sigma-delta 18 bits.

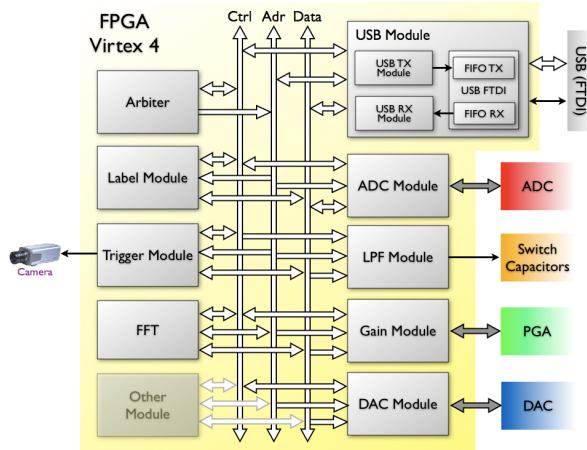


FIGURE IV.10 – Architecture modulaire VHDL implémentée sur le FPGA. Le module **Arbitre** gère le partage des trois bus et contrôle les autres modules. Certains modules sont dédiés à la paramétrisation de la chaîne d’acquisition (**Gain**, **LPF**, **ADC**), au traitement de l’information brute (*e.g.* **FFT**), ou à l’élaboration de primitives élaborées (détection et localisation de sources, filtrage spatial, etc.). Le module **USB** est en charge de la construction/transmission et réception/déconstruction des trames. Toute donnée peut être datée par le module **Label**. **Trigger** génère des signaux numériques calibrés. **DAQ** réalise la conversion numérique-analogique de données manipulées par le FPGA.

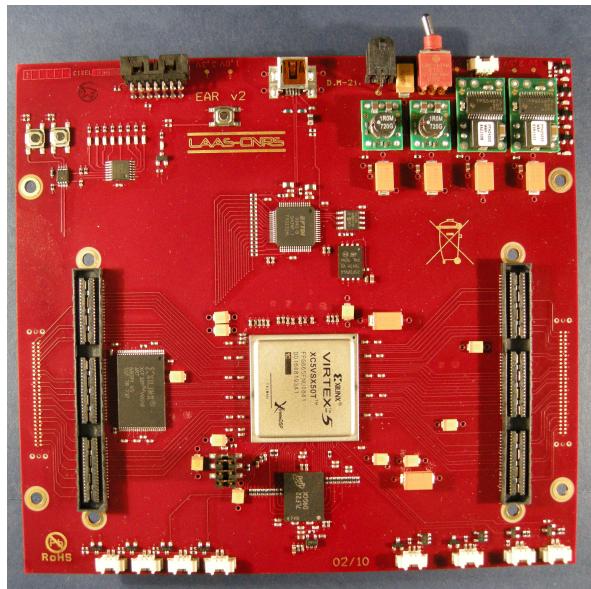


FIGURE IV.11 – Unité de traitement du capteur EAR2 (J. MANHÈS, LAAS).

(Figure IV.9-(b)).

L'une des deux librairies logicielles associées consiste en une architecture VHDL originale implémentée sur le FPGA. Les modules interconnectés qui la constituent effectuent des tâches d'acquisition/élaboration de primitives acoustiques ou bien de contrôle/communication, cf. Figure IV.10 pour davantage de détails. La conception modulaire permet l'adjonction de nouvelles fonctions sur la base de "templates" VHDL offrant les connections standard avec les trois bus. Enfin, le datage de données prévient toute erreur de transmission ou perte d'information.

D'autre part, sur l'hôte UNIX, la librairie C `libfpga` fournit des primitives bas-niveau pour le dialogue – configuration dynamique et échange de données – avec le capteur *via* la liaison USB. Sur cette base, un logiciel C++ multithreadé implémente une interface opérateur pour le réglage fin de l'acquisition, le codage de primitives auditives ainsi que leur exécution en temps réel, de même que la visualisation ou sauvegarde en temps réel de données brutes ou élaborées. Le développement de toute nouvelle routine s'effectue en trois étapes : prototypage MATLAB et tests hors-ligne sur des données simulées puis sur la base de signaux bruts acquis par le capteur EAR ; prototypage C++ et tests intensifs en ligne ; intégration en VHDL sur la cible FPGA, et validation [CIACL-21-AR-6].

Dans sa version actuelle, les données audio brutes sont échantillonnées à 15 kHz, des algorithmes de formation de voie codés en VHDL ou C++ s'exécutent à 15 kHz, et les primitives de détection et localisation codées en C++ s'exécutent à 15 Hz. Le FPGA Virtex-4sx35 ne contenant pas un nombre suffisant de cellules pour permettre leur implantation, un deuxième prototype EAR2 est en cours de réalisation par J. MANHÈS (Figure IV.11). Le kit AVNET est remplacé par une carte conçue et réalisée au LAAS, hébergeant un FPGA Xilinx Virtex-5sx50, des mémoires flash et RAM, et communiquant *via* une liaison USB2.0 high-speed (480Mbits/s). L'étage d'acquisition, de même que les librairies logicielles, demeurent inchangés.

IV.4 Références

- [Abhayapala 99] P.T.D. Abhayapala. *Modal Analysis and Synthesis of Broadband Nearfield Beamforming Arrays*. PhD thesis, Australian National University, 1999.
- [Akaike 74] H. Akaike. *A New Look at the Statistical Model Identification*. IEEE Trans. on Automatic Control, vol. 19, no. 6, pages 716–723, 1974.
- [Argentieri 06] S. Argentieri. *Conception d'un Capteur Sonore pour la Localisation de Source en Robotique Mobile*. PhD thesis, LAAS-CNRS et Université Paul Sabatier, 2006.
- [Argentieri 07] S. Argentieri, P. Danès & P. Souères. A tutorial on sound source localization in robotics : From biomimetic to array processing methods. Rapport LAAS N° 07696, 21pages, 2007.
- [Asano 99] F. Asano, H. Asoh & T. Matsui. *Sound Source Localization and Signal Separation for Office Robot Jijo-2*. In IEEE/SICE/RSJ Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems, pages 243–248, Taipei, Taiwan, 1999.
- [Balakrishnan 98] V. Balakrishnan & L. Vandenberghe. *Linear Matrix Inequalities for Signal Processing. An Overview*. In Annual Conf. on Information Sciences and Systems, 1998.
- [Borgerding 06] M. Borgerding. *Turning Overlap-Save into a Multiband Mixing, Downsampling Filter Bank*. IEEE Signal Processing Magazine, vol. 158, 2006.
- [Boyd 94] S. Boyd, L. El Ghaoui, E. Feron & V. Balakrishnan. Linear matrix inequalities in system and control theory. SIAM, 1994.
- [Bregman 90] A.S. Bregman. Auditory scene analysis : The perceptual organization of sound. MIT Press, Cambridge, MA, 1990.
- [Brière 08] S. Brière, J.M. Valin, F. Michaud & D. Létourneau. *Embedded Auditory System for Small Mobile Robots*. In IEEE Int. Conf. on Robotics and Automation (ICRA'08), pages 3463–3468, Pasadena, CA, 2008.
- [Brooks 99] R. Brooks, T.B. Senior & P.L.E. Uslenghi. *The Cog Project : Building a Humanoid Robot*. In C. Nehaniv, éditeur, Computations for Metaphors, Analogy, and Agents, volume LNCS 1562, pages 52–87. Springer-Verlag Heidelberg Berlin, 1999.
- [Brückmann 05] R. Brückmann, A. Scheidig, C. Martin & H.-M. Gross. *Integration of a Sound Source Detection into a Probabilistic-based Multimodal Approach for Person Detection and Tracking*. In Autonome Mobile Systeme, pages 131–137. Springer Berlin Heidelberg, 2005.
- [Cavassilas 91] J.F. Cavassilas. *Stochastic Matched Filter*. In Int. Conf. on Sonar Signal Processing, pages 194–199. Proceedings of the Institute of Acoustics, 1991.
- [Cherry 53] E.C. Cherry. *Some Experiments on the Recognition of Speech, with One and Two Ears*. Jour. of the Acoustical Society of America, vol. 25, pages 975–979, 1953.
- [De Cheveigné 02] A. De Cheveigné. La parole, des modèles cognitifs aux machines communicantes - analyse, synthèse et codage de la parole., chapitre L'analyse de Scènes Auditives Computationnelle, pages 175–196. Hermès, Paris, 2002.

- [Dufresne 09] D. Dufresne. *Localisation de Sources en Robotique : Étude de l'Antenne Semi-Circulaire*. Rapport technique, LAAS-CNRS, 2009. Stage de 2^e année Supélec, 138 pages.
- [Gonen 99] E. Gonen & J.M. Mendel. *Subspace-Based Direction Finding Methods*. In Madisetti, V.K. and Williams D.B., éditeur, *The Digital Signal Processing Handbook*, chapitre 62. Chapman & Hall, CRCnetBASE, 1999.
- [Haykin 05] S. Haykin & Z. Chen. *The Cocktail Party Problem*. Neural Computation, vol. 17, pages 1875–1902, 2005.
- [Hu 09a] J.-S. Hu, C.-Y. Chan, C.-K. Wang & C.-C. Wang. *Simultaneous Localization of Mobile Robot and Multiple Sound Sources Using Microphone Array*. In IEEE/RSJ Int. Conf. on Robotics and Automation, pages 29–34, Kobe, Japan, 2009.
- [Hu 09b] J.-S. Hu, C.-H. Yang & C.-K. Wang. *Estimation of Sound Source Number and Directions under a Multi-Source Environment*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pages 181–186, Saint-Louis, MO, 2009.
- [Ishi 09] C.T. Ishi, H. Ishiguro & N. Hagita. *Evaluation of a MUSIC-based Real-Time Sound Localization of Multiple Sound Sources in Real Noisy Environments*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pages 2027–2032, Saint-Louis, MO, 2009.
- [Kennedy 99] R.A. Kennedy, D.B. Ward & T.D. Abhayapala. *Nearfield Beamforming Using Radial Reciprocity*. IEEE Trans. on Signal Processing, vol. 47, no. 1, pages 33–40, 1999.
- [Knapp 76] C. Knapp & G. Carter. *The Generalized Correlation Method for Estimation of Time Delay*. IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 24, no. 4, pages 320–327, 1976.
- [Krim 96] H. Krim & M. Viberg. *Two Decades of Array Signal Processing Research : The Parametric Approach*. IEEE Signal Processing Magazine, vol. 13, no. 4, pages 67–94, 1996.
- [Kuratorri 05] S. Kuratori, N. Suzuki, K. Nakadai, H.G. Okuno & H. Amano. *Implementation of Active Direction-Pass Filter on Dynamically Reconfigurable Processor*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS’05), pages 515–520, 2005.
- [Lehmann 07] E.A. Lehmann & A.M. Johansson. *Particle Filter with Integrated Voice Activity Detection for Acoustic Source Tracking*. EURASIP Jour. on Advances in Signal Processing, vol. 2007, 2007.
- [Luo 03] Z.Q. Luo. Mathematical programming, volume 97, chapitre Applications of Convex Optimization in Signal Processing and Digital Communication. Springer-Verlag GmbH, 2003.
- [Löfberg 04] J. Löfberg. *YALMIP : A Toolbox for Modeling and Optimization in MATLAB*. In IEEE CACSD Conf. Taipei, Taiwan, 2004.
- [Mattos 04] L. Mattos & E. Grant. *Passive Sonar Applications : Target Tracking and Navigation of an Autonomous Robot*. In IEEE Int. Conf. on Robotics and Automation (ICRA’04), volume 5, pages 4265–4270, 2004.

- [Michaud 07] F. Michaud, C. Côté, D. Létourneau, Y. Brosseau, J.-M. Valin, É. Beaudry, C. Raïevsky, A. Ponchon, P. Moisan, P. Lepage, Y. Morin, F. Gagnon, P. Giguère, M.-A. Roux, S. Caron, P. Frenette & F. Kabanza. *Spartacus Attending the 2005 AAAI Conference*. Autonomous Robots, vol. 22, pages 369–383, 2007.
- [Nakadai 00] K. Nakadai, T. Matsui, H.G. Okuno & H. Kitano. *Active Audition System and Humanoid Exterior Design*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'00), pages 1453–1461, Takamatsu, Japon, 2000.
- [Nakadai 08] K. Nakadai, H.G. Okuno, H. Nakajima, Y. Hasegawa & H. Tsujino. *An Open Source Software System for Robot Audition HARK and Its Evaluation*. In IEEE-RAS Int. Conf. on Humanoid Robots, pages 561–566, Daejeon, Corée, 2008.
- [Nakamura 09] K. Nakamura, K. Nakadai, F. Asano, Y. Hasegawa & H. Tsuhino. *Intelligent Sound Source Localization for Dynamic Environments*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pages 664–669, Saint-Louis, MO, 2009.
- [Okuno 04] H.G. Okuno, T. Ogata, K. Komatani & K. Nakadai. *Computational Auditory Scene Analysis and its Application to Robot Audition*. In IEEE Int. Conf. on Informatics Research for Development of Knowledge Society Infrastructure (ICKS'04), pages 73–80, 2004.
- [Okuno 08] H.G. Okuno & K. Nakadai. *Computational auditory scene analysis and its application to robot audition*. In IEEE Wkshop on Hands-free Speech Communication and Microphone Arrays, pages 124–127, 2008.
- [Omologo 97] M. Omologo & P. Svaizer. *Use of the Crosspower-Spectrum Phase in Acoustic Event Location*. IEEE Trans. on Speech and Audio Processing, vol. 5, no. 3, pages 288–292, 1997.
- [Parra 02] L.C. Parra & C.V. Alvino. *Geometric Source Separation : Merging Convolutional Source Separation with Geometric Beamforming*. IEEE Trans. on Speech and Audio Processing, vol. 10, no. 6, pages 352–362, 2002.
- [Radich 94] B.M. Radich & K. Buckley. *The Effect of Source Number Underestimation on MUSIC Location Estimates*. IEEE Trans. on Signal Processing, vol. 42, no. 1, pages 233–235, 1994.
- [Rosenthal 97] D. Rosenthal & H.G. Okuno, éditeurs. Computational auditory scene analysis. Lawrence Erlbaum Associates, 1997.
- [Schmidt 86] R.O. Schmidt. *Multiple Emitter Location and Signal Parameter Estimation*. IEEE Trans. on Antennas Propagation, vol. 34, no. 3, pages 276–280, 1986.
- [Tamai 05] Y. Tamai, Y. Sazaki, S. Kagami & H. Mizoguchi. *Three Ring Microphone Array for 3D Sound Localization and Separation for Mobile Robot Audition*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'05), pages 903–908, 2005.
- [Toh 99] K.C. Toh, M.J. Todd & R.H. Tutuncu. *SDPT3 — A MATLAB Software Package for Semidefinite Programming*. Optimization Methods and Software, vol. 11, pages 545–581, 1999.

- [Valin 03] J.M. Valin, F. Michaud, J. Rouat & D. Létoumeau. *Robust Sound Source Localization using a Microphone Array on a Mobile Robot*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'03), volume 2, pages 1228–1233, 2003.
- [Valin 04a] J.M. Valin, F. Michaud, B. Hadjou & J. Rouat. *Localization of Simultaneous Moving Sound Sources for Mobile Robot Using a Frequency-Domain Steered Beamformer Approach*. In IEEE Int. Conf. on Robotics and Automation, pages 1033–1038, New Orleans, LA, 2004.
- [Valin 04b] J.M. Valin, J. Rouat & F. Michaud. *Enhanced Robot Audition Based on Microphone Array Source Separation with Post-Filter*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'04), pages 2123–2128, 2004.
- [Van Trees 02] Harry L. Van Trees. Optimum array processing, volume IV of *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, Inc., 2002.
- [Van Veen 88] B.D. Van Veen & K.M. Buckley. *Beamforming : A Versatile Approach to Spatial Filtering*. IEEE ASSP Magazine, vol. 5, no. 2, pages 4–24, 1988.
- [Wang 85] H. Wang & M. Kaveh. *Coherent Signal-Subspace Processing for the Detection and Estimation of Angles of Arrival of Multiple Wide-Band Sources*. IEEE Trans. on Acoustics, Speech and Signal Processing, vol. 33, no. 4, pages 823–831, 1985.
- [Wang 03] F. Wang, V. Balakrishnan, P.Y. Zhou, J.J. Chen, R. Yang & C. Franck. *Optimal Array Pattern Synthesis Using Semidefinite Programming*. IEEE Trans. on Signal Processing, vol. 51, pages 1172–1183, 2003.
- [Ward 95] D.B. Ward, R.A. Kennedy & R.C. Williamson. *Theory and Design of Broadband Sensor Arrays with Frequency Invariant Farfield Beampatterns*. Jour. of the Acoustical Society of America, vol. 97, pages 1024–1034, 1995.
- [Ward 04] D.B. Ward & T.D. Abhayapala. *Range and Bearing Estimation of Wideband Sources using an Orthogonal Beamspace Processing Structure*. In IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, pages 109–112, Montreal, Canada, 2004.
- [Wax 85] M. Wax & T. Kailath. *Detection of Signals by Information Theoretic Criteria*. IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 33, no. 2, pages 387–392, 1985.
- [Williams 99] D.B. Williams. *Detection : Determining the Number of Sources*. In Madisetti, V.K. and Williams D.B., éditeur, *The Digital Signal Processing Handbook*, chapitre 67. Chapman & Hall, CRCnetBASE, 1999.
- [Williams 07] M.I.Y. Williams. *Fundamental Limits in Space-Time Processing*. PhD thesis, Australian National University, 2007.
- [Xu 90] W. Xu & M. Kaveh. *Analysis of the Performance and Sensitivity of Eigendecomposition-based Detectors*. IEEE Trans. on Signal Processing, vol. 38, pages 1959–1971, 1990.

Chapitre V

Prospectives

Sommaire

V.1 Travaux en cours et prospectives associées à court terme	142
V.1.1 Détection et Estimation	142
V.1.1.1 Localisation ARGOS	142
V.1.1.2 Retour sur la détection de transitions entre modèles d'état hétérogènes	142
V.1.2 Asservissement visuel	143
V.1.2.1 Développements incrémentaux selon l'approche DAR	143
V.1.2.2 Prolongement de l'approche DAR et Caractérisation de son pessimisme	144
V.1.3 Audition en Robotique	144
V.1.3.1 EAR et HARK	144
V.1.3.2 Finalisation de EAR2 et implémentation VDHL de nouvelles fonctions	144
V.1.3.3 Nouvelle géométrie d'antenne pour la Robotique humanoïde	145
V.2 Axes structurants de notre recherche	145
V.2.1 Asservissement visuel	145
V.2.2 Audition en Robotique	146
V.2.2.1 Vers l'audition binaurale	146
V.2.2.2 Vers l'audition binaurale et active	147
V.2.2.3 Vers un suivi robuste de locuteur impliquant le capteur EAR	148
V.2.3 À plus long terme...	148
V.3 Quelques références significatives	150

Ce dernier chapitre recense quelques prolongements de nos travaux. Ceux qui seront réalisés dans le court terme sont d'abord détaillés. Nous discutons ensuite les axes de recherche qui structureront notre recherche à moyen ou long terme.

Nous rappelons que les encadrements et publications référencés ci-après sont codifiés et recensés dans les §I.2.3 (pages 15 et suivantes) et §I.2.4 (pages 19 et suivantes).

V.1 Travaux en cours et prospectives associées à court terme

Les prospectives présentées dans cette section s'inscrivent dans des thèses de Doctorat en cours ou dans la continuité de stages récents. Il s'agit de travaux planifiés, pouvant parfois présenter un caractère très incrémental. Leur horizon temporel est lié soit à ces thèses soit à des jalons (“milestones”) de projets de recherche engagés.

V.1.1 Détection et Estimation

V.1.1.1 Localisation ARGOS

Notre contribution à la localisation ARGOS dans le cadre de l'estimation d'état en contexte stochastique sera poursuivie. Notre filtre IMMHO [CIACL-24-DE-9] sera étendu de manière élémentaire au cas où les dynamiques relatives aux divers modes de fonctionnement sont non linéaires, *via* l'introduction de filtres de KALMAN unscented en lieu et place des filtres linéaires. Une évaluation sera ensuite menée sur des données réelles, afin de caractériser la pertinence de ce schéma de localisation multi-modèles.

Ensuite, deux problématiques complémentaires seront envisagées. D'une part, nous aborderons le problème de l'estimation conjointe des états et des paramètres de modèles à sauts, dont les dynamiques admettent des ordres hétérogènes. Dans un contexte de trajectographie d'animaux marins, ceci permettra par exemple de confronter un modèle de marche aléatoire expliquant un comportement erratique en zone de nourriture, avec un modèle de marche aléatoire corrélée relativement à un mouvement de migration au long cours, dont le coefficient de diffusion est *a priori* inconnu. Une étude bibliographique soignée sera effectuée, car une solution à ce problème ne se limite probablement pas à la simple application de notre schéma sur un système à sauts Markoviens augmenté dont certains états admettent une dynamique nulle. D'autre part, le lissage multi-modèle sera considéré. Le retraitement hors-ligne de relevés satellitaires effectués sur une plage temporelle donnée afin d'affiner les localisations est en effet une fonctionnalité très demandée par les clients de CLS.

V.1.1.2 Retour sur la détection de transitions entre modèles d'état hétérogènes

La solution proposée en collaboration avec M. Thierry SENTENAC pour la détection de transitions entre modèles d'état hétérogènes (§II.1.1) sera revisée. Nous avons prévu de la compléter, de poursuivre son évaluation, puis de la soumettre pour publication auprès de la communauté scientifique. Une participation ponctuelle de M. Rémy LOPEZ et des collègues co-encadrants de CLS n'est pas exclue. Au-delà d'une éventuelle comparaison de ce schéma de détection multi-modèles avec le filtre IMMHO sur des problèmes de trajectographie ARGOS, ceci permettrait un questionnement sur les différences fondamentales entre les deux approches à des fins de détection de ruptures. Le filtre IMMHO peut en effet être exploité pour la détection du mode déplacement le plus vraisemblable *a posteriori*, similairement à

l'exploitation des approches par simulation de Monte Carlo [Isard 98] et [Andrieu 03] relatée dans §II.1.2.1 pour le suivi visuel de gestes. Cependant, le fait qu'il repose sur la donnée *a priori* d'une chaîne de Markov exprimant les probabilités de sauts entre les divers modes suppose que la nature est « coopérative », au sens où elle affiche cette information sur son fonctionnement et s'y conforme. À l'inverse, selon l'approche développée avec M. SENTENAC, la nature est « neutre » au sens où elle n'affiche aucune information relative à l'enchaînement des modes. La portée de ces deux philosophies doit être approfondie, de même que les conséquences sur la robustesse des solutions obtenues vis à vis d'incertitudes de modèles. Ceci impliquera certainement de revenir aux fondements de la théorie de la détection [Nikiforov 09] et de ses variations dans le contexte du filtrage adaptatif [Gustafsson 01].

Dans le cadre de notre participation récente aux co-encadrements [DOC-0b][STAGE-11] dirigés par M^{lle} Viviane CADENAT, spécialiste de l'asservissement visuel dans le groupe, et par M. SENTENAC, nous avons posé les bases de l'exploitation de schémas de trajectographie multi-modèles pour l'asservissement visuel en environnement fortement dynamique. Nous poursuivrons cet objectif d'asservir un robot mobile sur une cible manœuvrante sur la base de la vision, en intégrant dans l'approche de [Folio 08] l'estimation de tous les éléments mobiles – cibles et obstacles – de la scène.

V.1.2 Asservissement visuel

V.1.2.1 Développements incrémentaux selon l'approche DAR

Comme indiqué au §III.3.3.4, notre approche de l'analyse multicritère par représentations algébro-différentielles (DAR) pourrait être étendue de façon à exploiter des fonctions de LYAPUNOV polyquadratiques, rationnelles, voire polyquadratiques/rationnelles par morceaux. Ce travail incrémental est certainement nécessaire afin de conclure définitivement sur ces méthodes. Au vu de la taille élevée et du conditionnement numérique parfois médiocre des programmes LMI obtenus pour des cas d'étude élémentaires au moyen de fonctions de LYAPUNOV relativement simples – *e.g.* bi-quadratiques par morceaux – il n'est hélas pas exclu que l'intérêt de telles extensions soit très limité.

Dans le projet STIC-AMSUD NCTVS [PROJ-R-2] (*Nonlinear Control Tools for Visual Servoing*), nous étudions avec nos collègues brésiliens les possibilités offertes par l'approche DAR pour la synthèse de stratégies de commande référencée vision multicritère originales et efficaces. La difficulté réside naturellement dans le fait les inégalités matricielles obtenues admettent une dépendance bilinéaire en les matrices de LYAPUNOV et les matrices définissant le correcteur recherché. Il s'agit alors d'identifier les classes de contrôleurs pour lesquelles le problème original de synthèse multicritère peut être ramené à un programme d'optimisation LMI, et d'isoler les éventuelles sources supplémentaires de pessimisme impliquées dans l'obtention d'une telle solution. La recherche d'une solution “tractable” peut être éventuellement limitée à la classe des modèles d'état en boucle ouverte intéressant l'asservissement visuel en contexte cinématique, où le vecteur d'état x vérifie $\dot{x} = J_{3D}(x)u$. En dernier recours, l'expression de $J_{3D}(\cdot)$ peut également être mise à profit, quoiqu'au détriment de la généricité.

Comme cela a déjà été vu au §III.3.1.1-E, la littérature en asservissement visuel propose souvent des contrôleurs de type « Jacobienne inverse », statiques ou séquencés. Le fait de se limiter à ces classes de commandes implique que l'une des clés d'un comportement satisfaisant en boucle fermée est une redéfinition adéquate des indices visuels. Nous pensons qu'il est pertinent d'envisager d'autres stratégies de commande, y compris sur la base d'indices visuels

élémentaires, et d'examiner si elles peuvent constituer une option viable. Ainsi, la proposition d'une méthode constructive de synthèse multicritère d'un retour de sortie dynamique, fût-il d'ordre plein – *i.e.* d'ordre 6 – rencontrerait certainement un écho positif dans la communauté. En effet, à notre connaissance, aucune solution ne permet actuellement d'améliorer des propriétés de la boucle fermée *via* une intégration temporelle de l'information visuelle.

D'autre part, toujours dans le cadre du projet NCTVS mais en partenariat avec nos collègues chiliens, nous prévoyons de comparer nos approches de la localisation visuelle selon les stratégies LMI [CIACL-17-AV-5] et [CIACL-06-AV-3] avec d'autres méthodes de la littérature, et de les compléter si besoin.

V.1.2.2 Prolongement de l'approche DAR et Caractérisation de son pessimisme

En parallèle, la fin de la thèse de Sylvain DUROLA [DOC-05] revisite les fondements de l'obtention de conditions LMI pour l'analyse de systèmes rationnels selon l'approche DAR. Il s'agit d'une part de proposer une alternative moins conservative au Lemme III.2 (page 93) sur la base d'annihilateurs quadratiques constants ou affines. Des premiers résultats sont en cours d'obtention, dont les avantages sont clairement identifiés, et qui présentent accessoirement des connexions intéressantes avec [Chesi 04].

Un second volet concerne l'identification de connexions entre l'approche DAR et des techniques récentes d'optimisation sur les polynômes. Les bases théoriques sont le lien entre l'optimisation LMI et les polynômes sommes de carrés (SOS = Sum Of Squares), ainsi que la caractérisation de la positivité d'un polynôme sur un ensemble semi-algébrique de \mathbb{R}^n [Chesi 10]. Pour une même formulation d'un problème d'analyse, éventuellement multicritère, dans le formalisme DAR, on recherche des conditions nécessaires et suffisantes – non convexes – d'obtention d'un bassin d'attraction par application du Positivstellensatz [Stengle 74]. Sur la base de cette étape, l'objectif est d'isoler précisément les sources de pessimisme des conditions suffisantes LMI issues de l'application de la S-procédure et du Lemme III.2 (page 93), voire d'en proposer des extensions. L'obtention d'un compromis acceptable entre le pessimisme des programmes obtenus et leur « tractabilité » – convexité, taille et conditionnement numérique – est essentielle pour convaincre la communauté Robotique du bien-fondé de nos travaux.

V.1.3 Audition en Robotique

V.1.3.1 EAR et HARK

Le portage de la librairie Open Source HARK [Nakadai 08] est prévu incessamment [STAGE-15]. Ceci nous permettra d'une part de contribuer à cette librairie par l'intégration de nos fonctions de détection et localisation de sources (cf. §IV.3.2.3, §IV.3.3.3, §IV.3.3.4), et, d'autre part, d'envisager à moyen terme l'interfaçage de EAR et HARK.

V.1.3.2 Finalisation de EAR2 et implémentation VDHL de nouvelles fonctions

Un volet de nos travaux actuels concerne la finalisation de la deuxième évolution EAR2 de notre capteur auditif intégré. Les aspects liés au matériel et à l'architecture VHDL sont pris en charge par M. Jérôme MANHÈS, ingénieur de recherche au LAAS-CNRS. Une partie du stage [M2R-10] a concerné le prototypage de l'algorithme MUSIC faible bande sur FPGA dans le cas d'un bruit spatialement blanc. Suivant en cela [Minseok 03], une transformation

unitaire permet de ramener la décomposition en éléments propres de la matrice de covariance centro-Hermitienne des signaux perçus par les microphones à la décomposition d'une matrice symétrique réelle. Celle-ci est effectuée par la méthode de JACOBI, qui consiste en une série de séquences de rotations planaires, ou « balayages ». Ces rotations s'implémentent simplement sur FPGA au moyen d'algorithmes CORDIC (COordinate Rotation Digital Computer) [Andraka 98]. Toutes les fonctions élémentaires de cet algorithme *MUSIC Unitaire* ont été prototypées sous MATLAB, et testées en précision finie. Ce travail sera repris durant l'année universitaire en cours, et étendu au cas large bande. Du code synthétisable VHDL sera ensuite prototypé, probablement en générant automatiquement le code relatif à certains sous-ensembles depuis MATLAB-SIMULINK au moyen de Synplify DSP, comme cela a été le cas lors du stage [STAGE-13]. L'objectif est de procéder à son implémentation sur le capteur EAR2 dès que celui-ci sera opérationnel, le nombre de cellules du FPGA Xilinx Virtex-4sx35 de EAR n'étant pas suffisamment élevé y compris pour MUSIC Unitaire faible bande.

V.1.3.3 Nouvelle géométrie d'antenne pour la Robotique humanoïde

Un aspect du projet AUDIO_HRP2 [PROJ-R-1] (*Capteur Audio Biomimétique et Fonctions Auditives pour le Robot Humanoïde HRP2*) concerne l'extension de nos primitives de localisation et de filtrage spatial à une nouvelle géométrie d'antenne, ainsi que leur implémentation sur EAR ou EAR2 et leur évaluation. Un réseau de microphones arc-circulaire sera conçu puis fixé en « bandana » sur une reproduction de la tête du robot humanoïde HRP-2 hébergé par le LAAS-CNRS. Les fondements théoriques de l'extension à une antenne circulaire ou arc-circulaire de notre formation de voie modale optimisée ont été posés lors de [STAGE-12], en vue d'obtenir un diagramme de directivité invariant en fréquences et polarisé selon un couple (azimuth, élévation) donné [Dufresne 09]. Ce travail sera également poursuivi durant l'année universitaire en cours.

V.2 Axes structurants de notre recherche

Cette section se propose de dresser la liste des axes qui structureront notre recherche à moyen terme, en nous projettant au-delà des objectifs cités précédemment. Certains de ces axes sont clairement identifiés car accompagnés par des projets. D'autres, plus prospectifs et moins formalisés, feront l'objet d'un travail bibliographique poussé, puis, si cela s'avère pertinent, de nouveaux encadrements.

V.2.1 Asservissement visuel

Nous poursuivrons notre activité dans la thématique de l'asservissement visuel multicritère dans le contexte des systèmes rationnels incertains soumis à des contraintes rationnelles. En vue d'obtenir une solution constructive, nous poursuivrons notre inflexion vers l'optimisation SOS et nous familiariserons avec son problème dual, dit des moments [Parrilo 03][Lasserre 01]. Il ne s'agira bien sûr pas de développer des recherches fondamentales dans ces domaines, mais d'en comprendre l'essence et d'en être des utilisateurs avertis. Nous poursuivrons une veille sur des techniques d'analyse ou de commande proches de nos préoccupations, *e.g.* [Papachristodoulou 05]. Notons qu'à l'heure actuelle, des travaux tels que [Tan 06] pour la détermination de bassins d'attraction de systèmes non linéaires à dynamiques polynomiales par des fonctions de LYAPUNOV polynomiales ou “max-polynomiales” ou “min-polynomiales”

conduisent à des programmes SOS bilinéaires, donc sujets à des optima locaux. Bien que le support théorique de prédilection demeure la théorie de LYAPUNOV, il sera probablement opportun d'étudier plus en profondeur toute une littérature ayant trait à l'analyse des domaines de stabilité des systèmes non linéaires autonomes par des méthodes alternatives [Genesio 85].

D'autres problèmes devront être traités, qui intéressent plus spécifiquement la Robotique. Il s'agit par exemple d'analyser ou de synthétiser des commandes référencées vision autorisant les saturations d'actionneurs, dont le bassin d'attraction maximal est plus étendu dans notre contexte cinématique.

Bien que nous remettions en cause les schémas de commande conventionnels de type « Jacobienne inverse » et que nous supposons implicitement qu'une loi de commande convenablement sélectionnée doit suppléer au caractère rudimentaire de certaines primitives visuelles – suivant en cela la position de [Papanikolopoulos 93] –, nous avons la volonté de « raccrocher » davantage notre approche à la littérature classique de l'asservissement visuel. Nous ne souhaitons pas remettre en cause *a priori* le principe de notre modélisation – *i.e.* une équation d'état qui unit le torseur cinématique de la caméra à la situation relative caméra-cible, et une équation de mesure qui rend compte du lien *statique* entre cette situation et les informations visuelles qui alimentent le contrôleur – car elle permet une position très propre du problème de stabilisation dans le respect des contraintes. Cependant, la littérature classique, basée sur des commandes linéarisantes pour un modèle en boucle ouverte exprimant le lien *différentiel* entre torseur cinématique et indices visuels, fait état de réflexions fondamentales que nous nous devons d'adapter à notre contexte. Parmi ces réflexions, citons : (1) la limitation de la connaissance *a priori*, *e.g.* est-il possible, dans le modèle de la boucle ouverte, de ne disposer que des primitives 2D de projection de la cible en la situation de référence ($\{\bar{x}_i^*, \bar{y}_i^*\}$ sur la Figure III.1 page 72) sans nécessiter son modèle complet ($\{a_i, b_i, c_i\}$ sur la Figure III.1) ? ; (2) l'importance du choix des indices visuels, *e.g.* un nouvel ensemble de primitives géométriques dont le lien à la situation caméra-cible s'exprime analytiquement, ou des combinaisons de primitives élémentaires réduisant le « degré de non-linéarité » de l'équation de mesure de notre modèle en boucle ouverte (et par conséquent le pessimisme de nos conclusions) ; et (3) la sélection *a priori* de couples (primitives visuelles, stratégies de commande) qui confient à la boucle fermée une robustesse structurelle par rapport à certaines incertitudes – *e.g.* sur les paramètres intrinsèques de la caméra – auquel cas le problème de synthèse se ramène à la synthèse des paramètres libres du contrôleur.

V.2.2 Audition en Robotique

V.2.2.1 Vers l'audition binaurale

Le projet AUDIO_HRP2 [PROJ-R-1] (*Capteur Audio Biomimétique et Fonctions Auditives pour le Robot Humanoïde HRP2*) nous donne également l'occasion de débuter une recherche vers la thématique beaucoup plus ambitieuse et risquée de l'*audition binaurale*. Bien que les approches binaurales conduisent encore des résultats mitigés, cette inflexion nous semble nécessaire. Aux arguments technologiques qui plaident en sa faveur – simplification de la géométrie et du système d'acquisition des signaux – s'ajoutent des notions plus subjectives liées à une meilleure acceptabilité par les humains de robots humanoïdes auditifs capables d'interaction lorsque ceux-ci respectent un certain biomimétisme. Une difficulté majeure concerne l'élaboration de modèles de propagation des ondes prenant en compte les phénomènes de diffusion engendrés par les épaules et la tête, pourtant nécessaires à la synthèse de tout algorithme

de localisation ou de séparation de sources. Dans le meilleur des cas, la connaissance *a priori* se limite généralement à un ensemble de relevés expérimentaux de la fonction de transfert de la tête (HRTF = Head Related Transfer Function), *i.e.* du transfert entre le point milieu de la tête en environnement libre – lorsque la tête est absente – et le point d’entrée de chaque conduit auditif. Ces mesures sont effectuées en chambre anéchoïque, pour une source monochromatique caractérisée par diverses valeurs du quadruplet (distance,azimuth,élévation,fréquence). L’obtention d’un modèle rendant compte du filtrage en élévation effectué par les pavillons, et qui puisse être exploité à des fins de conception, est tout aussi ardue.

En collaboration avec Alain SKAF, post-doctorant [POSTDOC-1] dans notre équipe, nous tentons de définir une méthodologie systématique permettant de doter une tête robotique nue quelconque d’un capteur binaural « satisfaisant ». Nous supposons que des relevés de la HRTF de cette tête peuvent être obtenus soit par l’expérimentation, soit par simulation à éléments finis de son modèle CAO. De nombreux travaux ont concerné la définition d’approximations continues individualisées d’une HRTF, *i.e.* spécifiques à chaque morphologie. Parmi ceux-ci, la référence récente [Zhang 10] démontre trois propriétés essentielles : l’expression d’une HRTF sur un continuum de distances-azimuths-élévations-fréquences peut être assimilée à la troncature de sa décomposition sur la base des harmoniques sphériques ; l’ordre de troncature s’exprime analytiquement en fonction de la bande de fréquences considérée ; le nombre de mesures ponctuelles de la HRTF suffisant à une approximation continue satisfaisante des coefficients modaux, et par là même de la HRTF globale, se déduit de cet ordre. Nous basons nos recherches sur ces résultats, car ils permettent d’envisager un positionnement des oreilles guidé par la minimisation d’un critère de performance, tel que la sensibilité moyenne, sur un cône spatial donné, des différences interaurales en amplitude ou en phase par rapport à l’azimuth de la source.

Lorsque cette méthodologie de positionnement des deux capteurs constituant le système binaural sera opérationnelle, nous nous focaliserons sur leur définition. Chez les humains, l’oreille est constitué de trois parties essentielles : l’oreille externe –ou « pavillon »– collecte l’information de pression acoustique et joue un rôle important dans la localisation verticale ; l’oreille moyenne opère une adaptation d’impédance ; enfin, l’oreille interne réalise la conversion mécano-électrique proprement dite, et par conséquent le codage de l’information brute. Il s’agira de reproduire ces fonctions, ce qui implique en particulier la conception et la réalisation de pavillons. Actuellement, très peu de travaux ont traité de ce problème difficile [Kumon 05].

V.2.2.2 Vers l’audition binaurale et active

La plupart des travaux en analyse de scènes auditives ont jusqu’ici considéré une vue statique du monde. Cette hypothèse est simplificatrice, du fait que parole et audition ont lieu dans des environnements où émetteurs et récepteurs sont mobiles [Cooke 07]. Du fait qu’elle couvre la conception et l’implémentation de systèmes intégrant des capacités de mobilité et de locomotion, la Robotique constitue un creuset particulièrement indiqué pour l’étude des connexions entre mobilité et perception auditive. C’est pourquoi nous nous sommes également engagés vers cette voie de recherche, appelée « audition active ».

Le projet ANR-JST BINAAHR [PROJ-R-3] (*Binaural Active Audition for Humanoid Robots*) aborde le caractère actif de l’audition binaurale. Un volet s’appuie sur la théorie récente en psychologie de la perception [O’Regan 01] selon laquelle la perception est un processus actif car exploratoire et intimement lié –entrelacé– avec l’action. Certains collègues du projet développent des algorithmes postulant des processus communs à la perception et l’action,

inspirés de données physiologiques et comportementales observées chez les humains et les animaux. La Robotique constitue une plateforme d'expérimentation de ces théories, à plusieurs niveaux : validation/réfutation de certains aspects ; comparaison avec des approches davantage ancrées dans les sciences de l'ingénieur ; évaluation de leurs potentialités en terme de contraintes temporelles, en conditions d'acquisition dégradées, en terme de robustesse au bruit et aux réverbérations, etc. Un second volet du projet concerne la définition et l'implémentation de fonctions auditives permettant une interaction – elle aussi active – avec un ou plusieurs humains sans équipement particulier et dans des environnements non contrôlés. Au niveau bas, cet ensemble recouvre des fonctions de localisation et de séparation de sources exploitant l'information de mouvement, mais aussi des fonctions de détection binaurale d'activité vocale, de reconnaissance de locuteur, etc. Chacun des volets sera implémenté sur un système auditif binaural doté de mobilité – tête KEMAR préalablement robotisée, robots humanoïdes HRP2 – et testé sur des scénarios de complexités variables, le but étant de disposer d'un système binaural auditif complet et robuste.

Notre contribution principale, en phase avec l'orientation globale de notre recherche, concernera la conception et l'implémentation d'un capteur binaural telle que présentée ci-dessus, et, d'autre part, la localisation binaurale active de sources mobiles multiples. Un exemple bien connu de l'apport du mouvement pour la localisation binaurale est la disambiguation avant-arrière [Kim 09]. La fusion spatio-temporelle de la dynamique *a priori* du capteur binaural et de l'information apportée par les indices auditifs binauraux mesurés fait naturellement appel à la théorie du filtrage. Pour des raisons voisines de celles justifiant son exploitation en suivi visuel, le filtrage particulaire a été identifié comme un outil particulièrement indiqué. Quelques réalisations existent [Lu 10], mais le problème est loin d'être clos. Nous réaliserons ainsi une première « jonction » avec nos activités passées en Détection et Estimation.

V.2.2.3 Vers un suivi robuste de locuteur impliquant le capteur EAR

Dès que les tous derniers développements présentés aux §IV et §V.1.3 auront atteint une maturité et une stabilité suffisantes, nous pourrons envisager le suivi visio-auditif de locuteur dans une scène dynamique, ainsi que l'extraction de son message sans qu'il soit nécessaire de l'équiper avec un appareillage spécialisé. Ce travail s'effectuera en collaboration avec des collègues de mon groupe de recherche, et constituera ainsi une deuxième « jonction » avec les compétences acquises en filtrage particulaire dans le cadre des travaux présentés au §II.1.2. La Figure V.1 illustre la manière dont pourraient interagir un tel schéma de suivi avec les fonctions de détection, localisation et extraction de sources intégrées dans le capteur EAR/EAR2, une détection d'activité vocale s'inscrivant dans le prolongement de nos travaux en détection de “patterns” acoustiques, et un module de reconnaissance de locuteur disponible « sur étagère », cette dernière fonction n'entrant pas dans nos compétences.

V.2.3 À plus long terme...

Très brièvement, nous concluons ce manuscrit par deux prospectives à plus long terme.

L'une concerne l'exploitation de la perception auditive et son couplage avec d'autres modalités sensorielles dans des lieux intelligents. Nous proposerons le déploiement d'une grande antenne de microphones – typiquement 32 transducteurs – dans une salle d'expérimentation du bâtiment intelligent ADREAM en cours de construction au LAAS-CNRS. Les traitements as-

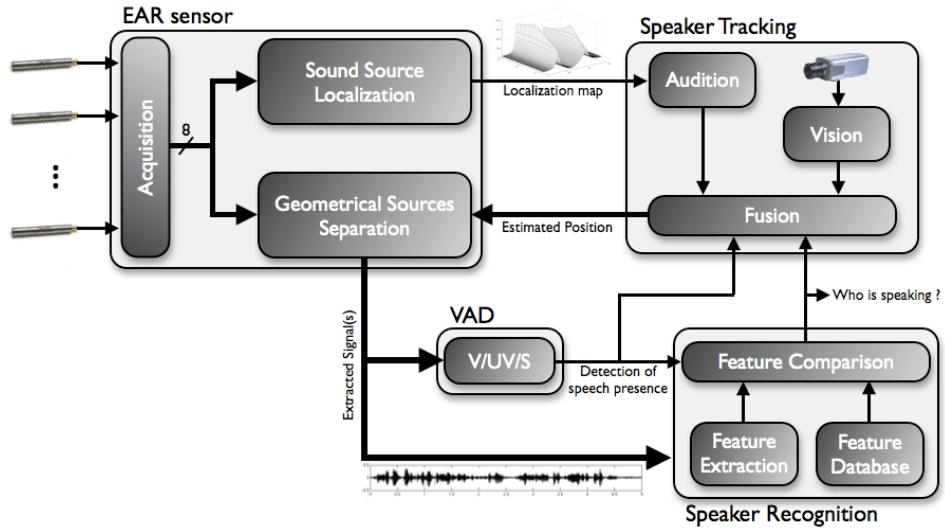


FIGURE V.1 – Proposition d'un schéma de suivi visio-auditif.

sociés devront satisfaire l'ensemble des contraintes imposées par la Robotique et mentionnées au §IV.2.1, excepté celle liée à l'embarquabilité. Dans cet environnement évolueront des robots mobiles, que nous équiperons d'audition binaurale. Les fonctions auditives développées dans ce contexte devront naturellement fusionner les primitives élaborées délivrées par le capteur d'ambiance avec celles élaborées par les robots. Notons que ce problème peut être contraint par le contexte : relative incohérence des informations à fusionner, bande passante maximale lors de l'échange d'informations entre les différents acteurs, etc. La mobilité ouvrira naturellement de multiples perspectives, telles la fusion de l'information de mouvement avec la perception lors de l'élaboration de primitives audio, mais également la reconfiguration dynamique du réseau constitué de l'antenne et des robots pour améliorer la qualité de ces primitives, l'éventuelle exploitation de connaissances *a priori* sur la scène, etc.[Martinson 07]. Comme la salle d'expérimentation sera également munie de caméras d'ambiance, la possible combinaison de primitives auditives et visuelles ouvrira également des perspectives en détection, suivi, navigation, etc.

Une dernière prospective qui nous tient particulièrement à cœur quoiqu'enore non formalisée concernera une formalisation « automaticienne » de la perception active. Le problème académique du positionnement d'un robot par rapport à une cible immobile sur la base de sa perception extéroceptive en environnement libre sera d'abord abordé. On supposera qu'à un instant donné, la perception du robot ne le renseigne que partiellement sur sa situation relative à la cible qu'il doit stabiliser : nombre d'indices visuels dans le plan image de la caméra ne permettant pas d'établir une correspondance bijective instantanée entre la perception et la situation effective caméra-cible ; information auditive intermittente, et réduite à l'azimuth de la source ; etc. Le problème de commande sera à nouveau posé en terme de la recherche d'un contrôleur dynamique stabilisant sous les contraintes de visibilité, saturations, etc. Cependant, nous souhaitons guider la synthèse d'un contrôleur par l'optimisation d'un critère relatif à l'*information* totale sur le vecteur d'état qu'apporte la mesure sur la durée de la tâche, la formulation exacte de ce critère devant être définie.

V.3 Quelques références significatives

- [Andraka 98] R. Andraka. *A survey of CORDIC algorithms for FPGAs.* In ACM/SIGDA 6th Int. Symp. on Field Programmable Gate Arrays (FPGA'1998), pages 191–200, 1998.
- [Andrieu 03] C. Andrieu, M. Davy & A. Doucet. *Efficient Particle Filtering for Jump Markov Systems. Application to Time-Varying Autoregressions.* IEEE Trans. on Signal Processing, vol. 51, pages 1762–1770, 2003.
- [Chesi 04] G. Chesi, A. Garulli, A. Tesi & A. Vicino. *Robust Analysis of LFR Systems Through Homogeneous Polynomial Lyapunov Functions.* IEEE Trans. on Automatic Control, vol. 49, no. 7, pages 1211–1216, 2004.
- [Chesi 10] G. Chesi. *LMI Techniques for Optimization over Polynomials in Control : a Survey.* IEEE Transactions on Automatic Control, to appear, 2010.
- [Cooke 07] M. Cooke, Y.C. Lu, Y. Lu & R. Horaud. *Active Hearing, Active Speaking.* In Int. Symp. on Auditory and Audiological Research, 2007.
- [Dufresne 09] D. Dufresne. *Localisation de Sources en Robotique : Étude de l'Antenne Semi-Circulaire.* Rapport technique, LAAS-CNRS, 2009. Stage de 2^e année Supélec, 138 pages.
- [Folio 08] D. Folio & V. Cadenat. *Dealing with Visual Features Loss during a Vision-Based Task for a Mobile Robot.* International Journal of Optomechatronics, vol. 2, no. 3, pages 185–204, 2008.
- [Genesio 85] R. Genesio, M. Tartaglia & A. Vicino. *On the Estimation of Asymptotic Stability Regions : State of the Art and New Proposals.* IEEE Trans. on Automatic Control, vol. 30, no. 8, pages 747–755, 1985.
- [Gustafsson 01] F. Gustafsson. Adaptive filtering and change detection. Wiley, 2001.
- [Isard 98] M. Isard & A. Blake. *A Mixed-State Condensation Tracker With Automatic Model-Switching.* In IEEE Int. Conf. on Computer Vision (ICCV'98), page 107, 1998.
- [Kim 09] H.D. Kim, K. Komatani, T. Ogata & H.G. Okuno. *Binaural Active Audition fo Humanoid Robots to Localize Speech over Entire Azimuth Range.* Applied Bonics and Biomechanics, 2009.
- [Kumon 05] M. Kumon, T. Shimoda, R. Kohzawa & Z. Iwai. *Audio Servo for Robotic Systems with Pinnae.* In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'05), pages 885–890, Edmonton, Canada, 2005.
- [Lasserre 01] J.B. Lasserre. *Global Optimization with Polynomials and the Problem of Moments.* SIAM Journal on Optimization, vol. 11, pages 796–817, 2001.
- [Lu 10] Y.C. Lu & M. Cooke. *Motion Strategies for Binaural Localization of Speech Sources in Azimuth and Distance by Artificial Listeners.* Speech Communication, 2010. disponible en ligne.

- [Martinson 07] E. Martinson. *Acoustic Awareness for Intelligent Robotic Action.* PhD thesis, Georgia Institute of Technology, 2007.
- [Minseok 03] K. Minseok, K. Ichige & H. Arai. *Implementation of FPGA based Fast DOA Estimator using Unitary MUSIC Algorithm.* In IEEE Vehicular Technology Conference, pages 213–217, 2003.
- [Nakadai 08] K. Nakadai, H.G. Okuno, H. Nakajima, Y. Hasegawa & H. Tsujino. *An Open Source Software System for Robot Audition HARK and Its Evaluation.* In IEEE-RAS Int. Conf. on Humanoid Robots, pages 561–566, Daejeon, Corée, 2008.
- [Nikiforov 09] I. Nikiforov. *Estimation et Détection pour les Systèmes Embarqués : Approche Statistique.* Rapport technique, Cours École Doctorale Systèmes, 136 pages, 2009.
- [O'Regan 01] J.K. O'Regan & A. Noë. *A Sensorimotor Account of Vision and Visual Consciousness.* Behavioral and Brain Sciences, vol. 24, pages 939–1031, 2001.
- [Papachristodoulou 05] A. Papachristodoulou & S. Prajna. *Analysis of Non-polynomial Systems using the Sum of Squares Decomposition.* In D. Henrion & A. Garulli, éditeurs, Positive Polynomials in Control. Springer, 2005.
- [Papanikolopoulos 93] N.P. Papanikolopoulos, P.K. Khosla & T Kanade. *Visual Tracking of a Moving Target by a Camera Mounted on a Robot : A Combination of Control and Vision.* IEEE Trans. on Robotics and Automation, vol. 9, no. 1, pages 14–35, 1993.
- [Parrilo 03] P.A. Parrilo. *Semidefinite Programming Relaxations for Semialgebraic Problems.* Mathematical Programming B, vol. 96, pages 293–320, 2003.
- [Stengle 74] G. Stengle. *A Nullstellensatz and a Positivstellensatz in Semialgebraic Geometry.* Mathematische Annalen, vol. 2, pages 87–97, 1974.
- [Tan 06] W. Tan. *Nonlinear Control Analysis and Synthesis using Sum-of-Squares Programming.* PhD thesis, University of California, Berkeley, 2006.
- [Zhang 10] W. Zhang, T.D. Abhayapala & R.A. Kennedy. *Insights into Head-Related Transfer Function : Spatial Dimensionality and Continuous Representation.* Jour. of the Acoustical Society of America, vol. 127, pages 2347–2357, 2010.

Troisième partie

SÉLECTION D'ARTICLES

Les 5 articles joints à ce manuscrit sont les suivants.

Détection-Estimation :

- [JSI-5-DE-1] L. BRÈTHES, F. LERASLE, P. DANÈS, M. FONTMARTY. *Particle Filtering Strategies for Data Fusion dedicated to Visual Tracking from a Mobile Robot*. Machine Vision and Applications, Vol.21, N.4, pp. 427–448, 2010 (en ligne depuis 2008). DOI : 10.1007/s00138-008-0174-7.
- [CIACL-24-DE-9] R. LOPEZ, P. DANÈS, F. ROYER. *Extending the IMM Filter to Heterogeneous-Order State Space Models*. IEEE Conf. on Decision and Control (CDC'2010). Atlanta, GA, 2010.

Asservissement visuel :

- [JSI-2-AV-1] P. DANÈS, D. BELLOT. *Towards an LMI Approach to Multicriteria Visual Servoing in Robotics*. European Journal of Control, Special Issue on “Linear Matrix Inequalities in Control”, Vol.12, N.1, pp. 86–110, 2006. DOI : 10.3166/ejc.12.86–110. “Companion Report” étendu (62p.) sur <http://www.laas.fr/~danes/RESEARCH/VS/companionEJC2k5.pdf>.
- [COCI-1-AV-1] P. DANÈS, D.F. COUTINHO, S. DUROLA. *Multicriteria Analysis of Visual Servos through Rational Systems, Biquadratic Lyapunov Functions, and LMIs*. Chapitre 10 de *Visual Servoing via Advanced Numerical Methods*, G. CHESI & K. HASHIMOTO Eds, pp. 169–188. Lecture Notes in Control and Information Sciences. Springer, Vol. 401/2010. DOI : 10.1007/978-1-84996-089-2_10.

Audition en Robotique :

- [CIACL-22-AR-7] P. DANÈS, J. BONNAL. *Information-Theoretic Detection of Broadband Sources in a Coherent Beamspace MUSIC Scheme*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'2010). Taipei, Taiwan, 2010. Session invitée “Robot Audition”.