



HAL
open science

Perception multimodale pour un robot mobile en milieu marin

Yan Guo

► **To cite this version:**

Yan Guo. Perception multimodale pour un robot mobile en milieu marin. Automatique / Robotique. Université Pierre et Marie Curie - Paris VI, 2011. Français. NNT: . tel-00637552

HAL Id: tel-00637552

<https://theses.hal.science/tel-00637552>

Submitted on 2 Nov 2011

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

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

UNIVERSITÉ PIERRE ET MARIE CURIE
ÉCOLE DOCTORALE SMAE
SCIENCES MECANIQUE ACOUSTIQUE
ET ELECTRONIQUE DE PARIS

THÈSE

pour obtenir le titre de

Docteur de l'Université Pierre et Marie Curie

Présentée et soutenue par

Yan GUO

Perception multimodale pour un robot mobile en milieu marin

soutenue le 4 octobre 2011

Jury :

<i>Rapporteurs :</i>	Pascal VASSEUR	- Prof à l'Université de Rouen
	René GARELLO	- Prof à l'Ecole Telecom Bretagne
<i>Directeur :</i>	Bruno GAS	- Prof à l'Université Pierre et Marie Curie
<i>Examineurs :</i>	Jean-luc ZARADER	- Prof à l'Université Pierre et Marie Curie
	Hervé GLOTIN	- Prof à l'Université Sud Toulon Var
	Eric LUCET	- CdP chez ROBOSOFT
	Frédéric PLUMET	- MdC à l'Université Versailles St Quentin en Yvelines

Résumé

Dans le domaine de la robotique, les véhicules autonomes de surface en milieu marin jouent un rôle important. Ils permettent de réaliser des opérations dangereuses, comme la surveillance d'environnements marins ou encore des relevés hydrographiques. Avant d'envisager le déplacement d'un véhicule autonome de surface, il est nécessaire d'assurer sa perception de l'environnement. Elle consiste à observer, localiser et éviter les obstacles. A cause des contraintes technologiques, la complexité de l'environnement naturel, et de la diversité des situations rencontrées, il est difficile d'effectuer une plate-forme parfaitement autonome et adaptée à des applications variées.

Cette thèse s'inscrit dans le cadre du projet ASAROME (Autonomous Sailing Robot for Oceanographic MEasurements), un projet de réalisation d'un voilier autonome pour des missions de mesures et d'observations de longues durées. Dans un premier temps, nous avons conçu une plate-forme de perception composée de plusieurs types de capteurs : hydrophones, caméra panoramique, centrale inertielle et sonar. A partir des signaux acquis pendant des campagnes de test, nous avons développé des méthodes de traitement du signal. Elles consistent à analyser et traiter des signaux sonores sous-marins et des images panoramiques pour la reconnaissance d'objets, la détection et la localisation d'obstacles. Pour améliorer la perception des obstacles, des méthodes de fusion de données multi-capteurs ont été développées. L'ensemble des algorithmes ont été validés expérimentalement, en lac puis en mer.

Ce travail n'est que le premier pas vers la réalisation d'un robot autonome en milieu marin capable de réaliser des missions complexes de longue durée, mais il montre sa faisabilité par le développement de capacités de perception adaptées.

Mots clés : Véhicule autonome de surface, reconnaissance de signal sonore, traitement du signal, système embarqué, vision panoramique, fusion de données, multi-capteurs, perception.

Abstract

In the field of robotics, the autonomous surface vehicles play an important role. They can perform dangerous operations such as marine environment monitoring or hydrographic surveys. Before considering the movement of an autonomous surface vehicle, it is necessary to ensure its perception of the environment. It consists of observing, locating and avoiding obstacles. Due to technological constraints, the complexity of the natural environment and the diverse situations encountered, it is difficult to make a platform which is perfectly autonomous and suited to various applications.

This thesis is in the project of ASAROME, a national project about the realization of an autonomous sailboat for long-duration missions. At first, we have designed a perception platform which consists of several types of sensors : hydrophones, panoramic camera, inertiel measurement unit and sonar. From the signals acquired in the experiments, we have developed several signal processing methods. It consists of underwater signal and panoramic image processing for recognition, detection and location of obstacles. To improve the obstacles perception, multi-sensors fusion methods have been developed. All these algorithms are validated in lake or sea.

This work is just the first step towards the realization of an autonomous marine robot which is able to achieve long-duration complex missions. But it shows the feasibility through the development of adapted perception capacities.

Keywords : Autonomous surface vehicle, sound signal recognition, signal processing, embedded system, panoramic vision, data fusion, multi-sensors, perception.

TABLE DES MATIÈRES

Table des figures	xi
Liste des tableaux	xiii
Introduction générale	1
1 Généralités sur les véhicules autonomes de surface	5
1.1 Introduction	5
1.2 Etat de l'art	6
1.3 Conclusion	14
2 Système embarqué et capteurs	15
2.1 Description du projet	16
2.2 Présentation du système	17
2.2.1 Architecture du système	18
2.2.2 Montage des capteurs	18
2.3 Capteurs de perception	20
2.3.1 Capteur de vision omnidirectionnelle	21
2.3.2 Hydrophone	22
2.3.3 Centrale inertielle	23
2.3.4 Sonar	24
2.3.5 Logiciel de manipulation	24

2.4	Conclusion	26
3	Traitements des signaux sonores	27
3.1	Introduction	28
3.2	Représentation des signaux	31
3.2.1	ERB	32
3.2.2	LPC	32
3.2.3	LPCC	33
3.2.4	MFCC	33
3.3	Classification des signaux	34
3.3.1	Modèle prédictif	35
3.3.2	Classification <i>k-ppv</i>	36
3.3.3	Décisions locales et globales	37
3.4	Expérimentations	38
3.4.1	Codage	39
3.4.2	Paramètres des classifieurs	41
3.4.3	Résultats	42
3.5	Conclusion	44
4	Vision panoramique	47
4.1	Introduction	48
4.1.1	Historique sur les panoramas	48
4.2	Calibrage de la caméra panoramique	50
4.2.1	Positionnement de la caméra relativement au miroir	51
4.2.2	Réflexion des rayons par la surface du miroir	57
4.3	Estimation de l'orientation du bateau	60
4.3.1	Algorithme	60
4.3.2	Résultats des tests	64
4.4	Détection d'obstacle	65
4.5	Estimation de l'état de la mer par la vision : détection de l'orientation	67
4.5.1	Segmentation du ciel et de l'eau	68
4.5.2	Détection du contour horizon	68
4.5.3	Détermination du plan d'incidence lié à l'horizon	70
4.6	Résultats expérimentaux	71
4.6.1	Test pour la détection d'obstacle	71
4.6.2	Test de détection de l'orientation dans un bassin	72

4.6.3	Test de détection de l'orientation en mer	76
4.7	Conclusion	79
5	Système de fusion	81
5.1	Introduction	82
5.1.1	Niveau de fusion	83
5.1.2	Architectures topologiques	83
5.1.3	Système Multi-capteurs	85
5.2	Bloc d'intégration de la vision panoramique et de la centrale inertielle	86
5.2.1	Traitement des données de la centrale inertielle	86
5.2.2	Détection de l'orientation du bateau	88
5.2.3	Définition de la position d'un objet	89
5.2.4	Résultats expérimentaux	90
5.3	Bloc de fusion	95
5.3.1	Données capturées par le sonar	95
5.3.2	Prétraitement des données du sonar	99
5.3.3	Détection d'obstacles par le sonar	99
5.3.4	Calibration du sonar par rapport à la caméra	102
5.4	Les méthodes de fusion	102
5.4.1	Fusion par pondération 1 (F1)	103
5.4.2	Fusion par pondération 2 (F2)	104
5.4.3	Fusion par estimation (F3)	104
5.4.4	Fusion par filtrage simple (F4)	105
5.4.5	Fusion par filtrage de Kalman (F5)	106
5.5	Résultats expérimentaux	110
5.6	Conclusion	115
	Conclusions et Perspectives	117
	A Liste de la base de données sonores	121
	Bibliographie	123

TABLE DES FIGURES

1.1	Un catamaran automatique AUTOCAT de MIT aux USA.	7
1.2	L'USV Owl MK II qui est utilisé pour la reconnaissance et les missions de surveillance.	8
1.3	SPRINGER est le catamaran automatique de 3m.	9
1.4	Trois SCOUT après une sortie dans l'eau.	10
1.5	ROAZ de LSA en portugal.	10
1.6	L'ASV ATLANTIS de l'Université de Floride.	11
1.7	Le prototype d'un USV de l'Université de Whales.	12
2.1	Institut Systèmes Intelligents et Robotiques.	17
2.2	Schéma du bateau et la positionnement des capteurs.	18
2.3	L'architecture du système du projet ASAROME.	19
2.4	La maquette utilisée dans le test de la perception.	21
2.5	Caméra gigabit.	22
2.6	Hydrophone utilisée dans le projet.	22
2.7	La centrale inertielle de XSENS utilisée sur la maquette.	23
2.8	Fusion des trois types de capteurs pour calculer l'orientation absolue précise.	24
2.9	Sonar Micron de Tritech.	25
2.10	Plate-forme utilisée pendant le test de la maquette.	25

3.1	Expériences faites sur le lac Léman avec deux bateaux distants de 17 km. Un bateau est équipé, une cloche de bronze frappée par un marteau articulé. Une bougie fixée au manche du marteau devant ce bateau est allumée en même temps. Dans un autre bateau, l'expérimentateur porte un cornet acoustique dirigé vers l'autre bateau.	28
3.2	Reginald A. Fessenden et l'oscillateur Fessenden.	29
3.3	Réponses impulsionnelles d'un ensemble de filtres réalisant le codage ERB des signaux audio.	32
3.4	Les étapes d'extraction des MFCCs.	34
3.5	Architecture du système de reconnaissance sonore.	35
3.6	Exemple de la classification k -ppv.	36
3.7	Spectres d'un événement sonore correspondant à un <i>goutte à goutte</i> enregistrés sous l'eau (en haut) et dans l'air (en bas).	39
3.8	Spectres d'un événement sonore correspondant à un frappement du métal enregistrés sous l'eau (en haut) et dans l'air (en bas).	40
3.9	Taux de reconnaissance en fonction du nombre de coefficients de codage en utilisant les codages MFCC sur la base de signaux long-termes.	41
3.10	Evolution des coefficients de codage ERB, LPC, LPCC et MFCC au cours du temps pour un signal long-terme (30s, fenêtres de 2048 points).	42
3.11	Evolution des coefficients de codage ERB, LPC, LPCC et MFCC au cours du temps pour un signal court-terme (1.6s, fenêtres de 1024 points).	43
3.12	Réseau de neurone utilisé dans le classifieur prédictif.	44
3.13	Taux de reconnaissance en fonction du nombre de cellules cachées en utilisant les codages LPCC sur la base de signaux long-termes.	45
3.14	Taux de reconnaissance en fonction du nombre de voisins en utilisant les codages MFCC sur la base de signaux long-termes.	45
3.15	Taux de reconnaissance des codes sur la base durable.	46
3.16	Taux de reconnaissance des codes sur la base transitoire	46
4.1	La première peinture panoramique en 1792.	49
4.2	Vue depuis le sommet de Lookout Mountain en 1864 par George N. Barnard.	49
4.3	Le tube de la caméra panoramique.	51
4.4	Les différents repères locaux du système de vision.	52
4.5	Le contour du miroir est utilisé pour calibrer la caméra.	53
4.6	Le profil du miroir utilisé dans nos études.	58
4.7	Un exemple de la réflexion des rayons par la surface du miroir.	58

4.8	(a) Les rayons réfléchis pour un profil du miroir. (b) La résolution de la distribution des rayons pour un profil du miroir avec une hauteur de 2 m.	59
4.9	L'axe du bateau est utilisé pour définir un système de coordonnées globales en supposant que la caméra est placée perpendiculairement à cet axe.	61
4.10	Les amers utilisés pour l'estimation de l'orientation.	62
4.11	Un calibrage complet du capteur catadioptrique fournit un mapping des pixels de l'image \mathbf{p}_i sur une ligne 3D $(\mathbf{m}_i, \mathbf{u}_i)$ et son intersection avec la surface définie par l'eau. \mathbf{P}_i peut être calculée si l'orientation relative du bateau par rapport à la surface de l'eau est connue.	62
4.12	Ajustement du plan sur le capteur calibré.	63
4.13	L'estimation des distances pour chaque pixel dans l'image.	64
4.14	Oscillation du bateau dans une mer agitée.	65
4.15	Modèle colorimétrique calculé à partir d'une image dans la base de référence inspiré par [Cha09].	66
4.16	Détection d'amers utilisant la signature colorimétrique.	67
4.17	Estimation de l'orientation du capteur omnidirectionnel via l'horizon.	68
4.18	(a) Image panoramique prise dans un lac . (b) Après la détection, la surface de l'eau est reconnue. (c) Les bruits sont supprimés en utilisant des techniques standard de morphologie mathématique. (d) Le contour horizon détecté.	69
4.19	Détermination du plan d'incidence lié à l'horizon	70
4.20	Expérimentation réalisée dans un lac avec un bateau.	71
4.21	Mouvements de la plateforme.	73
4.22	Détection de la position du bassin sur l'image pour simuler la position de l'horizon.	73
4.23	Les angles calculés par la caméra (ligne continue) et la centrale inertielle (ligne en pointillé).	75
4.24	Bateau à moteur utilisé.	76
4.25	Détection de la position de l'horizon sur une image prise en mer.	77
4.26	Orientation du bateau calculée par la caméra (ligne continue) et la centrale inertielle (ligne en pointillé) à chaque instant.	77
4.27	Une partie de la surface de la mer est hors de l'image.	78
5.1	Les trois niveaux de fusion : (a) fusion des données, (b) fusion des caractéristiques, et (c) fusion des décisions.	84
5.2	Un exemple d'un système de fusion comportant N capteurs-architecture : topologie parallèle.	84

5.3	Un exemple d'un système de fusion comportant N capteurs-architecture : topologie série.	85
5.4	Un exemple d'un système de fusion-architecture hiérarchique avec deux sous centres de fusion.	85
5.5	L'architecture de notre système de perception composée de plusieurs capteurs.	86
5.6	Changement de repère de $OXYZ$ à $O'X'Y'Z'$	89
5.7	Les intersections entre les rayons et le plan d'eau représentent les positions des objets calculées.	90
5.8	Les photos sont prises par la caméra panoramique avec les différentes orientations pour les distances variées. (a)-(c) sont les photos prises avec les bouées à deux mètres, et (d)-(f) pour des bouées à six mètres. (a) et (d) sont les cas stables et sans mouvement, les autres sont les cas en mouvement.	91
5.9	Le navire porte-conteneur de CMA CGM.	93
5.10	Les photos sont prises par la caméra panoramique. (a) est une photo du bateau de transport au loin, et (b) est une photo du bateau proche.	94
5.11	Un exemple d'un écho reçu par le sonar pour une seule direction.	96
5.12	Un exemple des signaux capturés par le sonar dans un bassin carré. (a) Les données bruitées reçues par le sonar avec une portée de 10 m (200 échantillons) et une résolution angulaire de 1.8°. (b) Les mêmes données présentées dans un système de coordonnées cartésiennes.	96
5.13	Un exemple de signal reçu pour $N_E = 200$ et une portée de 20 m. L'échelle des couleurs définissant la portée est réglée sur l'intervalle 13 - 40 dB.	97
5.14	Un exemple de calcul du paramètre P_E	97
5.15	Un exemple d'un balayage du sonar pour un secteur de 90 ° et une résolution angulaire de 1.8 °.	98
5.16	Un sonar balayant de 90 à 270 °, et détectant un obstacle. La fenêtre de la portée dynamique pour afficher l'échelle des couleurs est réglée entre 6 et 46 dB.	99
5.17	(a) Le bruit propre du sonar pour une seule détection. (b) Le même bruit affiché dans le système de coordonnées cartésiennes.	100
5.18	(a) Les données bruitées capturées par le sonar. (b) La détection d'obstacles.	101
5.19	Un exemple d'un balayage du sonar pour un secteur de 360 ° avec trois objets capturés.	101
5.20	Une bouée avec réflecteur détectable par le sonar et le capteur de vision est utilisée pour calibrer le sonar par rapport au système de vision.	103
5.21	Schéma de la fusion par pondération dans notre système de fusion.	103
5.22	Schéma de la fusion par estimation dans notre système de fusion.	105

5.23	Schéma de la fusion avec un filtre simple dans notre système de fusion.	106
5.24	Schéma de filtre de Kalman utilisé dans notre système de fusion.	107
5.25	(a) est les résultats obtenus par les données de la caméra. (b) est les résultats des données du sonar.	109
5.26	(a) les covariances obtenus par les données de la caméra. (b) les covariances des données du sonar.	110
5.27	(a) et (b) sont les résultats obtenus par les données de la caméra pour deux essais différents. (c) et (d) sont les résultats des données du sonar pour ces essais.	111
5.28	Les résultats obtenus par la méthode de fusion par l'estimation (F1) sur les deux essais correspondants.	112
5.29	Les résultats obtenus par la méthode de fusion par la pondération 1 (F2) sur les deux essais correspondants.	113
5.30	Les résultats obtenus par la méthode de fusion par la pondération 2 (F3) sur les deux essais correspondants.	113
5.31	Les résultats obtenus par la méthode de fusion avec un filtre simple (F4) sur les deux essais correspondants.	114
5.32	Les résultats obtenus par la méthode de fusion avec un filtre Kalman (F5) sur les deux essais correspondants.	114

LISTE DES TABLEAUX

4.1	Résultats des paramètres estimées	65
4.2	Erreurs estimées des données réelles	72
5.1	Estimation de la distance entre l'objet et la camera avec l'image panoramique en utilisant l'amer de 6 bouées	92
5.2	Estimation de la distance entre l'objet et la camera avec l'image panoramique en utilisant les informations de la centrale inertielle	92
5.3	Estimation de la taille d'un objet avec l'image panoramique en utilisant les informations de la centrale inertielle	93
A.1	La base de données transitoires	121
A.2	La base de données non transitoires	122

Introduction générale

Le premier véhicule autonome de surface en milieu marin a été développé pendant la seconde guerre mondiale pour des applications militaires. Ces dernières années, son utilisation s'est diversifiée, et touche maintenant de nombreux domaines comme le commerce, l'environnement, la recherche biologique marine et la robotique. Dans un cas général, les véhicules autonomes de surface sont utilisés pour les recherches et la surveillance de toute la partie immergée, émergée et aérienne de l'environnement marin.

L'autonomie est un critère majeur sans lequel de nombreuses applications seraient impossibles. En particulier, un grand nombre de missions irréalisables par des hommes peuvent être effectuées par les véhicules autonomes. C'est le cas notamment des opérations dangereuses pour l'homme, ou nécessitant une présence continue pendant des durées importantes en environnement hostile. Ces applications sont aussi diversifiées que la surveillance d'environnements marins, les campagnes de levés hydrographiques, l'étude d'animaux marins ou encore le recueil des données pour la météo marine.

Avant d'envisager le déplacement d'un véhicule autonome de surface, il est nécessaire d'assurer sa perception de l'environnement. Ceci consiste à observer, localiser et éviter les obstacles. Actuellement, les contraintes technologiques, la complexité de l'environnement naturel et la diversité des situations rencontrées rendent particulièrement complexe le développement d'un système de perception efficace et précis. La réalisation d'une plate-forme totalement autonome est donc un défi majeur. Cette thèse est une étape clef pour les véhicules autonomes de surface en milieu marin. Elle s'intéresse à l'acquisition multi-capteurs et au traitement des

signaux multi-modaux pour assurer la perception des véhicules autonomes de surface.

Des plate-formes utilisant un ou plusieurs capteurs existent pour des missions simples, mais leur capacité de perception est très limitée. L'objectif de cette thèse est de démontrer que la réalisation d'un robot autonome en milieu marin pour des missions complexes est possible grâce à des capacités de perception adaptées.

Ce rapport s'organise autour de cinq chapitres, chacun dédié à un aspect particulier de la perception pour un robot autonome en milieu marin.

Le premier chapitre présente un état de l'art des véhicules autonomes de surface depuis 1939, le début de la seconde guerre mondiale, jusqu'à maintenant. Ces travaux sont organisés selon trois axes : le temps, le pays et le domaine. Une cinquantaine d'exemples sont mentionnés pour les différents axes de développement.

Le second chapitre concerne la présentation du projet et la construction d'une plate-forme dédiée au milieu marin. Une description du projet et des partenaires est donnée au début de ce chapitre. Puis l'architecture du système et les périphériques utilisés sont présentés. Le choix des capteurs est très important pour ce système. Ils doivent être étudiés selon des critères de poids, de volume et de consommation d'énergie. Les capteurs choisis dans notre projet permettent de diriger le voilier par rapport à son environnement. Une plate-forme multicapteurs contenant deux hydrophones, une caméra panoramique, une centrale inertielle et un sonar est réalisée. Ces capteurs sont contrôlés par un ordinateur central. Chaque capteur fournit des mesures brutes pour une lecture directe des données en C sous linux sans passer par une interface intermédiaire. Chaque trame de signaux est ensuite datée par rapport à l'horloge commune d'un ordinateur afin d'assurer une synchronisation des données.

Le troisième chapitre traite des signaux sonores sous-marins pour la détection d'objets ou l'état de la mer. Nous proposons dans ce chapitre une étude des signaux transitoires et des signaux long-termes. La base de données utilisée contient treize signaux transitoires et trente signaux non transitoires. Tous les signaux dans la base sont réalisés à partir d'hydrophones avec une fréquence d'échantillonnage de 44 *kHz*. Quatre méthodes de codages (ERB¹, MFCC², LPC³ et LPCC⁴) sont appliquées sur notre base de données. Puis deux méthodes de classifications sont testées, la méthode *k*-ppv⁵ et la modélisation prédictive avec un réseau de type perceptron multicouche.

¹Equivalent Rectangular Bandwidth

²Mel Frequency Cepstral Coefficient

³Linear Predictive Coding

⁴Linear Prediction Cepstral Coefficients

⁵*k* plus proches voisins

Le quatrième chapitre expose l'utilisation d'un système de vision panoramique. Un historique sur les panoramas et le système de vision panoramique utilisé est d'abord présenté. Notre système de vision panoramique comporte une caméra Gigabit, un miroir et un tube acrylique transparent pour protéger l'ensemble des composants. La combinaison des ces différentes parties forme un capteur non central. Une phase de calibrage est nécessaire pour estimer la pose relative des composants. Une étape de traitement des images panoramiques basée sur la méthode de Hough est ensuite effectuée pour la détection de l'orientation du bateau. A partir de l'orientation du bateau par rapport à la surface de la mer, nous pouvons associer à chaque pixel de l'image panoramique une position 3D de la surface de la mer. Puis un critère colorimétrique est utilisé pour la segmentation et la localisation d'objets dans l'image. La détection de l'orientation du bateau est présentée à la fin de ce chapitre. La méthode utilisée est basée sur la localisation de l'horizon et sa projection sur la surface du miroir, qui permet de reformer un plan de l'horizon 3D et de déterminer l'assiette du bateau à partir d'une image panoramique.

Le dernier chapitre est consacré la fusion de données pour l'amélioration de la capacité de la détection d'obstacles. L'architecture de notre système de perception est présentée au début de ce chapitre. Notre système de fusion est constitué d'un bloc d'intégration qui comporte le système de vision et la centrale inertielle. Le bloc de fusion de données est effectué en couplant le bloc d'intégration à un sonar. Les informations concernant la présence d'obstacles seront obtenues en sortie de ce bloc. Une amélioration de la détection de l'orientation du bateau est ensuite présentée. Une centrale inertielle permettant de mesurer précisément l'orientation de bateau en temps réel est ajoutée dans cette étape. Cette méthode est plus simple et plus efficace que celle du chapitre précédent. L'utilisation d'un sonar est ensuite présentée pour compléter la détection d'obstacles immergés et émergés. Nous proposons cinq méthodes de fusion différentes pour la détermination de la position des obstacles détectés.

La conclusion résume les problèmes théoriques et techniques traités au cours de cette thèse. L'approche mise en oeuvre pour aboutir un robot autonome de surface en milieu marin est présentée. Les perspectives présentées soulignent notamment l'enrichissement de la base de donnée des signaux sonores pour perfectionner la détection des états de la mer et l'amélioration de la précision du sonar. Cette thèse n'est que le premier pas vers des systèmes de véhicule autonome de surface capable de réaliser des mission complexes de longue durée, mais elle montre sa faisabilité par le développement du système de perception.

CHAPITRE

1

Généralités sur les véhicules autonomes de surface

Sommaire

1.1	Introduction	5
1.2	Etat de l'art	6
1.3	Conclusion	14

1.1 Introduction

Dans le cadre de la recherche sur les véhicules autonomes de surface (ASV pour Autonomous Surface Vehicle), l'objectif est de détecter et d'éviter les obstacles dans un contexte maritime. Le développement de véhicules autonomes de surface (ASV pour Autonomous Surface Vehicle) ou (USV pour Unmanned Surface Vehicle) est inspiré par les défis de navigation et de communication des véhicules autonomes sous-marins (AUV).

1.2 Etat de l'art

Le développement d'ASV a commencé au début de la seconde guerre mondiale [Rob06, Ber08]. Pendant cette période, les Canadiens ont développé un premier ASV : une torpille de surface référencée COMOX [For04]. Ces anciens véhicules autonomes de surface ont été développés durant cette guerre comme des systèmes d'artillerie équipés de roquettes. Après la guerre, les applications d'USV ont augmenté : en 1946, l'United States Navy a utilisé des bateaux drônes pour recueillir les échantillons de l'eau radioactif après le souffle des bombes sur l'atoll de Bikini.

Pendant les années 50 et 60, l'US Navy a construit et utilisé un bateau de déminage télécommandé, dont elle s'est servie pour le test de tir de missiles. Le développement et l'utilisation de drônes ASV a continué et évolué au cours des années, le plus souvent pour réaliser des missions dangereuses. Les forces navales européennes ont commencé le développement des systèmes de nouvelle génération vers la fin des années 70 et le début des années 80 [Wer07]. Citons le système STANFLEX du Danemark, système composé d'armes et d'équipements, qui peut être installé sur un navire [Wikb], le système avec trois drones de Troïka allemand pour effectuer des missions difficiles, et le USV système SAM¹ II ACV [Koc] de Suède qui permet de réaliser le dragage des mines pour un véhicule de surface autonome ou télécommandé à distance.

En 1983, un engin semi-submersible est développé par l'International Submarine Engineering (ISE) au Canada : DOLPHIN²[Tho85], utilisé principalement comme engin téléguidé pour l'exploration ou la recherche de mines. La propulsion est assurée par un moteur diesel et ses dimensions ($7.3m \times 4.5m$) en font une plateforme capteur bien dimensionnée pour des applications dans des conditions de mer difficile. Plusieurs navires de ce type ont été construits et livrés par ISE à "Canadian Hydrographic Service" et à l'US Navy. Toutefois, peu de publications existent sur les aspects contrôle et sur les missions confiées à ce type d'engin. Dix ans après, au MIT, le premier navire robotisé ARTEMIS est réalisé par Manley qui a été utilisée pour étudier des architectures de commande, des systèmes de navigation et des acquisitions de données pour un véhicule autonome de surface en milieu marin. Ce véhicule a démontré avec succès la capacité de fonctionner de manière autonome et d'acquérir des données hydrographiques[Man97]. Une amélioration d'ARTEMIS (ACES [Man97]) a ensuite été développée. L'ACES offre une meilleure vitesse de croisière, une plus grande capacité de charge utile qui permet d'effectuer des missions plus longues, et une meilleure tenue de mer. La troisième version de ce véhicule autonome de surface AUTOCAT a été réalisée en 2000

¹Self-propelled Acoustic/magnetic Mine sweeper

²Deep Ocean Logging Platform with Hydrographic Instrumentation and Navigation

[MMCW00] (Fig. 1.1). Propulsé par deux moteurs électriques, il est utilisé pour cartographier le fond des océans avec une grande précision.



FIG. 1.1: Un catamaran automatique AUTOCAT de MIT aux USA.

Vers la fin des années 90, l'US Navy s'intéresse aux ASVs pour la reconnaissance et les missions de surveillance. Le Roboski, ayant une forme de motomarine, a été utilisé pour l'autodéfense de bateau, il sert encore maintenant comme véhicule de test pour réaliser l'acquisition des signaux en milieu marin. Durant la même période, Navtec Inc a développé OWL MK II (voir Fig.1.2) qui mesure 3 mètres de long et 1.5 mètre de large. Il est équipé d'une coque profilée qui a été modifiée pour augmenter la capacité de furtivité et alimenté par un système de propulsion de motomarine. Une version avec un sonar latéral et une caméra a été utilisée dans le Golfe Persique pour effectuer des missions de surveillance.

La recherche sur les USV en France commence en 1994, dans le cadre d'un ancien projet ENSR sur les relevés bathymétriques dans l'estuaire de la Gironde [LLS94]. Il mesure 5 mètres de long et utilise les équipements relatifs au sondage, à la navigation et à la sécurité. Il peut réaliser les missions en mode automatique pendant 5 heures avec un rayon d'action de 10 *km*. En 2003, des recherches de IBOAT commencent à Toulouse en France, c'est un bateau autonome équipé de capteurs qui mesure 2.4 mètres de long, 0.4 mètre de large, et 3 mètre de haut et pèse 35 *kg*. Les capteurs utilisés dans l'IBOAT comporte une boussole, un anémomètre, et un GPS qui permet de donner le trajet et la position du bateau. Mais le système de perception n'est pas conçu dans l'IBOAT. En Allemagne le projet MESSIN, démarre en 2000 [MBK00], autour du développement d'un catamaran à propulsion électrique pour des applications de mesures en eaux peu profondes. Dans le même temps, un projet européen débuté en 2000 à Lisbonne pour le développement de DELFIM[EP01, AOO⁺06,



FIG. 1.2: L'USV Owl MK II qui est utilisé pour la reconnaissance et les missions de surveillance.

[GSPC06](#)], se base également sur un catamaran. Le DELFIM a été conçu pour effectuer l'acquisition automatique de données marines et servir de relais entre un véhicule sous-marin et un bateau de soutien. Il mesure 3.5 mètre de long et 2 mètres de large. Le système de propulsion est constitué de deux hélices actionnées par des moteurs électriques. Un monocoque de 10 m CARAVELLA [[APA99](#)] a aussi été développé par le même laboratoire. Un système de sécurité basé sur un radar a été installé sur ce bateau qui permet de détecter et éviter les obstacles.

En Italie, l'université de Gênes développe depuis 2002 un programme de recherche pour l'étude de l'interface air-mer en Antarctique. Ces recherches ont conduit au développement de SESAMO, un catamaran à propulsion électrique qui mesure 2.4 mètres de long et 1.8 mètre de large. L'objectif de ce projet est de concevoir, développer et tester une plate-forme autonome pour l'acquisition des données hydrologiques. Cet ASV a été utilisé avec succès dans une mission de mesures en Antarctique en 2004 [[CBB⁺05](#)]. Dans la même année Curcio a développé une plate-forme de type Kayak sous le nom de SCOUT [[CLP05](#)] (Fig. 1.4). Cette plate-forme multicapteurs est équipé d'un système de propulsion, un GPS, une boussole et un système de communication pour l'évitement de collision. L'objectif de ce projet est de réaliser un véhicule autonome de surface performant avec une grande vitesse et à faible coût. En Angleterre, un projet de l'Université de Plymouth nommé SPRINGER [[NSC06](#)](2006), utilise un catamaran d'environ 3 m pour des applications liées à la recherche de polluants et à des relevés environnementaux et hydrographiques en rivière, bassin ou zones côtières (Fig. 1.3).



FIG. 1.3: SPRINGER est le catamaran automatique de 3m.

Au Japon, Yamaha a développé deux USVs, l'UMV-H³ et l'UMV-O⁴ [EYS⁺04]. L'UMV-H est composé d'une coque de bateau à grande vitesse de 4.44 m de long qui lui permet de se déplacer sur l'eau et de faire des observations. Celui-ci est propulsé par un hydrojet. Le système utilise un GPS pour localiser sa position avec une centaine de coordonnées. Il peut naviguer automatiquement avec une caméra et un sonar sous marin. L'UMV-H peut également apprendre et reproduire un trajet effectué par la navigation manuelle.

Depuis 2000, de plus en plus de projets portent sur le développement de véhicules autonomes de surface. Des compétitions ont été organisées pour stimuler le développement de voiliers autonomes en Europe :

- Microtransat : Toulouse 2006, Aberystwyth 2007, Transatlantic 2010
- Sailbot : Canada 2006, USA 2007, USA 2009

Le projet ROAZ a été démarré par LSA⁵ au Portugal [LSA]. ROAZ est un petit catamaran pour des opérations dans des rivières et des lacs. Il mesure 1.5 mètres de long, et 1 mètre de large et peut supporter 50 kg d'équipements. Une version améliorée le ROAZ II, est conçue pour des opérations dans l'océan. Il est plus grand que le ROAZ, avec une longueur de 4.5

³Unmanned Marine Vehicle High-Speed

⁴Unmanned Marine Vehicle Ocean type

⁵Autonomous Systems Laboratory



FIG. 1.4: Trois SCOUT après une sortie dans l'eau.



FIG. 1.5: ROAZ de LSA en portugal.

mètres et une largeur de 2.2 mètres. Depuis 2008, le OASIS ⁶ ASV développé par NOAA ⁷ a déjà été déployé dans l'océan Atlantique nord et le Chincoteague Bay [DPE⁺]. C'est un véhicule autonome de surface utilisant l'énergie solaire pour réaliser les missions de longue durée. De même en 2008, "Tianxiang One", un ASV développé en chine, est le premier bateau autonome appliqué dans la météorologie. "Tianxiang One" est également équipé d'un système d'alimentation fiable, permettant des déplacements de centaines de kilomètres et d'une durée de 20 jours [YPSP10]. En 2010, STE ⁸ a lancé le Venus. Cet USV est basé sur une plateforme de 9 m, et intègre le GNC ⁹ dans son système de capteurs.

Peu de projets concernent les ASVs à propulsion vélique. Le plus ancien étant le projet ATLANTIS de G.H. Elkaim de l'Université de Floride (Fig. 1.6). C'est un catamaran de 7 m avec un mât aile de 5.37 mètre de haut [Elk01, EWK06, EB08]. Son mât aile a été construit en trois parties :

- La partie du bas : Elle comporte les montages électronique (des batteries, un ballast), un microcontrôleur de l'anémomètre et un magnétomètre.
- La partie du milieu : Elle contient aussi un microcontrôleur et les modules électroniques qui permettent de contrôler les deux bômes attachés.
- La partie du haut : Elle utilise un microcontrôleur et les modules électroniques pour contrôler le bord de fuite en haut.

L'objectif de ce projet est de démontrer une avance de haute précision (1 mètre) pour contrôler un véhicule de surface à propulsion vélique en milieu marin.

Un autre USV à propulsion vélique a été créé par l'Université de Wales, Aberystwyth en 2006 [Nea06]. Le prototype construit est un monocoque de 1.5 m avec un mât aile de 1.3 mètres, l'USV final mesure 3 mètres de long (Fig.1.7). Le prototype à petite échelle a été conçu comme une preuve de faisabilité pour un véhicule autonome de surface à propulsion vélique. Son but était de vérifier la consommation d'énergie et justifier la capacité à réaliser des missions de longue durée en mode automatique. Mais aussi de tester la capacité de navigation dans un environnement naturel avec un large spectre de vent et la capacité de charge utile.

Plus récemment, en 2010, un projet porte sur un monocoque. Celui-ci fait 4.2 m de long avec un mât aile de 5 m, et est conçu pour fonctionner avec un vent de 7-10 km/h [RvE10]. Il est équipé d'un GPS, d'une boussole, et d'un anémomètre dans le système de contrôle qui permettent de contrôler la navigation et acquérir des données. Un capteur pour surveiller la qualité de l'eau, les capteurs pour mesurer la température de l'eau et de l'air, un capteur de

⁶Ocean Atmosphere Sensor Integration System

⁷National Oceanic and Atmospheric Administration

⁸Singapore Technologies Electronics

⁹Guidance Navigation Control



FIG. 1.6: L'ASV ATLANTIS de l'Université de Floride.



FIG. 1.7: Le prototype d'un USV de l'Université de Whales.

salinité et un capteur de Ph aquatique ont été installé dans l'USV pour mesurer les paramètres de la surface de l'eau.

La perception est un aspect important dans le domaine des ASV. Celle-ci concerne la localisation d'obstacles et la décision de trajectoire qui permet d'éviter les collisions. A la fin des années 70, le système ARPA ¹⁰ [IMO79] a été développé pour le suivi d'obstacle et la planification de trajectoires sécurisées. Il signifie le début des recherches dans le domaine de la perception. C'est un équipement associé au radar de navigation qui permet de suivre la trajectoire d'un objet et calculer la vitesse et la distance de cet objet. L'observation de données n'est valable qu'avec un objet stable, c'est à dire que l'obstacle observé doit conserver le même cap et la même vitesse.

Au début des recherches de la perception, les USVs ont été équipé d'un seul capteur. Par exemple, le DOLPHIN d'ISE au Canada, présenté précédemment, a été équipé d'un sonar pour réaliser la détection de l'environnement. Dix ans après, le système multi-capteurs avec les capteurs classiques (sonar, radar, loch, sondeur, GPS, boussole) est de plus en plus utilisé. L'ENSR en France, présenté dans ce chapitre, a embarqué un radar, un loch, un GPS et un sondeur pour assurer la sécurité du bateau.

Depuis les années 2000, les capteurs (microphone/hydrophone, caméra) typiquement utilisés dans le domaine des véhicules terrestres autonomes ont été aussi utilisés par les véhicules autonomes de surface en milieu marin. En 2004, Snyder et al. ont développé un système de perception pour un bateau autonome. Ce système contient un réseau circulaire de caméras et une caméra tournante qui permet de réaliser la perception de l'environnement et la localisation [SMH⁺04]. En 2006, L'DECE ¹¹ de Virginia Tech. a utilisé un monocoque de 3 m [SGR⁺06, GXR⁺08, Gon08], avec un gyromètre de 3 axes, un GPS, un capteur de profondeur et une caméra panoramique, pour réaliser la détection de la ligne de séparation ciel-mer, afin d'estimer le trajet du bateau, mais cet ASV ne comporte pas de capteur pour réaliser la détection d'obstacle sous l'eau, ce qui est dangereux quand aucune entrave n'apparaît à la surface. Dans la même année Cooney et al. ont conçu un ASV équipant un système acoustique de suivi [CSA⁺06]. Ce système comporte un transducteur, un transpondeur et trois hydrophones qui permet de localiser la cible acoustique immergée. Les positions des trois hydrophones sont importantes, il faut séparer chaque hydrophone de 2.4 mètre. La vitesse de la détection est très limitée. En 2009, un catamaran de 5 m est développé par ASL ¹² en Australie [DGU09], équipé de plusieurs capteurs pour la détection d'obstacles : un GPS, une boussole, un capteur de profondeur, un scanner laser, une caméra et un sonar. La caméra

¹⁰Automatic Radar Plotting Aid

¹¹Department of Electrical and Computer Engineering

¹²Autonomous Systems Laboratory

utilisée ici, est une caméra normale avec un champ limite, il y a donc toujours une zone morte dans laquelle la détection est impossible.

Les contraintes technologiques et la complexité de l'environnement naturel rendent particulièrement complexe le développement d'un système de perception efficace et précis. La réalisation d'un véhicule de surface totalement autonome en milieu marin est donc un défi majeur.

1.3 Conclusion

L'ASV s'est développé pour réaliser divers missions dans l'environnement marin. Dans ce chapitre, nous avons présenté une vue globale sur les recherches déjà réalisées en ASV. Dans le chapitre suivant nous faisons une description sur le projet ASAROME dans lequel j'ai travaillé pendant trois ans de thèse et le système embarqué utilisé dans nos études.

CHAPITRE

2

Systeme embarqué et capteurs

Sommaire

2.1	Description du projet	16
2.2	Présentation du système	17
2.2.1	Architecture du système	18
2.2.2	Montage des capteurs	18
2.3	Capteurs de perception	20
2.3.1	Capteur de vision omnidirectionnelle	21
2.3.2	Hydrophone	22
2.3.3	Centrale inertielle	23
2.3.4	Sonar	24
2.3.5	Logiciel de manipulation	24
2.4	Conclusion	26

L'environnement maritime propre aux véhicules autonomes de surface est souvent complexe et pose de nombreux défis pour la recherche. Par exemple, dans une mer agitée en condition naturelle, la perception de l'état de la mer est difficile à mettre en oeuvre. Le choix

des capteurs du système doit s'opérer en considérant les problèmes liés à l'environnement marin.

Ce chapitre donne une vue d'ensemble du système développé. Premièrement, nous présentons le projet ASAROME ¹ et les partenaires dans le groupe. Ensuite nous décrivons l'architecture du système utilisé. Enfin, les capteurs de perception sont présentés.

2.1 Description du projet

Cette thèse se place dans le cadre d'un robot autonome de surface. Nous avons travaillé dans le projet ASAROME démarré en 2008 qui se situe dans le domaine de la robotique autonome. Son objectif est de réaliser un petit voilier autonome pour des mesures et observations en milieu marin pendant une longue durée.

Le bateau autonome a été réalisé sur la base d'une coque de Miniji de 3.6 m de longueur et 300 kg dont 180 kg de lest. Il est équipé d'un mât aile de 3 m de longueur, d'un ensemble d'actionneurs et capteurs permettant une télé-navigation à distance, ainsi que des capteurs classiques de navigation, présentés à la partie 2.2 de ce chapitre.

Ce projet regroupe cinq laboratoires français, académiques et industriel. Ils ont chacun des compétences complémentaires :

- *ROBOSOFT* : Est un industriel français qui est le porteur du projet. Il apporte son expertise et son savoir faire dans le développement de solutions adaptées à la robotique mobile.
- *Institut des Systèmes Intelligents et de Robotiques (ISIR)* : L'ISIR intervient en matière de perception et de commande en robotique.
- *Laboratoire de Mécanique des Fluides (LMF)* : Le LMF contribue sur les aspects hydrodynamique navale et apporte ses compétences sur la complémentarité simulation numérique/modélisation expérimentale.
- *Institut Jean Le Rond d'Alembert (IJLRA)* : La contribution du IJLRA porte sur les aspects modélisation et simulations numériques en aérodynamique.
- *Institut Français de Navigation (IFN)* : L'IFN est l'expert maritime de ce projet. Il fournit l'interface avec l'utilisation d'engin de surface autonomes.

Cette thèse se déroule au sein de l'équipe Systèmes Intégrés Mobiles et Autonomes (SIMA) à l'ISIR.

¹Autonomous SAiling Robot for Oceanographic MEasurements

L'ISIR est un laboratoire de recherche de l'Université Pierre et Marie Curie (UPMC) associé au Centre National du Recherche Scientifique (CNRS), département Sciences et Technologies de l'Information et de l'Ingénierie (STII). L'ISIR a été créé au 1er Janvier 2007 par le regroupement de trois équipes de l'UPMC relevant de la nouvelle UFR des Sciences de l'Ingénieur :

- Le Laboratoire de Robotique de Paris dans son ensemble (LRP)
- Le groupe Perception et Réseaux Connexionnistes (PRC) du Laboratoire des Instruments et Systèmes d'Ile de France (LISIF)
- L'équipe AnimatLab du Laboratoire d'Informatique de Paris 6 (LIP6)

Ce regroupement vise à fédérer des compétences complémentaires et synergétiques dans les domaines de :

- La modélisation des systèmes et des interactions complexes
- Les commandes robustes des systèmes, la téléopération et l'identification
- La perception et le Traitement du Signal (vision, toucher et parole)
- Les architectures de contrôle : évolution, développement et apprentissage



FIG. 2.1: Institut Systèmes Intelligents et Robotiques.

L'équipe SIMA s'intéresse au développement de robots à haute mobilité et haute autonomie, spécialisés dans les domaines de la conception et la commande de robots mobiles en évolution dans des milieux naturels, la vision omnidirectionnelle et distribuée, et la perception active multimodale pour la navigation autonome.

2.2 Présentation du système

Le voilier autonome est équipé de plusieurs capteurs lui donnant la capacité de percevoir son environnement. La multiperception est nécessaire pour la sécurité et l'efficacité optimale de la navigation.

2.2.1 Architecture du système

La sécurité des bateaux est assurée par une détection d'obstacle rapide en utilisant une approche multisensorielle. Cette section présente l'architecture multi-capteurs intégrée sur le bateau avec la liste des capteurs embarqués présentés en détail ci-dessous : Anémomètre et girouette, appareil photo panoramique, capteur d'angle de bôme, capteur d'angle de barre, centrale inertielle, hydrophone, Sonar, et Global Positioning System (GPS) (Fig. 2.2).

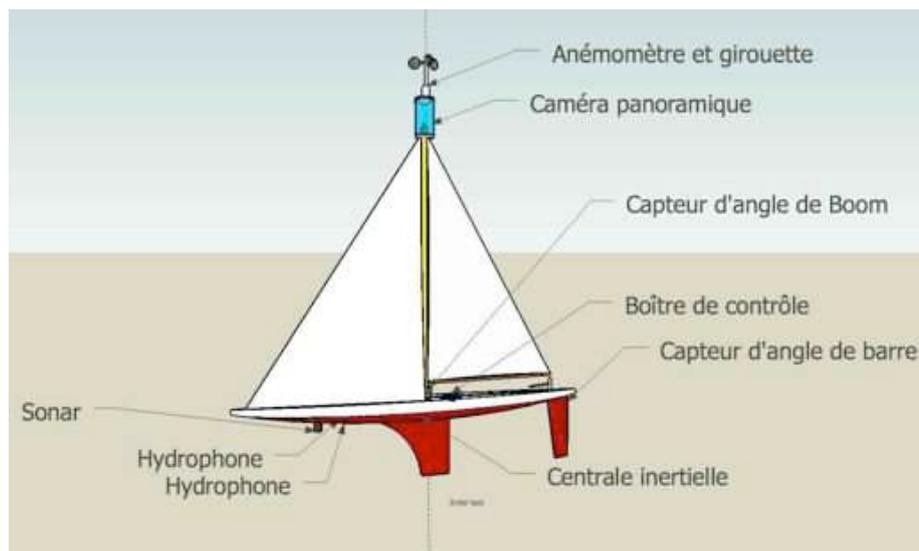


FIG. 2.2: Schéma du bateau et la positionnement des capteurs.

2.2.2 Montage des capteurs

Notre système est composé de cinq parties (Fig. 2.3) : Un bloc de perception, un bloc de navigation, un bloc proprioceptif, un bloc de fusion et un bloc de commande.

Le bloc des capteurs de navigation est un ensemble regroupant GPS ², anémomètre et girouette ³. Le bloc perception comprend une centrale inertielle ⁴, un sonar ⁵, deux hydrophones ⁶ et une caméra panoramique ⁷. S'ajoute à cela un capteur d'angle de bôme ⁸, d'angle

²Référence : Xsens MTi-G, <http://www.xsens.com>

³Référence :Furuno CV3F, <http://www.furuno.fr>

⁴Référence : Xsens Mt9, <http://www.xsens.com>

⁵Référence : Trittech Micron DST, <http://www.tritech.co.uk>

⁶Référence : Aquarian H2A, <http://www.aquarianaudio.com>

⁷Référence : Prosilica GE1350C, <http://www.datvision.com>

⁸Baumer IVO codeur absolu, <http://www.baumerivo.fr>

de barre⁹, et des capteurs de type proprioceptif.

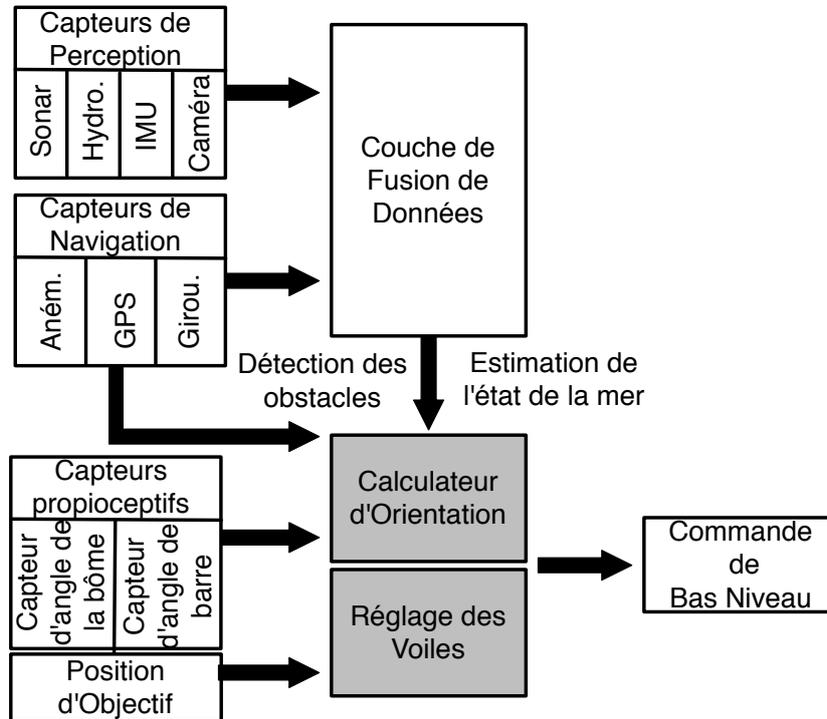


FIG. 2.3: L'architecture du système du projet ASAROME.

La couche de fusion de données récupère l'information à partir des blocs de navigation et perception pour localiser les obstacles et calculer leur position par rapport au voilier.

Les blocs "calculateur d'orientation" et "réglage des voiles" comportent les algorithmes décisionnels qui prennent en compte les caractéristiques intrinsèques du voilier pour calculer la meilleure direction et la position de la voile. Ces valeurs sont envoyées au bloc de commande de bas niveau qui contrôle la position de la voile et l'orientation du bateau. Le gouvernail est actionné par un vérin LINAK¹⁰ linéaire et le bôme est actionné par un moulinet de pêche électronique¹¹.

Le voilier a trois ordinateurs de bord afin d'optimiser l'acquisition et le traitement des données. Etant donné que le processus (Perception des capteurs, acquisition et traitement des données) implique une utilisation intensive des processeurs, nous utilisons un PC 104 dédié¹². Ce PC comporte un processeur ATOM N270 embarqué. Ce processeur est directement relié à la caméra IP à l'aide d'un port Ethernet Gigabit, au sonar par le port série et aux

⁹Référence : LINAK 300N LA12, <http://www.linak.de>

¹⁰Référence : LINAK 300N LAR, <http://www.linak.de>

¹¹Référence : Kristal fishing XL630, <http://www.kristal-fishing.fr>

¹²Référence : ARBOR mini ITX Box PC, <http://www.arbor.com.tw>

hydrophones par les prises jack standards de 3.5 mm. Il est également utilisé pour la fusion de données. Un autre PC 104 ¹³ avec ATOM N270 est utilisé pour le routage et la commande de haut niveau.

Un troisième PC 104 ¹⁴ embarqué avec un processeur RENESAS ¹⁵ est utilisé pour centraliser toutes les informations de la navigation et des capteurs proprioceptifs. La transmission des données de la centrale inertielle et du GPS se fait sur un bus série. Le reste des données des capteurs sont transmises sur un bus CAN. La commande actionnant le bateau est également envoyée à l'ordinateur par le biais du bus CAN. Ensuite, les communications entre ordinateurs, envoi et récupération de données de navigation telles que le cap désiré et la position de la voile, sont faites par UDP sur un réseau Ethernet.

Enfin, l'alimentation est basée sur une batterie (12V, 60A), qui est chargée par une éolienne ¹⁶ et un panneau solaire de 0,5 m² ¹⁷.

2.3 Capteurs de perception

Pour des véhicules autonomes de surface, le choix de capteurs pour construire le système est important. Les caractéristiques des capteurs, telles que le poids, le volume et la consommation d'énergie, doivent être choisies en fonction des conditions d'environnement et des paramètres physiques du bateau. Dans notre cas, le voilier est petit, donc nous devons choisir des capteurs qui consomment peu d'énergie et qui sont le plus petit et le plus léger possible.

Le bateau final du projet ASAROME ne pouvant pas être disponible tout de suite, une maquette embarquant les différents capteurs destinés à être intégrés sur le navire a été conçue afin de fournir un support physique autorisant les tests et autres validations des algorithmes de perception. Chaque capteur fournit des mesures sous leur forme la plus brute possible grâce à une lecture directe des données en C sous linux, c'est à dire sans passer par une interface intermédiaire. Chaque trame des signaux est ensuite datée par rapport à l'horloge d'un ordinateur afin d'assurer une synchronisation des données. Dans le détail, la maquette est composée de :

- un PC embarqué placé dans un boîtier étanche assurant la centralisation des relevés fournis par les différents capteurs et le traitement des données,

¹³Référence : LIPPERT, <http://www.lippertembedded.de>

¹⁴Référence : EMTRION, <http://www.emtrion.com>

¹⁵Référence : RENESAS SH7781@400mhz, <http://www.renesas.com>

¹⁶Référence : ATMB 1000 50W-12V, <http://atmbmarine.com>

¹⁷Référence : SUNSET AS60, <http://www.sunset-solar.de>

- un capteur de vision omnidirectionnelle fixé sur un mât central,
- un sonar placé à coté du capteur omnidirectionnel immergé,
- deux hydrophones fixés sur le bâti,
- une centrale inertielle 3 axes.

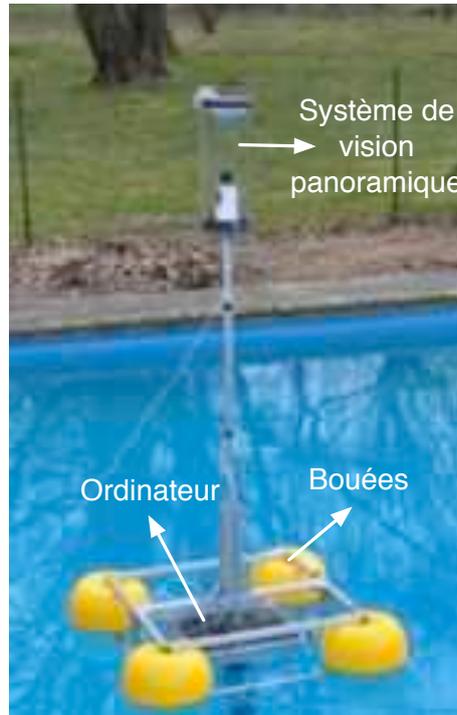


FIG. 2.4: La maquette utilisée dans le test de la perception.

2.3.1 Capteur de vision omnidirectionnelle

Un capteur de vision panoramique pour la détection d'obstacles et l'analyse de l'état de la mer est utilisé ici. Il est constitué d'une caméra numérique haute définition (Caméras industrielles Gigabit Ethernet 1.4 Méga pixels) couplée avec un miroir assurant la vision panoramique. Le capteur est dit catadioptrique.

Nous utilisons une caméra prosilica, tous les programmes sont écrits avec la bibliothèque PvAPI [Pro09] en langage C. Les images capturées sont encodées en TIFF, avec une taille de 1360×1024 pixels. Pour faire l'économie d'espace de stockage et aussi améliorer la vitesse de traitement, nous réduisons la taille d'image à 1024×1024 pixels. La communication entre l'ordinateur et la caméra est réalisée par le TCP/IP, afin d'établir la connexion, nous fixons l'adresse IP de la caméra : **169.254.21.1**.



FIG. 2.5: Caméra gigabit.

Le principe d'une vision panoramique a été retenu pour des raisons évidentes de perception globale et rapide de l'environnement immédiat de la structure. Ce capteur est placé au sommet d'un mat afin de faciliter une vision stéréoscopique pour faciliter la détection précise des obstacles de surface.

2.3.2 Hydrophone

Deux hydrophones sont utilisés pour l'analyse acoustique de l'état de la mer (Hydrophones H2a-3). L'hydrophone est un transducteur électroacoustique destiné à transformer, dans les liquides, des oscillations acoustiques en oscillations électriques.



FIG. 2.6: Hydrophone utilisée dans le projet.

Ces hydrophones utilisent les prises jack standards de 3.5 mm, le programme d'acquisition est écrit aussi en C utilisant la bibliothèque ALSA ¹⁸ [ALS09]. Les données sont enregistrées avec une fréquence d'échantillonnage de 44.1 kHz.

¹⁸Advanced Linux Sound Architecture

Un hydrophone placé à une distance de 50 *cm* à quelques mètres sous la surface de la mer permet des applications traditionnelles d'écoute du milieu marin. Après analyse du bruit acoustique, il fournit une estimation de l'état de la surface de la mer (agitation) et du taux de précipitation en temps de pluie.

2.3.3 Centrale inertielle

La maquette est équipée d'une centrale inertielle permettant un relevé rapide des différents paramètres dynamiques de la structure : accélérations et vitesses angulaires dans les 3 axes grâce à la présence d'un accéléromètre trois axes, d'un gyroscope trois axes et d'un magnétomètre à trois dimensions. Les centrales inertielles sont largement utilisées à bord des navires, d'aéronefs, de missiles ou de véhicules spatiaux pour assurer leur stabilité. Dans notre recherche, la centrale inertielle de XSENS MT9 (Fig. 2.7) est utilisée et fixée dans la boîte étanche pour mesurer l'assiette du bateau.

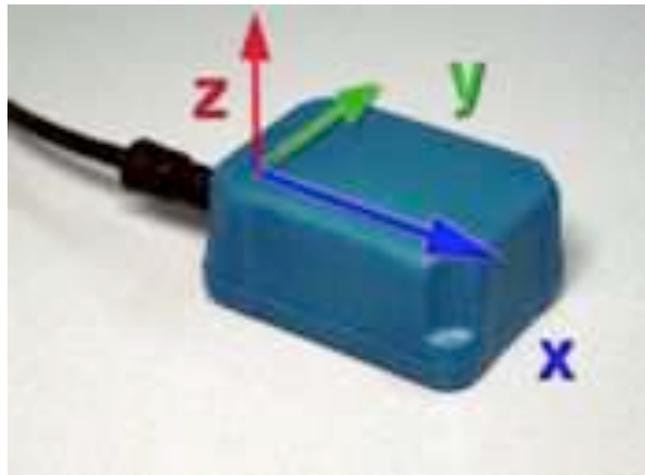


FIG. 2.7: La centrale inertielle de XSENS utilisée sur la maquette.

Le programme d'acquisition de cette centrale inertielle est fait avec la bibliothèque libm-tobjet¹⁹ [Xse03] fournie avec le capteur. Nous pouvons obtenir l'orientation absolue précise du capteur dans l'espace à trois dimensions à partir des données du gyroscope, de l'accéléromètre et du magnétomètre en temps réel. L'algorithme peut être expliqué comme étant la fusion des capteurs où la mesure de la gravité (accéléromètre) et le nord magnétique (magnétomètre) compensent les erreurs d'intégration en les minimisant (Fig. 2.8). Les données de sortie sont reçues sous la forme de quaternions avec une vitesse de 10 *trames/s*. Elles sont définies comme l'orientation entre le système de coordonnées du référentiel corporel et lequel

¹⁹library of Motion Tracker object

du référentiel terrestre. Le système de coordonnées de référentiel terrestre est utilisé comme le système de référence, qui est défini par un système de coordonnées cartésiennes :

- La coordonnée X est positive et indique vers le Nord magnétique local.
- La coordonnée Y suit un système de coordonnées cartésiennes.
- La coordonnée Z est positive et indique la verticale.

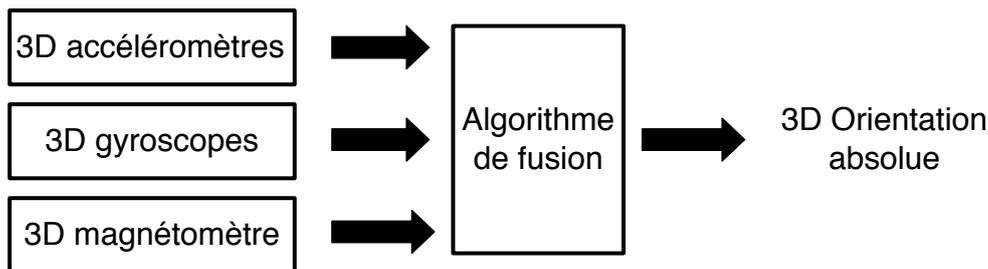


FIG. 2.8: Fusion des trois types de capteurs pour calculer l'orientation absolue précise.

2.3.4 Sonar

Le sonar utilise les propriétés particulières de la propagation du son dans l'eau pour détecter et situer les objets sous l'eau. Les sonars peuvent être actifs (émission d'un son et écoute de son écho) ou passifs (écoute des bruits). Dans la pratique, les sonars sont volumineux et lourds. Comme notre voilier est petit (3.6 m de longueur, 300 kg dont 180 kg de lest), nous utilisons ici un sonar actif de taille minimale (sonar Micron de Tritech) de dimension 5.6 cm × 7.85 cm à impulsions ultrasonores permettant la détection d'obstacles éloignés (Fig. 2.9).

Comme pour les autres capteurs, le programme d'acquisition est écrit en C. Le sonar est connecté avec l'ordinateur par le port série RS-232, et utilise le SeaNet RS-232 Protocol [Tri09]. La distance est obtenue par la mesure du temps écoulé entre l'émission et la réception de l'écho, la vitesse de propagation du son dans l'eau de mer étant égale à environ 1500 m/s. La vitesse de détection est de 36 °/s.

2.3.5 Logiciel de manipulation

Pour faciliter la manipulation et surveiller le processus d'acquisition des données, nous avons créé une plate-forme (Fig. 2.10). Elle comporte deux parties principales : une partie d'affichage des données reçues pour tous les capteurs, et une partie de contrôle, où nous pouvons démarrer/arrêter l'acquisition des données à l'aide de l'interface.



FIG. 2.9: Sonar Micron de Tritech.

Cette plate-forme est installée avec un ordinateur portable sous Windows. Le portable et l'ordinateur de perception sont connectés par un réseau local, la communication entre eux est réalisée par UDP avec une distance maximale de 20 m.



FIG. 2.10: Plate-forme utilisée pendant le test de la maquette.

2.4 Conclusion

L'ASV est développé pour réaliser diverses missions dans l'environnement marin. Dans ce chapitre, nous avons fait une description du projet ASAROME dans lequel j'ai travaillé pendant mes trois ans de thèse. Ensuite nous avons présenté le schéma du bateau et l'architecture du système que nous utilisons dans le projet. Enfin, la partie plus importante concernait le choix des capteurs. Dans notre cas, nous avons comparé différents capteurs selon leurs poids, volumes et consommations d'énergie, afin de les intégrer dans la maquette en respectant le mieux possible le cahier des charges (faible poids, faible consommation électrique, faible volume). En les couplant à des algorithmes de traitement du son et d'image, ils leur sera alors possible de détecter de façon fiable et précise des obstacles et de l'état de la mer.

Dans le chapitre suivant, nous abordons la reconnaissance des signaux sonores sous marin pour une méthode de modélisation prédictive utilisant les réseaux de neurones.

CHAPITRE

3

Traitements des signaux sonores

Sommaire

3.1	Introduction	28
3.2	Représentation des signaux	31
3.2.1	ERB	32
3.2.2	LPC	32
3.2.3	LPCC	33
3.2.4	MFCC	33
3.3	Classification des signaux	34
3.3.1	Modèle prédictif	35
3.3.2	Classification k -ppv	36
3.3.3	Décisions locales et globales	37
3.4	Expérimentations	38
3.4.1	Codage	39
3.4.2	Paramètres des classifieurs	41
3.4.3	Résultats	42
3.5	Conclusion	44

3.1 Introduction

Le son sous-marin est utilisé par les animaux marins depuis des millions d'années. Aristote est le premier qui a noté que le son pouvait être entendu dans l'eau ainsi que dans l'air. La science de l'acoustique sous-marine a fait ses débuts en 1490 avec Leonardo Da Vinci [Rob83]. En 1826, Daniel Colladon et Charles Sturm ont fait un développement sur l'acoustique sous-marine dans le lac Léman. Ils ont mesuré le temps écoulé entre un flash de lumière et le son de la cloche de bronze en utilisant un cornet acoustique sous-marin [CM77]. Ils ont ainsi mesuré une vitesse du son de 1435 m par seconde sur une distance 17 km , fournissant la première mesure quantitative de la vitesse du son dans l'eau. Dans la même période, un physicien français, François Sulpice Beudant a mesuré la vitesse du son avec une moyenne de 1500 m par seconde dans la mer, ce qui est très proche de la vitesse théorique prévue.

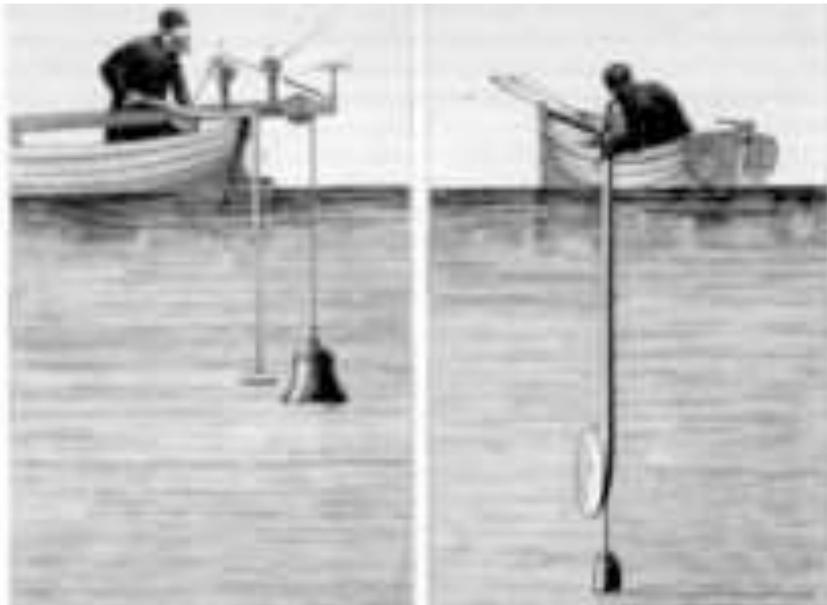


FIG. 3.1: Expériences faites sur le lac Léman avec deux bateaux distants de 17 km . Un bateau est équipé, une cloche de bronze frappée par un marteau articulé. Une bougie fixée au manche du marteau devant ce bateau est allumée en même temps. Dans un autre bateau, l'expérimentateur porte un cornet acoustique dirigé vers l'autre bateau.

Au début des années 1900, Thomas Edison et ses collaborateurs ont installé un téléphone dans une boîte étanche. Ils fabriquent ainsi le premier hydrophone permettant de recevoir des signaux provenant d'une cloche sous-marine. Malheureusement, cet hydrophone captait aussi le bruit de fond, y compris celui du moteur du bateau, des éclaboussures d'eau et des poissons. Il était très difficile d'entendre le son de cloche.

Le naufrage du paquebot Titanic le 14 avril 1912, a encouragé les recherches scientifiques dans la domaine de l'acoustique sous-marine. Un mois après, L. R. Richardson a développé une méthode pour déterminer les distances entre un bateau et un objet en utilisant le son et les échos sous l'eau. Entre 1912 et 1914, des dizaines de brevets ont été accordés en écholocalisation en Europe et aux États-Unis. Reginald A. Fessenden a conçu un hydrophone amélioré qui ressemblait beaucoup à un haut parleur de haute puissance sous-marin actuel et aussi connu sous le nom d'oscillateur Fessenden.



FIG. 3.2: Reginald A. Fessenden et l'oscillateur Fessenden.

Le développement de l'acoustique sous-marine a continué au même rythme au cours de la Première Guerre mondiale, poussé par les premiers déploiements à grande échelle de sous-marins. D'autres progrès en acoustique sous-marine ont permis le développement de mines acoustiques. En 1919, le premier article scientifique sur l'acoustique sous-marine est publié par H. Lichte [Lic19], un chercheur allemand. Il décrit la réfraction des ondes sonores produit par la température et les gradients de salinité dans l'océan. Pendant les années 1920, le sondeur est développé [Ric70]. Et dans les années 1930 des systèmes sonar comportant des transducteurs piézoélectriques à partir de matériaux synthétiques ont fait leur apparition. Pendant la Seconde Guerre mondiale, la technologie de sonar actif était bien développée, et jouait un rôle très important. Après la Seconde Guerre mondiale, le développement de systèmes sonar a été poussé principalement par la guerre froide, ce qui a entraîné des progrès dans

la théorie et la pratique de l'acoustique sous-marine, aidé par les techniques informatiques. En 1954, Middleton [MM54] et Peterson [PBF54] qui travaillent encore sur le traitement des signaux acoustiques sous-marins aujourd'hui, ont présenté certains principes fondamentaux de traitement statistique du signal. A la fin des années 1950, une nouvelle impulsion a été donnée par l'apparition de sous-marin nucléaire. Robert Price a conduit, pendant les années 1960, le développement de SUCOM, un système de transmission pour les données acoustiques sous-marine. Dans la fin des années 1960 et début des années 1970, des études expérimentales importantes de propagation acoustique, ont été menées pour les sonars actifs et passifs [SB66, WB71]. Beaucoup de recherches ont commencé à se concentrer sur l'adaptation et la résolution du traitement du signal acoustique, dans les années 1970 et début des années 1980 [PBC77, BP77, AB83]. Depuis les années 1980, la classification et la reconnaissance des signaux sous marin sont étudiées. Bouachache [BR84, BS90] et Yen [Yen98] ont utilisé la distribution de Wigner-Ville et la distribution croisée de Wigner-Ville pour réaliser la détection et la classification de son sous l'eau ; Chen [Che85] a utilisé le processus autorégressif sur la segmentation du signal ; le processus autorégressif d'ordre supérieur est apparu dans [LFAM88] ; Parks et Weisburn [PW92] ont trouvé qu'une perception par analyse des fréquences offre un avantage pour la classification des sons des baleines. Il y a aussi des approches basées sur l'analyse temps-fréquence [OLBMP02] et sur les modèles Markov [KCP94]. Tucker et Brown [TB05] ont proposé en 2005 une autre idée pour classer des sons transitoires sous-marins enregistrés par des sonars passifs. Ils proposent de considérer des caractéristiques acoustiques perceptuelles, c'est à dire qui encodent des informations que les humains ont l'habitude d'utiliser dans des tâches de classification de signaux transitoires.

Dans le domaine du traitement des signaux sonores, nous nous intéressons plus particulièrement aux données issues de capteurs sonores sous-marins avec pour objectif la détection d'engins motorisés proches et lointains [GG09]. Dans ce cadre, nous proposons une étude comparative d'algorithmes de codage et de classification habituellement utilisés dans le domaine audio pour la classification d'événements sonores sous-marins (bruits liés aux conditions météorologiques, bruits liés au trafic maritime ou encore liés à la proximité d'animaux marins, etc.). Lim et al. [LBHL07, LBHL08], ont récemment montré qu'il était possible de classer des événements sonores sous-marins transitoires par un codage MFCC ¹ des trames acoustiques. Ils ont proposé une classification des vecteurs caractéristiques par comparaison de distances euclidiennes (k -ppv ²) ou par l'apprentissage d'un PMC ³. Nous proposons dans ce chapitre de compléter cette étude selon les axes suivants :

- extension au cas des signaux *permanents* ou *long-termes*, c'est à dire dont les caracté-

¹Mel Frequency Cepstral Coefficients

² k Plus Proches Voisins

³Perceptron MultiCouche

- ristiques varient faiblement au cours du temps ;
- étude d'autres méthodes de codages (LPC ⁴, LPCC ⁵ et ERB ⁶) ;
 - classification par modélisation prédictive non-linéaire ;
 - plus grand nombre de classes de signaux (13 à 30 classes au lieu de 8 dans [LBHL07, LBHL08]).

Nous montrerons ainsi que les meilleurs taux de reconnaissance pour les signaux long-termes sont obtenus par un codage LPCC suivi d'une classification prédictive PMC, tandis que les meilleurs taux pour les signaux transitoires sont obtenus par une classification de type *k-ppv* précédée d'un codage LPC ou LPCC des trames acoustiques. En revanche, le codage MFCC traditionnellement utilisé en reconnaissance de la parole ne permet pas d'obtenir ici les meilleurs résultats.

En première partie de ce chapitre, nous décrirons les algorithmes de codage des signaux. En deuxième partie nous aborderons la classification des signaux sous-marins. La troisième section de cet article est consacrée à la description des expérimentations menées et à l'analyse des résultats obtenus.

3.2 Représentation des signaux

Pour le problème de classification abordé dans cet article, nous avons supposé que la forme, la durée et la réponse spectrale du signal n'étaient pas connues. Les caractéristiques spectrales peuvent être fortement variables dans le temps, ce qui a amené bon nombre d'auteurs à utiliser des méthodes de représentation temps-fréquence (transformée en ondelettes [MW97], distribution de Wigner-Ville et distribution croisée de Wigner-Ville [Yen98, BS90], ou transformée de Fourier à court terme (STFT) [EHPS98]).

Dans ce chapitre, nous suivons la démarche de Lim et al [LBHL07, LBHL08] qui proposent d'utiliser les méthodes de codage de signaux classiquement utilisées en traitement de la parole (MFCC) ainsi que des classifieurs de type *k-ppv* et perceptron multicouche pour catégoriser des signaux transitoires sous-marins. Pour cela, nous avons testé quatre types de codes : trois codes utilisés conventionnellement en traitement du signal audio (MFCC, LPC et LPCC) et un code inspiré du modèle de la cochlée de Patterson (ERB) [PRH⁺92].

⁴Linear Predictive Coding

⁵Linear Prediction Cepstral Coefficients

⁶Equivalent Rectangular Bandwidth

Tous les codages utilisés sont basés sur une description court-terme du spectre : le signal est découpé en trames successives d'environ 45 ms. Un prétraitement par fenêtrage de Hamming est ensuite effectué sur chacune des trames avant extraction de 14 paramètres selon les 4 méthodes de codage précédentes.

3.2.1 ERB

Le modèle cochléaire de Patterson est basé sur un ensemble de filtres passe-bande indépendants. Dans ce modèle, la bande passante de chaque filtre cochléaire est décrite par une bande passante rectangulaire équivalente (ERB). L'ERB est une mesure psychoacoustique, ce qui donne une approximation de la bande passante des filtres dans l'audition humaine. Chaque filtre modélise le signal présent en sortie d'un nerf de la cochlée.

Pour N filtres ERB dont les réponses fréquentielles s'étagent comme l'indique la figure 3.3, nous obtenons N signaux pour chaque trame, dont nous calculons l'énergie sur la période de temps correspondant à la trame. Ces signaux sont segmentés en trames successives sur lesquelles nous calculons une norme en valeur absolue :

$$\mathbf{x}_k = [x_1^k, x_2^k, \dots, x_N^k] \text{ avec } x_{i=1\dots N}^k = \sum_{q=1}^L |y_i^k(q)| \quad (3.1)$$

où x_i^k représente la composante i du vecteur de code \mathbf{x}_k calculé sur la trame k et les y_i^k sont les échantillons du signal avec une longueur de L en sortie du $i^{\text{ème}}$ filtre.

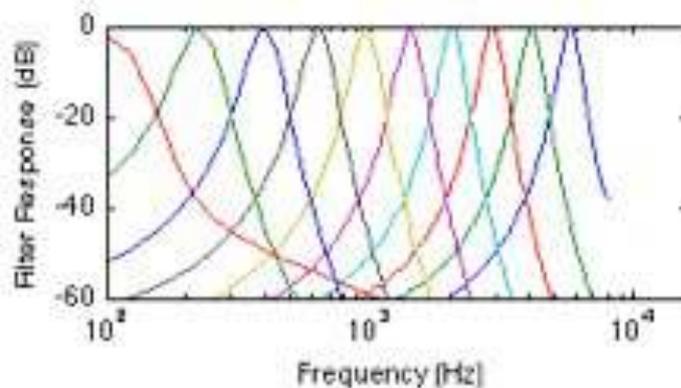


FIG. 3.3: Réponses impulsionnelles d'un ensemble de filtres réalisant le codage ERB des signaux audio.

3.2.2 LPC

Le codage LPC est une technique performante d'analyse de la parole. Il est utile pour encoder la parole en bonne qualité et à un faible bit rate. Le codage LPC est fondé sur la modélisation du conduit vocal, suivant en cela l'hypothèse du système source-filtre linéaire. Nous considérons un signal $s(n)$ issu d'un processus autorégressif :

$$s(n) = \sum_{i=1}^N a_i s(n-i) + Gu(n) \quad (3.2)$$

où $u(n)$ est le signal d'excitation, G est le gain et a_i avec $i \in \{1 \dots N\}$ sont les coefficients de prédiction linéaire LPC. Le prédicteur optimal est donné par :

$$\tilde{s}(n) = \sum_{i=1}^N a_i s(n-i) \quad (3.3)$$

Il existe de nombreuses façons de trouver des coefficients de LPC, comme les algorithmes de moindres carrés rapides [MG76] et l'algorithme de Levinson-Durbin [Lev47]. L'algorithme de Levinson-Durbin permet d'extraire par récurrence sur l'ordre de prédiction, les coefficients LPC.

3.2.3 LPCC

Le codage LPCC permet de calculer les coefficients de la représentation cepstrale du signal, à partir des coefficients LPC. Considérons $(a_0, \dots, a_r)^T$, les $r+1$ coefficients calculés par une analyse LPC. Les r premiers coefficients cepstraux peuvent être calculés par :

$$c_{i=1 \dots r} = a_i + \sum_{k=1}^{i-1} \frac{k}{i} c_k a_{i-k} \quad (3.4)$$

Les coefficients cepstraux de $r+1$ jusqu'au degré s désiré peuvent être calculés comme :

$$c_{i=r+1 \dots s} = \sum_{k=i-r}^{i-1} \frac{k}{i} c_k a_{i-k} \quad (3.5)$$

Comme le codage LPC, le codage LPCC appartient à la littérature de méthode d'extraction de caractéristique pour la compression ou la classification de signaux audio, en particulier de signaux de parole.

3.2.4 MFCC

Le codage MFCC extrait les caractéristiques fréquentielles du signal en suivant une échelle non linéaire appelée *échelle de Mel* et inspirée de la sensibilité auditive de l'oreille humaine.

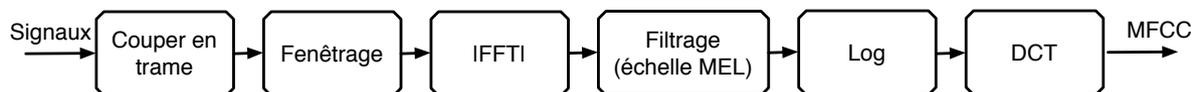


FIG. 3.4: Les étapes d'extraction des MFCCs.

La figure 3.4 montre le processus d'extraction des coefficients de MFCC. Le signal sonore est d'abord coupé en trames, puis la pondération de Hamming est appliquée sur chaque trame du signal. Ensuite, nous calculons la FFT des données qui permet d'obtenir une représentation fréquentielle du signal. La prochaine étape spécifiée comme mel-filtrage fournit un modèle d'audition réalisé par un banc de filtres triangulaires dans l'échelle de Mel. L'échelle de Mel est défini comme :

$$Mel(f) = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right) \quad (3.6)$$

où f est la fréquence en Hz . Dans l'étape suivante, un logarithme naturel est utilisé comme fonction de lissage des données. Enfin, les coefficients cepstraux peuvent être calculés en appliquant une DCT (Discrete Cosinus Transform) sur les coefficients en sorties des filtres.

Ce codage est intéressant à tester car il est le plus utilisé dans les systèmes de reconnaissance de la parole. En particulier Lim et al. [LBHL08] l'utilisent dans leur système de classification de signaux transitoires. Nous montrons qu'il n'est peut-être pas le mieux adapté dans tous les cas.

3.3 Classification des signaux

Nous proposons deux méthodes de classification des signaux audio. Cela nous permettra d'une part d'établir des comparaisons, et d'autre part d'apporter des informations plus robustes aux étages de décision puis de commande du système robotique complet.

La première technique de classification est une technique de classification par modélisation prédictive des signaux. Une méthode déjà utilisée dans des systèmes de reconnaissance de locuteur [Ma93] qui présente l'avantage de proposer une unique décision quelque soit le

nombre de trames considérées. La deuxième méthode des *k-ppv* fait partie des méthodes les plus simples et les plus classiques en reconnaissance des formes, d'où son intérêt ici. En revanche, elle nécessite de mettre en place un algorithme de décision global travaillant à partir de décisions locales opérées au niveau de la trame acoustique.

3.3.1 Modèle prédictif

La modélisation prédictive des sources sonores permet d'estimer une distance entre une source inconnue et un ensemble de modèles de sources. Ces sources sont chacune modélisées par un réseau de type perceptron multicouche utilisé en prédiction. Le perceptron multicouche est un classifieur linéaire organisé en plusieurs couches au sein desquelles une information circule de la couche d'entrée vers la couche de sortie uniquement. Chaque couche est constituée d'un nombre variable de neurones, les neurones de la couche de sortie correspondant toujours aux sorties du système.

Considérons le vecteur des paramètres \mathbf{x}_k extraits d'une trame k . La modélisation d'une source sonore consiste en l'apprentissage d'un réseau par minimisation de l'erreur quadratique de prédiction calculée sur l'ensemble des trames de la séquence :

$$Q(\Omega) = \sum_k \|F(\mathbf{x}_{k-2}, \mathbf{x}_{k-1}) - \mathbf{x}_k\|^2 \quad (3.7)$$

où F représente la fonction réalisée par le réseau qui, à deux trames successives \mathbf{x}_{k-2} et \mathbf{x}_{k-1} en entrée, associe la trame suivante \mathbf{x}_k .

Pour un problème de classification d'une source parmi N , on dispose donc de N réseaux préalablement appris à l'aide de bases d'apprentissages constituées de signaux appartenant à la même classe sonore (figure 3.5). En reconnaissance, une source inconnue voit ses trames constitutives présentées en entrée des N réseaux. N erreurs de prédictions sont ainsi calculées et la décision finale est obtenue selon le principe du maximum de vraisemblance : la classe de la source inconnue est celle du réseau présentant l'erreur de prédiction la plus faible.

3.3.2 Classification *k-ppv*

Le classification *k-ppv*⁷ est une méthode de classification d'objets dans l'espace des caractéristiques. Habituellement, la distance euclidienne est utilisée comme mesure de distance. Les classes des k exemples les plus proches (k est une constante étant choisie par l'utilisateur) sont vérifiées et la classe majoritaire est décidée comme le résultat de la classification.

⁷*k* Plus Proches Voisins

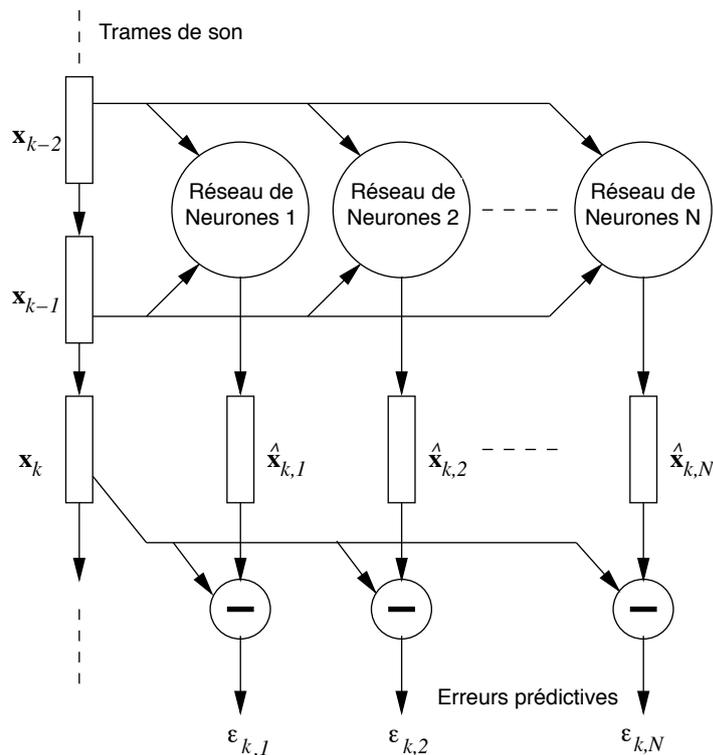
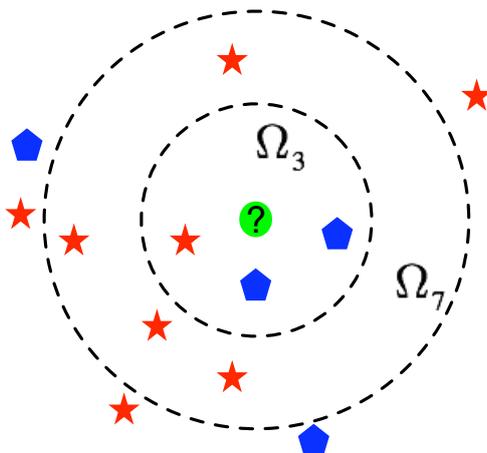


FIG. 3.5: Architecture du système de reconnaissance sonore.

L'algorithme de k -ppv est illustré sur la figure 3.6. Il existe plusieurs possibilités pour classifier le cercle marqué d'un point d'interrogation, ils doivent être classés soit dans la première classe qui contient les étoiles, soit dans la deuxième classe qui contient les pentagones. Si $k = 3$, notre objet est classé à la deuxième classe parce qu'il y a 2 pentagones et seulement une étoile à l'intérieur du cercle Ω_3 . Si $k = 7$, il est classé dans la première classe, les 5 étoiles et seulement 2 pentagones se trouvent à l'intérieur du cercle Ω_7 .

Cette méthode nous permet de réaliser une classification locale des trames acoustiques : pour une trame donnée, le vecteur caractéristique issu du codage est comparé à l'ensemble des vecteurs de codes d'une base de référence préalablement constituée. La distance euclidienne entre le code inconnu et les codes de référence est calculée. La décision consiste à affecter au code inconnu la classe majoritaire parmi les k vecteurs de référence les plus proches. Pour une séquence sonore on dispose donc d'une succession de décisions locales à partir desquelles une décision globale est prise. Cet algorithme a été utilisé par Lim et al. [LBHL07].

FIG. 3.6: Exemple de la classification k -ppv.

3.3.3 Décisions locales et globales

Les deux algorithmes de classification (le modèle prédictif et le modèle k -ppv) permettent de produire une décision locale au niveau de la trame.

Décisions par l'erreur globale (GD1)

Dans la suite, nous noterons c_k la décision locale effectuée sur la $k^{\text{ième}}$ trame acoustique. Pour le classifieur prédictif, nous avons :

$$c_k = \arg \min_{i=1, N} \{\varepsilon_{k,i}\} \quad (3.8)$$

où les $\varepsilon_{k,j}$ désignent les N erreurs calculées sur les N perceptrons multicouches. la décision globale est naturellement obtenue par :

$$c = \arg \min_{i=1, M} \left\{ \sum_{k=1}^K \varepsilon_{k,i} \right\} \quad (3.9)$$

où K représente le nombre total de trames de la séquence sonore inconnue. Nous désignons par $GD1$ cet algorithme de décision.

Décisions majoritaire (GD2)

En revanche, le classifieur k -ppv ne permet pas de construire aussi facilement une décision globale. Un moyen est d'utiliser un algorithme de décision de type *décision majoritaire* souvent utilisé dans la littérature :

$$c = \arg \max_{i=1,M} \{|U_i|\} \quad (3.10)$$

où U_i représente l'ensemble des trames appartenant à la classe $i, i = 1, \dots, M$. Nous désigneront par *GD2* cet algorithme de décision.

Décisions par la plus longue suite (GD3)

Le dernier algorithme, noté *GD3*, est une variante de l'algorithme *GD2* consistant à considérer la plus longue suite des décisions locales identiques successives.

$$c = \arg \max_{i=1,M} \{W_i\} \quad (3.11)$$

où W_i est la longueur de la suite successive des décisions locales pour la classe i .

3.4 Expérimentations

Nous présentons un ensemble de simulations réalisées à partir d'une base de signaux réels extraits des enregistrements *The Underwater Sound Effects Series* [Bas04]. Cette base comporte 500 enregistrements d'événements sonores échantillonnés à 44KHz et réalisés à partir d'hydrophones Brüel & Kjaer. Parmi ces enregistrements, certains correspondent à des événements sonores transitoires (d'une durée maximale de l'ordre de la seconde) et d'autres à des événements long-termes (jusqu'à une minute d'enregistrement). Cette base nous a donc permis de traiter les deux types d'événements. Le plus souvent, ce sont les signaux court-termes qui sont traités dans la littérature. Or dans le contexte robotique qui nous concerne, nous aurons également à traiter des signaux long-termes, par exemple pour apporter des informations concernant l'état de la mer et/ou du navire. Bon nombre de ces signaux ont également été enregistrés dans l'eau et au dessus de l'eau. A titre d'exemple, les figures 3.7 et 3.8 représentent respectivement les spectres des signaux obtenus par l'enregistrement d'un événement long-terme de type *goutte à goutte*, et d'un événement court-terme de type

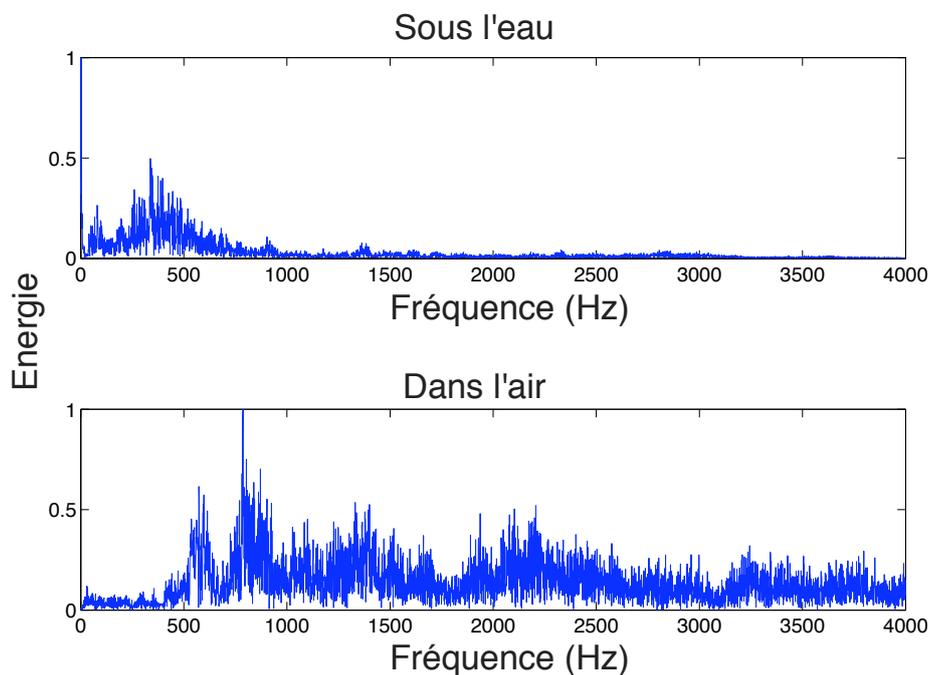


FIG. 3.7: Spectres d'un événement sonore correspondant à un *goutte à goutte* enregistrés sous l'eau (en haut) et dans l'air (en bas).

frappement d'un métal. En haut des figures sont représentés les spectres des enregistrements au dessous du niveau de l'eau en bas au dessus de ce même niveau, c'est à dire dans l'air.

Nous pouvons constater sur ces figures les différences spectrales qui interviennent : essentiellement une réduction de la largeur de bande des signaux sous l'eau, ainsi qu'un déplacement vers les basses fréquences. Ce résultat était attendu dans la mesure où l'on sait que les basses fréquences se propagent mieux que les hautes fréquences sous l'eau.

Nous pouvons déjà en conclure que le codage MFCC, classiquement utilisé en reconnaissance de la parole, est probablement moins bien adapté à l'analyse des signaux sous-marins. C'est ce que nous vérifions dans les expérimentations présentées ci-dessous.

3.4.1 Codage

Pour les quatre méthodes de codage utilisées, nous avons représenté l'évolution dans le temps des coefficients tout au long de la durée d'un événement. Dans toutes les simulations proposées, nous avons utilisé 14 coefficients de codage. En effet, une étude portant sur la dimension des vecteurs de code nous a montré que la dimension optimale varie dans une

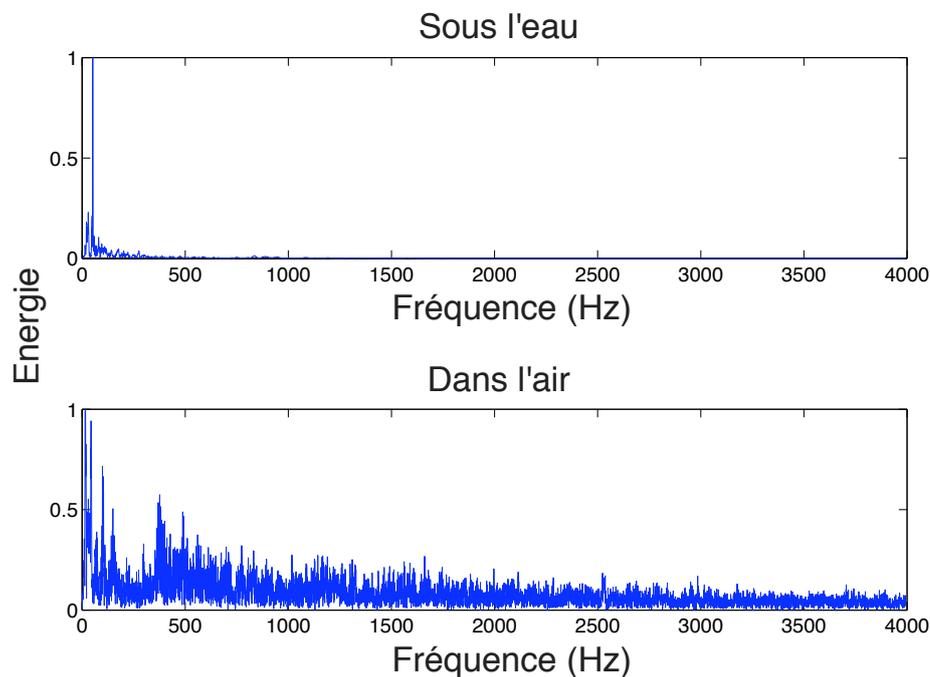


FIG. 3.8: Spectres d'un événement sonore correspondant à un frapement du métal enregistrés sous l'eau (en haut) et dans l'air (en bas).

plage allant de 10 à 18 avec une moyenne autour de 14 (Fig. 3.9, en fonction de la nature des signaux traités (court-termes ou long-termes) et des méthodes de codage utilisées. La figure 3.9 donne une exemple d'évaluation sur les différents nombres de coefficients de codage en utilisant les codages MFCC sur la base de signaux long-termes.

Les figures 3.10 et 3.11 représentent des exemples d'évolution pour un signal permanent et un signal transitoire. Dans le cas des signaux permanents, nous constatons une variance plus importante des coefficients MFCC relativement aux coefficients LPC et LPCC. Nous trouvons là une conséquence de l'inadaptation des coefficients MFCC. En effet, comme le montre l'évolution des coefficients LPC et LPCC, les propriétés long-terme du signal sont quasi-stationnaires. En revanche, une grande partie des bandes de fréquence des codes MFCC (les fréquences supérieures à 1500Hz comme le montre le spectre de la figure 3.7) ne véhiculent plus d'information utile. De la même façon, le codage ERB couvre l'ensemble des fréquences du spectre audible. Une bonne part des coefficients sont de faible amplitude car l'énergie sur les canaux de plus haute fréquence est faible.

Concernant les signaux court-termes, l'exemple du spectre présenté figure 3.8 montre qu'il s'agit de signaux à bande étroite. L'évolution des coefficients est présentée sur la figure 3.11.

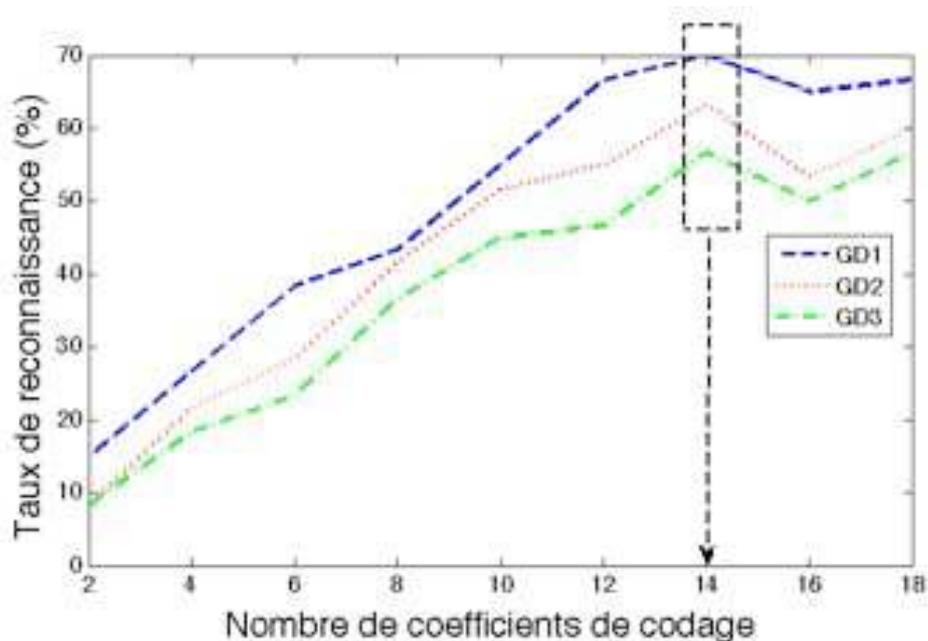


FIG. 3.9: Taux de reconnaissance en fonction du nombre de coefficients de codage en utilisant les codages MFCC sur la base de signaux long-termes.

Ils montrent que du point de vue spectral, les signaux courts-terme ne sont pas des signaux stationnaires. Le nombre de coefficients représentatifs d'une évolution est plus important dans le cas des codes LPC et LPCC que MFCC et ERB. Là encore, les résultats montrent que les coefficients LPC et LPCC pourraient être de meilleurs candidats pour la modélisation des signaux sous-marins. Ces résultats observés sur deux séquences sonores particulières sont généralisables à une grande partie des signaux disponibles dans la base.

3.4.2 Paramètres des classifieurs

Les classifieurs proposés dans ce chapitre (le classifieur prédictif neuronal et le classifieur k -ppv) nécessitent d'être correctement paramétrés. Pour cela nous avons effectué des séries de simulations portant sur le nombre de cellules cachées des réseaux de neurones d'une part et sur le nombre de voisins du classifieur k -ppv d'autre part.

Nous utilisons ici un réseau de neurone avec deux couches cachées pour le classifieur prédictif (Fig. 3.12), le nombre de cellules dans la couche cachée est un point intéressant à étudier. Deux bases ont été constituées, l'une composée de signaux court-termes (13 classes de signaux) et l'autre composée de signaux long-termes (30 classes de signaux). Nous avons ainsi réalisé deux séries d'expérimentations. La séparation des bases était nécessaire car les signaux

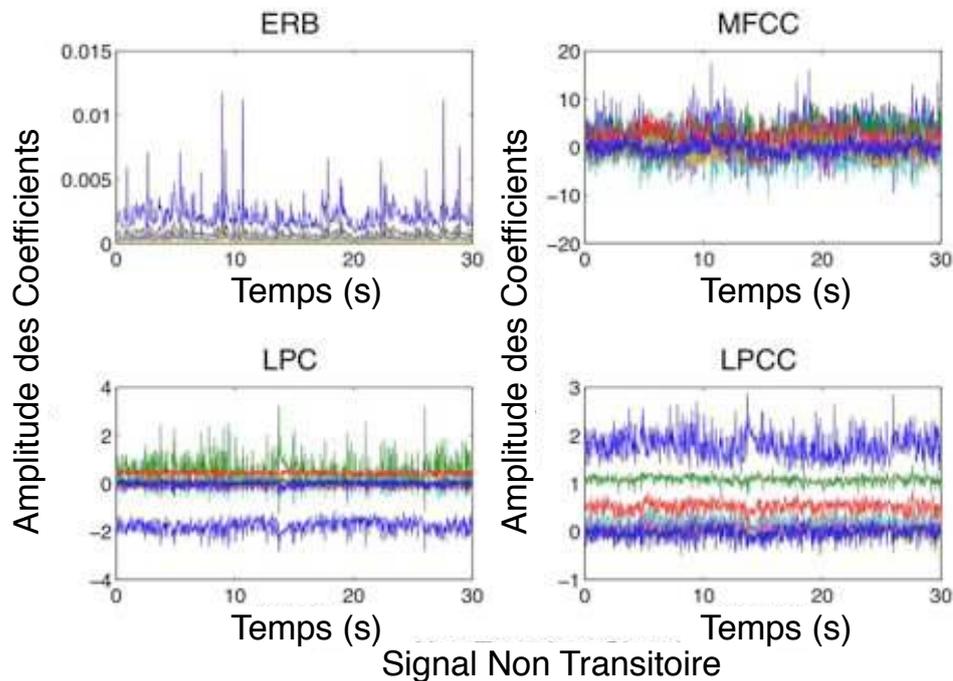


FIG. 3.10: Evolution des coefficients de codage ERB, LPC, LPCC et MFCC au cours du temps pour un signal long-terme (30s, fenêtres de 2048 points).

court-termes et long-termes ont des caractéristiques différentes qui peuvent justifier l'emploi de codes différents. Nous présentons ici deux exemples de simulations : l'une (figure 3.13) représente les scores obtenus avec le classifieur prédictif et les codes LPCC en fonction du nombre de cellules cachées, l'autre (figure 3.14) représente les scores obtenus avec le classifieur k -ppv et les codes MFCC en fonction du nombre de voisins. Pour ces tests, les algorithmes de décision $GD1$, $GD2$ et $GD3$ ont été testés avec le classifieur prédictif et $GD2$, $GD3$ avec le classifieur k -ppv. En effet, le k -ppv tel que nous l'avons implémenté ne permet pas de produire une décision du type $GD1$ (voir plus haut dans cet article).

En général, nous notons qu'un faible nombre de neurones ou qu'un faible nombre de voisins permet d'obtenir les meilleurs évaluations. Concernant les réseaux prédictifs, nous avons là un résultat lié au problème de sur-apprentissage : lorsque le nombre de cellules est trop important, les réseaux apprennent mieux les séquences d'apprentissage mais généralisent moins bien. Concernant le classifieur k -ppv, ceci est dû au faible nombre de références disponibles dans la base de signaux. Or nous savons que les k -ppv fonctionnent d'autant mieux que le nombre de références est important.

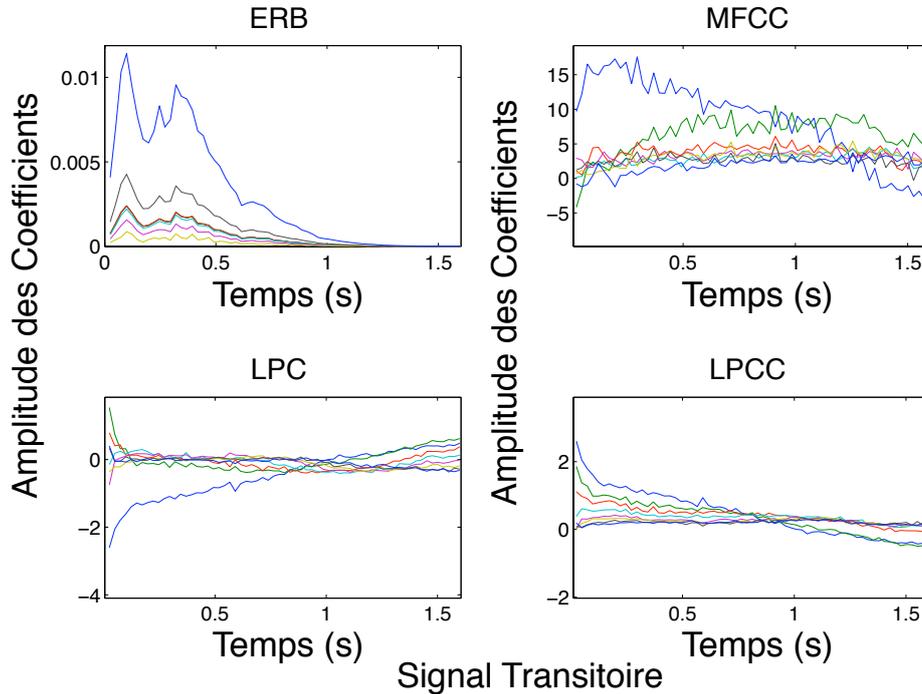


FIG. 3.11: Evolution des coefficients de codage ERB, LPC, LPCC et MFCC au cours du temps pour un signal court-terme (1.6s, fenêtres de 1024 points).

3.4.3 Résultats

Le dernier ensemble d'expérimentations que nous proposons dans cet article concerne les scores obtenus par les deux classifieurs testés sur les quatre codes ERB, LPC, LPCC et MFCC sur les deux bases de signaux court-termes et long-termes. Concernant les réseaux prédictifs, il s'agit de scores obtenus sur une base de test, c'est à dire sur des exemples non appris par les modèles. Les figures 3.15 et 3.16 représentent les scores obtenus. $C1$ désigne le classifieur prédictif et $C2$ représente le classifieur k -ppv.

Concernant les signaux longs, on note une supériorité des codes LPCC quelque soient les algorithmes de classification utilisés (entre 78% et 85%). En revanche, les résultats obtenus avec les codes MFCC sont plus mitigés : bons avec le classifieur prédictif (entre 79% et 81%), mais moins bons avec le classifieur k -ppv (entre 59% et 65%).

Pour ce qui concerne les signaux courts, la dépendance vis à vis du classifieur utilisé augmente. Les codes LPC et LPCC permettent toujours d'obtenir les meilleurs résultats, mais à condition toutefois d'utiliser les réseaux prédictifs. Les réseaux prédictifs présentent ici un intérêt particulier puisqu'ils permettent de modéliser l'évolution de signaux non stationnaires,

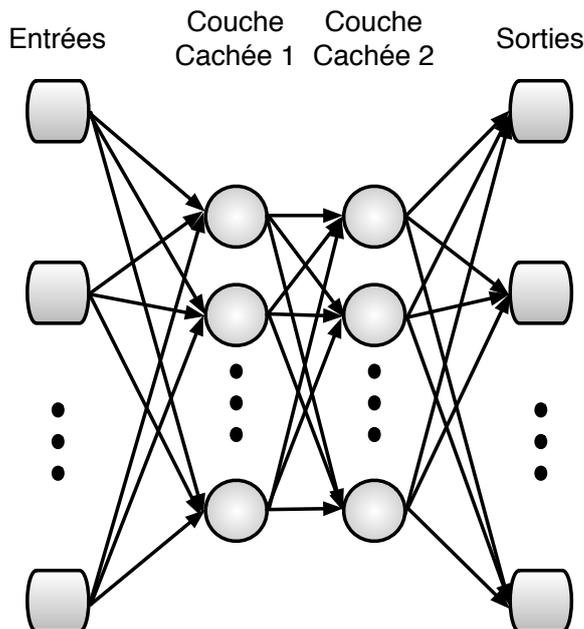


FIG. 3.12: Réseau de neurone utilisé dans le classifieur prédictif.

or les signaux court-termes sont non stationnaires. Enfin, les scores de 100% obtenus, supérieurs aux scores obtenus sur les signaux longs, s'expliquent par le nombre de classes : 13 pour les signaux courts et 30 pour les signaux longs.

Ces résultats montrent que les codes MFCC ne sont pas les codes les plus appropriés pour la reconnaissance de signaux audio sous-marins. Les codes LPCC apparaissent ici comme de meilleurs candidats. La suite de notre travail dans le cadre du projet ASAROME consiste en la fusion de données multimodales (sonores, visuelles et accélérométriques) pour l'aide à la décision et le développement de commandes appropriées pour la navigation du navire autonome.

3.5 Conclusion

Les signaux sonores sous-marins peuvent être classifiés en deux groupes principaux selon leurs durées : signal transitoire et signal non-transitoire. Dans ce chapitre, nous avons présenté des algorithmes de classification de signaux sonores pour l'aide à la navigation d'un navire autonome. Nous avons comparé deux classifieurs et quatre méthodes de codage des signaux, suivant en cela le travail entamé par Lim et al. dans [LBHL07, LBHL08]. Les codes MFCC classiquement utilisés en traitement du signal de la parole ne sont pas le mieux adaptés au traitement de signaux audio sous-marins. Les résultats obtenus montrent que, si les mé-

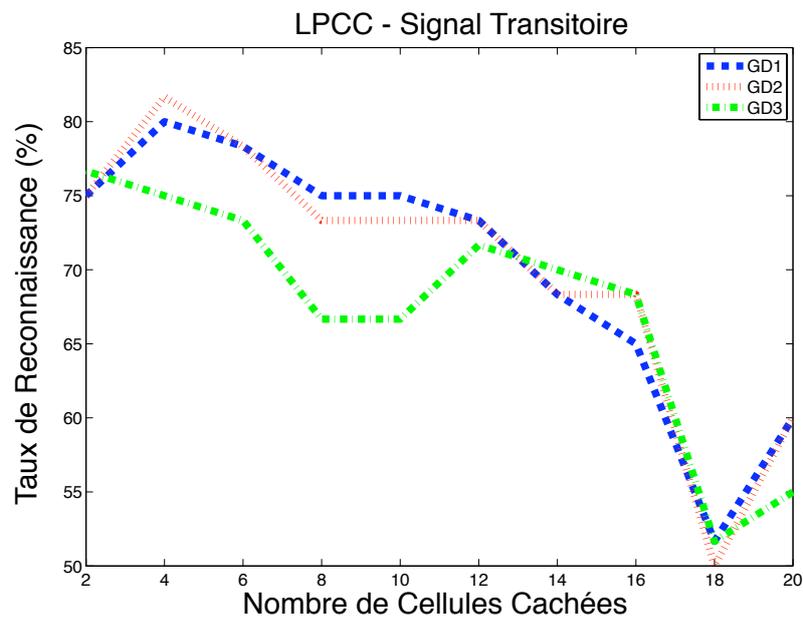


FIG. 3.13: Taux de reconnaissance en fonction du nombre de cellules cachées en utilisant les codages LPCC sur la base de signaux long-termes.

thodes de représentation des signaux sont bien choisies, il est possible d'obtenir des résultats prometteurs en classification.

Dans le chapitre suivant, nous présenterons le système de vision panoramique utilisé pour la détection d'obstacle et la détection de l'assiette.

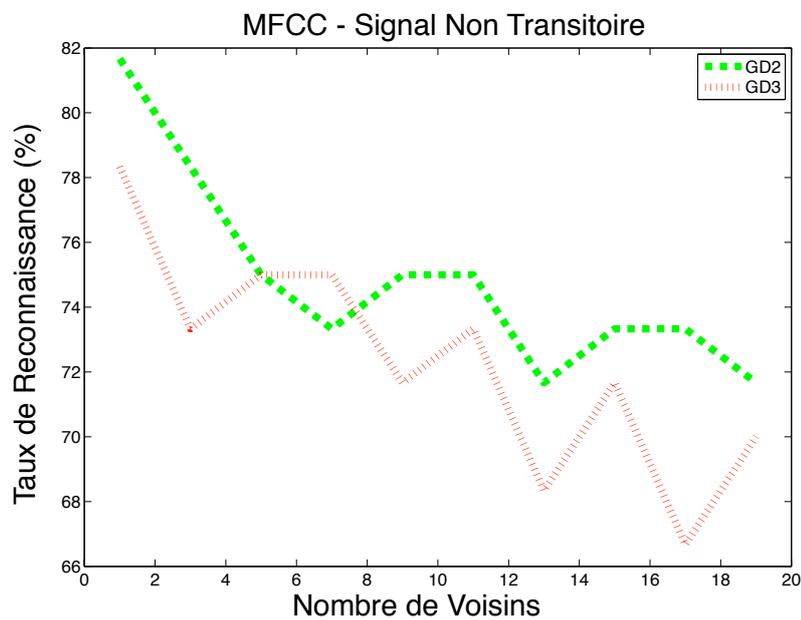


FIG. 3.14: Taux de reconnaissance en fonction du nombre de voisins en utilisant les codages MFCC sur la base de signaux long-termes.

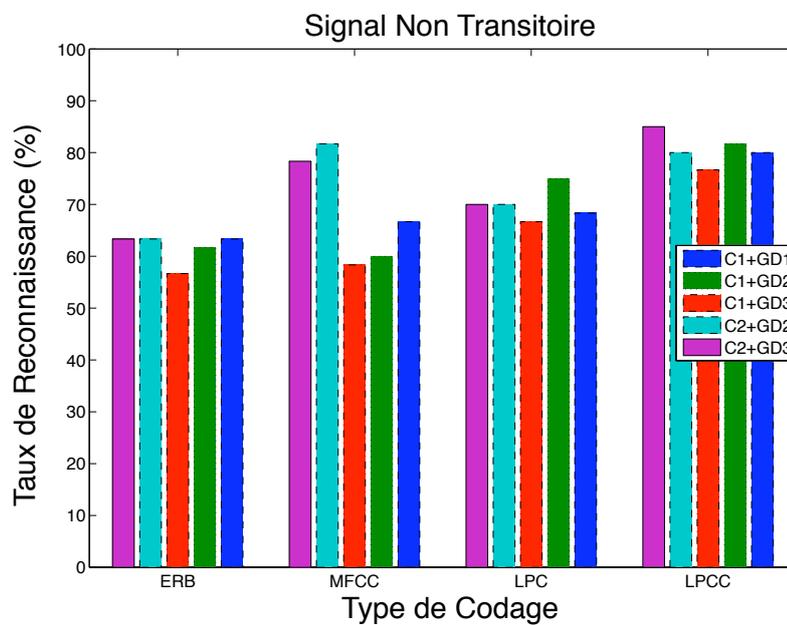


FIG. 3.15: Taux de reconnaissance des codes sur la base durable.

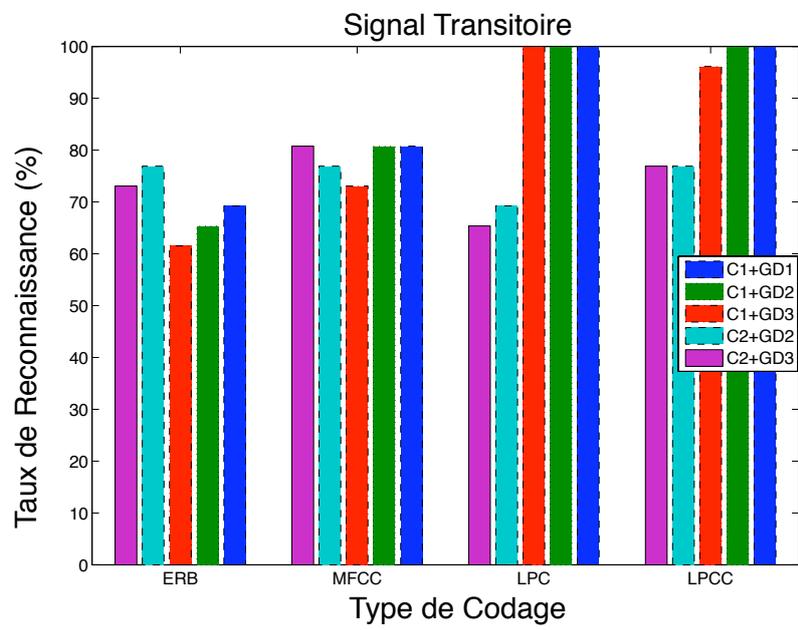


FIG. 3.16: Taux de reconnaissance des codes sur la base transitoire

CHAPITRE

4

Vision panoramique

Sommaire

4.1	Introduction	48
4.1.1	Historique sur les panoramas	48
4.2	Calibrage de la caméra panoramique	50
4.2.1	Positionnement de la caméra relativement au miroir	51
4.2.2	Réflexion des rayons par la surface du miroir	57
4.3	Estimation de l'orientation du bateau	60
4.3.1	Algorithme	60
4.3.2	Résultats des tests	64
4.4	Détection d'obstacle	65
4.5	Estimation de l'état de la mer par la vision : détection de l'orientation	67
4.5.1	Segmentation du ciel et de l'eau	68
4.5.2	Détection du contour horizon	68
4.5.3	Détermination du plan d'incidence lié à l'horizon	70
4.6	Résultats expérimentaux	71
4.6.1	Test pour la détection d'obstacle	71
4.6.2	Test de détection de l'orientation dans un bassin	72

4.6.3 Test de détection de l'orientation en mer	76
4.7 Conclusion	79

4.1 Introduction

La vision humaine est un processus très complexe qui permet d'observer l'environnement, détecter et localiser les objets, mais également de distinguer leurs formes et leurs couleurs. La vision par ordinateur est la reconstruction par un modèle algorithmique de la vision humaine. Elle possède des propriétés semblables, afin de réaliser les mêmes missions que celles de l'œil humain. Grâce à ses bonnes propriétés en termes de précision, et de détection, ainsi que la diversité des informations qu'elle peut fournir, elle est utilisée dans des domaines variés.

La vision par ordinateur joue un rôle très important dans le domaine de la robotique. L'installation d'une ou plusieurs caméras permet d'améliorer la perception. La limitation de l'angle de vue est un inconvénient majeur de la vision conventionnelle. Comme pour les humains, le champ de vision des yeux est d'environ 180°. Pour élargir le champ visuel, nous devons tourner la tête par un mouvement du cou. Pour un robot l'utilisation de la vision panoramique est nécessaire pour permettre une acquisition de l'environnement sur 360°. [BK01] donne une explication sur l'origine, la théorie et les applications de la caméra panoramique.

Ce chapitre est consacré au traitement d'images panoramiques. Nous expliquons tout d'abord la méthode de calibrage. Ensuite nous présentons l'estimation de l'orientation ¹ du bateau et la détection d'obstacles. Enfin, nous donnons un algorithme pour l'estimation de l'état de la mer par la vision panoramique.

4.1.1 Historique sur les panoramas

La vision panoramique est une réduction de la vision omnidirectionnelle, qui permet de faire une détection de vision en bande. Elle décrit une représentation de l'environnement avec un large champ de vision, balayant 360° d'azimut. Dans le domaine de l'art, Barker a créé la première peinture circulaire à 360° en largeur qu'il a nommée "Panorama" en 1792 (Fig. 4.1).

La première caméra panoramique a été inventée par Joseph Puchberger en Autriche en 1843. C'est une caméra à manivelle avec un champ de vision entre 120 et 150°. En 1844,

¹LAROUSSE : Orientation, direction des axes d'un engin spatial par rapport à un trièdre de référence.



FIG. 4.1: La première peinture panoramique en 1792.

une caméra panoramique améliorée a été assemblée par Fredrich von Martens en Allemagne avec un champ visuel de 150° . En 1857, l'invention de la caméra rotative de M. Garella en Angleterre a étendu le champ de vision à 360° . La figure 4.2 montre une photo prise par G. N. Barnard en 1864.



FIG. 4.2: Vue depuis le sommet de Lookout Mountain en 1864 par George N. Barnard.

Les capteurs panoramiques peuvent être classifiés selon les dispositifs utilisés :

- Les systèmes constitués de plusieurs caméras [Hen95, ISY03].
- Les systèmes mécaniques tournants composés d'une unique caméra [LKPP08, BMD96].
- Les système couplant une caméra et un miroir [Ree70, Gre89, YK90, YYY93, SBFR96, Pro97].

Dans le premier cas, l'acquisition des images est effectuée, suivie d'une étape de post-traitement nécessaire. Il faut relier les images des différente caméras lors de cette phase, mais cette étape est rapide. Cette solution est quasiment temps réel. Dans le second cas, la procédure de capture n'est pas immédiate puisque la caméra doit être orientée selon différentes directions. Dans le dernier cas, l'image panoramique est capturée grâce à la réflexion d'un miroir. C'est la méthode la plus rapide. Dans ce chapitre, nous nous concentrons sur ce système, qui peut être classifié en deux principaux types selon leur points de vue : les capteurs centraux et les capteurs non centraux. Un système présentant un point de vue unique est qualifié de capteur central, sinon il est classifié comme un capteur non central.

Le premier système de capteur omnidirectionnel central a été breveté par D.W. Rees [Ree70] en 1970. Il couple une caméra et un miroir hyperbolique. Ce système peut capturer l'environnement dans une seule photo. En 1993, Yamazawa et al. ait proposé un capteur omnidirectionnel central installé dans un robot utilisant un miroir hyperbolique [YYY93]. Nalwa a réalisé un système multi-capteurs avec un miroir planaire [Nal96] en 1996. Nayer a construit un système de capteur central avec un miroir parabolique en 1997 [Pro97, NB97]. Baker et Nayar ont analysé les propriétés géométriques des miroirs les plus couramment utilisés en 1998 [BN98]. Su et al. ont développé un nouveau système de capture stéréo omnidirectionnel en 2006 [SLZ06]. Jusqu'ici, tous les systèmes présentés ont une résolution d'image non linéaire. Nagahara et al. ont proposé un nouveau système de capteur omnidirectionnel avec un point de vue unique. La résolution de la distribution est linéaire [NYY07].

Les caméras n'ont pas besoin d'avoir un point de vue unique [SGN06], le premier capteur non central est créé par Yagi et al. [YK90] en 1990. Au même moment, plusieurs équipes ont travaillé sur le système panoramique de capteur non central [Hon91]. Chahl et al. ont développé un réflecteur avec une résolution d'image linéaire [CS97]. Les systèmes omnidirectionnels avec une résolution invariante ont été développés à partir 1999 [CM99, HB00, Sri03].

Au cours des dernières années, la vision panoramique est utilisée dans divers domaines. Dans le domaine industriel, les caméras de surveillance panoramique sont appliquées pour le contrôle de produits sur des chaînes de fabrication, ou pour la sécurité [MHTK05]. Dans le domaine public, le système panoramique est utilisé pour le tourisme (visite virtuelle [Ecl]) et des cartes (GoogleMaps [Goo]). Le domaine de la recherche est également concerné par le développement de véhicules autonomes de surface comportant un module de vision panoramique [SGR⁺06, GXR⁺08, Gon08, RGI⁺11, GRI⁺11].

Les capteurs catadioptriques possédant un centre unique de projection sont difficiles à réaliser en raison d'une exigence particulière sur la précision de la conception du capteur. L'utilisation de capteurs non centraux se développe donc de plus en plus. Le système de vision utilisé dans nos études comporte une caméra Gigabit, un miroir et un tube acrylique transparent protégeant le capteur (Fig. 4.3). La combinaison de ces différents composants forme un capteur non central. Dans toute la suite, les méthodes exposées s'appliquent à ce type de capteur.

4.2 Calibrage de la caméra panoramique

L'utilisation de caméras implique (presque) toujours un calibrage du dispositif. Comme le calibrage classique d'une caméra seule, l'opération consiste à estimer les paramètres du

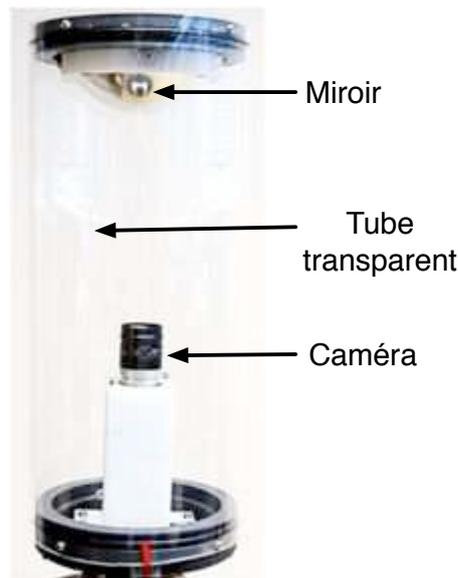


FIG. 4.3: Le tube de la caméra panoramique.

système optique permettant d'associer à chaque point de l'image le rayon lumineux incident qui l'engendre. Nous décrivons ici une méthode de calibrage plus générale. Ce travail sera décomposé en deux parties :

- Le calcul des paramètres de pose entre la caméra et le miroir, c'est à dire la position et l'orientation relative de la caméra par rapport au miroir.
- L'étude de la réflexion par la surface du miroir.

4.2.1 Positionnement de la caméra relativement au miroir

Dans le cas d'un capteur catadioptrique, le calibrage consiste à déterminer les placements (ou poses) relatifs de chacun des composants ainsi que les propriétés géométriques internes de la caméra (Fig. 4.4). La combinaison de ces différents composants produit une configuration matérielle dite "non centrale" c'est à dire qu'elle conduit à des images globalement non perspectives mais possédant des propriétés géométriques particulières comme celle de fournir une résolution angulaire constante des images. Nous devons estimer la pose relative de chacun des composants des capteurs : la caméra, le miroir et le tube acrylique transparent qui est utilisé comme boîtier étanche pour toute la structure.

Le calibrage consiste à estimer la pose de la caméra par rapport au miroir à l'aide d'un plan. Il s'agit d'utiliser ici une équivalence entre une homographie et la projection perspective

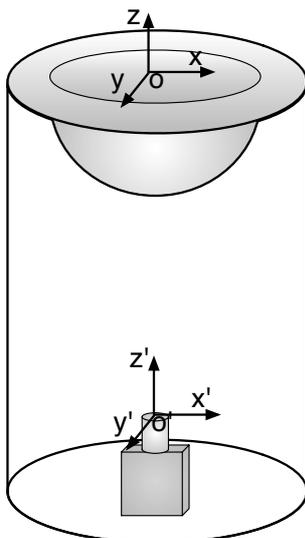


FIG. 4.4: Les différents repères locaux du système de vision.

de la caméra. Pour cette méthode, la matrice intrinsèque de la caméra doit être connue. D'une manière générale, nous utilisons la toolbox de Jean Yves Bouguet [Bou97] pour réaliser cette tâche. A partir de la connaissance de la matrice intrinsèque K_I , nous estimons la pose relative de la caméra et du miroir avec une méthode similaire à celle décrite dans [FTB02], basée sur une homographie entre deux plans judicieusement choisis d'après [Stu00].

En pratique, le plan utilisé pour déterminer cette pose doit être toujours visible. C'est donc le plan défini par le contour du miroir qui est utilisé (Fig. 4.5). Quatre marques faciles à détecter sont placées sur la monture du miroir. Elles permettent de localiser le plan défini par la partie supérieure du miroir. Connaissant la position réelle des marques et de leurs images, une estimation rapide par minimisation linéaire de l'homographie \mathbf{h}_m^I , de la matrice de rotation \mathbf{r}_m^I et de la matrice de translation \mathbf{t}_m^I est alors possible. Une description détaillée de l'algorithme de Sturm et son utilisation dans notre système de vision sont présentées dans la suite.

La partie supérieure du miroir est définie comme un plan \mathcal{P}_M . Nous plaçons quatre marques foncées sur la monture du miroir pour matérialiser des points de ce plan. Leur coordonnées sont déterminées par les centres des marques, nous avons $P_i' = (X_i, Y_i, 0)^T$ avec $i \in \{1 \dots 4\}$ dans le système de coordonnées $OXYZ$ (Fig. 4.5). Le plan \mathcal{P}_M et le plan OXY sont confondus. Les coordonnées dans le système de coordonnées globales $P_i^g = (X_i^g, Y_i^g, 0)^T$ sont calculées selon la procédure suivante :

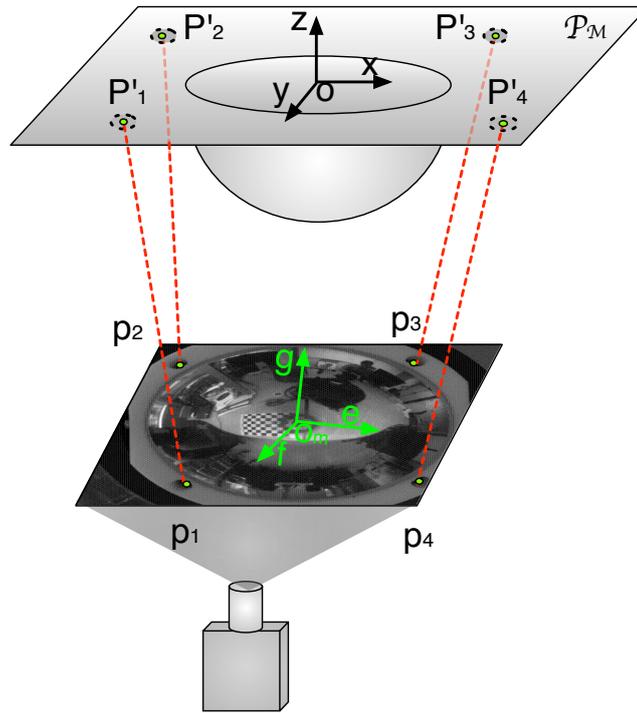


FIG. 4.5: Le contour du miroir est utilisé pour calibrer la caméra.

$$\begin{pmatrix} X_i^g \\ Y_i^g \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{r}_m^g & \mathbf{t}_m^g \\ 0_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix} \quad (4.1)$$

où \mathbf{r}_m^g est la matrice de rotation et \mathbf{t}_m^g est la matrice de translation. Elles permettent de déterminer la position et l'orientation du plan \mathcal{P}_M dans le repère global.

Les matrices de rotation et de translation de la caméra par rapport au système de coordonnées globales sont \mathbf{r}_c^g et \mathbf{t}_c^g . Les coordonnées des points P'_i dans le repère de la caméra $oefg$ correspondant à $P_i^c = (X_i^c, Y_i^c, 0)^T$ peuvent être déterminées selon l'équation suivante :

$$\begin{aligned}
 \begin{pmatrix} X_i^c \\ Y_i^c \\ 0 \\ 1 \end{pmatrix} &= \begin{pmatrix} \mathbf{r}_c^g & \mathbf{t}_c^g \\ 0_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r}_m^g & \mathbf{t}_m^g \\ 0_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix} \\
 &= \begin{pmatrix} \mathbf{r}_c^g \mathbf{r}_m^g & \mathbf{r}_c^g \mathbf{t}_m^g + \mathbf{t}_c^g \\ 0_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix}
 \end{aligned} \tag{4.2}$$

Les images obtenues par la caméra sont en 2D, le modèle de la caméra réalise une projection d'un point 3D sur une image 2D. Pour effectuer ce type de projection, nous introduisons la matrice de projection K_I pour déterminer le point $P_i^I = (X_i^I, Y_i^I, 0)^T$ dans l'image.

$$\begin{aligned}
 \begin{pmatrix} X_i^I \\ Y_i^I \\ 0 \\ 1 \end{pmatrix} &= \begin{pmatrix} K_I & 0 \\ 0_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r}_c^g \mathbf{r}_m^g & \mathbf{r}_c^g \mathbf{t}_m^g + \mathbf{t}_c^g \\ 0_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix} \\
 &= K_I \begin{pmatrix} \mathbf{r}_c^g \mathbf{r}_m^g & \mathbf{r}_c^g \mathbf{t}_m^g + \mathbf{t}_c^g \\ 0_{1 \times 3} & 1 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix} \\
 \text{avec } K_I &= \begin{pmatrix} k_e f & 0 & c_e \\ 0 & k_f f & c_f \\ 0 & 0 & 1 \end{pmatrix}
 \end{aligned} \tag{4.3}$$

où k_e et k_f sont les facteurs d'échelle d'image. f est la distance focale. c_e et c_f sont les coordonnées de la projection du centre optique sur l'image.

Pour simplifier le calcul envisagé, nous choisissons le système de coordonnées $OXYZ$ comme le repère global. Les matrices de rotation et de translation de la caméra par rapport au système de coordonnées globales \mathbf{r}_c^g et \mathbf{t}_c^g correspondent aux \mathbf{r}_m^I et \mathbf{t}_m^I . Nous avons donc la simplification de l'équation 4.4 :

$$\begin{aligned}
 \begin{pmatrix} X_i^I \\ Y_i^I \\ 0 \\ 1 \end{pmatrix} &= K_I \begin{pmatrix} \mathbf{r}_c^g \mathbf{r}_m^g & \mathbf{r}_c^g \mathbf{t}_m^g + \mathbf{t}_c^g \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix} \\
 &= K_I \begin{pmatrix} \mathbf{r}_c^g I_{3 \times 3} & \mathbf{r}_c^g \mathbf{0} + \mathbf{t}_c^g \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix} \\
 &= K_I \begin{pmatrix} \mathbf{r}_c^g & \mathbf{t}_c^g \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix} \\
 &= K_I \begin{pmatrix} \mathbf{r}_m^I & \mathbf{t}_m^I \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{pmatrix}
 \end{aligned} \tag{4.4}$$

La troisième coordonnée Z_i du point $P'_i = (X_i, Y_i, 0)^T$ est égale à zéro. La troisième colonne de l'élément \mathbf{r}_m^I est éliminée. Nous pouvons simplifier l'équation précédente :

$$\begin{pmatrix} X_i^I \\ Y_i^I \\ 1 \end{pmatrix} = K_I \begin{pmatrix} (\mathbf{r}_m^I)_{(1,2)} & \mathbf{t}_m^I \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 1 \end{pmatrix} \tag{4.5}$$

où $(\mathbf{r}_m^I)_{(1,2)}$ représente les deux premières colonnes de la matrice \mathbf{r}_m^I .

Nous obtenons les équations suivantes :

$$\begin{aligned}
 \mathbf{h}_m^I &\approx K_I \begin{pmatrix} (\mathbf{r}_m^I)_{(1,2)} & \mathbf{t}_m^I \end{pmatrix} \\
 \Rightarrow \underbrace{K_I^{-1} \mathbf{h}_m^I}_M &\approx \underbrace{\begin{pmatrix} (\mathbf{r}_m^I)_{(1,2)} & \mathbf{t}_m^I \end{pmatrix}}_{\mathbf{h}_t^I}
 \end{aligned} \tag{4.6}$$

Connaissant plusieurs points dans un plan et leurs projections dans l'image, nous pouvons calculer facilement l'homographie \mathbf{h}_m^I entre le plan et l'image. Une explication complète de cette méthode est présentée par Hartley dans sa publication [Har95]. Les mesures étant bruitées, M et \mathbf{h}_t^I ne peuvent pas être exactement égales. M donne seulement une approximation de \mathbf{h}_t^I .

Une constante λ est introduite pour que l'équation précédente soit établie :

$$\lambda M = \mathbf{h}_r^t \quad (4.7)$$

Nous cherchons à minimiser l'erreur entre λM et \mathbf{h}_r^t selon la norme de Frobenius pour calculer la constante λ et l'homographie \mathbf{h}_r^t :

$$\min_{\lambda, \mathbf{h}_r^t} \|\lambda M - \mathbf{h}_r^t\|_F^2 \text{ avec } (\mathbf{r}_m^I)^T \mathbf{r}_m^I = I_{3 \times 3} \quad (4.8)$$

Les deux premières colonnes de la matrice de rotation \mathbf{r}_m^I sont déterminées par les deux premières colonnes de \mathbf{h}_r^t . La troisième colonne de \mathbf{r}_m^I est calculée par le produit vectoriel des deux premières. La matrice de rotation est une matrice orthogonale de déterminant 1. Quel que soit λ , les deux premières colonnes de \mathbf{h}_r^t qui minimisent l'équation 4.8 sont toujours les mêmes. $(\mathbf{h}_r^t)_{(1,2)}$ peut être calculé par l'équation simplifiée de 4.8 :

$$\min_{(\mathbf{h}_r^t)_{(1,2)}} \|(M)_{(1,2)} - (\mathbf{h}_r^t)_{(1,2)}\|_F^2 \text{ avec } (\mathbf{r}_m^I)_{(1,2)}^T (\mathbf{r}_m^I)_{(1,2)} = I_{2 \times 2} \quad (4.9)$$

où $(\cdot)_{(1,2)}$ représente les deux premières colonnes de la matrice.

Nous considérons la décomposition en valeurs singulières de $(M)_{(1,2)} = U_{3 \times 2} \Sigma_{2 \times 2} V_{2 \times 2}^T$. Les deux premières colonnes de la matrice de rotation peuvent être calculées [Eve97] :

$$(\mathbf{h}_r^t)_{(1,2)} = U_{3 \times 2} V_{2 \times 2}^T \quad (4.10)$$

Nous obtenons également la matrice de rotation :

$$\mathbf{r}_m^I = \left((\mathbf{h}_r^t)_{(1)} \quad (\mathbf{h}_r^t)_{(2)} \quad (\mathbf{h}_r^t)_{(1)} \wedge (\mathbf{h}_r^t)_{(2)} \right) \quad (4.11)$$

où $(\mathbf{h}_r^t)_{(i)}$ exprime la $i^{\text{ième}}$ colonne de la matrice de rotation \mathbf{h}_r^t .

Nous cherchons λ optimum de l'équation 4.8. Selon la définition de la norme de Frobenius, nous obtenons :

$$\begin{aligned} \|\lambda(M)_{(1,2)} - (\mathbf{h}_r^t)_{(1,2)}\|_F^2 &= Tr(\lambda(M)_{(1,2)} - (\mathbf{h}_r^t)_{(1,2)}) Tr(\lambda(M)_{(1,2)} - (\mathbf{h}_r^t)_{(1,2)}) \\ &= \lambda^2 Tr((M)_{(1,2)}^T (M)_{(1,2)}) - 2\lambda Tr((\mathbf{h}_r^t)_{(1,2)}^T (M)_{(1,2)}) + Tr((\mathbf{h}_r^t)_{(1,2)}^T (\mathbf{h}_r^t)_{(1,2)}) \\ &= \lambda^2 Tr((M)_{(1,2)}^T (M)_{(1,2)}) - 2\lambda Tr((\mathbf{h}_r^t)_{(1,2)}^T (M)_{(1,2)}) + I_{2 \times 2} \\ &= B \end{aligned} \quad (4.12)$$

Quand $\frac{dB}{d\lambda} = 0$, nous pouvons retrouver le minimum de B , λ peut être donc déterminé :

$$\begin{aligned} \frac{dB}{d\lambda} &= 2\lambda \text{Tr}((M)_{(1,2)}^T (M)_{(1,2)}) - 2\text{Tr}((\mathbf{h}_r^t)_{(1,2)}^T (M)_{(1,2)}) = 0 \\ \Rightarrow \lambda &= \frac{\text{Tr}((\mathbf{h}_r^t)_{(1,2)}^T (M)_{(1,2)})}{\text{Tr}((M)_{(1,2)}^T (M)_{(1,2)})} \end{aligned} \quad (4.13)$$

Connaissant la matrice de rotation \mathbf{r}_m^I , la matrice de translation \mathbf{t}_m^I est calculée à partir l'équation 4.6 et 4.7 :

$$\begin{aligned} \lambda M &= \begin{pmatrix} (\mathbf{r}_m^I)_{(1,2)} & \mathbf{t}_m^I \end{pmatrix} \\ \Rightarrow \mathbf{t}_m^I &= \lambda M \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned} \quad (4.14)$$

Il existe deux solutions dans le cas général pour \mathbf{t}_m^I et \mathbf{r}_m^I , et la deuxième solution s'obtient par :

$$\begin{aligned} \mathbf{r}_m^{I'} &= \mathbf{r}_m^I \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \mathbf{t}_m^{I'} &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \mathbf{t}_m^I \end{aligned} \quad (4.15)$$

Nous pouvons choisir facilement la bonne solution convenant à notre problème à partir du troisième élément de \mathbf{t}_m^I selon le côté visible du plan. Par exemple, dans la figure 4.5, le côté visible de notre plan est négatif. Nous devons donc choisir la solution pour laquelle le troisième élément de \mathbf{t}_m^I est négatif.

4.2.2 Réflexion des rayons par la surface du miroir

Le miroir utilisé dans nos études est une surface de révolution. Nous disposons d'un nuage de points qui correspond à la moitié du profil du miroir [IEN05] (Fig. 4.6). La rotation de cette courbe autour de l'axe Z engendre la surface 3D du miroir. Comme la surface du miroir comporte un nombre infini de profils, et que tous les profils sont identiques, nos études sont basées sur un seul profil pour faciliter les calculs.

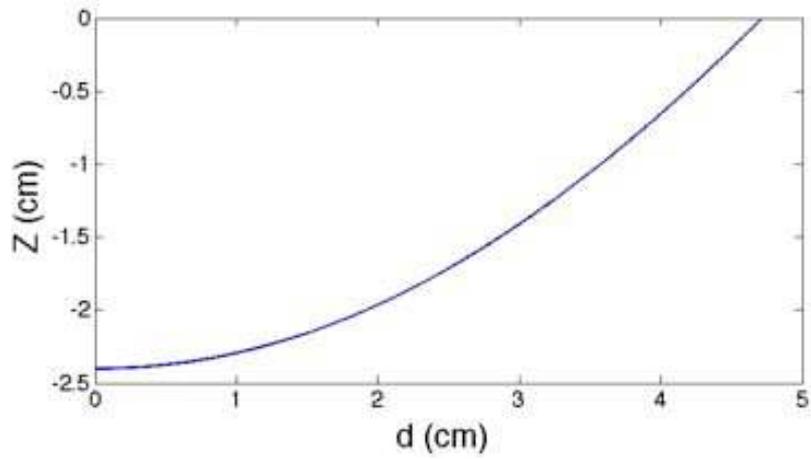


FIG. 4.6: Le profil du miroir utilisé dans nos études.

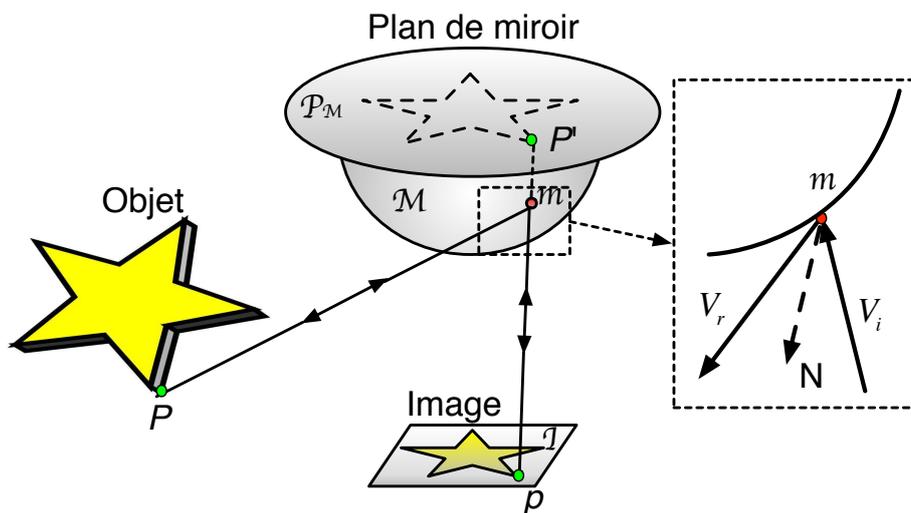


FIG. 4.7: Un exemple de la réflexion des rayons par la surface du miroir.

La figure 4.7 donne un exemple de réflexion de rayons par la surface du miroir. Avec l'homographie h_m^I calculée dans l'étape précédente, nous pouvons projeter tous les pixels d'image \mathcal{I} dans le plan du miroir \mathcal{P}_M . Le point incident m est obtenu par l'intersection de la droite \mathbf{pP}' et la surface du miroir \mathcal{M} . Le calcul des points incidents est expliqué précisément dans [IEN05]. La réflexion du rayon est calculée par symétrie orthogonale, comme représentée dans la partie zoomée de la figure 4.7. Pour tous les points d'un profil du miroir, nous pouvons reconstruire les rayons incidents et les rayons réfléchis (4.8a). Si nous connaissons la hauteur du système de vision, les intersections des rayons peuvent être calculées. Nous pouvons ainsi déterminer les positions des points observés. La figure 5.1b décrit un exemple de la résolution de la distribution des rayons pour un profil du miroir avec une hauteur de 2 m . Elle est non linéaire.

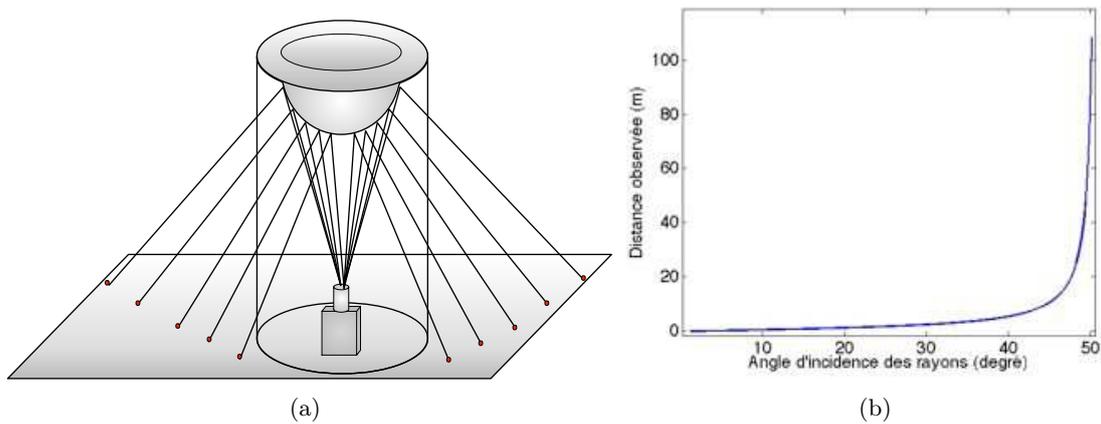


FIG. 4.8: (a) Les rayons réfléchis pour un profil du miroir. (b) La résolution de la distribution des rayons pour un profil du miroir avec une hauteur de 2 m .

IENG a présenté quelques propriétés sur les vecteurs réfléchis qui permettent d'écrire l'égalité suivante entre m et P dans l'image 4.7 :

$$P = m + \alpha \mathbf{V}_r \quad (4.16)$$

où α est un scalaire, P est l'intersection d'un rayon réfléchi avec le plan qui contient l'objet détecté.

Les trois coordonnées d'intersection P peuvent être calculées comme :

$$\begin{cases} P_x = m_x + \alpha \mathbf{V}_{\mathbf{rx}} & (4.17a) \\ P_y = m_y + \alpha \mathbf{V}_{\mathbf{ry}} & (4.17b) \\ P_z = m_z + \alpha \mathbf{V}_{\mathbf{rz}} & (4.17c) \end{cases}$$

Une paramétrisation standard d'un plan que nous pouvons utiliser est la suivante :

$$ax + by + cz + d = 0 \quad (4.18)$$

Si nous connaissons les quatre paramètres a, b, c, d du plan qui contient l'objet capturé par la caméra, en combinant l'équation 4.17 et 4.18, nous constatons que :

$$\alpha = \frac{-(am_x + bm_y + cm_z + d)}{a\mathbf{V}_{\mathbf{rx}} + b\mathbf{V}_{\mathbf{ry}} + c\mathbf{V}_{\mathbf{rz}}} \quad (4.19)$$

Nous obtenons aussi les coordonnées du point P que nous allons utiliser plus loin pour la détection des obstacles.

4.3 Estimation de l'orientation du bateau

Après calibrage, notre caméra panoramique peut être utilisée pour la phase de perception. Un système de coordonnées globales doit être définie dans un premier temps pour localiser tous les capteurs en relation avec le voilier. Le cadre choisi est celui défini par les axes du bateau qui est considéré comme elliptique. En utilisant la caméra, la position du bateau est déterminée à partir des images. Le bateau est d'abord segmenté à partir du fond de l'image et la matrice de covariance de la distribution des pixels est calculée. Les directions des axes principaux de cette ellipse sont alors données par les vecteurs propres de la matrice (Fig. 4.9). Toute structure ou tout objet de la scène seront désormais référencés dans ce cadre dans la suite de ce rapport.

4.3.1 Algorithme

Une fois le capteur calibré selon la méthode décrite plus haut et placé sur le mât, il est possible de calculer la projection de chaque pixel sur le plan défini par l'eau en faisant

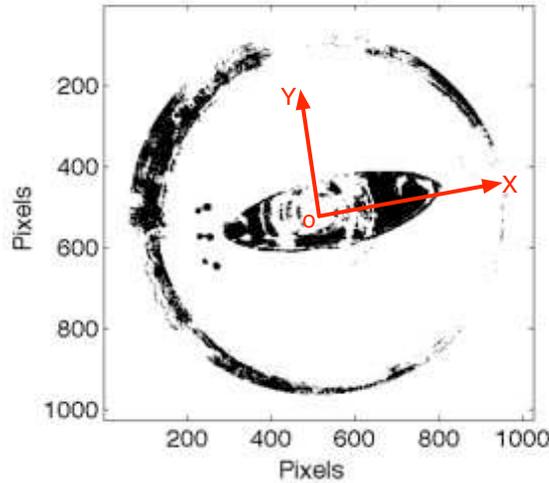


FIG. 4.9: L'axe du bateau est utilisé pour définir un système de coordonnées globales en supposant que la caméra est placée perpendiculairement à cet axe.

l'hypothèse que la mer est calme, si l'orientation relative de ce plan et du capteur de vision (identique à celui du bateau) peuvent être estimée.

Cette opération est nécessaire si nous voulons extraire des informations métriques à partir des images sans avoir besoin de la stéréovision. Comme nous l'avons fait pour l'estimation de la pose de la caméra par rapport à miroir, nous pouvons estimer l'orientation du capteur omnidirectionnel en détectant les projections de la surface de la mer sur l'image catadioptrique. Pour réaliser cette tâche, nous plaçons dans l'eau plusieurs bouées aisément détectables qui agissent comme des amers pour repérer le plan de la mer (Fig. 4.10).

Si la structure définie par les bouées est connue, c'est à dire si les distances relatives des bouées sont connues, suffisamment de contraintes géométriques peuvent être fournies pour l'estimation du plan. Nous allons aussi utiliser la paramétrisation standard d'un plan, donnée par l'équation 4.18. n bouées sont utilisées pour construire l'amer, la distance entre la $i^{\text{ème}}$ et la $j^{\text{ème}}$ bouée, D_{ij} , est mesurée à la main. Il y a $m = \binom{n}{2} = \frac{n!}{2!(n-2)!}$ distances ainsi définies. m est plus élevé que le nombre de paramètres pour $n \geq 4$ (dans notre cas $n = 6$). Les bouées projetées sur l'appareil photo sont également réduites à leurs centroïdes \mathbf{p}_i dans le plan d'image et, avec le calibrage du capteur, n lignes 3D $(\mathbf{m}_i, \mathbf{u}_i)$ sont calculées pour n centroïdes.

Avec un capteur non-central, il est probable que pour une distance donnée, un seul plan croise un faisceau de rayons qui dans notre cas est l'ensemble des rayons associés aux bouées.

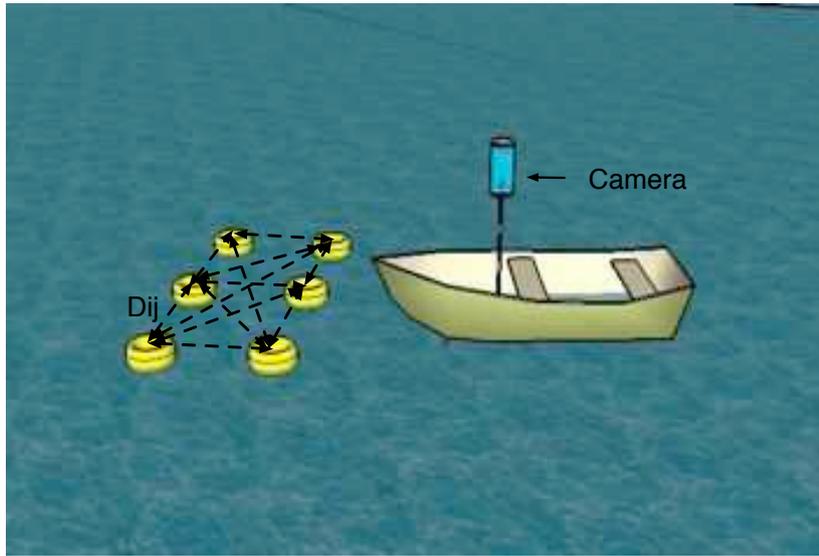


FIG. 4.10: Les amers utilisés pour l'estimation de l'orientation.

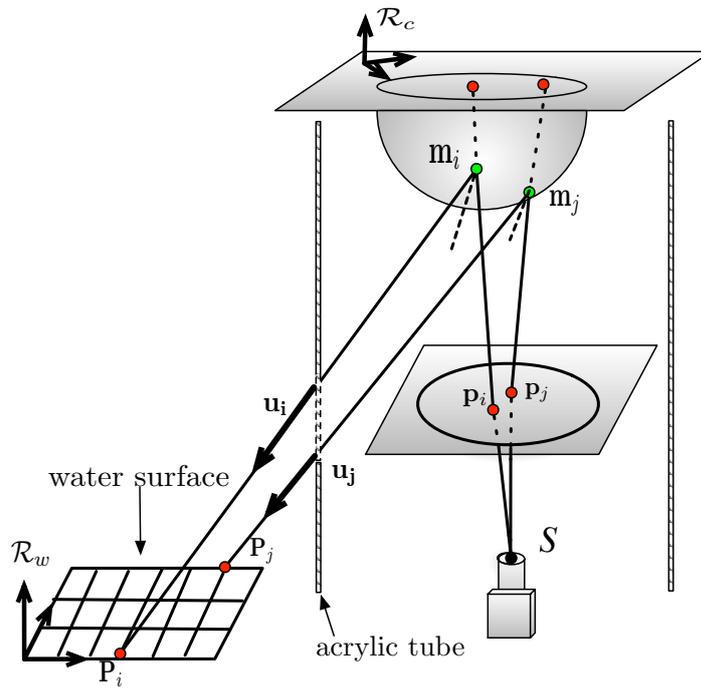


FIG. 4.11: Un calibrage complet du capteur catadioptrique fournit un mapping des pixels de l'image \mathbf{p}_i sur une ligne 3D $(\mathbf{m}_i, \mathbf{u}_i)$ et son intersection avec la surface définie par l'eau. \mathbf{P}_i peut être calculée si l'orientation relative du bateau par rapport à la surface de l'eau est connue.

Nous utilisons une méthode basée sur la méthode de Hough pour la détermination du plan. a, b, c sont calculées de -1 à 1 avec un pas constant de 0,002 et d est calculée, de 1 à 500 avec un pas constant de 1. Nous définissons π_k , le k^{ieme} plan qui coupe le faisceau de rayons. Chaque rayon coupe le plan et la distance entre deux intersections est :

$$D_{ij}^k = |\mathbf{P}_i^k - \mathbf{P}_j^k| \quad (4.20)$$

à \mathbf{P}_i^k et \mathbf{P}_j^k sont les positions des intersections pour les points i et j . Si π_k est le plan que nous cherchons, il vérifiera la condition $D_{ij} \approx D_{ij}^k$ (Fig. 4.12).

L'orientation de la surface de l'eau \mathbf{n}^{Eau} dans le système de coordonnées de la caméra est :

$$\mathbf{n}^{\text{Eau}} = (a, b, c)^T \quad (4.21)$$

En utilisant le changement de repère, l'orientation du bateau $\mathbf{n}^{\text{Bateau}}$ peut être calculée. A partir de la pose de la caméra par rapport à la surface de la mer, nous pouvons associer à n'importe quel pixel \mathbf{p}_i un point 3D \mathbf{P}_i sur le plan (Fig. 4.11).

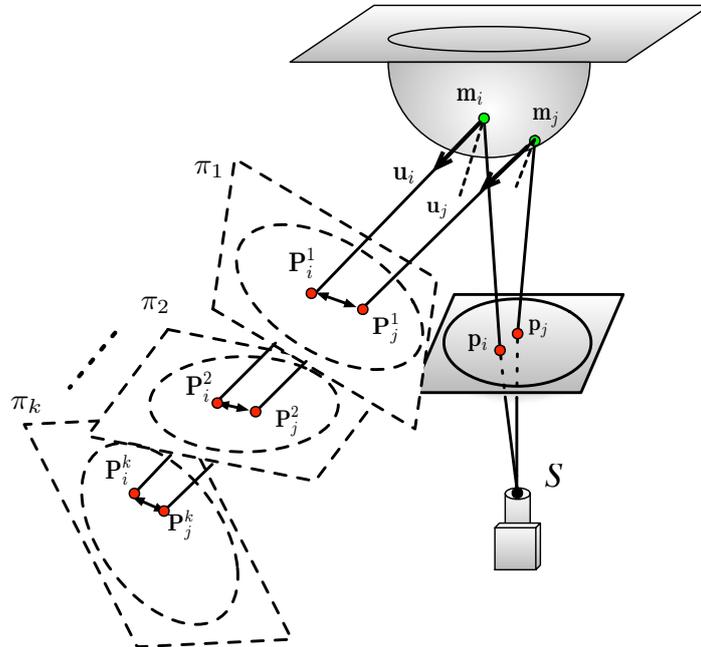


FIG. 4.12: Ajustement du plan sur le capteur calibré.

La distance entre chaque pixel et le bateau dans l'image peut être calculée comme : $d = |\mathbf{p}_i|$ selon les coordonnées globales introduites plus tôt dans ce chapitre. Ce dernier résultat permet de construire un plan de résolution du capteur sur la surface de la mer puisque nous pouvons projeter tous les pixels dessus (Fig. 4.3.1).

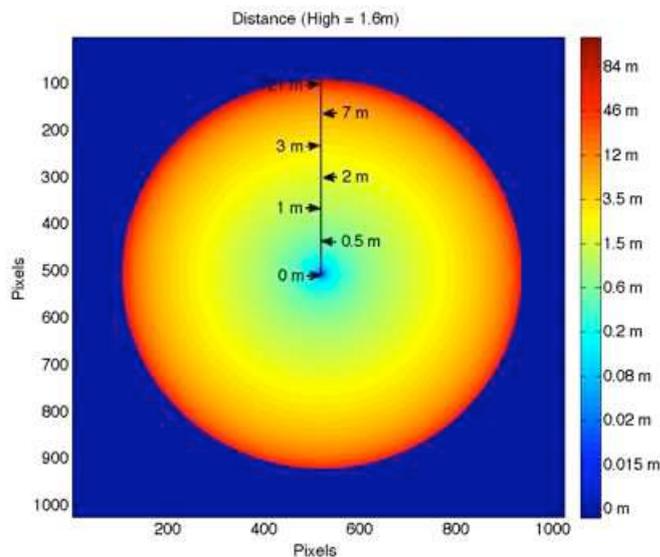


FIG. 4.13: L'estimation des distances pour chaque pixel dans l'image.

Le plan de résolution est construit en considérant un mât normal au plan de la surface de l'eau (mer calme). Cette hypothèse est cependant restrictive. Pour prendre en compte les oscillations présentées à la surface de l'eau (vagues), il convient de recalculer le plan de résolution pour différentes orientations du bateau par rapport à la surface de la mer (Fig. 4.14). Dans ce cas, un calcul lourd est nécessaire, car il y a de nombreuses possibilités, et la précision des résultats dépend également des orientations pour le voilier. Pour affranchir de ces inconvénients, nous expliquerons une nouvelle méthode utilisant la fusion de donnée provenant de la caméra et de la centrale inertielle dans le prochain chapitre. Cette approche permet de calculer rapidement l'orientation du bateau dans tous les cas. Les calculs sont alors réduits.

4.3.2 Résultats des tests

La méthode d'estimation du plan de l'orientation est testée avec des données synthétisées en générant aléatoirement 5 plans. Pour chaque plan, 6 points sont placés aléatoirement en supposant que les distances relatives entre chaque point sont connues. Pour chaque réglage, nous exécutons l'algorithme de détection du plan. Cette opération est répétée 100 fois, ce qui

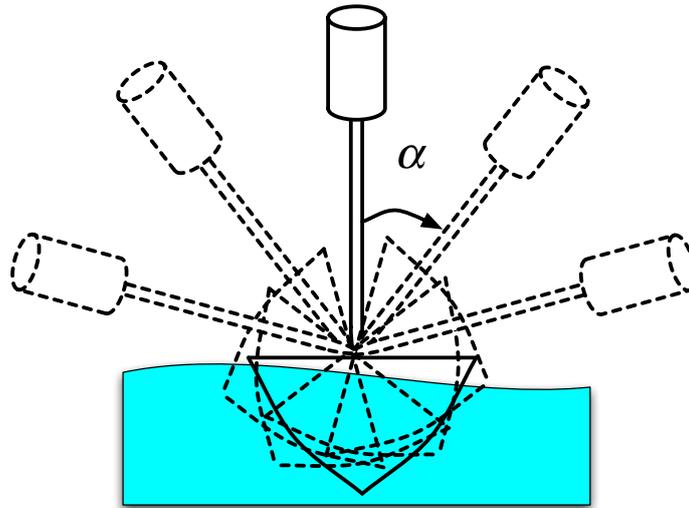


FIG. 4.14: Oscillation du bateau dans une mer agitée.

permet de fournir suffisamment de résultats pour produire des statistiques significatives. Le tableau 4.1 résume les résultats obtenus.

TAB. 4.1: Résultats des paramètres estimés

Paramètres	Taux d'erreur d'estimation des paramètres(%)			
	a	b	c	d
Cas 1	4	1	1	3.3
Cas 2	1	0	0	0
Cas 3	1	2	0	6.7
Cas 4	1	0	1	3.2
Cas 5	2	1	0	0.5

Selon le résultat, la précision de l'algorithme d'estimation du plan est raisonnable avec des erreurs qui sont inférieures à 7% et une valeur moyenne de 1.43%.

4.4 Détection d'obstacle

La détection d'obstacles est faite en deux étapes. La première consiste à étudier la différence entre deux images consécutives pour avoir une position grossière des objets en mouvement considérés par défaut comme des obstacles. Cela met en évidence les zones de l'image

à traiter lors de la deuxième étape.

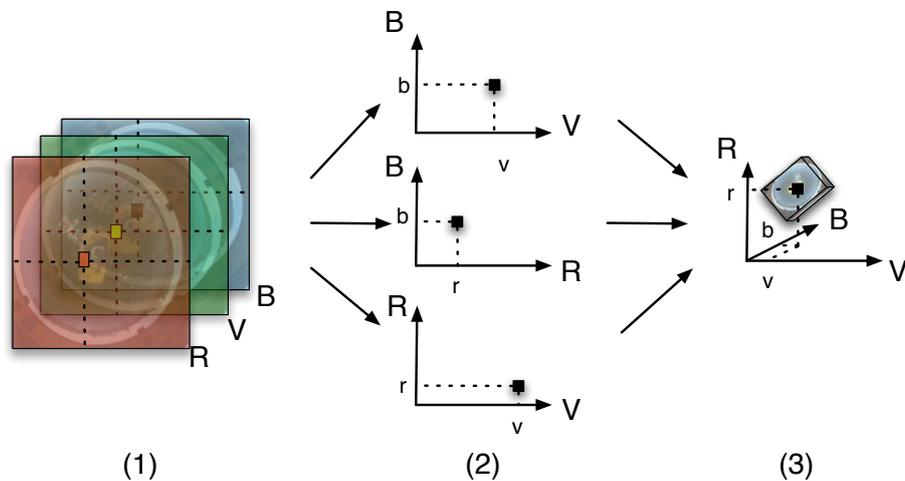


FIG. 4.15: Modèle colorimétrique calculé à partir d'une image dans la base de référence inspiré par [Cha09].

CHANG [Cha09] a proposé une méthode colorimétrique qui permet de segmenter un objet à partir du fond dans une image. Ce critère colorimétrique est utilisé ensuite pour réaliser la segmentation d'objets dans les régions mises en évidence par l'étude des différences entre deux images consécutives.

Un pré-test d'acquisition d'images doit être réalisé pour construire une base de référence. Pour toutes les images dans la base, une signature de couleur de la mer est calculée en sélectionnant des régions représentant de l'eau. Chaque pixel de ces régions est reprojété sur les trois plans des coordonnées RVB selon ses composantes de couleur. Cette opération produit un cluster dans chaque plan qui est représentatif de la distribution de la couleur du fond qui est dans notre cas, la mer. La figure 4.15 montre le modèle calculé à partir d'une image et les projections d'un point de l'espace RVB dans chaque plan.

Pour segmenter un pixel d'objet à partir du fond, nous le projetons dans le modèle de référence que nous avons créé à partir du pré-test. Il sera considéré comme un obstacle si au moins une de ses composantes RVB n'appartient pas aux clusters définis précédemment.

La figure 4.16 montre un exemple de segmentation d'un amer formé par six bouées jaunes en utilisant la combinaison de la différence d'images et la signature colorimétrique de l'eau.

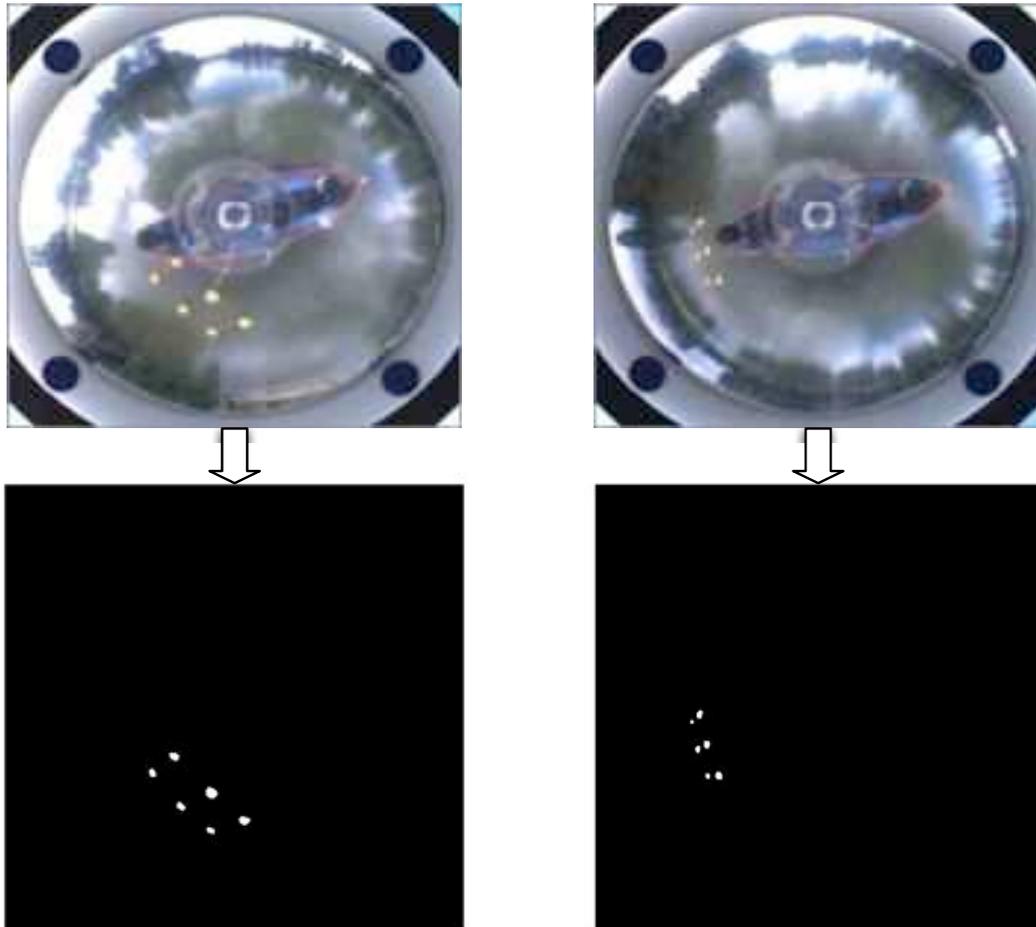


FIG. 4.16: Détection d'amers utilisant la signature colorimétrique.

4.5 Estimation de l'état de la mer par la vision : détection de l'orientation

Le mouvement de la houle définit l'orientation de la maquette. Une estimation de l'orientation de la maquette est possible grâce à la vision panoramique. Afin de réaliser cette détection, trois étapes sont nécessaires. La première est la localisation de l'horizon. La seconde est l'ajustement de la courbe de l'horizon. Sa projection dans l'image a une forme non spécifique courbe que nous approximations par une ellipse quelconque. La dernière concerne la détermination du plan d'incidence. Le capteur ne possédant pas un centre de projection unique, la projection sur la surface miroir de cette ellipse n'est en général pas une courbe plane. Cependant elle permet de définir un plan moyen. La détermination de ce plan et plus particulièrement de

sa normale nous permet ensuite de mesurer directement la variation de l'orientation de la maquette par rapport à l'eau (Fig. 4.17).

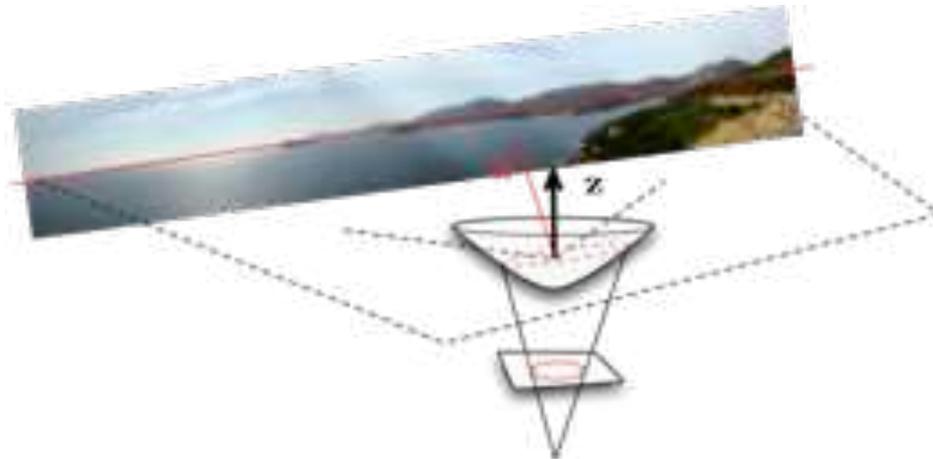


FIG. 4.17: Estimation de l'orientation du capteur omnidirectionnel via l'horizon.

4.5.1 Segmentation du ciel et de l'eau

En physique, l'horizon est un cercle centré sur l'observateur entre le ciel et la Terre (ou la mer) [Wika]. Dans le domaine du traitement d'images, la détection de l'horizon consiste à détecter la ligne de séparation entre le ciel et la Terre (ou la mer). Comme le montre l'image 4.17, le ciel se distingue de la mer principalement par une différence colorimétrique. Comme pour la détection d'obstacles, une première étape du traitement consiste à segmenter ces deux régions, comme nous l'avons présenté précédemment. Une trame d'images est choisie comme base de référence. Les couleurs des autres images seront normalisées par rapport à cette trame.

4.5.2 Détection du contour horizon

Une fois le modèle de référence construit, les images à traiter sont alors normalisées par rapport au modèle colorimétrique (Fig 4.18).

Comme sur l'image 4.18b, la surface de l'eau détectée est souvent bruitée, sa couleur n'est pas uniforme à cause de la réflexion. Nous utilisons ici des techniques standard de morphologie mathématique pour éliminer les bruits. Supposons que I_t est une image après la segmentation, α et β sont les nombres d'opérations réalisées (Fig. 4.18c). Nous utilisons :

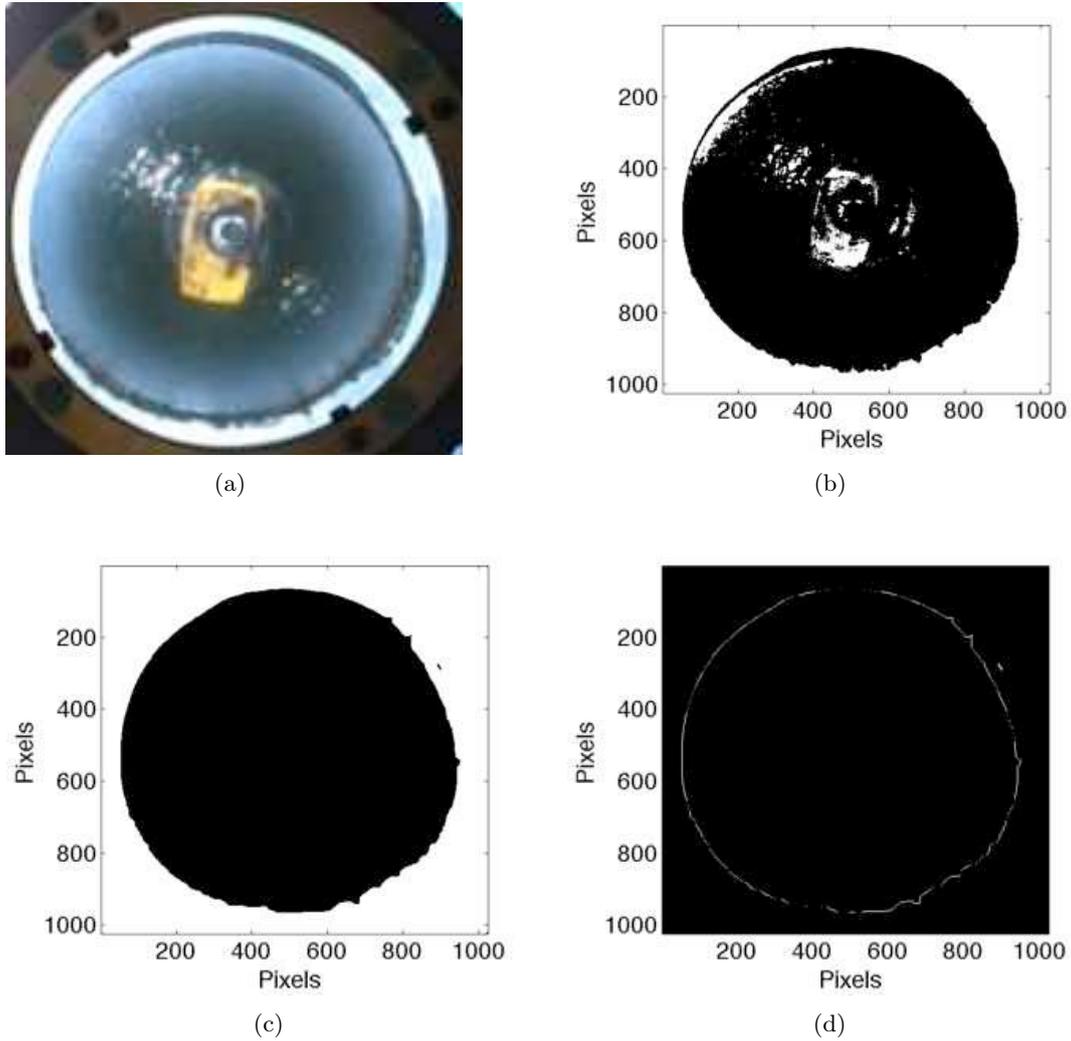


FIG. 4.18: (a) Image panoramique prise dans un lac . (b) Après la détection, la surface de l'eau est reconnue. (c) Les bruits sont supprimés en utilisant des techniques standard de morphologie mathématique. (d) Le contour horizon détecté.

$$I_t = \alpha \times (I_t \ominus S) + \beta \times (I_t \oplus S) \quad (4.22)$$

où S désigne l'élément structurant utilisé, \ominus et \oplus sont les opérations de dilatation et d'érosions morphologiques par l'élément structurant S . Le contour I_c définissant la ligne de séparation peut être déterminé selon l'équation :

$$I_c = I_t \oplus S - I_t \quad (4.23)$$

La forme du contour obtenu ici n'est pas spécifique (Fig. 4.18d), nous devons ajuster la courbe obtenue. Dans notre cas, c'est une ellipse. Halir et al. [HF98] ont proposé une méthode améliorée des moindres carrés basée sur l'algorithme de Fitzgibbon [FPF96] pour l'interpolation d'ellipse. Nous utilisons l'algorithme de Halir, avec les paramètres obtenus, nous pouvons donc ajuster l'horizon détecté, et reformer une ellipse spécifique.

4.5.3 Détermination du plan d'incidence lié à l'horizon

Dans la partie calibrage, nous pouvons mettre en correspondance les points du miroir avec les pixels de l'image. Avec l'ellipse de l'horizon obtenue précédemment, les points correspondants sur le miroir peuvent être calculés facilement. Ces points sont souvent sur différents plans. Pour définir un plan unique, nous devons calculer les valeurs propres pour tous les plans possibles. Le plan ayant la valeur propre minimale est choisi comme le plan \mathcal{I} contenant l'ellipse (Fig 4.19). En comparant avec le plan référence (le plan horizontal), l'orientation du bateau pourra être déterminée.

Soit deux vecteurs \mathbf{h} et \mathbf{n} correspondant respectivement à la direction horizontale et à la normale au plan de l'ellipse. L'angle θ entre les deux vecteurs est :

$$\theta = \arccos \frac{\mathbf{h}\mathbf{n}}{|\mathbf{h}||\mathbf{n}|} \quad (4.24)$$

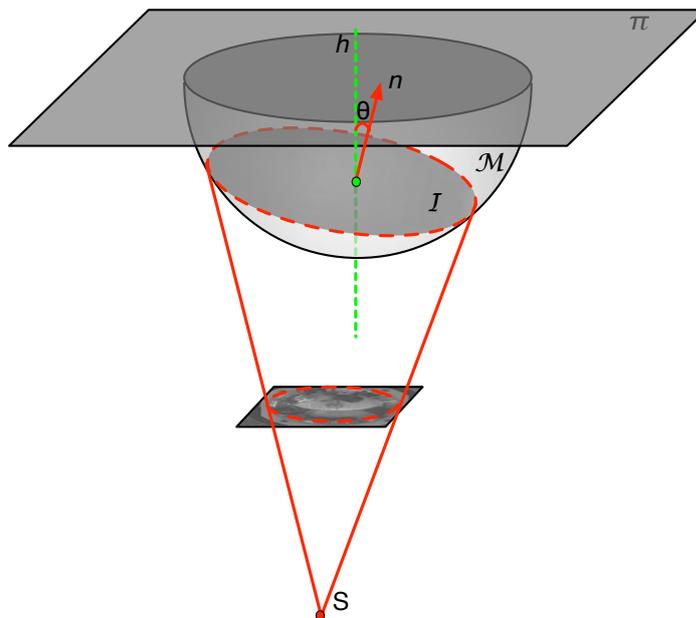


FIG. 4.19: Détermination du plan d'incidence lié à l'horizon

4.6 Résultats expérimentaux

Nous avons effectué deux tests indépendants : un premier test dans un lac calme pour valider la méthode de détection d'obstacle, un second test dans un bassin pour la détection d'horizon.

4.6.1 Test pour la détection d'obstacle

Nous avons réalisé cette expérimentation sur un lac calme (Fig. 4.20). Sa surface définit donc un plan plat. Un amer spécial avec 6 bouées (1 m de long \times 2 m de large) est utilisé pour ce test. Le tube de la caméra est fixé sur un trépied à 1 m de haut.



FIG. 4.20: Expérimentation réalisée dans un lac avec un bateau.

La position de l'amer est modifiée entre chaque enregistrement et la distance entre la caméra et le centre de l'amer est mesurée manuellement. Pour ce test, la distance varie de 2 m à 14 m.

A partir du tableau 4.2, notre algorithme semble avoir d'excellentes performances avec des taux d'erreur de moins de 8%. Les distances estimées sont plus petites que les distances mesurées. La valeur de l'erreur maximum d'estimation est de 0.63 m sur une distance de 14 m, et le minimum est de 0.19 m pour une distance de 8 m.

TAB. 4.2: Erreurs estimées des données réelles

	Distance mesurée (m)	Distance estimée (m)	Angle estimé (°)	Taux d'erreur distance estimée (%)
Cas 1	2	1.70	259	3.5
Cas 2	4	3.72	201	3.3
Cas 3	6	5.60	207	4.7
Cas 4	8	7.81	199	2.2
Cas 5	10	9.61	186	4.6
Cas 6	12	11.44	208	6.5
Cas 7	14	13.37	187	7.4

4.6.2 Test de détection de l'orientation dans un bassin

Un ensemble d'expériences a été réalisé au bassin de Saint-Cyr ². En premier lieu, l'étanchéité de la maquette a été contrôlée et plus particulièrement celle de tube de caméra ainsi que celle du boîtier embarquant l'ordinateur enregistrant les données, a été contrôlée. La flottabilité et la stabilité de la maquette ont ensuite été vérifiées avant de procéder aux tests d'acquisition des signaux. Ces derniers ont ensuite été effectués pour différentes situations (simulées) d'états de l'eau.

Deux types d'essais ont été réalisés pendant l'enregistrement des signaux. Le premier a été réalisé avec la plateforme au repos et le deuxième en la soumettant à un régime d'oscillation mécanique simulant une agitée (Fig. 4.21). Dans ce test, nous avons utilisé une caméra panoramique qui a été mise en haut du mât et une centrale inertielle pour comparer le résultat obtenu.

Le but de notre étude est de détecter l'horizon dans l'image. Cependant les images acquises dans le bassin sont peu adaptées pour cette opération en raison d'un environnement particulièrement encombré (par des bâtiments et de la végétation notamment). Une solution palliative fut d'utiliser le plan d'eau défini par le bassin et plus particulièrement ses bords qui permettent d'identifier le plan d'intersection avec le miroir mentionné en section 4.5.3.

L'image 4.22 présente un exemple de détection de la zone de transition eau-sol pour la caméra de surface symbolisées par les ellipses. La détection de surfaces est relativement efficace notamment grâce aux contrastes des milieux. Des résultats satisfaisants sont obtenus en utilisant le critère colorimétrique, et permet la détection de l'interface ciel-sol.

²Institut Jean Le Rond d'Alembert, 2 Place de la Gare de Ceinture, 78210 Saint Cyr l'Ecole France.

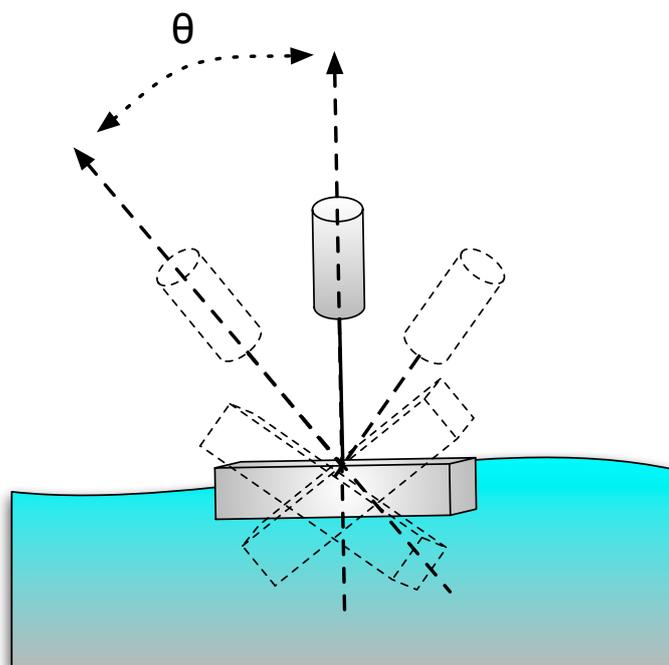


FIG. 4.21: Mouvements de la plateforme.

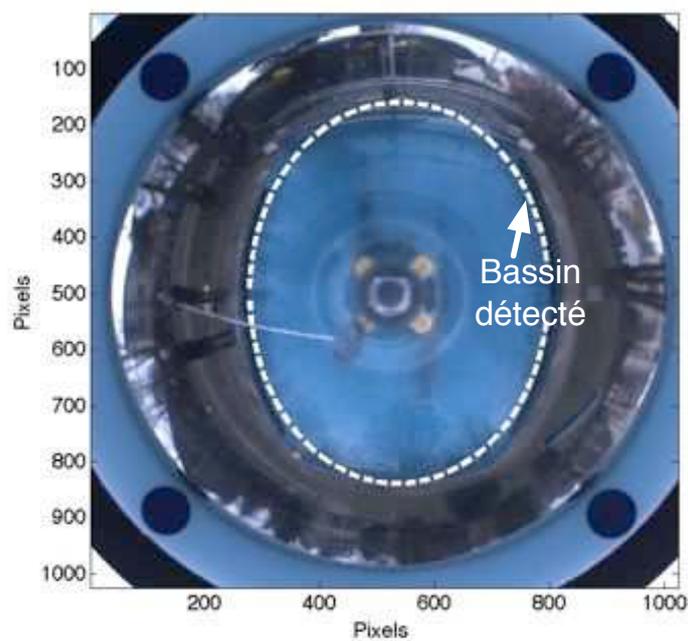


FIG. 4.22: Détection de la position du bassin sur l'image pour simuler la position de l'horizon.

L'orientation de référence choisie pour la maquette est celle pour laquelle elle est au repos, en absence de toute sollicitation mécanique. L'orientation relative est ensuite estimée à chaque instant en déterminant le plan d'intersection obtenu par projection sur le miroir de l'arc de la conique.

Une centrale inertielle simple d'Atomic³ est utilisée pour cet essai. Les orientations obtenues par la vision et par la centrale inertielle sont mises en corrélation dans la figure ci-dessous (Fig. 4.23). Des agrandissements sont effectués à cinq moments pour permettre l'analyse des données. En comparant les résultats obtenus, nous pouvons bien voir qu'ils ont presque la même fréquence. A cause de la discontinuité des pixels d'image, l'oscillation d'amplitude des résultats obtenus par la caméra panoramique est plus petite que celle de la centrale inertielle. Mis à part au début de la détection, il y a un décalage entre les résultats issus de la centrale inertielle et ceux de la caméra. Ce type d'erreur est causé par la faible performance de la centrale inertielle utilisée. Afin d'améliorer la précision des résultats obtenus, l'utilisation d'une centrale inertielle plus précise et plus performante est indispensable. La centrale inertielle XSENS⁴ est donc choisie et est utilisée pour les tests suivants.

³Centrale inertielle d'Atomic, <http://www.sparkfun.com/products/9184>

⁴XSENS MT9, <http://www.xsens.com>

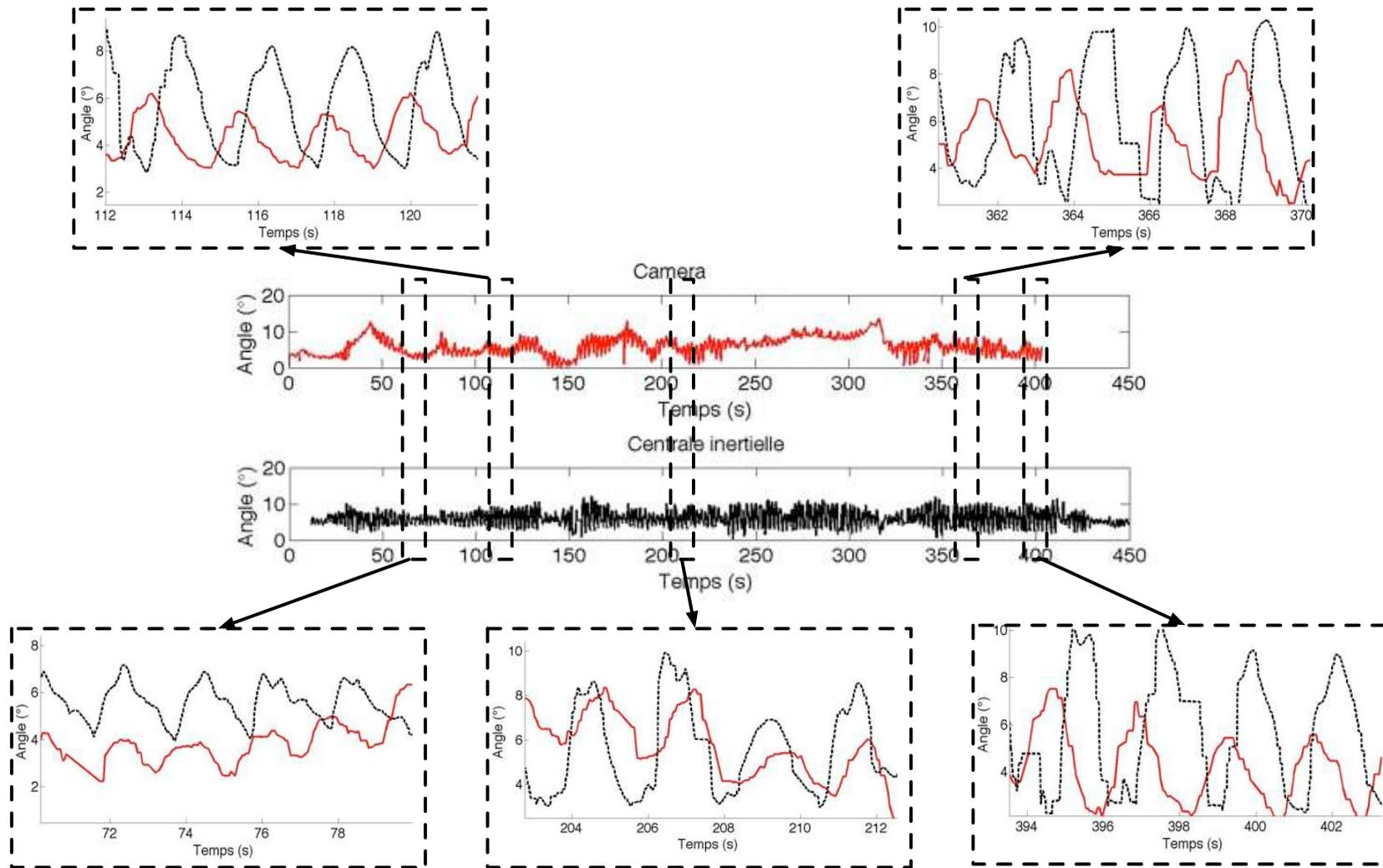


FIG. 4.23: Les angles calculés par la caméra (ligne continue) et la centrale inertielle (ligne en pointillé).

4.6.3 Test de détection de l'orientation en mer

Le test en mer a été réalisé au large de Deauville. Le voilier de notre projet n'étant pas disponible au moment de ce test, nous avons utilisé un bateau à moteur de type Antares 6.2 *m* (Fig. 4.24). Cependant, ceci ne modifie en rien la méthodologie utilisée, ni les algorithmes mis en oeuvre. Nous avons installé la caméra panoramique au milieu de ce bateau à 2.3 *m* de hauteur et fixé la centrale inertielle sous le tube transparent.



FIG. 4.24: Bateau à moteur utilisé.

Pendant la navigation, nous avons enregistré les signaux de mouvement du bateau avec la caméra panoramique et la centrale inertielle. La force du vent était de 5 à 6 beauforts, et la mer était agitée avec une hauteur de vague de 3 *m*. Le mouvement du bateau était donc significatif. Notre objectif est de tester les algorithmes développés pour la détection de l'orientation à l'aide d'images panoramiques. Les informations issues de la centrale inertielle sont utilisées comme référence pour ce test. L'image 4.25 donne un exemple de détection de l'horizon en utilisant le critère colorimétrique sur une image capturée en mer.

La figure 4.26 présente un extrait d'une douzaine de secondes de mesure. Les résultats obtenus par la caméra panoramique et par la centrale inertielle sont représentés sur la même figure. Nous voyons qu'ils ont la même fréquence, mais l'amplitude des variations des orientations obtenues par les images est plus faible que celle de la centrale inertielle. Il y a deux possibilités pour expliquer ce phénomène :

- A cause de la discontinuité des pixels, les images sont moins sensibles au changement d'orientation.
- A certains instants, une partie de la surface de la mer est hors de l'image (Fig. 4.27). Nous ne pouvons donc pas observer la surface entière, ce qui induit une erreur d'estimation de l'orientation.

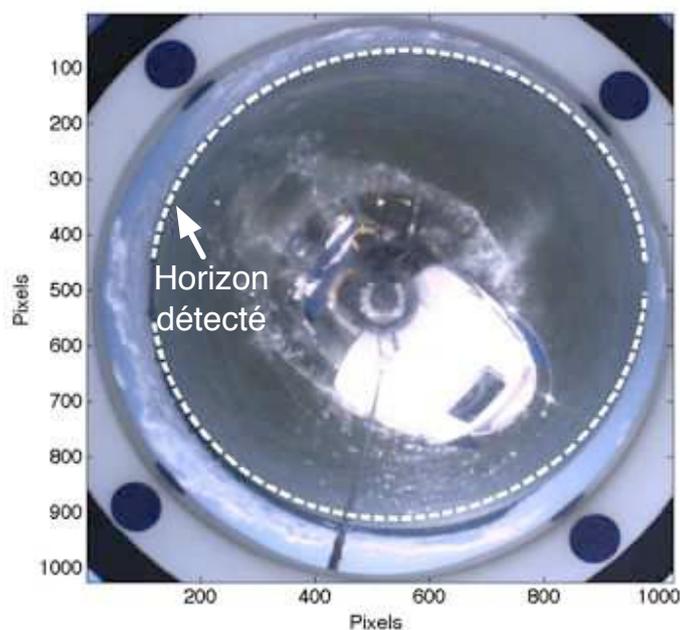


FIG. 4.25: Détection de la position de l'horizon sur une image prise en mer.

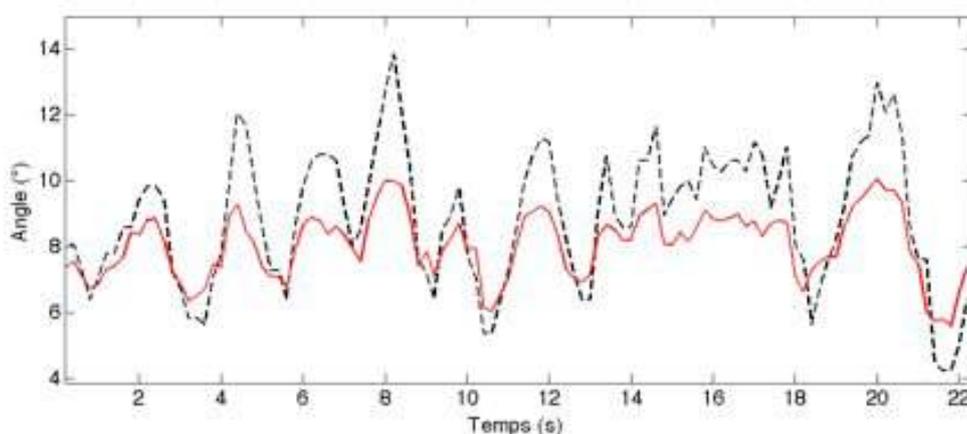


FIG. 4.26: Orientation du bateau calculée par la caméra (ligne continue) et la centrale inertielle (ligne en pointillé) à chaque instant.

Pour améliorer la précision de la détection de l'orientation par l'image panoramique une solution est d'augmenter la hauteur à laquelle est positionné le système de vision. Le champs de la caméra serait ainsi élargi. Dans ce test, selon le bateau utilisé, il était difficile d'installer

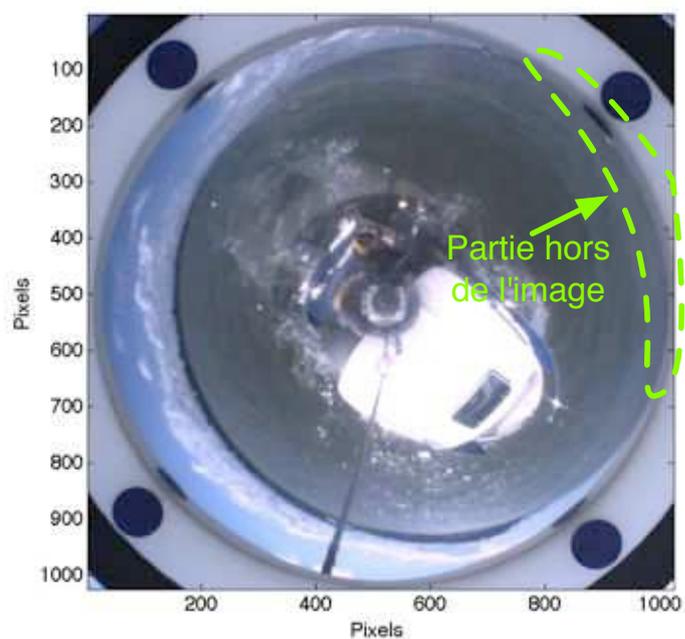


FIG. 4.27: Une partie de la surface de la mer est hors de l'image.

la caméra pour une hauteur de plus de 3 *m*. Mais pour notre voilier du projet, la caméra panoramique est positionnée à la hauteur de 5 *m*. Son champs de vision serait plus grand que celui dans le test, et sa performance de la perception pourrait ainsi être améliorée.

4.7 Conclusion

Dans ce chapitre, nous avons analysé le calibrage de la caméra pour un cas général. Puis une méthode d'estimation de l'orientation du bateau ainsi qu'une technique de détection d'obstacles utilisant un algorithme colorimétrique ont été proposés. Enfin, la méthode de détection de l'orientation a été présentée.

Nous avons effectué trois expériences. La première est réalisée dans un lac en région parisienne. Il met en évidence les bonnes performances de notre algorithme de détection d'obstacles. Cependant, une mire composée de six bouées est nécessaire, et la surface de l'eau doit être calme. Le deuxième test a été effectué dans un bassin. Les résultats obtenus par la caméra sont moins stables que ceux de la centrale inertielle qui est utilisée comme référence. Le troisième test a été réalisée en mer en utilisant un bateau à moteur. La détection de l'orientation du bateau par la caméra panoramique a été effectuée. Le chapitre suivant introduit une méthode plus générale pour la détection de l'orientation du bateau.

CHAPITRE

5

Systeme de fusion

Sommaire

5.1	Introduction	82
5.1.1	Niveau de fusion	83
5.1.2	Architectures topologiques	83
5.1.3	Systeme Multi-capteurs	85
5.2	Bloc d'integration de la vision panoramique et de la centrale inertielle	86
5.2.1	Traitement des donnees de la centrale inertielle	86
5.2.2	Detection de l'orientation du bateau	88
5.2.3	Definition de la position d'un objet	89
5.2.4	Resultats experimentaux	90
5.3	Bloc de fusion	95
5.3.1	Donnees capturees par le sonar	95
5.3.2	Pretraitement des donnees du sonar	99
5.3.3	Detection d'obstacles par le sonar	99
5.3.4	Calibration du sonar par rapport a la camera	102
5.4	Les methodes de fusion	102
5.4.1	Fusion par ponderation 1 (F1)	103

5.4.2	Fusion par pondération 2 (F2)	104
5.4.3	Fusion par estimation (F3)	104
5.4.4	Fusion par filtrage simple (F4)	105
5.4.5	Fusion par filtrage de Kalman (F5)	106
5.5	Résultats expérimentaux	110
5.6	Conclusion	115

5.1 Introduction

La fusion des capteurs est généralement définie comme l'utilisation de techniques qui combinent les données brutes ou les informations obtenues à partir des données brutes en provenant de plusieurs capteurs. Des informations sont recueillies afin de réaliser des inférences. Cette technique donne des résultats plus précis et plus robustes qu'avec un seul capteur [LW90, Hal92, Kle93, Var97, BRW07, Mit07]. Notre objectif est d'améliorer la qualité de l'information par rapport à celle qui serait obtenue avec un capteur individuel.

Les êtres humains et les animaux ont développé la capacité d'utiliser plusieurs sens pour améliorer leur capacité à survivre. Par exemple, il est difficile d'évaluer la qualité d'un gâteau en se basant uniquement sur le sens de la vue ou du toucher : Cependant une combinaison des données provenant de la vue, de l'odorat et du goût permet une meilleure évaluation. De la même manière, le sens de l'ouïe peut détecter des dangers que nous ne sommes pas forcément capable de percevoir par la vision uniquement. Ainsi, la fusion des données provenant de plusieurs capteurs est naturellement réalisée par les animaux et les humains pour obtenir une évaluation plus précise de leur environnement et identifier des menaces, afin d'améliorer leurs chances de survie.

Les applications de la fusion de données provenant de plusieurs capteurs sont très répandues. Elles incluent : la reconnaissance automatique de cible [PB08], le guidage par l'orientation de véhicules autonomes [SBD09, MNG⁺08], la télédétection [CSHL08], la surveillance des champs de bataille [ZH07, RS08], les systèmes de reconnaissance automatique tels que le système IFFN ¹ [HLL91], la surveillance des procédés de fabrication [AOM09, SS08], l'entretien de machines complexes [NYP10], la robotique [Lim07, LS08, GROJ08, MSB⁺09, GIR⁺11] et le domaine médical [Wu08]. Les techniques permettant de combiner ou de fusionner des données sont tirées d'un ensemble de disciplines, incluant le traitement du signal, l'estimation statistique, la théorie du contrôle, l'intelligence artificielle, et les méthodes

¹Identification Friend Foe Neutral

numériques [HLL91, KAB⁺92]. Historiquement, les méthodes de fusion de données ont été développés principalement pour des applications militaires. Toutefois, depuis quelques dizaines d'années ces méthodes sont appliquées dans le domaine civil.

5.1.1 Niveau de fusion

Les processus de fusion de données sont souvent classés en trois catégories [HL97, Law04] : *bas niveau*, *moyen niveau* et *haut niveau*, selon l'étape à laquelle s'effectue la fusion.

La fusion bas niveau (Fig. 5.1a) combine plusieurs sources de données brutes pour produire de nouvelles données brutes. Les données des capteurs peuvent être directement combinées si elles sont proportionnelles (par exemple, si les capteurs mesurent les mêmes phénomènes physiques comme deux capteurs de vision ou deux capteurs acoustiques). Les techniques de fusion de données incluent en général la détection classique et les méthodes d'estimation. Inversement, si les données des capteurs ne sont pas proportionnelles, elles doivent être fusionnées au niveau des caractéristiques ou au niveau de décision.

Le moyen niveau de fusion (Fig. 5.1b) est la fusion des caractéristiques qui sont extraites depuis des données provenant de différents capteurs. Elles sont concaténées en un vecteur unique de caractéristiques qui peut être traité par la reconnaissance des formes, les approches fondées sur les réseaux de neurones ou les algorithmes de clustering.

Enfin le haut niveau de fusion (Fig. 5.1c) est la fusion des décisions locales effectuées sur chacun des capteurs. Les méthodes de fusion des décisions comprennent la méthode des décisions pondérées, l'inférence classique et l'inférence bayésienne.

5.1.2 Architectures topologiques

Les systèmes de fusion peuvent être répartis en trois architectures topologiques différentes selon les noeuds de fusion fondamentale [Var97, Mit07] : la *topologie parallèle*, la *topologie série*, et la *topologie hiérarchique* incluant plusieurs sous fusions multi-capteurs.

Topologie parallèle Dans ce type de système, les informations des différents capteurs sont enregistrées en parallèle puis calculées par un centre de fusion (Fig. 5.2). Chaque capteur fournit les données enregistrées au processus de fusion.

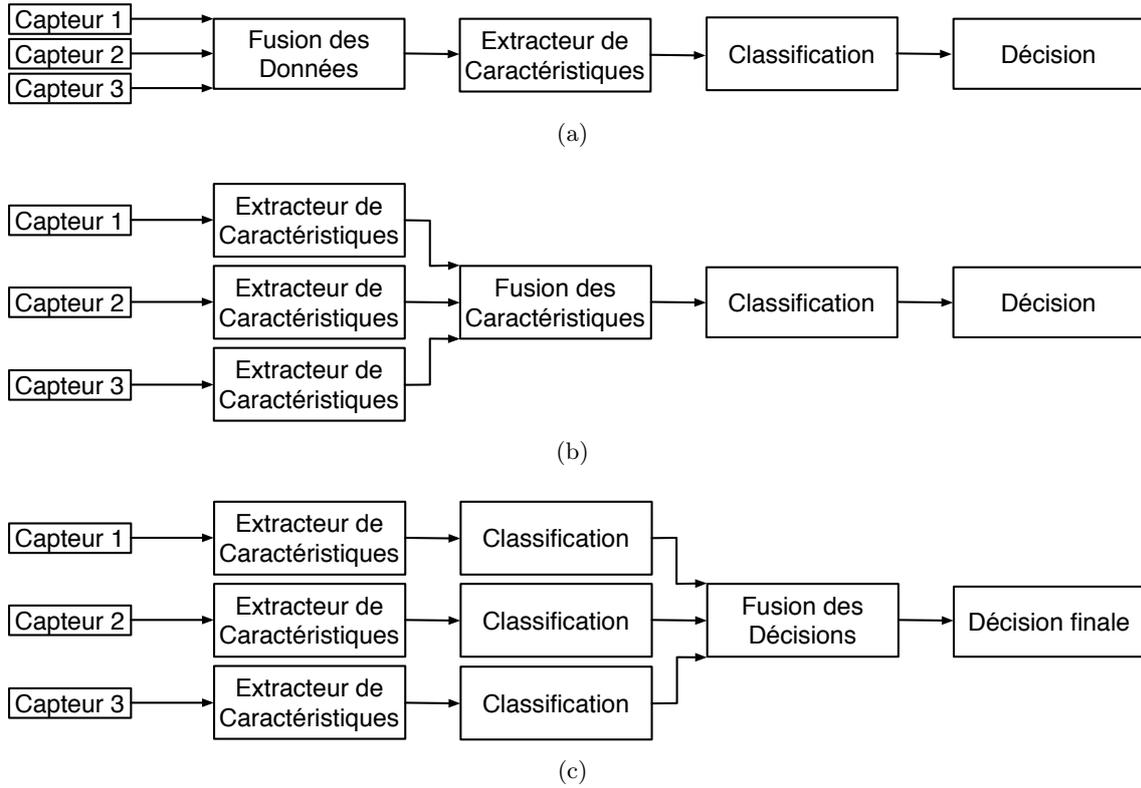


FIG. 5.1: Les trois niveaux de fusion : (a) fusion des données, (b) fusion des caractéristiques, et (c) fusion des décisions.

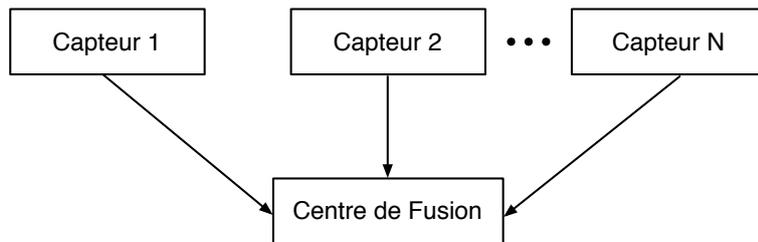


FIG. 5.2: Un exemple d'un système de fusion comportant N capteurs-architecture : topologie parallèle.

Topologie série La topologie série contient N cellules de capteurs. La $i^{\text{ième}}$ cellule de capteur avec, $i \in \{2, \dots, N\}$ reçoit les données de la $(i - 1)^{\text{ième}}$ cellule et envoie son résultat à la cellule suivante. Le centre de fusion du système est situé à la fin de cette chaîne (Fig. 5.3).

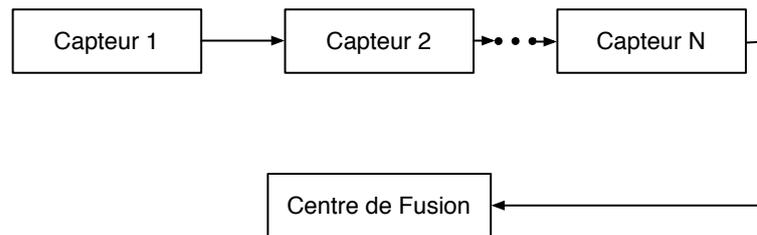


FIG. 5.3: Un exemple d'un système de fusion comportant N capteurs-architecture : topologie série.

Topologie hiérarchique Dans l'architecture hiérarchique il y a souvent plusieurs niveaux hiérarchiques. Le niveau supérieur contient un centre de fusion, et le niveau intermédiaire est constitué de plusieurs fusions locales. Chaque centre de fusion locale peut recevoir des données d'un groupe de capteurs (Fig. 5.4).

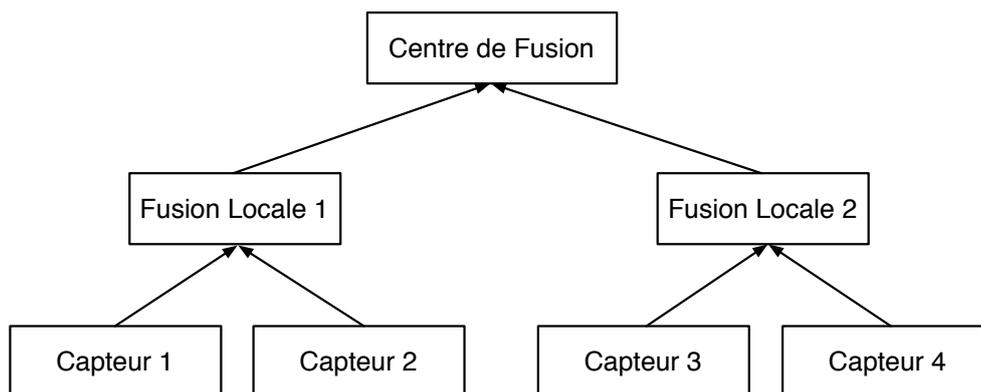


FIG. 5.4: Un exemple d'un système de fusion-architecture hiérarchique avec deux sous centres de fusion.

5.1.3 Système Multi-capteurs

Les tâches de perception sont importantes pour assurer la sécurité du bateau. Dans notre système de perception, plusieurs capteurs sont utilisés pour résoudre les problèmes de détection d'obstacles.

Notre système de perception est constitué de deux blocs principaux (Fig. 5.5) : le bloc d'intégration qui comporte le système de vision et la centrale inertielle. Les informations d'obstacle seront obtenues en sortie du bloc de fusion de données. Notre système peut être

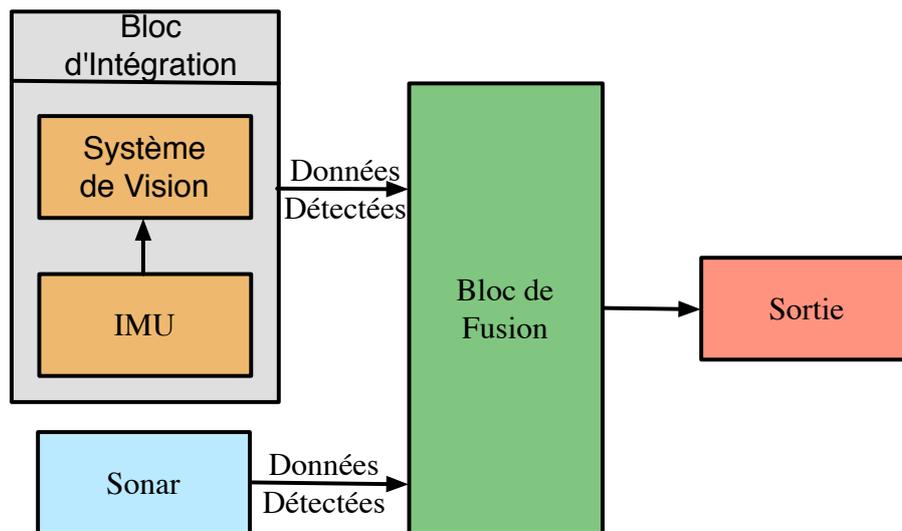


FIG. 5.5: L'architecture de notre système de perception composée de plusieurs capteurs.

considéré comme une architecture hybride dans laquelle nous mélangeons l'architecture des topologies parallèle, série et hiérarchique.

Dans ce chapitre, nous décrivons en détail ces deux blocs (le bloc d'intégration et le bloc de fusion) et expliquons les différentes méthodes de fusion que nous utilisons dans notre système.

5.2 Bloc d'intégration de la vision panoramique et de la centrale inertielle

En conditions naturelles, la surface des mers est souvent agitée. L'algorithme développé au chapitre 4 est difficile à appliquer car une mire composée d'au moins six bouées est nécessaire pour estimer chaque orientation du bateau et un calcul lourd est indispensable pour définir les plans de résolution selon l'orientation du bateau. Nous introduisons donc une centrale inertielle dans notre système pour détecter l'orientation du bateau et simplifier les calculs dans le cas d'un réel.

5.2.1 Traitement des données de la centrale inertielle

Les données renvoyées par la centrale inertielle sont des quaternions, qui fournissent une notation mathématique commode pour représenter l'orientation et la rotation d'objets en trois

dimensions [Ham47]. Dans notre cas, il faut détecter la matrice de rotation pour connaître le changement de positions du tube. Des calculs sont nécessaires pour convertir les quaternions en angles d'Euler, puis en matrice de rotation.

La conversion entre les quaternions et les angles d'Euler est donnée par [Die06] :

$$\begin{aligned}\phi &= \arctan\left(\frac{2(q_2q_3 + q_0q_1)}{q_0^2 - q_1^2 - q_2^2 + q_3^2}\right) \\ \theta &= \arcsin(-2(q_1q_3 - q_0q_2)) \\ \psi &= \arctan\left(\frac{2(q_1q_2 + q_0q_3)}{q_0^2 + q_1^2 - q_2^2 - q_3^2}\right)\end{aligned}\tag{5.1}$$

où q_0, q_1, q_2, q_3 sont les quaternions, et ϕ est l'angle de rotation autour de l'axe X , défini entre $[-180^\circ \dots 180^\circ]$, θ est l'angle de rotation autour de l'axe Y , défini entre $[-90^\circ \dots 90^\circ]$, ψ est l'angle de rotation autour de l'axe Z , définie entre $[-180^\circ \dots 180^\circ]$.

Dans un espace à trois dimensions, les matrices de rotation suivantes $R_x(\phi)$, $R_y(\theta)$ et $R_z(\psi)$ correspondent à des rotations autour des axes X, Y et Z [Sho94] :

$$\begin{aligned}R_x(\phi) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi & \cos \phi \end{bmatrix} \\ R_y(\theta) &= \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \\ R_z(\psi) &= \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}\end{aligned}\tag{5.2}$$

Chacune de ces matrices de rotation de base est définie dans le sens trigonométrique. Par exemple, R_z tourne de l'axe X vers l'axe Y .

La matrice de rotation R s'obtient dans le cas général à partir des matrices de base $R_x(\phi)$, $R_y(\theta)$ et $R_z(\psi)$ par multiplication matricielle. Dans notre cas, nous avons :

$$R = R_x(\phi)R_y(\theta)R_z(\psi)\tag{5.3}$$

qui représente une rotation composée du roulis ϕ , du tangage θ et du lacet ψ .

Si le vecteur initial \mathbf{U} est connu, le vecteur final \mathbf{U}' sera calculé avec la matrice de rotation R :

$$\mathbf{U}' = R\mathbf{U} \quad (5.4)$$

Une matrice de rotation peut être calculée à chaque instant avec les données obtenues depuis la centrale inertielle. Elle sera utilisée pour déterminer l'orientation du bateau.

5.2.2 Détection de l'orientation du bateau

En mathématiques, un plan peut être représenté par un vecteur normal à ce plan. Si le bateau est en mouvement, l'orientation initiale est importante pour déterminer l'orientation instantanée en temps réel par la méthode décrite dans le chapitre 4. Ce vecteur est ici noté $\mathbf{n}_I^{\text{Bateau}}$.

La centrale inertielle Xsens MT9 est fixée sous le tube de la caméra. Pour des raisons techniques, l'orientation de la centrale inertielle et celle de la caméra ne peuvent être rigoureusement identiques. Par contre, les changements d'orientation entre les deux capteurs sont équivalents. Ils ont la même matrice de rotation pendant le mouvement du bateau. Nous pouvons donc l'utiliser afin de déterminer l'orientation du bateau. La sortie de la centrale inertielle est un vecteur orientation de trois dimensions donnant l'orientation du repère lié au capteur par rapport à un repère fixe lié à la terre. Nous définissons \mathbf{E}_r comme le vecteur de référence dans le système de coordonnées lié à la centrale inertielle.

$\mathbf{n}_T^{\text{Bateau}}$ est le vecteur représentant l'orientation du bateau en temps réel. R_I^{Bateau} et R_T^{Bateau} sont les matrices de rotation du bateau. Elles sont identiques à celles de la centrale inertielle R_I^{IMU} et R_T^{IMU} .

$$\begin{cases} \mathbf{n}_I^{\text{IMU}} = R_I^{\text{IMU}} \mathbf{E}_r \\ \mathbf{n}_T^{\text{IMU}} = R_T^{\text{IMU}} \mathbf{E}_r \\ \mathbf{n}_T^{\text{IMU}} = R_G^{\text{IMU}} \mathbf{n}_I^{\text{IMU}} \end{cases} \Rightarrow R_G^{\text{IMU}} = R_T^{\text{IMU}} (R_I^{\text{IMU}})^{-1} \quad (5.5)$$

où $\mathbf{n}_I^{\text{IMU}}$ et $\mathbf{n}_T^{\text{IMU}}$ représentent l'orientation initiale et en temps réel de la centrale inertielle. Ils peuvent être calculés directement à partir des données détectées. R_G^{IMU} est la matrice de rotation globale.

Dans notre système, les mouvements de la centrale inertielle et du bateau sont synchronisés. Leurs matrices de rotations sont identiques. L'orientation du bateau $\mathbf{n}_T^{\text{Bateau}}$ en temps réel est calculée par :

$$\mathbf{n}_T^{\text{Bateau}} = R_G^{IMU} \mathbf{n}_I^{\text{Bateau}} \quad (5.6)$$

D'après l'équation 5.6, l'orientation instantanée du bateau peut être calculée en temps réel à partir de l'orientation initiale du bateau et de la matrice de rotation globale. Cette dernière est calculée par le rapport des matrices de rotation de la centrale inertielle.

5.2.3 Définition de la position d'un objet

Un système de coordonnées global $OXYZ$ est défini à partir du plan représentant la surface de l'eau, et l'axe Z perpendiculaire à cette surface. Un autre système de coordonnées global $O'X'Y'Z'$ référencé par le tube de la caméra est défini également (Fig. 5.6).

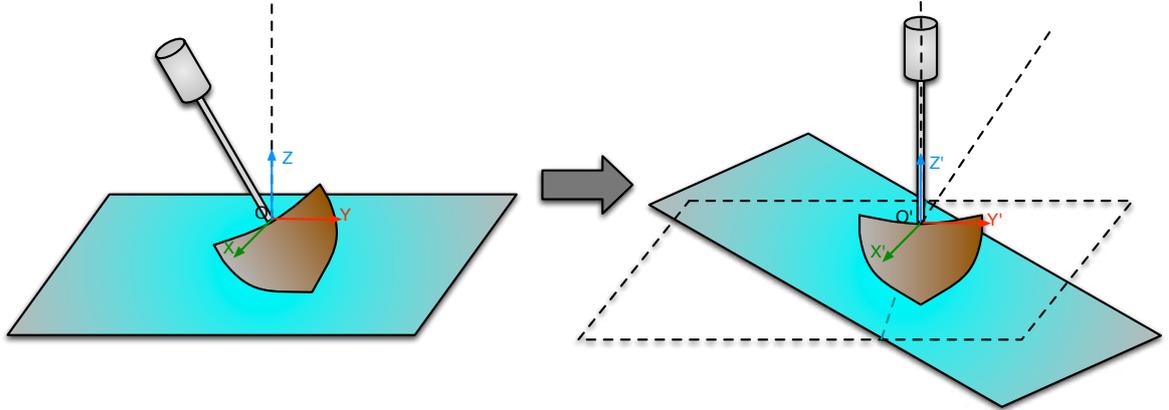


FIG. 5.6: Changement de repère de $OXYZ$ à $O'X'Y'Z'$.

L'orientation de la surface de l'eau $\mathbf{n}_T^{\text{Eau}}$ dans le système $O'X'Y'Z'$ peut être calculée à partir de l'orientation du bateau $\mathbf{n}_T^{\text{Bateau}}$ dans le système $OXYZ$ en utilisant la relation de changement de repère. De manière similaire à l'estimation de l'orientation initiale pour le bateau, lorsque la surface de l'eau est définie en utilisant les données de l'IMU, nous pouvons faire correspondre chaque pixel d'une image à un point 3D sur le plan d'eau en temps réel. Comme nous décrivons dans le chapitre 4, les intersections entre des rayons et le plan d'eau $\mathbf{n}_T^{\text{Eau}}$ sont donc les positions des objets détectés (Fig. 5.7).

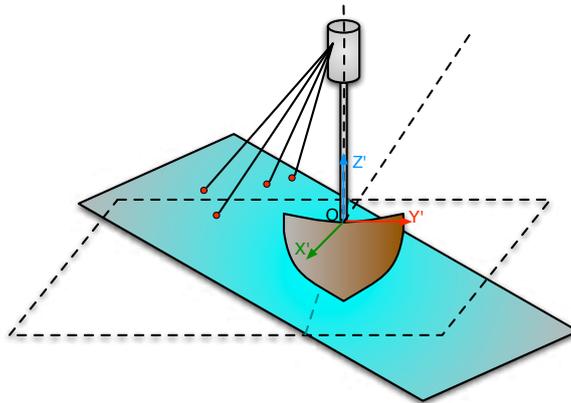


FIG. 5.7: Les intersections entre les rayons et le plan d'eau représentent les positions des objets calculées.

5.2.4 Résultats expérimentaux

Deux types de tests sont réalisés. Le premier est un test de simulation sur terre-ferme qui nous permet de réaliser toutes les étapes précisément et le deuxième est effectué en mer.

Test préliminaire sur la terre ferme

Nous mettons en pratique l'estimation de l'orientation du bateau par la centrale inertielle dans un endroit spacieux, simulant la surface de la mer. Le tube de la caméra est placé sur un trépied à 1.5 m de haut, et la centrale inertielle est fixée dessus.

Comme pour le test effectué dans le chapitre 4, un amer spécifique composé de 6 bouées (1 m de long \times 2 m de large) est utilisé par ce test. Cet amer est positionné à différentes distances, comprises entre 2 et 10 m. Pour chaque distance, nous avons testé plusieurs orientations de la caméra afin de simuler les différentes orientations du bateau. Deux types de orientation existent : l'orientation statique et l'orientation dynamique. L'orientation statique est l'orientation stable et sans mouvement. Le tube transparent de la caméra est perpendiculaire au sol (voir Fig. 5.8a et 5.8d). Pour l'orientation dynamique la caméra bouge vers des orientations arbitraires (voir Fig. 5.8b, 5.8c, 5.8e et 5.8f).

Nous avons calculé la position d'un obstacle avec deux méthodes différentes : l'une utilisant un amer de 6 bouées, et l'autre basé sur les données de la centrale inertielle. Les résultats sont présentés séparément dans les Tab. 5.1 et 5.2.

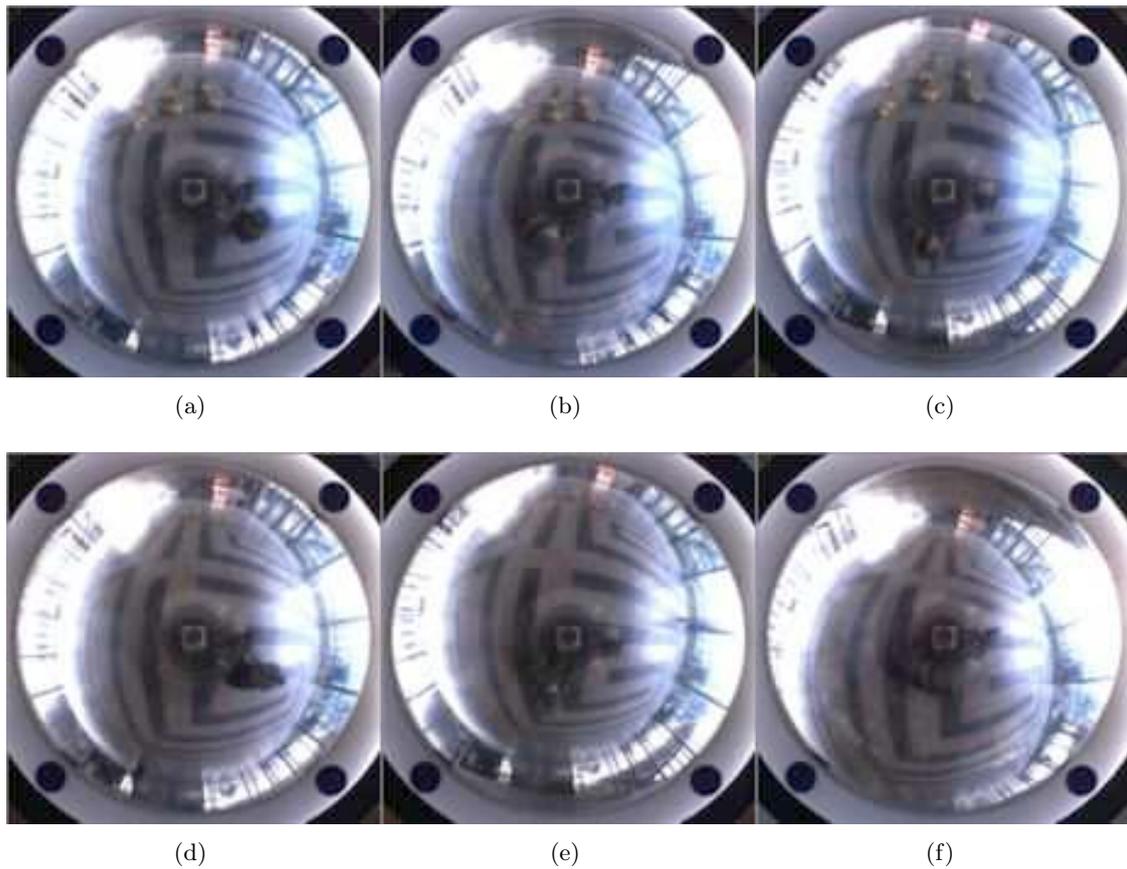


FIG. 5.8: Les photos sont prises par la caméra panoramique avec les différentes orientations pour les distances variées. (a)-(c) sont les photos prises avec les bouées à deux mètres, et (d)-(f) pour des bouées à six mètres. (a) et (d) sont les cas stables et sans mouvement, les autres sont les cas en mouvement.

TAB. 5.1: Estimation de la distance entre l'objet et la camera avec l'image panoramique en utilisant l'amer de 6 bouées

Distance mesurée(m)	Distance estimée, orientation statique(m)	Distance estimée, orientation dynamique(m)	Erreur d'estimation (%)
2	2.10	2.07	1.28
4	4.17	4.39	5.83
6	6.21	6.09	5
8	7.93	8.89	10.56
10	10.57	10.49	8.61

TAB. 5.2: Estimation de la distance entre l'objet et la camera avec l'image panoramique en utilisant les informations de la centrale inertielle

Distance mesurée(m)	Distance estimée, orientation statique(m)	Distance estimée, orientation dynamique(m)	Erreur d'estimation (%)
2	2.21	2.39	5.44
4	4.25	4.05	3.72
6	6.24	5.77	7.83
8	8.01	7.86	5.56
10	9.70	9.83	4.06

Le tableau 5.1 montre que toutes les erreurs d'estimation sont inférieures à 11%, avec une erreur maximale de 10.56 % à 8 m et une erreur minimale de 1.28 % à 2 m . Plus la distance entre l'objet et la caméra augmente, plus les erreurs d'estimation deviennent importantes.

Selon les résultats du tableau 5.2, la méthode utilisant la centrale inertielle semble avoir de bonnes performances avec une erreur d'estimation inférieure à 9 %. Toutes les distances estimées sont inférieures aux distances réelles. L'erreur maximale est de 7.83 % pour une distance de 6 m , et l'erreur minimale est de 3.72 % pour 4 m . La distance influence peu les résultats obtenus avec cette méthode.

Test en mer

Comme décrit dans la partie 4.6.3, un bateau à moteur a été utilisé pour le test en mer. Nous avons installé la caméra panoramique au milieu de ce bateau à 2.3 m de hauteur et fixé la centrale inertielle sous le tube transparent.



FIG. 5.9: Le navire porte-conteneur de CMA CGM.

La navigation d'un bateau de transport est enregistrée par la caméra panoramique. C'est un navire porte-conteneur de la compagnie française CMA CGM [CMA]. Sa longueur est de 200 m. Nous divisons la procédure de détection du bateau selon deux cas : loin ou proche. Notre objectif est d'estimer la taille du bateau en utilisant les images capturées et les signaux de la centrale inertielle pour les deux cas mentionnés.

La figure 5.10a (resp. 5.10b) représente le bateau de transport au loin (resp. proche). Pour chaque cas, nous avons traité dix images successives, et calculé la moyenne des résultats obtenus. La taille estimée est donnée dans le tableau 5.3.

TAB. 5.3: Estimation de la taille d'un objet avec l'image panoramique en utilisant les informations de la centrale inertielle

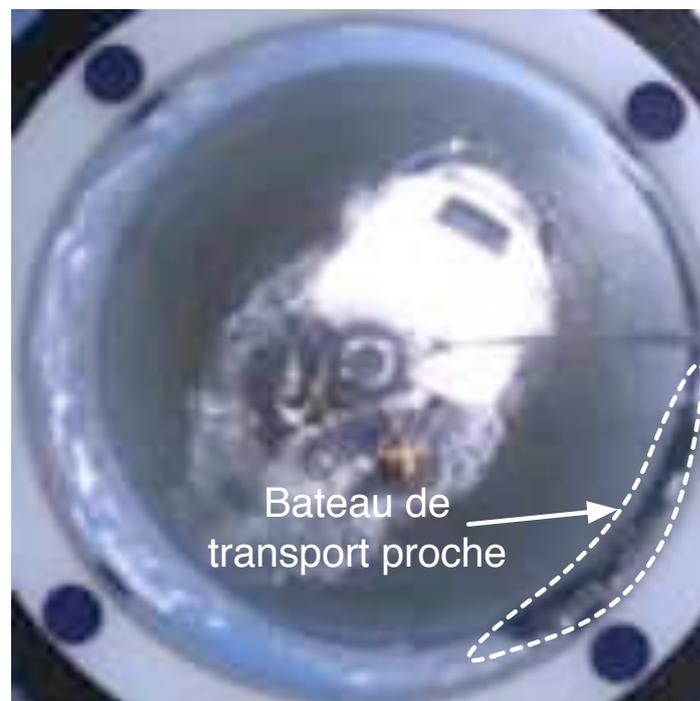
Distance	Taille estimée (m)	Erreur d'estimation (%)
Loin	230.62	15.31
Proche	172.44	13.78

L'erreur d'estimation est d'environ 15 %. Nous voyons que l'estimation de la taille d'un objet proche est meilleure que celle d'un objet éloigné. Ce phénomène peut être expliqué par la résolution du miroir. Plus les pixels sont éloignés du centre du miroir, plus ils représentent une distance importante. Pour un objet éloigné, une différence d'un pixel correspond à des résultats significativement différents.

En comparant les tableaux 5.1, 5.2 et 5.3, nous pouvons constater que les résultats du test



(a)



(b)

FIG. 5.10: Les photos sont prises par la caméra panoramique. (a) est une photo du bateau de transport au loin, et (b) est une photo du bateau proche.

en mer sont moins performants que ceux du test préliminaire. En effet, le test préliminaire a été réalisé avec des bouées aisément détectables, alors que le test en mer a été effectué sur des bateaux réels. La diversité de ses couleurs et de son volume rendent sa détection plus complexe. Les résultats obtenus pour ce test restent satisfaisants, et montrent la performance de notre algorithme.

5.3 Bloc de fusion

Dans le bloc de fusion nous voulons fusionner les données provenant du sonar et de la caméra pour détecter et localiser des obstacles en mer.

5.3.1 Données capturées par le sonar

Un sonar Tritech Micron DST est utilisé dans notre système de perception. Il est fixé sous la quille, à environ 1,5 m de profondeur. Le secteur analysé peut être réglé de 0 à 360 degrés, avec une fréquence de 36 °/s, une largeur de faisceau vertical de 30 ° et une largeur de faisceau horizontal de 3 °. Nous supposons que la vitesse du son dans la mer est d'environ 1500 m/s. Les données brutes capturées par le sonar sont obtenues sous la forme d'une matrice bilatérale, elles sont alignées verticalement selon le relèvement (de 0 à 360 °) (Fig. 5.12a). Chaque ligne représente l'orientation d'une détection (Fig. 5.11), chaque colonne représente une mesure des données dans la portée. Les données brutes sont difficiles à interpréter. Dans nos études, elles sont transformées en système de coordonnées cartésiennes (Fig.5.12b). Une analyse du secteur complet du sonar est décrite comme une période de balayage. Les données reçues de cette période de balayage sont considérées comme une trame de données brutes. La figure 5.12 montre une trame de données brutes et son affichage dans le système de coordonnées cartésiennes.

Le récepteur du sonar accepte un signal de retour dont l'énergie est comprise entre 0 et 80 dB (Fig.5.13). Un contrôle de la portée dynamique est utilisé pour ajuster la fenêtre d'affichage sur l'échelle des couleurs. Le signal reçu par le sonar contient N_E échantillons pour chaque détection. Ce sonar peut détecter des obstacles entre 2 et 100 m. Un paramètre P_E doit être réglé avant le balayage du sonar pour convertir les mesures en informations sur la distance de l'objet. Il est calculé selon le nombre d'échantillons pour chaque détection et il est défini selon une unité de 640 nanosecondes ($640 * 10^{-9}$ s). La figure 5.14 nous donne un exemple de calcul de P_E .

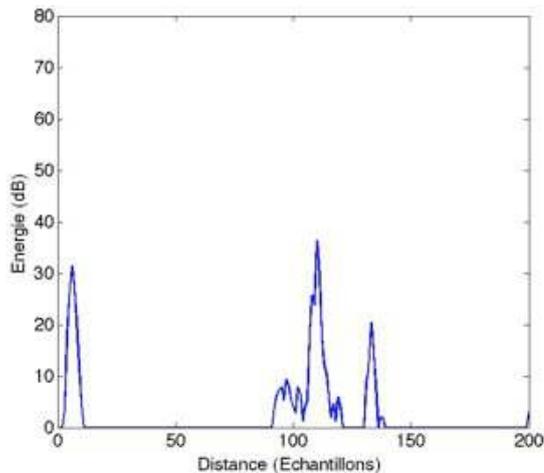


FIG. 5.11: Un exemple d'un écho reçu par le sonar pour une seule direction.

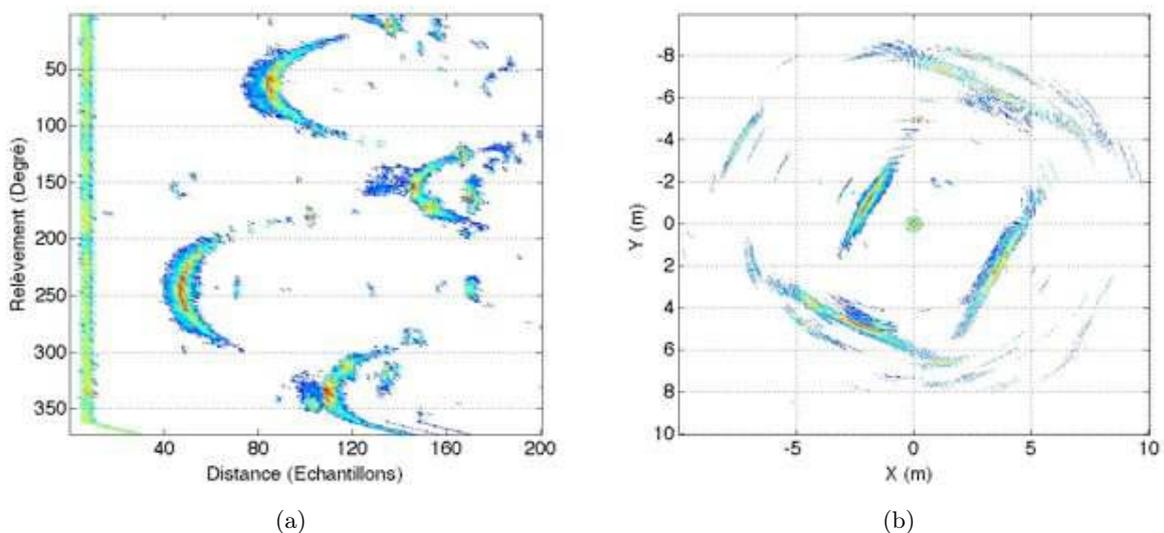


FIG. 5.12: Un exemple des signaux capturés par le sonar dans un bassin carré. (a) Les données bruitées reçues par le sonar avec une portée de 10 m (200 échantillons) et une résolution angulaire de 1.8° . (b) Les mêmes données présentées dans un système de coordonnées cartésiennes.

La distance obtenue d_i^{sonar} pour chaque angle $\alpha_i \in \{1 \dots 360^\circ\}$ est déterminée par la portée du sonar R^{sonar} , la quantité d'échantillons d'énergie N_E et le numéro d'échantillon :

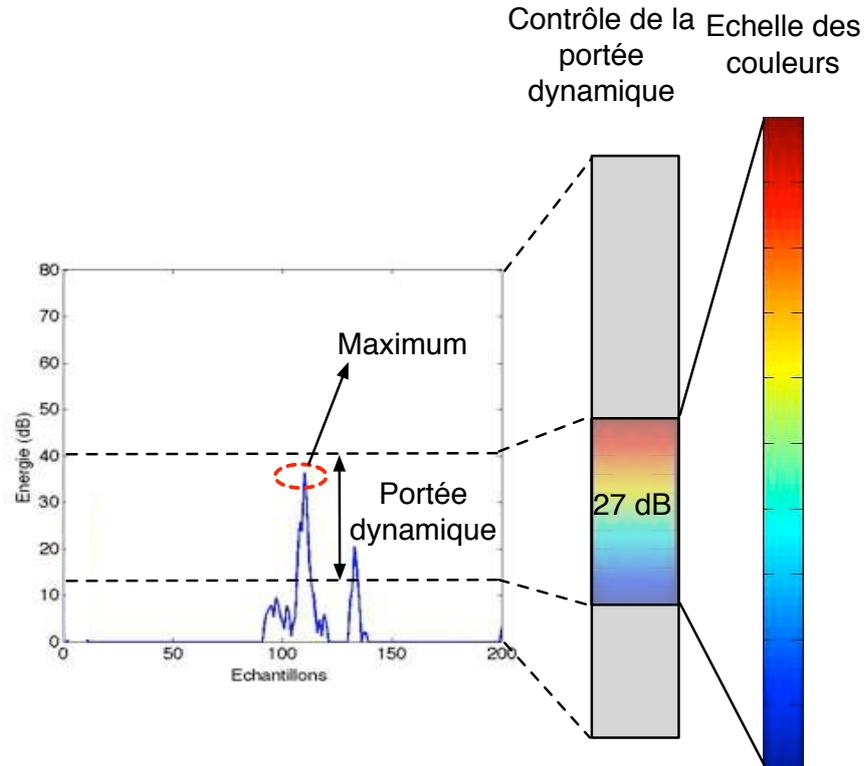


FIG. 5.13: Un exemple de signal reçu pour $N_E = 200$ et une portée de 20 m. L'échelle des couleurs définissant la portée est réglée sur l'intervalle 13 - 40 dB.

1. Calculer le temps de l'aller-retour t_{AR} pour une ligne de balayage du sonar avec une portée de 20 m et une vitesse de 1500 m/s.

$$\mathbf{t}_{AR} = 20 \times 2 \times \frac{1}{1500} = \mathbf{26.6 \text{ millisecondes}}$$

2. Calculer le temps d'échantillonnage t_E pour un certain nombre d'échantillons N_E , for exemple $N_E = 200$.

$$\mathbf{t}_E = \mathbf{t}_{AR} \times \frac{1}{200} = \mathbf{133.35 \text{ microsecondes}}$$

3. Convertir le temps de l'échantillon t_E en unité de 640 nanosecondes.

$$\mathbf{P}_E = \mathbf{t}_E \times \frac{1}{640}$$

FIG. 5.14: Un exemple de calcul du paramètre P_E .

$$d_i^{sonar} = \frac{R^{sonar}}{N_E} \times k_{\alpha_i} \quad (5.7)$$

où k_{α_i} est le $k^{\text{ième}}$ échantillon pour l'angle α_i avec $k \in \{1 \dots N_{bin}\}$. Les données capturées par

le sonar présentées dans un système de coordonnées cartésiennes peuvent être calculées selon la procédure suivante :

$$\begin{cases} x_i^{sonar} = d_i^{sonar} \cos \alpha_i \\ y_i^{sonar} = d_i^{sonar} \sin \alpha_i \end{cases} \quad (5.8)$$

Le secteur de balayage du sonar peut être réglé par deux paramètres : l'angle limite à gauche L_G et l'angle limite à droite L_D . La figure 5.15 montre un exemple sur un balayage de 90° du sonar, avec un angle limite à gauche de 135° , et un angle limite à droite de 225° .

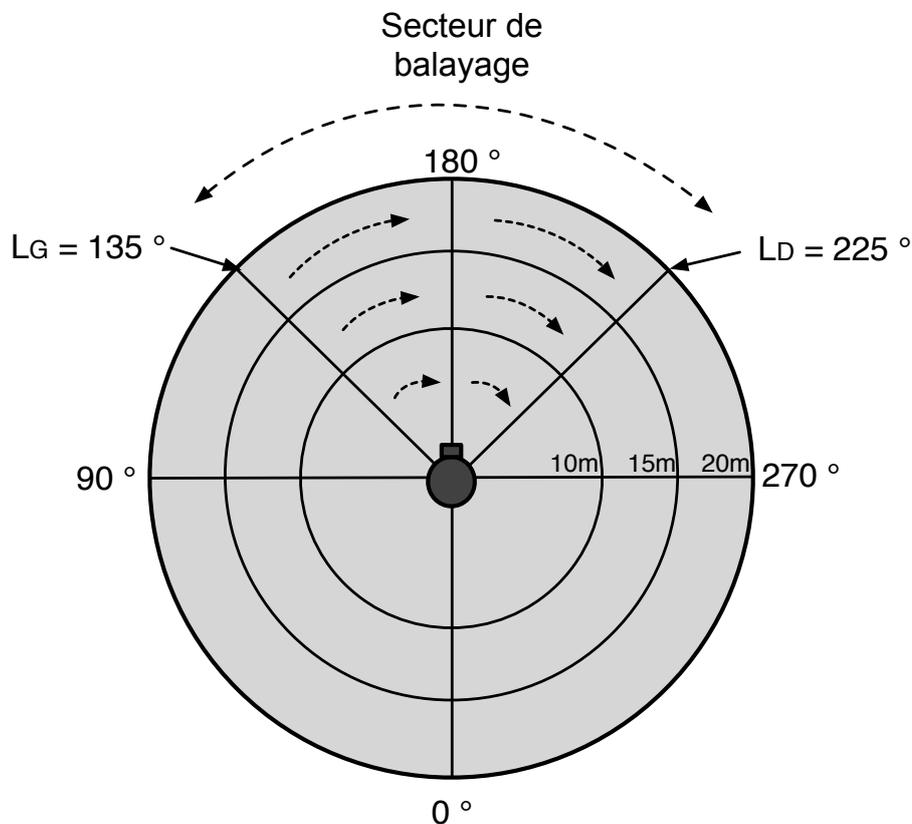


FIG. 5.15: Un exemple d'un balayage du sonar pour un secteur de 90° et une résolution angulaire de 1.8° .

La figure 5.16 montre un exemple de données enregistrées par le sonar au cours d'une expérience. Pour cet enregistrement, le sonar est immergé à 60 cm de profondeur, la portée de la détection est réglée à 20 m , la résolution angulaire est de 1.8° et son secteur angulaire est compris entre 90 et 270° . Un balayage d'un demi-plan dure 5 secondes. Un seul obstacle est détecté (un panneau réfléchissant de 80 cm) qui est affiché en bas à droite dans la figure.

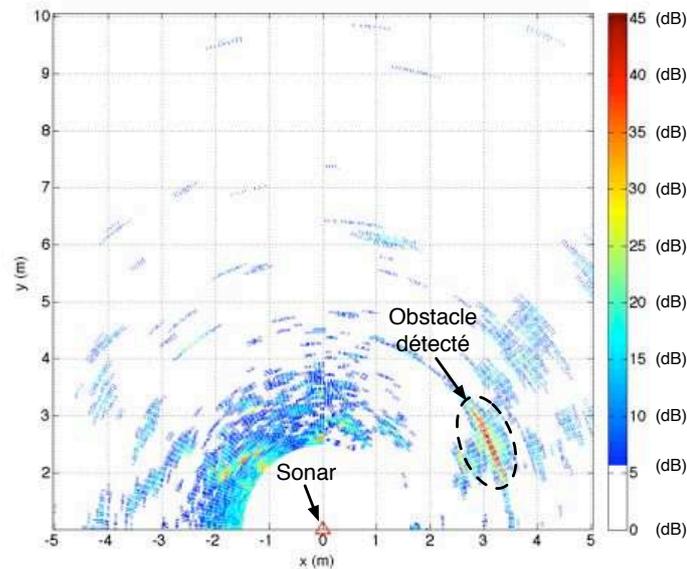


FIG. 5.16: Un sonar balayant de 90 à 270 °, et détectant un obstacle. La fenêtre de la portée dynamique pour afficher l'échelle des couleurs est réglée entre 6 et 46 dB.

5.3.2 Prétraitement des données du sonar

A cause des bruits électroniques et mécaniques à l'intérieur du sonar, et l'émission des ondes sonores, le sonar produit lui-même un bruit pendant la réception des signaux. En effet le récepteur du sonar capture le signal avant qu'il soit envoyé. Un prétraitement des données du sonar est nécessaire. Dans la détection pour une seule orientation, un segment de signaux avec une grande amplitude peut être trouvé dans les premiers échantillons (Fig. 5.17a). Lors de la reproduction des données dans un système de coordonnées cartésiennes, nous retrouvons ces bruits autour de la position cartésienne du sonar (Fig. 5.17b). Selon la spécification du sonar Tritech Micron DST, la dimension de la zone de bruit est connue et s'étend sur les 30 premiers échantillons de signal. Par conséquent, toutes les données de cette gamme sont considérées comme du bruit et seront éliminées. Leurs énergies seront remplacées par zéros, ce qui est considéré comme l'absence d'écho dans cette partie.

5.3.3 Détection d'obstacles par le sonar

Les données reçues par le sonar contiennent non seulement les obstacles, mais aussi du bruit ou encore des petits objets qui ne nous intéressent pas, comme montrés dans la figure 5.16. Ainsi, pour la détection d'obstacles, il faut filtrer les données reçues.

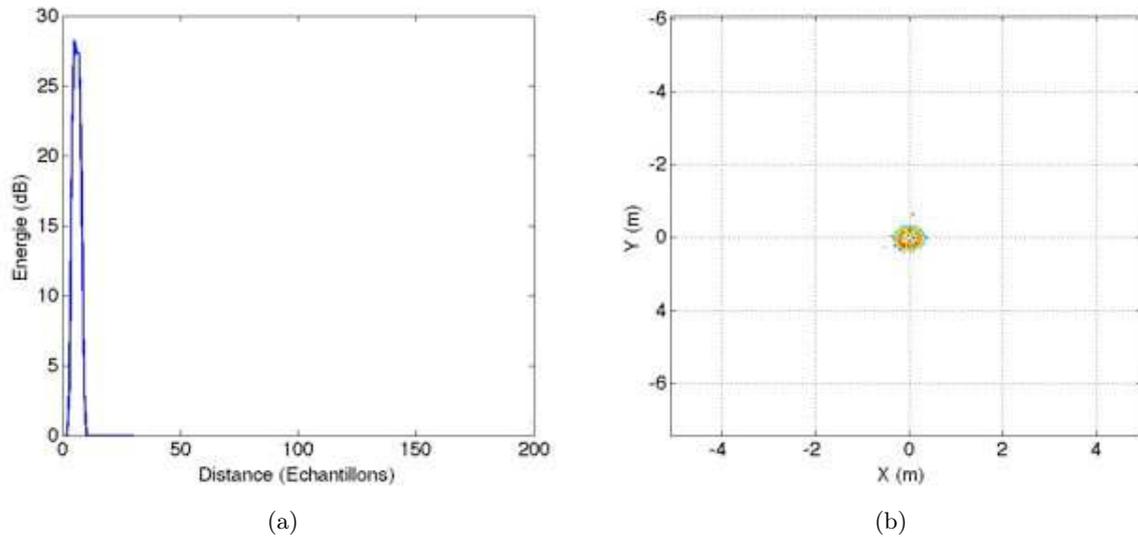


FIG. 5.17: (a) Le bruit propre du sonar pour une seule détection. (b) Le même bruit affiché dans le système de coordonnées cartésiennes.

Une amélioration de l'équation 5.7 est réalisée. Pour la détection de chaque direction, le numéro d'échantillon qui contient la plus grande énergie est sélectionné (Fig. 5.13) :

$$d_i^{obj} = \frac{R^{sonar}}{N_E} \times arg \max_{k=1 \dots N_{bin}} E_{\alpha_i}^k \quad (5.9)$$

où $E_{\alpha_i}^k$ est l'énergie du $k^{\text{ième}}$ échantillon pour l'angle α_i . Les points de l'obstacle $\mathbf{P}_i^{obj}(x_i^{obj}, y_i^{obj})$ peuvent être calculés selon la procédure suivante :

$$\begin{cases} x_i^{obj} = d_i^{obj} \cos \alpha_i \\ y_i^{obj} = d_i^{obj} \sin \alpha_i \end{cases} \quad (5.10)$$

La figure 5.18 donne un exemple de détection d'obstacles. Comme dans la figure 5.18b, les échos contenant moins d'énergie sont supprimés, mais les données restantes sont toujours bruitées. Un traitement plus sophistiqué doit être réalisé.

Dans le cas général, plusieurs objets sont détectés et les mesures sont fortement bruitées. Un traitement est nécessaire pour filtrer le bruit. Si pour des angles successifs, les mesures de distances sont proches, alors nous considérerons que les points détectés appartiennent à un même objet. Les autres points seront considérés comme du bruit. La figure 5.19 montre un exemple de détection d'obstacles par le sonar. Il y a deux objets détectés et deux mesures sont considérées comme du bruit. Le point à 357° est un bruit. Les points à 21° et 23° sont

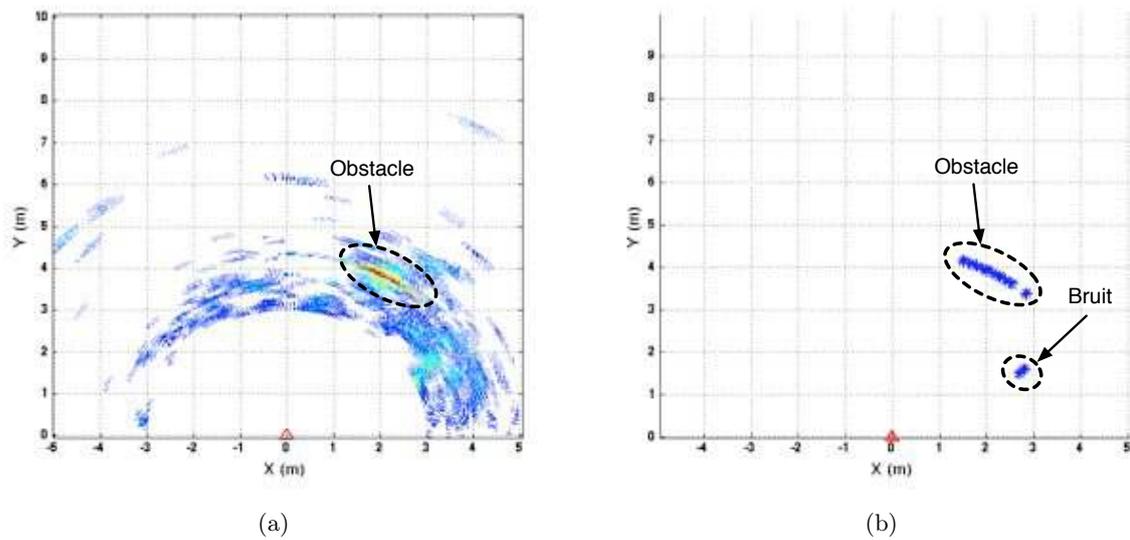


FIG. 5.18: (a) Les données bruitées capturées par le sonar. (b) La détection d'obstacles.

dans des angles successifs mais leurs distances sont très différentes (environ 5 m), ces mesures sont du bruit.

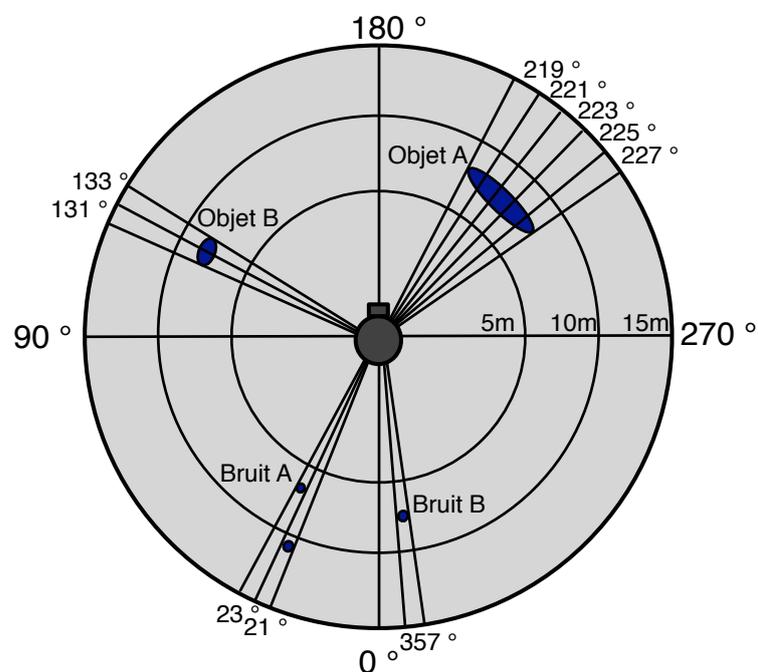


FIG. 5.19: Un exemple d'un balayage du sonar pour un secteur de 360 ° avec trois objets capturés.

Le centre de gravité de l'obstacle détecté par le sonar est $\mathbf{P}_c^{obj}(x_c^{obj}, y_c^{obj})$, il est calculé par l'angle au centre α_c et la moyenne des distances de tous les points de l'objet :

$$\begin{cases} d_m^{obj} = \frac{1}{N_{obj}} \sum_{i=1}^{N_{obj}} d_i^{obj} \\ x_c^{obj} = d_m^{obj} \cos \alpha_c \\ y_c^{obj} = d_m^{obj} \sin \alpha_c \end{cases} \quad (5.11)$$

où N_{obj} est le nombre des points d'un objet. La taille de l'objet l^{objet} peuvent être estimée par :

$$l^{obj} = (2\pi d_m^{obj}) \frac{1.8 N_{obj}}{360} \quad (5.12)$$

5.3.4 Calibration du sonar par rapport à la caméra

Le sonar doit être calibré par rapport au système de vision. La calibration est réalisée comme représenté figure 5.20. Un amer est placé dans un lac calme (la surface de l'eau est considérée comme un plan parfait). Il est constitué de deux parties : une partie émergée composée de deux bouées jaunes, une partie immergée comportant un réflecteur du sonar. Cet amer spécial est détectable à la fois par la caméra et par le sonar.

L'acquisition des données est réalisée en plaçant l'amer spécial à des distances différentes. A l'aide des données provenant du système de vision et du sonar sur les positions de la bouée, la matrice de transformation rigide entre ces deux systèmes de coordonnées peut être calculée.

5.4 Les méthodes de fusion

La caméra et le sonar ont des fréquences d'acquisition différentes : $f_{camera} = 12 \text{ image/s}$ et $f_{sonar} = 36 \text{ °/s}$. Nous réglons le secteur angulaire du sonar de 90 à 270 °. Une mesure complète du sonar dure donc 5 secondes. Dans nos études, une mesure de fusion est réalisée toutes les 5 secondes. Elle utilise les données reçues du sonar, et la moyenne des positions calculées à partir des 60 images. Nous proposons 5 méthodes de fusion différentes pour calculer la distance entre l'obstacle et le bateau à l'aide de n mesures successives.

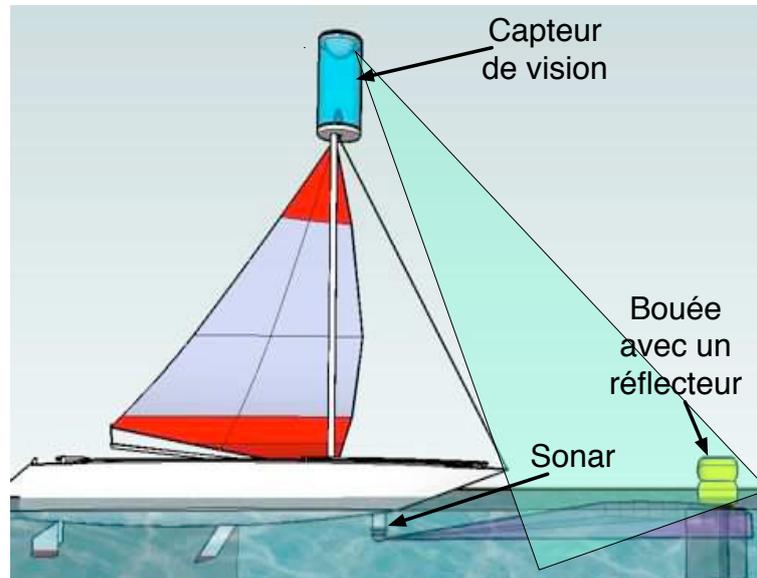


FIG. 5.20: Une bouée avec réflecteur détectable par le sonar et le capteur de vision est utilisée pour calibrer le sonar par rapport au système de vision.

5.4.1 Fusion par pondération 1 (F1)

Ce type de fusion utilise directement les n mesures en pondérant les données détectées [LGBG09]. Le schéma de cette méthode est montré dans la figure 5.21.

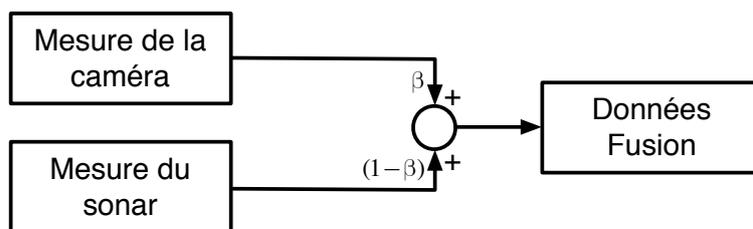


FIG. 5.21: Schéma de la fusion par pondération dans notre système de fusion.

La moyenne et le maximum sont estimés pour chaque modalité, et le coefficient de pondération est calculé comme :

$$\beta = \frac{\sigma_{camera}^2}{\sigma_{sonar}^2 + \sigma_{camera}^2} \quad (5.13)$$

où β est le facteur de pondération, σ_{camera} est l'écart type des distances obtenues par la caméra et σ_{sonar} est celui du sonar.

Le résultat de la fusion est obtenu par :

$$d_k^{fusion} = \beta d_k^{sonar} + (1 - \beta) d_k^{camera} \quad (5.14)$$

5.4.2 Fusion par pondération 2 (F2)

Cette méthode est similaire à la première méthode, mais le calcul du coefficient β est différent. Le maximum et le minimum de chaque mesure des deux modalités sont utilisés ici :

$$\alpha_{m \in \{camera, sonar\}} = \max_{k=1 \dots n} (d_k^m) - \min_{k=1 \dots n} (d_k^m)$$

$$\beta = \frac{\alpha^{camera}}{\alpha^{camera} + \alpha^{sonar}} \quad (5.15)$$

où α_m est la différence entre le maximum et le minimum de chaque mesure, d^m est la distance obtenue par le capteur m .

5.4.3 Fusion par estimation (F3)

Une simple estimation est réalisée dans cette méthode de fusion. La phase d'estimation des données est importante. Elle est réalisée selon l'équation 5.16 :

$$\begin{cases} d_1^{estimate} = \frac{1}{2} (d_1^{camera} + d_1^{sonar}) & \text{si } k = 1 \\ d_k^{estimate} = d_{k-1}^{fusion} & \text{si } k > 1 \end{cases} \quad (5.16)$$

où $k \in \{1 \dots n\}$ est le numéro de la mesure. La première donnée estimée $d_1^{estimate}$ est déterminée par la moyenne des données du sonar et des données de la caméra obtenues par la première mesure du test. Les données estimées pour les autres mesures $d_k^{estimate}$ avec $k > 1$ sont les données issues des fusions précédentes d_{k-1}^{fusion} comme décrit ci-après. La différence entre les mesures obtenues pour chaque modalité et la donnée estimée est calculée par :

$$D_k^{m \in \{camera, sonar\}} = |d_k^m - d_k^{estimate}| \quad (5.17)$$

où D_k^m est la différence entre les données obtenues par la caméra et le sonar et les données estimées pour la $k^{ième}$ mesure. La fusion est effectuée selon la relation entre D_k^m et ε_k :

$$\begin{cases} d_k^{fusion} = \frac{1}{2}(d_k^{camera} + d_k^{sonar}) & \text{si } D_k^m < \varepsilon_k \\ d_k^{fusion} = d_k^{camera} & \text{si } D_k^{camera} < D_k^{sonar} \\ d_k^{fusion} = d_k^{sonar} & \text{sinon} \end{cases} \quad (5.18)$$

où ε_k est un seuil adaptatif de la $k^{\text{ième}}$ mesure déterminé par la distance obtenue par le sonar :

$$\varepsilon_k = \frac{d_k^{sonar}}{\gamma} \quad (5.19)$$

γ est décidé expérimentalement au cours de nos essais en lac. Dans nos études, nous avons utilisé $\gamma = 10$.

Cette méthode de fusion est illustrée figure 5.22.

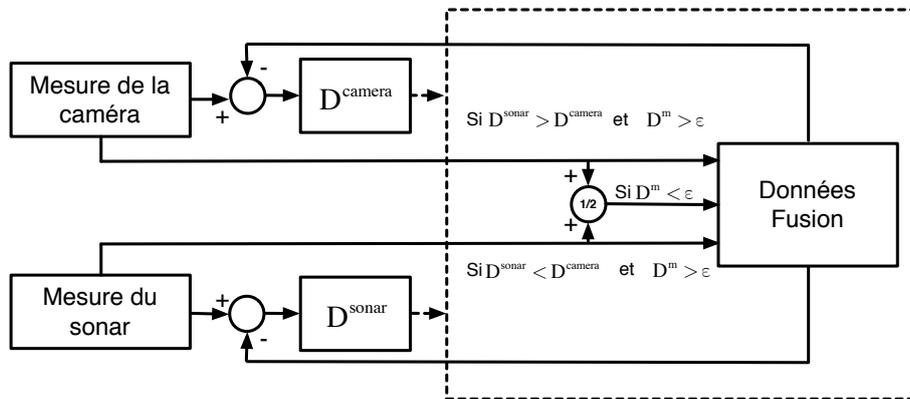


FIG. 5.22: Schéma de la fusion par estimation dans notre système de fusion.

5.4.4 Fusion par filtrage simple (F4)

Ce type de fusion utilise directement l'ensemble des n mesures. Les données de la caméra sont filtrées d'abord avant toute opération. La figure 5.23 donne un schéma de cette méthode de fusion.

Comme dans la figure 5.23, la moyenne de tous les types de données est d'abord calculée :

$$d_{moyenne} = \frac{1}{2n} \sum_{k=1}^n (d_k^{camera} + d_k^{sonar}) \quad (5.20)$$

La différence entre la $k^{\text{ième}}$ donnée de la caméra d_k^{camera} et $d_{moyenne}$ est calculée :

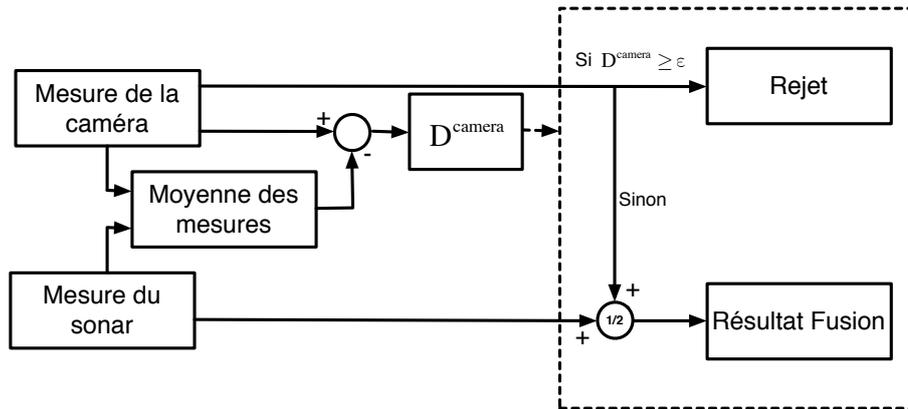


FIG. 5.23: Schéma de la fusion avec un filtre simple dans notre système de fusion.

$$D_k^{camera} = |d_k^{camera} - d_{moyenne}| \quad (5.21)$$

Selon la relation entre D_k^{camera} et le seuil ε_k , les données $d_k^{m \in \{camera, sonar\}}$ seront considérées comme des données rejetées ou des données de fusion :

$$\begin{cases} d_k^{m \in \{camera, sonar\}} \Rightarrow rejet & si \ D_k^{camera} \geq \varepsilon_k \\ d_k^{fusion} = \frac{1}{2}(d_k^{sonar} + d_k^{camera}) & sinon \end{cases} \quad (5.22)$$

5.4.5 Fusion par filtrage de Kalman (F5)

Le filtre de Kalman est fondé sur un schéma du type prédiction-correction, il estime les états d'un système dynamique à partir d'une série de mesures incomplètes ou bruitées. Il existe plusieurs méthodes de fusion en utilisant le filtre de Kalman [SH00], nous utilisons une méthode simple pour fusionner des données du sonar et de la caméra. Le schéma de notre méthode est montré figure 5.24.

Le filtre de Kalman permet d'estimer l'état d'un système. Ce système évolue au cours d'un temps discrétisé, il peut être représenté par une équation stochastique de l'espace :

$$X_{k+1} = AX_k + BU_k + W \quad (5.23)$$

avec A et B les paramètres du système. X_{k+1} est l'état du système à l'instant $k + 1$, et X_k est celui à l'instant k . U_k représente l'entrée de commande. W est le bruit d'état du système qui est un bruit blanc, sa matrice de covariance est Q .

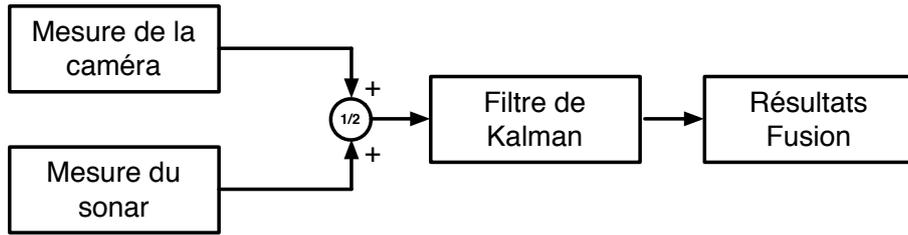


FIG. 5.24: Schéma de filtre de Kalman utilisé dans notre système de fusion.

Nous disposons d'un vecteur de mesure Z_k obtenues et reliées à l'état du système par la relation linéaire :

$$Z_k = HX_k + V \quad (5.24)$$

où V est le bruit de mesure, il est aussi un bruit blanc, sa matrice de covariance est R . H est une matrice déterministe qui relie l'état du système X_k et Z_k .

Afin d'obtenir l'état optimale du système nous devons combiner les observations Z_k avec l'information fournie par le modèle X_k . Dans notre système de fusion, selon l'équation (5.23) nous calculons d'abord l'estimation *a priori* $\hat{X}_{k|k-1}$ de l'état du système à l'instant k à partir des mesures disponibles à l'instant $k-1$:

$$\begin{aligned} \hat{X}_{k|k-1} &= \mathbb{E}(X_{k|k-1}) \\ &= A\hat{X}_{k-1|k-1} \end{aligned} \quad (5.25)$$

où $\hat{X} = \mathbb{E}(X)$ est l'espérance de X . $\hat{X}_{k-1|k-1}$ est l'estimation *a posteriori* de l'état du système à l'instant $k-1$. Nous supposons qu'il n'y a pas d'entrée de commande dans notre système, nous avons donc $\mathbb{E}(BU_{k-1|k-1}) = 0$. Pour un bruit blanc W_k , son espérance est toujours égale à 0.

Nous devons ensuite mettre à jour intermédiaire de la matrice de covariance de l'état du système en utilisant l'équation 5.23 :

$$\begin{aligned} P_{k|k-1} &= \mathbb{E}((X_{k|k-1} - \hat{X}_{k|k-1})(X_{k|k-1} - \hat{X}_{k|k-1})^T) \\ &= AP_{k-1|k-1}A^T + Q \end{aligned} \quad (5.26)$$

$P_{k|k-1}$ est la matrice de covariance qui correspond à $X_{k|k-1}$, et $P_{k-1|k-1}$ est la matrice de covariance de $X_{k-1|k-1}$.

Puis le gain du filtre optimal est calculé :

$$K_k = P_{k|k-1}H^T(HP_{k|k-1}H^T + R)^{-1} \quad (5.27)$$

L'estimation *a posteriori* de l'état du système à l'instant k est réactualisée :

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k(Z_k - H\hat{X}_{k|k-1}) \quad (5.28)$$

La matrice de covariance doit être mise à jour :

$$P_{k|k} = (I - K_kH)P_{k|k-1} \quad (5.29)$$

I est la matrice identité aux dimensions adéquates.

L'équations (5.25-5.28) sont les 5 équations basiques du filtre de Kalman, leurs développements sont décrits dans le rapport technique de Welch [WB95]. Dans notre système de fusion, les observations Z_k sont obtenues par les moyennes des données de la vision et du sonar :

$$Z_k = \frac{1}{2}(d_k^{camera} + d_k^{sonar}) \quad (5.30)$$

Les paramètres du filtre doivent être considérés à partir les observations des capteurs. La figure 5.25 donne une exemple de test, elle montre les observations obtenues par la vision et le sonar.

L'axe X représente les données réelles, et l'axe Y les données estimées. Pour chaque distance réelle, nous avons réalisé plusieurs observations, les incertitudes sont notées aussi dans la figure (la différence entre deux marqueurs verticaux). Les observations du sonar sont plus stables et plus proches des données réelles que celles de la vision. Nous donnons plus de confiance aux résultats du sonar que de la caméra. L'état du système est initialisé par la première donnée obtenue du sonar, sa matrice de covariance est peu importante pour le système, nous pouvons l'initialiser à 1 :

$$X_{0|0} = d_0^{sonar} \text{ et } P_{0|0} = 1 \quad (5.31)$$

Notre test est réalisé dans un lac calme, nous supposons que l'état du système $X_{k|k-1}$ est la même que $X_{k-1|k-1}$. Pour l'équation (5.25), dans notre système de fusion, nous avons :

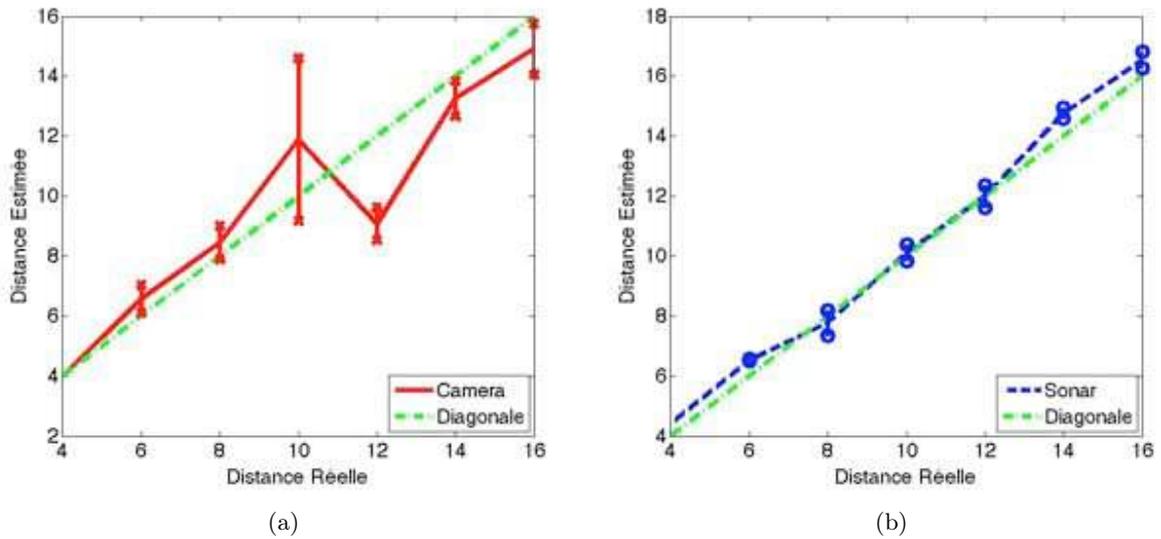


FIG. 5.25: (a) est les résultats obtenus par les données de la caméra. (b) est les résultats des données du sonar.

$$A = 1 \text{ et } H = 1 \quad (5.32)$$

La figure 5.26 montre que la plus grande valeur de la covariance apparaît dans les mesures de 10 m de la vision (Fig. 5.26a). Nous définissons R comme :

$$R = 11 \quad (5.33)$$

La matrice de covariance Q est difficile de définir. Ce paramètre explique le niveau de la confiance sur le système. Si nous donnons plus de confiance au système de fusion qu'aux capteurs, nous définirons $Q < R$. Inversement, si nous donnons moins de confiance au système, nous devons définir $Q > R$. Pour cette méthode de fusion, nous faisons confiance au modèle de notre système et nous définissons :

$$Q = 1 \quad (5.34)$$

Après la définition des paramètres du filtre de Kalman, nous mettons la moyenne des données de la vision et du sonar à l'instant k à l'entrée du filtre, nous pouvons obtenir au sortie les données de fusion.

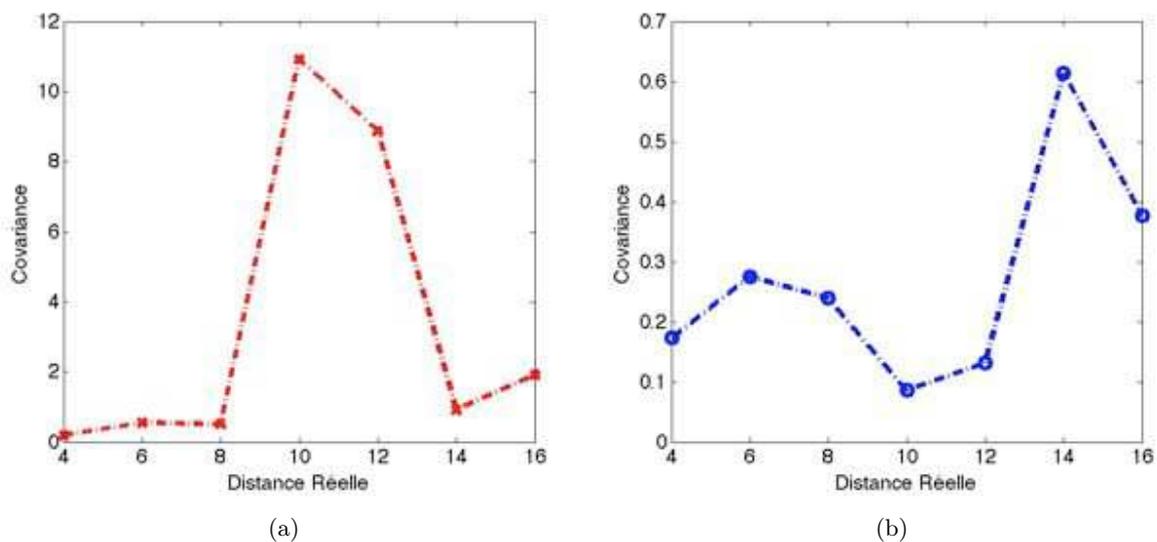


FIG. 5.26: (a) les covariances obtenus par les données de la caméra. (b) les covariances des données du sonar.

5.5 Résultats expérimentaux

Nous avons réalisé cette expérience dans un lac calme. Sa surface définit un plan. Un amer spécifique composé de 6 bouées (1 m de long \times 2 m de large) est utilisé pour ce test. Il comporte deux parties principales. La première est composée de bouées jaunes flottant sur la surface du lac, qui peuvent être détectées par le système de vision. La seconde partie est une plaque de métal immergée. Le sonar peut recevoir les signaux réfléchis par cette plaque. En raison de la structure spéciale des bouées, les deux types de capteurs permettent de détecter les parties immergée et émergée de l'amer. Le système de vision panoramique est fixé sur un trépied de 1.5 m de haut et la centrale inertielle est placée sous la caméra.

La position de l'amer est modifiée pour chaque enregistrement et la distance entre la caméra et le centre de l'amer est mesurée manuellement. Deux essais sont effectués. Pour chaque essai, différentes distances comprises entre 2 et 16 m sont testées. Pour chaque distance, 10 mesures successives sont enregistrées. Les résultats obtenus sont représentés dans une figure avec les données des distances réelles en abscisses et les données des distances estimées en ordonnées. Plus le résultat est proche de la diagonale dans la figure, meilleur il est.

La figure 5.27 représente les résultats obtenus par les capteurs, sans fusion. Les figures 5.27a et 5.27b sont les données obtenues par la caméra pour deux essais particuliers, les figures 5.27c et 5.27d donnent les mesures du sonar correspondantes. D'après les courbes,

nous déduisons que la détection du sonar est plus stable que celle de la caméra, car cette dernière est sensible aux conditions naturelles, comme la réflexion du soleil sur la surface de l'eau. Dans la figure 5.27a, les mesures pour une distance de 10 m donnent de mauvais résultats, qui varient de 9 à 15 m. Le même phénomène se retrouve pour les mesures de 12 m (Fig. 5.27a et Fig. 5.27b). De mauvais résultats pour le sonar apparaissent par des mesures de 8 m (Fig. 5.27d).

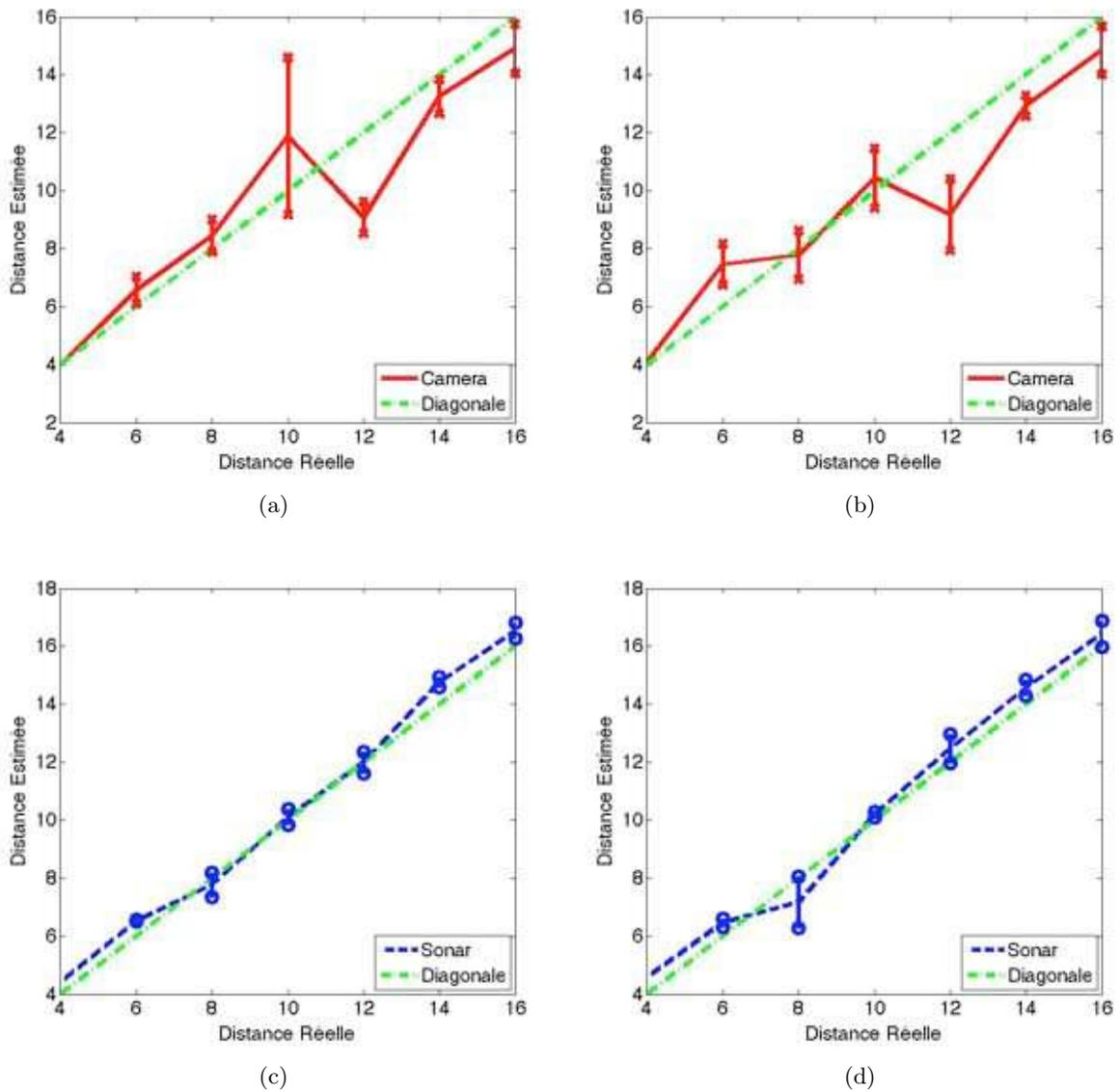


FIG. 5.27: (a) et (b) sont les résultats obtenus par les données de la caméra pour deux essais différents. (c) et (d) sont les résultats des données du sonar pour ces essais.

Les figures 5.28 à Fig. 5.32 donnent les résultats de la fusion réalisée selon les 5 méthodes

décrites précédemment. Les données sont plus exactes que celles obtenues directement par la caméra ou le sonar. D'après les courbes, nous notons que la troisième (F3) et la dernière méthodes de fusion (F5) sont plus performantes que les autres : la fusion avec une prédiction de donnée produit un résultat plus précis.

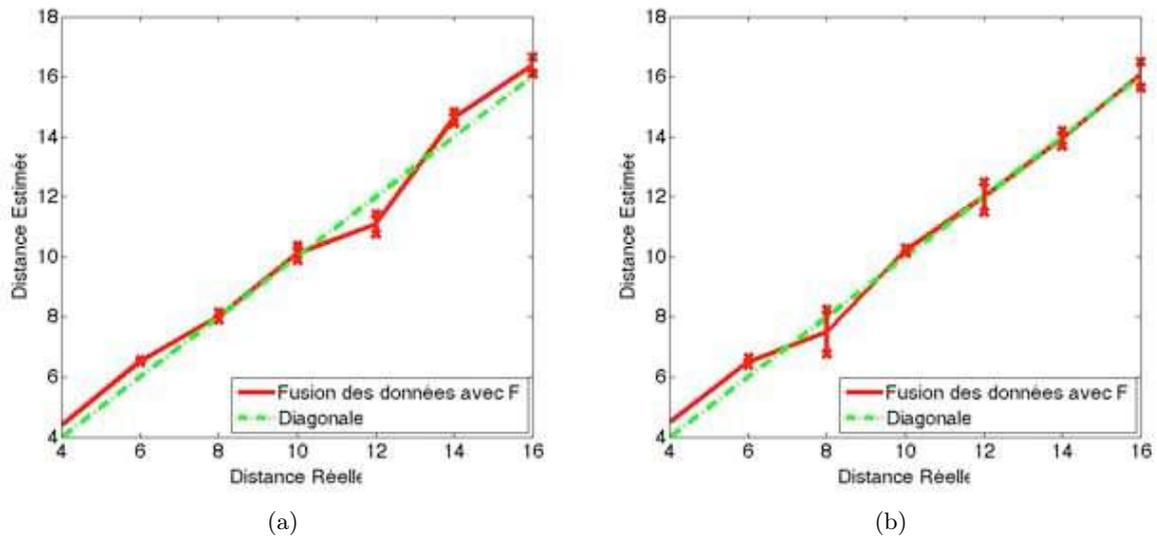


FIG. 5.28: Les résultats obtenus par la méthode de fusion par l'estimation (F1) sur les deux essais correspondants.

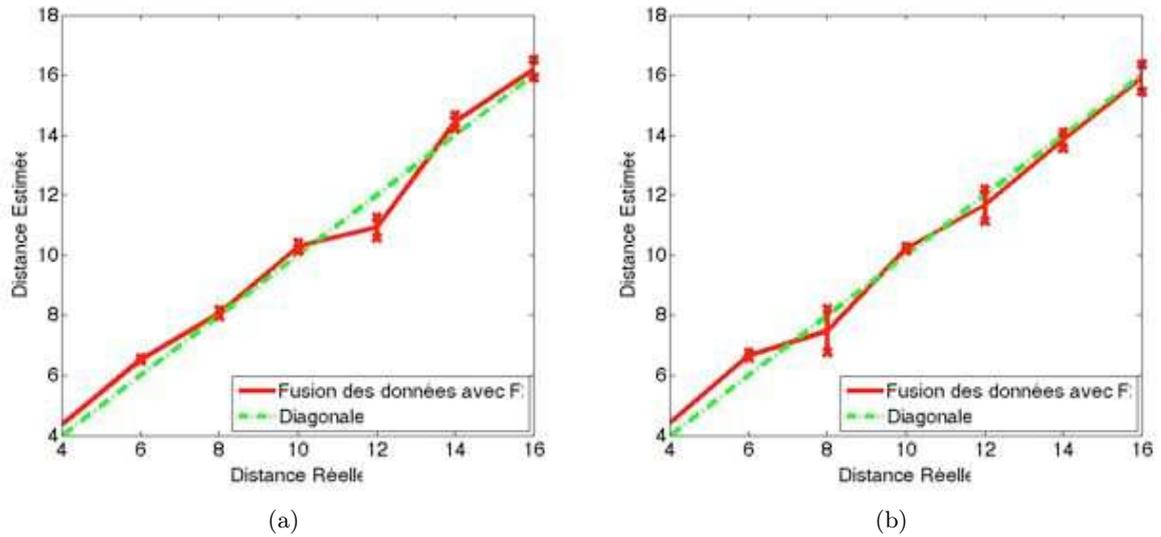


FIG. 5.29: Les résultats obtenus par la méthode de fusion par la pondération 1 (F2) sur les deux essais correspondants.

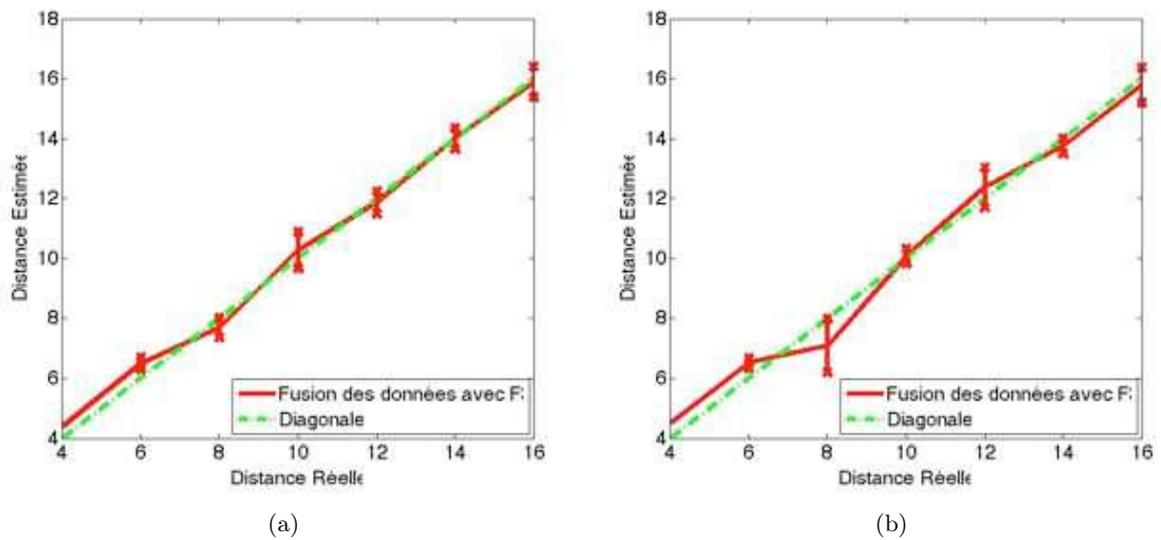


FIG. 5.30: Les résultats obtenus par la méthode de fusion par la pondération 2 (F3) sur les deux essais correspondants.

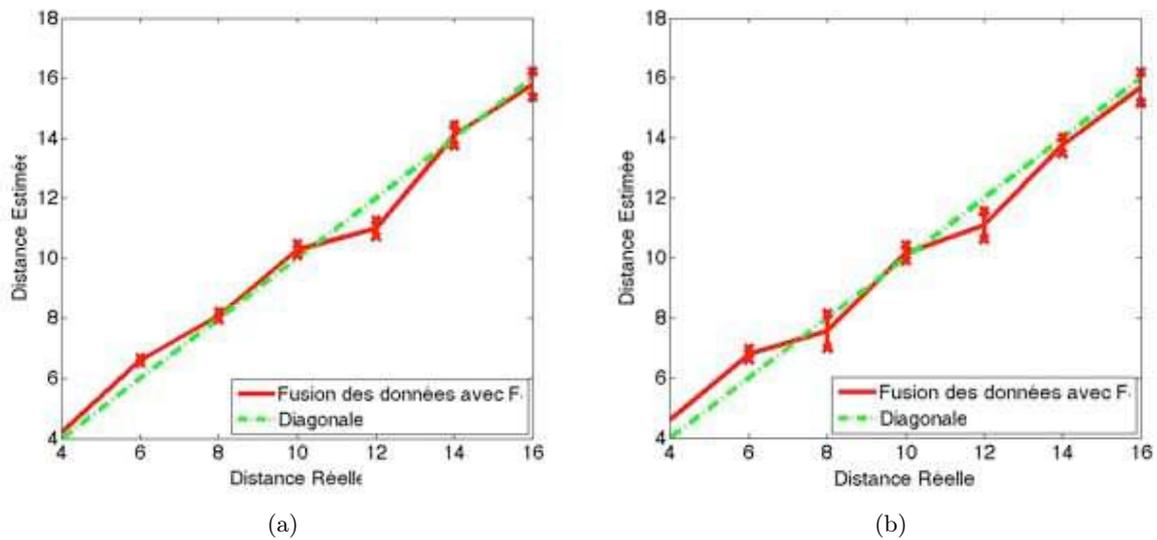


FIG. 5.31: Les résultats obtenus par la méthode de fusion avec un filtre simple (F4) sur les deux essais correspondants.

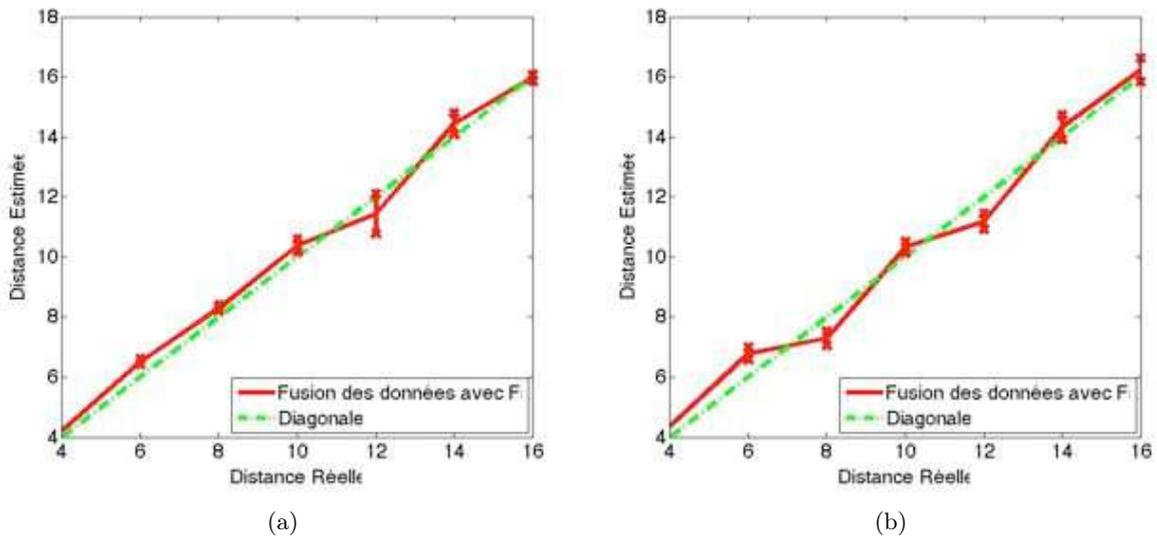


FIG. 5.32: Les résultats obtenus par la méthode de fusion avec un filtre Kalman (F5) sur les deux essais correspondants.

5.6 Conclusion

Les techniques de fusion des données combinent les informations issues de différents capteurs, pour atteindre des précisions accrues et des estimations meilleures que celles obtenues avec un capteur seul. Dans notre système de perception, nous avons réalisé un système multi-capteurs utilisant un système de vision panoramique, une centrale inertielle et un sonar. Nous avons expliqué dans ce chapitre l'architecture du système, et les différentes méthodes de fusion de données. Nos algorithmes de fusion sont performants, et les résultats issus de la fusion sont meilleurs que ceux donnés par les capteurs individuels.

Conclusions et Perspectives

Cette thèse concerne les études des systèmes de perception visuelle, sonore, inertielle et les signaux issus de l'ensemble des capteurs présents, en vue de traiter du problème de la détection d'obstacle, et de celui de l'estimation de l'état de la mer. Pour pouvoir réaliser ces tâches, plusieurs points ont été abordés au cours de cette thèse :

- La construction d'un système de perception multi-capteurs en milieu marin.
- Le traitement de signaux sonores sous-marins pour la reconnaissance d'objets et le traitement d'images panoramiques pour la détection et la localisation d'obstacles.
- La fusion de données avec différentes modalités pour l'amélioration de la perception des obstacles.

Plusieurs difficultés sont apparues lors de la construction d'un système de perception, comme : le choix des capteurs, l'étanchéité des dispositifs et la synchronisation des séquences d'acquisition des données. Le choix des capteurs utilisés dans le système est très important. Une étude technique sur le bateau a été d'abord réalisée pour définir la liste des capteurs nécessaires. Quatre types de capteurs ont été sélectionnés : deux hydrophones, une caméra panoramique, un sonar et une centrale inertielle. Ils permettent de traiter les différents aspects de la perception. Pour chaque type de capteur, des critères de poids, de volume et de consommation d'énergie ont été pris en compte.

Notre système de perception étant installé dans un véhicule autonome de surface, l'étanchéité

des dispositifs est indispensable. La caméra panoramique a été placée dans un tube acrylique transparent, qui permet de la protéger. Les hydrophones et le sonar sont des capteurs immergés, que nous pouvons utiliser directement sous l'eau. La centrale inertielle et les autres dispositifs (batteries, PC104...) ont été placés dans une boîte étanche. L'étanchéité de l'ensemble du système a été vérifiée expérimentalement.

Pour un système multi-capteurs, la synchronisation de l'acquisition des données est une tâche fondamentale. Nous avons réalisé l'acquisition des données des différents capteurs sous LINUX en utilisant sa fonctionnalité multi-tâches. Chaque séquence de données est ensuite datée par rapport à l'horloge commune d'un ordinateur afin d'assurer la synchronisation.

Le traitement de signaux sonores et d'images présente plusieurs difficultés : la construction d'une base de données, le calibrage du système de vision, la détection d'obstacles. La base de données utilisée est importante pour le développement d'algorithmes. Nous avons utilisé une base de signaux réels dans les enregistrements *The Underwater Sound Effects Series*. Cette base comporte 500 enregistrements d'événements sonores échantillonnés à 44 *kHz* et réalisés à partir d'hydrophones. Notre base de données comporte deux types de signaux : des signaux transitoires et des signaux long-termes. Deux classifieurs et quatre méthodes de codage de signaux ont été réalisés sur les deux types de signaux pour la reconnaissance d'objets sous-marins.

L'utilisation de caméras implique toujours un calibrage du dispositif, le calibrage de notre système de vision est plus complexe que celui d'une unique caméra. Le système de caméra panoramique utilisé dans nos études comportait une caméra Gigabit, un miroir et un tube transparent. Nous avons décomposé la procédure de calibrage en deux parties. La première partie étant le calcul de la pose relative de la caméra par rapport au miroir. La deuxième partie étant l'étude de la réflexion de rayons par la surface du miroir pour caractériser la projection de la scène vers le miroir.

La détection d'obstacles est une tâche essentielle pour des véhicules autonomes de surface. Dans nos études, nous avons réalisé la détection d'obstacles en deux étapes. La première étape a consisté à étudier la différence entre deux images consécutives pour avoir une position grossière des objets en mouvement. La deuxième étape était la segmentation d'objets en utilisant un critère colorimétrique. Nous avons considéré la surface de la mer comme un fond bleu uniforme. L'obstacle a pu être détecté par sa différence de couleur. Avec les informations inférées par le calibrage de la caméra, nous avons pu localiser la position des obstacles.

La fusion de données implique plusieurs étapes : le traitement de données des capteurs, la synchronisation du système multi-capteurs et le développement d'algorithmes de fusion. Nous avons utilisé une centrale inertielle et un sonar dans notre système de fusion. A partir des données brutes obtenues, nous avons réalisé un pré-traitement des données pour obtenir

des données exploitables.

L'utilisation de la centrale inertielle simplifie les calculs pour la détection de l'orientation du bateau. Sa fréquence d'acquisition est plus élevée que celle de la caméra panoramique. Pour intégrer ces deux types de signaux, nous avons échantillonné les données de la centrale inertielle pour qu'elle ait la même fréquence que ceux de la caméra. L'installation d'un sonar dans notre système de fusion introduit également un problème de synchronisation. La fréquence de balayage du sonar est beaucoup moins élevée que celle de la caméra. Pour chaque balayage du sonar, nous avons donc calculé une moyenne à partir des résultats obtenus par la caméra et la centrale inertielle sur la même période de détection.

Nous avons proposé cinq méthodes de fusion de complexité variable. La méthode la plus simple est la fusion par pondération. C'est la plus facile à mettre en oeuvre, et elle est largement utilisée dans le domaine de fusion de données. La fusion par estimation prend en compte les résultats expérimentaux pour valider les données prédites. Deux méthodes de fusion par filtrage ont également été développées. Les données des capteurs ont été filtrées avant la fusion, pour supprimer le bruit. Ces informations ont ensuite été utilisées pour l'étape de fusion.

Ces premiers travaux démontrent la faisabilité d'un système de perception avec de bonnes performances pour un véhicule autonome de surface en milieu marin. Ce travail n'est que la première version d'un système de perception pour un robot autonome en milieu marin, mais il ouvre de nombreuses perspectives dans le domaine de la perception des véhicules autonomes de surface pour la robotique.

Dans la partie de traitements des signaux sonores, nous avons montré que, si les méthodes de représentation des signaux sont bien choisies, il est possible d'obtenir des résultats prometteurs en classification. Afin de mieux reconnaître tous les états de la mer, le nombre de types de signaux doit être augmenté dans notre base de données. La précision de notre méthode pourrait ainsi être améliorée.

L'amélioration de la précision du sonar doit être poursuivie en ajoutant un capteur de température. Le sonar utilise des impulsions ultrasonores pour la détection d'obstacles éloignés. Selon le temps écoulé avant la réception des ondes ultrasonores réfléchies par l'objet, la distance entre le sonar et l'objet peut être calculée. Dans des conditions réelles, la vitesse des ultrasons est modifiée en fonction de la température et de la densité de l'eau. Afin d'améliorer la précision des informations issues du sonar, un capteur supplémentaire mesurant la température et la densité de l'eau doit être ajouté dans notre système de perception.

En se basant sur un système de perception efficace et précis qui comporte différents types de capteurs, le véhicule autonome de surface en milieu marin pourra être utilisé dans des

applications variées.

ANNEXE

A

Liste de la base de données sonores

C'est les listes des bases de données que nous avons utilisé dans le chapitre 3, pour réaliser le traitement des signaux sonores.

TAB. A.1: La base de données transitoires

Num	Type de signal	Num	Type de signal
1	Un frappement à basse fréquence.	8	L'éclaboussure de l'eau.
2	L'impact de la plongée humaine.	9	Le jeté d'une grosse pierre dans l'eau.
3	La pression sous l'eau.	10	L'explosion sous-marine.
4	La chute dans l'eau.	11	Des petites bulles sous l'eau.
5	Un seul coup de Harpon.	12	Un frappement du métal sous l'eau.
6	L'impact de l'eau.	13	L'alarme du sonar.
7	Le transvasement de l'eau.		

TAB. A.2: La base de données non transitoires

Num	Type de signal	Num	Type de signal
1	Un réservoir d'oxygène sous-marin avec une vitesse moyenne.	16	Une torpille passe à côté.
2	Un goutte à goutte.	17	Le démarrage du moteur.
3	Un réservoir d'oxygène sous-marin avec une vitesse moyenne.	18	La pluie forte.
4	Le mouvement des bulles.	19	Une plongée sous-marine.
5	Les bulles muettes à distance.	20	Un navire à côté.
6	Une avalanche.	21	Un bateau.
7	Un moteur du bateau.	22	Un sous-marin passe devant.
8	Un Pédal'eau.	23	Un missile passe à côté.
9	Les bulles intenses.	24	Un volcan.
10	L'eau de rivière.	25	Un faible rumble drone.
11	Un canoë.	26	Le son du fond dans l'eau.
12	Un tremblement de terre.	27	Les bulles en eau profonde.
13	L'explosion large d'une bombe.	28	La plongée sous-marine qui passe à côté.
14	Les bulles fortes.	29	Les grandes bulles.
15	Les fissures de glace.	30	L'impact de vagues.

BIBLIOGRAPHIE

- [AB83] U. Appel and A. V. BrandtFrance. Adaptive sequential segmentation of piecewise stationary time series. *Inform. Sci*, 29 :27–56, 1983. (Cité en page 30.)
- [ALS09] ALSA. *Advanced Linux Sound Architecture project homepage*, 2009. (Cité en page 22.)
- [AOM09] N. Attoh-Okine and S.A. Mensah. Sensor fusion and civil infrastructure systems monitoring : A valuation algebras analysis of output data. *Sensors Jouranal, IEEE*, 9, 2009. (Cité en page 82.)
- [AOO⁺06] J. Alves, P. Oliveira, R. Oliveira, A. Pascoal, M. Rufino, L. Sebastiao, and C. Silvestre. Vehicle and mission control of the delfim autonomous surface craft. In *Proc. of 2006 mediterranean conference on control and automation*, volume 1 and 2, pages 924–929, 2006. (Cité en page 7.)
- [APA99] T. Alves, A. Pascoal, and Al. The use of caravela 2000 vehicle in operational oceanography. In *Proc. 2nd EUROGOOS Conference*, 1999. (Cité en page 8.)
- [Bas04] Underwater sound effet series, 2004. (Cité en page 38.)
- [Ber08] V. Bertram. Unmaned surface vehicles - a survey. In *Skibsteknisk Selskab*, 2008. (Cité en page 6.)
- [BK01] R. Benosman and S.B. Kang. *Panoramic Vision : Sensor Theory and Application*. Springer Verlag, 2001. (Cité en page 48.)

- [BMD96] R. Benosman, T. Manière, and J. Devars. Multidirectionnal stereovision sensor, calibration and scenes reconstruction. In *proceeding of the international conference on pattern recognition*, 1996. (Cité en page 49.)
- [BN98] S. Baker and S. K. Nayar. A theory of catadioptric image formation. In *Proceedings of ICCV*, pages 35–42, 1998. (Cité en page 50.)
- [Bou97] J. Y. Bouguet. *Camera Calibration Toolbox for Matlab*. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html, 1997. (Cité en page 51.)
- [BP77] G. Bodenstern and H. M. Praetorius. Feature extraction from the electroencephalogram by adaptive segmentation. In *Proc. IEEE 65*, pages 642–652, 1977. (Cité en page 30.)
- [BR84] B. Bouachache and F. Rodriguez. Recognition of time-varying signals in the time-frequency domain by means of the wigner distribution. In *Proc. ICASSP 84*, 1984. (Cité en page 30.)
- [BRW07] E. Bosse, J. Roy, and S. Wark. *Concepts, models and tools for information fusion*. Artech House, 2007. (Cité en page 82.)
- [BS90] B. Boashash and P.O’ Shea. A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques. 38(11) :1829 – 1841, 1990. (Cité en pages 30 et 31.)
- [CBB⁺05] M. Caccia, R. Bono, G. Bruzzone, E. Spirandelli, G. Veruggio, A.M. Stortini, and G. Capodaglio. Sampling sea surfaces with sesamo : an autonomous craft for the study of sea-air interactions. *IEEE Robotics and Automation Magazine*, 12(3) :95–105, 2005. (Cité en page 8.)
- [Cha09] R. Chang. *Etude des réseaux de caméras non synchronisées ou non calibrées*. PhD thesis, UPMC- Paris6, 2009. (Cité en pages ix et 66.)
- [Che85] C. H. Chen. Recognition of underwater transient patterns. *Pattern recognition*, 18(6) :485–490, 1985. (Cité en page 30.)
- [CLP05] J. Curcio, J. Leonard, and A. Patrikalakis. Scout - a low cost autonomous surface platform for research in cooperative autonomy. In *Proc. Oceans’05*, volume 1, pages 725–729, 2005. (Cité en page 8.)
- [CM77] C. S. Clay and H. Medwi. *Acoustical Oceanography*. Wiley Interscience, 1977. (Cité en page 28.)
- [CM99] T.L. Conroy and J.B. Moore. Resolution invariant surfaces for panoramic vision systems. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 392–397, 1999. (Cité en page 50.)

-
- [CMA] (Cité en page 93.)
- [CS97] J. S. Chahl and M. V. Srinivasan. Reflective surfaces for panoramic imaging. *Applied Optics*, 36(31) :8275–8285, 1997. (Cité en page 50.)
- [CSA⁺06] L. Cooney, M.J. Stanway, P. Augenbergs, H. Brudage, B. Downey, T. Pennington, T. Stefanov-Wagner, and D. Tobias. Design of an acoustic-homing autonomous surface vessel. *OCEANS 2006*, 2006. (Cité en page 13.)
- [CSHL08] R.A. Chastain, M.A. Struckhoff, H.S. HE, and D.R. Larsen. Mapping vegetation communities using statistical data fusion in the ozark national scenic riverways. *Photogrammetric engineering and remote sensing*, 74(2) :247–264, 2008. (Cité en page 82.)
- [DGU09] Matthew Dunbabin, Alistair Grinham, and James Udy. An autonomous surface vehicle for water quality monitoring. In *Robotics and Automation (ACRA)*, 2009. (Cité en page 13.)
- [Die06] James Diebel. Representing attitude : Euler angles, unit quaternions, and rotation vectors. 2006. (Cité en page 87.)
- [DPE⁺] John M. Dolan, Gregg W. Podnar, Alberto Elfes, Stephen Stancliff, Ellie Lin, John Higinbotham, Jeffrey C. Hosler, John Moisan, and Tiffany A. Moisan. Ocean-atmosphere sensor integration system (oasis). (Cité en page 11.)
- [EB08] G.H. Elkaim and C.O. Boyce. An energy scavenging autonomous surface vehicle for littoral surveillance. In *ION Global Navigation Satellite Systems Conference*, 2008. (Cité en page 11.)
- [Ecl] Ecliptique. (Cité en page 50.)
- [EHPS98] K. Englehart, B. Hudgins, P. Parker, and M. Stevenson. Time-frequency representation for classification of the transient myoelectric signal. In *Engineering in Medicine and Biology Society*, volume 5, pages 2627–2630, 1998. (Cité en page 31.)
- [Elk01] G.H. Elkaim. *System identification for precision control of a wingsailed GPS-Guided catamaran*. Stanford University, 2001. (Cité en page 11.)
- [EP01] P. Encarnacao and A. Pascoal. Combined trajectory tracking and path following : an application to the coordinated control of autonomous marine craft. In *Proc. 40th IEEE Conference on Decision and Control*, pages 964–969, 2001. (Cité en page 7.)
- [Eve97] R. Everson. Orthogonal, but not orthonormal, procrustes problems. In *Advances in Computational Mathematics*, 1997. (Cité en page 56.)

- [EWK06] Gabriel H. Elkaim, Bruce R. Woodley, and Robert J. Kelbley. Model free subspace based h_∞ control (p). In *IEEE/ION Position Location And Navigation Symposim*, pages 1005–1016, 2006. (Cité en page 11.)
- [EYS⁺04] B. Enderle, T. Yanagihara, M. Suemori, H. Imai, and A. Sato. Recent developments in a total unmanned integration system. In *Proc. AUVSI Unmanned Systems Conference*, 2004. (Cité en page 9.)
- [For04] *Fastest in the World : The Saga of Canada's Revolutionary Hydrofoils*. Formac, 2004. (Cité en page 6.)
- [FPF96] A.W. Fitzgibbon, M. Pilu, and R.B. Fischer. Direct least squares fitting of ellipses. *the 13th International Conference on Pattern Recognition*, pages 253–257, 1996. (Cité en page 70.)
- [FTB02] J. Fabrizio, P. Tarel, and R. Benosman. Calibration of panoramic catadioptric sensors made easier. In *Proceedings of the IEEE Workshop on Omnidirectional Vision*, june 2002. (Cité en page 52.)
- [GG09] Y. Guo and B. Gas. Underwater transient and non transient signals classification using predictive neural networks. In *Proceedings of IROS'09*, 2009. (Cité en page 30.)
- [GIR⁺11] Y. Guo, S.H. Ieng, M. Romero, R. Benosman, and B. Gas. Obstacle detection using integration of omni-directional camera and inertial sensor. In *Proceedings of Clawar '11*, 2011. (Cité en page 82.)
- [Gon08] Xiaojin Gong. *Omnidirectional vision for an autonomous surface vehicle*. PhD thesis, Faculty of the Virginia Polytechnic Institute and State University, 2008. (Cité en pages 13 et 50.)
- [Goo] Google maps. (Cité en page 50.)
- [Gre89] P. Greguss. *Hugarian Patent 192125, usa patent 4566763, japan patent 1962784, german patent 3402847*, 1989. (Cité en page 49.)
- [GRI⁺11] Y. Guo, M. Romero, S.H. Ieng, F. Plumet, R.Benosman, and B. Gas. Reactive path planning for autonomous sailboat using an omni-directional camera for obstacle detection. In *2011 Iberoamerican Conference on Electronics Engineering and Computer Science (CIIIECC 2011)*, 2011. (Cité en page 50.)
- [GROJ08] J.G. Garcia, A. Robertsson, J.G. Ortega, and R. Johansson. Sensor fusion for compliant robot motion control. *Robotics, IEEE Transactions on*, 24, 2008. (Cité en page 82.)
- [GSPC06] P. Gomes, C. Silvestre, A. Pascoal, and R. Cunha. A path-following controller for the delfim autonomous surface craft. In *Proc. 7th IFAC conference on manoeuvring and control of marine vehicle*, 2006. (Cité en page 7.)

-
- [GXR⁺08] Xiaojin Gong, Bin Xu, Caleb Reed, Chris Wyatt, and Daniel Stilwell. Real-time robust mapping for an autonomous surface vehicle using an omnidirectional camera. *Applications of Computer Vision, IEEE Workshop on*, 0 :1–6, 2008. (Cité en pages 13 et 50.)
- [Hal92] D. Hall. *Mathematical Techniques in Multisensor Data Fusion*. MA : Artech House, 1992. (Cité en page 82.)
- [Ham47] William R. Hamilton. On quaternions. *Proceedings of the Royal Irish Academy*, 3(1847) :1–16, 1847. (Cité en page 86.)
- [Har95] R.I. Hartley. In defence of the eight point algorithm. In *International Conference on Computer Vision*, 1995. (Cité en page 55.)
- [HB00] R. Hicks and R. Bajcsy. Equi-areal catadioptric sensors. In *In Proc. CVPR*, pages 545–551, 2000. (Cité en page 50.)
- [Hen95] S.L. Henley. See you around. *Tom Halfhill-BYTE Magazine*, 1995. (Cité en page 49.)
- [HF98] R. Halir and J. Flusser. Numerically stable direct least squares fitting of ellipses. *Proceedings of WSCG'98*, pages 125–132, 1998. (Cité en page 70.)
- [HL97] D. L. Hall and J. Llinas. An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85 :6–23, 1997. (Cité en page 83.)
- [HLL91] D. L. Hall, R. J. Linn, and J. Llinas. A survey of data fusion systems. In *Proc. SPIE Conf. on Data Structure and Target Classification*, volume 1470, pages 13–36, 1991. (Cité en page 82.)
- [Hon91] J. Hong. Image based homing. In *proceedings of the international conference on robotics and automation*, 1991. (Cité en page 50.)
- [IEN05] Sio-Hoï IENG. *Etude des propriétés géométriques des capteurs panoramiques*. PhD thesis, UPMC, Paris 6, 2005. (Cité en pages 57 et 59.)
- [IMO79] *IMO Preference Standards for Automatic Radar Plotting Aids (ARPA)*. 1979. (Cité en page 12.)
- [ISY03] S. Ikeda, T. Sato, and N. Yokoya. High-resolution panoramic movie generation from video streams acquired by an omnidirectional multi-camera system. In *Multisensor Fusion and Integration for Intelligent Systems, MFI2003. Proceedings of IEEE International Conference on*, 2003. (Cité en page 49.)
- [KAB⁺92] O. Kessler, K. Askin, N. Beck, J. Lynch, and F. White. *Functional Description of the Data Fusion*. Tech Rep., Office of Naval Technol., 1992. (Cité en page 82.)
- [KCP94] A. Kundu, G. C. Chen, and C. E. Persons. Transient sonar signal classification using hidden markov models and neural nets. In *ICASSP '94*, volume 3, pages 325–328, 1994. (Cité en page 30.)

- [Kle93] L. A. Klein. *Sensor and Data Fusion Concepts and Applications*, volume 14. SPIE Opt. Engineering Press, 1993. (Cité en page 82.)
- [Koc] Kockums. *Kockums a part of ThyssenKrupp Marine Systems*. <http://www.kockums.se/en/products-services/naval-technologies/unmanned-systems/>. (Cité en page 6.)
- [Law04] A.K. Lawrence. *Sensor and data fusion : A tool for information assessment and decision making*. SPIE Press., 2004. (Cité en page 83.)
- [LBHL07] T. Lim, K. Bae, C. Hwang, and H. Lee. Classification of underwater transient signals using mfcc feature vector. pages 1– 4, 2007. (Cité en pages 30, 31, 37 et 44.)
- [LBHL08] T. Lim, K. Bae, C. Hwang, and H. Lee. Underwater transient signal classification using binary pattern image of mfcc and neural network. E91-A(3) :772– 774, 2008. (Cité en pages 30, 31, 34 et 44.)
- [Lev47] N. Levinson. The wiener rms (root mean square) error criterion in filter design and prediction. *J. Math. Phys.*, 25(4) :261–278, 1947. (Cité en page 33.)
- [LFAM88] K. Lashkari, B. Fredlander, J. Abel, and B. McQuiston. Classification of transient signals [aouctic signals]. In *Acoustics, Speech, and Signal Processing, ICASSP 88*, 1988. (Cité en page 30.)
- [LGBG09] L. Lachèze, Y. Guo, R. Benosman, and B. Gas. Audio/video fusion for objects recognition. In *The 2009 IEEE/RSJ International Conference on Intelligent RObots and Systems*, 2009. (Cité en page 103.)
- [Lic19] H. Lichte. On the influence of horizontal temperature layers in sea water on the range of underwater sound signals. *Physikalische Zeitschrift 17*, pages 385–389, 1919. (Cité en page 29.)
- [Lim07] P.U. Lima. A bayesian approach to sensor fusion in autonomous sensor and robot networks. *Instrumentation and Measurement Magazine, IEEE*, 10 :22–27, 2007. (Cité en page 82.)
- [LKPP08] H. D. Lee, D.G. Kim, M. C. Park, and G. T. Park. Augmented reality based vision system for network based mobile robot. *Lecture Notes in Computer Science*, pages 123–130, 2008. (Cité en page 49.)
- [LLS94] M. Chaumet Lagrange, H. Loebe, and S.Ygorra. Design of an autonomous surface vehicle. In *Proc. OCEAN'94*, pages 120–1225, 1994. (Cité en page 7.)
- [LS08] R.C. Luo and K.L. Su. Multilevel multisensor-based intelligent recharging system for mobile robot. *Industrial Electronics, IEEE Transactions on*, 55 :270–279, 2008. (Cité en page 82.)

-
- [LSA] LSA. *ROAZ Robot*. <http://www.lsa.isep.ipp.pt/pages/roazroaz.html>. (Cit  en page 9.)
- [LW90] J. Llinas and E. Waltz. *Multisensor Data Fusion*. MA : Artech House, 1990. (Cit  en page 82.)
- [Ma93] A. Mellouk and P. allinary. A discriminative neural prediction system for speech recognition. In *ICASSP*, volume 2, pages 533–536, 1993. (Cit  en page 34.)
- [Man97] J. Manley. Development of the autonomous surface craft "aces". In *Proc. Oceans' 94*, volume 2, pages 827–832, 1997. (Cit  en page 6.)
- [MBK00] J. Majohr, Th. Buch, and C. Korte. Navigation and automatic control of the measuring dolphin (messin). In *Proc. 5th IFAC Conference on Manoeuring and control of marine crafts*, pages 405–410, 2000. (Cit  en page 7.)
- [MG76] J. D. Markel and A. H. Gray. Linear prediction of speech. *Communication and Cy- bernetics*, 1976. (Cit  en page 33.)
- [MHTK05] J. Meguro, T. Hashizume, J. Takiguchi, and R. Kurosaki. Development of an autonomous mobile surveillance system using a network-based rtk-gps. *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 3096 – 3101, 2005. (Cit  en page 50.)
- [Mit07] H.B. Mitchell. *Multi-Sensor Data Fusion : An Introduction*. Springer-Verlag Berlin and Heidelberg GmbH and Co. K, 2007. (Cit  en pages 82 et 83.)
- [MM54] D. Middleton and D. Van Meter. Modern statistical approaches to reception in communication theory. *IRE Trans. Information Theory*, 1954. (Cit  en page 30.)
- [MMCW00] J. Manley, A. Marsh, W. Cornforth, and C. Wiseman. Evolution of the autonomous surface craft autocat. In *Proc. Oceans'00*, volume 1, pages 403–408, 2000. (Cit  en page 7.)
- [MNG⁺08] V. Milanes, J.E. Naranjo, C. Gonzalez, J. Alonso, and T.D. Pedro. Autonomous vehicle based in cooperative gps and inertial systems. *RObotica*, 26, 2008. (Cit  en page 82.)
- [MSB⁺09] P. Mishra, H.N. Shankar, J.S. Bhat, S.R. Kubair, S.H. Bharadwaja, S. Anudhan, D.K. Hundi, and G. Kamath. Multi-sensor autonomous robot based manipulation of valves for process control. In *Modelling and Simulation, 2009. AMS '09. Third Asia International Conference on*, pages 442–447, 2009. (Cit  en page 82.)
- [MW97] S. Del Marko and J. Weiss. Improved transient signal detection using a wavepacket-based detector with an extended translation-invariant wavelet transform. In *IEEE trans. signal process*, volume 45, pages 841–850, 1997. (Cit  en page 31.)

- [Nal96] V.S. Nalwa. A true omnidirectional viewer. Technical report, Bell Laboratories, Holmdel, NJ 07733, USA, February 1996. (Cité en page 50.)
- [NB97] S.K. Nayar and S. Baker. A theory of catadioptric image formation. Technical report, Department of Computer Science, Columbia University, USA, April 1997. (Cité en page 50.)
- [Nea06] M. Neal. A hardware proof of concept of a sailing robot for ocean observation. *IEEE Journal of Oceanic Engineering*, 31(462-469), April 2006. (Cité en page 12.)
- [NSC06] W. Naeem, R. Sutton, and J. Chudley. Soft computing design of a linear quadratic gaussian controller for an unmanned surface vehicle. In *Proc. 14th Mediterranean conference on control and automation*, pages 1–6, 2006. (Cité en page 8.)
- [NYP10] G. Niu, B.S. Yang, and M. Pecht. Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance. *Reliability Engineering and System Safety*, 95 :786–796, 2010. (Cité en page 82.)
- [NYY07] H. Nagahara, K. Yoshida, and M. Yachida. An omnidirectional vision sensor with single view and constant resolution. In *Computer Vision IEEE 11th International Conference on*, pages 1–8, 2007. (Cité en page 50.)
- [OLBMP02] P. M. Oliveira, V. Lobo, V. Barroso, and F. Moura-Pires. Detection and classification of underwater transients with data driven methods based on time-frequency distributions and non-parametric classifiers. In *OCEANS '02*, volume 1, pages 12–16, 2002. (Cité en page 30.)
- [PB08] S. Prasad and L.M. Bruce. Decision fusion with confidence-based weight assignment for hyperspectral target recognition. In *Geoscience and Remote Sensing, IEEE Transactions on*, volume 46, pages 1448–1456, 2008. (Cité en page 82.)
- [PBC77] H. M. Praetorius, G. Bodenstein, and O. D. Creutzfeldz. Adaptive segmentation of eeg records : a new approach to automatic eeg analysis. *Electroencephalogr Clin Neurophysiol.*, pages 84–94, 1977. (Cité en page 30.)
- [PBF54] W. Peterson, T. Birdsall, and W. Fox. The theory of signal detectability. *IRE Trans. Information Theory*, pages 171–212, 1954. (Cité en page 30.)
- [PRH⁺92] R. D. Patterson, K. Robinson, J. Holdsworth, D. McKeown, C.Zhang, and M.H. Allerhand. Complex sounds and auditory images. pages 429–446, 1992. (Cité en page 31.)
- [Pro97] Proceedings of the 1997 DARPA Image Understanding Workshop. *Catadioptric image formation*, New Orleans, May 1997. (Cité en pages 49 et 50.)

-
- [Pro09] Prosilica. *Prosilica PvAPI Programmers' Reference Manual*, 2009. (Cité en page 21.)
- [PW92] T. W. Parks and B. A. Weisburn. Classification of whale and ice sounds with a cochlear model. In *ICASSP '92*, volume 2, pages 481–484, 1992. (Cité en page 30.)
- [Ree70] D. Rees. Panoramic television viewing system. United States Patent No. 3,505,465, 1970. (Cité en page 49.)
- [RGI⁺11] Miguel Romero, Yan Guo, Sio-Hoi Ieng, Frederic Plumet, Ryad Benosman, and Bruno Gas. Omni-directional camera and fuzzy logic path planner for autonomous sailboat navigation. In *Iberoamerican Conference on Electronics Engineering and Computer Science*, 2011. (Cité en page 50.)
- [Ric70] G. S. Richie. Problems in bathymetric surveying problems in bathymetric surveying presented by modern trends in shipbuilding. *Radio and Electronic Engineer* 40 219-24, 1970. (Cité en page 29.)
- [Rob83] J. Urick Robert. *Principles of Underwater Sound*. McGraw-Hill(Tx), 1983. (Cité en page 28.)
- [Rob06] G. N. Roberts and Robert Sutton. Institution of Engineering and Thchnology, 2006. (Cité en page 6.)
- [RS08] H. Rababaah and A. Shirkhodaie. Fusion-based multi-target tracking and localization for intelligent visual surveillance systems. *Proceedings of SPIE, the International Society for Optical Engineering*, 2008. (Cité en page 82.)
- [RvE10] P.F. Rynne and K.D. von Ellenrieder. Development and preliminary experimental validation of a wind- and solar-powered autonomous surface vehicle. *IEEE Journal of Oceanic Engineering*, 35 :971–983, 2010. (Cité en page 12.)
- [SB66] J. C. Steinberg and T. G. Birdsall. Underwater sound propagation in the straits of florida. *J. ACOUST. SOC. AM.*, 39 :301–315, 1966. (Cité en page 30.)
- [SBD09] V. Subramanian, T.F. Burks, and W.E. Dixon. Sensor fusion using fuzzy logic enhanced kalman filter for autonomous vehicle guidance in citrus groves. In *Transactions of the ASAE*, volume 52, pages 1411–1422, 2009. (Cité en page 82.)
- [SBFR96] D. Southwell, A. Basu, M. Flala, and J. Rayda. Panoramic stereo. In *Proceedings of the international conference on pattern recognition*, 1996. (Cité en page 49.)
- [SGN06] Rahul Swaminathan, Michael D. Grossberg, and Shree K. Nayar. Non-single viewpoint catadioptric cameras : Geometry and analysis. *INTERNATIONAL JOURNAL OF COMPUTER VISION*, 66(3), 2006. (Cité en page 50.)

- [SGR⁺06] Anbumani Subramanian, Xiaojin Gong, Jamie N. Riggins, Daniel J. Stilwell, and Christopher L. Wyatt. Shoreline mapping using an omni-directional camera for autonomous surface vehicle applications. In *OCEANS*, number 1-6, 2006. (Cité en pages 13 et 50.)
- [SH00] J.Z. Sasiadek and P. Hartana. *Sensor data fusion using Kalman filter*, volume 2, pages 19–25. IEEE, 2000. (Cité en page 106.)
- [Sho94] K. Shoemake. Euler angle conversion. *Graphics gems IV*, pages 222–229, 1994. (Cité en page 87.)
- [SLZ06] L.C. Su, C.J. Luo, and F. Zhu. Obtaining obstacle information by an omnidirectional stereo vision system. In *Proc. International Conference on Information Acquisition*, 2006. (Cité en page 50.)
- [SMH⁺04] F. Snyder, D. D. Morris, H. P. Haley, T. R. Collins, and M. A. Okerholm. Autonomous river navigation. In *Mobile robots XVII*, 2004. (Cité en page 13.)
- [Sri03] M.V. Srinivasan. A new class of mirrors for wide-angle imaging. In *Proc. IEEE Workshop on Omnidirectional Vision and Camera Networks*, Juin 2003. (Cité en page 50.)
- [SS08] N. Subrahmanya and Y.C. Shin. Automated sensor selection and fusion for monitoring and diagnostics of plunge grinding. *Journal of Manufacturing Science and Engineering*, 130, 2008. (Cité en page 82.)
- [Stu00] P. Sturm. Algorithms for plane based pose estimation. In *Proceedings of the IEEE International conference on Computer Vision and Pattern Recognition*, pages 706–711, 2000. (Cité en page 52.)
- [TB05] S. Tucker and G.J. Brown. Classification of transient sonar sounds using perceptually motivated features. 30(3) :588 – 600, 2005. (Cité en page 30.)
- [Tho85] C. Thomas. Continuing program of the dolphin semi submersible. *Unmanned, Untethered Submersible Technology*, 4 :223–229, 1985. (Cité en page 6.)
- [Tri09] Tritech. *Software Notes for controlling and operating RS-232 Sonar Heads*, 2009. (Cité en page 24.)
- [Var97] P.K. Varshney. *Distributed detection and data fusion*. Springer, 1997. (Cité en pages 82 et 83.)
- [WB71] R. F. Williams and H. F. Battestin. Coherent recombination of acoustic multipath signals propagated in the deep ocean. *J. ACOUST. SOC. AM.*, 50 :1433–1442, 1971. (Cité en page 30.)
- [WB95] Greg Welch and Gary Bishop. An introduction to the kalman filter. Technical report, University of North Carolina at Chapel Hill Chapel Hill, NC, USA, 1995. (Cité en page 108.)

-
- [Wer07] *The Naval Institute Guide to Combat Fleets of the World : Their Ships, Aircraft, and Systems*. United States Naval Institute, 2007. (Cité en page 6.)
- [Wika] Wikipedia. *Horizon*. <http://fr.wikipedia.org/wiki/Horizon>. (Cité en page 68.)
- [Wikb] Wikipedia. *StanFlex*. <http://en.wikipedia.org/wiki/StanFlex>. (Cité en page 6.)
- [Wu08] W.H. Wu. *MEDIC : An end-to-end biomedical system based on active sensor fusion*. PhD thesis, UNIVERSITY OF CALIFORNIA, 2008. (Cité en page 82.)
- [Xse03] Xsens. *Technical Documentation MT9 and MT6*, 2003. (Cité en page 23.)
- [Yen98] N. Yen. Time and frequency representation of acoustic signals by means of the wigner distribution : Implementation and interpretation. 81(6) :1841 – 1850, 1998. (Cité en pages 30 et 31.)
- [YK90] Y. Yagi and S. Kawato. Panoramic scene analysis with conic projection. In *proceedings of the international conference on robots and systems*, 1990. (Cité en pages 49 et 50.)
- [YPSP10] R.J. Yan, S. Pang, H.B. Sun, and Y.J. Pang. Development and missions of unmanned surface vehicle. *Journal of marine science and application*, 9(4) :451–457, 2010. (Cité en page 11.)
- [YYY93] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional imaging with hyperboloidal projection. In *proceedings of the international conference on robotics and automation*, 1993. (Cité en pages 49 et 50.)
- [ZH07] Z. Zhu and T.S. Huang. *Multimodal surveillance : sensors, algorithms and systems*. Artech House Publishers, 2007. (Cité en page 82.)