



HAL
open science

Applications de la vision omnidirectionnelle à la perception de scènes pour des systèmes mobiles

Xavier Savatier

► **To cite this version:**

Xavier Savatier. Applications de la vision omnidirectionnelle à la perception de scènes pour des systèmes mobiles. Automatique / Robotique. Université de Rouen, France; ESIGELEC - IRSEEM, 2014. tel-01357470

HAL Id: tel-01357470

<https://hal.science/tel-01357470>

Submitted on 29 Aug 2016

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.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

MEMOIRE

Pour l'obtention d'une

Habilitation à Diriger les Recherches (HDR)

Spécialité : **Génie informatique, automatique et traitement du signal**

Présentée par

Xavier SAVATIER

Préparée au sein de l'Institut de Recherche en Systèmes Electroniques Embarqués (IRSEEM) de l'ESIGELEC

dans l'école doctorale Sciences Physiques, Mathématiques et Informatiques pour l'Ingénieur (SPMII) de l'Université de Rouen

« Applications de la vision omnidirectionnelle à la perception de scènes pour des systèmes mobiles »

HDR soutenue le **8 décembre 2014** devant le jury composé de :

M. Patrick RIVES

Directeur de Recherche, INRIA Sophia-Antipolis, président de jury

M. François CHAUMETTE

Directeur de Recherche, INRIA Rennes, rapporteur

M. El Mutapha MOUADDIB

Professeur des Universités, Université Picardie Jules Verne, rapporteur

M. Peter STURM

Directeur de Recherche, INRIA Rhône-Alpes, rapporteur

M. Bélahcène MAZARI

HDR, Directeur de la Recherche et de l'Innovation, CESI, examinateur

M. Pascal VASSEUR

Professeur des Universités, Université de Rouen, examinateur

Remerciements

Ce mémoire présente une synthèse des travaux que j'ai menés dans le domaine de la vision omnidirectionnelle à l'ESIGELEC au sein de son institut de recherche l'IRSEEM. Le thème de la vision omnidirectionnelle, appliquée au véhicule intelligent et à la robotique mobile, auront accompagné le développement de notre laboratoire ; rien n'aurait pu se faire sans la confiance et le soutien de son ancien directeur, Bélahcène Mazari. Je lui renouvelle ici mes plus vifs remerciements.

Le manuscrit résume un travail collectif réalisé par d'excellents doctorants et pour certains désormais collègues - Nicolas Ragot, Rémi Boutteau, Jean-François Layerle, Romain Rossi, Fengchun Dong et Yohan Dupuis. L'attrait originel de cette équipe pour la vision omnidirectionnelle n'aurait pas été aussi vif sans la passion et la créativité de Jean-Yves Ertaud. Je mesure la chance qu'il m'a été donnée de travailler à leurs côtés, je suis fier et honoré de présenter dans ce document une synthèse de leurs travaux respectifs. En dix ans, ce thème de recherche aura été un élément structurant de cette équipe. Les résultats que nous avons obtenus ensemble auront préfiguré les projets que nous menons désormais dans le domaine du véhicule autonome et de la robotique mobile.

Une part importante de mon travail a été dévolue à la mise en place d'une plateforme technologique pour la navigation autonome ; de nombreuses personnes m'ont accompagné dans cette tâche au premier plan desquelles Pierre Merriaux. Sans lui beaucoup de nos rêves n'auraient pu se concrétiser. Bravo et merci à lui pour son talent et son incroyable énergie au service dans tous ces projets.

Je remercie une nouvelle fois Pascal Vasseur de s'être porté garant de mon dossier d'habilitation. Je remercie également mes rapporteurs et les autres membres du jury pour le temps qu'ils ont consacré à l'examen de cette habilitation. Enfin je remercie toutes celles et ceux qui, à titre professionnel ou personnel, m'apporté soutien et encouragements dans la préparation de ce document.

Je dédie ce mémoire à mes parents et à mes enfants, Domitille et Lucas.

Table des matières

PRESENTATION DU MEMOIRE	9
I. SYNTHESE DE MES ACTIVITES D'ENSEIGNEMENT ET DE RECHERCHE	10
I.1 MON PARCOURS	11
FONCTIONS ACTUELLES	11
ACTIVITES PROFESSIONNELLES	11
FORMATION	11
I.2 DESCRIPTION DE MES ACTIVITES D'ENSEIGNEMENT	12
CADRE DES ACTIVITES D'ENSEIGNEMENT EN FORMATION INITIALE	12
LA DOMINANTE « INGENIERIE BIOMEDICALE »	12
LE LABORATOIRE DE SYSTEMES EMBARQUES	13
CONTRIBUTIONS AUX MODULES D'ENSEIGNEMENT EN SYSTEMES EMBARQUES	13
LES PROJETS INGENIEUR	14
FORMATION CONTINUE	15
I.3 SYNTHESE DES ENSEIGNEMENTS DISPENSES	16
ENSEIGNEMENTS EN FORMATION INITIALE A L'ESIGELEC	16
ENSEIGNEMENTS EN FORMATION CONTINUE A L'ESIGELEC	17
ENSEIGNEMENTS EN CYCLE PREPARATOIRE INTEGREE DE L'ESIGELEC	17
I.4 REDACTION D'OUVRAGES ET DE DOCUMENTS DE TRAVAUX PRATIQUES	18
I.5 RESUME DES ACTIVITES DE RECHERCHE	19
I.6 ENCADREMENT DE THESES EN RESPONSABILITE PARTAGEE	20
CO-ENCADREMENT DE THESES (SOUTENUES)	20
CO-ENCADREMENT DE THESES (EN COURS)	21
ENCADREMENT DE STAGES DE D.E.A ET MASTERES	21
ENCADREMENT DE STAGES DE FIN D'ETUDES	21
I.7 RESPONSABILITES AU SEIN DE L'INSTITUT DE RECHERCHE	22
GESTION D'EQUIPE	22
MONTAGE ET COORDINATION DE PROJETS	22
I.8 PARTICIPATION A DES RESEAUX ET RAYONNEMENT SCIENTIFIQUE	24

II. VISION OMNIDIRECTIONNELLE : ETAT DES LIEUX	25
<hr/>	
II.1 SYSTEMES EMBARQUES POUR LA PERCEPTION D'ENVIRONNEMENTS	26
INTRODUCTION	26
SOLUTIONS ACTIVES POUR LA PERCEPTION SPATIOTEMPORELLE D'UNE SCENE	27
SOLUTIONS PASSIVES POUR LA PERCEPTION DE L'ENVIRONNEMENT	29
TECHNOLOGIES POUR LA FORMATION D'IMAGE	30
INTERET D'UNE PERCEPTION OMNIDIRECTIONNELLE DE L'ENVIRONNEMENT EN VISION PAR ORDINATEUR	32
II.2 VISION OMNIDIRECTIONNELLE : TECHNOLOGIES-CLES	34
PRINCIPALES SOLUTIONS TECHNOLOGIQUES POUR L'AUGMENTATION DU CHAMP DE VUE	34
CAPTEURS CATADIOPTRIQUES	35
CONCLUSIONS	39
II.3 VISION OMNIDIRECTIONNELLE : PRINCIPES THEORIQUES ET MODELES ASSOCIES	40
APPROCHE GEOMETRIQUE DE LA VISION PAR ORDINATEUR	40
VISION PERSPECTIVE : MODELE STENOPE	41
MODELES ASSOCIES AUX CAPTEURS CATADIOPTRIQUES A PROJECTION CENTRALE	43
CAS DES CAPTEURS CATADIOPTRIQUES A PROJECTION NON-CENTRALE	45
II.4 COMPREHENSION DE SCENES ET A LA RECONSTRUCTION TRIDIMENSIONNELLE	47
DETECTION DE PRIMITIVES	47
RECONSTRUCTION 3D PAR STEREOVISION	49
RECONSTRUCTION TRIDIMENSIONNELLE PAR DEPLACEMENT DU CAPTEUR	52
II.5 CONCLUSIONS	54
III. PERCEPTION TRIDIMENSIONNELLE DE SCENES PAR VISION OMNIDIRECTIONNELLE ET APPLICATION A LA NAVIGATION AUTONOME DE VEHICULES	55
<hr/>	
III.1 CONTEXTE	56
III.2 CONCEPTION ET CALIBRAGE D'UN BANC DE STEREOVISION OMNIDIRECTIONNELLE	56
METHODOLOGIE DE CONCEPTION	56
MODELE GEOMETRIQUE ET CALIBRAGE	57
III.3 UTILISATION DU BANC DE STEREOVISION OMNIDIRECTIONNELLE POUR LA RECONSTRUCTION TRIDIMENSIONNELLE DE SCENES	60
STEREOVISION PAR APPARIEMENT DE POINTS	61
RECONSTRUCTION 3D DE LIGNES	62
RECONSTRUCTION 3D PAR STEREOVISION ET APPROCHE VOLUMETRIQUE	64

III.4	STRUCTURE FROM MOTION PAR VISION OMNIDIRECTIONNELLE	65
III.5	CONCLUSIONS	66
 IV. CAPTEURS NON CENTRAUX : CONCEPTION, MISE EN ŒUVRE ET APPLICATION A L'AIDE A LA CONDUITE ET A LA NAVIGATION AUTONOME		67
<hr/>		
IV.1	CAPTEURS CATADIOPTRIQUES NON-CENTRAUX : DESIGN PAR PROJECTION PRESCRITE ET CALIBRAGE	68
	CONTEXTE	68
	COMPARAISON DES METHODES DE DESIGN DE MIROIRS	68
	CONCEPTION D'UN MIROIR DOUBLE POUR LA PERCEPTION SIMULTANEE SCENE / CONDUCTEUR	70
	METHODE DE CALIBRAGE	71
	TRIANGULATION	71
IV.2	ODOMETRIE VISUELLE PAR UNE CAMERA PLENOPTIQUE	73
IV.3	CONCLUSIONS	77
 V. DETECTION DE VISAGES ET RECONNAISSANCE BIOMETRIQUE DE LA MARCHE A PARTIR D'IMAGES OMNIDIRECTIONNELLES		78
<hr/>		
V.1	DETECTION ET SUIVI DU VISAGE	79
	TRAVAUX INITIAUX SUR LE SUIVI DU VISAGE	79
	VISION OMNIDIRECTIONNELLE POUR LA DETECTION DE VISAGES ET LA BIOMETRIE	79
	ADAPTATION DU DETECTEUR DE VISAGES DE VIOLA-JONES AUX IMAGES OMNIDIRECTIONNELLES	80
V.2	RECONNAISSANCE BIOMETRIQUE PAR L'OBSERVATION VISUELLE DE LA DEMARCHE D'UN INDIVIDU	82
V.3	CONCLUSIONS	84
 VI. MISE EN PLACE D'UN PLATEAU DE MESURES DEDIE A LA NAVIGATION AUTONOME		85
<hr/>		
VI.1	LE 21^{EME} SIECLE OU L'ESSOR DES VEHICULES AUTONOMES	86
	ETAT DES LIEUX	86
	ENJEUX SCIENTIFIQUE ET TECHNOLOGIQUES DES VEHICULES AUTONOMES	88
VI.2	PLATEFORME DE MESURES POUR LA NAVIGATION AUTONOME	92
 VII. PERSPECTIVES DE RECHERCHE		95
<hr/>		
VII.1	BILAN GENERAL SUR MES ACTIVITES DE RECHERCHE	96

VII.2	ACTIVITES EN COURS ET PERSPECTIVES	96
	UTILISATION DE CAPTEURS « 2D+T »	97
	UTILISATION DE CAPTEURS 3D+T	98
VIII.	<u>LISTE DE DE MES PUBLICATIONS</u>	99
IX.	<u>BIBLIOGRAPHIE</u>	106
X.	<u>ANNEXE : MODÈLE DE DISTORTIONS</u>	110

PRESENTATION DU MEMOIRE

Ce mémoire, rédigé en vue d'obtenir l'Habilitation à Diriger les Recherches, présente une synthèse de mes activités d'enseignement et de recherche menées à l'ESIGELEC et au sein de son Institut de Recherche en Systèmes Electroniques Embarqués, l'IRSEEM, depuis l'obtention de mon doctorat à l'Université de Rouen en 2003.

Le mémoire est organisé en sept chapitres :

Le Chapitre I résume l'ensemble de mon parcours en tant qu'enseignant au sein de l'ESIGELEC, chercheur et responsable du pôle Instrumentation, Informatique et Systèmes à l'IRSEEM.

Le Chapitre II constitue un état de l'art des principaux travaux menés par la communauté scientifique sur le thème de la vision omnidirectionnelle. Ce chapitre expose le socle de connaissances sur lequel s'est appuyé mon travail de recherche et d'encadrement de thèses ; il préfigure l'exposé de mes principales contributions dans ce domaine.

Les chapitres III à V présentent une synthèse de mes contributions dans le domaine de la vision omnidirectionnelle ; elles concernent principalement l'utilisation de capteurs catadioptriques centraux et non centraux pour la perception tridimensionnelle de scènes, l'estimation de mouvement, la détection et suivi de personnes dans une image.

Le chapitre VI met l'accent sur mon travail de mise en place des moyens expérimentaux dévolus à l'étude de la navigation autonome de véhicules au sein du projet structurant de Campus en Intégration des Systèmes Embarqués, projet initié en 2008 et achevé début 2012.

Le chapitre VII conclut sur mes activités de recherche et ouvre une perspective sur les axes de travail que je souhaite suivre.

I. SYNTHÈSE DE MES ACTIVITÉS D'ENSEIGNEMENT ET DE RECHERCHE

Ce chapitre est une synthèse de mes activités en tant qu'enseignant-chercheur au sein de l'ESIGELEC. Dans un premier temps, j'y présente mon parcours de formation, les principaux enseignements dispensés et les responsabilités pédagogiques que j'ai occupées depuis l'obtention de mon doctorat. Dans une seconde partie, j'expose une synthèse de mes travaux de recherche et de mes responsabilités à l'IRSEEM, Institut de Recherche en Systèmes Electroniques Embarqués de l'ESIGELEC.

I.1 Mon parcours

Fonctions actuelles

- Enseignant-chercheur à l'ESIGELEC (Ecole Supérieure d'Ingénieurs)
- Chef-adjoint au Département Systèmes embarqués et Instrumentation de l'ESIGELEC
- Responsable du pôle Instrumentation, Informatique et Systèmes de l'IRSEEM (Institut de Recherche en Systèmes Electroniques EMbarqués)

Activités professionnelles

- Depuis 2004** : Responsable du pôle Instrumentation, Informatique et Systèmes de l'IRSEEM
- 2009 - 2011** : Responsable de la dominante « Ingénierie Biomédicale » de l'ESIGELEC
- Depuis 2001** : Enseignant en systèmes embarqués à l'ESIGELEC
- 1998 – 1999** : Service national : Adjoint au responsable du service informatique de l'Institut National de Invalides (Paris)

Formation

- 2003** : Thèse de doctorat de l'Université de Rouen : «*Etude et mise en œuvre d'un système d'observation de la qualité locomotrice chez l'homme* »
Mention très honorable
- 1997** : Diplôme d'ingénieur ESIGELEC
D.E.A. de Robotique - Université Pierre et Marie Curie (Paris VI)
- 1995** : Maitrise E.E.A. – Université des Sciences et Techniques de Lille (Lille I)

I.2 Description de mes activités d'enseignement

Cadre des activités d'enseignement en formation initiale

Depuis 1999, je suis enseignant à l'ESIGELEC : de 1999 à 2001 en tant que vacataire en parallèle de ma thèse, à partir de 2001 en tant que permanent ESIGELEC. Mon domaine d'enseignement couvre principalement les technologies et les outils pour la conception et la mise au point de systèmes embarqués à microprocesseurs. Les matières enseignées portent sur les architectures matérielles, la programmation en langage C pour les systèmes embarqués, les systèmes embarqués communicants et les méthodologies de conception et de développement. Je dispense aussi des Travaux Pratiques en systèmes embarqués permettant aux étudiants d'appréhender les méthodologies et outils - logiciels et matériels - nécessaires au développement et au test de tels systèmes. Ces enseignements sont dispensés dans les différentes formations proposées à l'école à savoir la filière initiale, l'apprentissage, la filière Fontanet et la formation continue.

Parallèlement à cette activité, j'ai mis en place et animé pendant deux ans (2009-2011) la dominante « Ingénierie Biomédicale », sous la responsabilité du directeur des Etudes de l'ESIGELEC.

La dominante « Ingénierie Biomédicale »

En 2009, j'ai été chargé par la direction des Etudes de structurer une nouvelle dominante destinée aux élèves souhaitant travailler dans le domaine des nouvelles technologies pour le biomédical. J'ai mis en place un comité de pilotage afin d'être en accord avec les attentes des professionnels en termes de formation afin de définir de façon pertinente le contenu des modules de cette dominante. Ce comité a été constitué de membres issus du monde médical et hospitalier (praticiens hospitaliers du CHU de Rouen, INSERM), ainsi que des PME et des porteurs de projet de création d'entreprise dans le domaine des technologies pour le biomédical (Quidd, Robocath, Ophtimalia).

Avec l'appui du comité de pilotage, nous avons structuré cette dominante en trois grands thèmes : instrumentation et systèmes embarqués pour la santé et le bien-être, compléments en sciences du vivant (physiologie, biomécanique), cycle de développement et aspects réglementaires et normatifs. Sa mise en œuvre a nécessité un important travail de recherche d'intervenants - en France et à l'étranger via le réseau à l'international de l'ESIGELEC – et la coordination des enseignements en lien avec le service des Etudes de l'ESIGELEC. Pour ce faire, j'ai sollicité des praticiens hospitaliers reconnus dans leur domaine mais aussi des industriels afin de couvrir des champs disciplinaires nouveaux à l'école (physiologie, biomécanique, microélectronique) et d'apporter aux étudiants une

vision extérieure sur l'utilisation des nouvelles technologies de l'électronique des systèmes embarqués pour la santé.

Depuis novembre 2011, devant le développement de l'activité de recherche, j'ai cédé la responsabilité de la dominante « ingénierie biomédicale » mais je continue d'y apporter mon soutien et mon expertise en tant que membre de son comité de pilotage.

Le laboratoire de systèmes embarqués

En accord avec le chef du département Systèmes Embarqués et Instrumentation de l'ESIGELEC, j'ai adapté son laboratoire d'enseignement en fonction des évolutions des contenus des dominantes qui y sont adossés : dominantes « Ingénierie des Systèmes Embarqués » et « Ingénierie Biomédicale ».

Mes principales contributions ont porté sur :

- Le développement d'outils de maintenance informatique
- Le développement et le déploiement de maquettes illustrant les concepts de communication internet sur systèmes embarqués (Freescale HS12 et pile OpenTCP)
- Le développement et le déploiement d'une maquette d'instrumentation biomédicale (chaîne d'instrumentation virtuelle représentative d'un Electrocardiogramme)

Contributions aux modules d'enseignement en systèmes embarqués

J'ai créé quatre modules spécifiques en rapport avec la conception de systèmes embarqués et le développement logiciel :

- *Compilation C pour les systèmes embarqués* (cours et travaux pratiques) de langage C, module créé en 2003 et destiné aux étudiants de deuxième année de la dominante « ingénierie des systèmes embarqués ».

Ce cours permet aux étudiants de compléter leurs connaissances en développement logiciel en langage C et d'appréhender les étapes de développement et de test d'un logiciel pour un système embarqué à microprocesseur. Depuis sa création en 2003, les cours et les travaux pratiques ont été modifiés à plusieurs reprises pour suivre les évolutions technologiques dans ce domaine (adaptation aux cibles Freescale Coldfire 5307, Coldfire 5213, HS12DP256 puis Texas Instruments MSP430).

- *Embarqué communicant* (cours et travaux pratiques), module créé en 2007 et destiné aux étudiants de troisième année de la dominante « ingénierie des systèmes embarqués ». Ce module introduit les concepts et solutions technologiques pour rendre un système à microprocesseurs communicant (Ethernet, TCP/IP, Zigbee, USB) en couvrant les aspects matériels et logiciels et avec une ouverture sur l'internet des objets.
- *Conception de systèmes embarqués* (cours et travaux pratiques), module conçu en 2011 à destination des étudiants de troisième année de la dominante « ingénierie biomédicale ». Ce module décrit les étapes-clés d'un développement en embarqué, de la conception jusqu'aux tests en partant d'une analyse système. Le propos est illustré par des études de cas de solutions embarquées de monitoring de constantes physiologiques.
- *Courants Porteurs en Ligne* (cours), module créé en 2004 à destination des étudiants de deuxième année de la dominante « Génie Electrique et Transport ». Ce cours introduit les principes généraux et les technologies clés de la communication sur courant porteur (CPL) en haut et bas débit.

Jusque 2009, je suis intervenu dans les modules de tronc commun portant sur la programmation de microprocesseurs (travaux dirigés et travaux pratiques). Enfin et depuis la même année, j'interviens dans le module « projet scientifique » à destination des étudiants de deux masters internationaux de l'ESIGELEC (*embedded systems* et *Control and Embedded Instrumentation*), interventions faites en anglais.

Au sein du département, mon action vise aussi à faciliter les passerelles entre les activités de recherche de l'IRSEEM et les activités du département SEI de l'ESIGELEC. A ce titre, j'ai participé avec d'autres collègues à la rédaction de deux papiers pour le CETSIS [ACTN2] [ACTN4] faisant la promotion d'une utilisation pour la pédagogie de méthodologies et d'outils utilisés en recherche.

Les projets ingénieur

La pédagogie par projets a été introduite au début des années 2000 à l'ESIGELEC ; les projets ingénieurs en sont une des déclinaisons. Constitués en équipe des six, les étudiants de deuxième année doivent mener un projet complet, de l'étude à la réalisation d'une maquette, et ce pour le compte d'un commanditaire, en général un industriel.

Je suis intervenu dans ce dispositif à plusieurs titres :

- en tant que binôme d'encadrement avec un rôle de contrôle et d'accompagnement sur le plan du management de projet (jusqu'à 2009, j'ai suivi en moyenne entre deux à trois équipes chaque année) ;
- en tant qu'instructeur chargé du suivi technique de projets (deux à trois équipes par an) ;
- en tant que commanditaire d'études et de réalisations dans le cadre des activités de recherche du pôle instrumentation, informatique et systèmes de l'IRSEEM.

Formation continue

Dans le cadre des actions de formation continue de l'ESIGELEC, j'ai conçu et dispensé plusieurs formations dans le domaine des systèmes embarqués :

- Langage C pour les systèmes embarqués
- Embarqué communicant

Je suis aussi intervenu dans des formations spécifiques sur les systèmes numériques.

En 2005, j'ai assuré en binôme une formation « langage C pour les systèmes embarqués » dans le cadre de l'action collective « filière électronique » pilotée par l'ESIGELEC.

Entreprises bénéficiaires : Whirpool, Coria, S2M, Jeulin, TCAR, Bayer,...

Une liste détaillée des formations que j'ai dispensées est donnée dans le chapitre suivant.

I.3 Synthèse des enseignements dispensés

Enseignements en formation initiale à l'ESIGELEC

2000 - 2001	TD de circuits de base, 1 ^{ère} année du cycle ingénieur - 56H ¹
2003 - 2006	TD et TP d'électronique numérique, 1 ^{ère} année du cycle ingénieur - 36H
2001 - 2009	TD et TP de systèmes numériques à microprocesseurs et interfaçage numérique, 2 ^{ème} année du cycle ingénieur (formation initiale et par l'apprentissage) – 140H
2004	TP de commande numérique de machines par DSP, 2 ^{ème} année de la dominante « Génie Electrique et transport » - 9H
2007 - 2008	TP de systèmes embarqués, 3 ^{ème} année de la dominante « Ingénierie des systèmes embarqués » - 42H
2004 - 2005	Cours de courants porteurs en ligne, 2 ^{ème} année de la dominante « Génie Electrique et transport » - 14H
2003 - 2011	Cours et TP de Compilation C pour l'embarqué, 2 ^{ème} année de la dominante « Ingénierie des systèmes embarqués » - 30H
2007 -...	Cours et TP d'embarqué communicant, 3 ^{ème} année de la dominante « Ingénierie des systèmes embarqués » - 48H
2009 -...	Cours et TP de conception de systèmes embarqués, 2 ^{ème} année de la dominante « Ingénierie Biomédicale » - 26H
2010-2011	Mini-projet, 3 ^{ème} année de la dominante « Ingénierie Biomédicale » - 10H
2001-2002	Projet scientifique, 3 ^{ème} année du cycle ingénieur - 45H
2002-2008	Encadrement de projets ingénieurs, 2 ^{ème} et 3 ^{ème} année du cycle ingénieur - 20H
2004-...	Instruction de projets ingénieurs, 2 ^{ème} et 3 ^{ème} année du cycle ingénieur - 16H
2009-...	Projets master, masters internationaux en systèmes embarqués - 10H

¹ Nombre d'heures équivalent TD en moyenne par an

La figure ci-dessous (Figure 1) résume mon activité d'enseignement depuis 2001 :

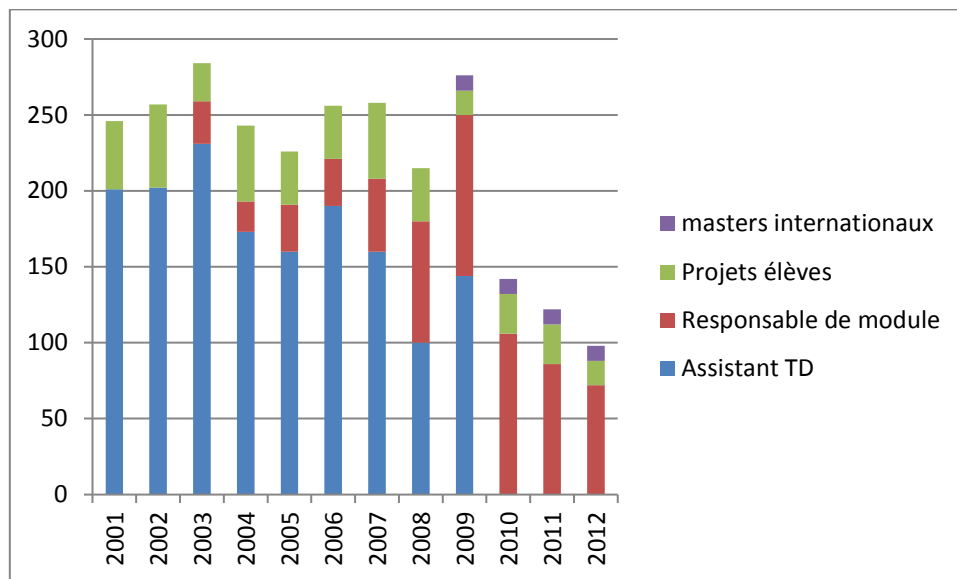


Figure 1 : histogramme de mon activité d'enseignement en heures équivalent TD et regroupé en type d'activités

Enseignements en formation continue à l'ESIGELEC

- 2003 Formation en logique combinatoire et séquentielle, BAYER
- 2003 Formation spécifique « centrale de tachymétrie du métro du Rouen », TCAR
- 2005-... Formation en langage C pour les systèmes embarqués
- 2012 Formation en embarqué communicant, S2M

Enseignements en Cycle Préparatoire Intégrée de l'ESIGELEC

- 2003-2004 Cours et TP en Sciences de l'Ingénieur (automatique, cinématique du solide), 1^{ère} et 2^{ème} année du cycle préparatoire intégrée de l'ESIGELEC

I.4 Rédaction d'ouvrages et de documents de travaux pratiques

[1] *Programmation avancée des systèmes à microprocesseurs*, ESIGELEC – 2010

- La chaîne de développement C pour un système embarqué
- Langage C pour l'embarqué
- Mécanisme de compilation et de test
- Qualité logicielle : règles de codage MISRA

[2] *Embarqué communicant*, ESIGELEC – 2011

- Modèles en couche, protocoles Internet (TCP, IP, ARP, Ethernet,...)
- Internet sur l'embarqué : concepts et solutions technologiques
- Communication sans fils : principes, protocole ZIGBEE
- Introduction à l'USB

[3] *Conception de systèmes biomédicaux embarqués*, ESIGELEC – 2011

- Classification des systèmes embarqués pour le biomédical
- Cycle de développement (conception, mise en œuvre, intégration, tests)
- Architectures matérielles et aspects logiciels
- Transmission de données sans fils

[4] *Transmission d'information par courants porteurs en ligne*, ESIGELEC, 2005

- Principe de la transmission numérique sur courants porteurs
- Perturbations du canal et techniques de fiabilisation de la transmission
- Courants porteurs bas débit, haut débit, indoor, outdoor : principes, usages, technologies, acteurs, réglementation

I.5 Résumé des activités de recherche

Mes activités de recherche ont débuté en tant que doctorant à l'Université de Rouen au sein de l'ESIGELEC, elles ont porté dans un premier temps sur la conception et l'évaluation de dispositifs de mesure de la dynamique de la marche de personnes atteintes de pathologies de la hanche. Ce travail a été mené en lien avec le Groupe de Recherche sur le handicap Locomoteur (GRHAL), associant des équipes du CHU de Rouen et du CETAPS de l'université de Rouen. En 2003 j'ai rejoint les équipes de recherche qui se constituaient avec la mise sur pieds de l'IRSEEM, Institut de Recherche en Systèmes Electroniques Embarqués, créé en 2001. En 2004, j'ai pris la responsabilité d'un des pôles de recherche de l'IRSEEM – le pôle Instrumentation, Informatique et Systèmes - avec pour mission d'y développer l'activité et comme principales actions la mise en place d'une animation scientifique, le montage et le suivi de projets de recherche partenariale et la mise en place des outils de suivi et de synthèse des activités. A ce poste, dans le prolongement de mes travaux dans le domaine de l'instrumentation sur l'homme, j'ai structuré et développé une recherche dans le domaine de la vision par ordinateur mais plus appliquée au monde de l'automobile et des transports terrestres (véhicule intelligent, robotique mobile) avec un focus sur la perception tridimensionnelle de scènes pour la navigation autonome de véhicules.

L'utilisation de la vision par ordinateur pour des tâches de navigation d'une voiture ou d'un robot mobile est un thème de recherche très exploré depuis plusieurs années. Dans la majorité des cas, la chaîne d'acquisition fait intervenir une ou plusieurs caméras que l'on peut qualifier de « conventionnelles » au sens où elles forment une image perspective semblable à notre propre perception visuelle. Mes travaux se sont inscrits dans une voie alternative, que la communauté a commencé à explorer fin des années 1990, partant du constat qu'une formation géométrique de l'image analogue à notre propre perception du monde n'est peut-être pas la meilleure représentation à fournir à une machine. J'ai donc concentré mes travaux sur l'étude de systèmes de vision omnidirectionnelle tels que les capteurs catadioptriques centraux et non centraux et leur utilisation pour des applications mobiles embarquées ou débarquées : modélisation et calibrage, reconstruction tridimensionnelle de scènes, calcul de pose, reconnaissance de primitives et d'objets. La suite du document présente une synthèse de mes travaux dans ces domaines.

Ces activités ont fait l'objet de six thèses soutenues et que j'ai co-encadrées. Une septième (Wan Fayu, [Th5]) a été encadrée sur un sujet différent en rapport avec mon expertise dans les systèmes électroniques embarqués.

I.6 Encadrement de thèses en responsabilité partagée

Co-encadrement de thèses (soutenues)

Nom du doctorant (Date de soutenance)	Titre de la thèse	Directeur de thèse	Ecole doctorale	Taux encadrement
[Th1] RAGOT Nicolas (Avril 2009)	Conception d'un capteur de stéréovision omnidirectionnelle : architecture, étalonnage et applications à la reconstruction de scènes 3D	B. MAZARI	SPMII (ED n° 351)	30%
[Th2] BOUTTEAU Rémi (Avril 2010)	Reconstruction tridimensionnelle de l'environnement d'un robot mobile, à partir d'informations de vision omnidirectionnelle, pour la préparation d'interventions.	B. MAZARI	SPMII (ED n° 351)	30%
[Th3] LAYERLE Jean-François (Juil. 2010)	Conception et calibration de systèmes de vision catadioptrique pour une application d'analyse du comportement de conduite	M. MOUADDIB	EDSS (ED n° 368)	50%
[Th4] ROSSI Romain (Juil. 2011)	Reconstruction 3D volumétrique par vision omnidirectionnelle sur architecture massivement parallèle	B. MAZARI	SPMII (ED n° 351)	30%
[Th5] FAYU Wan (Juil. 2011)	Software based approaches to improve EMC of a microcontroller using in situ EMI monitoring	A. LOUIS	SPMII (ED n° 351)	30%
[Th6] DONG Fengchun (Nov. 2012)	Conception d'un capteur polydioptrique pour la navigation autonome	R. BENOSMAN	SPMII (ED n° 351)	30%
[Th7] DUPUIS Yohan (Déc. 2012)	Authentification nomade par biométrie multimodale	P. VASSEUR	SPMII (ED n° 351)	50%

Co-encadrement de thèses (en cours)

Nom du doctorant (année de soutenance prévue)	Titre de la thèse	Directeur de thèse	Ecole doctorale	Taux encadrement
[Th8] BONARDI Fabien (2016)	Conception, réalisation et évaluation d'un système de vision embarqué et multimodal visible infrarouge pour la navigation autonome de véhicules	P. VASSEUR	SPMII (ED n° 351)	30%
[Th9] MERRIAUX Pierre (2016)	Localisation robuste par vision pour la navigation autonome	P. VASSEUR	SPMII (ED n° 351)	50%

Encadrement de stages de D.E.A et mastères

[Mast1] J.F. LAYERLE, « méthodes de conception d'interfaces homme – machine ergonomiques et adaptables », Rapport de projet de fin d'études, INSA de Toulouse, 2006

Encadrement de stages de fin d'études

[S1] L. BELHALLOUCHE, « Méthodes de reconnaissance faciale pour l'identification biométrique », stage de fin d'études, EMI (Maroc), 2009

[S2] V.N.V SARATH, « Design of inertial sensor using accelerometer & gyro meter data fusion », NIT Warangal (Inde), 2009

[S3] A. PINOL TRULLEN, « Geometrical analysis of mobile 3D vision system for autonomous navigation », ETSEIAT (Espagne), 2006

I.7 Responsabilités au sein de l'institut de recherche

Gestion d'équipe

J'ai rejoint l'équipe de direction de l'IRSEEM en 2004, laboratoire alors structuré en trois pôles : Electronique, Automatique et Instrumentation-Informatique ; j'ai pris la responsabilité de ce dernier. Le pôle est aujourd'hui constitué de seize enseignants-chercheurs en informatique, vision par ordinateur et robotique. J'y assure un rôle d'animation scientifique - réunions d'équipe, animation de groupes thématiques - la gestion administrative et budgétaire, et la coordination générale des programmes de recherche qui y sont rattachés (dix actuellement en cours). Comme les autres responsables de pôle, j'assure également une mission de représentation de l'IRSEEM par délégation de la direction générale de l'ESIGELEC / IRSEEM. Je suis membre du Conseil de l'IRSEEM.

Montage et coordination de projets

Depuis ma prise de poste en 2004, j'ai participé, en tant que rédacteur principal, au montage de plus d'une vingtaine de propositions de projets dans les appels européens (INTERREG, ARTEMIS), nationaux (ANR, ADEME, FUI) et régionaux dans les Grands Réseaux de Recherche Electronique, Energie, Matériaux (GRR EEM) et Logistique, Mobilité, Numérique (GRR LMN). Plus d'une dizaine de propositions ont été retenues pour financement (cf. Tableau 1 ci-après). Ces projets se sont inscrits dans le cadre de collaborations que j'ai développées en Europe (Université du Kent, université du Sussex) et en France (IFSTAR, Mines Paritech, INRIA, UPJV, Université de Rouen, etc.) en lien étroit avec le tissu industriel (automobile, défense, transports terrestres et maritimes) et avec l'appui de la direction de l'ESIGELEC et de l'IRSEEM.

Tableau 1 : liste des projets de recherche pour lesquels j'ai agi en tant que responsable scientifique

Nom du projet	Dispositif de financement (Budget IRSEEM)	Rôle dans le projet	Dates début/fin
OMNIVISS	INTERREG IIIA (690 K€)	Coordinateur du projet Porteur du projet : IRSEEM Partenaires : Université du Kent (UK)	2005-2008
COGITO	GRR TLTI (10 K€)	Chef de projet Porteur du projet : IRSEEM	2009
NOBA	INTERREG IVA + GRR EEM (640 K€)	Coordinateur du projet Porteur du projet : IRSEEM Partenaire : Université du Kent	2009-2013
MIRIADE	DGA RAPID (102 K€)	Chef de projet IRSEEM Porteur du projet : STARNAV Partenaires : IRSEEM, SOMINEX	2009-2011
ROROMAX	Etat + Région HN (84 K€)	Chef de projet IRSEEM porteur du projet : CRITT TL Plus de 20 partenaires régionaux	2011-2012
VIATIC	ANR ASTRID (193 K€)	Coordinateur de l'ANR Partenaires : MBDA et EHESS	2011-2014
NAVALIS	ADEME (522 K€)	Chef de projet IRSEEM Porteur : MSI Chantiers Allais	2012-2015
DESERVE	ARTEMIS (285 K€)	Chef de projet Porteur : VTT FINLANDE + de 20 partenaires R&D automobile	2013-2015
SAVEMORE	INTERREG IV + GRR EEM (253 K€)	Coordinateur du projet Partenaires : LITIS-INSA de ROUEN, Université du Kent	2013-2015
SERBER	CARNOT ESP (100 K€)	Coordinateur du projet	2013-2014
BLIDAR	Pôles de compétitivité (DERBI, Mer-Bretagne) (226 K€)	Chef de projet IRSEEM Porteur : EOLFI Partenaires : IFREMER, NKE	2014-2016
ARGOS	Challenge TOTAL/ANR (500 K€)	Coordinateur Partenaire : SOMINEX	2014-2016

I.8 Participation à des réseaux et rayonnement scientifique

Sur le plan international, outre la participation aux conférences du domaine en tant que co-auteur, je suis relecteur d'articles scientifiques pour des conférences internationales en comité de lecture (ICRA, IROS) ainsi que pour les revues IET-CVT et IEEE-CSVT. Je participe à un projet européen (DESERVE) avec des échanges réguliers avec des laboratoires et industriels du monde de la R&D automobile (CRF FIAT, CONTINENTAL, VTT Finlande, CTAG Espagne, IKA/FKA, INRIA, Mines Paritech, etc.). Je collabore entre autres avec le département d'électronique de l'Université du Kent (trois projets INTERREG depuis 2006), SupCom Tunis (collaborations avec Adel Ghazel, échanges d'étudiants, visites tant que professeur invité, co-encadrement d'une thèse avec SupCom depuis 2014), Beihang University en Chine (3 visites entre 2006 et 2009, dépôt d'une ANR blanc international avec l'équipe de Mao Xia), Nagoya Institute of Technology (visite en professeur invité par Jun Sato en 2011).

Au niveau national, je participe régulièrement aux journées de recherche du GDR ISIS et plus récemment du GDR robotique. Depuis 2006, je participe aux groupes de travail du Domaine d'Activités Stratégiques (DAS) Sécurité et Mobilité du pôle de compétitivité MOV'EO. Au sein de l'Institut de la transition Énergétique VEDECOM, je représente l'ESIGELEC, membre fondateur, à son Conseil d'Administration. Je suis également membre du Comité de Domaine 2 « délégation de la conduite et connectivité ». Depuis deux ans, j'ai renforcé mon action pour la promotion de nos travaux au niveau national et j'ai mené plusieurs actions de collaboration avec des laboratoires universitaires : laboratoire de microrobotique de Marseille (collaborations en cours de montage avec Frank Ruffier), Institut de mécanique Céleste de Calcul de Ephémérides de l'Observatoire de Paris (deux soumissions de projets européens en 2013) ou encore l'institut Carnot Armines.

Au niveau régional, je suis membre du conseil scientifique du sous-réseau électronique du GRR EEM et depuis cette année du GRR LMN. J'y soumetts régulièrement des projets de travaux de recherche en partenariat avec d'autres laboratoires. Je participe activement aux groupes de travail pilotés par la Région Haute-Normandie dans une démarche prospective en lien avec les enjeux du territoire (écomobilités, diffusion de l'électronique embarquée et des nouvelles technologies de l'information et de la communication, nouveaux usages). Je suis régulièrement mobilisé pour des actions d'accompagnement et d'expertise pour entreprises, notamment des PME normandes du domaine de l'électronique.

II. VISION OMNIDIRECTIONNELLE : ETAT DES LIEUX

Ce chapitre introduit, sous la forme d'un état de l'art, les principaux concepts associés à la vision omnidirectionnelle – capteurs, modèles, applications – domaine dans lequel s'inscrivent les contributions des thèses que j'ai co-encadrées et qui sont présentées dans les chapitres suivants.

II.1 Systèmes embarqués pour la perception d'environnements

Introduction

La perception de l'espace qui entoure un véhicule afin de le doter de capacités de mobilité, surveillance, interaction et plus largement d'autonomie, est un sujet amplement étudié depuis plus de trente ans. Cette problématique a été centrale dès les premières recherches en robotique mobile. La tâche de perception est en effet un maillon essentiel et critique pour aller vers des véhicules de plus en plus autonomes et qui seront au cœur des futurs moyens de transport en général, l'automobile en particulier. Dans ce type d'applications, on peut considérer que la perception de la scène est avant tout spatio-temporelle, le véhicule doit établir une représentation géométrique et tridimensionnelle d'un environnement dont les propriétés physiques varient dans le temps. La première raison de cette variation est le fait que la scène est composée d'objets en mouvement et que le véhicule est lui-même amené à se déplacer. En environnement extérieur et naturel, l'évolution des conditions environnementales et la diversité des scènes parcourues renforcent le caractère dynamique du monde à percevoir. Le véhicule pourra appréhender le monde qui l'environne essentiellement à partir de mesures physiques des rayonnements électromagnétiques émis ou réfléchis par les objets de la scène. Deux grandes familles de capteurs vont permettre cette mesure : les solutions par mesure active - émission d'une onde électromagnétique et mesure de sa réflexion par la scène - et les solutions passives – principalement en utilisant des caméras. Il faut souligner qu'il n'existe pas solutions « universelles » capables d'analyser, comme nous le faisons tous les jours, des scènes complexes, en particulier en environnement naturel.

Dans une première partie de ce chapitre, nous présentons brièvement ces deux grandes familles de capteurs. En effet, si ce mémoire porte exclusivement sur une synthèse de travaux exploratoires pour la perception omnidirectionnelle et tridimensionnelle de scènes par vision, cette étude ne peut être dissociée d'un état des lieux plus large sur d'autres modalités de perception qu'il faut aujourd'hui coupler pour pallier les limitations de chacune. L'autre raison est que le choix de la vision en embarqué est souvent guidé par des considérations de coût et de fiabilité du capteur ; d'autres technologies (Lidar en particulier, dont le principe est rappelé ci-après) donnent souvent de meilleures performances mais sont souvent écartées pour les mêmes considérations. Dans un travail expérimental, il est cependant utile de connaître ces technologies, leurs performances et conditions d'utilisation afin de les comparer entre elles et d'objectiver un choix.

Solutions actives pour la perception spatio-temporelle d'une scène

Les capteurs actifs émettent un signal pour mesurer la distance entre un capteur et une cible observée. On distingue principalement les capteurs à télémétrie et les capteurs actifs par triangulation. Les capteurs à télémétrie utilisent l'énergie d'une onde incidente rétrodiffusée par une cible pour mesurer la distance entre le capteur et l'objet observé. Les signaux émis peuvent être des ondes électromagnétiques lumineuses opérant dans le visible, le proche infrarouge ou l'ultraviolet, des ondes électromagnétiques radiofréquences ou des ondes mécaniques acoustiques. Le capteur peut fonctionner en régime impulsionnel ou en régime sinusoïdal à fréquence fixe ou variable. En régime impulsionnel, la mesure de distance est basée sur le calcul du temps de parcours d'un signal entre son émission et sa réception. En mode sinusoïdal, la mesure de distance est soit une mesure de déphasage, soit une différence de fréquence entre l'onde émise et l'onde reçue. Les capteurs actifs par triangulation sont composés d'une source lumineuse et d'un capteur vidéo. Cette méthode de mesure de distance est dite en lumière structurée. Dans sa forme la plus simple, la lumière structurée est un unique rayon. Les projections de plans lumineux ou de patrons bidimensionnels sont également utilisées. Le principal avantage des solutions actives par télémétrie ou triangulation est qu'elles fournissent immédiatement une mesure tridimensionnelle de l'environnement qui peut être exploitée pour générer une reconstruction dense de la scène.

Sur ce modèle, les capteurs actifs les plus connus sont ceux basés sur le principe du LIDAR (*Light Detection And Ranging*) qui, par analogie avec le RADAR, exploite l'information de lumière (visible ou infrarouge) renvoyée par un obstacle à son émetteur. Il existe de nombreuses solutions technologiques exploitant ce principe (Figure 2) à partir d'une mesure ponctuelle (télémètre 1D), une nappe par mise en rotation du faisceau (2D) ou sur plusieurs niveaux de nappes par un mouvement conjugué de rotation et de pivot de la tête optique (3D). Le faisceau lumineux du LIDAR se présentant sous la forme d'un cône, certains fabricants proposent désormais des solutions de type multi échos afin de mesurer les distances d'obstacles partiellement occultés mais encore visibles dans une partie du cône du laser ; cette technique rend aussi la mesure moins sensible à des phénomènes perturbateurs, tels que la pluie.

Parmi les principaux inconvénients des mesures par LIDAR pour des applications terrestres, on peut souligner :

- une portée limitée (mesure entre 100 et 200 mètres pour un télémètre laser essentiellement pour des questions d'énergie à embarquer, de taille de l'optique et de réglementation sur la puissance du signal) ;

- une zone aveugle en proximité du capteur, le temps de vol étant trop court pour être mesurés (1 mètre pour le vélocity par exemple) ;
- un champ de vue qui peut être limité (parfois inférieur à 90°) ;
- une sensibilité aux conditions d'environnement : la pluie ou un brouillard vont perturber la propagation du faisceau lumineux ; l'effet de cette perturbation peut être contenu si le Lidar est doté d'une technologie multi-échos ;
- la complexité de certaines technologies dont le coût et la fragilité les rendent incompatibles avec une utilisation en masse.



Figure 2 : exemple de capteurs de perception 3D par mesure active : a) UTM-30LX Hokuyo © ; b) Lidar 8 nappes IBEO © ; c) Lidar 64 nappes 360° Velodyne © ; d) caméra ToF Mesa Imaging ©

Les caméras dites RGB-D (RGB pour la couleur, D pour Depth) sont l'autre technologie suscitant un fort engouement depuis quelques années. Aussi appelées caméras *Time of Flight* (ToF) ou *flash lidar*, elles exploitent le signal optique émis par la caméra, retourné par les obstacles et mesuré via une matrice photosensible (Figure 2). Les pixels de l'image contiennent non seulement une information sur l'image (RGB) mais aussi sur la profondeur de la scène (D pour Depth). L'avantage par rapport à un LIDAR est d'éviter la mise en rotation de l'optique pour réaliser le scan d'une scène. Le principe sous-jacent de formation et de mesure du signal varie d'une technologie à l'autre. La mesure est généralement cantonnée à quelques mètres même si des solutions utilisant une source laser permettent une perception jusqu'à plusieurs kilomètres [59]. L'autre solution est d'utiliser un projecteur de lumière structurée (infrarouge) décalé par rapport au capteur d'images. La méthode est analogue à la stéréovision, la lumière structurée facilitant l'appariement des points et leur triangulation. De manière générale, tous ces systèmes de vision active sont de plus en plus utilisés pour des applications en robotique mobile ou pour l'automobile. Cependant, la mesure de distance est généralement entachée d'un bruit fonction du matériau réfléchissant l'onde lumineuse et de la structure 3D de la scène. La modélisation de ce bruit est difficile, il est quasiment impossible de l'éliminer en post-traitement [15] [22].

Solutions passives pour la perception de l'environnement

Par analogie avec notre propre façon de percevoir notre environnement et de nous y déplacer, les systèmes de vision fournissent de nombreuses informations exploitables pour des tâches de navigation du véhicule ; ils sont donc de plus en plus utilisés comme moyen de perception embarquée. L'utilisation d'une ou plusieurs caméras permet en théorie la détection et la reconnaissance d'informations utiles (obstacles, panneaux, autres véhicules), la compréhension du mouvement propre du véhicule, le calcul de la structure tridimensionnelle de l'environnement à des fins de cartographie ou de localisation. Les méthodes de traitement sous-jacentes se basent soit sur une analyse directe de l'image (détection et suivi de primitives, segmentation de l'image, reconnaissance de formes), soit impliquent de retrouver la structure géométrique tridimensionnelle de la scène observée. Enfin, les progrès technologiques rendent possibles l'intégration de caméras à faibles coût et avec un haut niveau de fiabilité ; l'évolution des puissances de traitement et de stockage des calculateurs embarqués autorisant l'exécution en temps réel et sur cible embarquée d'algorithmes de plus en plus complexes.

Les capteurs passifs vont utiliser les propriétés de réflexion du spectre lumineux des objets d'une scène ou encore exploiter le rayonnement électromagnétique (infrarouge) de la chaleur des corps. Cette mesure se fait à l'aide de photodétecteurs généralement organisés sous la forme d'une matrice photosensible au sein d'une caméra. Ces cellules vont permettre l'acquisition et la quantification de la lumière de l'environnement observé, la lumière pouvant se décrire comme un ensemble d'ondes électromagnétiques ou comme un flux de particules (photons). Cette dualité autorise une représentation soit comme la propagation d'une onde lumineuse, soit sous forme de rayons lumineux. C'est sur cette seconde représentation que se fondent les approches géométriques de la vision par ordinateur.

Notre œil possède une sensibilité spectrale située dans les longueurs d'onde allant de 400nm à 700nm, domaine spectral de la lumière que l'on qualifie de visible. Les capteurs d'image sont généralement conçus pour avoir une réponse spectrale centrée autour de 550nm, valeur qui correspond à un pic de sensibilité de l'œil (vert/jaune). Les composés chimiques qui entrent dans la fabrication des capteurs définiront le domaine de sensibilité de la caméra. Comme le montre le Tableau 2, le spectre visible représente une bande étroite du spectre lumineux. Les gammes NIR et SWIR ont un comportement physique similaire à la lumière visible, c'est-à-dire que le capteur doit détecter le reflet de l'éclairage de la scène. De jour, la lumière naturelle émise par le soleil est tout à fait exploitable, de nuit une lumière reflétée par la pleine lune suffira à avoir une vision nocturne correcte. Les éclairages artificiels (incandescent, LED,...) offre généralement un rayonnement

suffisant à une bonne exploitation à courte distance. En outre, ces deux classes d'infrarouges sont particulièrement adaptées à la vision en environnement contraint par des effets atmosphériques (diffusion de la lumière) tels que la fumée ou le brouillard. Cependant, les caméras NIR et SWIR présentent une forte sensibilité aux éclairages (urbains, phares des véhicules) avec des effets de saturation locale importante (blooming).

Tableau 2 : spectre électromagnétique des ondes lumineuses

Onde lumineuse	Longueur d'onde (nm)
Ultra Violet (UV)	100 – 400
Visible (Vis)	400 – 700
Proche infrarouge (NIR)	750 – 1400
Infrarouge ondes courtes (SWIR)	1400 – 3000
Infrarouge ondes moyennes (MWIR)	3000 – 8000
Infrarouge longues ondes (LWIR)	8000 – 15000
Infrarouge lointain (FIR)	> 15000

Les gammes LWIR et MWIR peuvent travailler avec des scènes non éclairées puisque tout corps émet un rayonnement propre et de ce fait « visible ». C'est ce que l'on nomme par extension l'imagerie thermique. L'image rend alors compte de deux constituant principaux, d'une part la température de l'objet (plus l'objet est chaud, plus il rayonne), d'autre part, de l'émissivité de l'objet qui est une caractéristique intrinsèque de son rayonnement. Le comportement physique est différent du spectre visible - par exemple le verre devient opaque - ce qui impose l'utilisation d'éléments optiques particuliers. Dans la navigation, du fait de l'opacité de l'eau dans ces longueurs d'ondes, la pluie peut alors masquer une scène. L'imagerie thermique est relativement constante quelles que soient les conditions d'illumination.

Technologies pour la formation d'image

Aujourd'hui, le marché est dominé par deux technologies : CCD (Charge-Coupled Device) et CMOS (Complementary Metal Oxide Semi-conductor). Toutes les deux exploitent le même phénomène physique, l'effet photoélectrique, et suivent le même principe : les photons arrivant sur le silicium de l'imageur vont créer des charges électriques qui s'accumulent pendant un temps dit d'intégration. Les capteurs CMOS ont des avantages déterminants : un faible coût avec des résolutions élevées. Néanmoins ils trouvent leur limite dans des environnements à faible luminosité où l'on constate un bruit important sur les images. Cette technologie étant quatre fois moins coûteuse à fabriquer que le CCD, elle a inondé le marché (95% des matrices fabriquées en 2014). Dans les deux cas, les capteurs sont intrinsèquement monochromatiques, la restitution de la couleur des images est réalisée soit par l'adjonction d'un filtre de Bayer - le cas de la plupart des caméras couleur - soit par l'utilisation d'un

capteur par couleur principale (TriCCD, Foveon,...), solution technologique haut de gamme et onéreuse.

Les caméras sont majoritairement conçues pour nous fournir des images de haute qualité. Leur principe de fonctionnement est issu de la photographie, les caractéristiques de l'image étant fonction de la quantité de lumière que laisse passer l'optique (l'ouverture en photographie) et du temps pendant lequel l'imageur sera exposé à ce flux de lumière (la vitesse en photographie aussi appelée temps d'intégration). Lorsque la scène ou la caméra sont en mouvement, le temps d'exposition doit être adapté pour respecter la théorie de Shannon et ne pas avoir de phénomènes de repliement de spectre. Cet effet peut être observé sur certaines caméras utilisant des matrices CMOS avec obturateur de type *rolling shutter* c'est-à-dire que la conversion d'une ligne est effectuée alors que les suivantes sont exposées. Dans le cas de mouvements importants, ce mode de contrôle de l'exposition peut introduire des artefacts comme montré à la différence des obturateurs de type *global shutter* natifs sur les capteurs CCD.

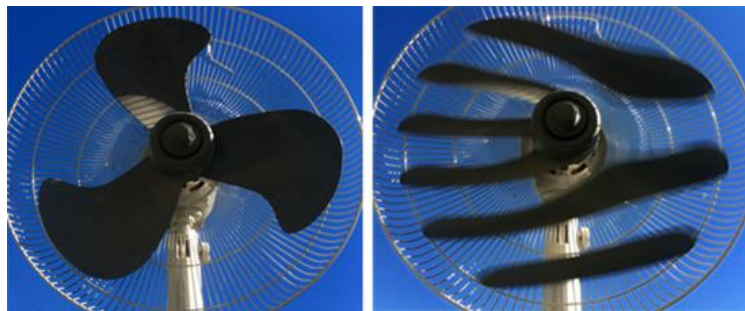


Figure 3 : l'exposition de type *global shutter* et native sur les caméras CCD est mieux adaptée pour l'observation de mouvements rapides (image de gauche) en comparaison à la technique de *rolling shutter* équipant la plupart des caméras CMOS (image de droite).(source : <http://www.digitalbolex.com/>)

La qualité de l'image (contraste, saturation, profondeur de champ...) étant fonction des paramètres d'exposition, de nouveaux types de caméras HDR (*High Dynamic Range*) ou TDI (*Time Delay Integration*) sont apparues qui pallient la limitation ouverture-vitesse en cumulant sur une seule image des vues prises avec des expositions différentes. La perception de scènes à forte dynamique et en présence de peu d'éclairage reste cependant encore un challenge en vision par ordinateur. Cette limitation est essentiellement liée à une acquisition synchrone de tous les pixels pour chaque image. L'autre inconvénient des technologies de caméras actuelles est l'important flot de données qu'elles génèrent, le goulot d'étranglement étant alors du côté de la ressource de calcul dont il faudra doter la machine pour lui permettre de traiter ces données à la volée. Récemment a été introduit un nouveau concept de caméra asynchrone inspirée du fonctionnement de l'œil humain. Ces caméras sont dites à évènement (*event-based*), aussi connues sous l'acronyme DVS (*Dynamic Vision*

System)[17] [48] ; le principe consiste à mesurer pour chaque pixel non pas l'intensité de lumière reçue mais sa variation dans le temps sous forme d'une réponse logarithmique. L'intérêt est qu'elle encode directement une information sur la dynamique de la scène qui peut être vu comme une compression du signal. Cette technologie connaît un certain essor essentiellement due à l'amélioration des rétines photosensibles permettant des rapports signal sur bruit et une dynamique acceptables. Il est aussi à mettre en parallèle avec des avancées récentes en théorie de l'information et l'introduction du concept de *compressed sensing* [54]. Le concept découle de travaux plus fondamentaux sur l'étude des systèmes de vision en particulier chez l'insecte (*compound eye*) et le développement de puces neuromorphiques [21]. Des applications pour le calcul du flot optique [7] ou encore la commande à retour d'effort [9] ont été récemment proposées.

Intérêt d'une perception omnidirectionnelle de l'environnement en vision par ordinateur

Les systèmes de vision les plus utilisés sont basés sur des caméras dites perspectives, conçues pour renvoyer une image analogue à notre propre perception du monde. En première approximation, l'image de la scène est formée en projetant les rayons lumineux sur une surface photosensible (en pratique le plan image) via un centre de projection unique aussi appelé point de vue (*viewpoint*). La distance entre le plan image et le centre de projection, connue sous le nom de distance focale (*focal length*), définit le champ de vue de la caméra (qui sera en pratique aussi conditionné à la taille du plan image). Ce modèle est connu sous le nom de modèle sténopé (*pinhole model*) (cf. Figure 4) en référence au système bien connu en photographie.

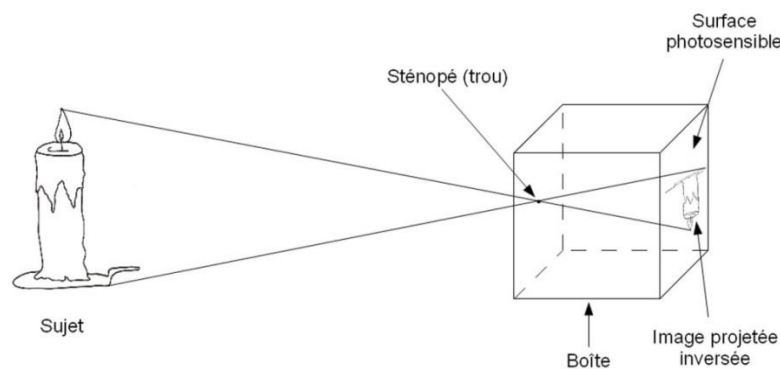


Figure 4 : Capture d'une image par sténopé.

En pratique, la convergence des rayons est faite par une lentille ce qui permet de concentrer l'intensité lumineuse et diminuer le temps d'exposition. Ce modèle est couramment utilisé en vision par ordinateur [61] car il représente en première approximation le principe de formation d'image d'une caméra classique. Un modèle plus réaliste doit a minima prendre en compte les effets de distorsion qu'introduit la lentille. Cependant ce modèle a l'avantage d'amener la notion de caméra

centrale, c'est-à-dire que les rayons lumineux de la scène passeront tous par un point unique appelé centre de projection. Il est à la base des principes géométriques sous-jacents à de nombreux algorithmes en vision par ordinateur pour des systèmes monoculaires et multi caméras [27]. Les travaux menés sur d'autres méthodes de formation de l'image où l'hypothèse de centralité n'est plus respectée ont cependant amené la communauté à introduire d'autres modèles.

Le principal inconvénient des caméras perspectives est le champ de vision réduit. Depuis la fin des années 90, de nombreux travaux en vision par ordinateur ont été menés au sein de la communauté scientifique pour concevoir et évaluer des systèmes de vision permettant une observation sur un champ de vue panoramique voire omnidirectionnel (Figure 5).

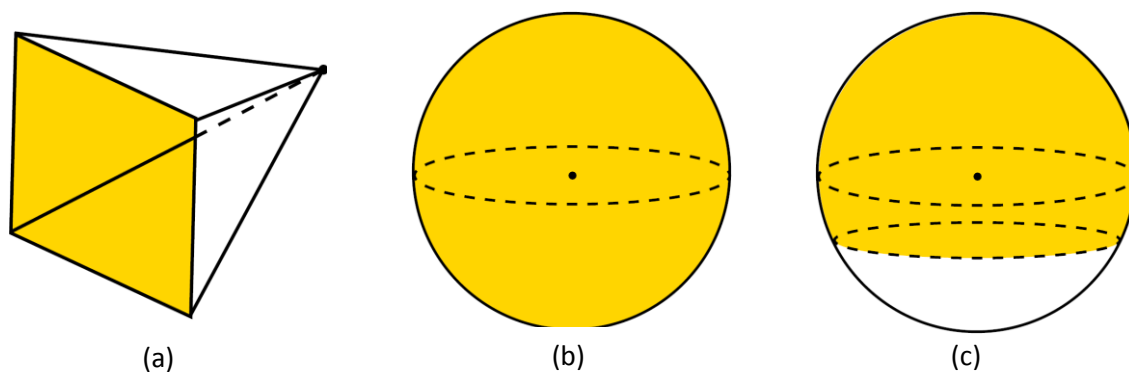


Figure 5 : Représentation des différents champs de vision
(a) directionnel - (b) panoramique - (c) omnidirectionnel

Ces systèmes se sont révélés particulièrement appropriés dans des applications de navigation autonome de véhicule car ils vont fournir une vue de l'environnement indépendante d'une direction et faciliter l'appréciation de la scène par la couche décisionnelle. Il a ainsi été démontré que dans des applications où le capteur de vision est en mouvement et où le traitement consiste à déterminer simultanément le mouvement propre du capteur et la structure tridimensionnelle de la scène (problème bien connu sous le nom de *Structure From Motion* ou *sfm*), l'élargissement du champ de vue à une observation omnidirectionnelle rendait le calcul plus stable que dans le cas classique d'un champ de vision directionnel [20].

Dans les chapitres suivants, nous dressons un rapide état des lieux des technologies qui ont été explorées afin de former une image omnidirectionnelle de la scène. Un recensement plus exhaustif peut être trouvé dans Sturm et al. [44]. Dans un deuxième temps nous présenterons une synthèse des principaux modèles et leur utilisation pour la perception tridimensionnelle de scènes à partir de systèmes mobiles.

II.2 Vision omnidirectionnelle : technologies-clés

Principales solutions technologiques pour l'augmentation du champ de vue

Une méthode intuitive pour former une image panoramique est de capturer plusieurs images à partir d'une caméra perspective voire une caméra linéaire (*pushbroom*) exécutant un mouvement de rotation. Ce type de capteur est appelé caméra cylindrique car la formation de l'image peut être vue comme la projection des points 3D de la scène sur un cylindre [4] [58]. Ce type de capteur forme une image à projection centrale uniquement si l'axe de rotation contient le centre de projection de la caméra. L'avantage de ce type de capteur est qu'il fournit des images de haute résolution. La caméra peut aussi être motorisée sur un axe horizontal. Associé à une fonction zoom, ce type de capteur est communément appelé caméra PTZ (Pan-Tilt-Zoom). Dans tous les cas, la nécessité de mettre en mouvement la caméra implique une mécanique spécifique induisant un encombrement parfois incompatible pour des applications embarquées. De plus son utilisation comme capteur omnidirectionnel nécessite la concaténation de plusieurs images perspectives prises à des instants différents ce qui pose problème si le capteur est lui-même en mouvement dans une scène dynamique. Récemment, Belbachir et al. [6] ont proposé un concept innovant en formant une image panoramique 3D à partir d'une caméra DVS mise en rotation sur un axe vertical.

Une autre approche consiste à multiplier le nombre de points de vue à partir d'un réseau de caméras. L'image omnidirectionnelle est alors formée par mosaïcage des images prises de manière synchrone ou asynchrone suivant la technologie du capteur. L'exemple le plus connu est celui de la Ladybug (Figure 6). Ces dispositifs restent relativement coûteux et encombrants. Le système de vision étant composé de plusieurs imageurs, l'image formée peut être de haute résolution ce qui explique qu'il soit largement utilisé sur des systèmes de cartographie embarquée sur véhicule (*mobile mapping*).



Figure 6 : camera Ladybug (Point Grey) et exemple d'image panoramique obtenue
(Source : Point Gray ; <http://www.glassworksbarcelona.com/>)

Sans arriver à une vision omnidirectionnelle, une autre approche consiste à utiliser un objectif grand-angle dit *fisheye* avec une caméra perspective (Figure 7). Ce type d'optique permet de former une vue un peu plus large qu'un hémisphère. L'utilisation des caméras *fisheye* s'est largement démocratisée avec une baisse importante du coût et une forte miniaturisation de l'optique. Ces caméras sont désormais largement utilisées dans le domaine automobile pour des fonctions de surveillance de l'environnement du conducteur (Figure 7).

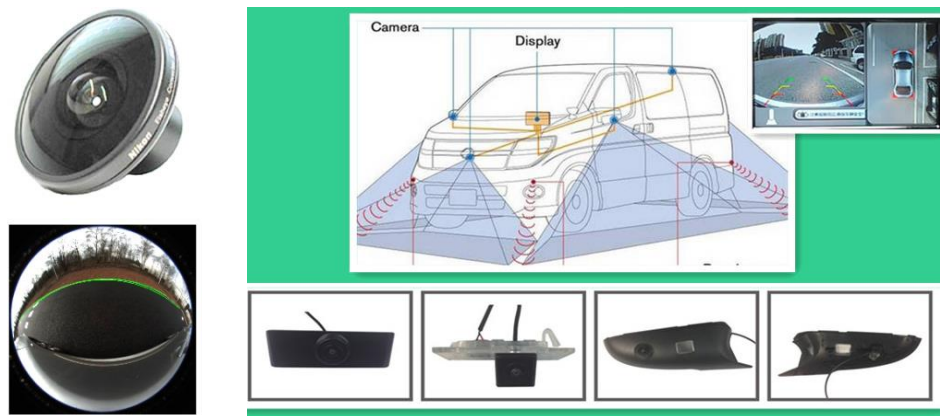


Figure 7 : Un objectif *fisheye* permet l'observation d'un champ de vue un peu plus large qu'un hémisphère (figure de gauche ; la caméra est placée sur le côté latéral d'un véhicule). Ces capteurs commencent à se diffuser largement pour des applications d'aide à la conduite automobile (ici, le système Birdview d'Audi illustré figure de droite). Source : <http://www.carreversecamerasystem.com/>

La diffusion des caméras *fisheye* dans le secteur automobile va s'accélérer avec les nouvelles dispositions de la réglementation américaine en matière de protection des usagers vulnérables. D'ici 2017, les constructeurs automobiles se verront dans l'obligation d'intégrer une caméra dans tous les véhicules circulant aux Etats-Unis afin de visualiser sur l'écran de bord les piétons situés derrière le véhicule².

Capteurs catadioptriques

Un système catadioptrique se définit comme un système combinant au moins un dispositif de réflexion « catoptrique » et un dispositif de réfraction « dioptrique ». Un capteur de vision catadioptrique est alors généralement constitué de trois éléments : une caméra, un objectif et un miroir. L'utilisation d'un miroir pour augmenter le champ de vue est un procédé assez intuitif qui a été assez vite exploré dès l'arrivée de la photographie. En 1911, L. H. Kleinschmidt [34] propose certainement le premier concept de catadioptrique panoramique : une caméra rotative est couplée avec

² « La voiture autonome remplit déjà les caisses de Valeo » - LES ECHOS | LE 07/04/2014

un miroir conique pour l'acquisition de vues topographiques (Figure 8). Plusieurs inventions ont alors suivi ce principe.

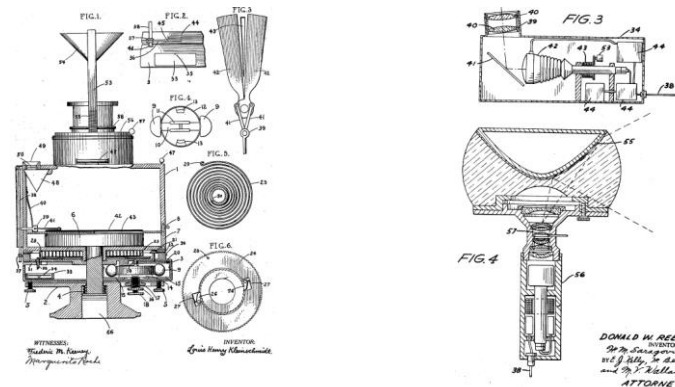


Figure 8 : figure de gauche : dispositif de vision catadioptrique panoramique couplant un miroir conique et une caméra rotative propose par Kleinschmidt en 1911 [34]. Figure de droite : en 1966, Rees mentionne l'obtention d'un point de vue unique par l'association d'un miroir hyperboloïde et d'une caméra [53].

Fin des années 1990, ces systèmes optiques ont suscité un vif intérêt de la communauté scientifique comme alternative aux systèmes de vision perspective avec les avantages que nous avons décrit quand l'observation d'un large champ de vue de la scène est requise. Visant des applications de robotique mobile ou télésurveillance, les systèmes catadioptriques sont alors conçus pour l'acquisition d'une vue panoramique de leur environnement. Sur ce principe de nombreuses configurations ont été étudiées ; une liste exhaustive est hors de portée de ce document. Quelques exemples sont montrés Figure 9.

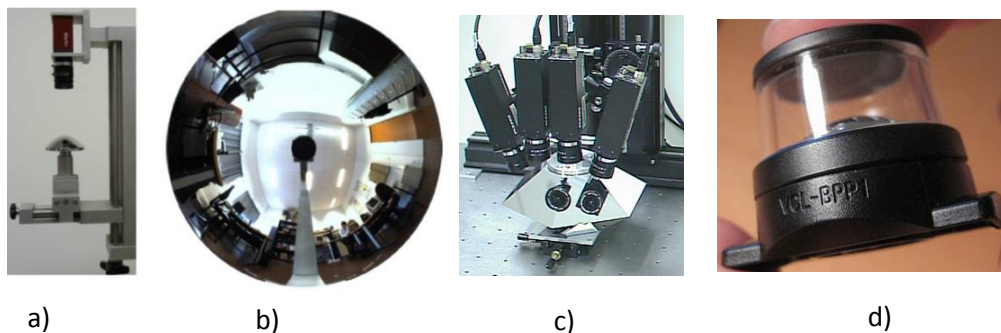


Figure 9 : Exemple de capteurs catadioptriques : a) capteur hyper catadioptrique b) image formée c) utilisation de miroirs plan formant une pyramide d) Panoramic Annual Lens (PAL) développée pour la caméra Bloggie (Sony)

La majorité des articles propose l'utilisation de miroirs présentant une symétrie de révolution sur un des axes pour obtenir un champ de vue à 360° car cette configuration simplifie l'usinage du miroir et son assemblage. Plusieurs travaux ont été proposés dans ce cas pour définir les profils de miroir répondant à des contraintes géométriques telles que l'unicité du point de vue ou l'uniformité de résolution. Il faut souligner qu'un capteur catadioptrique peut être constitué de plusieurs miroirs. Il

est ainsi possible de former une image panoramique en utilisant plusieurs miroirs plans placés sous une forme pyramidale (Figure 9). En plaçant de manière appropriée les caméras, le système peut conserver une projection centrale [67]. L'autre avantage est l'obtention d'images avec une résolution importante et constante. Cependant ces solutions restent complexes à prototyper et à embarquer sur un système mobile. Les capteurs catadioptriques disponibles sur la marché sont parfois assemblés sous la forme d'une lentille panoramique appelée PAL (*Panoramic Annular Lens*), intégrant des dispositifs de réfraction et de réflexion et initialement introduit par Greguss [26]. Un exemple est montré Figure 9.

Les capteurs catadioptriques peuvent eux-mêmes être scindés en deux familles, ceux à projection centrale et ceux à projection non-centrale. Dans le cas d'un capteur central, les rayons lumineux réfléchis par le miroir passent par un point de projection unique, aussi appelé point de vue unique. Les capteurs non-centraux, par définition ne respectant pas cette contrainte, sont plus difficiles à utiliser ; mais ils présentent certaines propriétés qu'il peut être intéressant d'exploiter : points de vue multiples sur une même image, equi-résolution,... Dans un article fondateur, Baker et al. [3] ont montré que seule une catégorie de formes de miroirs respectait cette contrainte de projection appelée contrainte de point de vue unique (*Single View Point constraint* ou SVP). Il faut noter que la propriété du point de vue unique de certains types de capteurs catadioptriques avait déjà été abordée par Rees en 1966 [53] avec l'utilisation d'un système (Figure 8) combinant un miroir hyperboloïde et une caméra placée au foyer de l'hyperboloïde inverse. Sans détailler ici la démonstration, nous en présentons une synthèse Figure 10.

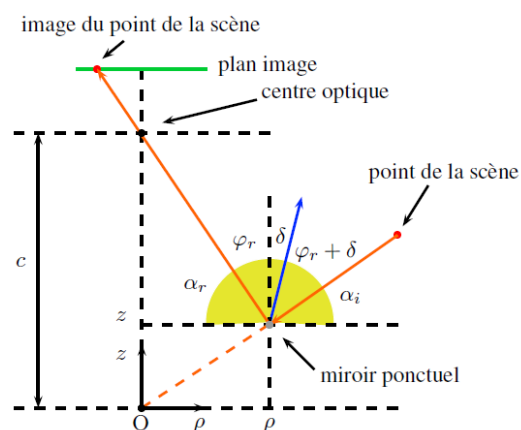


Figure 10 : formulation du problème de respect de la contrainte d'unicité du point de vue dans les cas des capteurs catadioptriques à projection centrale. Les rayons réfléchis sur le miroir doivent se prolonger en un point virtuel appelé point de vue effectif, ici placé à l'origine du repère. L'équation paramétrée du miroir peut être obtenue en considérant la loi de réflexion liant les angles incident α_i et réfléchi α_r et la normale à la surface δ . Le paramètre c représente la distance entre le point de vue effectif et le centre optique.

La contrainte de point de vue peut être exprimée en considérant un miroir ponctuel dans le plan cartésien (O, ρ, z) , en posant $\rho = \sqrt{x^2 + y^2}$, ce qui est représentatif d'un miroir de révolution. L'équation du miroir vérifiant la contrainte est définie par $\rho = f(z)$. On aboutit alors des familles de miroirs paramétrées par les constantes c et k et décrites par les équations :

$$\left(z - \frac{c}{2}\right)^2 - \rho^2 \left(\frac{k}{2} - 1\right) = \frac{c^2}{4} \left(\frac{k-2}{k}\right) \quad (k \geq 2)$$

$$\left(z - \frac{c}{2}\right)^2 + \rho^2 \left(1 + \frac{c^2}{2k}\right) = \left(\frac{2k + c^2}{4}\right) \quad (k > 0)$$

A partir de ces familles d'équations, on peut décrire plusieurs classes de miroirs respectant la contrainte de point de vue unique. Elles sont résumées sur le Tableau 3.

Tableau 3 : formes de miroirs respectant la contrainte de point de vue unique

Forme du miroir	Conditions	Equations du miroir
Plan	$k = 2$ et $c > 0$	$z = \frac{c}{2}$
Conique	$k \geq 2$ et $c = 0$	$z = \sqrt{\frac{k-2}{2}} \rho^2$
Sphérique	$k > 0$ et $c = 0$	$z^2 + \rho^2 = \frac{k}{2}$
Ellipsoïde	$k > 0$ et $c > 0$	$\frac{1}{a^2} \left(z - \frac{c}{2}\right)^2 + \frac{1}{b} \rho^2 = 1$ avec $a = \sqrt{\frac{(2k+c^2)}{4}}$ et $b = \sqrt{\frac{k}{2}}$
Hyperboloïde	$k > 2$ et $c > 0$	$\frac{1}{a^2} \left(z - \frac{c}{2}\right)^2 + \frac{1}{b} \rho^2 = 1$ avec $a = \frac{c}{2} \sqrt{\frac{k-2}{k}}$ et $b = \frac{c}{2} \sqrt{\frac{2}{k}}$
Paraboloïde	$k \rightarrow \infty$ et $\frac{c}{k} = h$	$z = \frac{h^2 + \rho^2}{2h}$

Comme on peut le constater en étudiant ce tableau, deux formes de miroirs sont impossibles à exploiter tout du moins si l'on veut garantir une projection centrale, il s'agit des miroirs pour lesquels le centre optique de la caméra et celui du point de vue effectif doivent être confondus, soit $c = 0$. Pour le miroir sphérique, le respect de la contrainte imposerait de placer la caméra au centre de la sphère. Pour le miroir conique, le centre optique doit être placé rigoureusement au sommet du cône

ce qui est difficile en pratique puisque le centre optique physique de la caméra se trouve derrière l'objectif. Récemment Lin et al. [36] ont montré que la limite théorique imposée par la SVP pouvait être reconsidérée en partant du principe que la lentille d'un objectif perspectif ne focalisait pas les rayons en un point unique mais plutôt dans un volume. Cette reformulation permet de modéliser un tel capteur comme un système de vision à projection centrale. Néanmoins, l'inconvénient majeur des capteurs catadioptriques à miroir conique est la forte disparité de résolution dans l'image avec une quasi-impossibilité de l'exploiter en totalité [36]. Le miroir plan quant à lui ne permet pas une augmentation du champ de vue ; de même pour le miroir ellipsoïdal.

En définitif, les capteurs catadioptriques à projection centrale à retenir sont ceux utilisant un miroir hyperboloïde (capteur hyper-catadioptrique) ou un miroir parabololoïde (capteur para-catadioptrique). Ce dernier requiert cependant l'utilisation d'une caméra orthographique car le centre optique doit être placé à l'infini (cf. Figure 11). Dans le cas du miroir hyperboloïdal, il sera impératif de placer le centre optique de la caméra au niveau du foyer (virtuel) de l'hyperbole ce qui en pratique exige un assemblage mécanique précis et robuste. Dans le cas du capteur para-catadioptrique, cette contrainte technique est levée puisque le centre optique est rejeté à l'infini. Cette simplification explique que ces capteurs ont souvent été privilégiés dans les études antérieures malgré un encombrement important de l'objectif télécentrique.

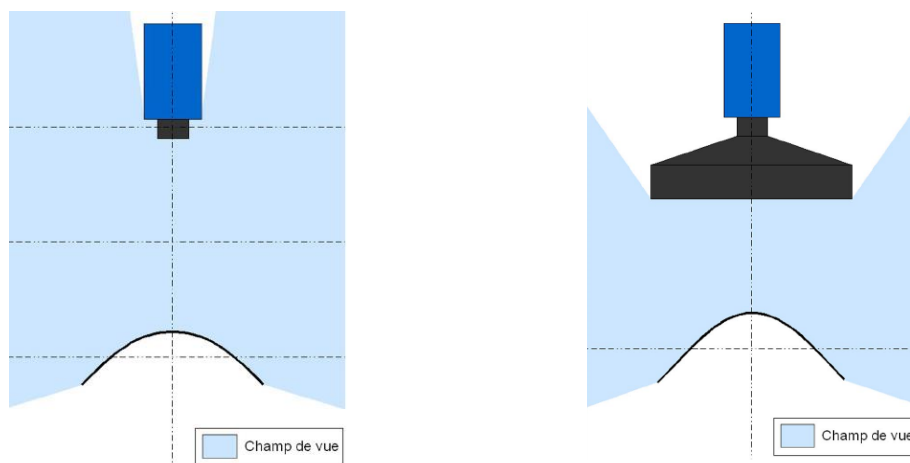


Figure 11 : Les capteurs hyper catadioptriques (figure de gauche) et para catadioptriques (figure de droite) sont les deux classes de capteurs catadioptriques à projection centrale (source : [Th2])

Conclusions

Dans la suite de ce manuscrit, nous présentons plusieurs applications exploitant des capteurs catadioptriques à projection centrale (triangulation, *structure from motion*, détection de visages...). Des travaux de conception et d'évaluation d'un dispositif à plusieurs miroirs permettant d'observer

des champs de vue avec une projection prescrite seront présentés. Avant d'aborder ces travaux, nous parcourons ci-après les principes théoriques et les modèles associés de la vision omnidirectionnelle.

II.3 Vision omnidirectionnelle : principes théoriques et modèles associés

Approche géométrique de la vision par ordinateur

Les traitements utilisés en vision par ordinateur s'appuient pour leur grande majorité sur une approche géométrique en considérant la scène observée comme un ensemble de rayons lumineux émis par une ou plusieurs sources (soleil, éclairages artificiels) et réfléchis par les objets de la scène. Chaque rayon lumineux peut être vu comme un objet théorique facilitant la formalisation des problèmes d'imagerie optique par une simple construction géométrique en se basant sur les lois fondamentales de l'optique géométrique : propagation rectiligne (dans un milieu homogène), indépendance des rayons, changement de direction soumis aux lois de la réfraction et de la réflexion. A partir de ces hypothèses et suivant le dispositif de formation géométrique de l'image, il sera possible d'établir un modèle informatique décrivant comment les rayons issus de points 3D de la scène se projettent sur les pixels de l'image (projection directe ou *forward projection*) ou inversement comment associer à chaque pixel un rayon dans la scène 3D (rétroprojection ou *backward projection*).

L'établissement de tels modèles a fait l'objet de nombreux travaux en vision par ordinateur et a donné lieu à une littérature scientifique abondante. Sturm et al. [44] en ont proposé une taxinomie intéressante. Ils distinguent trois types de modèles :

- Les modèles globaux : la projection ou la rétroprojection seront décrites par une fonction unique pour tout le champ de vue ou le plan image ;
- Les modèles locaux : les projections seront décrites par des fonctions différentes suivant le champ de vue considéré ou les parties du plan image ;
- Les modèles discrets : ils visent à associer à chaque pixel un rayon de la scène ou inversement à chaque rayon un pixel.

L'application d'un modèle à un système de vision en particulier imposera une étape dite de calibration ; dans le cas des modèles globaux et locaux il s'agira d'une estimation de paramètres. Dans le cas de modèles discrets on sera plutôt sur une approche de type table de correspondance (*Lookup Table*).

Du choix d'un modèle et de la méthode de calibration sous-jacente découleront les traitements permettant de retrouver la structure géométrique de la scène (*3D reconstruction*), la position des objets par rapport à la caméra (*absolute pose estimation*) et la compréhension du mouvement de la caméra (*relative pose estimation* ou *motion estimation*).

Vision perspective : modèle sténopé

Le modèle sténopé est largement décrit dans la littérature. Nous le reprenons ici uniquement afin d'introduire les notations utiles pour comprendre l'analogie avec les modèles de projection pour les capteurs omnidirectionnels. La formation d'une image perspective peut être très simplement décrite par une projection centrale d'un point de l'espace sur un plan (Figure 12).

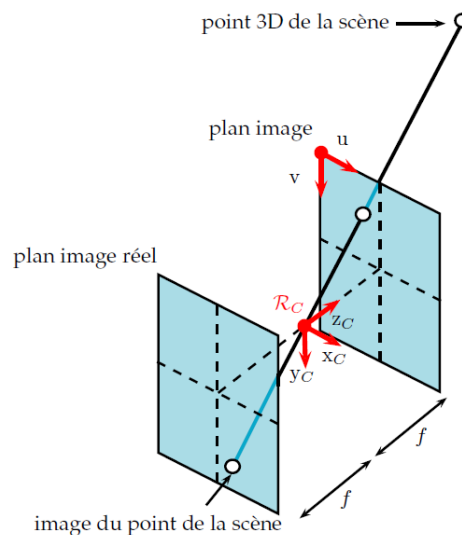


Figure 12 : projection perspective d'un point 3D sur le plan image

Par convention, l'origine du repère de la caméra \mathcal{R}_C est placée au centre de projection C (Figure 12). Il est orienté de sorte que le plan de projection ait pour équation $z = f$ où f représente la distance focale. Le plan peut aussi être représenté dans une position plus représentative de la réalité physique, c'est-à-dire en $z = -f$.

En notant

$M = [X \ Y \ Z \ 1]^T$: le point 3D de la scène dans l'espace projectif P^3

$m = [u \ v \ 1]^T$: le point projeté sur le plan image dans le plan projectif P^2

La projection du point 3D de la scène sur le plan image est décrite par la relation bien connue :

$$m = P \cdot M \quad (1)$$

$$\text{avec } P = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Cette transformation peut être généralisée considérant la transformation rigide $[R \ t]$ où R représente la matrice de rotation 3x3 et t la matrice de translation 3x1 entre un repère quelconque de l'espace 3D et celui de la caméra. Le passage aux coordonnées pixeliques sera obtenu par une transformation affine dans une forme simplifiée en considérant le plan caméra parallèle au plan image et prenant en compte les facteurs d'échelle des pixels (k_u, k_v) et les coordonnées du centre de l'image C exprimées dans le repère image $[u_0 \ v_0]^T$. La projection d'un point 3D sur l'image est finalement exprimée par la relation bien connue :

$$m = K [R \ t] M \quad (2)$$

$$\text{avec } K = \begin{bmatrix} k_u \cdot f & 0 & u_0 \\ 0 & k_v \cdot f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Les paramètres f, k_u, k_v, u_0, v_0 sont les paramètres intrinsèques de la caméra, ils sont regroupés au sein de la matrice K appelée matrice de calibration. La transformation $[R \ t]$ exprime la transformation rigide entre le repère caméra et le repère monde et représente les paramètres extrinsèques de la caméra. L'identification des paramètres intrinsèques requiert une étape dite de calibration dont le principe sera d'utiliser un objet de calibration (grille, damier, trièdre...) fournissant des points de contrôle (l'utilisation d'autres primitives est possible). Les coordonnées 3D et images des points de contrôle étant connues, les paramètres seront identifiés par des méthodes d'optimisation globale. En général, il est nécessaire d'estimer les paramètres $[R \ t]$ qui représentent la pose de la caméra par rapport au repère de l'objet de calibration. Dans certaines applications, on peut se contenter d'utiliser la relation (1) et se placer dans le cas général d'une projection centrale.

Comme cela est mentionné dans [60], la contrainte de point de vue unique peut tout à fait être obtenue par des miroirs ne rentrant pas dans les classes que nous venons de rappeler. Il faut alors

associer une caméra non-centrale avec un miroir dont la courbure devra faire converger les rayons réfléchis vers la caméra en passant par un point unique [60].

Modèles associés aux capteurs catadioptriques à projection centrale

Comme nous l'avons montré précédemment, il est possible de construire un capteur d'image omnidirectionnel respectant (par construction ou en première approximation) une projection centrale. Cette configuration permet une analogie immédiate avec le modèle de projection perspective (Figure 13) ce qui présente de nombreux avantages, en particulier la possibilité d'appliquer la théorie de la géométrie épipolaire.

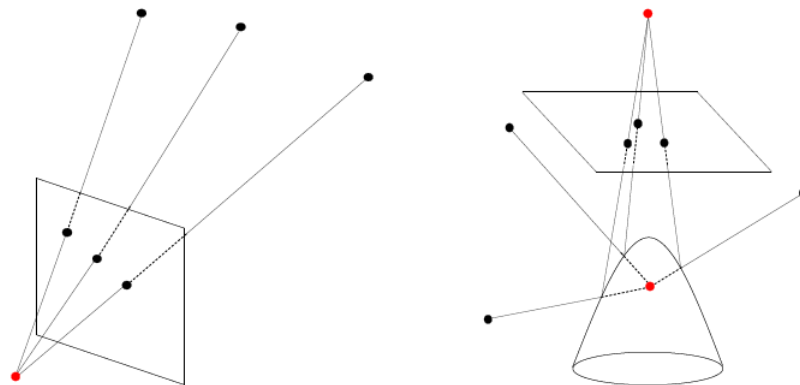


Figure 13 : analogie entre un système de projection centrale perspective (à gauche) et omnidirectionnelle (à droite ; ici le cas d'un capteur hyper catadioptrique). (Source : [55])

De nombreux travaux ont porté sur l'étude des modèles de projection des capteurs catadioptriques à projection centrale : utilisation d'un modèle spécifique à chaque miroir encore appelé modèle ad hoc [24] [64], modèles polynomiaux [56] et plus récemment le modèle unifié dit de la sphère d'équivalence. Successivement introduit par Geyer et al. [23] puis Barreto et al. [5], il a conduit à un modèle généralisé pour les capteurs catadioptriques à projection centrale en se basant sur l'équivalence projective entre une surface quadrique et une sphère. Ces capteurs peuvent être alors modélisés par une projection sur une sphère dite sphère d'équivalence puis sur un plan dont la distance φ par rapport au centre de projection est un paramètre dépendant du type de miroir (Figure 14).

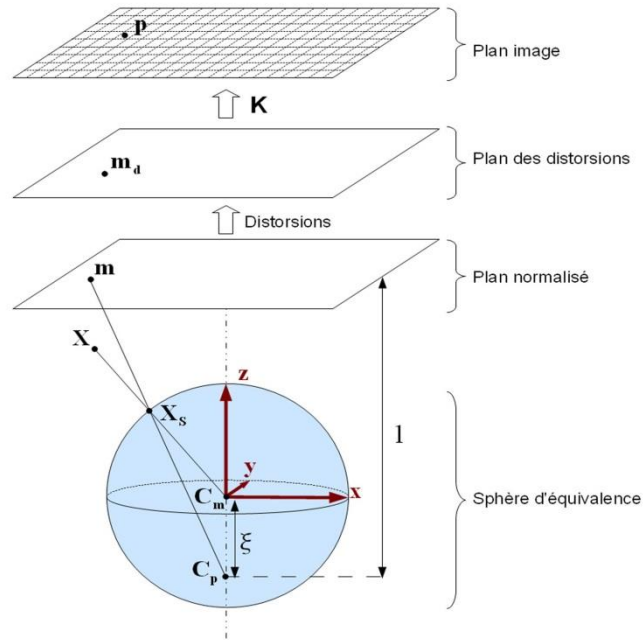


Figure 14 : modèle de la sphère d'équivalence

En notant $M = [X \ Y \ Z]^T$ un point 3D de l'espace et q_N la projection de Q sur le plan unitaire situé à une distance ξ du centre de la sphère, on obtient la relation suivante :

$$m_N = \begin{bmatrix} X \\ Y \\ Z + \xi \sqrt{X^2 + Y^2 + Z^2} \end{bmatrix}^T$$

La projection sur la caméra notée m_p est obtenue en utilisant le modèle sténopé décrit équation (7) :

$$m_p = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varphi - \xi & 0 & 0 \\ 0 & \xi - \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} m_N \quad (4)$$

Christopher Mei [44] a complété ce modèle pour y introduire un modèle de distorsions et en introduisant la notion de focale généralisée $\gamma = (\varphi - \xi)$ afin de réduire le nombre de paramètres à estimer. Le passage des coordonnées du plan normalisé aux coordonnées dans le plan des distorsions peut être décrit par la simple relation :

$$m_D = m_N + \Delta_R(m_N) + \Delta_T(m_N)$$

où Δ_R et Δ_T représentent les distorsions radiales et tangentielles (les formules associées sont données en annexe de ce mémoire). Le point q_p du plan image est alors défini par la relation :

$$m_P = \begin{bmatrix} \gamma & 0 & u_0 \\ 0 & \gamma & v_0 \\ 0 & 0 & 1 \end{bmatrix} m_D$$

Dans cette relation, on ne tient pas compte du changement de signe de γ (en référence à l'équation (7)) qui représente l'effet d'inversion d'image qu'introduit un miroir [51].

Le tableau ci-après résume les liens entre les paramètres du modèle de la sphère d'équivalence et les paramètres physiques des capteurs catadioptriques pour les miroirs paraboloides et hyperboloides.

	ξ	φ	γ
Paraboloïde	1	$1 + h$	$-hf$
Hyperboloïde	$\frac{df}{\sqrt{(d^2 + h^2)}}$	$\frac{d - h}{\sqrt{(d^2 + h^2)}}$	$\frac{-hf}{\sqrt{(d^2 + h^2)}}$

d : distance entre les points focaux

h : demi lactus rectum

Le modèle de la sphère d'équivalence est valable pour toutes les caméras catadioptriques centrales ($\xi = 0$ pour une caméra perspective, $\xi = 1$ pour un capteur para catadioptrique, $\xi < 1$ pour un capteur hyper catadioptrique).

L'étape de calibrage consistera à identifier les paramètres du modèle, ce qui revient à un problème d'optimisation non-linéaire. Les travaux menés dans les thèses de Nicolas Ragot [Th1] et Rémi Bouteau [Th2] ont porté sur l'utilisation de modèles locaux et globaux de capteurs catadioptriques pour leur élaboration ; ils sont présentés dans le chapitre III.

Cas des capteurs catadioptriques à projection non-centrale

Le respect de la contrainte de point de vue unique peut s'avérer difficile à mettre en œuvre dans la pratique car il faut réussir à faire coïncider les centres de projection de la caméra et du miroir. D'autre part, les capteurs hyper et para catadioptriques présentent l'inconvénient d'une résolution non uniforme entraînant des distorsions dans les images qui posent ensuite problème lorsque l'on cherche à y appliquer les traitements habituellement utilisés en vision par ordinateur. En levant la contrainte de la SVP, il est en fait possible de concevoir des systèmes optiques augmentant le champ de vue tout en maintenant constante la résolution [28] ; la notion de résolution constante doit elle-même être précisée en distinguant résolution angulaire, horizontale et verticale (Figure 15).

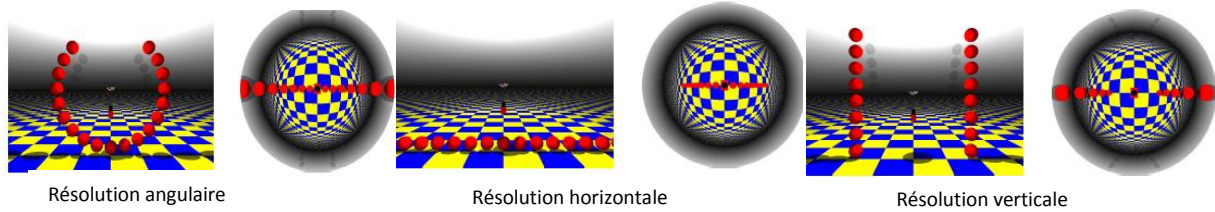


Figure 15 : illustration des non-uniformités de résolution pour un capteur hyper catadioptrique
(Source : [Th3])

Pour conserver une résolution constante, une solution consiste à utiliser plusieurs miroirs dont la surface doit être calculée en fixant une résolution constante dans la projection 3D/2D. Köse et al. [35] ont récemment proposé une famille de capteurs catadioptriques à double miroirs (*folded catadioptric sensors*) large champ et sans distorsion (Figure 16).

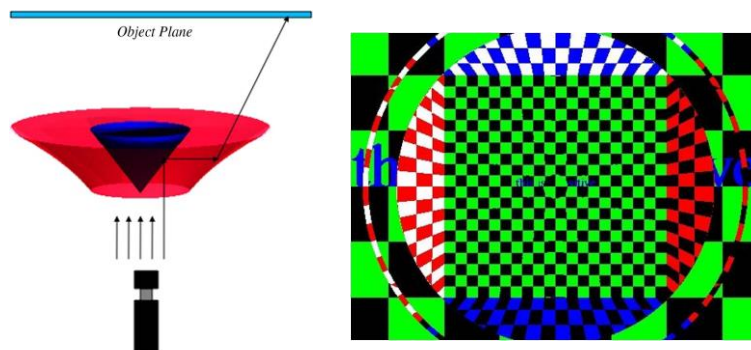


Figure 16 : exemple de capteur catadioptrique conservant une résolution uniforme. Le système est composé d'un miroir conique entouré d'un miroir elliptique ou parabolique. La figure de droite montre la projection orthographique obtenue (ici dans le cas d'un miroir extérieur de forme parabolique).
(Source : [35])

Comme indiqué dans [60], le modèle de projection directe est alors difficile à obtenir et nécessite sauf cas particulier l'utilisation de méthodes d'optimisation non-linéaires. Le modèle de rétroprojection peut être trouvé en considérant pour chaque pixel l'intersection du rayon lumineux qui l'a formé sur la caméra avec la surface du miroir et en utilisant les lois de la réflexion pour en déduire les rayons 3D. Ces points seront abordés plus en détail dans la suite de ce manuscrit dans le chapitre portant sur la conception de miroirs à projection prescrite que nous avons abordé dans le cadre de la thèse de Jean-François Layerle [Th3] et décrit dans le chapitre IV.

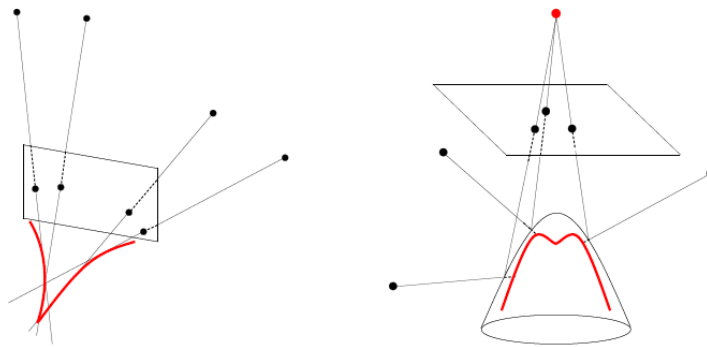


Figure 17 : Caméras non centrales. L'enveloppe des rayons optiques forment une caustique. [55]

La cas de l'utilisation d'un miroir conique est intéressant et continue de susciter un intérêt de la communauté scientifique. Comme expliqué précédemment, la contrainte de SVP ne peut pas être respectée dans le cas d'un miroir conique car cela supposerait de placer le centre optique de la caméra au sommet du cône. Aussi ce miroir est-il généralement utilisé comme un capteur non-central. Ce type de capteur a été un des premier étudié et prototypé par Yagi et al. [71] dès les années 1990 avec une application pour la navigation autonome d'un robot puis étudié dans le cadre du système de navigation autonome SYCLOP [12] [42]. Sur cette base, Lopez et al. [39] ont proposé une extension originale du modèle unifié de la sphère d'équivalence au cas des capteurs catadioptriques à miroir conique. Cette extension consiste à considérer la projection non plus sur une sphère mais sur un tore. Ce modèle est obtenu en s'inspirant du modèle de la sphère dont la projection dans un plan vertical donne un cercle.

II.4 Compréhension de scènes et à la reconstruction tridimensionnelle

Détection de primitives

En vision par ordinateur, de nombreux algorithmes - estimation de pose, stéréovision, estimation du mouvement, etc. - nécessitent la mise en correspondance (*matching*) d'informations issues d'images prises à des échelles et des angles de vues différents. Cette étape requiert une détection de primitives (*features*) de type points, descripteurs locaux, droites... Les images obtenues par des systèmes large champ présentent des distorsions géométriques qui excluent assez souvent l'utilisation des méthodes classiques car elles sont généralement peu robustes aux distorsions que l'on trouve dans ce type d'images. Lorsque le capteur conserve une projection centrale, il est possible de déplier l'image pour former une vue panoramique mais avec un compromis entre conservation des distances (résolution) ou des angles (Figure 18).



Figure 18 : illustration du dépliement d'une image issue d'un capteur catadioptrique à projection centrale. La figure de gauche montre l'image formée par le capteur. Le capteur étant calibré, l'image source peut être rétroprojetée sur une sphère qu'il suffit de rééchantillonner pour obtenir une vue panoramique. L'image en haut à droite représente la projection de la sphère sur un cylindre. Dans tous les cas une phase d'interpolation est nécessaire et l'image formée présente des distorsions par rapport à une image perspective. (Source : Y. Dupuis, rapport interne projet NOBA, 2010)

On peut chercher à contourner ces aberrations en réadaptant les méthodes de traitement qui ont été conçues pour des images plans à une image formée sur une sphère. Parmi les travaux menés sur ce thème, on peut citer ceux de Daniilidis et al. [16] sur l'adaptation d'opérateurs locaux ou encore ceux de Demonceaux et al. sur la notion de voisinage [18] et l'utilisation de la géodésique sur une sphère [19]. D'autres auteurs ont cherché à adapter des algorithmes de détection de primitives de type point [40] ou droite [69]. Dans la thèse Rémi Boutteau [Th2] nous avons proposé une contribution sur la détection de droites sur images omnidirectionnelle pour la reconstruction 3D dont nous présentons les principaux résultats dans le chapitre III.

La déformation de descripteurs locaux par projection sur une sphère pose aussi des difficultés dans le cas de l'extraction de primitives géométriques utilisées pour la classification. Un exemple peut être trouvé dans [57], les auteurs ayant proposé l'utilisation de descripteurs polygonaux pour la classification de scènes routières et la détection de piétons à partir de caméras fisheye. Dans la thèse de Yohan Dupuis [Th7], nous avons exploré cette problématique dans le cadre de la détection de visages, travaux détaillés dans le chapitre V.

Reconstruction 3D par stéréovision

Cas général

L'image formée par une caméra est une représentation du monde 3D réduite à un plan 2D. Retrouver la structure tridimensionnelle de la scène implique l'utilisation de deux vues (au minimum) prises depuis des points de vue différents. De manière analogue à notre système visuel, ces vues peuvent être acquises simultanément, on parle alors de stéréovision. Celle-ci consiste à mettre en correspondance des points (ou d'autres primitives comme des droites) prises par des caméras écartées d'une distance (*baseline*) qui doit être un compromis entre précision de reconstruction et champ de vue commun entre les images. La reconstruction 3D est obtenue par triangulation.

De nombreux travaux ont été menés à la fin des années 1990 sur la mise au point et l'utilisation de bancs de stéréovision avec désormais des implémentations temps réel pour la robotique mobile et le véhicule intelligent [8] [37]. L'information de profondeur de la scène obtenue en stéréovision peut être renforcée par l'utilisation d'une solution de mesure active. Ces systèmes sont aujourd'hui largement diffusés avec des performances de reconstruction temps réel à partir de calculateurs dédiés. A titre d'illustration, la Figure 19 montre deux exemples de systèmes de stéréovision temps réel, l'un pour une application en vidéosurveillance et l'autre pour le véhicule autonome.



Figure 19 : quelques exemples de solutions de stéréovision sur étagère : a) Xovis PT23D : caméra 3D intelligente dédiée au tracking de personnes. b) banc de stéréovision VisLab 3DV-E et utilisation pour la construction en temps réel d'une carte d'occupation 3D depuis un véhicule

La mise en correspondance de primitives et la triangulation permettent de construire une carte de disparité (exemple Figure 19 à gauche). La recherche de points homologues entre images est facilitée par l'utilisation de contraintes en particulier celle introduite par la géométrie relative entre deux vues d'une scène, aussi appelée géométrie épipolaire (Figure 20).

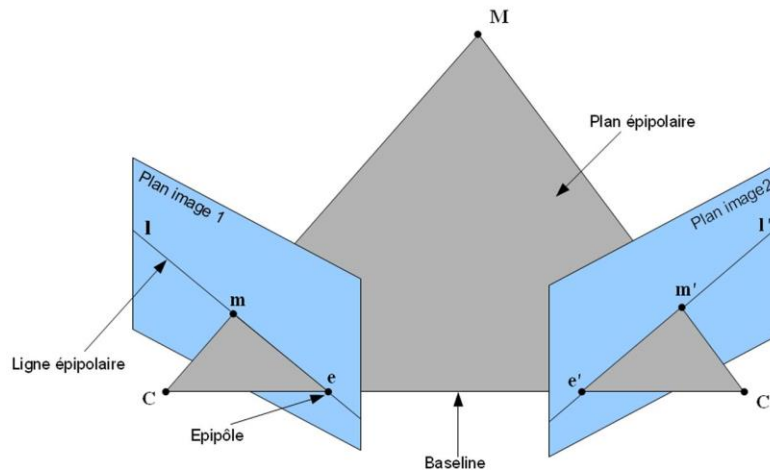


Figure 20 : géométrie d'un système stéréoscopique classique (Source : [Th2])

Cette contrainte est contenue dans deux matrices, la matrice fondamentale (dans le cas d'une caméra non calibrée) et la matrice essentielle (pour une caméra calibrée). En notant $(m ; m')$ et $(x ; x')$ les points projetés dans les deux images exprimés respectivement dans le plan image et dans le repère caméra, on a les relations bien connues :

$$\begin{aligned} x^T E x' &= 0 \\ m^T F m' &= 0 \end{aligned} \quad (5)$$

Avec

$$\begin{aligned} F &= (K)^{-T} E (K')^{-1} \\ E &= [t]_x R \end{aligned} \quad (6)$$

La transformation rigide entre les deux caméras est représentée par une translation $[t]_x$ (sous la forme d'une matrice antisymétrique) et une rotation R . La relation (5) est exprimée à un facteur d'échelle. Une solution linéaire nécessite donc 8 points au minimum [27]. Dans la pratique l'estimation (aux moindres carrés) est faite à partir de plus de points de contrôle. L'avantage d'une méthode d'estimation linéaire est son faible temps de calcul ; l'inconvénient c'est que sans une méthode de rejet des points aberrants, l'estimation n'est pas robuste ce qui pose des problèmes d'implémentation en pratique, car il est généralement impossible de ne pas avoir de faux appariements lors de l'étape de mise en correspondance. Les approches de type RANSAC permettent d'éliminer les points aberrants et garantir la robustesse des méthodes d'estimation des paramètres d'un modèle. En présence de bruit de mesure, il est de toute façon préférable de réduire le nombre d'échantillons nécessaires. La sur paramétrisation diminue donc l'efficacité d'un estimateur robuste. C'est donc toujours un compromis entre temps de calcul et robustesse.

La forme de l'équation (7) implique que $\det(E) = 0$ ce qui réduit encore le degré de liberté des matrices fondamentale et essentielle. En utilisant cette nouvelle contrainte, elles peuvent être exprimées directement à partir de la mise en correspondance de 7 points (3 solutions possibles) en la couplant à une méthode de type RANSAC pour avoir une estimation robuste [27]. Lorsque la caméra est calibrée, l'estimation de la matrice essentielle par résolution linéaire nécessite 6 points au minimum. La méthode la plus connue est celle basée sur une résolution non-linéaire utilisant 5 points [49] basé sur RANSAC ce qui la rend particulièrement robuste aux erreurs de mesure. Une méthode ne nécessitant qu'un point entre images a été récemment proposée mais elle fait une hypothèse forte sur le type de mouvement de la caméra (véhicule non-holonome et se déplaçant dans un plan).

Cas de la vision omnidirectionnelle

Les premiers travaux sur les capteurs catadioptriques ont rapidement permis d'étendre cette notion à la vision omnidirectionnelle avec la notion de géométrie épipolaire pour les capteurs catadioptriques [65], l'introduction de la matrice fondamentale para catadioptrique [23] et de la méthode d'estimation associée dans le cas de capteurs centraux non calibrés en travaillant directement sur la sphère (méthode des 9 points). Plus récemment, López-Nicolás et al. [39] ont introduit la notion de matrice fondamentale conique, nécessitant 20 points de contrôle au minimum dans le cas d'un capteur non central. Lorsque le capteur est calibré, la détermination de la matrice essentielle peut alors être faite avec les mêmes méthodes que pour les modèles de caméras perspectives car elle est en fait indépendante du modèle à partir du moment où les directions des rayons peuvent être exprimés à partir du plan image [60].

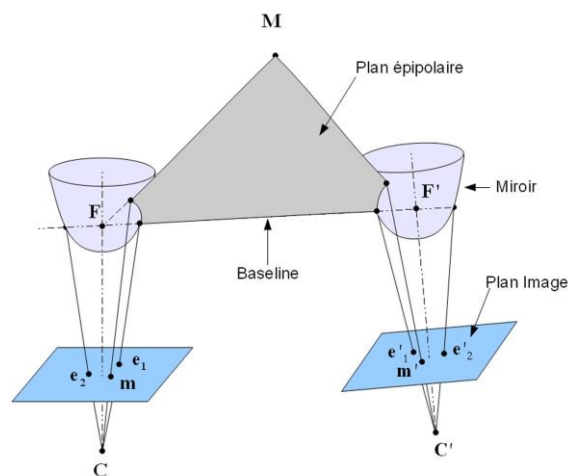


Figure 21 : contrainte épipolaire dans le cas de capteurs catadioptriques. (source : [Th2])

Banc de stéréovision omnidirectionnelle

De manière analogue à un banc de stéréovision classique, une tête stéréoscopique peut être formée de deux capteurs catadioptriques montés en configuration horizontale ou verticale [32] [41] [62]. Ce type de configuration fera l'objet d'une étude approfondie dans le cadre des thèses de Nicolas Ragot [Th1], Rémi Boutteau [Th2] et Romain Rossi [Th4] présentée dans la suite de ce manuscrit.

D'autres travaux ont cherché à former une image composite en projetant à l'aide de miroirs des points de vue différents sur une seule caméra. Ces solutions de stéréovision omnidirectionnelle monoculaire se basent sur l'utilisation de miroirs hyperboliques [10], coniques [33], paraboliques [13], parfois associant des types de miroirs différents [14] ou encore par la formation d'une double image à l'aide d'un miroir convexe et d'une lentille concave [72]. Quelques exemples sont montrés Figure 22. Dans la thèse de Jean-François Layerle [Th3], nous avons travaillé sur la conception d'un capteur à double miroir afin de générer une vue omnidirectionnelle et une vue à champ de vue prescrit que nous avons appliqué au suivi en 3D du visage d'une personne (travaux présentés dans le chapitre IV).

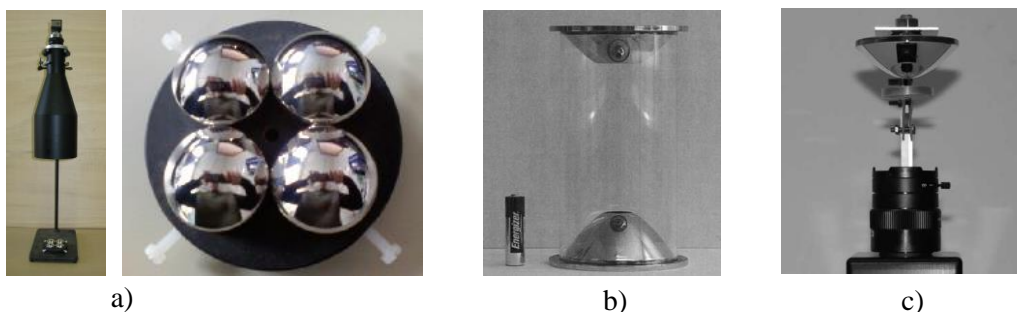


Figure 22 : exemple de système de stéréovision omnidirectionnelle avec une seule camera : a) miroir parabolique [13] b) miroir conique [33] c) utilisation d'un miroir convexe et d'une lentille concave [72].

L'utilisation d'un lidar 2D pour renforcer la reconstruction 3D fourni par un système de stéréovision omnidirectionnelle a aussi été étudiée [45]. Plus récemment Rameau et al. [52] ont proposé un système de stéréovision innovant associant une caméra PTZ et capteur catadioptrique, configuration que nous avons également explorée dans le cadre de la thèse de Yohan Dupuis [Th7].

Reconstruction tridimensionnelle par déplacement du capteur

Principes et transposition à la vision omnidirectionnelle

La structure tridimensionnelle d'une scène peut aussi être retrouvée à partir des images fournies par une caméra en mouvement, méthodes regroupées sous le terme de SLAM (*simultaneous Localization*

and Mapping) et de *Structure from Motion* (sfm). Une abondante littérature existe sur ce sujet. Schématiquement, ce traitement suppose la résolution de problèmes bien connus en vision par ordinateur que sont l'estimation de la pose entre l'objet et la caméra, l'estimation du déplacement propre de la caméra et la triangulation de l'objet à reconstruire entre les vues [60].

L'estimation du mouvement entre deux caméras calibrées (*motion estimation, egomotion* ou odométrie visuelle) revient à estimer la matrice essentielle entre les deux déplacements. Les méthodes évoquées précédemment s'appliquent ici, incluant les capteurs omnidirectionnels centraux dès lors que ceux-ci sont calibrés.

L'estimation de pose entre deux vues et la mise en correspondance de points permet leur triangulation. La résolution de l'un étant liée à l'autre, la solution n'est pas optimale. Ces erreurs doivent être atténuées soit par une étape de filtrage (approche à la base des méthodes de type SLAM car permettant une estimation en ligne et temps réel) ou par une méthode d'optimisation globale appelée ajustement de faisceaux. Celle-ci consiste à recalculer l'ensemble des inconnues (position des points 3D à reconstruire et paramètres extrinsèques de la caméra) par une phase d'optimisation non-linéaire. La fonction de coût utilise généralement l'erreur de reprojection des points 3D sur les images. Jusqu'à peu cette technique était utilisée hors ligne ce qui ne permettait pas une reconstruction 3D temps réel nécessaire à la navigation en robotique et qui explique que ce sont les techniques de type SLAM qui sont utilisées. Cependant des travaux récents ont modifié cette vision des choses avec des approches de type hybride où l'ajustement de faisceaux est appliqué de manière incrémentale. Dans la thèse de Rémi Boutteau, nous avons exploré l'utilisation de cette nouvelle approche au cas d'un banc de stéréovision omnidirectionnelle [Th2] et qui sera abordée plus en détails dans le chapitre III.

Une autre approche est l'estimation de la pose d'une caméra (calibrée) à partir de points 3D dont les coordonnées sont connues. Les méthodes permettant cette résolution sont regroupées sous le nom de problème *perspective-N-Points*. En utilisant une caméra calibrée, 3 points sont nécessaires dans le cas général, 2 dans le cas d'un mouvement dans un plan. Comme indiqué dans [60], les méthodes proposées dans la littérature pour des caméras perspectives exploitent la rétroprojection des points d'intérêt des images sous forme de rayons 3D ; elles sont donc transposables au cas des capteurs omnidirectionnels centraux et non-centraux à partir du moment où ils sont calibrés.

Il faut souligner l'abondance de travaux dans ce domaine. Les travaux les plus récents cherchent désormais à étendre ces méthodes en environnement extérieur, avec des problématiques liées à la

robustesse des descripteurs face à la forte variabilité des scènes, la gestion de cartes métriques ou topologiques importantes ou encore le fait que les scènes présentent d'importantes dynamiques.

II.5 Conclusions

Nous avons introduit dans ce premier chapitre les principales notions sous-jacentes à la vision omnidirectionnelle, elles-mêmes étendues des travaux en vision par ordinateur : capteurs pour la formation d'une image omnidirectionnelle, projection centrale, théorie du point de vue unique et la transposition de la géométrie multi-vues à la vision omnidirectionnelle. C'est dans ce cadre théorique et expérimental que se sont situés les travaux de thèse des doctorants que j'ai co-encadrés depuis 2006. Le chapitre suivant expose les travaux ayant trait à la conception et la mise en œuvre d'un banc de stéréovision omnidirectionnelle à partir de capteurs catadioptriques centraux.

III. PERCEPTION TRIDIMENSIONNELLE DE SCENES PAR VISION OMNIDIRECTIONNELLE ET APPLICATION A LA NAVIGATION AUTONOME DE VEHICULES

Ce chapitre est une synthèse des travaux de recherche des thèses de Rémi Boutteau, Nicolas Ragot et Romain Rossi ([Th1] [Th2] [Th4]) et qui ont toutes portées sur l'utilisation de capteurs catadioptriques à projection centrale pour l'observation omnidirectionnelle et tridimensionnelle d'une scène. Les principales contributions concernent les méthodes de calibrage, l'utilisation de la stéréovision et des méthodes de reconstruction 3D de la scène par déplacement du capteur (*Structure From Motion*).

NB : La production scientifique associée à ces travaux est référencée en [bleu](#) sous la forme [REF] dans le document.

III.1 Contexte

Nos premiers travaux sur les capteurs catadioptriques et la vision omnidirectionnelle ont été initiés en 2004 au sein de deux programmes de recherche : le programme INTERREG III-A OMNIVISS mené en partenariat avec une équipe de l'Université du Kent et un contrat de recherche mené conjointement avec la société d'ingénierie IPSIS pour le compte de la DGA. Dans les deux projets, il s'agissait de concevoir un système de vision omnidirectionnelle permettant la reconstruction tridimensionnelle de l'environnement. Dans le cas du projet OMNIVISS, le système était plutôt destiné à une utilisation fixe (surveillance de sites par exemple). Dans le cas du projet pour la DGA, ce système avait vocation à être embarqué sur un robot mobile pour lui permettre une navigation autonome dans un environnement inconnu. Le système devait fournir une perception tridimensionnelle et métrique de l'environnement aussi bien dans des situations où le mobile était fixe que lorsqu'il devait se déplacer.

Ces deux projets nous ont amenés à travailler sur la calibration et l'utilisation d'un banc de stéréovision omnidirectionnelle. Ils ont donné lieu à trois thèses ([Th1] [Th2] [Th4]) dans lesquelles ont été abordées la méthodologie de conception d'un tel système, les méthodes de calibrage reposant sur des modèles globaux et locaux, la reconstruction 3D par stéréovision, la reconstruction 3D voxelique et son optimisation sur processeur massivement parallèle, la reconstruction 3D avec utilisation du déplacement (*structure from motion*). Ce chapitre présente une synthèse de ces différents travaux.

III.2 Conception et calibrage d'un banc de stéréovision omnidirectionnelle

Méthodologie de conception

Si les capteurs catadioptriques ont été largement étudiés dès la fin des années 1990, leur mise en œuvre peut s'avérer complexe car du choix approprié des différents constituants du banc (miroir, objectif, caméra) et de leur bon assemblage dépendront la performance des traitements logiciels pour la calibration et la reconstruction 3D. Les travaux précurseurs à notre étude avaient essentiellement porté sur des capteurs para catadioptriques car l'utilisation d'une projection orthographique permet de simplifier le respect de la contrainte de point de vue unique, le point de projection étant rejeté à l'infini. Cependant ce type de configuration impose l'utilisation d'une lentille télécentrique dont l'encombrement n'était pas compatible avec notre application. Notre choix s'est donc porté sur un capteur hyper catadioptrique monté sur une structure mécanique nous permettant de régler la distance caméra-objectif. Deux capteurs ont été assemblés en configuration

verticale afin de former un banc de stéréovision (Figure 23). Cette solution facilite en effet l'utilisation de la géométrie épipolaire pour la reconstruction 3D par stéréovision puisque les courbes épipolaires sont des lignes radiales sur l'image omnidirectionnelle.

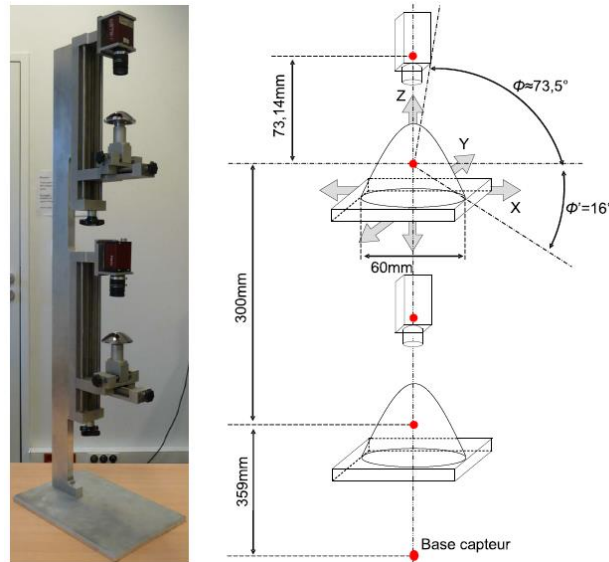


Figure 23 : banc de stéréovision omnidirectionnelle utilisé dans les thèses de [Th1] [Th2] [Th4]

Nous avons formalisé une méthodologie de conception sous la forme d'un rapport interne pour le projet DGA qui permet de définir l'ensemble des composants (courbure du miroir, objectif et focale, taille de l'imageur) en fonction de spécifications techniques (champ de vue de chaque capteur, champ de vue commun, résolution 3D fonction de la *baseline*).

Modèle géométrique et calibrage

Comme présenté précédemment, les approches classiques partent d'un modèle géométrique (modèle ad hoc, modèle de la sphère d'équivalence) représentant le parcours d'un point de l'espace réfléchi par le miroir avant d'être projeté sur le plan image. Ce processus n'est pas linéaire ce qui introduit des distorsions et une résolution non uniforme. Dans le cas des capteurs centraux, l'équation du miroir est une conique et on peut alors décrire la projection d'un point de l'espace sur un pixel de l'image par un modèle analytique (ou modèle global). D'autres approches consistent à représenter la correspondance 3D/2D sous forme locale en s'affranchissant d'un modèle.

Sur ces deux approches – modèle global et modèle local - plusieurs travaux ont été menés au sein de deux thèses ([Th1] [Th2]) qui nous amenés à utiliser deux méthodes de calibrage des capteurs :

- une méthode de calibrage globale basée sur le modèle de la sphère d'équivalence ;
- une méthode de calibrage locale faisant abstraction de tout modèle analytique.

Dans le cadre de la thèse de Rémi Boutteau [Th2], nous avons repris les travaux de Mei et al. [44] afin de développer une toolbox en C interfacée avec la librairie de traitement d'images OpenCV. En effet, l'étape de calibrage étant particulièrement fastidieuse, il nous est apparu nécessaire de disposer d'un outil de calibrage semi-automatique (la détection des points de contrôle ne nécessite que l'entrée manuelle des quatre coins de la mire) et ne nécessitant pas de logiciel tiers (Figure 24). Cet outil a été mis à la disposition de la communauté scientifique et est désormais utilisé dans tous nos développements nécessitant le calibrage d'un capteur à projection centrale.

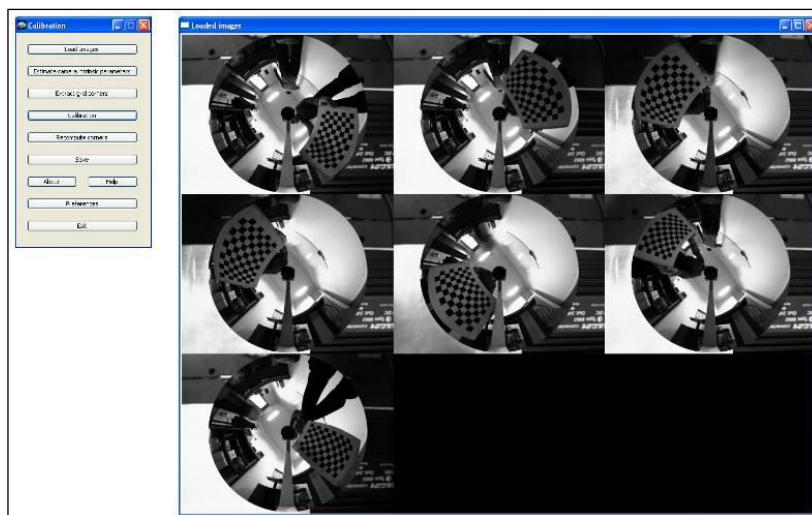


Figure 24 : Outil logiciel développé par Rémi Boutteau dans [Th2] et utilisé dans le laboratoire pour le calibrage de capteurs catadioptriques centraux.

Dans un deuxième temps, les techniques de calibrage faible ont été considérées afin d'établir la mise en correspondance inter caméra, étape indispensable pour la phase de reconstruction 3D. Les méthodes sans connaissance a priori (algorithmes des huit points, RANSAC) et utilisant une mire de calibrage ont été évaluées.

Dans la thèse de Nicolas Ragot [Th1], nous avons opté pour une orientation différente basée sur une approche non-paramétrique. Cela nous a conduits à développer un outil de calibrage basé sur une mire cylindrique.



Figure 25 : mire cylindrique pour la calibration 3D/2D et logiciel de contrôle réalisés dans le cadre de la thèse de Nicolas Ragot [Th1]

Les résultats obtenus ont été comparés à ceux issus d'une approche paramétrique, les précisions obtenues se révélant du même ordre. En fait, l'avantage de la méthode non paramétrique est principalement dans le formalisme décrivant la correspondance 3D/2D. Les points de l'espace sont représentés par leurs coordonnées sphériques (θ, ϕ, Γ) . Le calibrage consiste à utiliser une mire de calibrage cylindrique et interpoler les correspondances 3D/2D en utilisant un maillage par les triangles de Delaunay autour des points de référence de la mire de calibrage (cf. Figure 26).

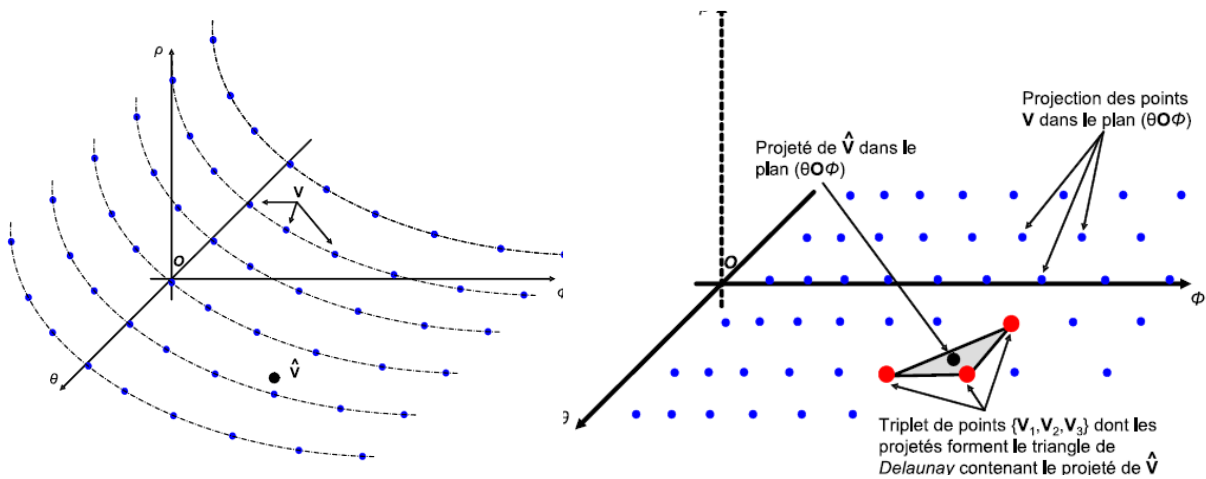


Figure 26 : méthode de calibrage non-paramétrique pour capteur catadioptrique en utilisant un maillage autour des points de calibrage.

Cette technique permet une transcription immédiate sous la forme d'une table de correspondance ou *Look Up Table* (LUT), ce qui est particulièrement bien adapté pour des applications embarquées, comme cela sera évoqué dans un chapitre suivant traitant de la reconstruction 3D par vision omnidirectionnelle. La méthode a été publiée dans [ACTI21] [ACTI22] [COM3].

III.3 Utilisation du banc de stéréovision omnidirectionnelle pour la reconstruction tridimensionnelle de scènes

La reconstruction 3D a pour but la modélisation de la forme des objets et des scènes. La reconstruction 3D est un terme général qui englobe toute la chaîne de traitements, de l'acquisition des données à la visualisation du modèle 3D. Elle se décompose essentiellement en trois étapes :

- L'acquisition des données
- La modélisation à partir des données acquises
- Le rendu ou visualisation de ce que l'on a modélisé

Dans le domaine du transport intelligent, les applications sont nombreuses : détection d'obstacles, cartographie, navigation autonome,... Actuellement, les meilleures performances sont obtenues avec des systèmes dits actifs (LASER, LIDAR, mesure de déphasage, lumière structurée) mais qui sont aussi les plus coûteux. Les systèmes passifs exploitent une ou plusieurs caméras, utilisent des techniques plus ou moins performantes (stéréovision par photogrammétrie, *shape from silhouettes*,...) mais présentent l'avantage d'exploiter des capteurs désormais bas coût (imageurs CCD et CMOS).

Dans les thèses de Rémi Boutteau [Th2] et Romain Rossi [Th4], nous avons exploré une approche originale consistant à utiliser deux capteurs catadioptriques superposés (Figure 27) et embarqué sur un robot mobile afin de percevoir et reconstruire en 3D l'environnement sur 360° et lui permettre une navigation autonome.

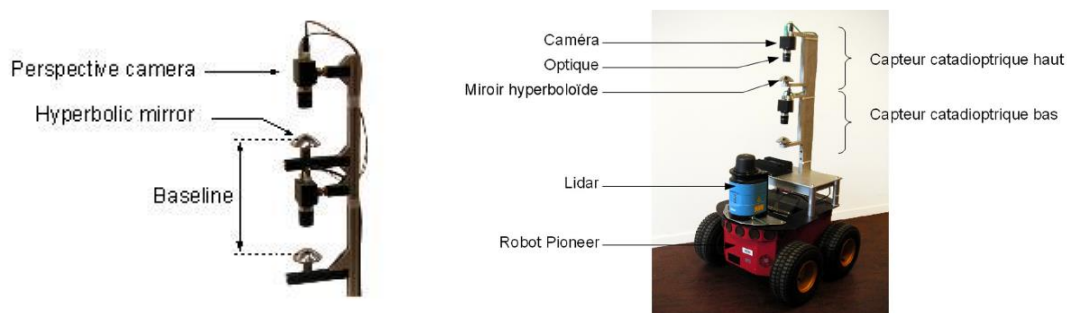


Figure 27 : A gauche : banc stéréoscopique constitué par deux capteurs catadioptriques centraux placés l'un au-dessus de l'autre (figure de gauche) ; à droite : intégration sur un robot mobile pour la navigation autonome.

Ce démonstrateur a été réalisé dans le cadre d'un contrat de recherche avec la DGA ([Th2]) et que nous avons conduit en cotraitance avec la société d'ingénierie logicielle IPSIS. L'enjeu était d'imaginer un système de perception passif pouvant renseigner un véhicule autonome sur son environnement et l'aider dans sa recherche d'une trajectoire. Cet objectif impliquait des contraintes fonctionnelles et

opérationnelles sur le système de perception de l'environnement, principalement liées à des applications de type défense :

- Capteur passif uniquement
- contrainte forte sur le temps de calcul : mise à jour de la reconstruction 3D toutes les 10s
- pas de modèle (cartographie) *a priori*
- contraintes physiques de la maquette : dimension, poids...

La solution retenue était basée sur le système stéréoscopique de vision omnidirectionnelle décrit précédemment. Le démonstrateur qu'il s'agissait de réaliser devait permettre d'évaluer la performance des algorithmes pour différentes missions opérationnelles : dimensionnement d'ouvertures dans des environnements intérieurs, extérieurs, estimation du volume d'un obstacle, estimation du volume libre d'une clairière, reconstruction 3D en environnement dynamique. Dans la thèse de [Th2], nous avons exploré plusieurs techniques : stéréovision, reconstruction 3D de droites et reconstruction par une technique de type *Structure from Motion*.

Stéréovision par appariement de points

La reconstruction par stéréovision passe par deux étapes-clés : l'appariement de points dans les deux images et qui représentent le même point 3D et la reconstruction 3D obtenue par triangulation, possible à partir du moment où la transformation rigide entre les caméras est connue (étape de calibrage). La recherche des bons appariements peut être réduite en considérant la contrainte géométrique qu'induit la géométrie épipolaire. Sur une caméra perspective, le plan épipolaire – lieu contenant le point 3D et les centres de projection – se projette en une droite sur les images (qui peut coïncider avec une ligne de l'image si les caméras sont correctement alignées). Nous nous sommes intéressés à la transposition de cette méthode bien connue en vision conventionnelle aux cas d'une paire de capteurs hyper catadioptriques. Comme illustrées sur la figure ci-après, dans le cas général les courbes épipolaires ne sont plus des droites mais des coniques, correspondant à l'intersection d'un plan avec une sphère, ce qui rend difficile l'utilisation de cette contrainte.



Figure 28 : courbes épipolaires pour une caméra perspective (vue de gauche) et un capteur catadioptrique (vue de droite) (source : [61])

Dans le cas particulier de deux capteurs catadioptriques placés l'un au-dessus de l'autre, ces courbes sont en fait les rayons des cercles de chaque image omnidirectionnelle. Nous avons exploré cette propriété en étudiant aussi les différentes méthodes de corrélation (SAD, ZNCC,...). Un important travail a été mené quant à l'utilisation de l'information de couleur associé au choix d'un modèle de représentation (RGB, H1H2H3). Les tests que nous avons réalisés ont montré que le nombre d'appariements corrects était multiplié par deux entre une méthode de corrélation exploitant la couleur et une méthode en niveaux de gris. Enfin nous avons adapté à la stéréovision omnidirectionnelle les méthodes usuelles de programmation dynamique prenant en compte des contraintes globales (ordre, unicité). Quelques résultats sont montrés sur la figure ci-dessous (Figure 29). Ces travaux ont été publiés dans [\[ACTI1\]](#).

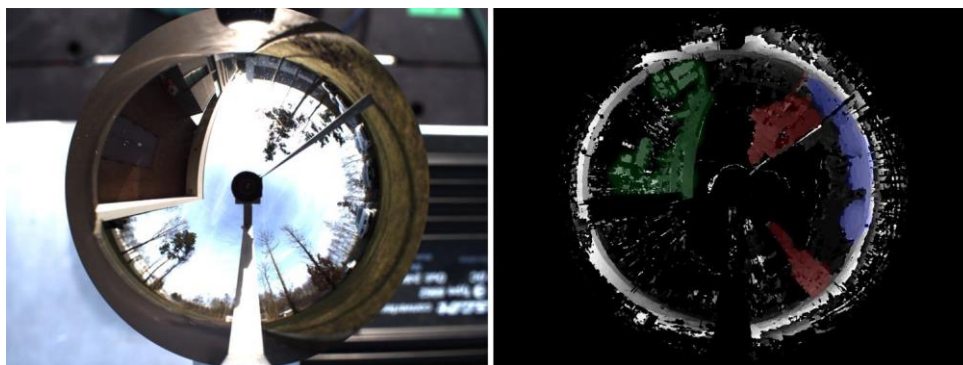


Figure 29 : carte de disparité obtenue par stéréovision en utilisant des contraintes globales sous forme d'une méthode de programmation dynamique

Reconstruction 3D de lignes

La méthode précédente permet une reconstruction dense de l'environnement sous la forme d'un nuage de points. Elle suppose que chaque pixel de l'image contient une information suffisamment discriminante pour pouvoir le corrélérer avec son homologue dans l'autre image. Dans des environnements faiblement texturés (couloirs par exemple), cette mise en correspondance peut

alors échouer. L'autre solution est alors de chercher à détecter des droites afin de retrouver la structure géométrique de la scène.

La détection de droites dans les images issues d'un capteur catadioptrique n'est pas un problème trivial car comme évoqué précédemment celles-ci vont se projeter sous la forme d'une conique. Plusieurs méthodes ont été proposées dans la littérature mais leur utilisation était restreinte à des configurations particulières (droites horizontales et verticales, miroir conique ou parabolique). Dans le cas général d'un capteur central omnidirectionnel, il faut considérer l'intersection du plan avec une sphère, lieu géométrique appelé grand cercle. Des méthodes utilisant la transformée de Hough peuvent alors être utilisées [69].

Nous avons proposé une méthode originale consistant à éliminer les courbes ne correspondant pas à la projection d'une droite 3D par un RANSAC. Pour ce faire, après n'avoir conservé que les pixels formant une chaîne connectée, on tire aléatoirement deux points candidats ce qui permet de calculer le plan devant contenir la droite 3D (plan passant par les deux points candidats et le centre de la sphère) ; on teste ensuite l'appartenance des autres points de la chaîne à ce plan. Les lignes restantes correspondent nécessairement à une droite dans l'espace (Figure 30). Les segments sont ensuite triangulés pour obtenir une reconstruction 3D des droites détectées.

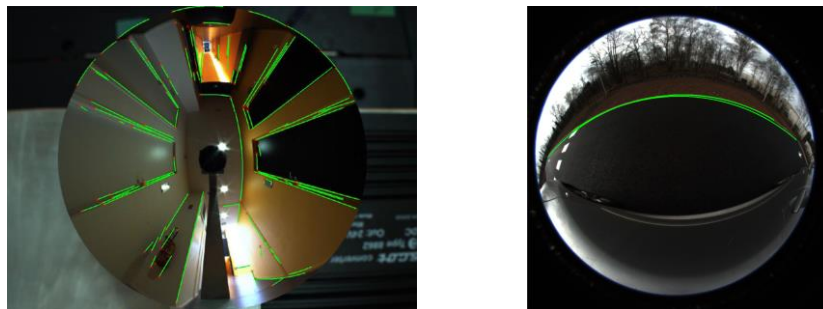


Figure 30 : deux exemples de détection de droites sur images omnidirectionnelles

Nous avons récemment appliqué cette méthode à la détection de marquage horizontal depuis un véhicule en utilisant un capteur fisheye. La méthode a été testée en environnement réel avec différents niveaux de qualité de marquage au sol et des fortes variations de texture dans la scène ; sur 180 images traitées, le taux de bonnes détections était de 87%. Les résultats ont été présentés dans [ACT15].

Reconstruction 3D par stéréovision et approche volumétrique

Dans les thèses de Nicolas Ragot et Romain Rossi [Th1] [Th4], nous avons considéré une autre approche pour la reconstruction 3D qui consiste à discrétiser l'espace en volumes élémentaires (voxels) dont la projection (pixels) peut être calculée sur les images fournies par les deux caméras. Une décision est alors prise sur la contenance du voxel (plein ou vide) via un test dit de *photo consistence* entre les pixels correspondants (Figure 31).

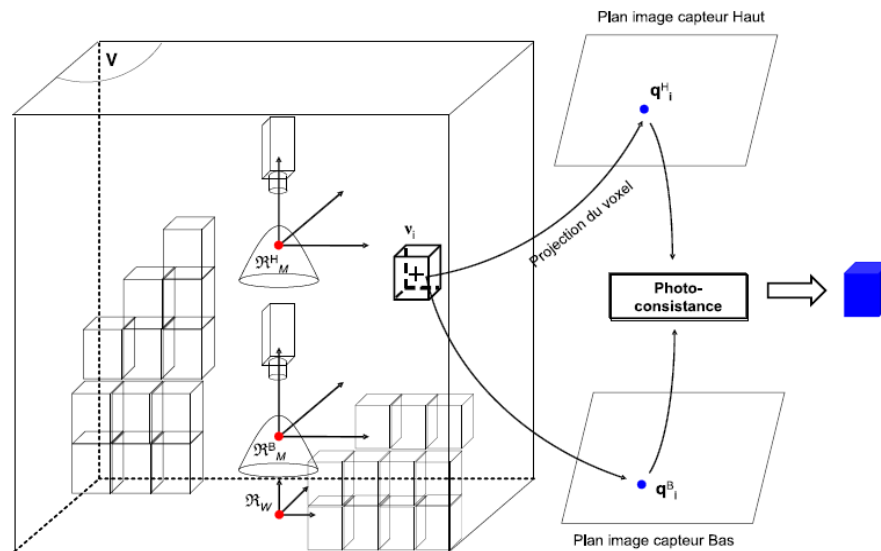


Figure 31 : Reconstruction 3D volumétrique par stéréovision omnidirectionnelle (source : [Th1])

L'approche a été présentée dans [ACTI20] et [ACLN3]. Dans le cadre de [Th4] des travaux ont été menés sur la conception et l'optimisation d'algorithmes de reconstruction 3D voxélique à partir d'images panoramiques. La méthode s'appuie sur une projection des voxels en prenant en compte le rectangle englobant de la projection du voxel considéré sur l'image. Les résultats de la reconstruction 3D d'une scène de synthèse sont montrés sur la figure ci-après.



Figure 32 : exemple de résultats obtenus par reconstruction 3D voxélique par vision omnidirectionnelle : image panoramique de la scène (figure de gauche) – zoom sur une partie de la scène (figure centrale) – reconstruction 3D de l'objet (figure de droite)

Nous avons aussi exploité le fait qu'une méthode de reconstruction voxélique est massivement parallélisable et avons adapté l'algorithme à une architecture GPU, travail précurseur au moment de cette thèse. Avec cette approche il a été possible d'obtenir des reconstructions 3D denses

« temps réel » (quelques secondes de calcul). Une comparaison des temps de calcul entre implémentation sur CPU et GPU est donnée dans la figure ci-dessous. Les résultats ont fait l'objet d'un article dans [ACT15].

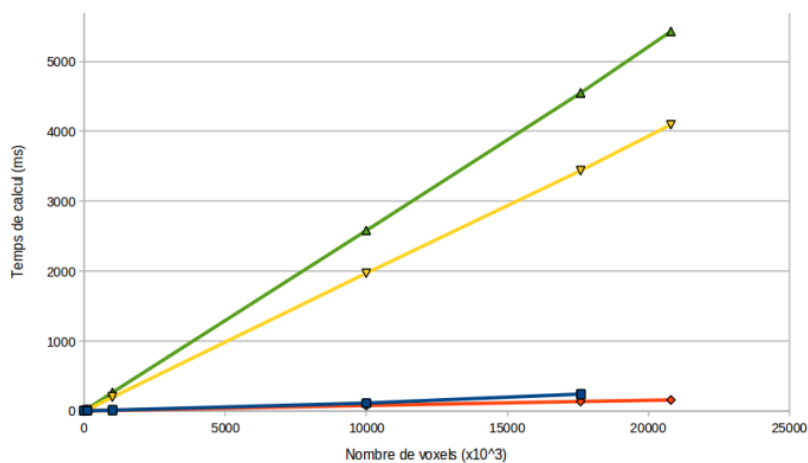


Figure 33 : comparaison des temps de calcul de reconstruction volumétrique entre une implémentation sur CPU (courbes vertes et jaunes) et une implémentation sur GPU (courbes bleues et rouges)

III.4 Structure from Motion par vision omnidirectionnelle

Lorsque le capteur de vision est embarqué sur un mobile (robot, voiture,...), il est possible d'exploiter les vues successives pour obtenir une reconstruction tridimensionnelle de l'environnement. Cette méthode est à la base des techniques de *structure From Motion* (sfm) et de SLAM. Le calcul de pose absolue du système de vision entre deux déplacements se fait en utilisant des points remarquables dans la scène (amers), les erreurs de reconstruction sont corrélées aux erreurs de localisation et inversement ; il est donc nécessaire d'utiliser des méthodes de minimisation de type filtrage prédictif (filtre de Kalman par exemple dans le cas du SLAM) ou par optimisation globale (ajustement de faisceaux). Nous avons appliqué cette dernière méthode à notre dispositif dont la configuration stéréoscopique permet une reconstruction euclidienne de l'environnement ce qui n'est pas le cas dans le cas d'un système monoculaire où la reconstruction se fait à un facteur d'échelle près. La méthode repose sur l'estimation de la matrice essentielle – les capteurs étant calibrés – en utilisant une estimation robuste associant l'algorithme des 8 points décrit précédemment et un RANSAC. En travaillant sur la sphère, le critère de rejet choisi était l'erreur angulaire entre la normale au plan épipolaire, calculée à partir de la première caméra, avec le rayon du point correspondant sur la sphère équivalente de la deuxième caméra. Le choix de la matrice est ensuite renforcé en utilisant ses propriétés lors de sa décomposition en valeurs singulières [27]. L'extraction des paramètres de pose entre deux vues donnant quatre solutions dans le cas de capteurs omnidirectionnels, nous avons proposé une méthode de rejet des projections aberrantes par triangulation. L'optimisation

globale par ajustements de faisceaux a été faite en considérant une fonction de projection stéréoscopique fournissant pour un point 3D les coordonnées images des deux points projetés.

Un exemple de résultats est montré Figure 34.

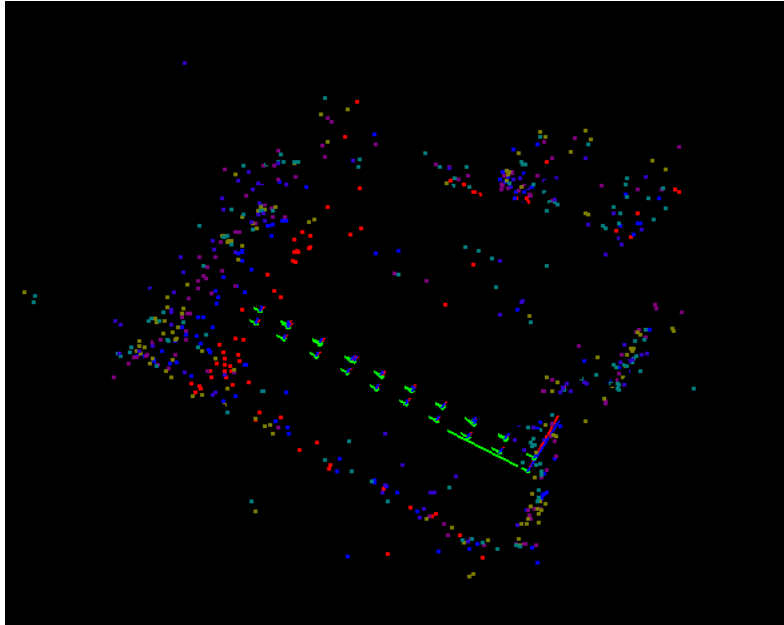


Figure 34 : Ajustement de faisceaux sur images omnidirectionnelles. Le capteur est déplacé par pas de 50 cm. L'ajustement de faisceaux permet le calcul simultané de la pose du capteur (en vert) et la reconstruction 3D des points de référence (en rouge et bleu).

III.5 Conclusions

Ce chapitre a présenté les principales contributions issues des thèses de Rémi Boutteau, Nicolas Ragot et Romain Rossi. Ces travaux ont permis d'aboutir à la production de nouvelles connaissances sur le volet applicatif des capteurs catadioptriques centraux pour la robotique mobile autonome. Ils ont aussi alimenté les livrables de deux projets, en particulier celui mené pour le compte de la DGA et exécuté dans un cadre contractuel. L'étude que nous avons conduite sur les capteurs hyper catadioptriques nous a amenés à nous intéresser aux méthodes de conception des miroirs dans le but de former des images composites d'une scène avec plusieurs miroirs. Ces travaux, faits dans le cadre de la thèse de Jean-François Layerle [Th3] sont présentés dans le chapitre suivant. Dans un deuxième temps, nous nous sommes intéressés à une autre approche permettant la capture de rayons lumineux d'une scène basée sur la fonction plénoptique. Ce travail a fait l'objet de la thèse de Fengchun Dong [Th6] au cours de laquelle nous avons exploré l'apport d'une caméra plénoptique (*lighfield camera*) pour l'estimation de mouvement d'un robot mobile. Les principaux résultats sont également rassemblés dans le chapitre suivant qui résume nos travaux sur les capteurs omnidirectionnels non-centraux.

IV. CAPTEURS NON CENTRAUX : CONCEPTION, MISE EN ŒUVRE ET APPLICATION A L'AIDE A LA CONDUITE ET A LA NAVIGATION AUTONOME

Ce chapitre est une synthèse des travaux de recherche des thèses de Jean-François Layerle [Th3] et Fengchun Dong [Th6] visant à explorer les capacités offertes par des systèmes omnidirectionnels non centraux. Dans un premier temps, ces travaux nous ont amené à évaluer différentes méthodes de design de miroirs, à concevoir et à mettre en œuvre un système de double miroir. Nous nous sommes ensuite intéressés à l'apport des caméras dites plénoptiques dans le domaine de la robotique mobile avec des applications pour l'estimation du mouvement par vision.

La production scientifique associée à ces travaux est référencée en [bleu](#) sous la forme [REF] dans le document.

IV.1 Capteurs catadioptriques non-centraux : design par projection prescrite et calibrage

Contexte

Dans la thèse de Jean-François Layerle [Th3] menée en collaboration avec l'IFSTTAR (Laboratoire de Psychologie de la Conduite LPC) et en partenariat avec le laboratoire MIS de l'Université de Picardie Jules Verne, nous avons imaginé un dispositif permettant une observation simultanée du visage du conducteur et de la scène de conduite. Il s'agissait de prolonger les travaux d'autres équipes [30] sur l'utilisation d'un capteur de vision omnidirectionnelle pour l'observation panoramique depuis l'habitacle d'un véhicule et le monitoring du conducteur (détection de l'hypovigilance par exemple). La trop faible résolution et les distorsions qu'engendrent un capteur catadioptrique analogue à celui étudié dans le chapitre précédent étaient cependant rédhibitoires pour mettre en œuvre les traitements d'images nécessaires au monitoring du conducteur en particulier le suivi de son regard (*gaze tracking*). Nous avons donc orienté nos travaux sur la conception d'une nouvelle configuration catadioptrique basée sur deux formes distinctes de miroirs : l'une pour la capture d'une vue panoramique de la scène de conduite et l'autre pour une vue du visage du conducteur avec une résolution suffisante pour la détection et le suivi de ses yeux.

Comparaison des méthodes de design de miroirs

Afin de trouver le meilleur design de miroirs, nous nous sommes intéressés aux méthodes permettant de décrire l'équation de la surface du miroir sous une forme analytique en imposant une projection entre le plan objet et le plan image (Figure 35). En considérant une réflexion spéculaire sur le miroir et pour laquelle les rayons lumineux incidents et réfléchis sont coplanaires, on peut montrer que l'équation du miroir vérifie l'équation suivante :

$$\begin{cases} f_y \frac{\lambda}{\alpha} x + \left(\cos\theta + \frac{\lambda}{\alpha} y \sin\theta \right) f_x = 0 \\ 2f_y \left(-\sin\theta + \frac{\lambda}{\alpha} y \cos\theta \right) - \left(\cos\theta + \frac{\lambda}{\alpha} y \sin\theta \right) (1 - f_x^2 - f_y^2) = 0 \end{cases} \quad (7)$$

où f représente l'équation de la surface du miroir et f_x f_y les dérivées partielles.

Il est admis que la forme polynomiale :

$$f(x, y) = \sum_{i=1}^n c_i \left(-\frac{\lambda}{2} x^2 + \alpha y \cos\theta - \frac{\lambda}{2} y^2 \sin\theta \right)^i$$

est une solution de l'équation (7). Connaissant la géométrie de la projection prescrite, les coefficients c_i peuvent alors être déterminés par une méthode d'optimisation non linéaire.

Nous avons comparé deux méthodes analytiques de conception de miroirs aboutissant à une projection prescrite (Figure 35) : la méthode de la surface fixe [28] et la méthode des champs de vecteurs [29]. Plusieurs configurations de miroirs courbes ont été étudiées et en jouant sur des combinaisons concaves/convexes entre les axes du miroir. Nous avons introduit plusieurs critères de comparaison entre les méthodes : distorsions, rapport entre surface projetée et surface totale, résolution. Les deux méthodes ont donné des résultats similaires ; la méthode de la surface fixe s'est avérée être plus simple à mettre en œuvre mais la méthode des champs de vecteur offre plus de possibilités en terme de choix de design de miroirs. Concernant les distorsions, l'étude a permis d'établir qu'un polynôme $f(x,y)$ d'ordre 4 était le meilleur compromis dans tous les cas entre niveaux de distorsions et complexité du modèle. D'autre part, il a été constaté une influence importante de l'angle d'élévation Φ_s (Figure 35) sur les distorsions engendrées par la courbure du miroir adapté à un champ de vue placé à cette hauteur.

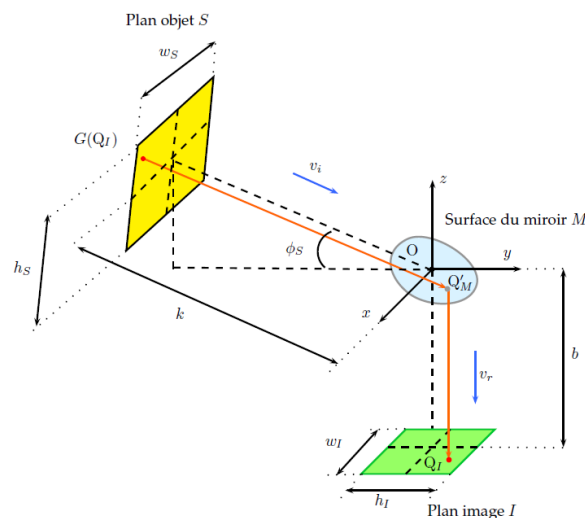


Figure 35 : Projection prescrite entre le plan objet et le plan image (ici dans le cas d'une projection orthographique entre le miroir et le plan image)

Enfin les résultats ont été comparés avec une autre approche utilisée pour le design des miroirs et passant par un maillage en éléments finis [66]. La comparaison a montré la supériorité des méthodes analytiques ; de plus, les méthodes par maillage sont difficiles à modéliser car elles ne fournissent pas d'équation analytique de la surface globale du miroir ce qui complique l'étape de calibrage.

Conception d'un miroir double pour la perception simultanée scène / conducteur

Dans un second temps, une étude a été entreprise visant à déterminer la surface la mieux adaptée à la capture de la scène routière en vue panoramique tout en tenant compte des contraintes de l'environnement (position dans le véhicule, optiques des capteurs, gêne pour le conducteur). La définition de la surface dédiée à l'observation du visage du conducteur a été faite à partir de la méthode de la surface fixe [28].



Figure 36 : à gauche : utilisation d'une méthode de projection prescrite pour le design d'un miroir double ; au centre : deux prototypes de miroirs doubles ; à droite : système catadioptrique complet.

Cette méthode a été appliquée au problème d'observation de l'activité de conduite automobile. Une étape de conception et simulation a permis d'aboutir aux spécifications d'un double miroir (Figure 36) renvoyant sur une image composite formée par une vue panoramique dans le véhicule (intérieur, scène de conduite, angles morts) et une vue du visage du conducteur présentant une résolution plus élevée (Figure 37).

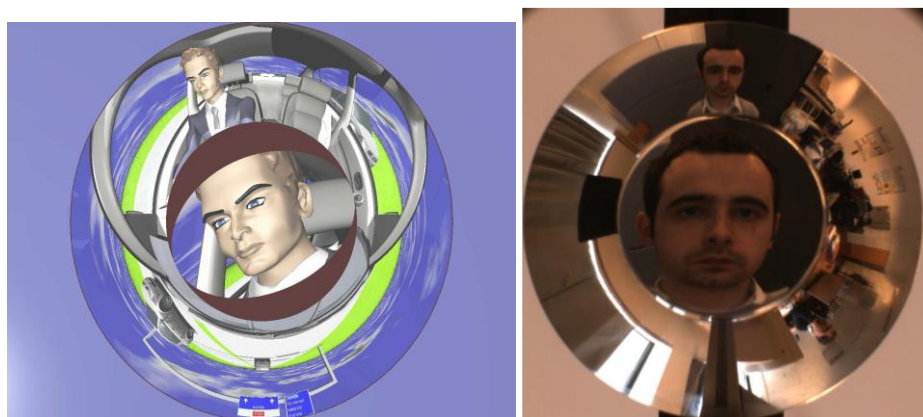


Figure 37 : à gauche : simulation du dispositif catadioptrique pour le monitoring d'activités de conduite ; à droite : exemple d'image obtenue avec notre prototype.

Méthode de calibrage

Les deux miroirs renvoyant une vue de la scène avec un point de vue différent, nous avons envisagé de travailler en stéréovision pour la localisation 3D du visage du conducteur et le calcul de la pose de la tête, étape préalable pour retrouver la direction du regard dans un référentiel 3D. Pour ce faire, il était nécessaire de mettre au point une méthode de calibrage unifiant les projections omnidirectionnelles et à champ de vue prescrit.

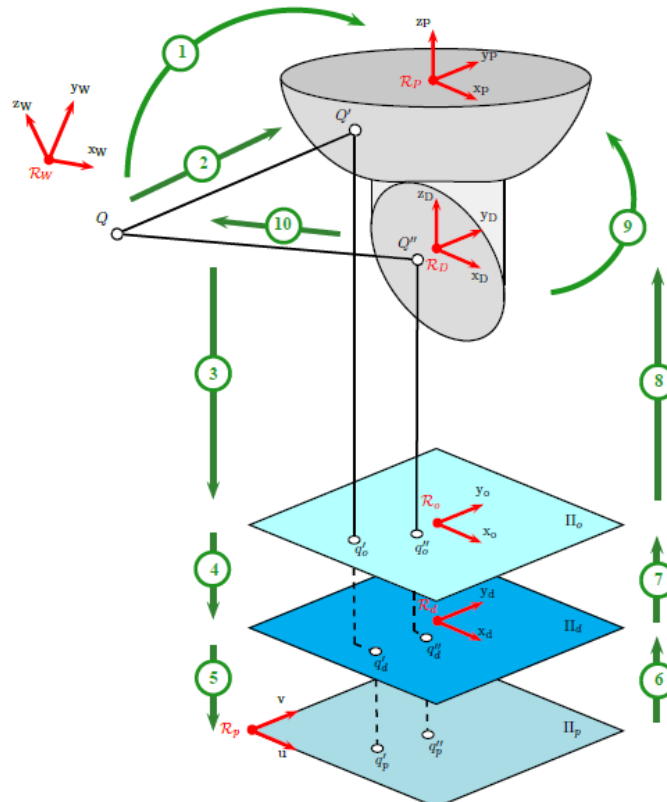


Figure 38 : Méthode de calibrage du capteur para catadioptrique à double miroir. Dans un premier temps le capteur central est calibré par la méthode de mise en correspondance d'une mire 3D en utilisant le modèle ad hoc du miroir (étapes 1 à 5). Les pixels de l'image sont ensuite rétro projetés sur les rayons du miroir à projection prescrite en utilisant l'équation analytique de sa surface (étapes 6 à 10).

Triangulation

Un algorithme a été développé pour déterminer les intersections des projections de points d'intérêts observés sur les deux vues du conducteur (panoramique et vue dédiée). Une évaluation de la précision de reconstruction d'une grille de points assimilés à ces points d'intérêt (Figure 39) a été proposée afin de valider la conception du système.

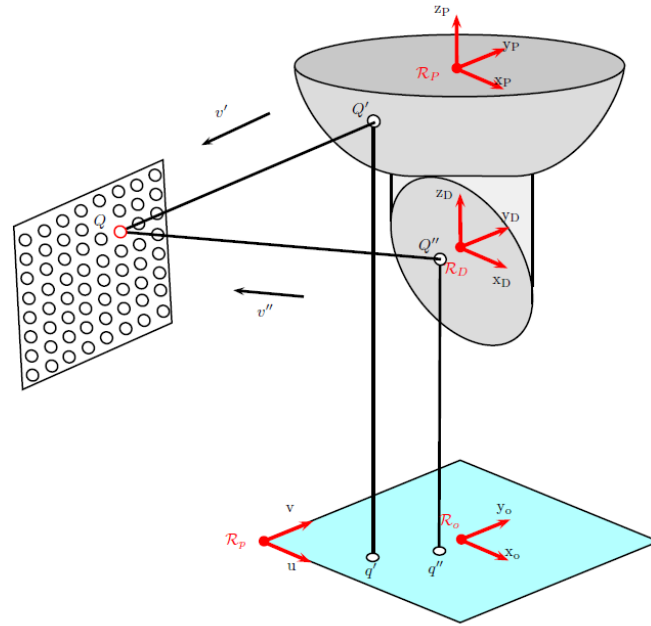


Figure 39 : Projection d'une grille assimilée à des points d'intérêts du visage sur les 2 surfaces du miroir.

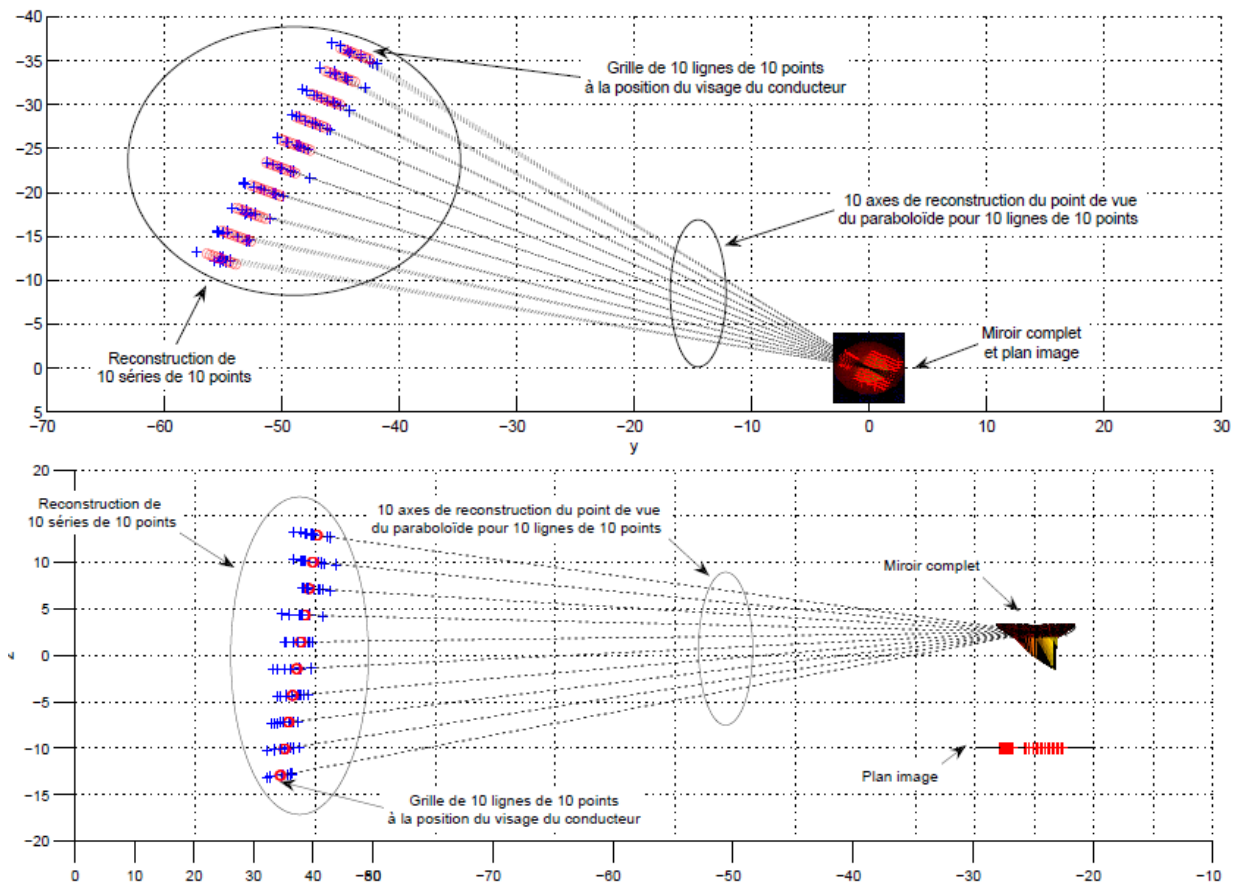


Figure 40 : Reconstruction 3D en cm d'une grille de 100 points avec un bruit sur le plan image

Ces travaux ont fait l'objet de plusieurs articles dans des conférences nationales et internationales [ACTI17] [ACTI18] [ACTI19] [ACTN5].

IV.2 Odométrie visuelle par une caméra plénoptique

La vision omnidirectionnelle, en augmentant le champ de vue d'un mobile, va permettre une perception dense de son environnement. Lorsque le capteur se déplace, il est possible d'utiliser les vues successives pour faire une reconstruction dense comme cela a été étudié dans [Th2] et [Th4]. Ces données seront ensuite passées aux couches décisionnelles du véhicule pour réaliser des tâches de détection et d'évitement d'obstacles, de planification et d'exécution de trajectoires... et plus globalement permettre une navigation autonome du véhicule. Un des points critiques est alors la mesure du mouvement propre (*egomotion*) du véhicule, aussi appelé odométrie visuelle.

L'observation de l'environnement par une caméra peut être décrite comme un processus de capture des rayons lumineux générés par les sources de lumière (soleil, éclairage) et dont la propagation est modifiée par les objets de la scène. La représentation d'une scène tridimensionnelle par un ensemble de rayons lumineux, aussi appelé champs de lumière (*lightfields*), est particulièrement bien adaptée pour la synthèse d'images [38] et est à la base de bon nombre de logiciels de rendu 3D (technique de lancer de rayons). On peut considérer qu'inversement un système de vision aura pour but de capturer de tout ou une partie du champ de lumière et que les traitements en vision pour ordinateur auront pour tâche d'extrapoler, extraire et interpréter les informations mesurées.

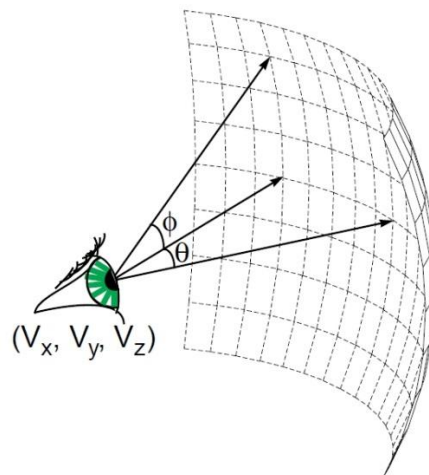


Figure 41 : Illustration de la fonction plénoptique. Un système de vision placé en un point de l'espace (V_x, V_y, V_z) peut y observer l'ensemble des rayons lumineux en faisant varier son orientation (θ, ϕ) .

(source : McMillan, M. & Bishop, G., 1995)

Les champs de lumière d'une scène peuvent être représentés par une fonction à 7 dimensions appelée fonction plénoptique³ en référence à la notion de vision, formalisée dans Adelson et al. [1]. Le principe de cette représentation est montré Figure 41.

Si l'observation se fait par un système monochromatique, la longueur d'onde λ est constante. Enfin, en considérant une propagation dans un milieu isotrope, les rayons de lumière peuvent être représentés par des droites orientées dans l'espace (*brightness invariance*). La fonction plénoptique est alors de dimension 5 dans le cas de scènes dynamiques et de dimension 4 dans le cas de scènes statiques car elle est alors indépendante du temps [11] [25]. Cette hypothèse n'est cependant plus valable si le milieu est hétérogène, par exemple en cas de pluie, brouillard ou encore si l'observation se fait à travers une vitre. Dans le cas d'une scène statique, Levoy et al. [38] ont proposé une paramétrisation de la fonction plénoptique 4D par deux plans fronto-parallèles. Cette paramétrisation est montrée Figure 42.

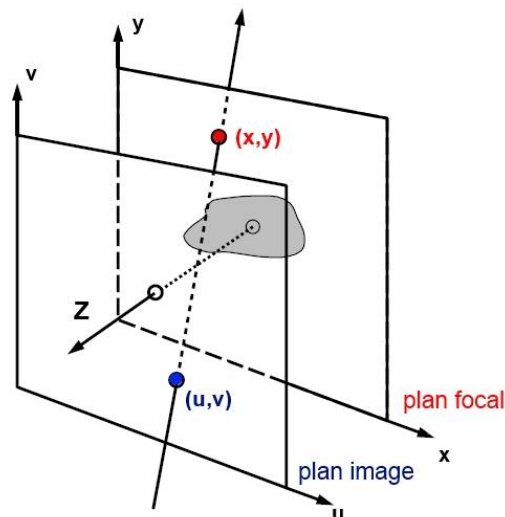


Figure 42 : Paramétrisation de la fonction plénoptique par deux plans parallèles.

(source : Gortler et al. , 1996)

On considère que la scène est englobée par une surface convexe qui pour simplifier est matérialisée par un cube et on s'intéresse aux rayons traversant une des paires de faces parallèles du cube. Cette approche est sans perte de généralité puisqu'il suffit de la reproduire pour les deux autres paires de face. Les rayons ayant une direction constante dans l'espace libre contenu entre les deux faces, ils sont entièrement définis par leur intersection avec les deux plans. Les plans sont fixés arbitrairement orthogonaux à l'axe Z du repère monde. Un des plans représentant le plan focal est paramétré par

³ du latin *plenus* signifiant complet et optique

les coordonnées x,y ; l'autre plan correspond au plan image, il est paramétré par les coordonnées (u,v) définis dans un repère local relatif à (x,y) . Les deux plans sont séparés d'une distance focale f .

Neumann [47] utilise cette paramétrisation pour étudier les systèmes de vision les mieux adaptés à la perception de scènes en mouvement (Figure 43). A partir d'une analyse algébrique des sous-espaces de la paramétrisation 4D de la fonction plénoptique, il a proposé une intéressante hiérarchie des systèmes de vision. Si la mesure de la fonction plénoptique est réduite au sous-espace (x,y) , nous sommes dans le cas d'une caméra perspective capturant un sous-ensemble des rayons lumineux de la scène passant par un point de vue unique (x,y) . Dans le cas où la direction de vue est fixée par (u,v) , les rayons qui forment l'image sont parallèles, la vue est orthographique ; lorsque le plan focal est placé à l'infini, les rayons arrivent orthogonalement au plan image (cas des objectifs orthographiques). La conception d'un système de vision peut donc être abordée comme la définition du sous-espace de la fonction plénoptique nécessaire à la réalisation d'un traitement donné.

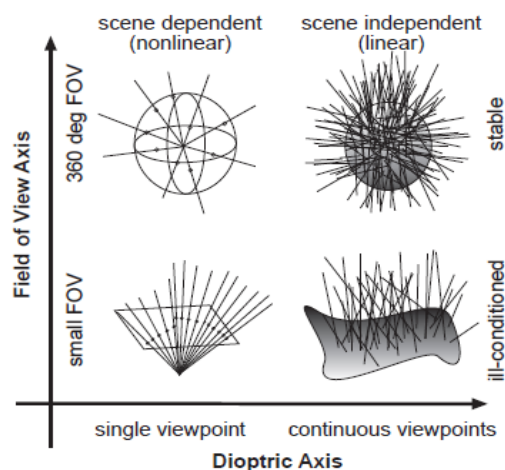


Figure 43 : Classification des caméras dans un problème de Structure From Motion. Les caméras omnidirectionnelles non-centrales (en haut à droite) permettent en théorie une résolution linéaire.

Dans le cadre de la thèse de Fengchun Dong [Th6] nous avons travaillé sur une architecture de vision dite « polydioptrique » en reprenant les travaux théoriques de Neumann sur la modélisation de la perception d'une scène par la fonction plénoptique. Cette fonction définit l'ensemble des rayons lumineux dans l'espace et les caméras plénoptiques sont adaptées à son analyse. Ces caméras sont capables de fournir simultanément un grand nombre de points de vue, permettant de lever ainsi de nombreuses ambiguïtés que l'on rencontre usuellement lors de l'utilisation d'une caméra. Cette approche permet de garantir plus aisément l'existence d'une solution au problème. Le système complet a été testé à l'UPMC sur un robot Pioneer de l'ISIR (Figure 44). La Figure 45 montre les performances de l'estimation en fonction des paramètres du capteur et du mouvement.

Ces travaux ont fait l'objet d'un article dans *International Journal of Robotics Research* [ACL1].

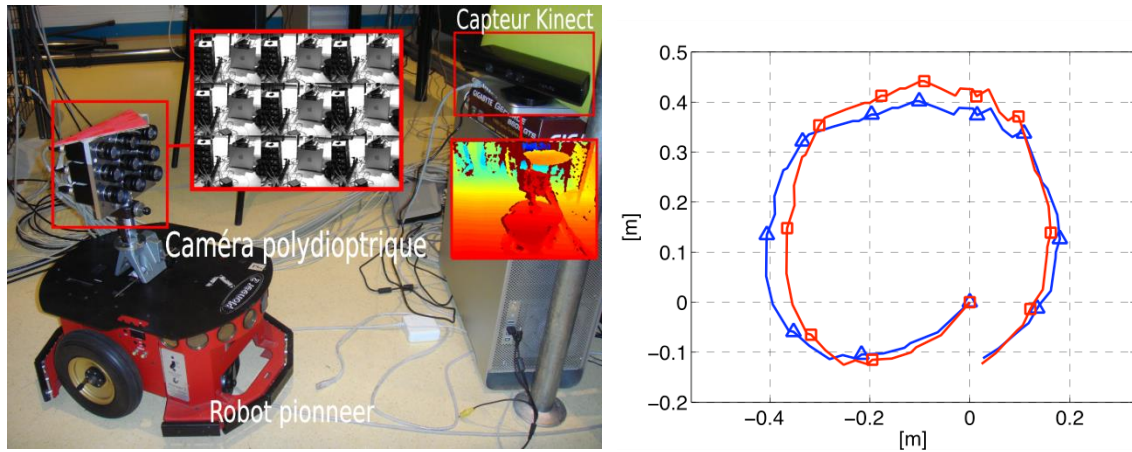


Figure 44 : à gauche : prototype de la caméra plénoptique (grille de 3 × 3 caméras) et capteur Kinect® qui fournit la vérité terrain ; à droite : trajectoire de référence (bleue) et trajectoire estimée (rouge).

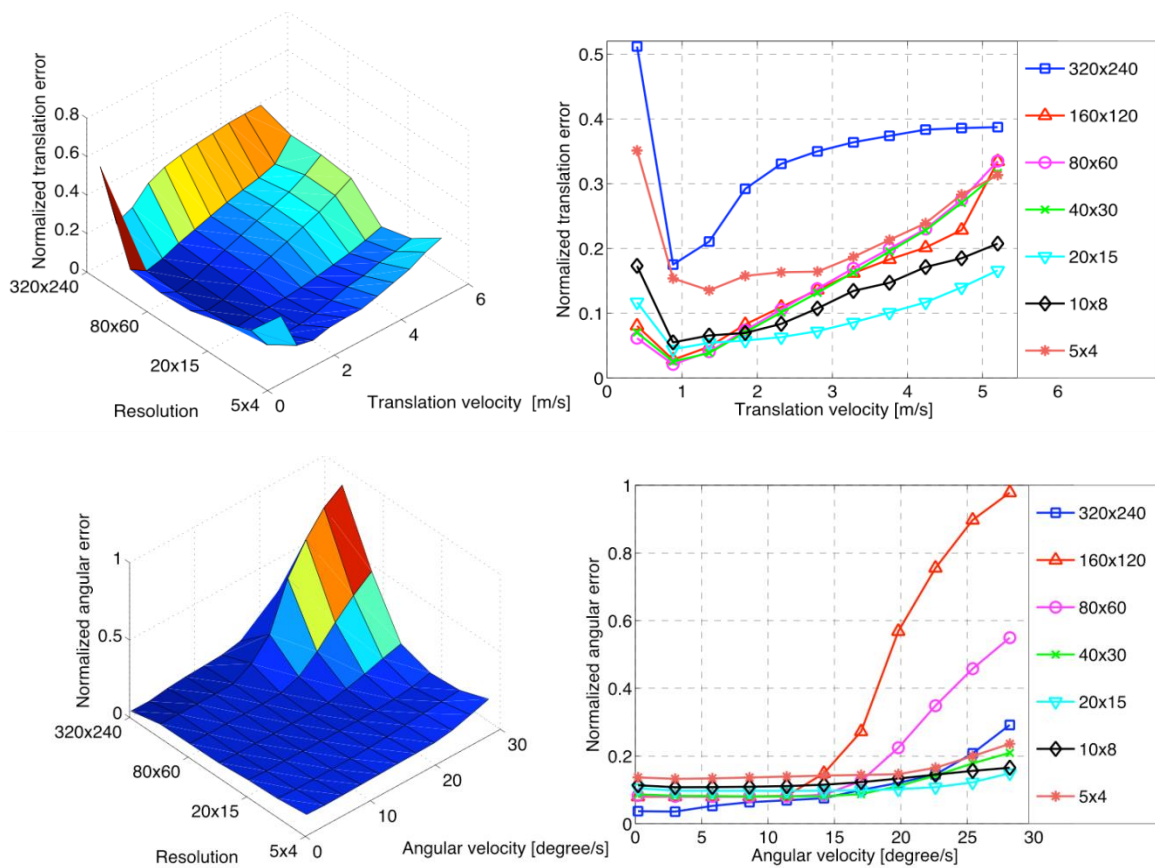


Figure 45 : Erreurs d'estimation en fonction de la résolution d'image et des vitesses. En haut, erreur pour des mouvements de translation uniquement ; en bas, erreurs pour des mouvements de rotation uniquement.

IV.3 Conclusions

Dans ce chapitre, nous avons présenté les travaux effectués sur des capteurs omnidirectionnels où la contrainte de centralité de la projection n'était plus respectée. Les avantages de ces solutions sont :

- la possibilité de former un champ de vue prescrit tout en maintenant une résolution constante, limitant ainsi les distorsions habituellement rencontrées sur les capteurs catadioptriques à projection centrale ;
- la possibilité d'une perception tridimensionnelle de la scène avec une seule caméra ce qui peut faciliter la résolution de problèmes de type *Structure From Motion*.

Ces capteurs restent cependant complexes à prototyper et à modéliser. Dans le cas des caméras plénoptiques, il faut cependant souligner que ces caméras commencent à se démocratiser, plusieurs dispositifs sont aujourd'hui commercialisés ; initialement cantonnées à la photographie (*refocusing*), des solutions permettant l'acquisition vidéo existent, qui devraient amener la communauté à explorer de nouveaux paradigmes en vision par ordinateur.

Les travaux que nous avons menés dans la thèse de Jean-François Layerle [Th3] nous ont conduits à nous intéresser à la détection et suivi de visages dans une image omnidirectionnelle. Nous avons prolongé ces travaux dans le cadre d'un programme de recherche sur la biométrie nomade mené avec l'Université du Kent (Interreg IV NOBA). Cette étude a abouti à des résultats intéressants sur l'adaptation d'algorithmes de détection de visages sur images omnidirectionnelles et sur un champ de recherche encore peu exploré, celui de la reconnaissance d'une personne à partir de sa démarche. Nous en présentons une synthèse dans le chapitre suivant.

V. DETECTION DE VISAGES ET RECONNAISSANCE BIOMETRIQUE DE LA MARCHE A PARTIR D'IMAGES OMNIDIRECTIONNELLES

Ce chapitre est une synthèse des travaux de recherche des thèses de Jean-François Layerle [Th3] et Yohan Dupuis [Th7] dans lesquels nous nous sommes attachés à adapter certaines méthodes récentes de détection et classification utilisées en vision par ordinateur au cas de la vision omnidirectionnelle : modèles actifs d'apparence, détection du visages, reconnaissance de la marche.

La production scientifique associée à ces travaux est référencée en [bleu](#) sous la forme [REF] dans le document.

V.1 Détection et suivi du visage

Travaux initiaux sur le suivi du visage

Dans la thèse de Jean-François Layerle [Th3] et dont les travaux sur le design de miroirs ont été présentés dans le chapitre précédent, nous nous sommes intéressés aux techniques de suivi 3D du visage en utilisant les modèles d'apparence active (AAM) qui présentent l'avantage d'une plus grande robustesse aux déformations du visage par rapport aux modèles rigides (Figure 46). Cette technique avait été utilisée avec succès pour le monitoring du conducteur en 2004 [31].

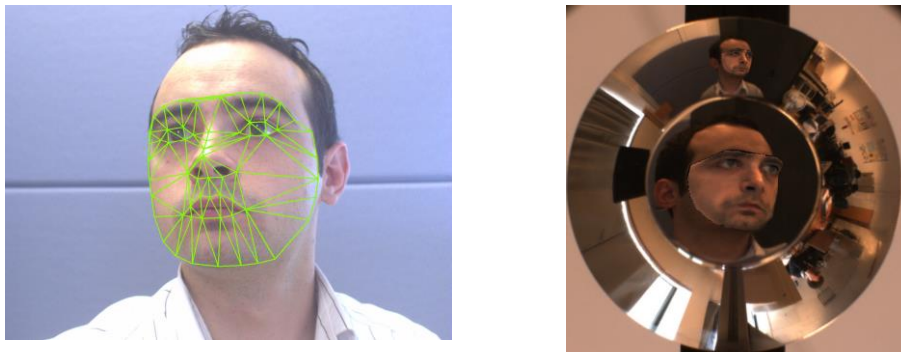


Figure 46 : Méthodes de suivi 3D du visage par modèle d'apparence active sur les images omnidirectionnelles

Vision omnidirectionnelle pour la détection de visages et la biométrie

Dans le cadre du projet Interreg IV NOBA (*Nomad Biometric Authentication*) et de la thèse de Yohan Dupuis qui y était rattachée [Th7], nous avons exploré un thème quasiment vierge, celui de l'utilisation de la vision omnidirectionnelle et des capteurs catadioptriques pour la reconnaissance d'individus par leur visage et leur démarche.

Dans un premier temps, nous avons cherché à contourner un des défauts majeurs des capteurs catadioptriques à savoir leur faible résolution, en concevant un système d'acquisition (Figure 47) associant un capteur catadioptrique (fournissant une vue omnidirectionnelle mais de faible résolution) et une caméra Pan-Tilt-Zoom (fournissant une vue dans une direction de vue donnée en haute résolution) ; ce dispositif a été présenté dans [ACTI6] [ACTI13] [ACLN1].

Dans un second temps et afin de déterminer de manière omnidirectionnelle les zones d'intérêt dans lesquelles diriger la caméra PTZ, nous nous sommes intéressés à la problématique de la détection de visages sur images omnidirectionnelles. La détection de visages a été abondamment traitée dans le cas des images perspectives (voir en particulier Viola & Jones [70] qui fait office de référence en matière de détection de visages dans des images) mais elle n'avait pas été abordée pour des images

omnidirectionnelles, sauf à appliquer une opération de « désanamorphose » et à considérer l'image obtenue comme une image perspective, ce qui permet de se replacer dans le cadre des méthodes classiques.

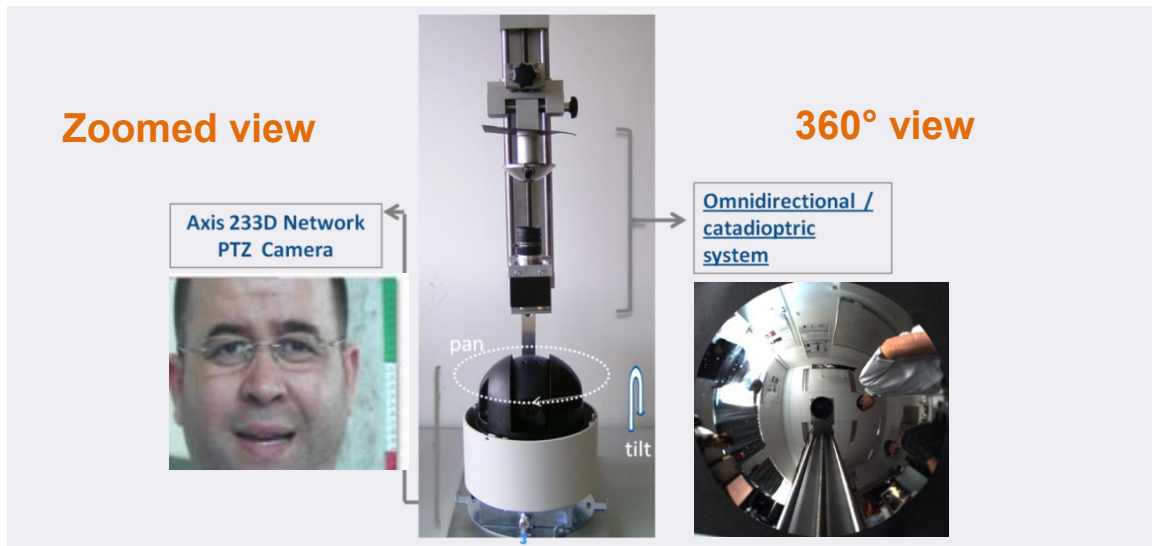


Figure 47 : système de vision hybride conçu pour le projet NOBA et associant un capteur catadioptrique et une camera PTZ.

Adaptation du détecteur de visages de Viola-Jones aux images omnidirectionnelles

La détection de visages dans les images est un thème qui a été largement étudié ; la maturité des algorithmes permet aujourd'hui d'avoir ce type de traitements embarqués dans un appareil photo ou un smartphone. Parmi les méthodes existantes, la plus connue et la plus utilisée pour ses performances de détection en temps réel et celle initialement proposée par Viola et Jones [70]. Cependant, nous nous sommes aperçus que ce détecteur voyait ses performances dégradées lorsqu'il était appliqué sur des images omnidirectionnelles. Nous nous sommes donc attachés à adapter les différentes étapes de ce type de détecteur au cas des images omnidirectionnelles.

La détection de visages implique trois étapes-clés : l'extraction de caractéristiques suffisamment discriminantes (*features*), le choix d'une méthode de classification et la phase d'entraînement. A partir de la méthode de classification de Viola-Jones, nous avons reconsidéré ces trois points dans le cas des capteurs catadioptriques. Dans un premier temps, nous avons contourné l'absence de bases de données d'images omnidirectionnelles de visages en les synthétisant par une projection à partir de plus de 2 500 images perspectives issues de la base FDDB⁴.

⁴ Face Detection Data Set and Benchmark

Dans un deuxième temps, nous nous sommes intéressés au choix des descripteurs dans le cas d'images omnidirectionnelles de visages. Les détecteurs de visages temps réel tel que celui de Viola-Jones utilisent comme descripteurs les caractéristiques pseudo-Haar (*Haar-like features*) construites à partir de masques rectangulaires. Dans le cas des images omnidirectionnelles et comme illustré Figure 48, un rectangle d'un plan de la scène se projette en un tétragone sur la sphère d'équivalence dont la forme discrétisée est alors un polygone. Ce constat nous a conduits à privilégier l'utilisation de caractéristiques pseudo-Haar polygonales récemment introduites par Pham et al [50].



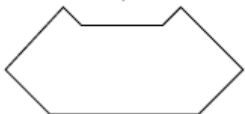
	Shapes	Integral Image	Integral Type
Rectangle		$I(i,j)$	Discrete
Projection on omnidirectional image		$I(\theta, \varphi)$	Continuous
Polygon		$I(i,j)$	Discrete

Figure 48 : Les descripteurs de Haar sont basés sur des sommes de masques rectangulaires, opérations qui peuvent être accélérées par le calcul d'une image intégrale. La projection d'un rectangle sur une image omnidirectionnelle étant un tétragone dont deux côtés sont formés par des arcs de cercle, l'utilisation de formes polygonales est plus appropriée dans le cas de la détection de visages sur images omnidirectionnelles.

Enfin nous avons examiné les apports des récentes variantes de l'algorithme Adaboost du détecteur de Viola et Jones (*Real, Gentle et Modest Adaboost*).



Figure 49 : détection de visages sur images omnidirectionnelles. En jaune, un seul visage est correctement détecté par l'algorithme de Viola-Jones ; en rouge, les visages correctement détectés par notre méthode. Sans adaptation, la méthode de Viola-Jones échoue du fait de la forte distorsion et de la rotation dans le plan inhérente à l'observation omnidirectionnelle.

Ces travaux ont démontré qu'il était possible de détecter des visages directement dans les images omnidirectionnelles (Figure 49) en conservant un schéma compatible avec des applications temps réel. Ces travaux ont fait l'objet de plusieurs publications internationales et nationales [ACL1] [ACL4] [ACTI9] [ACTI10] [ACTI12].

V.2 Reconnaissance biométrique par l'observation visuelle de la démarche d'un individu

La reconnaissance biométrique d'une personne à partir de traits comportementaux caractéristiques de sa démarche est un axe de recherche assez récent [43]. Cette modalité a retenu notre attention dans le cadre du projet NOBA car cette reconnaissance peut être faite à distance et de manière ambiante. Dans le cadre de ma thèse, j'avais été amené à étudier différentes approches pour la modélisation biomécanique de l'homme en utilisant des modèles polyarticulés [T2]. L'application était cependant différente puisqu'il s'agissait d'en déduire des indicateurs de la performance cinématique et dynamique d'une personne pendant sa démarche en conjuguant modélisation biomécanique et mesure du mouvement. Dans le cadre du projet NOBA où l'idée était d'utiliser l'observation de la marche cette fois dans le but d'identifier la personne, notre choix s'est porté sur les méthodes de reconnaissance sans modèle (model-free) qui sont moins lourdes en temps de calcul comparées aux méthodes nécessitant le placage d'un modèle polyarticulé.

Les approches sans modèle impliquent d'extraire la silhouette d'une personne lorsqu'elle marche ; nous avons donc été amenés à aborder le problème de la segmentation d'images issues d'un capteur catadioptrique et obtenues en environnement non contrôlé. Afin de séparer les pixels d'arrière-plan

de ceux de la silhouette, une méthode originale a été appliquée à partir de séquences vidéos couleur en utilisant l'algorithme des K-moyennes (*k-means*) que nous avons publiée dans [ACTI11] .

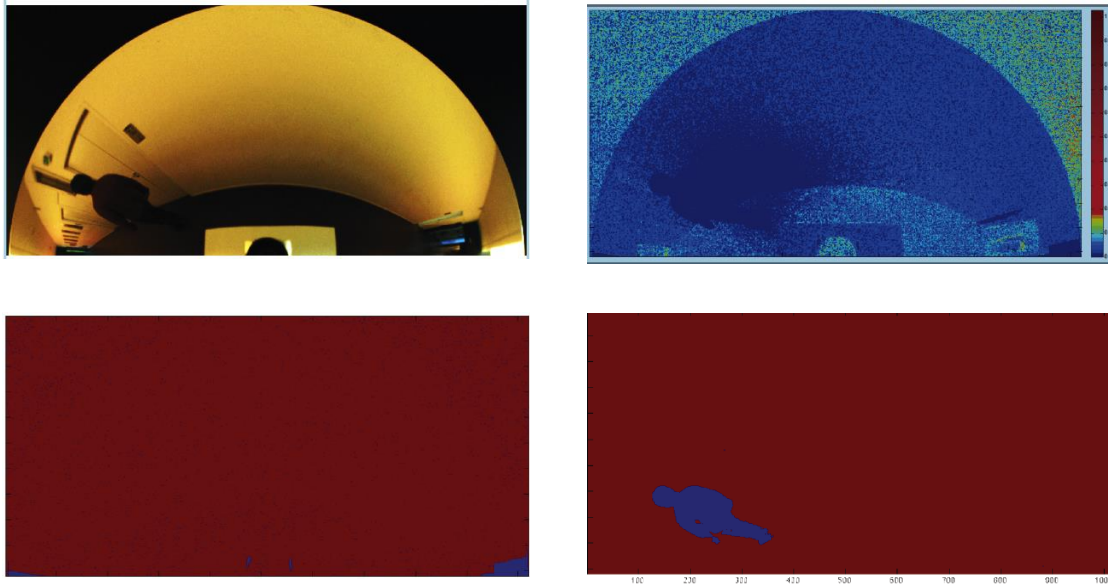


Figure 50 : Segmentation de silhouettes sur image catadioptrique pour la reconnaissance de la marche. La séquence vidéo est enregistrée sur une personne marchant dans un couloir (en haut à gauche). La figure en haut à droite représente la probabilité qu'un pixel appartienne à la silhouette. Comparée à une méthode classique par soustraction d'images (en bas à gauche), notre approche (en bas à droite) permet une segmentation robuste aux effets de variation d'éclairage.

A partir d'une revue exhaustive des méthodes les plus récentes pour la reconnaissance biométrique par la marche, nous avons retenu celles consistant à calculer, à partir d'une séquence vidéo, une représentation de l'énergie contenue dans le mouvement de la marche (*gait energy image* ou GEI). A partir de cette représentation, nous avons expérimenté une méthode de sélection d'attributs (*feature selection*) utilisant la technique des forêts aléatoires (*random forest*). Nous avons montré qu'avec cette étape, la signature GEI était moins sensible aux changements d'apparence. Les tests ont été effectués sur les bases de données biométriques CASIA (*Chinese Academy of Sciences, Institute of Automation*, Figure 51).



Figure 51 : exemple de scenarios de marche disponibles dans la base de données CASIA (source : [Th7])

Les résultats obtenus ont été particulièrement concluants en égalant voire dépassant les scores de reconnaissance des dernières références dans ce domaine. Notre approche et les résultats que nous avons obtenus ont été publiés dans *Image and Computer Vision* [ACL3].

V.3 Conclusions

Dans ce chapitre, nous avons rassemblé des travaux visant l'application de la vision omnidirectionnelle et des capteurs catadioptriques à un domaine connexe à celui de la robotique mobile, la détection et la reconnaissance de personnes dans une scène. Nous avons mené des travaux pour améliorer les méthodes de détection et de reconnaissance elles-mêmes basées sur des traits physiologiques (visage) ou comportementaux (marche) au cas de la vision omnidirectionnelle. Ces études ont aussi été l'occasion de travaux pluridisciplinaires dans les domaines de la vision par ordinateur, du traitement d'images et du *Machine Learning*.

Les travaux menés en vision par ordinateur avec des applications principalement tournées vers les systèmes embarqués, la perception tridimensionnelle et l'estimation du mouvement - le tout conduit dans le cadre de projets collaboratifs associant laboratoires académiques et industriels - ont été le socle d'un projet structurant aujourd'hui abouti : le Campus En Intégration de Systèmes Embarqués (CISE). Il a consisté en une extension du laboratoire au sein duquel j'ai été chargé de mettre en œuvre un plateau technique dévolu à la Navigation Autonome. La genèse de ce projet, le travail préparatoire et de mise en œuvre que nous avons conduits et les premiers développements en termes de projets collaboratifs sont l'objet du chapitre suivant.

VI. MISE EN PLACE D'UN PLATEAU DE MESURES DEDIE A LA NAVIGATION AUTONOME

Cet avant-dernier chapitre présente mon travail d'analyse, spécifications et pilotage de la mise en œuvre d'un plateau expérimental dévolu à l'étude de la navigation autonome de véhicules, plateforme technologique déployé dans le cadre du projet structurant de Campus en Intégration de Systèmes Embarqués (CISE) de l'IRSEEM entre 2008 et 2012.

VI.1 Le 21^{ème} siècle ou l'essor des véhicules autonomes

Etat des lieux

Le véhicule autonome, longtemps sujet d'étude cantonné au monde de la recherche académique en particulier dans le domaine de la robotique mobile, est promis à une diffusion forte dans les prochaines années dans de nombreux secteurs industriels au premier plan desquels celui du monde automobile. C'est dans ce secteur qu'on voit se multiplier des projets et démonstrateurs de véhicules automatisés voire autonomes : challenges DARPA, navettes automatisées pour le transport individuel (projets Praxitèle, Cybercars, Vipalab), démonstrateurs de conduite automatisée des constructeurs automobiles (projets Bertha Benz de Daimler, Drive Me de Volvo, Next 2 de Renault...).

La diffusion de véhicules capables de se mouvoir dans un environnement réel et complexe avec peu voire pas d'intervention humaine ne se limite pas aux domaines de l'automobile et des transports ; ces véhicules trouveront aussi un usage dans le domaine de la logistique⁵, la sécurité ou encore l'aide à la personne.

Deux facteurs ont largement contribué à cette évolution :

- Une abondante production de connaissances dans les domaines connexes de la robotique - informatique, automatique, mécatronique et intelligence artificielle - mettant ainsi à disposition des ingénieurs un vaste panel de méthodes et d'algorithmes pour concevoir des véhicules intelligents et capables d'agir de plus en plus en autonomie d'action ;
- La diffusion massive des Technologies de l'Information et de la Communication, au premier rang desquelles ce que l'on a coutume d'appeler les systèmes électroniques embarqués. Associant puces électroniques et logiciel enfoui, et dont l'augmentation exponentielle des performances repousse chaque jour les limites de l'intelligence dont on peut doter tous les objets de notre quotidien. Cette intelligence se caractérise par une autonomie de fonctionnement qui implique que le système ait été doté de capacités de perception, d'analyse, de décisions et de partage d'informations en temps réel.

⁵ Google a récemment fait l'acquisition de plusieurs startups leader dans le domaine de la robotique mobile et projette de nouveaux services tels que la livraison automatisée de colis en ville.

A titre d'exemple, le domaine automobile est assez illustratif de cette évolution. Après une diffusion massive de l'électronique embarquée dans les véhicules fin des années 90, dans un premier temps pour des fonctions de contrôle-commande, les constructeurs ont rapidement cherché à assister le conducteur dans sa tâche de conduite et développé des systèmes dits ADAS (Advanced Driving Assisted Systems). Pour comprendre les situations, alerter, voire agir à la place du conducteur, ces systèmes doivent disposer d'informations extérieures au véhicule, effectuer la détection et l'identification des autres obstacles qui se trouvent autour (infrastructure, autres véhicules, piétons), utiliser des données cartographiques, connaître l'état du trafic. Ces informations peuvent être perçues depuis le véhicule par des capteurs embarqués (caméra, radar, laser,...), ou communiquées par l'infrastructure voire, dans un futur proche, par les autres véhicules (communication Car to X).

Tableau 4 : Définition des niveaux d'automatisation pour le véhicule autonome
(source : Standard SAE J3016)

	Niveau	Intitulé	Description
Supervision de la conduite par le conducteur	0	Pas d'automatisation	Le conducteur effectue toutes les tâches de conduite avec l'aide éventuelle de systèmes d'avertissement
	1	Assistance à la conduite	Le système d'assistance agit soit sur la direction (contrôle latéral) soit sur la vitesse (contrôle longitudinal) en utilisant des informations sur la situation de conduite ; le conducteur effectue toutes les autres tâches
	2	Automatisation partielle	Le système d'assistance agit à la fois sur la direction et sur la vitesse en utilisant des informations sur la situation de conduite ; le conducteur effectue toutes les autres tâches
Supervision de la conduite par la machine	3	Automatisation sous conditions	Dans des scénarios de conduite définis, le système d'assistance agit sur l'ensemble des tâches de conduite mais le conducteur doit rester en mesure de reprendre le contrôle dans un délai approprié
	4	Automatisation élevée	Dans des scénarios de conduite définis, le système d'assistance agit sur l'ensemble des tâches de conduite sans action du conducteur
	5	Automatisation complète	Le système d'assistance agit sur l'ensemble des tâches de conduite sans action du conducteur et en toute situation

Avec l'électrification des fonctions du véhicule (colonne de direction, groupe motopropulseur, freinage), l'architecture d'une automobile est de plus en plus proche de celle d'un robot mobile. Le concept futuriste d'un véhicule conduisant à notre place n'est donc plus très loin ; actuellement les véhicules dits autonomes nécessitent toujours un conducteur dans la boucle (la SAE a établi une standardisation illustrant bien cette problématique et présenté Tableau 4) et il reste malgré tout de nombreux verrous pour aller vers une totale autonomie du véhicule.

Cet exemple peut bien entendu se généraliser à d'autres domaines d'application ayant trait à la mobilité intelligente, la gestion de l'énergie embarquée dans les secteurs de l'automobile (aides à la conduite, électromobilité, véhicule autonome), le transport et la logistique mais aussi le secteur de la construction navale (intelligence embarquée pour le navire du futur) ou encore le médical et les services à la personnes (robotique d'assistance, aides technologiques pour le maintien à domicile).

Enjeux scientifique et technologiques des véhicules autonomes

Toute l'intelligence dans la navigation autonome d'un véhicule va résider dans le triptyque perception-décision-action, le système devant être capable de comprendre un environnement complexe et dynamique (objets mobiles avec des vitesses relatives importantes, changements récurrents des conditions de visibilité et de propagation des ondes) et interagir avec. Cette interaction doit aussi se faire avec l'homme, en particulier le conducteur ou l'opérateur car à ce stade des avancées en intelligence artificielle, l'homme doit rester dans la boucle pour solutionner des situations singulières et non apprises par la machine. Cette problématique pose la question de l'interface homme-machine et des stratégies de délégation et de reprise ; là encore la question de la perception par la machine des actions ou intentions de l'humain sera un élément-clé.

La tâche de perception va consister à appréhender l'environnement dans lequel évolue le véhicule pour lui permettre de s'y localiser et d'y naviguer. La perception de l'environnement passe par l'utilisation d'un ou plus capteurs extéroceptifs devant caractériser la scène de manière exhaustive en appréhendant sa dimension spatiale (scène en trois dimensions) et temporelle (scène et/ou véhicule observant en mouvement) et en fournissant des informations de haut niveau à la couche décisionnelle (détection des régions d'intérêt, extraction de primitives, classification des objets de la scène). Pour des applications dans le domaine automobile ou de la robotique des services s'ajoute une contrainte sur le coût des capteurs et sur les capacités de calcul du calculateur embarqué.

Plusieurs stratégies sont possibles pour permettre la navigation autonome d'un robot, elles diffèrent essentiellement selon que l'on a une connaissance a priori ou non de l'environnement dans lequel va

évoluer le robot. Dans une approche où le véhicule découvre l'environnement, il lui faut à la fois s'y localiser et reconstruire une carte d'occupation locale typiquement par des méthodes de type SLAM pour que la couche décisionnelle puisse planifier et exécuter une trajectoire sûre et optimale.

En environnement routier, une autre approche va consister à utiliser la signalisation horizontale (bords de voie, marquages au sol) pour réaliser le contrôle latéral et longitudinal. La perception de l'environnement peut être simplement représentée par une carte d'occupation en deux dimensions pour permettre la planification de la trajectoire. Dans ces deux approches, l'exigence qui pèse sur la tâche de perception est très forte car il faut pouvoir retrouver la structure statique de l'environnement et la discerner de la partie dynamique, mais aussi comprendre la scène à un haut niveau sémantique (respect de la signalisation).

Une autre solution va consister à utiliser une connaissance exhaustive de la scène parcourue, représentée sous forme de cartes métriques et topologiques et qui peuvent être annotées hors ligne (signalisation), voire enrichies dynamiquement (par d'autres véhicules ou à partir d'information remontée par l'infrastructure elle-même). L'environnement étant parfaitement connu et numérisé, les actions du véhicule pourront être établies hors ligne et corrigées en temps réel en fonction des obstacles détectés par le véhicule. La localisation pourra être basée sur une mise en correspondance de ce qui est perçu depuis le véhicule avec les données cartographiques à partir d'une information de localisation approximative (GPS bas coût par exemple). C'est l'approche retenue pour le véhicule de Google ; étendue à des grands environnements, elle exige d'importantes ressources de stockage et une coopération entre le véhicule et un serveur distant, la masse de données ne pouvant être embarquée dans le véhicule.

La navigation autonome a été un sujet très étudié en robotique ces vingt dernières années mais essentiellement sur des véhicules évoluant dans des environnements contrôlés, principalement en intérieur [46] [63] [68]. Comme déjà évoqué, l'extension de ces travaux à des véhicules évoluant dans des environnements extérieurs est possible, elle a été démontrée ces dernières années avec de nombreux projets et démonstrateurs (Figure 52) en particulier dans le cadre des challenges du Darpa avec la réalisation de véhicules autonomes capables d'évoluer dans des environnements complexes et de grande dimension.



Figure 52 : historique non exhaustif des véhicules autonomes. En haut à gauche, véhicule automatisé de l'Université de Tsukuba (1977). En haut à droite, California Path : démonstration d'un conduite en peloton sur autoroute (1997). En bas à gauche, DARPA Urban Challenge (2007). En bas à droite, Google autonomous car (aujourd'hui). (Source : Jan Becker, Advanced Driver Assistance, Stanford University)

Cependant si on étudie de près l'architecture de ces véhicules, on constate que les systèmes électroniques qui y sont embarqués, en particulier les capteurs proprioceptifs et extéroceptifs - télémètres laser, radar, caméras, GPS différentiel et centrale inertielle - utilisés pour localiser le véhicule, lui fournir une perception exhaustive de la scène l'environnant et lui permettre de planifier et exécuter une trajectoire, ne sont pas compatibles pour une diffusion de masse du fait de leur coût élevé et de leur complexité de mise en œuvre (Figure 53). L'autre enjeu concerne l'architecture de calcul qui doit être compatible avec les contraintes de l'embarqué dans l'automobile.

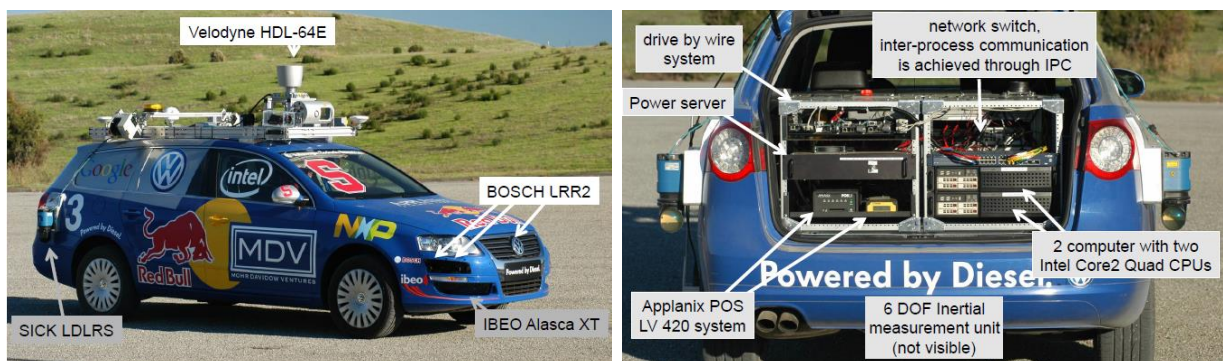


Figure 53 : véhicule de l'université de Stanford pour le Darpa Challenge : plus de 6 télémètres laser sont nécessaires pour la perception de l'environnement et la localisation. La photographie de droite en dit long sur la difficulté à rendre compatible un tel système avec les contraintes de l'embarqué en automobile (Source : Jan Becker, Advanced Driver Assistance, Stanford University)

La navigation autonome de véhicule en condition réelle et dans des environnements extérieurs de grande dimension présente encore de nombreux challenges en particulier sur la tâche de perception du fait de :

- la grande variabilité des scènes rencontrées : environnements urbains, périurbains, ruraux, cartes associées de très grande dimension
- la forte dynamique des scènes : déplacement du véhicule de quelques km/h à plus de 100km/h sur autoroute, présence de nombreux objets en mouvement de nature très différente : piétons, deux-roues, voitures et poids-lourds,...
- des conditions d'exposition très changeantes : fort ensoleillement, pluie, brouillard, vision de nuit...

D'autre part, l'autonomie décisionnelle de la machine suppose un niveau de sécurité et de fiabilité dans les fonctions de perception ; cette exigence pousse à l'utilisation de solutions reposant sur la fusion de plusieurs modalités (Figure 54).

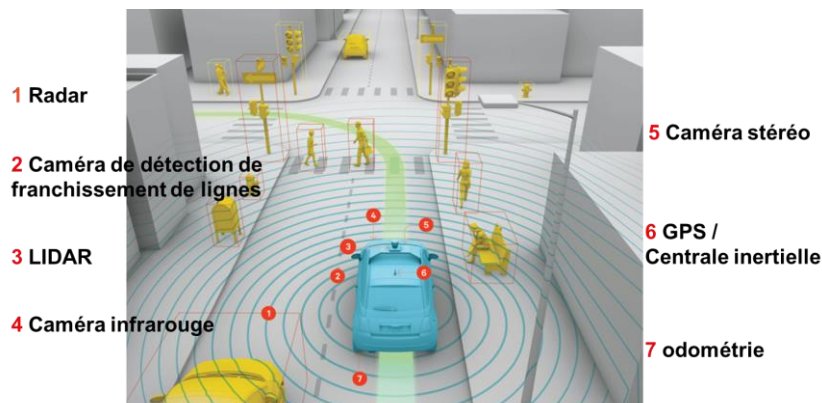


Figure 54 : illustration des enjeux autour de l'intégration des systèmes embarqués pour la navigation autonome : pas moins de sept systèmes de perception de l'environnement et du mouvement peuvent être requis dans la Google Autonomous Car.

Dans un tel schéma il est alors primordial de caractériser la performance et les modes de défaillance de chaque modalité. Face à la très grande diversité de scénarios à laquelle un véhicule autonome sera confronté en environnement extérieur, les erreurs liées aux ambiguïtés de perception et la variabilité perceptuelle seront inévitables. Il sera donc essentiel pour chaque capteur d'être capable de s'auto-diagnostiquer et fournir un indice de confiance de sa mesure.

VI.2 Plateforme de mesures pour la navigation autonome

Dans le cadre du projet de Campus en Intégration de Systèmes Embarqués de l'IRSEEM, j'ai travaillé à la mise en place d'une plateforme technologique pour les futurs systèmes d'aide à la conduite et la robotique mobile autonome. Initié en 2008, ce travail a porté sur la définition des besoins immobiliers et matériels pour la mise en œuvre d'un plateau technique dévolu à l'étude de la navigation autonome des systèmes mobiles. Ce dispositif a pour ambition de renforcer les travaux de l'IRSEEM dans les domaines de la perception de scènes et l'intelligence embarquée. Ce projet s'est inscrit dans un contexte de forte pénétration de l'électronique et des logiciels embarqués dans des systèmes de contrôle pour l'automobile, l'aéronautique et la défense.

Afin de pouvoir concevoir et évaluer ces nouveaux concepts et systèmes dans des scénarios en lien avec les nouveaux usages de la mobilité - éco-mobilité et systèmes de transport durable, robotique des services - la plate-forme « Navigation Autonome » rassemble un ensemble d'outils de simulation et de mesure pour la caractérisation des systèmes de perception, leur intégration mécatronique et leur évaluation en situation contrôlée (Figure 55).

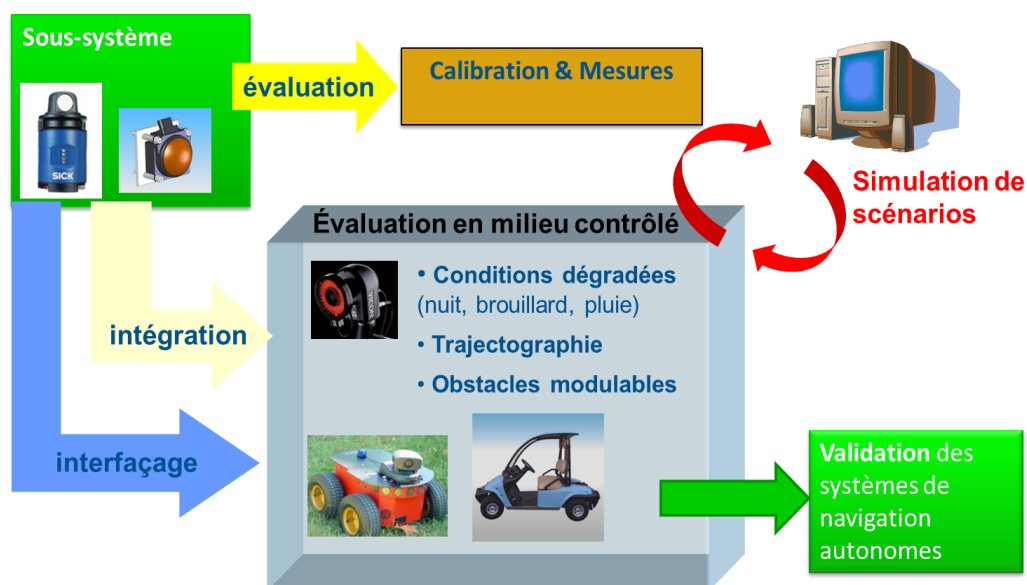


Figure 55 : Les moyens de caractérisation et de tests de la plate-forme « Navigation Autonome » couvrent les étapes-clés pour l'intégration de systèmes embarqués en robotique mobile autonome

Au sein de ce plateau, la mise en place des équipements s'est faite en deux grandes phases :

Phase 1 : moyens de test et essais pour la navigation autonome en environnement intérieur contrôlé

Phase 2 : moyens de test et essais en environnement extérieur

Pour chacune de ces deux phases, mon champ d'intervention a été le suivant :

- étude, définition de cahier des charges, rédaction et analyse des offres ;
- suivi des acquisitions des matériels ;
- pilotage des équipes de chercheurs, ingénieurs et techniciens pour la mise en œuvre des équipements au sein du plateau.

Le plateau technique est constitué de trois salles dont une dédiée aux expérimentations (150m², sous 5m de plafond, cf. Figure 56).



Figure 56 : vue de la zone d'expérimentation de la plate-forme «Navigation Autonome»

La particularité de ce plateau est de rassembler un ensemble de moyens de tests pour l'étude de systèmes mobiles en environnement intérieur contrôlé puis en environnement extérieur. Pour ce faire, les principaux équipements mis en œuvre sont un système de capture 3D du mouvement, des flottes de robots mobiles et un système complet de localisation précise et de perception 3D embarqué sur un véhicule ou un robot mobile.

Ces moyens vont permettre de nombreux travaux de recherche sur la perception de scènes dynamiques avec la possibilité d'établir des bases de référence (cartes 3D précises), de mesurer voire rejouer des trajectoires de mobiles et d'évaluer les performances des systèmes d'acquisition (capteurs et algorithmes) sur des scénarios représentatifs de situations réelles.

La mise en fonction de cette plateforme a été achevée en janvier 2012 pour un budget total de 1,1M€ environ.

Cette plateforme a permis le démarrage de plusieurs projets de recherche impliquant plusieurs enseignants-chercheurs et ingénieurs de l'IRSEEM. Les principaux sont :

- **ARGOS**, challenge international de robotique, organisé par Total et l'ANR. L'IRSEEM s'est associé à la PME SOMINEX. L'IRSEEM et SOMINEX pour réaliser un prototype répondant aux normes de sécurité (ATEX) et capable de réaliser des missions de surveillance et de détection d'incidents (fuite de gaz, flaques d'huile...) avec un niveau de complexité croissant allant du mode téléopéré au mode entièrement autonome.
- **DESERVE**, *DEvelopment platform for Safe and Efficient dRIVE*, JU ARTEMIS, programme européen réunissant 26 partenaires du monde de la R&D automobile, pour la mise en place d'outils de développement pour les ADAS. Les équipes de l'IRSEEM y développent des fonctions pour la localisation et la cartographie ainsi que des méthodes de test virtuel.
- **NAVALIS**, *crewboat innovant pour l'éolien offshore*, ADEME « véhicule du futur », projet porté par un consortium de PME du domaine de la construction navale et visant à l'amélioration des conditions de vie et de sécurité des personnels de maintenance à bord des crewboats. L'IRSEEM est chargé de l'étude et du prototypage de systèmes de stabilisation de passerelles de débarquement incluant le positionnement dynamique du navire.
- **VIATIC**, *Viabilité et AuTonomie des systèmes en environnement Incertain et Contraint*, ANR ASTRID avec MBDA, EHESS et LASTRE. Ce projet vise à étudier les applications de la théorie de la Viabilité pour le contrôle de robots mobiles avec une application duale défense/civile, planification de trajectoires de missiles vs. allocation de bornes de recharge pour des véhicules électriques. L'IRSEEM a réalisé les expérimentations en s'appuyant sur les moyens du Laboratoire de Navigation Autonome.

Ces nouveaux moyens ont aussi renforcé les collaborations de l'équipe au niveau académique et industriel ; des travaux concernant de la perception d'environnement et la localisation pour la robotique et mobilisant cette plateforme ont déjà pu faire l'objet de publications [\[ACTI1\]](#) [\[ACTI2\]](#) [\[ACTI3\]](#) [\[ACTN1\]](#).

VII. PERSPECTIVES DE RECHERCHE

Dans ce dernier chapitre, après un rapide bilan des travaux exposés dans les chapitres précédents, je présente les axes de recherche que je souhaite suivre dans mes travaux futurs et du renforcement et développement des collaborations académiques et industriels que les sous-tendent.

VII.1 Bilan général sur mes activités de recherche

Ma contribution dans le développement des différentes activités de recherche de l'IRSEEM s'est surtout axée sur la constitution d'un thème de recherche dans le domaine de la vision par ordinateur et plus spécifiquement les approches non conventionnelles (capteurs catadioptriques, système de vision plénoptique) pour répondre à des problématiques de perception d'environnement et de localisation pour des systèmes mobiles.

Les travaux se sont dans un premier temps orientés autour des capteurs catadioptriques en considérant les aspects design de miroirs, modélisation géométrique et méthodes d'étalonnage. Nous avons ensuite cherché à exploiter ces approches non conventionnelles pour des applications en lien avec la robotique mobile et le véhicule intelligent : *egomotion*, stéréovision omnidirectionnelle, *structure from motion*. Plus récemment nous avons exploré le potentiel des systèmes de vision plénoptique pour des applications en robotique mobile autonome. Enfin nous nous sommes intéressés à l'application des techniques d'apprentissage (*machine learning*) pour la compréhension de scènes observées par des systèmes de vision omnidirectionnelle.

En parallèle de cette activité, j'ai maintenu une activité de recherche en instrumentation et systèmes embarqués pour l'étude du mouvement humain en participant à des groupes de travail associant le CETAPS de l'Université de Rouen et le Département de chirurgie orthopédique du CHU de Rouen au sein de différents dispositifs (GHRAL, ERT). Ces travaux menés ont donné matière à deux publications dans des revues [ACLN2] [ACL8]. Enfin, dans une logique de pluridisciplinarité et en lien avec les activités de recherche du pôle électronique de l'IRSEEM, j'ai participé à des travaux sur le renforcement de l'immunité aux agressions électromagnétiques des systèmes à microcontrôleur par des techniques logicielles. Dans ce cadre, j'ai co-encadré une thèse du pôle électronique de l'IRSEEM [Th5] et ai été co-auteur de plusieurs articles [ACL5] [ACL6] [ACTI14].

VII.2 Activités en cours et perspectives

Actuellement, mes travaux sont concentrés sur l'utilisation des capteurs de vision omnidirectionnelle dans des systèmes mobiles dotés d'autonomie de déplacement. En effet l'utilisation de la vision pure pour de telles fonctions reste encore aujourd'hui un challenge. Les verrous sont liés à la complexité des scènes et des scénarios qui les sous-tendent : forte dynamique liée au mouvement du véhicule et des obstacles, environnement complexe, gestion des variations de luminosité et enjeu de robustesse dans des situations dégradées (pluie, brouillard, éclairage naturel vs. éclairage artificiel, nuit, etc.)

Fort de ce constat, mes travaux se structurent aujourd'hui autour de deux thèmes rassemblés sous le nom « capteurs 2D+t » et capteurs « 3D+t » et détaillés ci-après.

Utilisation de capteurs « 2D+t »

Les travaux menés par l'équipe et plus largement par la communauté scientifique sur la vision omnidirectionnelle permettent aujourd'hui de considérer comme acquises les étapes de modélisation géométrique et de calibration. L'étude de ces systèmes a aussi démontré qu'ils étaient bien appropriés pour des tâches de navigation d'un véhicule, en particulier celles consistant à retrouver simultanément le mouvement propre du mobile et la structure tridimensionnelle de l'environnement et s'y localiser par des approches de type détection et suivis d'amers géoréférencés. Ces problèmes restent encore ouverts en particulier dans le cas de systèmes basés vision en condition extérieure et face à des scènes complexes et à forte dynamique. Nos efforts vont plus spécifiquement se concentrer sur l'utilisation de systèmes de vision non conventionnels - capteurs catadioptriques et polydioptriques, objectifs fish-eyes, caméras plénoptiques - qui représentent une opportunité de reformuler le problème SFM d'une manière plus simple pour sa résolution. Ce travail va prendre corps dans le cadre du projet ROADTRAC retenu à l'appel à projet régional du Grand Réseau de Recherche en Electronique et déposé conjointement avec le LITIS et le CEREMA. Dans ce projet, l'IRSEEM sera chargé d'étudier le potentiel de ce type de capteurs pour caractériser de manière exhaustive la scène routière (détection et classification des types de véhicules, reconstruction tridimensionnelle des trajectoires).

D'autres modalités d'acquisition telles que la perception dans le domaine infrarouge et plus largement les caméras hyperspectrales sont intéressantes à étudier. Les évolutions technologiques de ce type de caméra rendent crédibles l'idée de les utiliser en application embarquée. On trouve malgré tout peu de travaux sur leur utilisation pour des applications de localisation et de navigation autonome bien que cette modalité présente l'avantage de s'affranchir des conditions d'éclairage et qu'elle soit la seule possible pour réaliser une tâche de navigation par vision de nuit. La coexistence de systèmes de vision fonctionnant dans des spectres différents posent des problèmes dans la mise en correspondance, soit basée sur des descripteurs (locaux ou globaux) communs entre les images, soit par recalage entre reconstruction 3D (ICP par exemple). Ce travail fait l'objet de la thèse de Fabien Bonardi [Th8]. Ce travail s'inscrit dans un cadre plus large du projet INTERREG IV-A SAVEMORE (IRSEEM, LITIS, Université du Kent) que je coordonne et qui vise à faire la preuve de concept d'un véhicule autonome pour la livraison de marchandises en ville.

Enfin, le thème de l'optimisation des algorithmes sur des architectures embarquées que nous avons abordé dans le cadre de la thèse de Romain Rossi [Th4] reste porteur avec la diffusion de solutions embarquées couplant architectures multicœurs et architectures massivement parallèles. Des premiers échanges avec ST-Microelectronics Grenoble ont eu lieu sur le sujet afin d'engager des travaux de R&D sur de nouvelles architectures processeur dédiées pour le traitement d'images et la vision par ordinateur en temps réel. Cette collaboration a fait l'objet d'un dépôt de projet européen avec le CEA-LETI et l'INRIA Grenoble.

Utilisation de capteurs 3D+t

Les progrès en électronique et traitement du signal embarqué entraînent une diffusion croissante de capteurs mesurant directement la structure géométrique de la scène et la restituant sous forme de nuages de points (caméras temps de vol, LIDARS 2D, LIDARS multi-nappes). Ces systèmes peuvent être utilisés pour obtenir une reconstruction dense et précise de l'environnement. Nous allons nous intéresser à l'utilisation de ce type de capteurs pour deux projets (ARGOS et NAVALIS) pour lesquels les conditions environnementales peuvent être sévères (vision nocturne, embruns, brouillard, atmosphères explosives...) excluant la possibilité d'utiliser des caméras passives. De nouveau, nos travaux consisteront à sélectionner et évaluer des méthodes d'estimation robuste pour implémenter des algorithmes de recalage entre nuages de points mesurés autour du véhicule et données cartographiques de l'environnement (CAO d'un bâtiment industriel par exemple) via la mise en œuvre de simulateurs puis de démonstrateurs complets. Ce travail est au cœur de la thèse de Pierre Merriaux [Th9].

VIII. LISTE DE DE MES PUBLICATIONS

MEMOIRES

- [T1] **Savatier, X.**, « Etude en simulation d'un modèle dynamique de la marche », *Rapport de DEA de Robotique*, Université Pierre et Marie Curie, 1997
- [T2] **Savatier, X.**, « Etude et mise en oeuvre d'un système d'observation de la qualité locomotrice chez l'homme », thèse de doctorat de l'Université de Rouen, 2003

Articles dans des revues internationales ou nationales avec comité de lecture répertoriées par l'AERES ou dans les bases de données internationales (ISI Web of Knowledge, Pub Med...)

- [ACL1] Dupuis, Y.; **Savatier, X.**; Ertaud, J. & Vasseur, P. (2014), 'Détection radiale de visages sur images omnidirectionnelles', *Traitement du Signal* **1**, 143-173.
- [ACL2] Dong, F.; Ieng, S.-H.; **Savatier, X.**; Etienne-Cummings, R. & Benosman, R. (2013), 'Plenoptic cameras in real-time robotics', *The International Journal of Robotics Research* **32**(2), 206-217.
- [ACL3] Dupuis, Y.; **Savatier, X.** & Vasseur, P. (2013), 'Feature Subset Selection Applied to Model-Free Gait Recognition', *Image and Vision Computing*, **31**(8), 580-591.
- [ACL4] Dupuis, Y.; **Savatier, X.**; Ertaud, J. & Vasseur, P. (2013), 'Robust Radial Face Detection for Omnidirectional Vision', *Image Processing, IEEE Transactions on*, **22**(5), 1808-1821.
- [ACL5] Fayu, W.; Duval, F.; **Savatier, X.** & Mazari, B. (2011), 'Effects of conducted electromagnetic interference on an analog-to-digital converter', *IET Electronic Letters* **47**(1), 23-25.
- [ACL6] Fayu, W.; Duval, F. and Cao, H.; **Savatier, X.** & Mazari, B. (2010), 'Increase of immunity of microcontroller to conducted continuous-wave interference by detection method', *IET Electronic Letters* **46**(6), 1113-1114.

- [ACL7] Boutteau, R.; **Savatier, X.**; Ertaud, J. & Mazari, B. (2009), 'An Omnidirectional Stereoscopic System for Mobile Robot Navigation', *Sensors & Transducers Journal, Special Issue on Robotic and Sensors Environments* 5, 3-17.
- [ACL8] L'hermette, M.; **Savatier, X.**; Baudry, L.; Tourny-Chollet, C. & Dujardin, F. (2008), 'A New Portable Device for Assessing Locomotor Performance', *International Journal of Sports Medicine* **29**(4), 322-326.

Articles dans des revues internationales ou nationales avec comité de lecture non répertoriées par l'AERES ou dans les bases de données internationales (ISI Web of Knowledge, Pub Med...)

- [ACLN1] Khemmar, R.; Ertaud, J. Y. & **Savatier, X.** (2013), 'Face Detection and Recognition based on Fusion of Omnidirectional and PTZ Vision Sensors and Heterogenous Database', *International Journal of Computer Applications* **61**(21), 35-44.
- [ACLN2] Dujardin, F.; Mouilhade, F.; Lhermette, M.; **Savatier, X.**; Ertaud, J. & Tourny-Chollet, C. (2006), 'Evaluation quantitative du handicap locomoteur par mesure accélérométrique de la cinématique pelvienne', *Revue de chirurgie orthopédique et réparatrice de l'appareil moteur* **92**(6), 71.
- [ACLN3] Ragot, N.; Ertaud, J.; **Savatier, X.** & Mazari, B. (2005), 'Modelling of the environment of a mobile robot by catadioptric sensors. State of the art and proposal for an innovative architecture', *Best-of book of AMSE Modeling conferences*.

Communications avec actes dans un congrès international

- [ACTI1] Merriaux, P.; Dupuis, Y.; Vasseur, P. & **Savatier, X.** (2014), 'Wheel Odometry-based Car Localization and Tracking on Vectorial Map', *IEEE Transportation Systems Conference (ITSC 2015)*, Qingdao, China.
- [ACTI2] Dupuis, Y.; Merriaux, P.; Subirats, P.; Boutteau, R.; **Savatier, X.** & Vasseur, P. (2014), 'GPS-based Preliminary Map Estimation for Autonomous Vehicle Mission Preparation', *IEEE/RSJ International Conference on Intelligent Robots (IROS)*, Chicago, USA.
- [ACTI3] Merriaux, P.; Boutteau, R.; Vasseur, P. & **Savatier, X.** (2014), 'IMU/LIDAR based positioning of a gangway for maintenance operations on wind farms', *IEEE/RSJ International Conference on Intelligent Robots (IROS)*, Chicago, USA.

- [ACTI4] Boutteau, R.; **Savatier, X.** & Ertaud, J. (2013), 'A dynamic programming algorithm applied to omnidirectional vision for dense 3D reconstruction', Recent Advances in Computer Vision and Pattern Recognition (RACVPR), Okinawa, Japan.
- [ACTI5] Boutteau, R.; **Savatier, X.**; Bonardi, F. & Ertaud, J.-Y. (2013), 'Road-line detection and 3D reconstruction using fisheye cameras', IEEE Transportation Systems Conference (ITSC 2013), Den Haag, The Netherlands.
- [ACTI6] Raj, A.; Khemmar, R.; Ertaud, J. & **Savatier, X.** (2013), 'Face Detection and Recognition under Heterogeneous Database based on Fusion of Catadioptric & PTZ Vision Sensors', Eighth International Conference on Computer Recognition Systems (CORES 2013), Milkow, Poland.
- [ACTI7] Belbachir, A.; Boutteau, R.; Merriaux, P.; Blosseville, J. & **Savatier, X.** (2013), 'From Autonomous Robotics Toward Autonomous Car', IEEE Intelligent Vehicle Symposium (IV 2013), Gold Coast, Australia.
- [ACTI8] Liu, J.; Khemmar, R.; Ertaud, J. & **Savatier, X.** (2012), 'Compressed Sensing Face Recognition Method in Heterogeneous Database with Small Sample Size Problem', Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS 2012), Naples, Italy.
- [ACTI9] Dupuis, Y.; **Savatier, X.**; Ertaud, J.-Y. & Vasseur, P. (2011), 'A New Approach for Face Detection with Omnidirectional Sensors', IEEE International Conference on Robotics and Biomimetics (ROBIO 2011), Phuket Island, Thailand.
- [ACTI10] Dupuis, Y.; **Savatier, X.**; Ertaud, J.-Y. & Vasseur, P. (2011), 'A direct approach for face detection on omnidirectionnal images', IEEE International workshop on robotic and sensors environments (ROSE 2011), Montreal, Canada.
- [ACTI11] Dupuis, Y.; **Savatier, X.**; Ertaud, J. & Hoblos, G. (2010), 'A Segmentation Method of Catadioptric Images for Gait Recognition in Unconstrained Environment', International Conference on Emerging Security Technologies (EST2010), Canterbury, United Kingdom.

- [ACTI12] Dupuis, Y.; **Savatier, X.**; Ertaud, J.-Y. & Hoblos, G. (2010), 'A Framework for Face Detection on Central Catadioptric Systems', IEEE International workshop on robotic and sensors environments (ROSE 2010), Phoenix, United States.
- [ACTI13] Iraqui, A.; Dupuis, Y.; Boutteau, R.; Ertaud, J.-Y. & **Savatier, X.** (2010), 'Fusion of Omnidirectional and PTZ cameras for face detection & tracking', International Conference on Emerging Security Technologies (EST2010), Canterbury, United Kingdom.
- [ACTI14] Fayu, W.; Duval, F.; **Savatier, X.** & Mazari, B. (2009), 'Characterization of the Susceptibility of Microcontroller to Conducted Continuous-wave Interference', 7th International Workshop on Electromagnetic Compatibility of Integrated Circuits (EMC COMPO09), Toulouse, France.
- [ACTI15] Rossi, R.; **Savatier, X.**; Ertaud, J. & Mazari, B. (2009), 'Real-Time 3D Reconstruction for Mobile Robot Using Catadioptric Cameras', IEEE International workshop on robotic and sensors environments (ROSE 2009), Leco, Italia.
- [ACTI16] Boutteau, R.; **Savatier, X.**; Ertaud, J. & Mazari, B. (2008), 'An omnidirectionnal stereoscopic system for mobile robot navigation', IEEE International workshop on robotic and sensors environments (ROSE 2008), Ottawa, Canada.
- [ACTI17] Layerle, J.; **Savatier, X.**; Ertaud, J. & Mouaddib, E. M. (2008), 'An Innovative Catadioptric Vision Device for the Driver Assistant Systems', International Conference on Vehicle and Infrastructure Safety Improvement in Adverse Conditions and Night Driving (VISION'08), Versailles Satory, France.
- [ACTI18] Layerle, J.; **Savatier, X.**; Ertaud, J. & Mouaddib, E. M. (2008), 'Catadioptric Sensor for a Simultaneous Tracking of the Driver's Face and the Road Scene', 8th Workshop on on Omnidirectional Vision, Camera Networks and Non-classical Cameras (OMNIVIS'08), Marseille, France.
- [ACTI19] Layerle, J.; **Savatier, X.**; Mouaddib, E. & Ertaud, J. (2008), 'Catadioptric vision system for an optimal observation of the driver face and the road scene', IEEE Intelligent Vehicles Symposium, Eindhoven, The Netherlands.

- [ACTI20] Ragot, N.; Rossi, R.; **Savatier, X.**; Ertaud, J. & Mazari, B. (2008), '3D volumetric reconstruction with a catadioptric stereovision sensor', IEEE Industrial Electronics (ISIE08), Cambridge, United Kingdom.
- [ACTI21] Ragot, N.; Ertaud, J.; **Savatier, X.** & Mazari, B. (2006), 'An efficient catadioptric sensor calibration based on a low-cost test-pattern', First International Conference on Computer Vision Theory and Applications (VISAPP), Setúbal, Portugal.
- [ACTI22] Ragot, N.; Ertaud, J.; **Savatier, X.** & Mazari, B. (2006), 'Calibration of a Panoramic Stereovision Sensor: Analytical vs Interpolation-Based Methods', IEEE Industrial Electronics (IECON), Paris, France.

Communications avec actes dans un congrès national

- [ACTN1] Merriaux, P.; Bouteau, R.; Vasseur, P. & **Savatier, X.** (2014), 'Algorithme de positionnement d'une passerelle à mouvements compensés à partir de mesures inertielles et lidar pour les opérations de maintenance des parcs éoliens offshore', Reconnaissance des Formes et Intelligence Artificielle (RFIA), Saint Etienne du Rouvray, France.
- [ACTN2] Ragot, N.; Merriaux, P.; Rossi, R.; Leclercq-Delapierre, D.; **Savatier, X.** & Delarue, J.-J. (2014), 'Le projet COALAS : un exemple de formation des élèves-ingénieurs par la Recherche ', 11^{ème} Colloque sur l'Enseignement des Technologies et des Sciences de l'Information et des Systèmes (CETSIS), Besançon, France.
- [ACTN3] Dupuis, Y.; **Savatier, X.**; Ertaud, J.-Y. & Vasseur, P. (2013), 'Détection radiale de visages sur images omnidirectionnelles', Actes du congrès des jeunes chercheurs en vision par ordinateur (ORASIS'13), Abbaye de Cluny, France.
- [ACTN4] Merriaux, P.; Delarue, J.-J. & **Savatier, X.** (2013), Un environnement attractif, ludique et SIMPLE pour des travaux pratiques, c'est possible, Actes du Colloque sur l'Enseignement des Technologies et des Sciences de l'Information et des Systèmes (CETSIS), Caen, France.

- [ACTN5] Layerle, J.; **Savatier, X.**; Ertaud, J. & Mouaddib, E. M. (2009), 'Conception d'un capteur catadioptrique pour une observation optimale du conducteur et de la scène de conduite', Actes du congrès des jeunes chercheurs en vision par ordinateur (ORASIS'09), Trégastel, France.

Communications orales sans actes dans un congrès international ou national

- [COM1] **Savatier, X.**; Boutteau, R.; Merriaux, P. & Le Méneç, S. (2012), 'Demonstration of the Salvo Enhanced No Escape Zone Concept Using Ground Mobile Robots', Fifteenth International Symposium on Dynamic Games and Applications (SDG 2012), Byšice, Czech
- [COM2] Virmani, M.; Dupuis, Y. & **Savatier, X.** (2010), 'Augmented Reality on the Web: Development using Computer Vision; Problems & Proposed Solutions with the present', W3C workshop: Augmented Reality on the Web, Barcelona, Spain.
- [COM3] Ragot, N.; Ertaud, J.; **Savatier, X.** & Mazari, B. (2007), 'Calibrage non paramétrique d'un capteur de stéréovision panoramique', Journée thématique du GDR ISIS et du Club EEA « Journée Vision Omnidirectionnelle », Amiens, France.

IX. BIBLIOGRAPHIE

- [1] Adelson, E. H. & Bergen, J. R. (1991), The Plenoptic Function and the Elements of Early Vision, in 'Computational Models of Visual Processing', pp. 3--20.
- [2] Aycard, O.; Baig, Q.; Bota, S.; Nashashibi, F.; Nedevschi, S.; Pantilie, C.; Parent, M. N.; Resende, P.; Vu, T.-D. & others (2011), Intersection safety using lidar and stereo vision sensors, in 'IV'2011-IEEE Intelligent Vehicles Symposium', pp. 863--869.
- [3] Baker, S. & Nayar, S. (1999), 'A Theory of Single-Viewpoint Catadioptric Image Formation', *International Journal of Computer Vision (IJCV)* **35**(2), 175-196.
- [4] Barth, M. & Barrows, C. (1996), A fast panoramic imaging system and intelligent imaging technique for mobile robots, in 'Intelligent Robots and Systems' 96, IROS 96, Proceedings of the 1996 IEEE/RSJ International Conference on', pp. 626--633.
- [5] Barreto, J. (2006), 'A Unifying Geometric Representation for Central Projection Systems', *Computer Vision and Image Understanding (CVIU)* **103**(3), 208-217.
- [6] Belbachir, A.; Schraml, S.; Mayerhofer, M. & Hofstatter, M. (2014), A Novel HDR Depth Camera for Real-time 3D 360-degree Panoramic Vision, in 'Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops', pp. 419--426.
- [7] Benosman, R.; Clercq, C.; Lagorce, X.; Ieng, S.-H. & Bartolozzi, C. (2014), 'Event-based visual flow', *Neural Networks and Learning Systems, IEEE Transactions on* **25**(2), 407 - 417.
- [8] Bertozzi, M. & Broggi, A. (1998), 'GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection', *Image Processing, IEEE Transactions on* **7**(1), 62--81.
- [9] Bolopion, A.; Ni, Z.; Agnus, J.; Benosman, R. & Régnier, S. (2012), Stable haptic feedback based on a dynamic vision sensor for microrobotics, in 'Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on', pp. 3203--3208.
- [10] Cabral, E. L.; de Souza, J. & Hunold, M. C. (2004), Omnidirectional stereo vision with a hyperbolic double lobed mirror, in 'Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on', pp. 1--9.
- [11] Camahort, E.; Leries, A. & Fussell, D. (1998), Uniformly sampled light fields'Rendering Techniques', Springer, pp. 117--130.
- [12] Cauchois, C.; Brassart, E.; Drocourt, C. & Vasseur, P. (1999), Calibration of the omnidirectional vision sensor: SYCLOP, in 'Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on', pp. 1287--1292.
- [13] Caron, G.; Marchand, E. & Mouaddib, E.-M. (2009), 3D model based pose estimation for omnidirectional stereovision, in 'Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on', pp. 5228--5233.
- [14] Chen, L.; Wang, W.; Zhang, M.; Bao, W. & Zhang, X. (2011), 'Complementary-structure catadioptric omnidirectional sensor design for resolution enhancement', *Optical Engineering* **50**(3), 033201--033201.
- [15] Chiabrando, F.; Piatti, D. & Rinaudo, F. (2010), 'SR-4000 ToF camera: further experimental tests and first applications to metric surveys', *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **38**(5), 149--154.
- [16] Daniilidis, K.; Makadia, A. & Bulow, T. (2002), Image processing in catadioptric planes: Spatiotemporal derivatives and optical flow computation, in 'Omnidirectional Vision, 2002. Proceedings. Third Workshop on', pp. 3--10.

- [17] Delbruck, T. (2008), Frame-free dynamic digital vision, in 'Proceedings of Intl. Symp. on Secure-Life Electronics, Advanced Electronics for Quality Life and Society', pp. 21--26.
- [18] Demonceaux, C. & Vasseur, P. (2006), 'Markov random fields for catadioptric image processing', *Pattern Recognition Letters* **27**(16), 1957--1967.
- [19] Demonceaux, C.; Vasseur, P. & Fougerolle, Y. (2011), 'Central catadioptric image processing with geodesic metric', *Image and Vision Computing* **29**(12), 840--849.
- [20] Fermüller, C. & Aloimonos, Y. (2000), 'Observability of 3D motion', *International Journal of Computer Vision* **37**(1), 43--63.
- [21] Floreano, D.; Pericet-Camara, R.; Viollet, S.; Ruffier, F.; Brückner, A.; Leitel, R.; Buss, W.; Menouni, M.; Expert, F.; Juston, R. & others (2013), 'Miniature curved artificial compound eyes', *Proceedings of the National Academy of Sciences* **110**(23), 9267--9272.
- [22] Foix, S.; Alenya, G. & Torras, C. (2011), 'Lock-in Time-of-Flight (ToF) Cameras: A Survey', *IEEE Sensors Journal* **11**(3), 1.
- [23] Geyer, C. & Daniilidis, K. (2001), 'Catadioptric projective geometry', *International Journal of Computer Vision (IJCV)* **43**(3), 223-243.
- [24] Gonzalez-Barbosa, J. (2004), 'Vision panoramique pour la robotique mobile : stéréovision et localisation par indexation d'images', PhD thesis, Thèse de doctorat, Université Toulouse III.
- [25] Gortler, S. J.; Grzeszczuk, R.; Szeliski, R. & Cohen, M. F. (1996), The lumigraph, in 'Proceedings of the 23rd annual conference on Computer graphics and interactive techniques', pp. 43--54.
- [26] Greguss, P. (1986), 'Panoramic imaging block for three-dimensional space', Google Patents, US Patent 4,566,763.
- [27] Hartley, R. & Zisserman, A. Press, C. U., ed. (2004), *Multiple View Geometry in Computer Vision*, Cambridge University Press.
- [28] Hicks, R. A. (2005), 'Designing a mirror to realize a given projection', *JOSA A* **22**(2), 323--330.
- [29] Hicks, R. A. & Perline, R. K. (2004), The method of vector fields for catadioptric sensor design with applications to panoramic imaging, in 'Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on', pp. 11--143.
- [30] Huang, K. S.; Trivedi, M. M. & Gandhi, T. (2003), Driver's view and vehicle surround estimation using omnidirectional video stream, in 'Intelligent Vehicles Symposium, 2003. Proceedings. IEEE', pp. 444--449.
- [31] Ishikawa, T. (2004), 'Passive driver gaze tracking with active appearance models', .
- [32] Kim, J.-H. & Chung, M. J. (2003), SLAM with omni-directional stereo vision sensor, in 'Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on', pp. 442--447.
- [33] Jang, G.; Kim, S. & Kweon, I. (2006), 'Single-camera panoramic stereo system with single-viewpoint optics', *Optics Letters* **31**, 41-43.
- [34] Kleinschmidt, L. (1911), 'Aotabatto Fob', Google Patents, US Patent 994,935.
- [35] Köse, E. & Perline, R. K. (2014), 'Double-mirror catadioptric sensors with ultrawide field of view and no distortion', *Applied optics* **53**(4), 528--536.
- [36] Lin, S.-S. & Bajcsy, R. (2001), True single view point cone mirror omni-directional catadioptric system, in 'Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on', pp. 102--107.
- [37] Labayrade, R.; Aubert, D. & Tarel, J.-P. (2002), Real time obstacle detection in stereovision on non flat road geometry through " v-disparity" representation, in 'Intelligent Vehicle Symposium, 2002. IEEE', pp. 646--651.

- [38] Levoy, M. & Hanrahan, P. (1996), Light field rendering, in 'Proceedings of the 23rd annual conference on Computer graphics and interactive techniques', pp. 31--42.
- [39] López-Nicolás, G. & Sagüés, C. (2014), 'Unitary torus model for conical mirror based catadioptric system', *Computer Vision and Image Understanding* **126**, 67--79.
- [40] Lourenço, M.; Barreto, J. P. & Vasconcelos, F. (2012), 'srd-sift: Keypoint detection and matching in images with radial distortion', *Robotics, IEEE Transactions on* **28**(3), 752--760.
- [41] Lui, W. L. D. & Jarvis, R. (2010), 'Eye-Full Tower: A GPU-based variable multibaseline omnidirectional stereovision system with automatic baseline selection for outdoor mobile robot navigation', *Robotics and Autonomous Systems* **58**(6), 747--761.
- [42] Marhic, B.; Mouaddib, E. M.; Pegard, C. & Hutin, N. (1999), Localisation based on invariant-models recognition by SYCLOP, in 'Intelligent Robots and Systems, 1999. IROS'99. Proceedings. 1999 IEEE/RSJ International Conference on', pp. 793--798.
- [43] Matvoski, D.; Nixon, M. & Carter, J. (2014), 'Gait analysis', *Computer Vision: A Reference Guide*.
- [44] Mei, C. & Rives, P. (2007), Single View Point Omnidirectional Camera Calibration from Planar Grids, in 'IEEE International Conference on Robotics and Automation (ICRA)', pp. 3945-3950.
- [45] Miura, J.; Negishi, Y. & Shirai, Y. (2002), Mobile robot map generation by integrating omnidirectional stereo and laser range finder, in 'Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on', pp. 250--255.
- [46] Murray, D. & Little, J. J. (2000), 'Using real-time stereo vision for mobile robot navigation', *Autonomous Robots* **8**(2), 161--171.
- [47] Neumann, J. (2004), 'Computer vision in the space of light rays: plenoptic videogeometry and polydioptric camera design', PhD thesis, University of Maryland.
- [48] Ni, Z. (2013), 'Asynchronous event based vision: algorithms and applications to microrobotics', PhD thesis, Paris 6.
- [49] Nister, D. (2004), 'An efficient solution to the five-point relative pose problem', *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* **26**, 756-770.
- [50] Pham, M.-T.; Gao, Y.; Hoang, V.-D. & Cham, T.-J. (2010), Fast polygonal integration and its application in extending haar-like features to improve object detection, in 'Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on', pp. 942--949.
- [51] Puig, L.; Bermúdez, J.; Sturm, P. & Guerrero, J. J. (2012), 'Calibration of omnidirectional cameras in practice: A comparison of methods', *Computer Vision and Image Understanding* **116**(1), 120--137.
- [52] Rameau, F.; Demonceaux, C.; Sidibé, D.; Fofi, D. & others (2014), Control of a PTZ camera in a hybrid vision system, in 'International Conference on Computer Vision Theory and Applications'.
- [53] Rees, D. W. (1966), 'Hyperbolic Ellipsoidal Real Time Display Panoramic Viewing Installation for Vehicles', Google Patents, US Patent 3,229,576.
- [54] Sankaranarayanan, A. & Baraniuk, R. (2014), 'Compressive Sensing', *Computer Vision: A Reference Guide*.
- [55] Scaramuzza, D. (2008), 'Omnidirectional Vision : From calibration to robot motion estimation', PhD thesis, Thèse de doctorat, ETH Zurich.
- [56] Scaramuzza, D.; Martinelli, A. & Siegwart, R. (2006), A Toolbox for Easily Calibrating Omnidirectional Cameras, in 'IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)', pp. 5695-5701.

- [57] Silberstein, S.; Levi, D.; Kogan, V. & Gazit, R. (2014), Vision-based pedestrian detection for rear-view cameras, in 'Intelligent Vehicles Symposium Proceedings, 2014 IEEE', pp. 853--860.
- [58] Smadja, L.; Benosman, R. & Devars, J. (2004), Cylindrical sensor calibration using lines, in 'Image Processing, 2004. ICIP'04. 2004 International Conference on', pp. 1851--1854.
- [59] Stettner, R. (2010), Compact 3D flash lidar video cameras and applications, in 'SPIE Defense, Security, and Sensing', pp. 768405--768405.
- [60] Sturm, P.; Ramalingam, S.; Tardif, J.-P.; Gasparini, S. & Barreto, J. (2011), 'Camera models and fundamental concepts used in geometric computer vision', *Foundations and Trends in Computer Graphics and Vision* **6**(1--2), 1--183.
- [61] Sturm, P. (2014), 'Pinhole camera model', *Computer Vision: A Reference Guide*.
- [62] Su, L.; Luo, C. & Zhu, F. (2006), Obtaining obstacle information by an omnidirectional stereo vision system, in 'Information Acquisition, 2006 IEEE International Conference on', pp. 48--52.
- [63] Surmann, H.; Nüchter, A. & Hertzberg, J. (2003), 'An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments', *Robotics and Autonomous Systems* **45**(3), 181--198.
- [64] Svoboda, T. (1999), 'Central Panoramic Cameras, Design, Geometry, Egomotion', PhD thesis, Thèse de doctorat, Czech Technical University.
- [65] Svoboda, T. & Pajdla, T. (2002), 'Epipolar Geometry for Central Catadioptric Cameras', *International Journal of Computer Vision (IJCV)* **49**(1), 23-37.
- [66] Swaminathan, R.; Nayar, S. K. & Grossberg, M. D. (2003), 'Framework for designing catadioptric projection and imaging systems', *Proc. ICCV-PROCAMS*.
- [67] Tan, K.-H.; Hua, H. & Ahuja, N. (2004), 'Multiview panoramic cameras using mirror pyramids', *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **26**(7), 941--946.
- [68] Thrun, S. (1998), 'Learning metric-topological maps for indoor mobile robot navigation', *Artificial Intelligence* **99**(1), 21--71.
- [69] Vasseur, P. & Mouaddib, E. (2004), Central catadioptric line extraction, in 'British Machine Vision Conference (BMVC)', pp. 57-66.
- [70] Viola, P. & Jones, M. J. (2004), 'Robust real-time face detection', *International journal of computer vision* **57**(2), 137--154.
- [71] Yagi, Y. & Kawato, S. (1990), Panorama scene analysis with conic projection, in 'IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)', pp. 181-187.
- [72] Yi, S. & Ahuja, N. (2006), An omnidirectional stereo vision system using a single camera, in 'Pattern Recognition, 2006. ICPR 2006. 18th International Conference on', pp. 861--865.

X. ANNEXE : MODÈLE DE DISTORTIONS

Le modèle sténopé suppose une projection perspective parfaite. En réalité la projection des rayons lumineux sur l'imageur est altérée par des aberrations géométriques (essentiellement des distorsions, cf. Figure 57) et chromatiques.

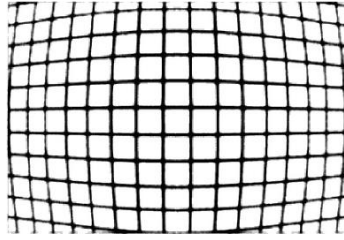


Figure 57 : exemple de distorsions géométriques

Le modèle de projection peut être affiné en modélisant les distorsions géométriques sous une forme polynomiale. En notant q_s la projection estimée de Q et q_d la projection réelle, on introduit deux types de distorsions :

$$q_d = q_s + \Delta_R + \Delta_T$$

avec :

$$\Delta_R = [\delta_{R_x} \quad \delta_{R_y}]^T \text{ représentant les distorsions radiales}$$

$$\Delta_T = [\delta_{T_x} \quad \delta_{T_y}]^T \text{ représentant les distorsions tangentielles}$$

$$\delta_{R_x} = (u - u_0) \sum_{i=1}^{\infty} k_i \rho^{2i} \quad ; \quad \delta_{R_y} = (v - v_0) \sum_{i=1}^{\infty} k_i \rho^{2i}$$

$$\delta_{T_x} = \left(p_1(\rho^2 + 2(u - u_0)^2) + 2p_2(u - u_0)(v - v_0) \right) \left(1 + \sum_{j=1}^{\infty} p_2 + j\rho^{2j} \right)$$

$$\delta_{T_y} = \left(p_2(\rho^2 + 2(v - v_0)^2) + 2p_1(u - u_0)(v - v_0) \right) \left(1 + \sum_{j=1}^{\infty} p_2 + j\rho^{2j} \right)$$

$$\rho^2 = (u - u_0)^2 + (v - v_0)^2$$

k_i représentent les coefficients de distorsion radiales, p_j les coefficients de distorsions tangentielles.

Ces paramètres devront donc être estimés lors la phase de calibration. En pratique on constate que l'ordre 2 ($i = j = 2$) est suffisant.

RESUME

Ce mémoire présente une synthèse des travaux que j'ai menés à l'ESIGELEC au sein de son institut de recherche l'IRSEEM. Mes activités de recherche ont porté dans un premier temps sur la conception et l'évaluation de dispositifs de mesure de la dynamique de la marche de personnes atteintes de pathologies de la hanche, dans le cadre de ma thèse effectuée à l'université de Rouen en lien le Centre Hospitalo-Universitaire de Rouen. En 2003, j'ai rejoint les équipes de recherche qui se constituaient avec la mise sur pieds de l'IRSEEM, Institut de Recherche en Systèmes Electroniques Embarqués, créé en 2001. Dans ce laboratoire, j'ai structuré et développé une activité de recherche dans le domaine de la vision par ordinateur appliquée au véhicule intelligent et à la robotique mobile autonome. Dans un premier temps, j'ai concentré mes travaux à l'étude de systèmes de vision omnidirectionnelle tels que les capteurs catadioptriques centraux et leur utilisation pour des applications mobiles embarquées ou débarquées : modélisation et calibrage, reconstruction tridimensionnelle de scènes par stéréovision et déplacement du capteur. Ces recherches ont fait l'objet de trois thèses co-encadrées à l'IRSEEM dont une en partenariat avec la DGA. Dans un second temps, je me suis intéressé à la conception et la mise en œuvre de systèmes de vision à projection non centrale (capteurs catadioptriques à miroirs composés, caméra plénoptique). Ces travaux ont été effectués au travers de deux thèses en collaboration avec le MIS de l'Université Picardie Jules Verne et l'ISIR de l'Université Pierre et Marie Curie. Enfin, dans le cadre d'un programme de recherche en collaboration avec l'Université du Kent, j'ai consacré une partie de mes travaux à l'adaptation de méthodes de traitement d'images et de classification pour la détection de visages sur images omnidirectionnelles (adaptation du détecteur de Viola et Jones) et à la reconnaissance biométrique d'une personne par analyse de sa marche. Aujourd'hui, mon activité s'inscrit dans le prolongement du renforcement des projets de l'IRSEEM dans le domaine de la robotique mobile et du véhicule autonome : mise en place d'un plateau de mesures pour la navigation autonome, coordination de projets de recherche en prise avec les besoins industriels. Mes perspectives de recherche ont pour objet l'étude de nouvelles solutions pour la perception du mouvement et la localisation en environnement extérieur et sur les méthodes et moyens nécessaires pour objectiver la performance et la robustesse de ces solutions sur des scénarios réalistes. Elles prennent corps au travers de deux thèses co-encadrées en collaboration avec le LITIS de l'Université de Rouen.