



HAL
open science

Perception de l'environnement marin et de la menace à partir d'images sonar haute fréquence : du fond marin à la surface de mer

Isabelle Quidu

► **To cite this version:**

Isabelle Quidu. Perception de l'environnement marin et de la menace à partir d'images sonar haute fréquence : du fond marin à la surface de mer. Traitement du signal et de l'image [eess.SP]. Université de Bretagne Occidentale (UBO), Brest, 2018. tel-01840696v1

HAL Id: tel-01840696

<https://hal.science/tel-01840696v1>

Submitted on 16 Jul 2018 (v1), last revised 24 Sep 2018 (v2)

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.

HABILITATION A DIRIGER DES RECHERCHES

DE
L'UNIVERSITE DE BRETAGNE OCCIDENTALE
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : Signal, Image, Vision

Par

Isabelle QUIDU, *Maître de Conférences*

**« Perception de l'environnement marin et de la menace à partir
d'images sonar haute fréquence : du fond marin à la surface de mer »**

présentée et soutenue à l'ENSTA Bretagne, Brest, le 26 janvier 2018
Unité de recherche : Lab-STICC UMR 6285

Rapporteurs avant soutenance :

Philippe BLONDEL, Senior lecturer, University of Bath
Christophe COLLET, Professeur, Université de Strasbourg
Basel SOLAIMAN, Professeur, IMT Atlantique, Brest

Composition du Jury :

Président
Jean-Marie NICOLAS, Professeur, Telecom ParisTech
Directeur de recherche
Gilles BUREL, Professeur, Univ. de Bretagne Occidentale

Philippe BLONDEL, Senior lecturer, University of Bath
Christophe COLLET, Professeur, Université de Strasbourg
Basel SOLAIMAN, Professeur, IMT Atlantique, Brest
Ali KHENCHAF, Professeur, ENSTA Bretagne

Invité
Michel LEGRIS, Enseignant-Chercheur, ENSTA Bretagne

Remerciements

Mes tout premiers remerciements vont à Gilles Burel qui après avoir été mon Directeur de thèse de 1998 à 2001 a accepté d'être mon Directeur de recherche en 2016. Ses conseils avisés et sa bienveillance ont contribué à la qualité de cette synthèse.

Je tiens à remercier chaleureusement les rapporteurs, Philippe Blondel, Christophe Collet et Basel Solaiman, pour leur travail d'expertise et la qualité de nos échanges. Je remercie également Ali Khenchaf qui m'a fait confiance en 2005 en m'accueillant à l'école dans son laboratoire et pour l'examen de ce travail de synthèse. Enfin, un grand merci à Jean-Marie Nicolas, Président du jury, pour avoir mené la soutenance avec justesse et bonne humeur, et pour son éclairage sur la cohérence.

Les travaux présentés sont le fruit de diverses collaborations et rencontres. Je remercie le DRDC-Atlantique et tout particulièrement John Fawcett qui m'a accueillie en 2011 pour un séjour d'étude et de recherche qui a été à l'origine de nombreux échanges avec la communauté mondiale de la reconnaissance de cibles sous-marines. J'ai une pensée amicale pour mes anciens collègues de Thales avec qui j'ai débuté dans le domaine de l'imagerie sonar et qui m'ont touchée par leur présence à la soutenance ou leurs sympathiques messages. Je tiens à souligner la qualité du travail des stagiaires et doctorants que j'ai eu l'honneur d'encadrer, et des ingénieurs de recherche dont les contributions tout au long de ce manuscrit sont significatives. Je termine par remercier l'ensemble de mes collègues de l'ENSTA Bretagne et du Lab-STICC avec qui je travaille et j'échange avec plaisir depuis plusieurs années sous diverses formes, projets, co-encadrements, séminaires, colloques et pauses-café...

Mille mercis aux proches collègues et amis qui m'ont encouragée dans cette ambitieuse démarche pour l'obtention de l'HDR. Je n'oublie pas les petites attentions ou coups de main d'Alex, Gilles, Natacha, Michel I., Ali M., Annick..., le précieux soutien de ma chère amie Naig, l'indulgence d'Hélène, ma collègue de bureau, qui m'a (sup)portée au gré de mes humeurs, et bien sûr la disponibilité de Michel L., toujours prompt à répondre à mes questions et à partager son expérience du domaine avec enthousiasme.

Enfin merci à Richard pour sa présence,... et sa contribution à l'abstract !

Table des matières

1	INTRODUCTION	11
2	CURRICULUM VITAE ETENDU	15
2.1	DIPLOMES ET ETUDES SUPERIEURES	15
2.2	EXPERIENCE PROFESSIONNELLE	15
2.3	ACTIVITES PEDAGOGIQUES	17
2.3.1	<i>Enseignements</i>	17
2.3.2	<i>Responsabilités, jurys et suivi</i>	18
2.4	REPRESENTATION DU PERSONNEL DANS DIFFERENTS CONSEILS	18
2.5	ACTIVITES D'ENCADREMENT	19
2.5.1	<i>Stagiaires (stage de 5 mois en moyenne)</i>	19
2.5.2	<i>Co-encadrement doctoral</i>	19
2.6	ACTIVITES DE RECHERCHE CONTRACTUELLE (DEPUIS SEPTEMBRE 2005)	21
2.6.1	<i>Responsabilité de contrat</i>	21
2.6.2	<i>Participation dans divers contrats</i>	23
2.7	COLLABORATIONS INTERNATIONALES	23
2.8	RAYONNEMENT	24
2.9	ANIMATION SCIENTIFIQUE (DEPUIS SEPTEMBRE 2005)	25
2.10	EXPERTISES	26
2.11	PUBLICATIONS	27
3	SYNTHESE DES TRAVAUX	35
3.1	PRESENTATION DE L'IMAGERIE SONAR HAUTE FREQUENCE ET DE SON USAGE MILITAIRE	35
3.2	RECONNAISSANCE AUTOMATIQUE DE CIBLE SUR LE FOND MARIN	36
3.2.1	<i>Chasse aux mines</i>	37
3.2.1.1	Un exemple de chaîne classique de détection de cibles sur images sonar latéral	42
3.2.1.1.1	Contexte	42
3.2.1.1.2	Description de l'algorithme	42
3.2.1.1.3	Résultats	45
3.2.1.2	Nouvelle approche de l'ATR	48
3.2.1.2.1	Prise en compte de l'environnement proche	48
3.2.1.2.2	Lien avec l'état de l'art	48
3.2.1.2.3	Une autre qualification de l'environnement : travaux du doctorant Laurent Picard	50
3.2.1.3	Classification multi-vue	54
3.2.1.3.1	Etat de l'art	54
3.2.1.3.2	Géométrie floue pour l'analyse des données multi-aspects du DUBM44	55
3.2.2	<i>Surveillance</i>	58
3.2.2.1	Positionnement des travaux	59
3.2.2.2	Détection incohérente	60
3.2.2.2.1	Travaux réalisés lors d'un séjour de recherche au Canada en 2011 [QUI12a]	60
3.2.2.2.2	Travaux du doctorant Florian Nicolas (Thèse CIFRE)	63
3.2.2.3	Détection cohérente	64
3.2.2.3.1	Travaux réalisés en collaboration avec le DRDC-A et FFI [QUI12b]	64
3.2.2.3.2	Travaux du doctorant Vincent Myers	71
3.2.3	<i>Un point délicat : dimensionnement de la fenêtre de calcul</i>	73
3.3	DETECTION D'OBSTACLES POUR VEHICULE SOUS-MARIN AUTONOME	75
3.3.1	<i>Détection, suivi et caractérisation des obstacles dans la colonne d'eau</i>	76
3.3.2	<i>Reprise de vue du véhicule</i>	83
3.3.3	<i>Suivi de fond</i>	86
3.3.4	<i>Commentaires sur ces travaux consacrés à la détection d'obstacles</i>	88
4	PROJET DE RECHERCHE	89

4.1	ANIMATION EQUIPE PRASYS.....	89
4.1.1	<i>Historique et positionnement</i>	89
4.1.2	<i>Contribution</i>	90
4.2	EMERGENCE DU ROBOT MOBILE SOUS-MARIN POUR LE MILITAIRE ET LE CIVIL.....	92
4.2.1	<i>Reconnaissance de cibles (détection/classification)</i>	93
4.2.2	<i>Identification de cibles et inspection d'ouvrage</i>	95
5	CONCLUSION	101
6	BIBLIOGRAPHIE	103
	ANNEXE 1: SEGMENTATION ORIENTEE CLASSIFICATION	115
	ANNEXE 2 : PROCESSUS ATR DEVELOPPE CHEZ THALES EN 2005	117
	ANNEXE 3 : DETECTION INCOHERENTE DE CHANGEMENT	119
	ANNEXE 4 : DETECTION COHERENTE DE CHANGEMENT	121
	ANNEXE 5 : DETECTION ET SUIVI D'OBSTACLES DANS LA COLONNE D'EAU	123
	ANNEXE 6 : AIDE A LA REPRISE DE VUE DES AUV	125

Abréviations

AMI	Acoustic Mine Imaging
ASM	Acoustique Sous-Marine
ASA	Acoustical Society of America
ATR	Automatic Target Recongnition
AUV	Autonomous Underwater Vehicle
BEGM	Bâtiment d'Expérimentation de la Guerre des Mines
BF	Basse Fréquence
BRS	Bâtiment Remorqueur de Sonar
CAD	Computer Aided Detection
CAC	Computer Aided Classification
CFAR	Constant False Alarm Rate
CI	Cycle Ingénieur
CID	Connaissance Information Décision
CIFRE	Convention Industrielle de Formation par la Recherche
CMM	Caractérisation du Milieu Marin
CMRE	Centre for Maritime Research and Experimentation
CMT	Chasseur de Mine Tripartite
DAURADE	Démonstrateur d'AUV de ReA DiscrEt
DCL	Détection Classification Localisation
DCUT	Detection and Classification of Underwater Targets
DESS	Diplôme d'Etudes Supérieures Spécialisées
DGA	Direction Générale de l'Armement
DGA TN	DGA Techniques Navales
DIDSON	Dual-frequency IDentification SONar
DLS	Downward Looking Sonar
DPCA	Displaced-Phase Centre Array
DRDC(-A)	Defence Research and Development Canada (Atlantic)
DSTO	Australian Defence Science & Technology Organisation
DUBM	Détection Ultrason Bâtiment Mines
DVL	Doppler Velocity Log

EGS	Etudes Générales Sonar
ENSIETA	Ecole Nationale Supérieure des Ingénieurs des Etudes et Techniques de l'Armement
ENSTA	Ecole Nationale Supérieure de Techniques Avancées
ERE	Etudes et Recherche à l'Etranger
ESCO	Electronique des Systèmes Communicants
ESPADON	Evaluation incrémentale de Solutions Potentielles d'Automatisation de Déminage pour les Opérations Navales
E ³ I ²	Extraction et Exploitation de l'Information en Environnements Incertains
FFI	<i>Forsvarets ForskningsInstitutt</i> (Norwegian Defence Research Establishment)
FIPA	Formation d'Ingénieurs Par Alternance
FLS	Forward Looking Sonar
FUI	Fonds Unique Interministériel
GESMA	Groupe d'Etudes Sous-Marines de l'Atlantique
HCERES	Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur
HDR	Habilitation à Diriger des Recherches
HF	Haute Fréquence
HMS	Hull Mounted Sonar
ICoURS	International Conference on Underwater Remote Sensing
IED	Improvised Explosive Device
IHM	Interface Homme Machine
IHSEV	Interaction Humain Système et Environnement Virtuel
IMT	Institut Mines-Télécom
INS	Inertial Navigation System
IRIS	securItly and ReResilience of Information Systems
iSAAM	Intersections in Signal processing, Acoustics & ATR for Maritime applications
JPDAF	Joint Probabilistic Data Association Filter
Lab-STICC	Laboratoire des sciences et techniques de l'information, de la communication et de la connaissance
LED	Light-Emitting Diode
MCM	Mine Counter Measures

MILCO	MIne Like COntact
MOQESM	MOntoring Quantitatif de l'Environnement Sous-Marin
MUSE	Multi Usage Sonar Element
NATO	North Atlantic Treaty Organization
NRL	Naval Research Laboratory
NURC	NATO Undersea Research Centre
OAS	Obstacle Avoidance System
PAP	Poisson Auto-Propulsé
PEA	Programme d'Etudes Amont
PRASYS	Perception, Robotics and Autonomous SYStems
RAPID	Régime d'Appui Pour l'Innovation Duale
RDDC	Recherche et Développement Défense Canada
REA	Rapid Environment Assessment
REMUS	Remote Environmental Monitoring Unit
ROV	Remotely Operated Vehicle
SA	Système Autonome
SAMDIS	Synthetic Aperture and Mine Detection Imagery Sonar
SAMM	Systèmes Autonomes en Milieu Maritime
SAS	Synthetic Aperture Sonar
SFIIS	Sécurité, Fiabilité et Intégrité de l'Information et des Systèmes
SHOM	Service Hydrographique et Océanographique de la Marine
SIFT	Scale Invariant Feature Transform
SMF	Sondeur MultiFaisceau
SNR	Signal to Noise Ratio
SPIV	Sonar Propulsé à Immersion Variable
SSS	SideScan Sonar
SURF	Speeded Up Robust Features
TFAC	Taux de Fausses Alarmes Constant
TOPVISION	Test Opérationnels de Vidéos Sous-marines pour l'Identification d'Objets Nuisibles
T-SAS	Towed SAS
TUS	Thales Underwater Systems
ULS	Upward Looking Sonar

UMR	Unité Mixte de Recherche
USV	Unmanned Surface Vehicle
UV	Unité de Valeur
UXO	UneXploded Ordnance

1 Introduction

Les ondes sonores sont des ondes mécaniques qui de fait se propagent naturellement bien dans l'eau. A l'instar des mammifères marins, les hommes ont mis en œuvre dès le début du 20^{ème} siècle des systèmes capables de percevoir l'océan sur des distances plus ou moins longues et avec plus ou moins de détails en fonction de leurs besoins. Ces systèmes ont été baptisés SONAR « SOund NAvigation and Ranging » en raison de leur rôle premier qui visait à déterminer la présence d'un « objet » vu comme un obstacle à la propagation de l'onde sonore émise qui en se réfléchissant permettait d'en déduire la distance compte tenu du temps de trajet aller-retour effectué. La cible recherchée à cette époque est en tout premier lieu un sous-marin ou un iceberg. En parallèle des avancées technologiques, l'usage des systèmes sonar s'est évidemment diversifié, pour la détection de cibles beaucoup plus petites comme les mines ou pour la caractérisation des fonds marins par exemple.

Pour des raisons physiques, l'intensité de l'onde décroît avec la distance et d'autant plus que sa fréquence est élevée. Malheureusement, une basse fréquence équivaut à une grande longueur d'onde et par suite, potentiellement peu de détails sur l'environnement. D'une part, parce que toute aspérité plus petite que la longueur d'onde ne renverra aucune signature acoustique perceptible et d'autre part, puisque la directivité des transducteurs est proportionnelle à la longueur d'onde. Si l'on ne considère que les « objets » acoustiquement visibles c'est-à-dire grands devant la longueur d'onde, leur reconnaissance exige cependant une résolution sonar suffisante. Avec une antenne linéaire de 1m de longueur par exemple, compte tenu de la célérité moyenne de l'onde dans l'eau (1500m/s), une résolution transversale minimale de 15cm à 40m de distance nécessite une fréquence minimale de 400kHz, qui correspond à une longueur d'onde de 3.75mm. Cette résolution angulaire, qui se dégrade par conséquent en distance, peut être améliorée avec une plus grande antenne mais se posent des problèmes de coût et d'encombrement. Ces contraintes physiques (atténuation en distance et/ou longueur de l'antenne) peuvent néanmoins être en partie dépassées par la mise en œuvre de traitements avancés de synthèse d'ouverture issus du radar et adaptés au sonar au début des années 90 pour en faire des systèmes sonar synthétiques qui, outre le fait d'améliorer la résolution transversale, la garantit constante sur l'ensemble de la portée. Malgré cela, une fréquence minimale de 100 kHz semble nécessaire pour atteindre un niveau de détails suffisant à l'exploration du milieu marin par sonars latéraux ou frontaux. Ce sont ces systèmes sonar d'imagerie dits haute fréquence, qu'ils soient conventionnels ou synthétiques, qui nous intéressent dans ce document pour percevoir l'océan du fond jusque la surface en

passant par la colonne d'eau dans un but essentiellement militaire de neutralisation de la menace.

Pour chaque zone prospectée et type de menace visé, des systèmes et des algorithmes spécifiques doivent être développés. Pour la chasse aux mines, les systèmes les plus utilisés ont été les sonars de coque et les sonars latéraux tractés. Mais avec l'évolution de la menace (mines furtives, dispositifs artisanaux), les systèmes sont de plus en plus souvent déportés sur de petits véhicules sous-marins, plus maniables et plus sécurisants. Au cours d'une passe, un système d'imagerie fournit une succession d'images ou données relatives à chaque récurrence sonar où seuls un œil expérimenté ou un automate entraîné détecteront une signature acoustique (écho, ombre) ou toute autre anomalie (statistique...) indiquant la présence d'une éventuelle menace, le tout compte tenu d'un contexte opérationnel donné. S'agissant de passes répétées sur des intervalles de temps relativement courts, nous verrons que, même en haute fréquence et donc sans possibilité de pénétrer le sédiment, il est possible de détecter automatiquement une menace imperceptible à l'œil comme un mouvement de sédiment. Si le véhicule sous-marin est autonome, il doit percevoir son environnement pour mener à bien sa mission mais aussi assurer sa propre survie, en cours puis en fin de mission. Ceci nous amène à la détection et au suivi d'obstacles dans le sens d'avance du véhicule à l'aide d'un sonar frontal, utilisé en mode horizontal le plus souvent. Les sonars frontaux sont des sonars sectoriels ou des sondeurs multifaisceau disposant du mode frontal. Outre les contraintes d'encombrement et de poids, le système d'évitement d'obstacles est classiquement choisi pour ses performances (portée, résolution...) moyennant un prix bien plus faible que les sonars de chasse aux mines. L'angle de dépointage est adapté en fonction de la position des obstacles (devant, sur le fond ou en surface) et des algorithmes spécifiques à chaque mission (évitement d'urgence, suivi de fond, reprise de vue) sont développés. Ces algorithmes sont robustes et efficaces pour fournir en quelques récurrences seulement l'information nécessaire à la manœuvre d'urgence.

Nous apprendrons au fil de ce manuscrit que le traitement d'images sonar n'exige pas seulement des compétences en traitement d'image et en acoustique mais également la connaissance et la pratique des systèmes, avec les principes de formation des images et les modes d'acquisition des données. Par ailleurs, le système sonar n'est qu'une partie d'un ensemble de systèmes qui comprend son porteur et les autres capteurs dans un environnement complexe et variable. En d'autres termes, l'intensité du pixel de l'image sonar est le résultat combiné de la scène « insonifiée » et d'un ensemble de facteurs, technologiques, opérationnels et physiques, qu'il convient de maîtriser au mieux.

Ce manuscrit regroupe l'ensemble des éléments significatifs de mon parcours professionnel m'amenant aujourd'hui à demander l'Habilitation à Diriger des Recherches.

Après plus de 15 ans de recherches industrielle et académique dans la problématique de la reconnaissance de cibles sous-marines, l'encadrement d'étudiants sur divers projets ou stages, et plus récemment de doctorants et d'ingénieurs de recherche, j'ai l'honneur de solliciter l'avis de mes pairs sur mon expérience et ma capacité à diriger des thèses.

Plan et organisation du manuscrit

Le chapitre 2 est une présentation synthétique de mon parcours et de mes activités d'enseignement et de recherche sous forme de Curriculum Vitae étendu. Le chapitre 3 présente de manière synthétique un ensemble de contributions à deux sous-domaines de l'imagerie sonar: le premier concerne la reconnaissance de cibles sous-marines sur le fond (paragraphe 3.2) et le second, la détection d'obstacles pour véhicule sous-marin autonome (paragraphe 3.3). La contribution des doctorants, ingénieurs de recherche et stagiaires y est clairement mentionnée, de manière plus ou moins détaillée en fonction de la proximité des sujets avec la sélection de travaux choisie. Cette synthèse est complétée par quelques réflexions personnelles sur le positionnement et les points durs de ces travaux. Mes activités initiales de développement d'algorithmes de reconnaissance de cibles pour l'aide aux opérateurs sonar des chasseurs de mines se réorientent aujourd'hui vers l'automatisation de cette tâche pour des véhicules sous-marins autonomes avec un corollaire en détection d'obstacles pour la sécurisation de la mission. Ceci est cohérent avec ma nouvelle responsabilité d'animatrice d'équipe au sein du Lab-STICC UMR 6285 qui a souhaité encourager l'activité en robotique mobile autonome de perception de l'environnement physique en créant l'équipe « Perception, Robotics and Autonomous SYStems » (PRASYS). Le challenge du lancement de cette équipe et le renforcement de mes activités en développement d'algorithmes de perception pour les véhicules sous-marins autonomes sont développés dans le chapitre 4 de ce document.

2 Curriculum Vitae étendu

2.1 Diplômes et études supérieures

2001 **Doctorat** de l'Université de Bretagne Occidentale, *Mention Très Honorable*

1997 **DESS Electronique** à l'Université de Bretagne Occidentale, *Mention Bien (major de la promotion)*

Domaines étudiés: l'électronique des supports de transmissions, le traitement et l'exploitation des signaux associés, les communications et l'électronique du milieu marin.

1996 **Maîtrise EEA (Electronique Electrotechnique et Automatique)** à l'Université de Rennes 1, *Mention Très Bien (major de la promotion)*

2.2 Expérience professionnelle

- Depuis le 1^{er} septembre 2005, **Maître de Conférences**, section 61, à l'Université de Bretagne Occidentale mise à disposition de l'ENSTA Bretagne à Brest,
 - membre permanent de E³I² (Extraction et Exploitation de l'Information en Environnements Incertains), équipe d'accueil 3876, de 2005 à 2011,
 - membre permanent du Lab-STICC (Laboratoire des sciences et techniques de l'information, de la communication et de la connaissance), unité mixte de recherche 6285 du CNRS, à partir de 2012.

Activités de recherche : traitement d'images sonar appliqué au domaine de la guerre des mines, perception par moyens acoustique et optique de l'environnement marin à partir de véhicules sous-marins autonomes (détection et classification d'objets immergés sur le fond, dans la colonne d'eau et sub-surface), connaissance des systèmes sonar haute fréquence, expérimentation à la mer.

Relations industrielles : contrats et/ou thèse CIFRE avec les sociétés Thales Underwater Systems et iXblue.

Relations étatiques : Contrats de recherche avec la DGA Techniques Navales sur les moyens acoustiques de perception et de détection d'obstacles pour véhicule sous-marin autonome.

Collaborations internationales : 7 mois de séjour d'études et de recherche à l'étranger (bourse DGA ERE) au DRDC-A, Canada ; Campagne d'essais et échange de données avec FFI, Norvège.

Mots clés : imagerie sonar, traitement d'images, perception, reconnaissance et pistage de cibles sous-marines.

Enseignements (moyenne annuelle de 192h équivalent TD) : [cours, TD, TP] dispensés au sein de l'ENSTA Bretagne dans le cadre des formations initiale, par alternance et continue, ainsi qu'à l'Université de Bretagne Occidentale et à l'Ecole Navale.

Responsabilités collectives: co-organisation de colloques/workshops, rapporteur (revues, conférences), responsabilité de matières, représentation du personnel dans différents conseils.

- De novembre 2001 à août 2005, **Ingénieur Traitement du Signal au Service Etudes Générales Sonar** chez Thales Underwater Systems à Brest (10 personnes).
 - Réalisation de travaux d'étude amont (Contrats DGA, études autofinancées),
 - Support fonctionnel (définition algorithmique et validation) aux programmes des lignes de produit.

- De novembre 1998 à octobre 2001, **préparation d'une thèse de doctorat sous contrat CIFRE** (Convention Industrielle de Formation par la Recherche) au Service Etudes Générales Sonar chez Thales Underwater Systems à Brest en partenariat avec le Laboratoire d'Electronique et Systèmes de Télécommunications (LEST - UMR CNRS 6165), soutenue le 11/12/2001 à l'Université de Bretagne Occidentale devant le jury :

Directeurs de thèse :

Pierre VILBÉ	Université de Bretagne Occidentale Laboratoire d'Electronique et Systèmes de Télécommunications	Professeur des Universités Section 63
Gilles BUREL		Professeur des Universités Section 61

Rapporteurs :

Joseph RONSIN	INSA Rennes Laboratoire « Architecture et Traitement des Images Spatiales et Temporelles »	Professeur des Universités Section 61
Christophe COLLET	Université de Starasbourg 1 Laboratoire des Sciences de l'Image, de l'Informatique et de la Téledétection	Professeur des Universités Section 61

Examineurs :

Jean-Marie NICOLAS	ENST Paris Département Traitement du Signal et des Images	Professeur
Benoît ZERR	Groupe d'études sous-marines de l'Atlantique	Chef du Service d'Acoustique Haute Fréquence
Jacques MEUDEC	Thales Underwater Systems	Directeur Affaires et Projets Guerre des Mines
Jean-Philippe MALKASSE	Thales Underwater Systems, Service Etudes Générales Sonar	Chef de groupe Traitement du Signal

2.3 Activités pédagogiques

2.3.1 Enseignements

- de 1999 à juin 2005 (130 heures éq. TD), enseignante vacataire :
 - à l'IUP GEII Télécommunications et Réseaux de Brest : 100 heures de Travaux Dirigés effectuées (Traitement du signal, Techniques mathématiques de l'ingénieur),
 - ponctuellement, en DESS Electronique à l'Université de Bretagne Occidentale (UBO) à Brest (Particularités et exploitation de l'imagerie sonar), à l'Ecole Nationale Supérieure de Techniques Avancées (ENSTA) ParisTech, en DESS Modélisation Simulation et Optimisation de l'Université de Bretagne Sud à Vannes.
- Depuis septembre 2005 (moyenne annuelle de 192 heures éq. TD), l'essentiel des enseignements (mathématiques appliquées au traitement de l'information) sont effectués à l'ENSTA Bretagne dans le cadre de :
 - formation initiale d'ingénieurs CI 1, 2 et 3,
 - formation Ingénieurs Grandes Ecoles destinée aux étudiants Saoudiens,
 - formation par alternance FIPA 1, 2 et 3,
 - formations continues en acoustique sous-marine ou traitement du signal.

Autres enseignements :

- à l'Université de Bretagne Occidentale, aux étudiants du master Spécialité Professionnelle "Electronique des Systèmes Communicants" (7h de cours de « sonar HF – application à la chasse aux mines ») jusque janvier 2017,
- à l'Ecole Navale aux élèves officiers de la voie d'approfondissement ASM (16h de traitement d'images en 2008, 2009 et 2010).

2.3.2 Responsabilités, jurys et suivi

Responsabilité à l'ENSTA Bretagne :

- matière Traitement du Signal de la Formation d'Ingénieurs Par Alternance (FIPA) 1^{ère} année depuis septembre 2007,
- matière Traitement du Signal Analogique de la formation initiale d'Ingénieurs 2^{ème} année de 2005 à 2010,
- UV de 3^{ème} année cycle ingénieur intitulée « Signal, image, vidéo, vision » depuis septembre 2012.
- profil PSO (Perception et Systèmes d'Observation) de l'option SPID (Systèmes, Perception, Information, Décision), année 2014-2015.

Membre de jury :

- de recrutement pour le cursus FIPA avec participation à la sélection des dossiers et aux auditions des candidats à cette formation (depuis l'année 2008/2009),
- FIPA de l'ENSTA Bretagne (depuis l'année 2008/2009),
- de projets de fin d'études d'ingénieurs (un ou deux par an),
- de l'épreuve d'entretien pour le concours commun des écoles des mines et maintenant concours mines-télécom (2 jours par an depuis juin 2013),
- Master 2 STS Spécialité ESCO de l'UBO (de 2009-2010 à 2016-2017).

Suivi :

- Tutorat d'élèves en cursus FIPA (un élève par an en moyenne).
- Correspondant école d'un stage de fin d'études d'ingénieurs par an.

2.4 Représentation du personnel dans différents conseils

- conseil du laboratoire E³I²-EA3876 de septembre 2008 à décembre 2011,
- conseil d'administration de l'ENSTA Bretagne de mars 2009 à octobre 2012,
- conseil de la recherche de l'ENSTA Bretagne de juillet 2014 à juin 2017.

2.5 Activités d'encadrement

2.5.1 Stagiaires (stage de 5 mois en moyenne)

De 2002 à 2004 chez Thales Underwater Systems : encadrement d'étudiants en stage de fin d'études pendant cinq mois de l'année (un par an).

Depuis septembre 2005 à l'ENSTA Bretagne :

- NGUYEN Lan Chi, étudiante en master 2 professionnel Mathématiques et informatique spécialité Ingénierie Mathématique de l'Université de Bretagne Sud, sur le sujet « Mise en œuvre d'un algorithme de pistage multi-cibles à partir d'une séquence d'images sonar » en 2008,
- HUBERT William, étudiant en master 2 ETR spécialité Signaux et Circuits de l'UBO, sur le sujet « Du modèle 3D de l'objet à son écho sur image sonar haute résolution » en 2009,
- PICARD Laurent, étudiant en master 2 ETR spécialité Signaux et Circuits de l'UBO, sur le sujet « Détection de trajectoires rectilignes de cibles entre deux eaux par transformée de Hough tridimensionnelle » en 2013 [36].
- HMISSI Mohamed Zied, étudiant en master 2 ETR spécialité Signaux et Circuits de l'UBO, sur le sujet « Suivi du fond marin à partir d'un flux d'images sonar frontal », en 2016.

2.5.2 Co-encadrement doctoral

Depuis mon arrivée à l'ENSTA Bretagne en Septembre 2005, j'ai participé à l'encadrement de six thèses dont l'une se termine actuellement et une autre est encore en cours. L'un des premiers doctorants que j'ai suivi n'a malheureusement pas soutenu malgré une bonne quantité de travail et qualité de rédaction. Le tableau de page suivante résume les informations relatives à chacune de ces thèses.

Prénom NOM du doctorant	Sujet	Ecole doctorale	Date de démarrage	Directeur de thèse	Autres co- encadrants	Taux d'encadrement	Date de soutenance	Financement	Position à l'issue du doctorat
Vincent MYERS	Le traitement, l'interprétation et l'exploitation d'images sonar à antenne synthétique obtenues à partir de trajectoires répétitives [43, 51]	SICMA	01/12/15	Benoit ZERR*	×	50%	2018	DRDC	Reste en poste au DRDC
Florian NICOLAS	Amélioration des techniques de détection de changement pour la lutte contre les mines sous-marines en situation de surveillance [38,41,42]	SICMA	01/02/15	Benoit ZERR*	×	50%	2018	ANRT	NA
Laurent PICARD	Prise en compte de l'environnement marin dans le processus de reconnaissance automatique de cibles sous-marines [37,39,40]	SICMA	01/10/13	Alexandre BAUSSARD*	Gilles LE CHENADEC*	30%	18/05/17	DGA	Post-Doc à l'ENSTA Bretagne
Ayda ELBERGUI	Amélioration des techniques de reconnaissance automatique de mines marines par analyse de l'écho à partir d'images sonar haute résolution [27,30,31]	EDSM	01/12/09	Pr Basel SOLAIMAN**	Benoît ZERR*	30%	10/12/13	DGA	Congé de maternité
Stéphane BAZEILLE	Reconnaissance automatique d'objets en vidéo sous-marine [4,9,22,46]	SMIS	01/10/05	Pr Luc JAULIN*	×	50%	01/10/08	Région	Post-Doc à l'ENSTA ParisTech
Lionel CROS (co-encadré à partir de janvier 2006)	Développement d'une architecture de traitement adapté à la Tomographie Acoustique Discrète [20,47]	SMIS	01/09/04	Pr Ali KHENCHAF*	Cédric GERVAISE*	30%	Non soutenue	Contrat	Ingénieur en SSII

* ENSTA Bretagne, ** IMT Atlantique (anciennement Telecom Bretagne)

2.6 Activités de recherche contractuelle (depuis septembre 2005)

2.6.1 Responsabilité de contrat :

Après mon départ de Thales Underwater Systems (TUS), les bonnes relations que j'ai entretenues avec mes anciens collègues du service Etudes Générales Sonar du site de Brest ont été à l'origine de quelques collaborations. Par ailleurs, l'ENSTA Bretagne a un lien privilégié avec sa tutelle en ce qui concerne l'enseignement mais également la recherche. DGA Techniques Navales Brest (ex-GESMA) nous a tout naturellement sollicités pour soutenir le développement d'algorithmes de perception pour ses AUVs démonstrateurs (AUV pour Autonomous Underwater Vehicle). Cette collaboration a commencé de manière informelle par une invitation à une campagne de quatre jours d'essais en mer en avril 2006 à bord du BEGM « Thétis » afin d'évaluer « à chaud » la qualité des données sonar d'évitement RESON 8101 pour leur démonstrateur Redermor. S'en est suivie une série de contrats de recherche sur la problématique de la détection d'obstacles pour AUV.

Contrat industriel avec Thales Underwater Systems :

« Etude d'Algorithme de Classification et de Reconnaissance Automatique (ACRA) » [61,62]

Date de démarrage: 21/12/05

Durée: 36 mois

L'objectif de l'étude était de définir une méthode de référence en ce qui concerne la classification et la reconnaissance d'objets basées sur l'extraction de contour à partir d'une image de perception sous-marine (sonar ou vidéo), et de comparer les performances de cette méthode à celle d'autres méthodes.

Projet FUI « AUV de Sécurité Maritime (ASEMAR) » avec Thales Underwater Systems :

Sous-étude « Etude de la classification multi-vues d'objets immergés à partir d'images SAS » [48,64].

Suivi d'une ingénieure de recherche : I. Leblond.

Date de démarrage: juin 2008

Durée: 18 mois

Le projet ASEMAR avait pour objectif la réalisation d'un prototype innovant d'AUV de sécurité maritime. Un des aspects était l'amélioration des performances des sonars classificateurs haute fréquence par la mise en œuvre d'une classification multi-vue sur ombre à partir d'images SAS (Synthetic Aperture Sonar). L'un des apports qui a fait

l'objet d'une publication fut la mise en œuvre de stratégies prédictives sur les vues à acquérir.

Contrat de recherche avec Groupe d'Etudes Sous-Marines de l'Atlantique (DGA/GESMA) :

« Etudes sur le traitement de données acquises par de nouvelles méthodes de perception sous-marine pour la chasse aux mines et la détection d'obstacles » [5,21,24,25,26,49,63,65,66]

Date de démarrage: 15/12/07

Durée: 36 mois

Ces études concernaient le traitement de données acquises par de nouveaux outils de perception sous-marine pour la chasse aux mines et la détection d'obstacles. L'objectif de ces nouveaux moyens de perception et du traitement de données associé était double : remplir la mission principale de chasse aux mines (à l'aide d'un SAS) et permettre au véhicule d'assurer sa survie (à l'aide d'un sonar frontal). Le traitement d'images SAS pour la reconnaissance de cibles a été traité par l'un de mes collègues J.C. Cerus afin qu'il n'y ait pas d'interaction avec les travaux que je menais encore à l'époque pour Thales. Je me suis principalement consacrée au traitement de données du sonar frontal Reson Seabat 8101 pour la détection d'obstacles dans la colonne d'eau et sur le fond.

Contrat de recherche avec DGA Techniques Navales Brest :

« Détection et suivi d'obstacles pour un AUV de REA Discret » [35,68,69,70]

Suivi d'une ingénieure de recherche : I. Karoui. Location et essai de capteurs.

Date de démarrage: 15/01/14

Durée: 36 mois

Ce second contrat financé par DGA TN visait à contribuer au Programme d'Etudes Amont « Rapid Environment Assessment Discret » proposant d'étudier les briques technologiques permettant de collecter des données d'environnement à l'aide d'un AUV (Autonomous Underwater Vehicle) sur des zones où il y a peu, voire pas de connaissance a priori. Pour les missions d'un AUV de REA Discret sur une zone inconnue, il est impératif de veiller à la sécurité du drone lui-même. Les risques sont les objets entre deux eaux, un autre AUV, un filet, les navires de surface lors d'un recalage GPS ou lors du retour en zone de récupération, les épaves et le relief sous-marin constituant les dangers les plus probables. Ce contrat a ainsi permis d'étudier le potentiel de différents sonars frontaux dans des configurations diverses répondant à des missions différentes : évitement d'obstacles dans la direction d'avance du véhicule, suivi de fond et surtout reprise de vue du véhicule. Une partie des travaux est présentée au paragraphe 3.3.

2.6.2 Participation dans divers contrats :

Projet RAPID (financement DGA) avec iXblue :

Projet MUSE pour « Multi Usage Sonar Element » [7,34,67]

Date de démarrage: 01/01/12

Durée: 24 mois

Le projet MUSE proposait d'étudier, de développer et d'expérimenter la mutualisation de fonctions principales de traitements avancés, au sein d'un élément de sonar multi usage, dont la versatilité pourrait à terme mener à plusieurs configurations permettant de répondre à un ensemble de fonctionnalités, réparties en 3 grandes familles : la navigation, la détection d'objets et/ou d'obstacles, l'imagerie bathymétrique. Ma contribution s'est portée, d'une part sur la détection d'obstacles avec des premiers travaux sur la problématique de la reprise de vue en collaboration avec une ingénieure, Imen Karoui, d'autre part sur une extension de l'algorithme de pistage de cibles dans la colonne d'eau développé pour DGA-TN et qui avait été publié [5].

Contrat de recherche avec DGA Techniques Navales Brest :

« Etudes de traitement de données 3D pour l'aide à l'identification de mines sous-marines »

Date de démarrage: 10/11/15

Durée: 36 mois

Dans le cadre d'une réflexion sur les futurs moyens de guerre des mines (voir paragraphe 4.2), DGA TN développe le concept d'AUV-I ou AUV d'identification, permettant de déterminer le type de mine en passant à l'aide d'un véhicule autonome à moins d'une dizaine de mètres de l'objet étudié. La faisabilité de l'utilisation de l'acoustique de façon complémentaire à l'optique est étudiée dans le contexte très particulier de l'imagerie des petits objets de forme complexe.

2.7 Collaborations internationales

Collaboration avec le DRDC-A, NS, Canada :

Participation à l'Arrangement Spécifique (AS) n°42 « *Systèmes de Chasse aux Mines – Amélioration de la fonction Reconnaissance Automatique de Cibles* » du comité directeur de recherche de défense au titre de l'Arrangement entre le Ministre de la défense de la République française et le Ministre de la défense nationale du Canada relatif à la coopération en matière de recherche de défense signé le 3 novembre 1993. L'objectif principal de cet AS, signé en avril 2010 pour une durée de 4 ans, était de prendre en compte les nouvelles fonctionnalités et possibilités des sonars de chasse aux mines afin d'améliorer la performance des traitements pour la reconnaissance automatique des cibles.

- De Janvier à Juillet 2011 : séjour d'étude et de recherche au centre de R&D pour la Défense Canada Atlantique d'Halifax au sein du groupe "Défense Portuaire et Antimines" de la section scientifique de la "Protection des Biens Maritimes".

Financement : bourse ERE de DGA MRIS

Thème de la recherche : Mise au point d'algorithmes de détection et de classification pour les nouveaux systèmes sonar de chasse aux mines. Plus précisément les travaux ont concerné deux axes prioritaires. Le premier consiste à détecter les éventuels changements entre deux acquisitions du même capteur séparées d'un intervalle de temps variable. Cet aspect, à l'époque peu étudié en imagerie sonar contrairement à l'imagerie radar, suscite un récent mais grand intérêt des marines en général. C'est en effet un des seuls moyens de détection de dispositifs explosifs artisanaux désignés par « IEDs : Improvised Explosive Devices ». Le second concerne le développement d'algorithmes de reconnaissance automatique de cibles par les nouveaux systèmes de chasse aux mines et en particulier le DUBM44 qui permet d'acquérir trois aspects de la cible.

Collaboration avec FFI, Norvège :

En avril 2011, invitation à une campagne d'essais en Norvège dans le cadre d'un consortium entre le Norwegian Defence Research Establishment (FFI), Defence Research and Development Canada (DRDC) et Naval Research Laboratory (NRL), US. Un accord franco-norvégien a ensuite été engagé pour une collaboration durable et l'exploitation commune des données de cette campagne.

Ateliers de travail sur l'ATR :

Invitation par l'ONR à la première (Décembre 2012, Washington) puis seconde (Mai 2015, Atlanta) éditions du workshop iSAAM (International Workshop on Intersections in Signal Processing, Acoustics & ATR for Maritime Applications). Chaque édition a comporté sur trois journées une alternance de présentations sous forme de tutoriels et de discussions en petits groupes sur des problèmes posés. Parmi ces derniers, on peut citer les limitations des approches classiques et les perspectives d'amélioration d'un point de vue traitement de la donnée mais aussi système d'acquisition, ou encore l'analyse des performances (prédiction, estimation, optimisation).

2.8 Rayonnement

Participation à des groupes de recherche :

- Avril 2006 : séminaire « Technologies d'imagerie et d'acoustique pour le milieu marin » organisé par le G2RA (Groupement Régional de Recherche en Acoustique),

- Juin 2007 : workshop « Navigation sur amers en milieu sous-marin - Vers un partage des bases de données » organisé par le GESMA,
- Juin-juillet 2007 : dans le cadre du GIS (Groupe d'Intérêt Scientifique) Europôle Mer, participation au groupe de travail autour de l'axe de recherche 5 « Système de mesures, d'observations et d'interventions »,
- Février 2008 : journée thématique « Robotique Navale & Lutte Sous-Marine » organisée par le GESMA.

Séminaires invités :

- Juin 2007, School of Engineering and Physical Science, Heriot-Watt University, Edinburgh : « ATR for Mine Warfare Perception means for AUV »,
- Décembre 2007, Laboratoire de Mathématiques et Applications des Mathématiques EA 3885, Université de Bretagne Sud, Vannes: « AUV (Redermor) Obstacle Detection and Avoidance Experimental Evaluation »,
- Décembre 2011, Laboratoire de Mathématiques et Applications des Mathématiques EA 3885, Université de Bretagne Sud, Vannes: « Détection de changement entre deux passes sonar ».

2.9 Animation scientifique (depuis septembre 2005)

Membre des comités d'organisation de colloques :

- Colloque « Caractérisation du milieu marin » dans la cadre de la semaine internationale des sciences et technologies de la mer (Sea tech Week) en octobre 2006,
- Workshop MOQESM (MONitoring Quantitatif de l'Environnement Sous-Marin) des éditions 2010, 2014 et 2016 de la Sea Tech Week,
- Colloque ICoURS'12 « International Conference on Underwater Remote Sensing » à Brest en octobre 2012 avec un focus sur la détection et la classification de cibles sous-marines à la suite des éditions canadienne et écossaise, respectivement en 2001 (CAD-CAC organisée par RDDC-A) et 2007 (DCUT organisée par l'université d'Heriot-Watt).

Montage et animation d'une nouvelle équipe de recherche (Voir également le paragraphe 4.1) : Depuis décembre 2014, participation active au montage du projet d'une nouvelle équipe (PRASYS pour « Perception, Robotics and Autonomous SYStems ») au sein du Lab-STICC sur le thème de la perception pour la robotique mobile autonome. Les évaluateurs de l'HCERES ont émis au printemps 2016 un avis favorable à la création de cette équipe. J'ai en charge la responsabilité et l'animation de cette équipe qui est officiellement née au début de l'actuel quinquennal du laboratoire, soit le 1^{er} janvier 2017.

2.10 Expertises

Évaluatrice pour la Commission Européenne :

(<http://ec.europa.eu/research/participants/portal/desktop/en/experts/index.html>)

- Novembre 2012 : inscription dans la base des experts externes de la commission européenne pour la recherche et l'innovation,
- Novembre 2017 : évaluation de deux propositions soumises suite à l'appel H2020-FETOPEN-2016-2017-RIA_27-09-2017.

Rapporteur pour des revues :

- Revue « Traitement du Signal » en 2007,
- Journal « Ocean Engineering » en 2010,
- Journal IEEE “Transactions on Geoscience and Remote Sensing” en 2014.

Rapporteur pour les conférences :

- Caractérisation du milieu marin (CMM) en 2006 (membre du comité scientifique)
- International Conference on Information & Communication Technologies: from Theory to Applications (ICTTA) en 2006,
- Advanced Concepts for Intelligent Vision Systems (ACIVS) en 2006,
- International Conference on Intelligent Robots and Systems (IROS) en 2013 et 2017.

Examinatrice de jurys de thèse :

- 01/10/08 : M. Stéphane Bazeille à l'ENSTA Bretagne sur « Reconnaissance automatique d'objets en vidéo sous-marine »,
- 10/02/12 : Mlle Haijuan Hu à l'Université de Bretagne Sud sur « Quelques contributions en débruitage des images en lien avec la variation totale, les moyennes non-locales et des représentations dans des bases adaptées »,
- 10/12/12 : Mme Ayda Elbergui à l'ENSTA Bretagne sur « Amélioration des techniques de reconnaissance automatique de mines marines par analyse de l'écho à partir d'images sonar haute résolution »,
- 18/05/17 : M. Laurent Picard à l'ENSTA Bretagne sur « Prise en compte de l'environnement marin dans le processus de reconnaissance automatique de cibles sous-marines ».

2.11 Publications

Thèse

- [1] I. Quidu, « Classification multi-vues d'un objet immergé à partir d'images sonar et de son ombre portée sur le fond », *thèse de doctorat*, Université de Bretagne Occidentale, Brest, soutenue le 11 décembre 2001.

Ouvrages

- [2] I. Quidu, V. Myers and B. Zerr, *Proceedings of the 2012 International Conference on Detection and Classification of Underwater Targets*, Cambridge Scholars Publishing, 2014.
- [3] B. Zerr, L. Jaulin, V. Creuze, N. Debese, I. Quidu et B. Clément, *Quantitative Monitoring of the Underwater Environment: Results of the International Marine Science and Technology Event MOQESM*, Springer International Publishing. *MOQESM: Monitoring Quantitatif de l'Environnement Sous-Marin*, Oct 2014, Brest, France. 6, 2016.

Dans la liste qui suit, les noms des doctorants et ingénieurs de recherche encadrés sont soulignés.

Reuves internationales avec comité de lecture

- [4] S. Bazeille, I. Quidu and L. Jaulin, « Color-based underwater object recognition using water light attenuation », *Intelligence Service Robotics*, 2012.
- [5] I. Quidu, L. Jaulin, A. Bertholom and Y. Dupas, « Robust Multitarget Tracking in Forward Looking Sonar Image Sequences using Navigational Data », *IEEE Journal of Oceanic Engineering*, Vol. 37, No 3, July 2012.
- [6] I. Quidu, “Incoherent change detection using amplitude sidescan sonar images”, *Proceedings of Meetings on Acoustics (POMA)*, 2012.
- [7] I. Karoui, I. Quidu and M. Legris, « Automatic sea surface obstacle detection and tracking in Forward Looking Sonar image sequences”, *IEEE Transactions on Geoscience and Remote Sensing (TGRS)*, Vol. 53, no 8, August 2015.

Reuves nationales avec comité de lecture

- [8] I. Quidu, J.P. Malkasse et G. Burel, « Classification de mines sous-marines à partir de l'image sonar brute : caractérisation du contour de l'ombre portée par algorithme génétique », *Revue Traitement du Signal*, Vol. 20, No 1, 2003.
- [9] S. Bazeille, I. Quidu, L. Jaulin & J.P. Malkasse, « Une méthode de pré-traitement automatique pour le débruitage des images sous-marines », *Revue Traitement du Signal*, Vol. 25, No. 1 & 2, 2008/02-04.

Congrès internationaux avec actes et comité de lecture

- [10]I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé, « Mine Classification based on raw sonar data: an approach combining Fourier Descriptors, Statistical Models and Genetic Algorithms », *Proc. OCEANS'2000 MTS/IEEE*, Providence, Rhode Island, Sept 2000.
- [11]I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé, « Mine classification using a hybrid set of descriptors », *Proc. OCEANS'2000 MTS/IEEE*, Providence, Rhode Island, Sept 2000.
- [12]I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé, « Mine classification based on a multiview characterisation », *Undersea Defence Technology (UDT) Europe 2001*, Hamburg, Germany, 26-28 june 2001.
- [13]I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé, « A 2-D Filter Specification for Sonar Image Thresholding », *Advanced Concepts for Intelligent Vision Systems (ACIVS'2001) conference*, Baden-Baden, Germany, July 30 – August 3, 2001.
- [14]I. Quidu, J.P. Malkasse, P. Vilbé and G. Burel, « Mine Classification based on a Fuzzy Characterisation », *Undersea Defence Technology (UDT) Europe 2002*, La Spezia, Italy, 18-20 june 2002.
- [15]F. Florin, F. Van Zeebroeck, I. Quidu et N. Le Bouffant, « Classification performances of Mine Hunting Sonar : Theory, practical results and operational applications », *Undersea Defence Technology (UDT) Europe 2003*, Malmö, Sweden, 24-26 june 2003.
- [16]F. Florin, I. Quidu et N. Le Bouffant, « Shadow classification performances evaluation of Mine Hunting Sonar », *Undersea Defence Technology (UDT) Asia 2003*, Singapore, 12-14 november 2003.
- [17]F. Florin, F. Fohanno, I. Quidu et J.-P. Malkasse, « Synthetic Aperture and 3D Imaging for Mine Hunting Sonar », *Undersea Defence Technology (UDT) Europe 2004*, Nice, 22-25 juin 2004.
- [18]I. Quidu, N. Burlet, J.-P. Malkasse et F. Florin, « Automatic classification for MCM systems », *IEEE Conf. Oceans'05 Europe*, Brest, 20-23 juin 2005.
- [19]F. Florin, I. Quidu, J.-P. Malkasse and N. Le Bouffant, « From UUV to AUV: the SCM challenge », *Undersea Defence Technology (UDT) Europe 2005*, Amsterdam, juin 2005.
- [20]L. Cros, C. Gervaise and I. Quidu, « A signal synthesis procedure designed for discreet ocean acoustic tomography », *32nd International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Hawaii Convention Center in Honolulu, April 15 - 20, 2007.

- [21] I. Quidu, A. Hétet, Y. Dupas and S. Lefèvre, « AUV (REDERMOR) obstacle detection and avoidance experimental evaluation », *IEEE OCEANS'07 conference*, Aberdeen (Scotland), 18-21 June 2007.
- [22] S. Bazeille, I. Quidu, & L. Jaulin, "Identification of underwater man-made object using a colour criterion", *International Conference on Detection and Classification of Underwater Targets*, Edinburgh, UK, 18-19 September 2007.
- [23] A. Martin & I. Quidu, "Decision Support with Belief Functions Theory for Seabed Characterization", *Information Fusion*, Cologne, Germany, 30 June-3 July 2008.
- [24] I. Quidu, Y. Dupas & L. Jaulin, "Theoretical optical flow for target position prediction on FLS images", *Underwater Acoustic Measurement (UAM'09)*, Nafplion, Greece, 21-26 June 2009.
- [25] I. Quidu, A. Bertholom and Y. Dupas, "Ground obstacle tracking on forward looking sonar images", *European Conference on Underwater Acoustics (ECUA) 2010*, Istanbul, Turkey, July 2010.
- [26] I. Quidu, A. Bertholom and Y. Dupas « Deferred estimation of vertical position of a floating obstacle by minimising defects of tracking », *invited paper, Underwater Acoustic Measurement (UAM'11)*, Kos Island, Greece, June 2011.
- [27] A. Elbergui, I. Quidu and B. Zerr, "Model Based Target Classification for High Frequency – High Resolution Imaging Sonar", *Underwater Acoustic Measurement (UAM'11)*, Kos Island, Greece, June 2011.
- [28] T. O. Sæbø, Ø. Mitdgaard, V. Myers, J. R. Dubberley and I. Quidu, "Change Detection Using Synthetic Aperture Sonar: Preliminary Results from the Larvik Trial", *OCEANS'2011 MTS/IEEE*, Kona, Hawaiï, Sept 19-22, 2011.
- [29] I. Quidu, "Incoherent change detection using amplitude sidescan sonar images", *Proceedings of the 11th European Conference on Underwater Acoustics*, Edinburgh, Scotland, UK, 2-6 July 2012.
- [30] A. Elbergui, I. Quidu, B. Zerr and B. Solaiman, "Model Based Classification of Mine-Like Objects in Sidescan Sonar using the Highlight Information", *Proceedings of the 11th European Conference on Underwater Acoustics*, Edinburgh, Scotland, UK, 2-6 July 2012.
- [31] A. Elbergui, I. Quidu, B. Zerr et B. Solaiman, « Amélioration d'une méthode de classification des mines sous-marines basée sur l'écho acoustique dans des images sonar latéral », *Proceedings of TAIMA'13*, Hammamet, Tunisie, 13-18 mai 2013.
- [32] V. Myers, I. Quidu, T. Sæbø, Ø. Mitdgaard and R. Hansen, "Results and analysis of coherent change detection experiments using repeat-pass synthetic aperture sonar images", *1st Underwater Acoustics Conference*, Island of Corfu, Greece, Sunday 23rd through Friday 28th June, 2013.

- [33] I. Quidu, V. Myers, Ø. Midtgaard and R. E. Hansen, “Subpixel image registration for coherent change detection between two high resolution sonar passes”, in *Proc. of the International Conference on Underwater Remote Sensing (ICoURS) - Detection and Classification of Underwater Targets, Brest, France, 8-11 October 2012*, Cambridge Scholars publishing, 2014.
- [34] I. Karoui, I. Quidu and M. Legris, « Acoustic obstacle detection for safe AUV surfacing », *2nd Underwater Acoustics Conference*, Island of Rhodes, Greece, Sunday 22nd through Friday 27th June, 2014.
- [35] I. Karoui, I. Quidu and M. Legris, « Automatic spatial clustering and tracking of sea surface obstacles in Forward Looking Sonar sequences », *3rd Underwater Acoustics Conference*, Greek island of Crete, 21st – 26th June, 2015.
- [36] L. Picard, I. Quidu, M. Legris, “Rectilinear paths detection of target in the water column by using a three-dimensional Hough transform of forward looking sonar data”, *3rd Underwater Acoustics Conference*, Greek island of Crete, 21st – 26th June, 2015.
- [37] L. Picard, A. Baussard, G. Le Chenadec, and I. Quidu, “Potential of the intrinsic dimensionality for characterizing the seabed in the ATR context,” in *IEEE/OES & MTS OCEANS conference*, Genoa, Italy, May 2015.
- [38] F. Nicolas, A. Arnold-Bos, I. Quidu and B. Zerr, “Rigid Sonar Tracks Registration For MCM Survey Missions”, *proceedings of Undersea Defence Technology conference*, Oslo, Norway, June 1-3, 2016.
- [39] L. Picard, A. Baussard, G. Le Chenadec, and I. Quidu, “Seafloor characterization for ATR applications using the monogenic signal and the intrinsic dimensionality”, in *IEEE/OES & MTS OCEANS conference*, Monterey, CA, USA, September 2016.
- [40] L. Picard, A. Baussard, G. Le Chenadec, and I. Quidu, “Detection of MLO in sand ripple seafloor using the monogenic signal and intrinsic dimensionality”, in *IEEE/OES & MTS OCEANS conference*, Monterey, CA, USA, September 2016.
- [41] F. Nicolas, A. Arnold-Bos, I. Quidu and B. Zerr, “Fast Fourier-Based Block-Matching Algorithm for Sonar Tracks Registration in a Multi-Resolution Framework”, in *MOQESM’16 conference*, Brest, October 2016.
- [42] F. Nicolas, A. Arnold-Bos, I. Quidu, B. Zerr, "Markov-based approaches for ternary change detection between two high resolution synthetic aperture sonar tracks", *MTS/IEEE OCEANS Conference*, Aberdeen, United Kingdom, 19-22 June 2017.
- [43] V. Myers, J. Groen, H. Schmaljohann, I. Quidu and B. Zerr, “Multi-look processing for Coherent Change Detection with Synthetic Aperture Sonar”, *Underwater Acoustics Conference*, Skiathos, Greece, 2017.

Congrès nationaux avec actes et comité de lecture

- [44] I. Quidu, J.P. Malkasse, P. Vilbé and G. Burel, « Fusion multi-attributs d'images sonar », *colloque GRETSI'01*, Toulouse, 10-13 septembre 2001.
- [45] A. Hétet, I. Quidu and Y. Dupas, « Obstacle Detection and Avoidance for AUV: problem analysis and first results (Redermor) », *CMM'06 – Sea Tech Week 2006*, Brest, 16-19/10/06.
- [46] S. Bazeille, I. Quidu, L. Jaulin, « Automatic underwater image pre-processing », *CMM'06 – Sea Tech Week 2006*, Brest, 16-19/10/06.
- [47] L. Cros, C. Gervaise, I. Quidu, L. Jaulin and Y. Stephan « A new concept of ocean acoustic tomography », *CMM'06 – Sea Tech Week 2006*, Brest, 16-19/10/06.
- [48] I. Leblond, I. Quidu et C. Berron, « Mise en oeuvre de stratégies prédictives sur les vues à acquérir pour la classification multi-vues d'objets immergés à partir d'images SAS », *17ème conférence en Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, 19-22 janvier 2010, Caen, France.
- [49] L. Picard, A. Baussard, G. Le Chenadec, I. Quidu, "Détection de cibles sous-marines dans des champs de rides de sable par imagerie sonar", 26e GRETSI, Juan-les-Pins, France, Sep 2017.

Congrès internationaux sans actes

- [50] I. Quidu and Y. Dupas, "Forward Looking Techniques for environment modelling, obstacle detection and characterization", *Acoustics'08 conference*, Paris, June 29-July 4, 2008.
- [51] V. Myers, I. Quidu, T.O. Sæbø, R.E. Hansen and B. Zerr, "Computing navigation corrections for co-registration of repeat-pass synthetic aperture sonar images", *Acoustics'17 Boston, Journal of the Acoustical Society of America*, 141:5, 4042-4042, 2017.

Rapports techniques

- [52] I. Quidu, « Reconnaissance de formes en imagerie sonar », rapport TMS SAS 99/B/EGS/NC/087-IQ, juillet 1999, 32 pages.
- [53] I. Quidu, « Classification multi-vues d'un objet immergé à partir de son ombre portée sur le fond », rapport TMS SAS 00/B/EGS/RG/06-IQ, janvier 2000, 23 pages.
- [54] I. Quidu et J.P. Malkasse, « Classification automatique des ombres. Rapport d'étude – PEA sonar multi-cibles en guerre des mines », rapport TMS SAS 00/B/EGS/DR/104-IQ-JPM, septembre 2000, 105 pages.
- [55] I. Quidu, « Identification 3D. Rapport d'étude – PEA sonar multi-cibles en guerre des mines », rapport TUS 02/B/EGS/DR/043-IQ, avril 2002, 79 pages.

- [56] I. Quidu, « Fonction détection de transitoires – Document de validation fonctionnelle sur démonstrateur - PEA Aéro phase 3 », DVF No 61 373 267 279, PEA Aéro Phase 3, contract number A 00 77 167 00 470 53 75, février 2003, 52 pages.
- [57] I. Quidu, « Aide Opérateur », rapport d'étude auto-financée TUS 02/B/EGS/DR/164-IQ, avril 2003, 36 pages.
- [58] I. Quidu et N. Le Bouffant, « Rapport final de la phase technique de l'étude de probabilité de classification - Etude d'un sonar multi-cibles pour guerre des mines », contrat DGA no A 99 77 120 00 470 75 53 », rapport TUS 03/B/EGS/DR/043-IQ-NLB, mars 2003, 109 pages.
- [59] L. Schwab, A. Lemer, I. Quidu, « SFR&D Support Vector Machines », rapport d'étude auto-financée, TUS SAS 04/S/EGS/DR/498-LS du 10 décembre 2004, 72 pages.
- [60] I. Quidu, « Document de définition d'algorithmes : CAD/CAC pour sonar latéral – SFR&D MCM Classification », rapport d'étude auto-financée, n°61472219-1 du 24 janvier 2005, 81 pages.
- [61] I. Quidu, « Etude d'algorithmes de classification et de reconnaissance automatique (ACRA) - Poste 1 : Etude théorique », *contrat TUS SAS 05/S/DA/275/DS*, 19/12/06, 54 pages.
- [62] I. Quidu, « Etude d'algorithmes de classification et de reconnaissance automatique (ACRA) - Poste 2 : Application au sonar », *contrat TUS SAS 05/S/DA/275/DS*, 21/12/07, 58 pages.
- [63] I. Quidu, « Etudes sur le traitement de données acquises par de nouvelles méthodes de perception sous-marine pour la chasse aux mines et la détection d'obstacles – rapport intermédiaire », *CONTRAT GESMA – ENSIETA*, décembre 2008, 43 pages.
- [64] I. Leblond, C. Berron et I. Quidu, « Projet ASEMAR - Lot 6 'Etudes Technologiques' - Tâche 6.1 : 'Etude de la classification multi-vues d'objets immergés à partir d'images SAS' – rapport d'avancement », février 2009, 94 pages.
- [65] I. Quidu et J.C. Cexus, « Etudes sur le traitement de données acquises par de nouvelles méthodes de perception sous-marine pour la chasse aux mines et la détection d'obstacles – rapport intermédiaire », *CONTRAT GESMA – ENSIETA*, Février 2010, 141 pages.
- [66] I. Quidu, « Etudes sur le traitement de données acquises par de nouvelles méthodes de perception sous-marine pour la chasse aux mines et la détection d'obstacles – rapport final », *CONTRAT GESMA – ENSIETA*, Décembre 2010, 31 pages.
- [67] M. Legris, S. Penneç, N. Cusseau, I. Quidu et I. Karoui, « Rapport d'essais – Illustration des algorithmes & traitements », *MUSE-05-01*, 15/02/2014, 199 pages.
- [68] I. Quidu et I. Karoui, « Etudes relatives à l'évitement d'obstacles pour AUV dans le cadre du PEA REA Discret – rapport intermédiaire », 12/03/2015, 62 pages.

- [69] I. Karoui et I. Quidu, « Etudes relatives à l'évitement d'obstacles pour AUV dans le cadre du PEA REA Discret – rapport intermédiaire 2 », 15/03/16, 30 pages.
- [70] I. Quidu, « Etudes relatives à l'évitement d'obstacles pour AUV dans le cadre du PEA REA Discret – rapport final », 06/02/17, 54 pages.

3 Synthèse des travaux

3.1 Présentation de l'imagerie sonar haute fréquence et de son usage militaire

L'imagerie sonar dite 'haute fréquence' (sonar HF dans la suite) ou 'haute résolution' désigne les systèmes de fréquence centrale d'utilisation allant de 100Khz à 900Khz environ. A ces fréquences, l'onde sonore ne pénètre pas dans le sédiment et interdit de fait la détection d'objets enfouis. Les présents travaux se limitent donc aux objets posés sur le fond, flottant entre deux eaux ou encore à la surface de mer et se déclinent suivant deux fonctions bien différentes : la chasse aux mines et l'évitement d'obstacles.

En chasse aux mines, la contre-mesure marine (ou Mine Counter-Measure MCM) désigne un ensemble de tâches depuis la détection (à plusieurs centaines de mètres) jusqu'à la neutralisation des mines ou dispositifs explosifs artisanaux que nous regroupons sous le terme générique de cibles. La première partie du processus dédiée à la reconnaissance des cibles est encore largement menée depuis des navires de surface appelés chasseurs de mines qui sont aujourd'hui concurrencés par de *petits* drones sous-marins, plus maniables et surtout sécurisants pour les marins. Pour pouvoir visualiser rapidement une large zone du fond marin, le système privilégié est le sonar latéral, en mode conventionnel (SideScan Sonar SSS) ou synthétique (Synthetic Aperture Sonar SAS), permettant de balayer le fond des deux côtés (un sonar latéral sur chaque flanc) d'un *poisson* remorqué (par exemple le T-SAS de Thales remorqué par les bâtiments remorqueurs de sonars BRS de la Marine Française depuis 2007) ou plus récemment d'un véhicule sous-marin autonome (Autonomous Underwater Vehicle AUV). Toutefois, c'est le sonar sectoriel, de coque ou sur véhicule téléopéré (Remotely Operated Vehicle ROV) comme le Sonar Propulsé à Immersion Variable (SPIV) de Thales, qui fut utilisé par les chasseurs de mines tripartites CMT (de l'alliance tripartite Française, Belge et Hollandaise), refondus en France au début des années 2000, dont le renouvellement est prévu à partir de 2020 (cf. programme SLAMF évoqué dans le projet de recherche au paragraphe 4.2) [PUY00]. La reconnaissance automatique de cible sur le fond marin à partir de ces systèmes d'imagerie fait l'objet du paragraphe suivant et se décline en deux axes liés à deux missions différentes : la chasse aux mines d'une part et la surveillance d'autre part.

La détection d'obstacles sur le trajet d'un AUV vise à prévenir toute collision qui pourrait endommager le véhicule et constitue donc la première étape d'un processus plus général d'évitement d'obstacles. Le système privilégié cette fois est le sonar frontal sectoriel, dit Obstacle Avoidance System OAS, avec différents dépointages possibles

(tilts) menant à trois configurations pour des missions différentes : Forward Looking Sonar FLS vers l'avant pour la détection des obstacles dans la colonne d'eau, Downward Looking Sonar DLS vers le bas pour le suivi de fond et enfin, Upward Looking Sonar ULS vers le haut pour la reprise de vue du véhicule. Nous reviendrons sur ces trois fonctions dans le paragraphe 3.3.

3.2 Reconnaissance automatique de cible sur le fond marin

La reconnaissance automatique de cible (Automatic Target Recognition ATR) comprend à la fois la détection d'un objet potentiellement dangereux sur des critères assez simples (persistance d'un écho à grande distance par exemple) et sa classification en MILCO/non MILCO (voir Figure 1) où MILCO, pour MIne Like COntact, désigne la présence d'une signature acoustique susceptible d'être celle d'une mine (de par la silhouette de l'ombre portée le plus souvent). La reconnaissance peut aussi comprendre une phase supplémentaire d'identification, étape visant à confirmer si le MILCO est bien une mine (dans ce cas, le type peut être déterminé à ce stade). Le diagramme de la Figure 1 est générique et l'ensemble des tâches se décline différemment selon les moyens mis en œuvre. Les opérateurs des chasseurs de mines tripartites par exemple appliquent une procédure séquentielle performante à l'aide de différents systèmes sonar parmi lesquels un sonar frontal de coque (dit HMS pour Hull Mounted Sonar) disposant d'un premier mode destiné à la détection à grande distance, le mode sonar détecteur, et un second mode de fréquence plus élevée destiné à la classification, le mode sonar classificateur à plus courte distance. D'une manière très différente, un poisson remorqué ou un AUV dispose d'un sonar latéral pour collecter un jeu de données utilisé à la fois en détection et en classification. Dans le cas du HMS en mode détecteur, seul l'écho est exploitable et des détecteurs, populaires en radar, comme le CFAR sont adaptés. Sur une image sonar latéral en revanche, écho et surtout ombre acoustique des objets sont visibles. Le filtrage adapté à cette paire écho-ombre est un des algorithmes de détection les plus connus et largement utilisé en sonar [DOB97]. Pour notre part, nous proposerons dans le paragraphe 3.2.1.1, une approche statistique. Dans les deux cas cependant, l'algorithme de classification traite une imagerie sonar centrée sur la position de l'objet détecté.

L'approche classique de contre-mesure marine (ou Mine Counter-Measure MCM) en amont d'un bâtiment sensible (sous-marin par exemple) pour garantir une navigation en toute sécurité incombe classiquement aux chasseurs de mines. Cette approche a des limites en cas d'ambiguïté qui nécessiterait un déroutage du navire pour acquérir un second point vue sous un angle différent. Cette manœuvre est bien plus aisée avec un porteur plus léger comme un ROV (le SPIV par exemple) ou un AUV. Ceci nous amène à l'approche multi-vue, objet principal de mes travaux de thèse [QUI01a], par ailleurs

largement étudiée d'un point de vue algorithmique [ZER09] [FAW10] [LER11] [MYE12] et plus marginalement d'un point de vue système [COO01] [CHA14].

Les deux stratégies de chasse aux mines ci-dessus, destinées à l'exploration de zones inconnues donc mono-passe mais mono- ou multi-vue, sont abordées au paragraphe 3.2.1. Le cas un peu différent de détection d'objet potentiellement dangereux, mine ou plus particulièrement IED, apparu entre deux passes sonar est destiné à des missions de surveillance ou *monitoring*. Cette problématique abordée au paragraphe 3.2.2 a, depuis 2010, suscité beaucoup d'intérêt. Le principe général est de détecter des changements entre les données acquises selon deux trajectoires similaires dites répétitives ce qui nécessite en amont un recalage précis des données.

3.2.1 Chasse aux mines

A bord des chasseurs de mines, l'ATR se déroule classiquement selon une approche séquentielle : détection, segmentation (binaire fond-ombre ou ternaire en ajoutant la classe écho), extraction de descripteurs puis classification supervisée. Supposons que l'on se place dans un contexte opérationnel connu et maîtrisé, cette stratégie a deux principales faiblesses que nous avons pointées dès les années 2000 [QUI01a] : le risque de propagation des erreurs sans rétroaction possible et la nécessité de disposer d'une large base de données capable de couvrir l'ensemble des types de mines possibles et leurs signatures acoustiques sur un fond marin variable. Nous reviendrons sur ces faiblesses et une alternative possible dans le paragraphe suivant.

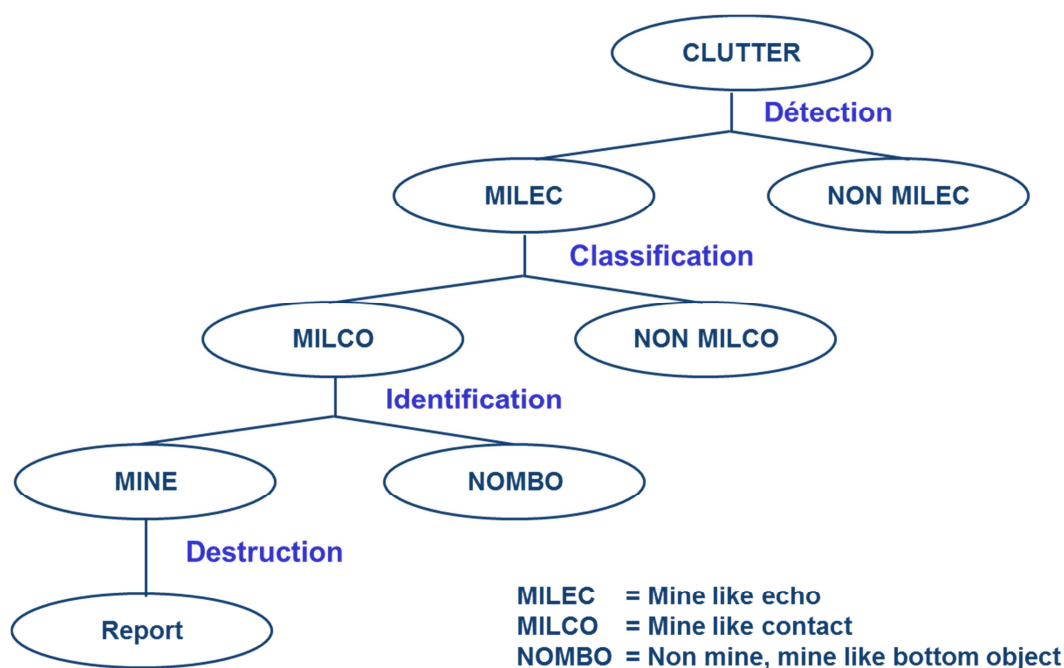


Figure 1 – Tâches successives de chasse aux mines sous-marines

La détection en MILEC/NON MILEC porte sur le niveau de l'écho visible sur une portée de plusieurs centaines de mètres et sa persistance (éventuellement appréciée à l'aide d'un échogramme représentant l'historique des échos). La classification en MILCO/non MILCO exploite les données d'un sonar dit classificateur d'une portée de l'ordre d'une centaine de mètres, du fait d'une fréquence plus élevée que le sonar détecteur. Grâce à une résolution latérale de quelques dizaines de centimètres, le sonar classificateur (et d'autant plus après traitement antenne synthétique) met en évidence l'ombre portée de l'objet ciblé. Le processus ATR développé avec mes collègues de Thales Underwater Systems présenté en 2005 exploite ces deux modes sonar [QUI05]. Cependant avec l'émergence de drones sous-marins, capables de s'approcher des objets tout en laissant le bâtiment de surface à distance du danger, le sonar latéral a tendance à s'imposer en guerre des mines.

Avec un minimum de trois voies sur l'objet, la classification a été principalement faite 'sur ombre', c'est-à-dire à partir de l'ombre acoustique de l'objet [JAN87] [THO96] [MIG98] [QUI01a]. Néanmoins, l'écho peut être utilisé conjointement à l'ombre pour lever les ambiguïtés ou apporter des informations complémentaires [ZER96] [DAN98] [MYE02] [REE04a]. Les travaux de notre doctorante, Ayda Elbergui, dirigée par Basel Solaiman (IMT Atlantique) et également co-encadrée par mon collègue Benoit Zerr, ont montré la pertinence de l'écho, et en particulier sa version synthétique basée sur l'A-scan¹, dans un processus de classification supervisée, y compris pour des résolutions assez faibles [ELB12b]. La « qualité photographique » des images haute résolution (5cm et moins), obtenues par traitement antenne synthétique, pose plus récemment la question de l'exploitation de la forme 2D de l'écho, tout comme cela se fait depuis des décennies avec la forme de l'ombre acoustique, pour la classification des cibles [PAI10].

Jusqu'aux années 2010, l'automatisation de la reconnaissance de cibles sous-marines à partir d'images sonar HF passe très souvent par une étape préliminaire comprenant un ensemble de traitements visant à préparer les données aux traitements ultérieurs comme une segmentation tout en préservant au mieux la signature des cibles [PER00] [FAW12]. Ces pré-traitements consistent le plus souvent à corriger l'atténuation en distance (même si cette fonction est en partie intégrée dans le capteur), filtrer le bruit multiplicatif de *speckle* et à lisser l'arrière-plan, allant parfois même jusqu'à supprimer les rides de sable, connues pour poser problème à nombre d'algorithmes de segmentation [DOB10] [NEL12]. La signature des cibles est alors mieux contrastée, leur détection et segmentation facilitées [QUI05]. D'autres artéfacts d'acquisition comme la trace de réflexion de la surface en petit fond (liée aux lobes secondaires et donc identifiables), la présence de

¹ L'A-scan (pour Amplitude-scan) est la quantité d'énergie renvoyée au cours du temps par l'objet ciblé.

banc de poissons ou encore de thermocline sont des phénomènes connus [CHA08] (voir Figure 2). Les problèmes de couverture incomplète ou d'illumination liés aux mouvements du porteur sont enfin abordables, ou bien du point de vue image seul ou mieux, par prise en compte des données de navigation (voir [DOB05] et paragraphe 3.2.1.1). Outre ces prétraitements destinés à faciliter la détection de cibles potentiellement dangereuses, est également pratiquée lors de la classification une normalisation géométrique des images centrées sur une cible détectée pour rendre la classification plus robuste. La normalisation en rasance ainsi proposée dans mes travaux de thèse visait à pallier le manque de données d'entraînement des algorithmes de classification et limiter la dispersion intra-classe des différents types de mines (sphériques, cylindriques, tronconiques) [QUI01a]. La normalisation de toute l'image n'est malheureusement pas satisfaisante d'un point de vue physique car elle est une inexacte approximation de la cible vue sous une rasance différente. Ce phénomène est en particulier visible pour la mine tronconique Manta ou la mine à géométrie complexe Rockan, de l'ordre du mètre à la base et de hauteur inférieure à 40cm, dont la proportion d'ombre 'propre', correspondant à la zone de fond occultée par ces mines, est non négligeable (voir Figure 3).

Toutes ces tentatives de préparation des données pour un traitement automatisé et systématique visant un traitement universel ont malheureusement démontré une erreur d'appréciation en passe d'être corrigée. A cela deux causes principales : la grande variété des environnements possibles (point pénalisant surtout la détection) et le manque de données réelles pour évaluer les algorithmes. De nombreux outils de simulation, plus ou moins réalistes, ont été développés pour fournir des données sans avoir à organiser de lourdes campagnes d'acquisition en mer mais le problème de l'exhaustivité des fonds conjugué à la généralisation des algorithmes reste posé [PET06] [GUE07]. Une voie intermédiaire proposée pour la classification (c'est-à-dire après détection) de mines consiste à utiliser la simulation pour comparer les données issues de l'image de l'objet détecté à une image simulée dans les mêmes conditions (système, géométrie d'acquisition) : à l'instar de ce qui avait été fait pour les ombres par S. Reed, A. Elbergui, dont on a évoqué plus haut les travaux, l'a fait pour les échos [REE04b] [ELB12a]. Une autre limite des approches classiques ATR est leur nature séquentielle qui présente le risque de propager une erreur faite à un point donné dans le reste de la chaîne sans supervision ou rebouclage possible. Cette limite et le problème précédent concernant la non exhaustivité des bases d'apprentissage ont été pointés lors de mes travaux de thèse en ce qui concerne les étapes ultérieures à la détection. Une alternative a été proposée avec le développement d'un algorithme génétique de segmentation orientée classification [QUI00, QUI03]. En se concentrant sur l'objectif ultime de la chaîne de reconnaissance, une segmentation discriminante du contour est effectuée à partir de

l'imagette sonar brute. Le résultat est double : tout en caractérisant le contour de l'ombre à l'aide de descripteurs de Fourier, on offre la possibilité de classifier l'objet détecté à partir d'une population réduite d'individus de départ. *L'annexe 1 (page 115) est une reproduction de l'article paru dans la revue Traitement du Signal qui donne les détails algorithmiques et une évaluation de la méthode [QUI03].*

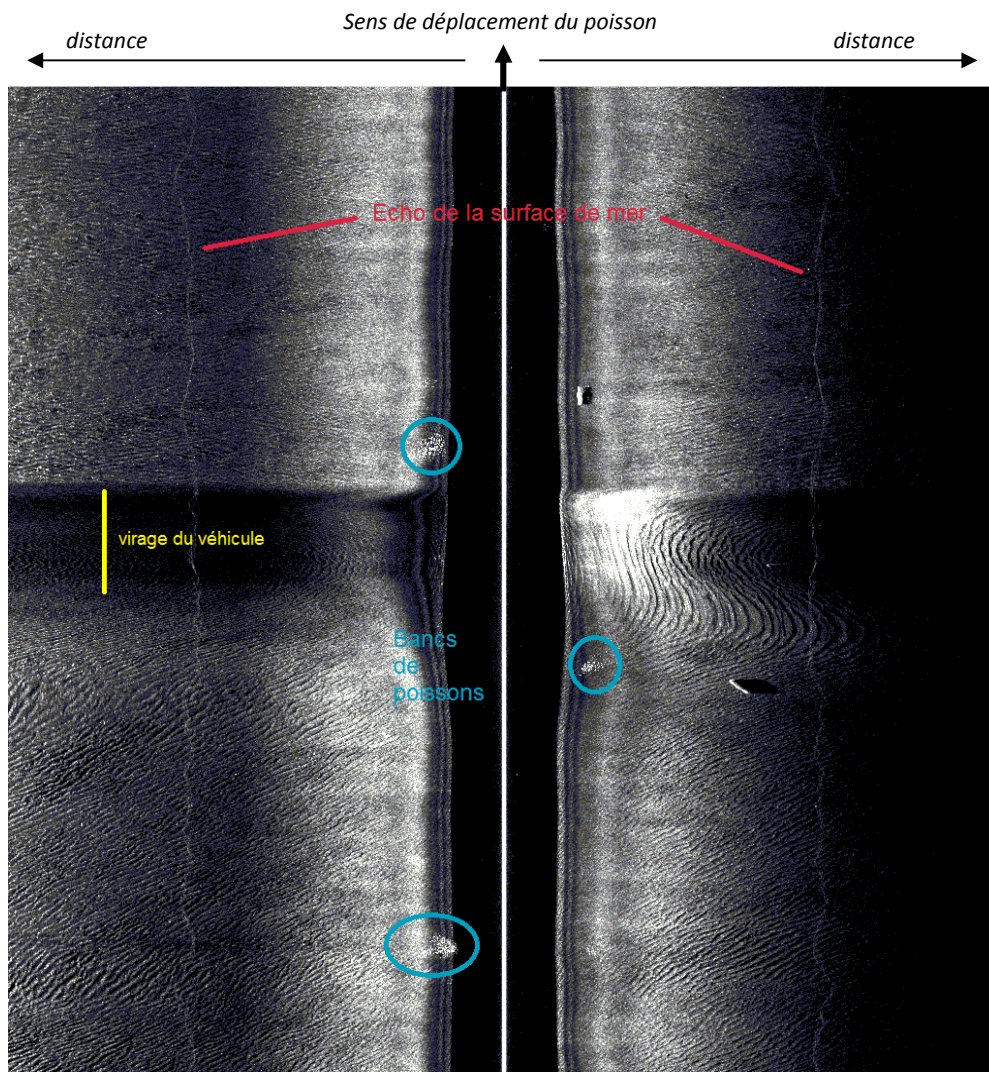


Figure 2 – Image sonar latéral (bâbord et tribord) à 900kHz de Marine Sonic Technology Ltd avec différents artéfacts d’acquisition (source : ATR Challenge 2013)

La classification mono-vue souffre d’ambiguïtés inhérentes à la géométrie des cibles recherchées et à la projection perspective des ombres acoustiques. La nature du fond, comme les rides de sable, peut également impacter les performances du processus ATR. Très tôt est donc née l’idée d’acquérir plusieurs points de vue de l’objet. S’il est envisageable de tourner autour de la cible (cas du ROV SPIV par exemple), les différentes pistes théoriques de combinaison des informations présentées dans mon manuscrit de thèse démontrent un réel bénéfice à partir de seulement 3 points de vue

[QUI01a]. Plus récemment sont apparus des systèmes multi-aspects (systèmes développés pas THALES : DUBM44 suivi de SAMDIS) permettant l'acquisition latérale de trois points de vue sur l'objet suivant une trajectoire rectiligne. THALES n'a, à ma connaissance, pas publié d'études sur l'exploitation de ces différentes vues pour améliorer la performance de classification des cibles. La courte étude théorique que j'ai pu mener sur ce sujet en 2011 lors de mon séjour d'étude au Canada sera résumée au paragraphe 3.2.1.3.

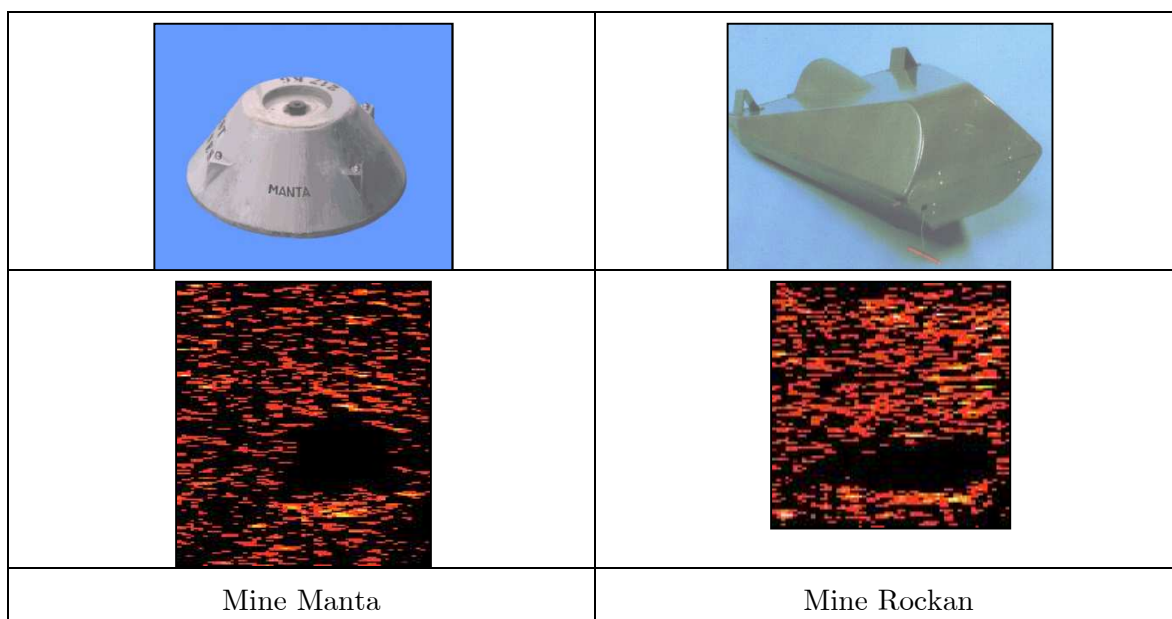


Figure 3 – Photos et images sonar des mines furtives Manta et Rockan (source : Thales)

Avant cela, en 3.2.1.1, une chaîne de détection d'objets potentiellement dangereux que nous avons proposée en 2013 lors du Challenge ATR organisé par le DSTO australien illustrera une approche classique de la détection de cibles sur le fond. Comme pour la chaîne développée en 2005 mais propriété de Thales Underwater Systems (voir [QUI05]), les traitements sont développés pour un capteur donné et sans prise en compte particulière de l'environnement. Nous présenterons ensuite, au 3.2.1.2, une approche en rupture qui s'inspire des réflexions de la communauté ATR des années 2010 et qui remet l'environnement au cœur du processus ATR.

La classification de cible est de fait peu discutée dans ce manuscrit. D'une part, parce qu'elle faisait l'objet de mes travaux de thèse en mono et multi-vue ; d'autre part, parce qu'il me semble que l'effort doit être porté 1) sur la phase de détection impossible à généraliser quel que soit le contexte opérationnel et 2) sur les perspectives qu'offre le portage des algorithmes sur véhicules sous-marins.

3.2.1.1 Un exemple de chaîne classique de détection de cibles sur images sonar latéral

3.2.1.1.1 Contexte

En mai 2013, l'organisme de recherche de la défense Australienne, the Australian Defence Science & Technology Organisation (DSTO), a lancé une compétition dit Sonar ATR Challenge dans le but de faire un point sur l'état de l'art de la reconnaissance automatique de cibles sous-marines à partir de données sonar latéral. Le DSTO et la Marine Royale Australienne ont ainsi collecté près de 11000 images à partir du sonar latéral de Marine Sonic Technology Ltd à la fréquence centrale de 900kHz porté par l'AUV REMUS 100 et ce, dans diverses régions autour de l'Australie ainsi qu'au nord des Etats-Unis côté Atlantique. Leurs spécialistes de guerre des mines avaient inspecté visuellement l'ensemble des données et pointés les MILCOs. L'enjeu du challenge consistait alors à développer et tester en moins d'un mois une chaîne ATR adaptée à ces données à partir des 654 images d'entraînement avant de la valider à l'aveugle sur les 11202 images de test. Les positions et cap du véhicule étaient également disponibles afin notamment de détecter les virages sur lesquels les données sonar sont inexploitable.

A l'issue du challenge, le DSTO donna quelques indices sur la manière dont leurs officiers de marine avaient distingué un MILCO d'un autre objet. Leur appréciation se basait tout d'abord sur leur expérience dans leur environnement local et l'allure, en matière de taille et forme de l'écho et de l'ombre, des cibles connues sur ces fonds familiers. Par ailleurs, le DSTO a noté que, sur un fond encombré, ils avaient été moins enclins à labelliser un objet comme MILCO que si ce même objet avait été sur un fond sableux que l'on pourrait qualifier de « facile ». Ceci montre toute la difficulté de l'automatisation d'une tâche auparavant faite par l'homme et la prise en compte inconsciente de l'environnement proche de la cible.

Nous avons relevé le défi et proposé une chaîne fondée sur notre propre expérience. Outre la possibilité de nous mesurer à d'autres équipes, le challenge nous donnait l'opportunité de disposer d'un grand nombre de données que nos doctorants pourraient traiter, dont une partie (les données d'entraînement) avec la vérité terrain, c'est-à-dire la position de cibles potentielles repérées par des experts.

3.2.1.1.2 Description de l'algorithme

La première étape de notre algorithme consiste à détecter automatiquement les zones de l'image à ignorer. Il s'agit d'écarter tout d'abord les pixels sombres de la colonne d'eau, donc en tout début de portée, liés au temps de trajet le plus court de l'onde entre l'émetteur et le fond. La limite de colonne d'eau est obtenue en détectant un saut d'intensité dans le profil moyen en distance (en supposant l'altitude du porteur à peu

près constante sur la zone parcourue appelée trace). Une marge de 20% est ajoutée afin de négliger également l'emprunte des lobes secondaires sur les premiers mètres. Il s'agit également d'étudier la trajectoire du véhicule et de détecter les récurrences sonar, communément appelées « pings », correspondant aux acquisitions au cours d'un virage, là aussi avec une marge fixée empiriquement à 50 pings. Un exemple de détection de ces zones à ignorer est donné Figure 4 sur une image de la base de données dont la trajectoire correspondante de l'AUV est illustrée Figure 5.

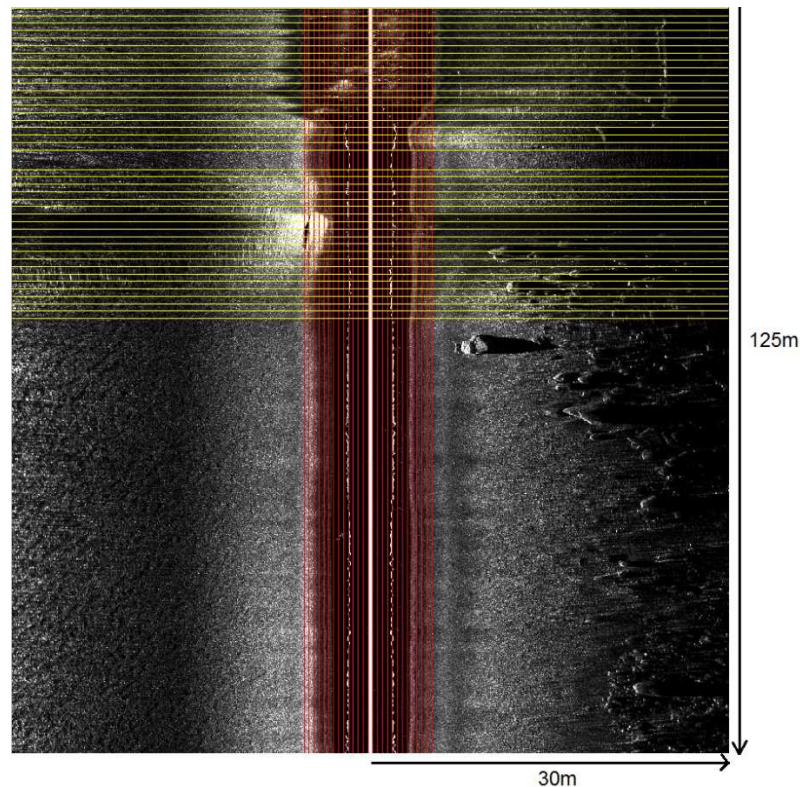


Figure 4 – Détection des zones à ignorer sur image SSS de dimension 1000 lignes par 1024 colonnes : hachuré rouge, la colonne d'eau ; hachuré jaune, deux virages.

En ce qui concerne la détection des objets posés sur le fond, les histogrammes d'un échantillon représentatif d'images ont fait apparaître des niveaux de gris de pixels quantifiés sur 8 bits et distribués selon une loi de Rayleigh, ce qui permettait d'envisager une approche statistique. Pour les zones exploitables de l'image, nous avons donc choisi de développer un détecteur d'anomalies statistiques sous l'hypothèse simple que la présence d'un objet perturbe localement la statistique des niveaux de gris des pixels de l'image. Le paramètre statistique calculé en chaque pixel, sur un voisinage défini par une fenêtre de taille équivalente à 1m en distance et 1m en azimuth, est le coefficient de variation, noté c_v , qui est, par définition, obtenu par le rapport entre l'écart-type des pixels et leur moyenne. Tout en garantissant un nombre de pixels suffisants pour une estimation correcte du paramètre statistique, cette taille de fenêtre est bien adaptée à la dimension des anomalies recherchées, c'est-à-dire des mines marines.

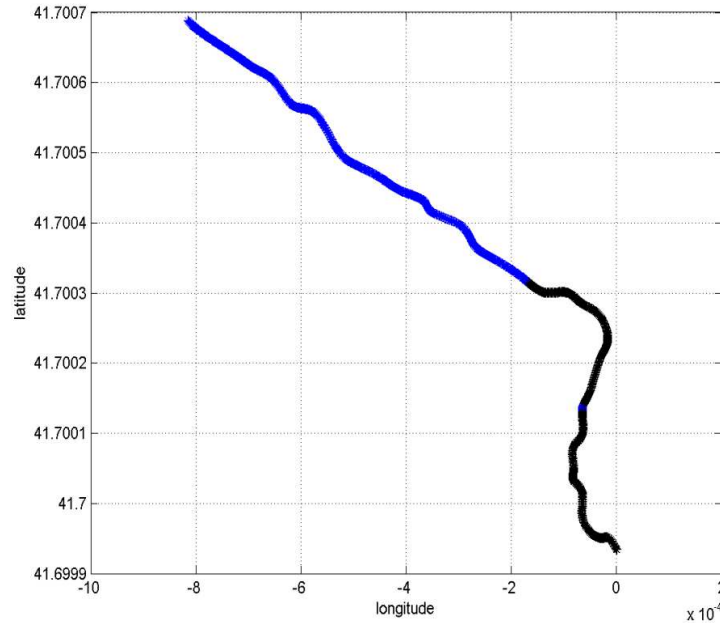


Figure 5 – Trajectoire de l’AUV REMUS 100 pour l’acquisition de la Figure 4². Les positions des pings repérés sur des virages sont en noir.

Nous avons introduit le coefficient de variation quelques années plus tôt pour détecter une anomalie statistique dans la colonne d’eau (voir travaux publiés dans [QUI12], et paragraphe 3.3) mais la transposition de l’algorithme sur le fond marin exige quelques ajustements (seuil local, post-traitements,...) en grande partie liés à l’hétérogénéité et la complexité de l’arrière plan, c’est-à-dire le fond marin. Compte tenu de la taille du pixel évaluée à 0.059m en distance et 0.125m en azimuth pour une portée de 30m, la fenêtre glissante de calcul du coefficient de variation est rectangulaire. Notons I_{cv} l’image obtenue par calcul du coefficient c_v en chaque pixel. La détection proprement dite des anomalies se fait par seuillage de cette image. Afin de rendre l’estimation du seuil plus fiable, les niveaux des pixels sont préalablement bornés au 99^{ème} centile. Un seuil global par rapport au niveau c_v moyen de l’image entière n’est pas satisfaisant sur des zones encombrées (dans le jargon, on peut trouver le terme anglo-saxon de fort *clutter*). Par ailleurs, un seuil théorique est inenvisageable pour plusieurs raisons dont la possible troncature d’histogramme par le constructeur pour rehausser la dynamique, le changement sédimentaire ou encore la présence de végétation... Nous avons donc estimé un seuil local, en chaque pixel, en appliquant à l’image I_{cv} un filtre moyenneur de dimension 5 fois plus grande que la fenêtre de calcul, soit I_{cvs} l’image obtenue. La détection d’anomalie consiste ainsi simplement à comparer les deux images. L’image binaire I_{cvdet} des détections est obtenue par l’opération suivante, pixel à pixel :

² Les coordonnées GPS ne correspondent pas aux zones géographiques indiquées par les organisateurs, probablement pour des raisons de confidentialité.

$$I_{cvdet} = I_{cv} > 2 \times I_{cvs}.$$

Cette image I_{cvdet} est traitée par fermeture binaire à l'aide d'un élément structurant deux fois plus long (axe distance) que la fenêtre de calcul de manière à connecter des zones connexes appartenant au même objet (cas des objets dont l'ombre portée de forte élongation a été fragmentée), puis par ouverture avec un simple élément structurant en croix pour éliminer des fausses alarmes qui encombreraient la suite du traitement. Parmi les anomalies ainsi détectées, certaines sont dues à la nature du fond (bosquets de végétation par exemple) ou de taille incompatible avec une mine. Aussi, des post-traitements peuvent, par des critères assez simples, éliminer ces fausses alarmes dues à des macro-structures du fond ou des « objets » trop petits ou au contraire trop gros pour être des mines. Il faut cependant bien noter que ce faisant, certaines IEDs de forme ou taille atypique pourront aussi être écartées. Les critères retenus ici sont les suivants :

1. Le contraste réverbération sur ombre C_{ro} minimum est de 20dB
2. Le contraste écho sur réverbération C_{er} minimum est de 8dB
3. Dimension « standard » attendue à une rasanee donnée : l'objet (sur la base de son ombre segmentée) doit avoir une hauteur minimum de 20cm et une largeur minimum de 30cm, mais aussi une hauteur maximum de 1m.

Pour évaluer chacun de ces critères, une imagerie de taille variable en fonction de la rasanee pouvant contenir un objet de 1m de hauteur est découpée autour de l'anomalie détectée. Cette imagerie sonar est ensuite débruitée par filtrage médian de dimension 3×3 puis segmentée en 3 classes : fond, écho et ombre en utilisant deux seuils égaux aux centiles de 5% et 99%. Les zones connexes les plus importantes de l'imagerie correspondent pour chaque seuil respectivement à la zone écho (en avant de l'ombre) et la zone d'ombre du potentiel objet détecté. Ces masques permettent l'estimation des contrastes (critères 1 et 2) à partir de l'imagerie sonar brute :

$$C_{ro} = 20 \log_{10} \left(\frac{m_r}{m_o} \right) \text{ et } C_{er} = 20 \log_{10} \left(\frac{m_e}{m_r} \right)$$

où m_x désigne le niveau de gris moyen des pixels de la région x .

Afin de satisfaire le 3^{ème} et dernier critère, le masque de l'ombre est d'une part érodé pour savoir s'il correspond à un objet de taille minimale. D'autre part, si l'imagerie ne peut contenir tout le masque de l'ombre (dont l'élongation dépend de la rasanee courante), cette détection correspondant à un objet trop haut est écartée.

3.2.1.1.3 Résultats

Notre performance a été en deçà de nos espérances avec certes le meilleur taux de détection mais aussi un taux de fausse alarme inacceptable. Cela est en grande partie dû à une différence d'appréciation. Notre algorithme s'est attaché à retenir tout objet dont

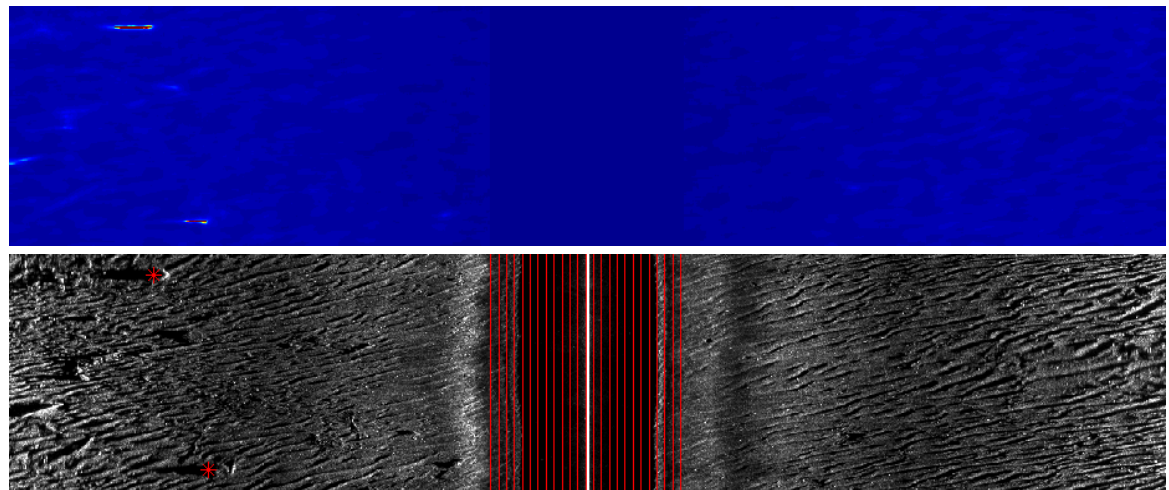
la signature acoustique « ressemblait » en matière de gabarit et de contraste à la signature possible d'un MILCO mais sans aller jusqu'à identifier cet objet en mine ou non mine qui aurait nécessité une analyse précise de la signature de chaque objet détecté. Ce point montre que les notions de classification MILCO/non MILCO et identification mine/non mine (telle qu'indiquée sur la Figure 1) peuvent être diversement comprises. Enfin, les détections ont été stockées image par image sans recouplement ce qui a mathématiquement gonflé les taux pour un objet perçu plusieurs fois.

Le tableau 1 donne différents exemples de détections obtenues sur des zones de difficulté variables. Notez que la taille du pixel est environ deux fois plus grande pour une portée de 50m comparée à celle qui est observée sur une portée de 30m. Sur les cibles étendues, le coefficient de variation est élevé au niveau des frontières ombre/fond ou ombre/écho. Le dernier exemple montre bien l'intérêt du seuillage local qui permet la détection d'une cible malgré une faible valeur du coefficient de variation à cet endroit de l'image.

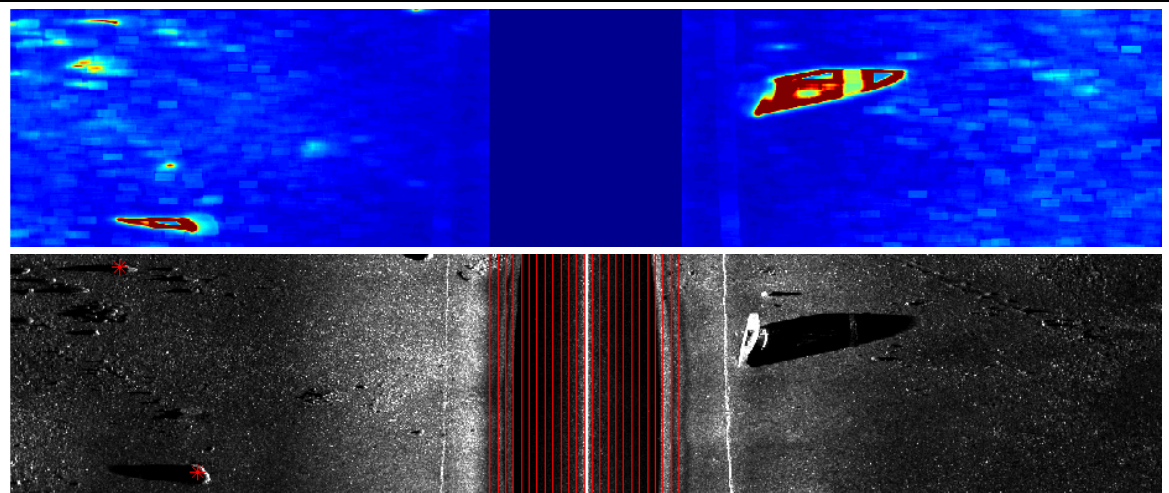
Tableau 1 – Echantillon de résultats de détections sur des zones variées (image *Icv* en table de couleur 'jet' de MatLab et détections rapportées croix en rouge sur l'image sonar)

Zone très encombrée (image Training00208) sur une portée de 50m et une trace de 45m	
Zone de fond texturé (image Training00001) sur une portée de 30m et une trace de 30m	

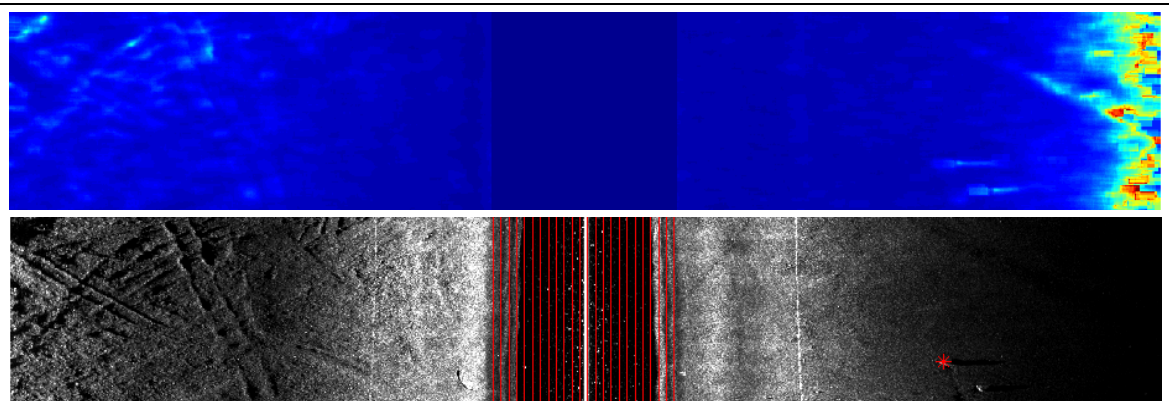
Zone de rides de sable (image Training00640) sur une portée de 30m et une trace de 30m



Zone d'objets divers (image Training00020) sur une portée de 30m et une trace de 30m



Fond avec trace de chalut (image Training00129) sur une portée de 30m et une trace de 25m



3.2.1.2 Nouvelle approche de l'ATR

3.2.1.2.1 Prise en compte de l'environnement proche

Avec l'avènement de véhicules sous-marins autonomes de plus en plus stables, conjugué à des capteurs de perception de plus en plus résolus, il y a une forte exigence en matière de performances quelle que soit la zone explorée. L'enjeu consiste à garantir un bon taux de détection tout en maintenant un taux de fausse alarme assez faible pour ne pas surcharger la suite du processus ATR et couvrir la zone en un temps minimal. Posé comme tel, l'étude des performances fait naturellement intervenir le système mais aussi l'environnement dans lequel il évolue. Depuis le début des années 2010, la communauté de reconnaissance de cibles sous-marines cherche, non plus à gommer l'environnement proche de la cible mais, au contraire, à le caractériser pour, ou bien intégrer des informations environnementales aux algorithmes de reconnaissance, ou bien orienter le choix d'un traitement adapté au contexte opérationnel [SAW10] [WIL10a, WIL11]. La nécessité de prendre en compte les données environnementales dans le processus ATR pour augmenter les performances en détection était aussi l'une des principales conclusions de la première édition du workshop iSAAM (International Workshop on Intersections in Signal Processing, Acoustics & ATR for Maritime Applications) près de Washington en 2012 auquel notre équipe a été invitée (cf. paragraphe 2.7).

3.2.1.2.2 Lien avec l'état de l'art

Pour qualifier l'environnement proche de la cible dans l'image sonar latéral (conventionnelle ou synthétique), deux critères s'imposent rapidement dans la communauté : l'anisotropie et la complexité. Ces critères seront introduits presque simultanément par le CMRE [FAK13] et l'université de Heriot-Watt [DAN12], puis repris par FFI [GEI14] et nous-mêmes à l'ENSTA Bretagne via la thèse de Laurent Picard d'une manière un peu différente [PIC15b, PIC16]. L'anisotropie est définie en 2013 par E. Fakiris et D. Williams [FAK13] comme le carré du coefficient de variation calculé en sortie d'un banc de filtres de Haar. Il se trouve que le carré du coefficient de variation, coefficient que nous avons introduit au paragraphe précédent, fut déjà utilisé sous le terme de coefficient de lacunarité par R. Kessel pour détecter des mines dans les images sonar au début des années 2000 et repris plus tard par J. Fawcett [KES02] [FAW12]. Contrairement à l'anisotropie, la lacunarité est calculé en chaque pixel à partir des niveaux de gris des pixels de son voisinage sans orientation particulière. La thèse de M.J. Chantler situe l'origine du mot dans les années 80 [CHA94]: "*The term lacunarity was derived by Mandelbrot from the Latin term for gap (lacuna) and he used it to describe the size of holes in images of galaxies [MAN83]. L. M. Linnett [LIN91] likened this characteristic to "structure" in texture, which is dependent on phase information*

[CLA92]”. L. M. Linnett [LIN91] qui s’intéresse à la segmentation des fonds marins en matière de texture ajoute en effet dans son manuscrit que : “*This (lacunarity) appears to be similar to the use of phase (after a Fourier transform) in helping to describe the positions of primitives in an image*”, ce qui en fait un paramètre adapté à la détection d’information sémantique, comme par exemple des débris autour d’une épave sous-marine [CLA92]. Pour revenir à B. B. Mandelbrot, il aurait introduit ce paramètre pour distinguer des fractales de même dimension mais d’apparences ou textures différentes [KEL89]. R. Kessel note en effet qu’une zone de fond homogène (caractérisée en imagerie sonar par l’alternance aléatoire d’un niveau de gris fort à faible d’un pixel à l’autre, typique du phénomène de chatoiement ou *speckle*) est peu lacunaire car n’a pas de ‘structure’ apparente alors que la présence d’écho ou d’ombre induit une forte lacunarité. Le seuil de détection de région d’intérêt qu’il utilise est le carré de la valeur théorique du coefficient de variation pour une loi de Rayleigh égale à 0.523 soit un seuil de 0.273. Le bruit multiplicatif de speckle fournit ainsi une borne basse pour la lacunarité. Plus la lacunarité augmente, plus la région est hétérogène. Si on fait le parallèle avec la méthode des boîtes pour le calcul de la dimension fractale d’une image en niveaux de gris considérée comme une surface tridimensionnelle, l’hétérogénéité désigne la présence de monticules/trous (ou de leurs frontières si la fenêtre ne peut les contenir entièrement) dans une fenêtre dont la dimension doit être choisie en fonction de la taille des structures ou objets recherchés (par exemple, 1m×1m dans le paragraphe 3.2.1.1 décrivant notre détecteur d’anomalie inspiré de [QUI12]). Au lieu de mesurer l’écart entre la masse³ courante et la masse attendue dans la géométrie fractale, nous mesurons ici un étalement des niveaux de gris par rapport au niveau moyen. Plus récemment, D. Williams reprend également ce calcul de la lacunarité en chaque pixel (cette fois sans les fenêtres de Haar) pour distinguer un type de fond particulier : pointant les limites de l’anisotropie et de la complexité pour caractériser le fond marin dans les images sonar et en particulier, pour distinguer un fond facile homogène et un fond couvert de posidonie, il démontre l’intérêt de la lacunarité pour une meilleure caractérisation et un outil de mesure du niveau de difficulté en chasse aux mines [WIL15].

Revenons toutefois aux deux critères, anisotropie et complexité, pour en détailler les différentes déclinaisons trouvées dans la littérature et tenter de pondérer le récent propos de D. Williams. Rappelons que le coefficient d’anisotropie introduit en 2013 par E. Fakiris et D. Williams est la lacunarité de R. Kessel mais calculé en sortie du filtrage (adapté) de fenêtres orientées (12 orientations) de Haar (motifs sensés être sensibles à la différence de niveau entre écho et ombre) de taille variable fonction de la résolution, ce

³ La masse peut être définie par le nombre de boîtes nécessaires pour couvrir l’amplitude des niveaux de gris des pixels de la fenêtre glissante [DON00].

qui en fait un outil adapté aux rides et de leur point de vue spécifiquement dans le cas défavorable de faible contraste ombre sur écho. On passe ainsi d'une mesure isotrope de région d'intérêt ou d'anomalie (lacunarité, coefficient de variation) à une mise en évidence de zones de rides par une mesure anisotrope (l'anisotropie)... En ce qui concerne l'origine exacte de la notion d'anisotropie, E. Fakiris et D. Williams disent s'être inspirés du papier de O. Daniell *et autres* en 2012 mais lui-même dit s'être inspiré d'un article de D. Williams où est introduite une mesure de *ripplicity* [WIL10b] [DAN12]... Quoiqu'il en soit, O. Daniell semble le premier à utiliser le terme d'anisotropie mais sans faire le lien avec la lacunarité (ou le coefficient de variation) puisque son implémentation de l'anisotropie est basée sur le calcul de l'écart-type des réponses énergétiques de bancs de filtres orientés. Concernant la complexité, O. Daniell et D. Williams utilisent cette fois une approche bien différente. E. Fakiris et D. Williams en 2013 le définissent comme le produit d'un SNR avec la moyenne qui a l'avantage de mettre en valeur les zones encombrées (*clutter*) via le SNR et les zones de niveau énergétique suffisant grâce au produit par la moyenne. D'après le même papier, cette définition a l'avantage de ne pas réagir aux rides contrairement à la définition de O. Daniell qui est juste une mesure de l'énergie moyenne des sorties du banc de filtres.

La finalité du calcul de ces descripteurs est bien différente selon les auteurs. O. Daniell les utilise en post-traitement pour écarter quelques fausses alarmes [DAN14]. D. Williams organise la phase de classification en fonction de la valeur de ces descripteurs [WILL14]. Les norvégiens se sont également penchés sur la qualification de l'environnement afin d'intégrer ces coefficients à leur outil d'estimation de performances [MID14]. Leur calcul des descripteurs d'anisotropie et complexité est proche de O. Daniell et de sa dernière implémentation, c'est-à-dire sur la base de réponses énergétiques des coefficients d'ondelettes complexes [GEI14].

3.2.1.2.3 Une autre qualification de l'environnement : travaux du doctorant Laurent Picard

Laurent Picard démarre sa thèse consacrée à la « Prise en compte de l'environnement marin dans le processus de reconnaissance automatique de cibles sous-marines » dans ce contexte en octobre 2013 grâce à un financement de la Direction Générale de l'Armement et sous la direction d'Alexandre Baussard [PIC17]. Notre objectif était d'intégrer au plus tôt des informations environnementales au processus ATR, notamment dès la phase de détection. Gilles Le Chenadec, expert de la caractérisation des fonds marins, a participé avec moi au co-encadrement des travaux.

La contribution de Laurent à la qualification des fonds marins dans les images sonar latéral s'est déroulée en deux temps (voir Figure 6). Tout d'abord, une nouvelle représentation isotrope (*i.e.* sans avoir, comme dans les approches basées sur les

ondelettes ou les bancs de filtres orientés, à choisir un nombre d'orientations pour analyser l'image) de l'image sonar a été obtenue par le biais du signal monogène [FEL01]. Cet outil, pendant 2D du signal analytique en 1D, permet une séparation orthogonale des informations énergétiques, géométriques et structurales d'une image définies en chaque pixel par l'amplitude, l'orientation et la phase du signal monogène (voir Figure 7).

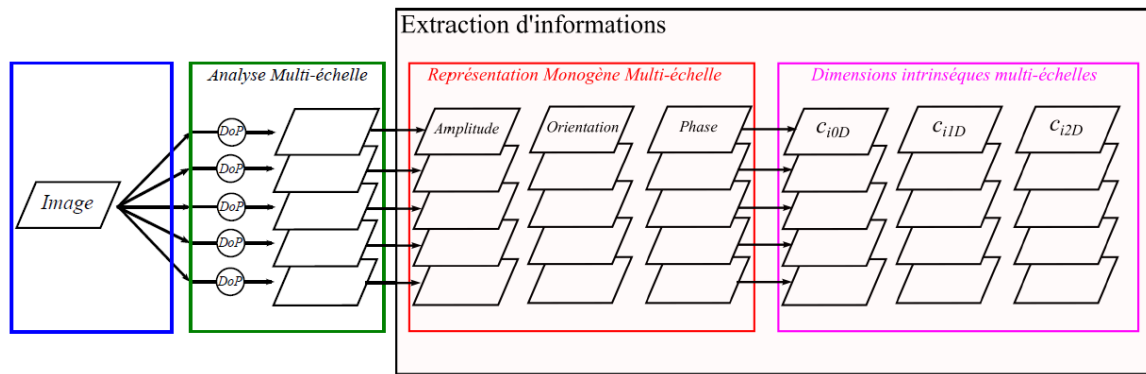


Figure 6 – Synoptique des traitements développés par Laurent Picard

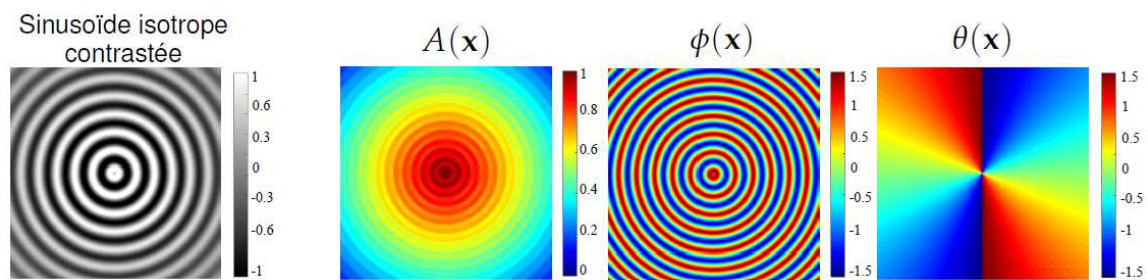


Figure 7 – Exemple de représentation monogène d'une image synthétique (à gauche) : amplitude locale $A(x)$, phase locale $\phi(x)$ et orientation locale $\theta(x)$ où x désigne un pixel

Le calcul du signal monogène a été fait dans un cadre multi-échelle basé sur la mise en œuvre de filtres différence de noyaux de Poisson notés DoP (*Difference of Poisson*) afin de décomposer l'image en différents niveaux de détails permettant de cibler les dimensions typiques des mines marines. Un second outil, capable de quantifier les degrés de libertés d'un signal 2D, a été choisi par Laurent pour mener la qualification des fonds souhaitée : le concept de dimension intrinsèque. La représentation monogène (phase locale exceptée) multi-échelle permet l'estimation des dimensions intrinsèques qui sont au nombre de 3 pour chaque échelle. Depuis les travaux de Krüger, il en existe une formulation continue qui permet de fournir en chaque pixel 3 coefficients bornés (entre 0 et 1) et complémentaires que l'on peut voir comme des indices de confiance en chacune des dimensions intrinsèques [KRU03]. L'exemple Figure 8 de valeurs d'indices calculés sur une image sonar où figure une zone de rides démontre que la dimension intrinsèque de 1, notée C_{i1D} , réagit bien à l'échelle des rides. Cette observation est entrée en compte en fin de thèse dans le développement d'un détecteur de cibles sur ces fonds

anisotropiques problématiques en chasse aux mines. Une analyse des valeurs de ces indices pour différents types de fond a mené à la définition de nouveaux descripteurs d'homogénéité, d'anisotropie et de complexité également bornés chacun dans un intervalle $[0; 1]$, dont la complémentarité facilite l'interprétation et l'usage pour un opérateur ou un AUV. Ces trois descripteurs peuvent être utilisés comme aides à l'opérateur ou comme données d'entrée de la boucle perception-action d'un véhicule autonome. Sur les fonds homogènes ou anisotropes, nous avons démontré un réel bénéfice pour les opérateurs en certifiant d'une part l'absence d'objets sur un fond homogène, et en proposant d'autre part un algorithme performant (au vu de la littérature) de détection d'objets de type mine sur les zones anisotropes comme les champs de rides de sable (un article décrivant cet algorithme a été soumis à la fin de la thèse à la revue IEEE Transactions on Geoscience and Remote Sensing). Cet algorithme combine la dimension intrinsèque de 1 pour trouver le niveau de détail (échelle) qui caractérise le mieux les rides, et l'amplitude et la phase du signal monogène pour détecter un objet sur la base de son écho (persistance énergétique dans les échelles supérieures à celle des rides de sable) et de son ombre (où la valeur de la phase locale monogène est très faible à l'échelle des rides de sable). En mode embarqué, cette cartographie du fond peut également guider ou dérouter le véhicule sous-marin afin de remplir au mieux sa mission de déminage [WIL12].

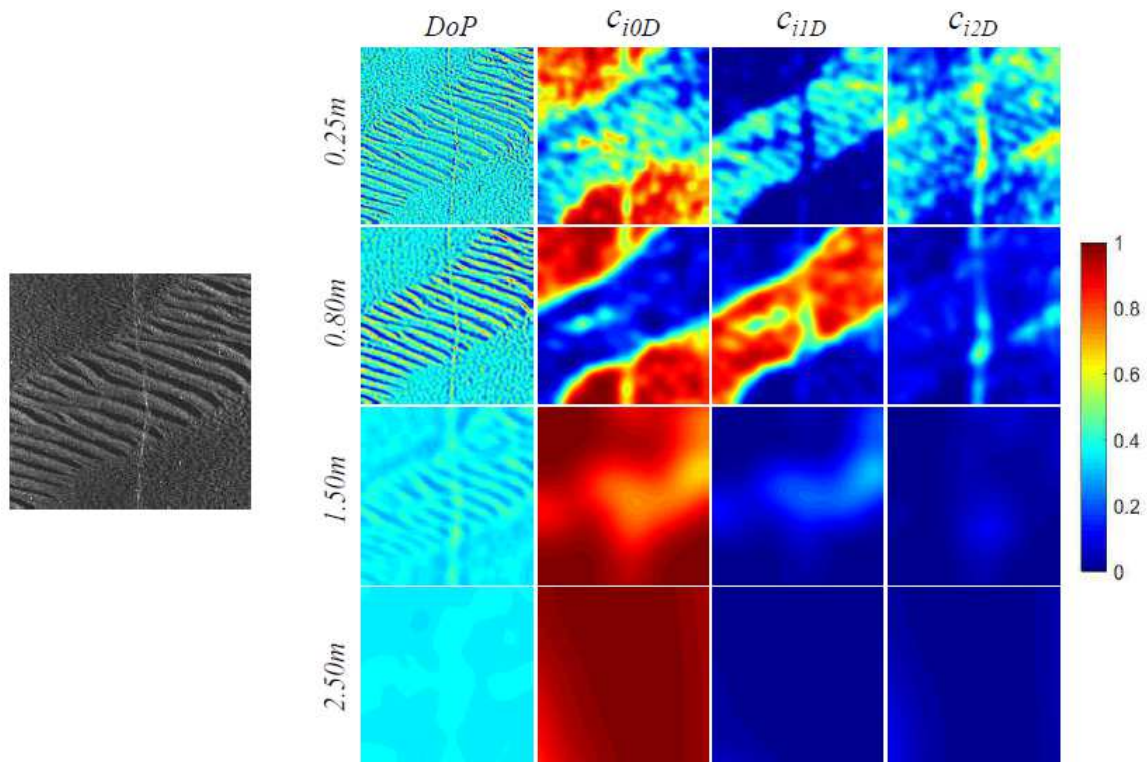


Figure 8 – Dimensions intrinsèques sur une zone mixte rides/sable

La Figure 9 donne un exemple de cartes d'anisotropie et de complexité (la carte d'homogénéité étant complémentaire des deux) obtenues sur une image sonar latéral Klein 5000. Notez bien la dynamique des cartes comprises entre 0 et 1. Les zones de rides de sable sont bien à un niveau maximal d'anisotropie et minimal de complexité. Les zones homogènes ont bien des niveaux faibles sur les deux cartes. Par contre, les grandes zones d'ombre sont aussi étiquetées homogènes (elles le sont effectivement du point de vue des dimensions intrinsèques calculées) alors qu'un objet peut s'y cacher. Laurent a ainsi dans le contexte ATR souligné la nécessité de détecter ces ombres dans les zones homogènes, ce qui implicitement introduit une zone de fond supplémentaire, distincte des zones homogènes sans danger.

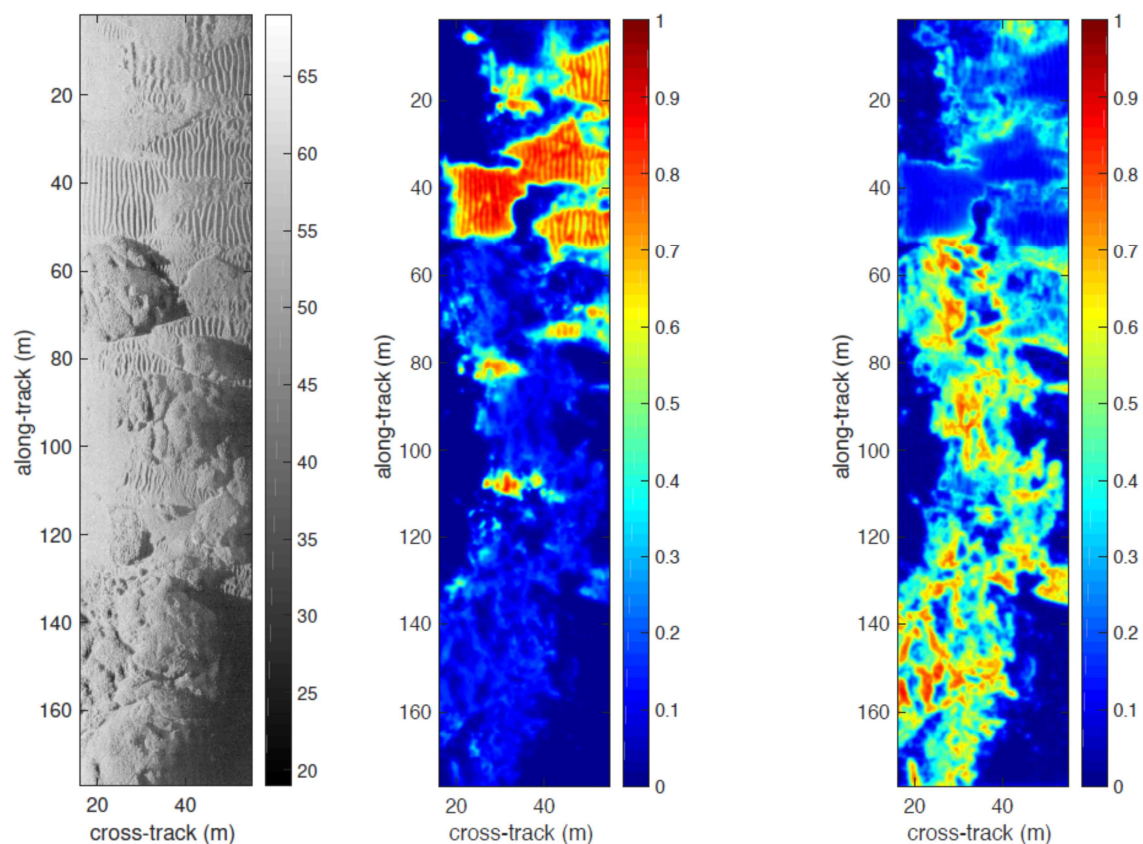


Figure 9 – Exemple de cartes d'anisotropie (au centre) et de complexité (à droite) obtenues sur une image sonar latéral Klein 5000 (données DGA) [PIC17]

Bien que destiné à des images sonar, le jury de soutenance a souligné la généricité de l'outil développé. Il peut en effet s'appliquer à diverses bases de données sonar (y compris si la dynamique ou la statistique des images ont été affectées par des prétraitements dont nous n'avons pas la source) mais également d'autres données comme des données de télédétection spatiale par exemple.

Les perspectives de ce travail sont multiples et font l'objet d'une prochaine offre de thèse.

En premier lieu, comme mentionné ci-dessus pour les zones homogènes, et plus encore s'agissant des zones dites 'complexes', il conviendra de se poser la question d'une qualification plus fine que la seule qualification des fonds en zones homogènes/anisotropes/complexes. Par ailleurs, une étude approfondie des informations monogènes dont seule l'amplitude est sensible au contraste doit être menée. La phase locale doit en particulier être mieux comprise même si son potentiel pour la détection de cibles dans les zones anisotropes a déjà été mis en évidence (voir plus haut). L'analyse multi-échelle nécessite également un effort particulier : quelles échelles d'analyse pour quel type d'information ? pour décrire le fond ? pour décrire les cibles ?

En second lieu, la détection de cibles sur des fonds complexes qui regroupent une grande variété de fonds en termes de géométrie et de structure est à étudier. Raisonnablement, un détecteur unique est inenvisageable. Une première étape nécessite d'analyser plus finement ces fonds. Toujours dans une approche multi-échelle, on peut par exemple s'intéresser à caractériser une zone particulièrement texturée, à évaluer l'encombrement de la zone, modéliser le fond pour l'« apprendre », voire effectuer une partition des fonds complexes suivant un ensemble de critères à définir... Ces éléments serviront d'entrée dans une seconde étape pour développer des détecteurs plus ou moins évolués en fonction du niveau de complexité du fond.

En troisième et dernier lieu, dans une problématique de déminage par véhicule autonome, les informations remontées sur le type de fond doivent être prises en compte pour replanifier la mission en cas de zones d'ombre susceptibles de cacher une cible ou encore de faible niveau de confiance sur la classification d'une cible nécessitant l'acquisition d'un second point de vue. Ainsi, cet ensemble porteur-environnement-capteur promet des gains importants de performance (taux de détection et classification MILCO/non MILCO) qu'il convient de quantifier, ce qui à notre connaissance n'a été fait que partiellement, sur la base du résultat de classification [ZER09], de la portée effective du sonar [COU12] ou encore de la qualité du fond [MID14].

3.2.1.3 Classification multi-vue

3.2.1.3.1 Etat de l'art

Ce thème d'intérêt est concomitant au développement des véhicules sous-marins (autonomes ou pas) capables de faire le tour de l'objet en toute sécurité afin d'en acquérir plusieurs points de vue. Thales Underwater Systems a cofinancé mes travaux de thèse sur ce sujet à partir de novembre 1998 pour automatiser la classification d'objets au moyen du sonar propulsé à immersion variable (SPIV), un équipement novateur apparu lors de la refonte des chasseurs de mines tripartites [QUI01a]. Bien que les algorithmes n'aient été testés alors que sur données simulées, plusieurs stratégies sont

encore d'actualité et ce, d'autant plus que la technologie des véhicules et les performances des sonars et des capteurs de navigation sont aujourd'hui bien meilleures qu'à l'époque. Dès les années 90, plusieurs travaux font appel à des techniques de fusion basés sur la théorie de Dempster-Shafer ou la théorie des possibilités au NURC (aujourd'hui CMRE) et au GESMA (aujourd'hui DGA-TN) d'abord, puis à l'université de Heriot-Watt et enfin au RDDC-A [STA96] [DAN98] [REE04b] [FAW10]. Un autre sujet concerne l'étude des performances fonction du nombre de vues [WIL09]. Plus récemment, et vu le gain obtenu dès le second point de vue, l'intérêt s'est porté sur le choix optimal de cette seconde vue en cas d'ambiguïté sur la première vue [JI05] [ZER09] [MYE12] [LEB10] [LER11]. L'exploitation des différents aspects peut permettre la reconstruction tridimensionnelle de l'objet, par rétroprojection des ombres [MAL00] ou par des techniques de *shape from shading* [COI09a]. Une autre possibilité est de combiner l'ensemble des informations disponibles pour effectuer une classification sur une seule image au lieu de fusionner les résultats de classification sur chaque vue. Ce fut le cas sur des données simulées dans la dernière partie de ma thèse [QUI01a], un peu comme l'avait fait un peu plus tôt STN Atlas (aujourd'hui ATLAS Elektronik) [HOE98] et qui a été ensuite proposé au NURC [GRO10]. La difficulté avec des données réelles est dans le géoréférencement des différents points de vue et/ou leur recalage si les images sont de qualité suffisante. La question est ensuite de pouvoir classifier l'objet sur la base de cette nouvelle image : [QUI01a] et [GRO10] ont fourni des pistes dans leurs cas respectifs. Cette piste a été reprise en 2011 lors d'un séjour de recherche à l'étranger financé par la DGA (cf. paragraphe 2.7) pour mettre au point des algorithmes de détection et de classification pour les systèmes sonar nouvelle génération de Thales qui offrent une acquisition multi-aspects sur une seule passe (DUBM44 puis SAMDIS). L'intérêt des systèmes multi-aspects est de fournir en une seule passe trois aspects sur une cible sans devoir dérouter le véhicule, d'où un gain de temps considérable [CHA14]. Le gain de performances est accru par la combinaison des différents points de vue bien sûr mais également du fait de la limitation du risque d'occultation de la cible.

3.2.1.3.2 Géométrie floue pour l'analyse des données multi-aspects du DUBM44

L'objectif de cette étude effectuée dans le cadre de l'AS 42 et d'un séjour de recherche au Canada était l'évaluation des capacités du système sonar DUBM44 à discriminer différentes classes d'objets. On se place donc ici dans l'étape d'identification de la Figure 1. Le système dispose de 2 modes : le mode SAS et le mode multi-aspects. Le second mode permet de former 3 images conventionnelles selon 3 aspects à -30, 0 et 30 degrés (cf. Figure 10).

Nous avons choisi de reprendre le principe de combinaison des différents aspects dans une seule image proposé lors de ma thèse mais, à la différence des données supposées acquises

alors selon une trajectoire circulaire autour de l'objet (cas de mes travaux de thèse), nous ne disposons ici que de trois aspects selon des points de vue bien définis. Par ailleurs, nous souhaitons cette fois également exploiter l'écho. Pour se doter d'une base de données représentative, nous avons utilisé le simulateur SIGMAS développé au CMRE et disponible lors de mon séjour au RDDC-A, pour créer une base de triplets d'images pour 7 classes d'objets le long de la portée du sonar pour différentes orientations des objets vis-à-vis du capteur [COI09b] et pour des résolutions supposées proches du système réel.

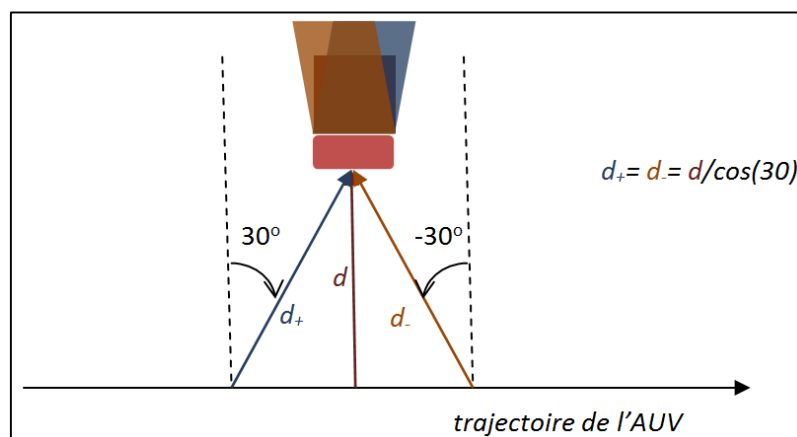


Figure 10 - Configuration du DUBM44 en mode multi-aspects sur un objet cylindrique vu par le travers sur l'aspect central

Nous avons formé deux images par sommation, après recalage sur la position de l'objet déterminée par la simulation et normalisation des 3 imageries à une rase de référence :

- L'une des images est obtenue par sommation des ombres préalablement segmentées par un seuillage percentile équivalent à 20% des pixels,
- L'autre est obtenue par sommation des échos bruts puis seuillage percentile équivalent à 1% des pixels de l'image somme globale (des post-traitements utilisant la position relative de la zone d'ombre et de la zone d'écho permettent de supprimer des fausses alarmes).

En effet, sommer les ombres brutes n'a pas de sens du fait du faible nombre d'aspects et du bruit de *speckle* affectant cette région de l'image. Ce n'est plus le cas des zones d'échos pour lesquelles l'amplitude et la position relative des pixels (*i.e.* la forme bidimensionnelle de l'écho) sont porteurs d'information sur la forme, l'orientation et parfois la nature (fonction de la fréquence du sonar) de l'objet.

Les niveaux de gris des deux images ainsi obtenues sont normalisés entre 0 et 1. La valeur du pixel est donc fonction de :

- pour l'ombre, la fréquence d'apparition sur les différents aspects (ensemble discret de valeurs possibles : $\{0,1,2,3\}/3$),

- pour l'écho, le niveau relatif du pixel sur l'ensemble des trois aspects (ensemble continu de valeurs possibles entre 0 et 1).

La Figure 11 donne au centre et à droite l'exemple d'une paire de ces images sur lesquelles ont été extraits des attributs flous introduits dans [QUI01a]. La géométrie floue permet de caractériser une région d'une image en niveaux de gris, comme la géométrie classique le fait pour une image binaire. Cette approche est née de l'extension d'algorithmes standards de traitement d'images à la logique floue (géométrie floue, mesures floues de l'information de l'image, morphologie mathématique floue,...) [TIZ98]. Dans ce formalisme, la pertinence de chaque pixel est proportionnelle à son niveau de gris vu comme un degré d'appartenance à ladite région. L'application de cette géométrie semble adaptée à la caractérisation d'une part, de la zone d'ombre segmentée pour laquelle la dispersion des pixels est d'autant plus importante que l'ombre portée de l'objet diffère d'un point de vue à l'autre et d'autre part, de la zone d'écho brut (après sommation ou pas) pour laquelle le niveau de gris du pixel est lié à l'énergie rétrodiffusée par la facette correspondante de l'objet.

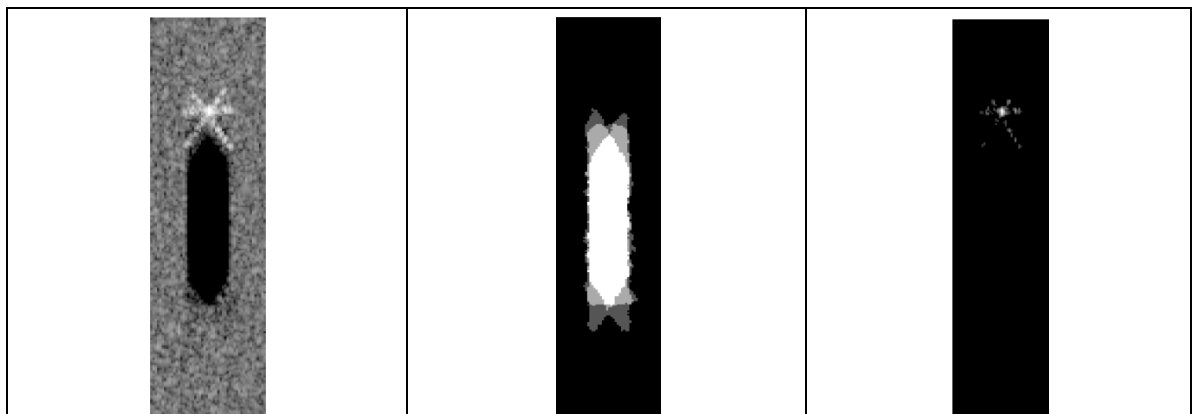


Figure 11 – Cas du cylindre par le travers (schématisé Figure 10) : image de sommation brute à gauche, et image sommée des ombres segmentées au centre, et image sommée sur échos bruts à droite.

L'étude ayant été financée par la DGA et non publiée, la définition des attributs flous et l'étude de leur potentiel pour discriminer différents types d'objets acquis par un système de type DUBM44 ne peuvent être détaillées ici. Cependant, une analyse factorielle discriminante a montré une bonne partition des différents groupes d'objets sur la base de 11 attributs flous et son interprétation a permis de dire quels étaient les attributs les plus influents sur les deux premiers axes factoriels. Une validation sur données réelles n'a malheureusement pas pu être menée à ce jour du fait du niveau de classification des données du système DUBM44.

3.2.2 Surveillance

Le monitoring acoustique actif de zones sensibles (chenal d'accès militaire, port, ...) à des fins de détection de mines ou d'explosifs est encore à ses balbutiements. L'actuelle menace terroriste met pourtant cet enjeu au premier plan. L'acquisition répétée, et suivant le même point de vue, de données sonar latéral (ou SAS) par un AUV est l'une des solutions alternatives au déploiement d'un poisson remorqué depuis un navire de surface⁴. Pour la surveillance de chenal, à la manière de ce qui se fait déjà en radar pour le monitoring de zones terrestres ou maritimes sensibles, les études portent sur la détection d'éventuels changements entre deux acquisitions ou passes du même système sonar séparées d'un intervalle de temps allant de quelques heures à plusieurs mois : ce qu'on appelle la détection de changement. La paire d'images acquises contient une image référence (de la première passe) et une image esclave ou répétée (de la seconde passe). Un « changement » peut être une modification de la structure du fond ou bien l'apparition ou disparition d'un objet (naturel ou manufacturé). Dans le contexte de la surveillance de chenaux ou de zone portuaire, on s'intéresse bien évidemment à savoir si un ou plusieurs objets suspects sont apparus dans l'intervalle de temps qui sépare les deux passes. Ceci relève de la contre-mesure appliquée aux mines marines mais c'est également un des moyens privilégiés pour la détection de dispositifs explosifs artisanaux, désignés dans le domaine par « IEDs : Improvised Explosive Devices », la détection d'obus qui peuvent affleurer au bout de quelques années et désignés par « UXO : UneXploded Ordnance », ou encore de nouvelles mines furtives de forme similaire à un rocher (les PDM par exemple). Dans ce contexte, la nature et la texture du fond ne peuvent être négligées car, d'une part elles jouent un rôle primordial dans l'étape de recalage des images entre elles et d'autre part, elles influencent les performances de la détection de changement. Dès lors, on comprend bien que recalage et détection de changement sont deux points indissociables. Les méthodes de recalage peuvent se diviser en quatre catégories suivant deux critères : méthode cohérente ou incohérente combinée à un recalage rigide ou élastique. La cohérence des données entre les deux passes n'a été que très récemment utilisée, en partie parce qu'elle est limitée temporellement et spatialement [LYO10] [LYO13] [BRO14]. Les méthodes de recalage incohérent sont habituellement classées en deux catégories : les méthodes symboliques et les méthodes iconiques. Les méthodes symboliques nécessitent une étape préliminaire d'extraction d'amers géométriques ou de segmentation de l'image avant une étape d'appariement de ces éléments entre les deux passes [GUE96] [DAN98] [MIG05] [ZER05] [LEB06] [FER12] [VAN11] [BER14]. Les

⁴ La surveillance du chenal d'accès de la base de l'Île Longue en rade de Brest à la haute mer est actuellement réalisée par de multiples passes d'un sonar latéral mis à l'eau depuis un BRS (<http://www.ouest-france.fr/bretagne/brest-29200/l-antares-ange-gardien-des-sous-marins-nucleaires-4556782>) et permet aux sous-marins nucléaires lanceurs d'engins (SNLE) de naviguer en toute sécurité.

méthodes iconiques mettent en œuvre des outils de similarité entre les niveaux de gris des deux images [CHA05] [CHA11] [QUI11].

Les travaux de recherche décrits dans ce paragraphe ont été réalisés en collaboration avec le DRDC-A pendant et à la suite du séjour de recherche de 7 mois en 2011 financé par la DGA via une bourse ERE (voir paragraphe 2.7). Au tout début de l'étude, des algorithmes de recalage et de détection incohérente de changement ont été développés et testés sur un jeu de données en amplitude KLEIN 5500 établi par le DRDC-A en 2008. Ensuite, en avril 2011, les défenses canadienne et norvégienne ont conjointement mené une campagne d'essai au large de la Norvège pour collecter des données du système norvégien HISAS 1030. L'objectif était en particulier de pouvoir vérifier la cohérence du fond entre deux passes acquises avec un intervalle de temps relativement court (entre 2 et 8 jours) pendant lequel plusieurs cibles avaient été posées puis récupérées et d'envisager leur détection par perte locale de cohérence.

3.2.2.1 Positionnement des travaux

A ses débuts, le recalage de scènes sous-marines a eu pour objectifs le mosaïquage, l'aide à la navigation autonome ou encore de fusion de données multi-capteurs [GUE96] [DAN98] [ZER05] [LEB06] [CHA07]. L'intérêt du recalage à des fins de détection de changement en imagerie sonar n'est apparu que vers 2010 et bien plus tard qu'en imagerie radar où les applications à la fois civiles et militaires sont nombreuses [SIN89] [RIG93] [PRE06].

Le RDDC-A avait organisé une campagne d'essais durant l'hiver 2008 pour recueillir des données sonar latéral KLEIN 5500 lors de deux levés séparés d'une durée d'un mois [MYE09a]. Ce sont ces données que nous avons utilisées pour développer notre méthode de recalage et de détection incohérente de changement, publiée en 2012 [QUI12]. Une première étude avait été réalisée fin 2008 sur ces données pour évaluer une méthode de recalage globale par optimisation des corrélations locales et une technique de détection de changement par différence des corrélations locales [FOR08]. S. Daniel de l'université de Laval au Québec avait ensuite travaillé sur ces mêmes données en 2010 pour le RDDC-A et étudié différents algorithmes (corrélation également, mais aussi adaptation au contexte des images sonar des algorithmes SIFT et SURF) de recalage sur ces mêmes données [*rapport non disponible*]. Les approches proposées à l'issue de ces premiers travaux du RDDC-A s'appuyaient en particulier sur les points saillants c.-à-d. les objets présents, ce qui en fait des approches symboliques. C'est une stratégie que nous n'avons pas choisie à mon arrivée en janvier 2011 car l'objectif était justement de détecter les éventuels nouveaux objets entre deux passes et que la zone pouvait ne pas contenir d'objets ou structures saillantes. Au contraire, nous avons développé une approche iconique basée sur

la mesure de similarité entre les distributions des pixels de différentes zones homogènes de chacune des images.

Comme indiqué plus haut, peu de travaux avant 2010 font mention de détection de changement en imagerie sonar. A partir de passes recalées, V. Myers a étudié divers critères de similarité (statistique, métrique et entropique) entre les échantillons de pixels obtenus par fenêtre glissante sur chaque passe. Une comparaison de ces critères et d'une autre méthode de détection haut niveau (niveau des objets détectés) a été faite dans [MYE09b]. Beaucoup ont ainsi fait le choix d'appliquer la détection de changement à un autre niveau que l'image, le niveau des objets détectés et classifiés dans chaque image [DUB08] [COI08] [WEI09] [FER12]. De notre côté, nous avons d'abord voulu investiguer la détection pixel à pixel une fois les images recalées. Pour ce faire, et en raison des similitudes entre images radar et sonar, nous nous sommes intéressés aux méthodes de détection de changement en usage en imagerie SAR et en particulier à la division d'images, opération préférée à la différence du fait de la correction intrinsèque de niveau moyen [RIG93] [PRE06]. Cette technique permet de déceler les changements de niveau d'énergie rétrodiffusée entre les deux passes.

Toutes les techniques de détection de changement évoquées ci-dessus sont dites incohérentes car elles n'utilisent pas l'information portée par la phase des données. Cette information étant disponible pour les données norvégiennes, il a été possible d'envisager de faire une détection cohérente de changement comme elle existe déjà en imagerie SAR où l'on parle de décorrélation de speckle ou perte de cohérence [SIN89]. Dans la suite, nous décrirons nos contributions à ces deux approches, détection incohérente (paragraphe 3.2.2.2) puis détection cohérente (paragraphe 3.2.2.3).

3.2.2.2 Détection incohérente

3.2.2.2.1 Travaux réalisés lors d'un séjour de recherche au Canada en 2011 [QUI12a]

Ces travaux ont fait l'objet d'une communication dont l'article est donné en annexe page 119. Par conséquent, dans ce paragraphe, nous en donnons seulement un résumé.

Les données du RDDC-A issues du sonar latéral KLEIN 5500 figuraient à l'inventaire des données à disposition dans le cadre de l'AS n°42 [MYE09a]. Ce sonar opère à une fréquence de 455kHz avec une bande de 20kHz. La résolution nominale est de 3.29cm en distance et 10 à 20cm en gisement mais les images sont géoréférencées (avec une précision d'environ 10m) sur une grille de résolution 11cm×11cm. Deux passes ont eu lieu respectivement les 30 janvier et 18 février 2008. Chaque image couvre une zone bâbord et une zone tribord de 75m de portée sur une trace de 150m en moyenne.

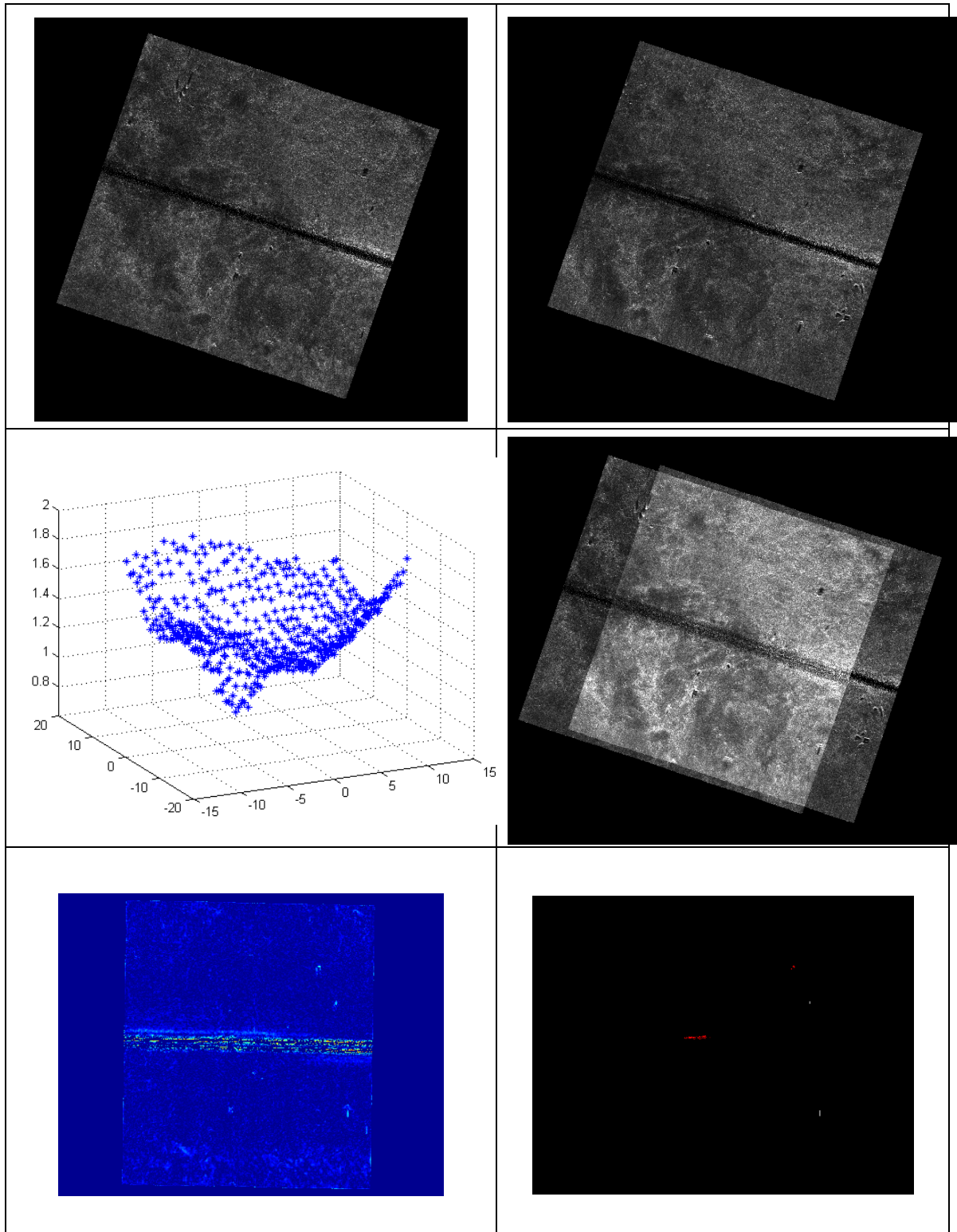


Figure 12 - Résultat du recalage pour un jeu d'images comprenant beaucoup d'objets ; en haut : images à recaler correspondant à chacune des passes ; au milieu à droite : images recalées ; au milieu à gauche : graphe des valeurs de divergences globales pour différentes translations au niveau le plus grossier de la pyramide ; en bas à gauche : log-image en valeur absolue (après rotation) ; en bas à droite : image post-traitées morphologiquement des détections (ensemble des zones connexes rouges et blanches) et les détections finales en blanc retenues après analyse locale.

Pour ces données géoréférencées et dans un objectif de détection incohérente de changement, un recalage rigide prenant en compte une simple translation s'est avéré suffisant pour superposer les zones communes à une paire d'images. L'approche statistique choisie vise à tester différentes translations possibles dans un périmètre raisonnable et à retenir celle qui minimise la somme des divergences de Kullback-Leibler entre distributions de régions homogènes (obtenues par segmentation de Fisher [FIS58]) de part et d'autre. Une approche multirésolution permet d'accélérer le traitement. La Figure 12 donne un résultat complémentaire des résultats donnés dans [QUI12a].

Une fois les images recalées, les « différences » apparaissent plus clairement et l'objectif est de les détecter automatiquement. En imagerie SAR, la division d'images pixel à pixel est pratiquée depuis plus de 25 ans [SIN89] [RIG93] [BAZ05]. L'avantage de la division par rapport à la différence est de s'affranchir intrinsèquement du niveau absolu des pixels dans chacune des images et de ne s'intéresser qu'au changement relatif. On cherche ici à détecter les changements de niveau d'énergie rétrodiffusée entre les deux passes, ou plutôt la puissance moyenne rétrodiffusée par le fond. Une estimée de cette puissance est donnée en chaque pixel par son intensité, c'est-à-dire le carré de l'amplitude. Néanmoins, compte tenu du bruit multiplicatif (*speckle*) affectant les images SAR, l'estimateur précédent est modifié de sorte qu'à la place de l'intensité ponctuelle d'un pixel, on moyenne celle-ci sur les différents pings intégrés pour l'obtention de l'image SAR. Le rapport des puissances ainsi estimées est un indicateur des changements intervenus sur le fond. Le rapport est ensuite passé en échelle logarithmique afin de compresser l'échelle de valeurs et de mieux équilibrer les valeurs au-dessous et au-dessus de 1 [PRE06]. L'image obtenue est appelée log-image. La transposition de ce principe à nos données est assez aisée, la principale adaptation consiste simplement à moyenner les images puisque nous ne disposons que d'un seul ping. Nous définissons donc la log-image formée à partir des images I_1 et I_2 de la manière suivante, où \bar{I}_i désigne l'image $I_i, i = 1, 2$, obtenue après filtrage moyenné :

$$\begin{cases} I_{\log}(i, j) = 20 \log_{10} \frac{\bar{I}_1(i, j)}{\bar{I}_2(i, j)} & \text{si } I_1(i, j) \text{ et } I_2(i, j) \text{ sont non nuls} \\ I_{\log}(i, j) = 0 & \text{si } I_1(i, j) \text{ ou } I_2(i, j) \text{ est nul} \end{cases}$$

Dans la log-image obtenue, les pixels nuls indiquent les zones homogènes entre les deux images (ou bien, conformément à la seconde condition ci-dessus, les zones de non recouvrement ou zones au nadir pour lesquelles les pixels sont nuls pour l'une des deux images) tandis que plus le niveau s'éloigne de zéro, plus les pixels relatifs à ces deux images se distinguent. Deux options se présentent : soit ne s'intéresser qu'à la seule différence, soit conserver en plus l'origine (image I_1 ou I_2) de cette différence. Dans le premier cas, on passe la log-image en valeur absolue. Dans le second cas, les valeurs

positives de la log-image sont relatives à des échos présents dans I_1 seule et les valeurs négatives relatives à des échos présents dans I_2 seule. Nous avons choisi de publier le premier cas, même si le second cas a l'intérêt de tracer l'origine des différences et donc de savoir pertinemment si la différence est due à une apparition ou une disparition d'« objet ». La détection proprement dite nécessite un seuillage percentile fixé à 99% des pixels. L'image binaire obtenue est ensuite post-traitée par morphologie mathématique pour éliminer les détections dues à la présence d'objets trop petits pour être une mine potentielle. Les dernières fausses alarmes sont rejetées, ou bien sur des critères de forme comme la compacité, ou bien en raison d'un recouvrement trop faible entre cette forme détectée par recalage global et la forme correspondante obtenue par recalage local. Notons toutefois que la présence de ces fausses alarmes est en grande partie due à des déformations locales qui ne sont pas gérées par le recalage rigide.

Une seconde méthode a été proposée pour une détection de changement par différence absolue des images obtenues après détection des anomalies statistique. Cette méthode avait des performances équivalentes à celle que l'on vient de résumer (plus d'informations en annexe dans [QUI12a]).

3.2.2.2.2 Travaux du doctorant Florian Nicolas (Thèse CIFRE)

Florian Nicolas étudie l'« Amélioration des techniques de détection de changement pour la lutte contre les mines sous-marines en situation de surveillance » sous la direction de Benoît Zerr depuis février 2015. Andreas Arnold-Bos chez Thales Underwater Systems et moi-même le co-encadrons. Dans ce cadre industriel, Florian dispose de données réelles du dernier système SAS de Thales SAMDIS avec de nombreux couples d'images acquises selon approximativement le même point de vue mais une rasance différente. L'intervalle de temps et/ou la géométrie d'acquisition entre chaque passe ainsi que le géoréférencement des données ne permettent pas d'utiliser la cohérence entre les images.

L'année 2015 a été consacrée au recalage iconique rigide puis élastique. Dans une première communication, différentes métriques de similarité (somme des différences au carré, corrélation normalisée, information mutuelle,...) ont été comparées pour obtenir un recalage rigide (translation seule) mettant en œuvre différentes optimisations (recherche exhaustive dans un périmètre ou simplexe) [NIC16a]. La comparaison s'est faite à différentes résolutions sonar à l'aide de critères de temps de calcul, de performance et de robustesse au bruit. Cette communication a aussi pointé les limites d'un recalage rigide qui peut donner de bons résultats visuels mais échoue à modéliser les déformations locales comme cela a été noté précédemment. La suite des travaux a alors consisté à développer une méthode de recalage non-rigide (à résolution unique cette fois) et a fait l'objet d'une seconde communication [NIC16b]. Cette méthode consiste dans un premier temps à apparier dans un cadre multirésolution des blocs issus de chacune des deux

images à recaler par une technique de corrélation rapide (dans le domaine de Fourier) et autorisant le masquage de certaines régions susceptibles de pénaliser le recalage, technique inspirée de [PAD12]. Dans un second temps, le champ de vecteurs obtenu est filtré sur la base de données de navigation et une fonction de déformation (dite *warping function* à base de B-Splines comme dans [QUI12b] ou triangulation de Delaunay) est appliquée dans un troisième temps à l'image esclave. Cette méthode de recalage étant satisfaisante, la détection de changement proprement dite a pu être étudiée en 2016. Un article détaillant une approche basée Markov de détection de changement ternaire permettant de pister l'origine des changements (apparition ou disparition) est en préparation et va être soumis à IEEE Journal of Oceanic Engineering. Outre la rédaction du manuscrit, la troisième année de thèse a commencé par une réflexion différente du recalage et de la détection de changement avec une approche symbolique.

3.2.2.3 Détection cohérente

3.2.2.3.1 Travaux réalisés en collaboration avec le DRDC-A et FFI [QUI12b]

Dérivée de l'optique, la cohérence complexe entre deux signaux complexes centrés x et y est définie par

$$\gamma = \frac{\Gamma_{xy}(0)}{\sqrt{\Gamma_{xx}(0)}\sqrt{\Gamma_{yy}(0)}}$$

où $\Gamma_{xy}(\tau) = E\{x(t)y^*(t-\tau)\}$ est le coefficient de corrélation complexe entre deux processus aléatoires stationnaires (à l'ordre 2) x et y , avec $E\{\cdot\}$ l'espérance mathématique et x^* le conjugué de x . Le module et l'argument de γ sont respectivement appelés degré de cohérence et différence de phase effective [TOU99]. L'application pratique à deux séries temporelles discrètes nécessite de faire l'hypothèse d'ergodicité des processus et de redéfinir l'intercorrélacion comme : $\Gamma_{xy}(m) = \sum_n x(n)y^*(n-m)$. Dans ces conditions, et sur une durée finie de N échantillons, $|\gamma|$ est un estimateur de la cohérence et s'écrit:

$$|\gamma| = \left| \frac{\sum_{n=1}^N x(n)y^*(n)}{\sqrt{\sum_{n=1}^N |x(n)|^2} \sqrt{\sum_{n=1}^N |y(n)|^2}} \right|$$

La cohérence a souvent été discutée en imagerie SAS puisque la synthèse d'ouverture repose sur la cohérence physique entre deux émissions sonar (cf. [HAY09] pour une

revue). Dans le cadre de la détection cohérente de changement, nous mesurons localement la cohérence physique du fond entre les imagerie I_1 et I_2 de chacune des passes (chaque matrice est mise sous une forme vectorielle).

L'idée de l'application de la mesure de cohérence à la détection de changement entre deux passes sonar date de 2010 environ [STE09] [LYO10] [GMI10]. La détection d'un nouvel objet peut théoriquement se faire sur perte de la cohérence estimée localement. Toutefois, elle est conditionnée par la bonne cohérence globale du fond entre les deux passes qui peut être affectée par la décorrélation temporelle du speckle [RIG93] [LYO13]. En effet, une décorrélation du speckle intervient si la distribution des diffuseurs du fond a changé. Ce phénomène peut être rapide sous l'effet des courants marins qui modifient la couche sédimentaire au cours du temps. Bien entendu, cette réflexion ne tient que si l'on est certain de ne pas subir une décorrélation spatiale, c'est-à-dire que l'acquisition des données s'est faite selon des points de vue proches : faible décalage des antennes relatives aux deux passes dans les trois dimensions [BRO14]. La première chose à faire avant de tenter une détection de changement est donc de s'assurer de la bonne cohérence générale entre les données des deux passes.

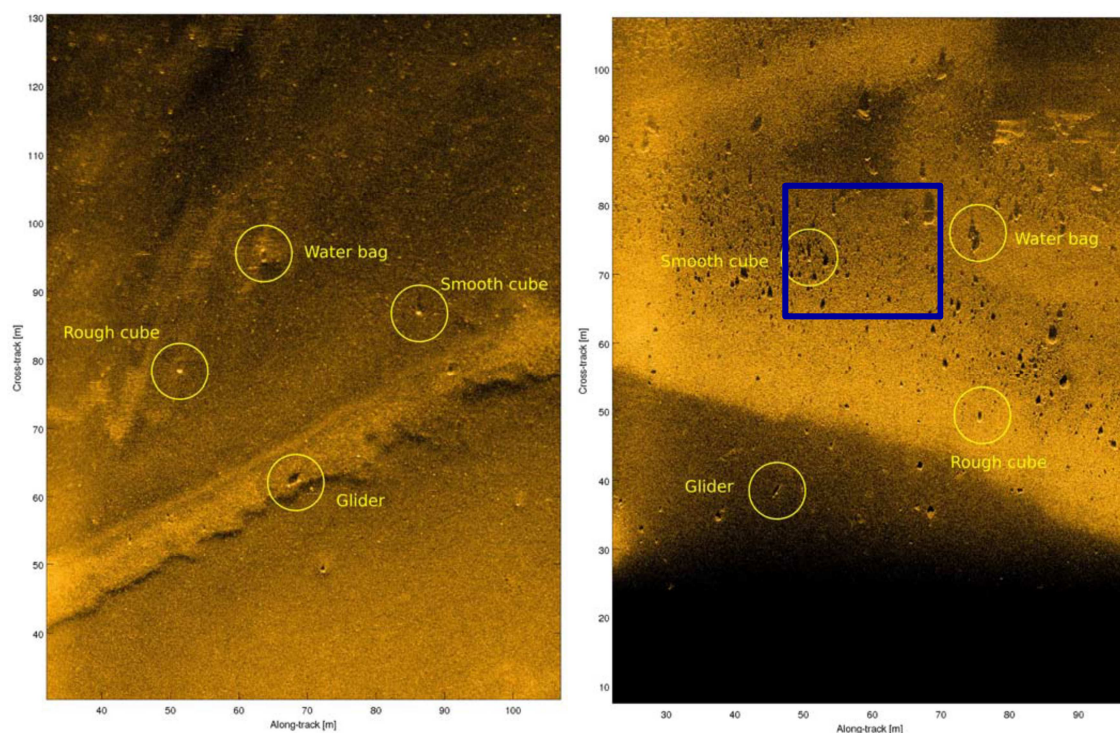


Figure 13 – Données HISAS 1030 (Avril 2011) : un exemple de fond « facile » à gauche et un exemple de fond « complexe » à droite, avec la nature des objets déployés (tiré de [MID11]).

Les données utilisées pour ces travaux ont été collectées en avril 2011 lors d'une campagne norvégienne co-organisée par FFI (Norvège), DRDC-A (Canada) et NRL (USA), au large de Larvik, en Norvège. L'AUV norvégien HUGIN 1000-HUS AUV est

équipé du système HISAS 1030 interférométrique qui permet d'obtenir des images SAS de résolution inférieure à 4cm à une fréquence de 100kHz. Pour chaque image traitée, la zone couverte est d'une centaine de mètres en portée pour une trace de 75m environ. Les images ont été acquises sur deux types de fond : un fond relativement facile et un fond complexe avec du cailloutis (cf. Figure 13). L'intervalle de temps s'échelonne de 2 à 8 jours, ce qui peut être un peu long pour conserver une cohérence sur certains jeux de données. Huit objets ont été placés : 2 cubes en béton (de 20cm de côté), l'un lisse, l'autre rugueux, un *glider*, une poche d'eau lestée de sable, 2 cônes tronqués type mine Manta et 2 cylindres de taille différente.

Pour chaque passe, la position de l'antenne physique ping par ping était disponible. Nous avons donc pu calculer le décalage en hauteur, l'embarquée et le cavalement entre les deux passes, ceci est illustré pour un jeu de deux passes sur la Figure 14. D'après une étude théorique faite par Michel Legris de l'ENSTA Bretagne, la cohérence s'avère bien plus robuste à l'embarquée qu'au cavalement. Cette conclusion est basée sur la formulation des limites d'embarquée et de cavalement avant une perte de corrélation de 3dB. Ainsi, la limite sur l'embarquée évolue avec le carré de la longueur de l'antenne alors que la limite sur le cavalement est linéaire avec cette longueur. La Figure 15 donne, pour les spécifications du HISAS 1030 et une altitude de 15m, une idée des écarts admissibles en fonction de la distance avant une perte de corrélation de 3dB. Compte tenu du fait que les formules théoriques de décorrélation ont été développées pour une antenne physique, nous avons calculés les écarts pour une longueur d'antenne physique « équivalente » qui permet de maintenir la résolution transversale de l'antenne synthétique du HISAS 1030 sur toute la portée, ce qui explique l'accroissement de la longueur de l'antenne avec la portée sur le graphe du bas de la Figure 15. Si l'on compare Figure 14 et Figure 15, seul le cavalement mesuré est d'un ordre de grandeur proche du cavalement maximal et ce, en début de portée seulement. Cependant il faut avoir à l'esprit que la technique de synthèse d'ouverture exige en amont une continuité des capteurs et que par conséquent, le cavalement est a priori moins problématique pour les images SAS que pour les images conventionnelles. Ces analyses, malgré les approximations faites, nous montrent que d'un point de vue géométrique, la cohérence devrait être maintenue entre les deux passes dont nous disposons.

Les travaux de recalage élastique et détection de changement cohérente ont fait l'objet d'une communication dont l'article est donné en annexe page 121. Par conséquent, dans la suite de ce paragraphe, nous ne restituons qu'une partie des éléments.

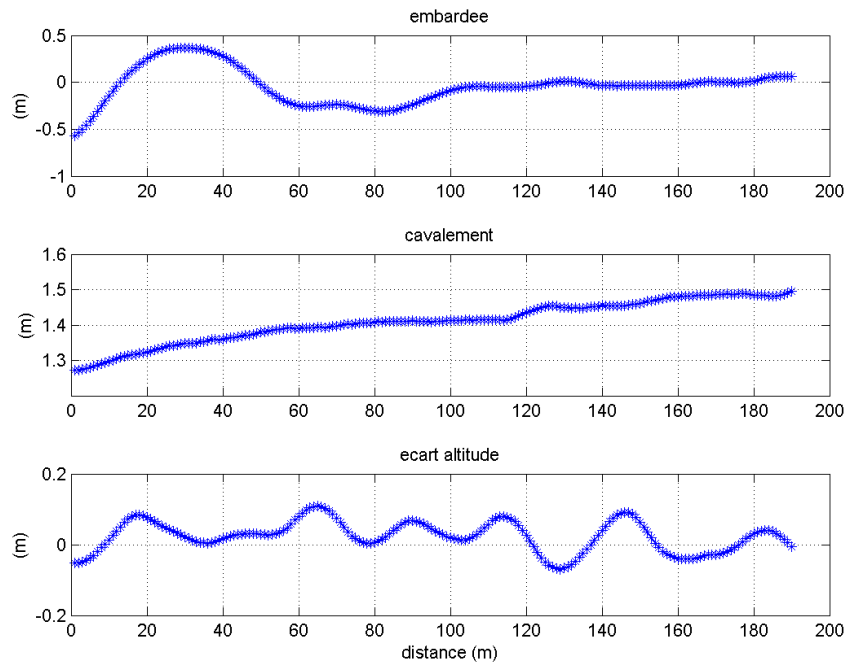


Figure 14 - Ecart entre les antennes physiques des deux passes sur la trace utilisée pour former les images SAS

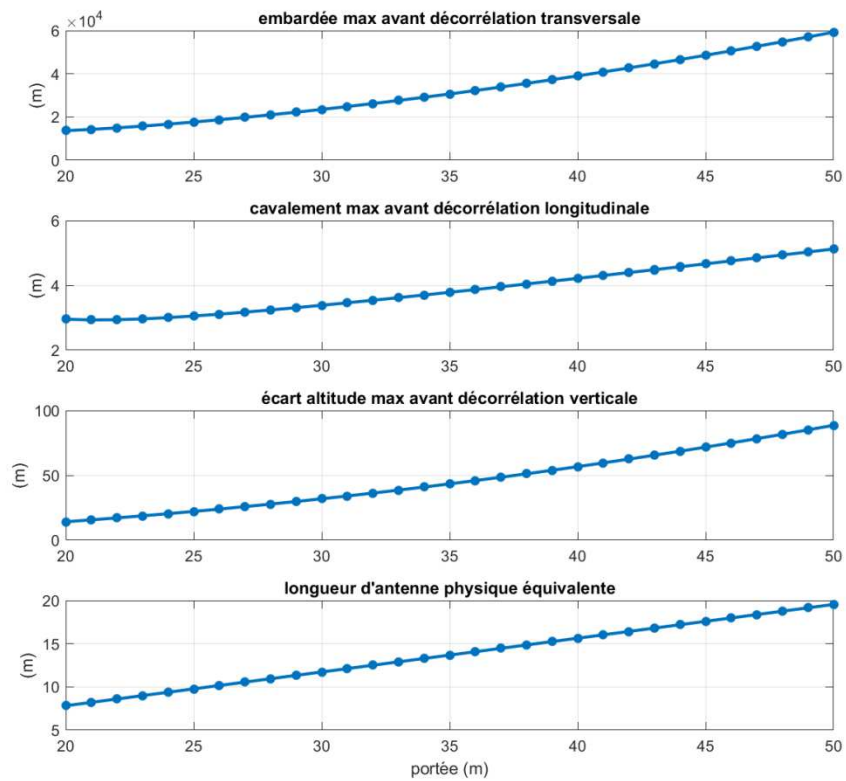


Figure 15 - Ecart admissible avant décorrélation en fonction de la portée pour une antenne physique de longueur équivalente garantissant les performances du HISAS 1030 sur toute la portée

Les jeux de données fournis par nos collègues norvégiens avaient déjà été recalées de manière rigide par extraction sur chaque image de points caractéristiques grâce à l'algorithme SURF et appariement de ces deux ensembles de points [MID11]. La translation ainsi trouvée est appliquée à la correction de navigation utile pour régénérer la seconde passe SAS, également nommée passe répétée [SAE11b]. Notre contribution débutait donc à ce stade avec deux étapes : un recalage élastique et subpixelaire suivi de la détection cohérente de changement. Le recalage élastique consiste tout d'abord à détecter des points de contrôle par corrélations (de données complexes) subpixelaires locales entre l'image référence de la première passe et l'image esclave de la passe répétée. Le champ de déplacement résultant est ensuite utilisé pour estimer au sens des moindres carrés à l'aide de B-splines hiérarchiques le champ de déformation à appliquer à l'image répétée : minimisation de l'erreur entre la fonction d'approximation (somme pondérée de bases B-splines) et les points de mesure et ce, à différentes échelles. La fonction d'approximation permet de déplacer chaque pixel de l'image esclave afin de corriger la déformation mais un post-traitement par interpolation est nécessaire entre les points de la grille de destination.

Pour mesurer la qualité du recalage, trois matrices, calculées à partir des images recalées, illustrent respectivement :

1. la différence de phase (calculée ici pixel à pixel mais un filtrage 2D peut réduire le bruit [SAE13]),
2. la log-image (introduite au paragraphe 3.2.2.2.1) (pixel à pixel),
3. la cohérence (calculée localement à l'aide d'une fenêtre glissante).

L'intérêt du sonar latéral interférométrique par rapport au sonar latéral est de fournir un moyen d'estimer la bathymétrie du fond grâce à une seconde antenne. La différence de phase (soit l'argument de γ défini plus haut) entre les signaux reçus sur deux capteurs distincts (ou deux signaux de voie formée) dépend de la différence trajets, appelée différence de marche, entre le réflecteur du fond à l'origine des signaux et chacun des capteurs, de sorte qu'un « tour » de phase de 2π correspond à une différence de marche égale à la longueur d'onde. Cette différence de marche évolue rapidement sur les premiers mètres puis de plus en plus lentement jusqu'à atteindre une asymptote. Sur un fond plat, cela se matérialise par une rotation de phase, c'est-à-dire une différence de phase qui ondule de manière continue entre $-\pi$ et $+\pi$ avec une période spatiale de plus en plus grande. Comme pour le sonar latéral, la base interférométrique d'un SAS permet d'estimer à chaque passe la bathymétrie de la zone et cette donnée est quasi indispensable à la synthèse d'ouverture effectuée à partir d'un véhicule suivant une trajectoire non parfaitement rectiligne (c'est la raison des performances élevées des systèmes HISAS 1030 norvégien et SAMDIS français). Contrairement à un sonar latéral

(ou SAS) interférométrique, les deux antennes d'une base interférométrique multi-passe ne sont pas solidaires et leur écart, soit la longueur de la base (*baseline* en anglais), évolue en 3 dimensions : en altitude, dans le sens de déplacement (le cavalement) et perpendiculairement au déplacement (l'embarquée). La différence de phase, estimée cette fois à partir des images recalées et due à la différence de marche résiduelle des deux passes, est par conséquent difficile à prédire ou à imputer à la bathymétrie du fond sauf à connaître de manière très précise la navigation du porteur lors de chacune des passes ainsi que la bathycélérité (une étude séparée et approfondie de chacun de ces effets est une perspective). La rotation de phase ne s'observe alors plus uniquement selon l'axe des distances mais ondule suivant des motifs complexes. En effet, le recalage parfait (*i.e.* celui qui remet les signaux de capteurs en phase) est impossible à obtenir entre deux passes, d'une part parce que le traitement SAS (formation de voies synthétique et DPCA) mélange les signaux pour une célérité donnée et éventuellement une bathymétrie estimée par interférométrie classique donnée, d'autre part parce que les erreurs de navigation et la modification du milieu (fond, célérité) sont des phénomènes impossibles à séparer. Ainsi une rotation de phase est bien visible si les images des deux passes sont bien recalées. On peut simplement supposer que les ondulations lentes de la différence de phase sont plus probablement dues à la différence de marche entre les deux antennes (sauf si le fond est en pente régulière) tandis que les ondulations rapides à des macrostructures du fond (un peu à l'instar d'une image floutée dont on ne retient que les hautes fréquences pour en lire le contenu informationnel).

Ainsi, un recalage réussi entre deux passes n'ayant subi aucun changement doit faire apparaître d'une part, dans l'image de différence de phase une pseudo-rotation de phase liée à l'écart physique des deux antennes, d'autre part une log-image de faible niveau de gris et enfin, un niveau de gris fort pour la cohérence.

La Figure 16, correspondant à la zone symbolisée par un rectangle bleu sur la Figure 13 (moyennant une rotation de 90°), illustre l'effet du recalage élastique par rapport au recalage rigide selon trois critères : la différence de phase entre l'imagette référence et l'imagette esclave, la log-image et enfin, la cohérence (dont une baisse locale permet de détecter un changement). Les résultats de la Figure 16 font apparaître, après recalage élastique : en haut, la rotation de phase est plus nette (moins bruitée) ; au milieu, une log-image plus propre facilitant la détection de la cible repérée par un fort niveau local (méthode de détection de changement incohérente présentée au paragraphe 3.2.2.2.1) ; en bas, une cohérence globalement plus élevée qui met bien en évidence la perte locale de cohérence liée à l'apparition de la cible.

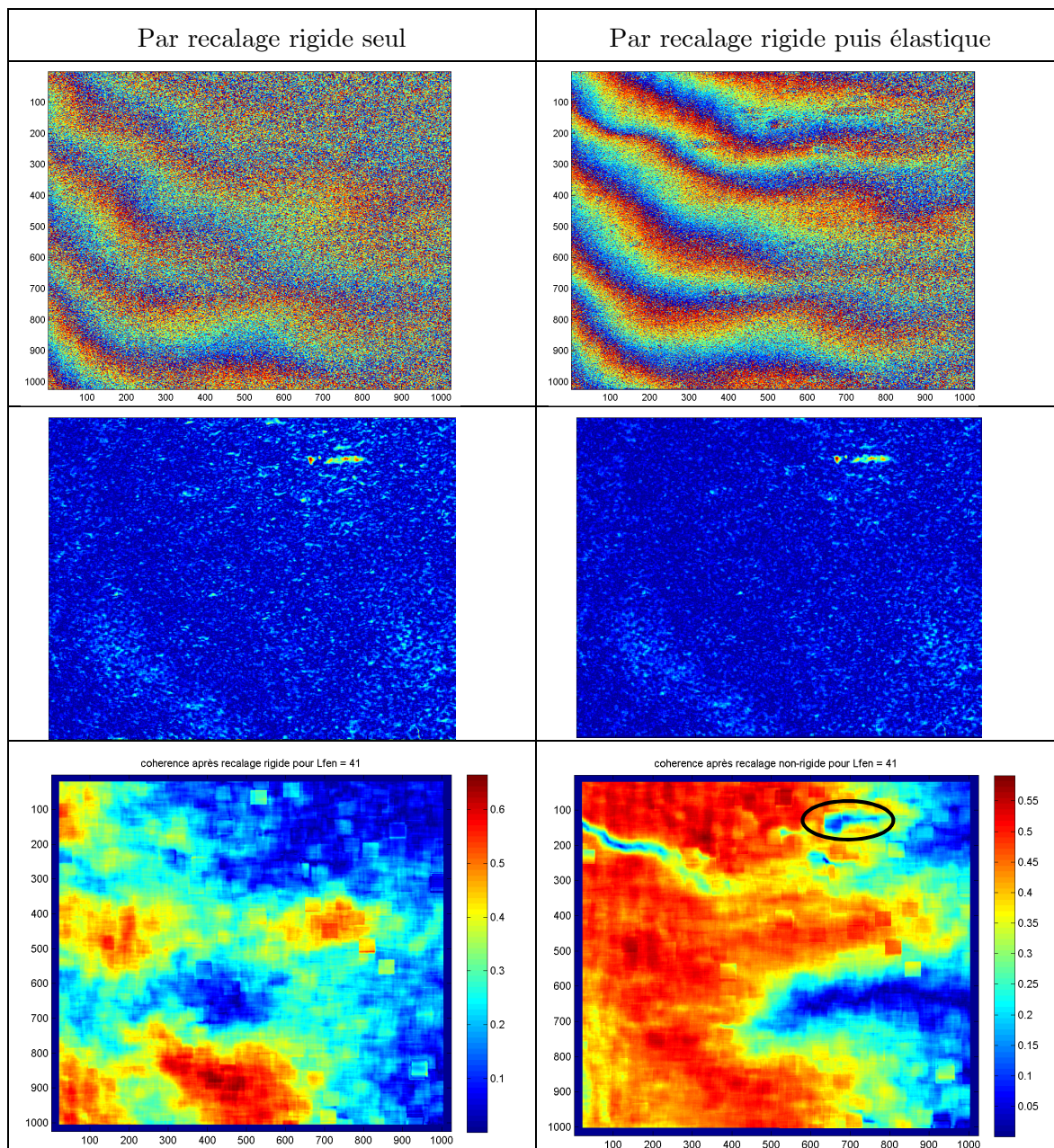


Figure 16 – Effet du recalage élastique sur la différence de phase et les détections incohérente puis cohérente de changement (de haut en bas) ; la position de la cible apparue entre les deux passes est indiquée par l'ellipse noire.

Ces résultats bien que novateurs et encourageants montre encore une perte résiduelle de cohérence que l'on ne peut imputer de manière certaine à une décorrélation d'origine naturelle (temporelle ou type de sédiment par exemple) et ce, malgré une tentative d'amélioration du recalage par un algorithme itératif [MYE13], également validé sur des données du SAS interférométrique AquaPix de Kraken Sonar Inc. [DIL14]. Il nous fallait clarifier ce point en étudiant davantage ce type de données. C'est tout l'objet des travaux de thèse de Vincent Myers.

3.2.2.3.2 *Travaux du doctorant Vincent Myers*

Vincent Myers, scientifique au DRDC-A, a débuté sa thèse intitulée « Le traitement, l'interprétation et l'exploitation d'images sonar à antenne synthétique obtenues à partir de trajectoires répétitives » fin 2015 sous la direction de Benoît Zerr. Du fait de notre collaboration passée, j'ai le plaisir de co-encadrer ses travaux dont l'objectif principal consiste à mieux comprendre et exploiter les données SAS multi-passe dites RP-SAS (Repeat-Pass Synthetic Aperture Sonar). Le traitement RP-SAS est un champ de recherche relativement nouveau et concerne le traitement cohérent de données SAS acquises à partir de trajectoires ou passes répétitives sur une même zone [HUN09] [SYN10] [SAE11a] [LYO13]. Comme pour un sonar latéral interférométrique (c'est-à-dire formé de deux antennes parallèles), il s'agit de combiner les données capteurs ou SAS de chaque passe afin d'obtenir des informations sur une zone impossibles à obtenir ou imperceptibles lors d'une seule passe. Exploiter l'amplitude et la phase des données RP-SAS permet la détection de changement par perte de cohérence comme nous l'avons vu au paragraphe précédent pour détecter des objets mais également des changements subtils de l'ordre du mouvement de sédiment ou encore du changement de célérité dans la colonne d'eau. Certains auteurs pensent que la création d'une carte interférométrique à partir des deux passes SAS peut également être exploitée pour obtenir une bathymétrie de précision meilleure qu'avec deux passes réalisées avec un sonar latéral conventionnel (du fait d'une résolution latérale meilleure pour le SAS) [HUN16]. Comme nous l'indiquons plus tôt au sujet de la différence de phase des passes recalées, cette bathymétrie n'est envisageable que si le milieu est parfaitement stable (célérité, morphologie du fond) et moyennant une navigation parfaitement maîtrisée (comme avec deux antennes solidaires) ou réestimée si l'on a accès aux données capteurs brutes [SAE11b] [HUN16].

Le traitement des données RP-SAS exige donc de fortes contraintes sur la navigation dont les erreurs peuvent s'accumuler significativement au fil des passes et par des écarts entre deux passes (cavalement) ou mouvements de la plateforme (lacet d'après [HUN16]) trop importants. Des baisses ponctuelles de cohérence difficiles à expliquer limitent encore aujourd'hui la détection de changement à partir d'images SAS préalablement recalées. Des explications possibles sont l'utilisation d'une fonction de déformation inadaptée ou une conséquence de la méconnaissance du relief du fond sur lequel les données ont été acquises (si on ne dispose pas d'un sonar interférométrique et que les trajectoires des antennes sont chahutées). Deux pistes ont été proposées dans cette thèse, toujours sur la seule base des images SAS: utiliser des paramètres permettant de mesurer la qualité du recalage et donc des déformations estimées dans un contexte multi-look d'une part [MYE17b] ; aborder le problème de manière plus globale, des images elles-

mêmes au traitement synthétique sous-jacent, en estimant les éléments résiduels de correction à la navigation permettant un recalage exact d'autre part.

Pour la première approche, L images SAS de résolution latérale plus basse sont générées en filtrant les données dans le domaine du nombre d'onde de l'image SAS pleine résolution (voir Figure 17 extraite de [MYE17b] avec $L=4$) de sorte à découper l'ouverture synthétique en L sous-ouvertures. D'un point de vue incohérent d'une part, le *speckle* peut être considéré comme décorrélé entre ces sous-images ce qui permet d'obtenir une image débruitée par sommation. Il en résulte une corrélation plus forte des images des deux passes ainsi traitées. D'un point de vue cohérent d'autre part, sur les zones de rides observées, un gain de cohérence est observé si on garde le maximum de cohérence par paire de sous-images. La Figure 18 illustre le bénéfice des deux méthodes sur des données RP-SAS du système Vision 2000 développé par Atlas Elektronik : deux cibles déposées entre les deux passes sont mises en évidence avec, pour l'image de cohérence, la possibilité de détecter un changement subtil de sédiment à proximité des objets, invisible à l'œil nu.

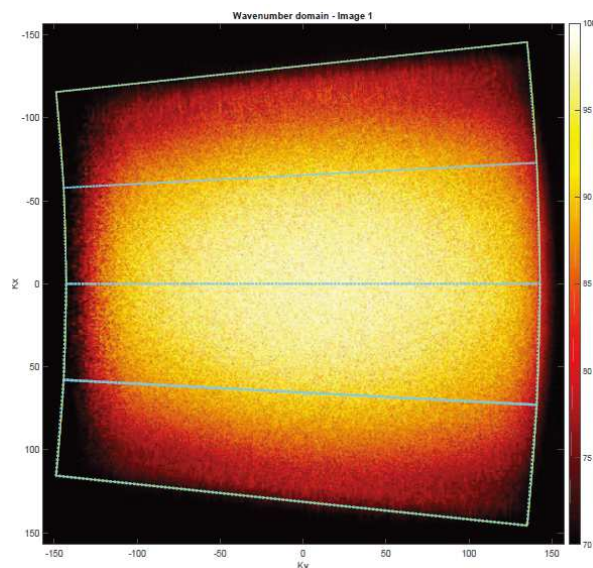


Figure 17 – Spectre dans le domaine du nombre d'onde d'une image SAS et découpage en 4 sous-bandes (extrait de [MYE17b])

Pour la seconde approche, l'idée est de s'inspirer du principe de l'algorithme DPCA (Displaced-Phase Centre Array [BEL02]) utilisé classiquement pour la formation d'une seule image SAS et de l'appliquer à l'ensemble des données acquises durant les deux passes différentes. Cette technique a été utilisée très récemment sur les données capteurs brutes (voir [HUN16]) mais, outre le temps de calcul considérable nécessaire pour faire du DPCA sur les deux ensembles de données, ces dernières ne sont pas toujours conservées (car encombrantes) et par suite indisponibles pour les post-traiter. La démarche proposée par Vincent Myers est donc d'inverser le traitement SAS pour recalculer

géométriquement les deux trajectoires et estimer la différence de vitesse de la plateforme afin de fournir les corrections nécessaires (sur les centres de phase de la seconde passe) à un recalage parfait de la seconde passe sur la première (un article a été soumis à IEEE Journal of Oceanic Engineering sur ce sujet, à la suite de résultats déjà présentés à la conférence de l'ASA en juin 2017 [MYE17a]). Ces approches sont supposées meilleures que les approches de traitement d'images, comme l'approche par déformations itératives de l'image utilisée précédemment dans [QUI12b], parce qu'elles autorisent cette fois une déformation de l'antenne et ce, en se basant sur des principes physiques qui facilitent une interprétation des résultats.

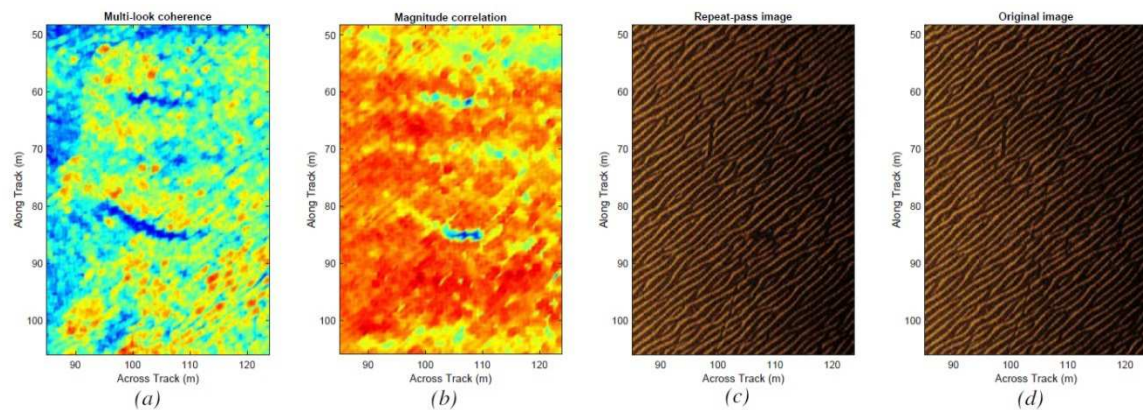


Figure 18 – Mise en évidence de l'apparition de deux objets par perte de cohérence (a) et corrélation (b) entre la première passe (d) et la seconde (c) (tiré de [MYE17b])

Un objectif plus général de cette thèse est de modéliser et expliquer les phénomènes physiques qui permettent le traitement des images RP-SAS. En ce sens, la seconde partie de la thèse est consacrée à la modélisation par TDFD (Time-Domain Finite Difference) du champ acoustique pour avoir une meilleure compréhension de la cohérence entre deux signaux et des effets de changement de célérité, nature de sédiment, relief... sur celle-ci.

3.2.3 Un point délicat : dimensionnement de la fenêtre de calcul

Dans les paragraphes précédents 3.2.1 et 3.2.2, les résultats sont conditionnés au réglage, trop souvent empirique, de la taille d'une certaine « fenêtre de calcul ».

Pour la détection de cibles ou la qualification des fonds (paragraphe 3.2.1), la difficulté est de fixer une taille de fenêtre adaptée à la dimension des 'structures' recherchées. D. Williams calcule la lacunarité en chaque pixel sur un voisinage de dimension $2m \times 2m$. Comme pour le détecteur d'anomalie décrit au paragraphe 3.2.1.1, D. Williams justifie cette dimension de manière empirique : « This size is chosen to be large enough to capture distinguishing (textural) characteristics of the seafloor but small enough to respect the short length scale at which seafloor conditions are known to vary » [WIL15]. Beaucoup plus tôt, R. Kessel a proposé de définir de manière objective la taille de la

fenêtre de calcul en fonction de la surface de la structure recherchée et de la proportion de pixels appartenant à la structure dans cette fenêtre, elle-même dépendant du contraste entre les pixels de la structure et le fond [KES02]. Bien que fondée théoriquement sur les travaux de T. Stanton [STA84] sur l'effet de la microstructure du fond sur le signal sonar, cette définition exige des informations a priori qui, en pratique, en font une approche empirique également limitée à la taille des structures recherchées. L. Picard utilise pour sa part un modèle simple de structure sous-marine pour dimensionner son espace multi-échelle. La réponse maximale des filtres DoP lui a permis de relier les paramètres de définition du filtre et la dimension de la structure à mettre en évidence. Ainsi, il garantit que son espace d'échelle balaie l'ensemble des objets depuis une dimension minimale (0.25m) jusqu'à une dimension maximale (2.5m). L'intervalle [0,25m ; 2,5m] est discrétisé par découpage en 3 octaves des filtres DoP, elles-mêmes divisées en 4 sous-bandes fréquentielles, soit 12 échelles [LOW99]. Le seul réglage empirique est le choix des 4 sous-bandes par octave, nombre plutôt élevé mais qui permet d'obtenir une description fine de l'information contenue dans l'image, quitte à avoir certaines redondances. Pour le recalage de deux passes sonar (paragraphe 3.2.2), un réglage empirique est généralement utilisé mais J. Groen a récemment étudié l'influence du paramétrage du recalage (y compris la dimension de la fenêtre) sur la détection de changement. Il conclut de manière surprenante en « the optimum block size is not a constant, *i.e.* image registration using adaptive block sizes is advantageous » [GRO17]. Cette communication à la quatrième édition de la conférence Underwater Acoustics n'a malheureusement pas fait l'objet d'un article.

Ce point délicat d'analyse des données me semble mériter une étude approfondie. La recherche de critères objectifs est essentielle pour valider toute démarche scientifique et conclure mais curieusement, certains aspects comme le dimensionnement de la fenêtre de calcul restent à l'appréciation du développeur. Fonctionner « à tâtons » biaise inévitablement le problème. Les quelques exemples donnés ci-dessus montrent que le dimensionnement doit être guidé à la fois par la physique des phénomènes en jeu, la géométrie d'acquisition (formation de l'image) et enfin les éléments à mettre en évidence. Lors de mes travaux de thèse par exemple, l'optimisation d'un filtre bidimensionnel visant à bimodaliser l'histogramme d'une imagerie sonar centrée sur une cible présentant une ombre portée a montré que la dimension optimale de la fenêtre était exactement liée au suréchantillonnage de la cellule de résolution sonar [QUI01b]. Ce fut une manière inattendue mais objective de remonter à la résolution sonar à partir des données.

3.3 Détection d'obstacles pour véhicule sous-marin autonome

Les véhicules autonomes sont supposés pouvoir réaliser une mission dans un environnement inconnu. La mission principale d'un AUV de REA (Rapid Environment Assessment) par exemple est de collecter en toute discrétion des données sur son environnement, généralement à l'aide d'un sondeur multifaisceau ou d'un sonar latéral et de ramener ces données au navire depuis lequel il a été déployé. En d'autres termes, il faut l'équiper d'un système anti-collision lui permettant d'assurer seul sa sécurité pendant la phase critique d'autonomie depuis la plongée jusqu'à la reprise de vue. Grâce à ce système naturellement placé sur son 'nez', l'AUV doit pouvoir détecter assez tôt tout obstacle dans sa trajectoire qui pourrait l'endommager. Compte tenu de la distance minimale de sécurité d'une dizaine de mètres, l'acoustique est privilégiée et le système dédié à cette tâche est appelé Obstacle Avoidance Sonar (OAS). Malgré une résolution sonar moindre qu'un sonar latéral, l'OAS doit ainsi permettre la détection et le pistage de l'obstacle, de même que l'extraction des informations nécessaires à une manœuvre d'évitement. Cette capacité à détecter les obstacles dangereux intervient à toutes les immersions possibles, près du fond jusque la surface. La configuration du sonar peut bien sûr être adaptée pour maximiser le champ de vision dans le cas de la reprise de vue ou bien détecter au mieux les changements de bathymétrie dans le cas du suivi de fond.

L'ENSTA Bretagne a un lien privilégié avec sa tutelle en ce qui concerne l'enseignement mais également la recherche. DGA Techniques Navales Brest (anciennement GESMA) nous a tout naturellement sollicités pour soutenir le développement d'algorithmes de perception pour ses AUVs démonstrateurs. Cette collaboration a commencé de manière informelle par une invitation à une campagne de quatre jours d'essais en mer en avril 2006 à bord du BEGM « La Thétis » afin d'évaluer « à chaud » la qualité des données de l'OAS Reson Seabat 8101 pour leur démonstrateur Redermor (voir une copie écran de l'interface développée à cette occasion sur la Figure 19). Cette collaboration s'est formalisée par deux de contrats de recherche de 3 ans chacun sur la problématique de la détection d'obstacles pour AUV. Un premier contrat dès 2007 fut consacré aux « Etudes sur le traitement de données acquises par de nouvelles méthodes de perception sous-marine pour la chasse aux mines et la détection d'obstacles » dans la colonne d'eau. Le second contrat débuté en 2014 visait la « Détection et suivi d'obstacles pour un AUV de REA Discret » du fond jusque la surface. Pour ces contrats, les données ont été majoritairement collectées par le démonstrateur DAURADE, véhicule développé par le SHOM en coopération avec DGA TN et destiné à réaliser des missions diverses de REA en autonomie.

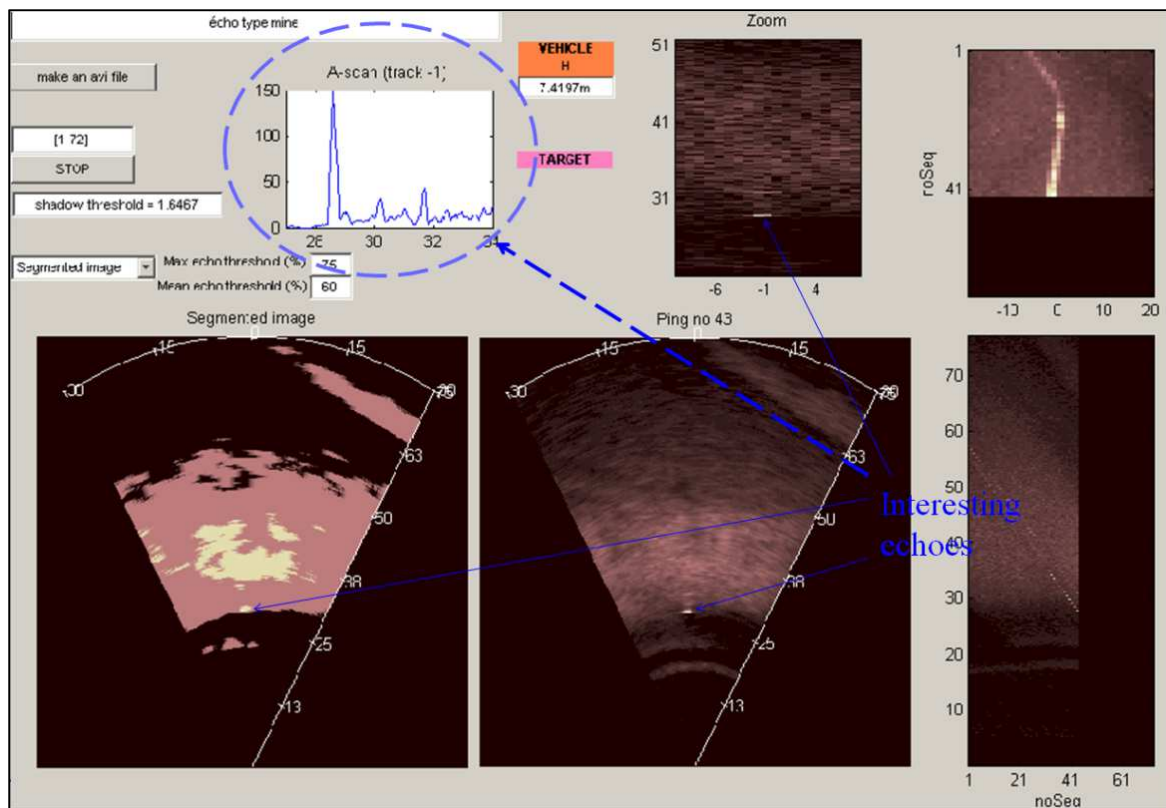


Figure 19 – IHM développé pour la visualisation et le traitement rapide des données Reson Seabat 8101 – cas de la mine à orin avec deux échos ponctuels (En bas : image segmentée à gauche, image brute au centre, historique des échos maximaux par ligne distance dit « échogramme » à droite ; En haut : informations sur la position du véhicule et l’obstacle dont l’amplitude sur une voie dit « A-scan » à gauche, zoom au centre, historique des échos maximaux par voie) [QUI07]

3.3.1 Détection, suivi et caractérisation des obstacles dans la colonne d’eau

Ces travaux ont fait l’objet d’un article publié dans IEEE Journal of Oceanic Engineering en 2012 donné en annexe page 123 [QUI12c]. Par conséquent, dans ce paragraphe, nous ne donnons qu’une description des extensions/améliorations apportées par la suite, mais de manière brève étant donné leur confidentialité.

L’évitement d’obstacles vise à garantir la sûreté du véhicule au cours de sa mission. La plupart des OAS sont des sonars frontaux ou Forward Looking Sonars (FLS). Depuis les années 2000, la technologie « blazed array » que nous détaillerons plus loin a été largement utilisée et en particulier par son inventeur, la société BlueView. Le “Center for Autonomous Underwater Research” de la “Naval Postgraduate School, Monterey, CA” a beaucoup exploité les sonars BlueView entre 2000 et 2010 avec plusieurs « Master’s Thesis » consacrés à la robotique mobile sous-marine et tout particulièrement l’évitement d’obstacles [CHU03] [HEA04] [FUR06] [MAS08]. De notre côté, il s’agit plus exactement d’exploiter les séquences d’images sonar pour induire une manœuvre d’évitement.

Le FLS dédié à une détection dans le sens d'avance du véhicule n'a pas a priori de dépointage ou alors de seulement quelques degrés vers le bas ce qui permet d'intercepter le fond si l'altitude du porteur est assez faible. Ce système est configuré en mode horizontal c'est-à-dire qu'il a une bonne discrimination dans le plan horizontal (ouverture en azimut de 45 à 130 degrés selon les systèmes) grâce à la formation de voies mais aucune discrimination dans le plan vertical (ouverture en site de 15 à 30 degrés). Par conséquent, un obstacle est précisément repéré en distance et gisement (ou azimut) mais une forte incertitude persiste sur sa localisation dans le plan vertical. La Figure 20 schématise le champ de vision de l'AUV DAURADE et donne un exemple d'image sectorielle formée. La Figure 21 donne schématiquement le volume balayé selon les trois profils et les voies formées.

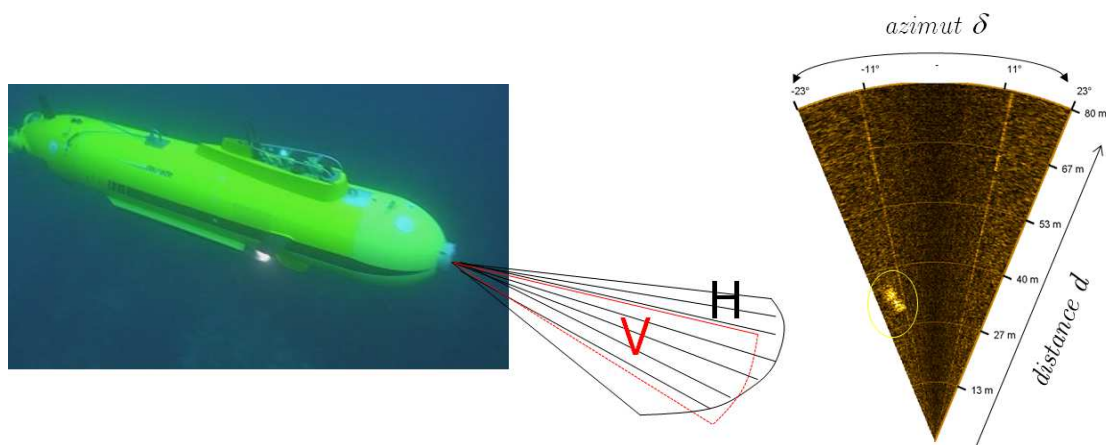


Figure 20 – A gauche : AUV DAURADE et champ de vision du FLS (H = horizontal, V = vertical) ; A droite : exemple d'image sectorielle obtenue par le FLS BlueView P450-45 où la signature d'un obstacle est entourée.

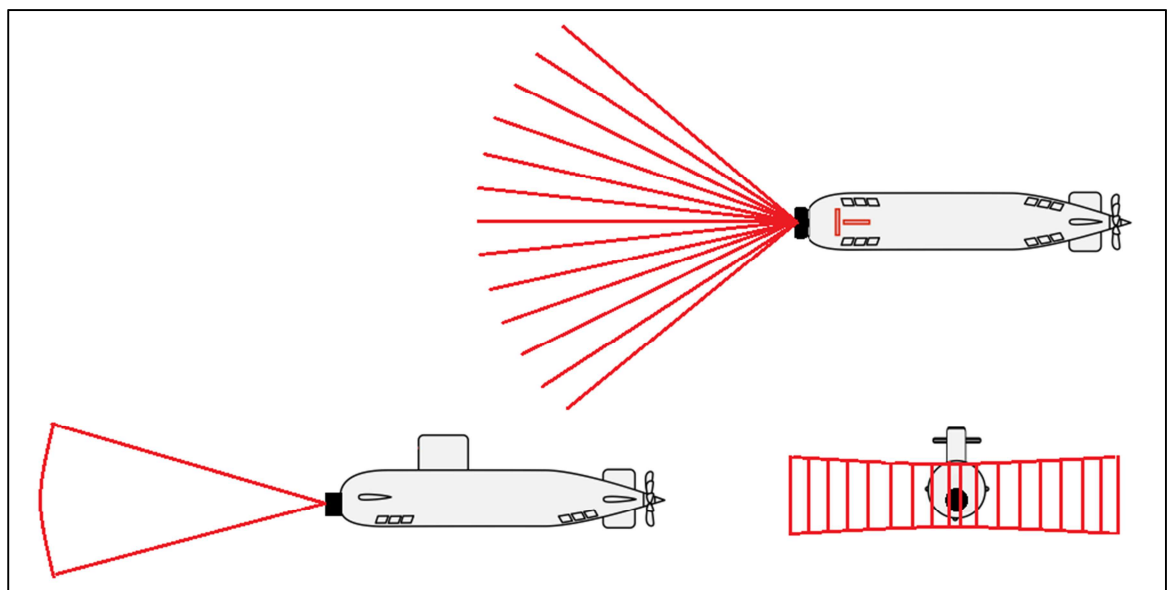


Figure 21 – Volume balayé par le FLS en configuration horizontale

L'objectif est de détecter, pister et éventuellement caractériser tout type d'obstacle dans la colonne d'eau de manière à induire une manœuvre d'évitement si nécessaire. Jusqu'aux années 2010, le pistage des obstacles sur une séquence d'images sonar frontal par filtre de Kalman était effectué sur la base des images seules, approche que l'on peut qualifier « d'approche signal » opposée à l'« approche robotique » que nous proposons. L'originalité de notre approche est en effet de prendre en compte les données de navigation, attitude (donnée de la centrale inertielle INS) et vitesse du véhicule (donnée du DVL), en entrée de l'équation d'état du modèle d'évolution du système afin de prédire à chaque récurrence sonar la nouvelle position de la cible dans l'image compte tenu de sa position précédente. Cette approche est donc robuste à un changement brusque de cap par exemple, ce qui intervient lors d'un virage de l'AUV sensé quadriller une zone. Cette position prédite est bien entendu corrigée ensuite par la mesure que constitue la nouvelle position détectée dans l'image et associée selon une approche classique type plus proche voisin au sens de la distance euclidienne entre détections courantes et positions estimées des obstacles pistés. Ce processus itératif de prédiction et correction est formalisé par un filtre de Kalman sans parfum compte tenu de la forte non linéarité du modèle d'évolution détaillé dans [QUI12c].

Pour la phase de détection, nous avons choisi de mettre en œuvre un test d'ajustement de manière à détecter une anomalie statistique, principe déjà évoqué (avec quelques ajustements liés à un arrière plan plus complexe et variable que la colonne d'eau que nous avons ici) au paragraphe 3.2.1.1. Ce test, qui utilise le coefficient de variation, est performant dans la colonne d'eau étant donné la stationnarité de l'arrière plan. Le coefficient de variation est théoriquement constant pour les images sonar d'amplitude en vertu des propriétés du bruit de *speckle* [SCH96]. Un écart absolu du coefficient de variation par rapport à la valeur théorique (0.523 pour une distribution de Rayleigh) observé localement indique la présence d'une anomalie, soit un obstacle potentiel.

Par rapport à l'article [QUI12c], et compte tenu de l'utilisation lors du second contrat d'un nouveau capteur d'évitement, le BlueView P900, un prétraitement des images a été nécessaire pour limiter le nombre de fausses alarmes liés à la technologie de ce capteur. La technologie particulière « *blazed array* » des capteurs BlueView utilise en effet des barrettes acoustiques large bande qui peuvent émettre et/ou recevoir de multiples faisceaux acoustiques angulairement dispersés [BAC04]. Contrairement aux sonars frontaux classiques qui font, à la réception, une formation de voies par retard et déphasage des signaux de capteurs de l'antenne pour pointer dans différentes directions, les sonars *blazed array* utilisent un principe similaire au réseau échelle de diffraction en optique pour décomposer le signal émis large bande en signaux bande étroite émis dans des directions propres. A la réception, le champ sonore multi-fréquentiel rétrodiffusé est recombinaison par la même technologie *blazed array* en un seul signal large bande et l'image

est formée par traitement temps-fréquence. Les voies « formées » ont ainsi chacune une fréquence et par conséquent une résolution propres.

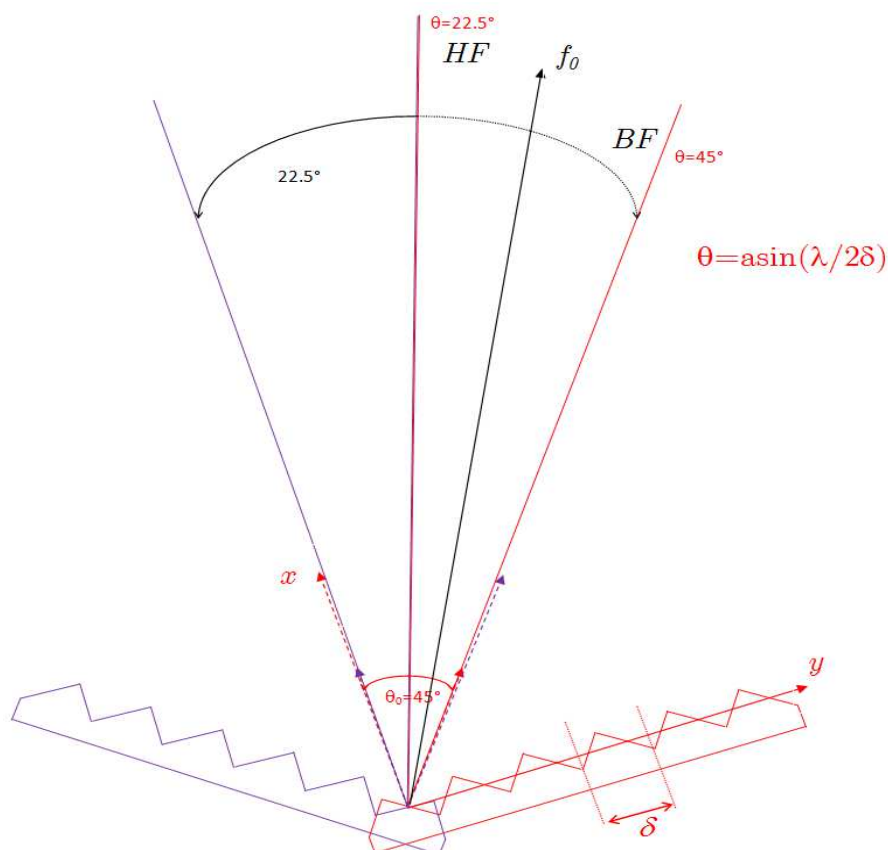


Figure 22 - Principe de formation de voies selon la technologie *blazed array* d'un secteur de $\theta_0=45^\circ$ avec deux barrettes « en V » (θ désigne les directions d'émission des signaux bandes étroites, λ la longueur d'onde, δ le pas intercapteur et f_0 la fréquence d'utilisation du sonar).

Le sonar BlueView P450-45 utilise deux barrettes acoustiques pour former un secteur de 45° horizontalement, chacune des barrettes couvre un secteur de 22.5° comme le schématise la Figure 22. Malheureusement, cette technologie induit deux effets délétères sur les images : d'une part une fluctuation du niveau moyen rétrodiffusé le long des voies (du fait du changement de fréquence), d'autre part la présence ponctuelle d'interférences propres ou rayonnées (une interférence signe sur l'ensemble des voies correspondant à la bande émise du capteur qui interfère avec le sonar BlueView et sur un intervalle de distance fonction de sa durée d'émission). La Figure 23 donne un exemple d'image fournie par le sonar BlueView P900-90 qui utilise la juxtaposition de quatre barrettes acoustiques donnant une dispersion fréquentielle de 600 à 1200kHz. On peut y noter la présence nette d'une interférence vers la basse fréquence (BF) et à 45m environ. Pour corriger ces effets, une normalisation selon les voies par division avec le profil moyen d'une part, et une détection puis neutralisation des interférences grâce à la présence de

forte corrélation pour une ou un ensemble de voies (de fréquences identiques) pour les quatre barrettes du P900 d'autre part, sont effectuées.

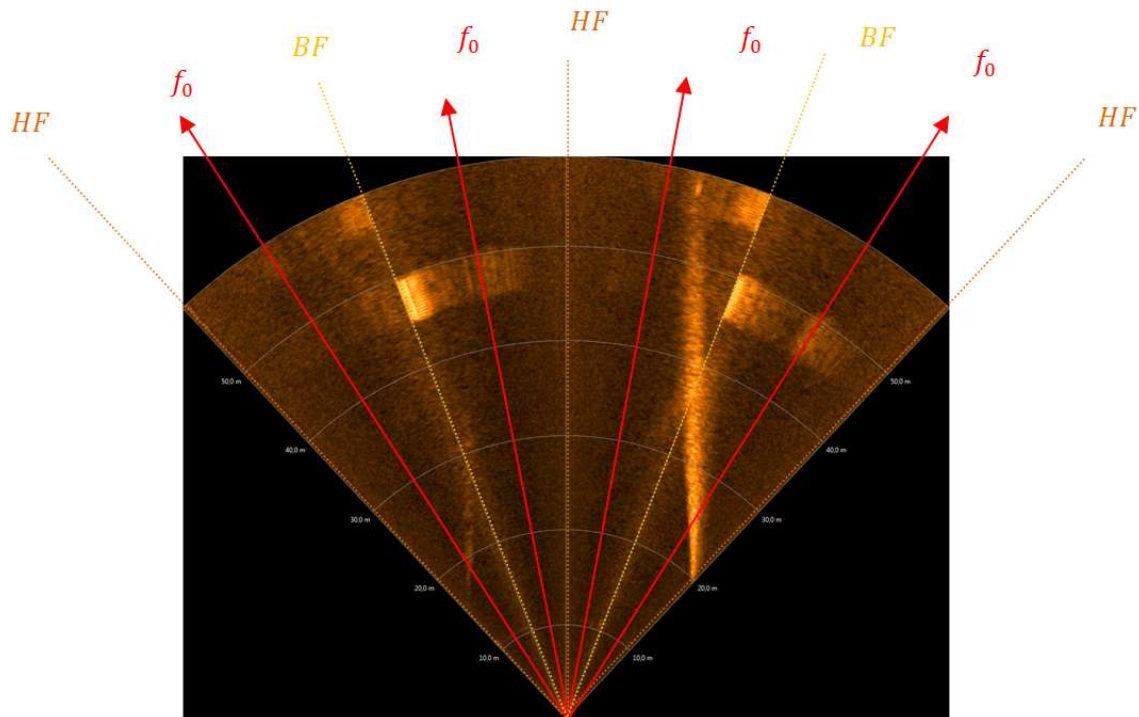


Figure 23 – Image sonar du BlueView P900-90 (mis en position verticale) de fréquence centrale $f_0=900$ kHz obtenue à partir des 4 barrettes acoustiques avec présence d’une interférence par un signal large bande d’un autre capteur porté par le véhicule DAURADE

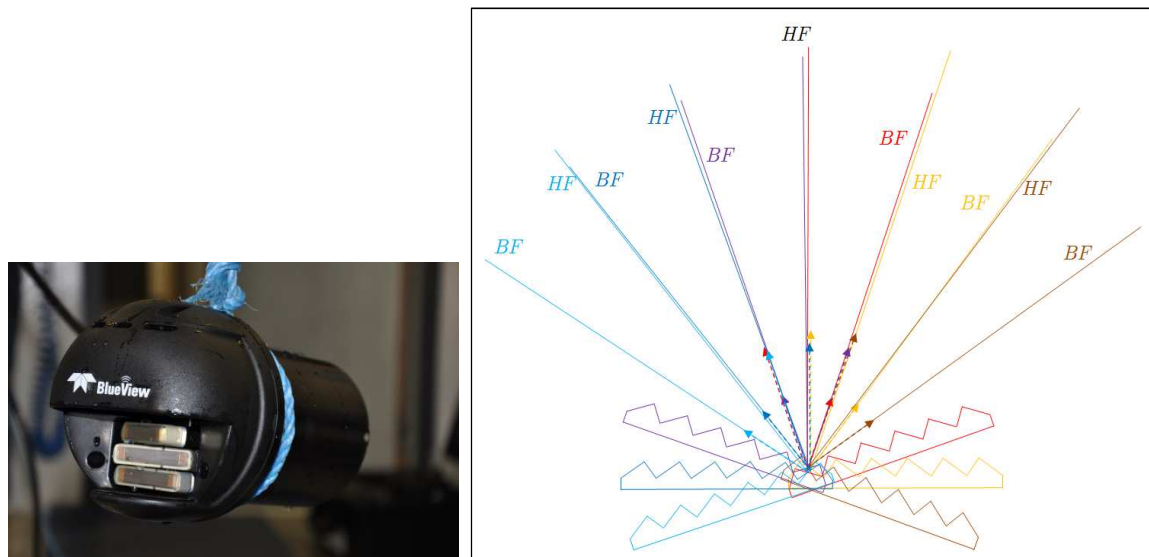


Figure 24 – Sonar BlueView M900-130 à gauche et un agencement possible des 6 barrettes (juxtaposées 2 à 2) à droite

Plus récemment, nous avons observé de près les barrettes haute fréquence du M900-130 et les premières images collectées en bassin : il s’avère qu’il n’y a plus de continuité en

fréquence entre deux voies voisines mais formées par deux barrettes distinctes (voir Figure 24). Une explication possible est une volonté du constructeur de limiter le risque de diaphonie (*cross-talk* en anglais) lié au couplage acoustique entre transducteurs. Ce nouvel agencement sera à prendre en compte pour adapter l’algorithme de neutralisation automatique des interférences.

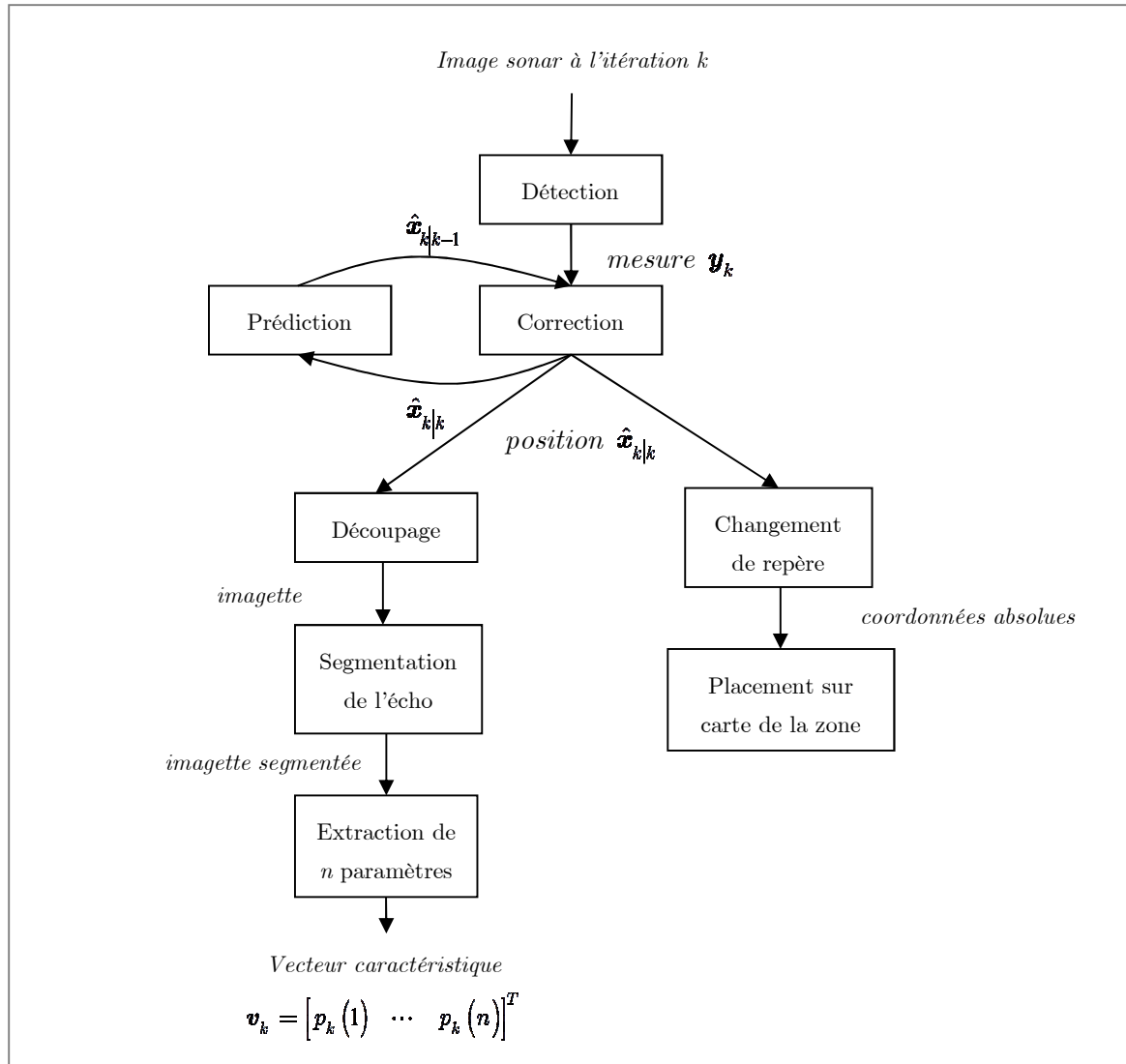


Figure 25 – Diagramme de la chaîne de traitement d’une image sonar frontal

La chaîne de traitement mise en œuvre pour chaque image de la séquence est résumée à la Figure 25 avec quelques améliorations apportées lors du second contrat de recherche. Le modèle d’évolution du système a tout d’abord été enrichi pour pouvoir estimer l’altitude de l’obstacle, en plus de ses coordonnées polaires (distance, azimuth). A la différence de la méthode d’estimation différée proposée dans l’article [QUI12c], l’estimation se fait ainsi en temps réel au fil de la séquence et de manière plus précise à mesure que la cible approche. La position de l’obstacle est alors tridimensionnelle malgré l’incertitude en site du capteur évoquée plus haut. La figure montre bien l’étape de

détection suivie du filtrage de Kalman (prédiction, correction) ainsi que le traitement de chaque imagerie découpée autour du point de détection afin de stocker quelques caractéristiques après segmentation de son écho. Une nouvelle exigence de DGA TN fut de fournir une cartographie de la zone prospectée. L'une des missions de DAURADE d'une durée de deux heures en octobre 2015 a été dédiée à cette tâche. Grâce à un changement de repère, la position absolue (longitude, latitude) de l'obstacle est reportée sur une carte afin de visualiser et trier l'ensemble des obstacles rencontrés. Un regroupement hiérarchique ascendant des détections est effectué *a posteriori* (hors ligne) afin de déterminer le nombre d'obstacles pertinents et les analyser sur la base des données de navigation (altitude du véhicule, revisite de l'obstacle) et de quelques caractéristiques visant à distinguer un obstacle manufacturé d'un obstacle naturel inoffensif comme un banc de poissons. La quasi-totalité des obstacles factices ont été retrouvés mais le traitement de la mission a mis en évidence d'autres objets non référencés. La nature contractuelle de ce travail ne permet pas d'en dire plus ici.

Dans tout ce travail, la cinématique des obstacles n'est pas prise en compte. Même si la majorité des obstacles que l'on rencontre sont fixes ou peu mobiles (cas des bancs de poissons sur quelques récurrences), le pistage des obstacles mobiles est un challenge intéressant. Une première piste, testée dans le cadre du projet RAPID MUSE (Multi Usage Sonar Element) avec la société iXblue financé par la DGA pour le développement d'un élément sonar polyvalent, a consisté à compléter le vecteur d'état avec les coordonnées du vecteur vitesse relative de l'obstacle. Malheureusement, l'augmentation du nombre de composantes du vecteur d'état et donc des degrés de liberté ne garantit pas des estimations fiables malgré un pistage correct de l'obstacle sur les images. Une autre solution a été proposée et validée lors d'un stage de Master 2 en 2013. Lors de ce stage, Laurent Picard a adopté une approche dite « track before detect » afin de détecter des trajectoires rectilignes entre deux eaux à partir d'une séquence d'images sectorielles par la mise en œuvre d'une transformée de Hough tridimensionnelle. La première contribution de ce stage fut le développement d'un algorithme de transformée de Hough en niveaux de gris pour capitaliser les données images au fil de la séquence. La seconde fut la prise en compte du comportement du véhicule pour accélérer le temps de calcul de la transformée tridimensionnelle malgré l'incertitude inhérente du sonar frontal en site. Ces résultats de stage d'un bon niveau ont fait l'objet d'une communication [PIC15a].

3.3.2 Reprise de vue du véhicule

La sécurisation de la manœuvre de reprise de vue d'un AUV par moyen acoustique est une thématique peu répandue dans la littérature. Quelques exemples désastreux de reprise de vue de sous-marins (par exemple, la collision entre *l'Ehime Maru* et *l'USS Greeneville* en 2001) et la multiplication des plateformes sur et sous l'eau démontrent pourtant la nécessité pour le véhicule de visualiser son environnement avant de refaire surface. L'AUV étant autonome, DGA TN a souhaité investiguer le potentiel d'un OAS incliné vers le haut, soit un Upward Looking Sonar ULS, pour sécuriser cette manœuvre. La Figure 26 donne schématiquement le volume balayé selon les trois profils et les voies formées.

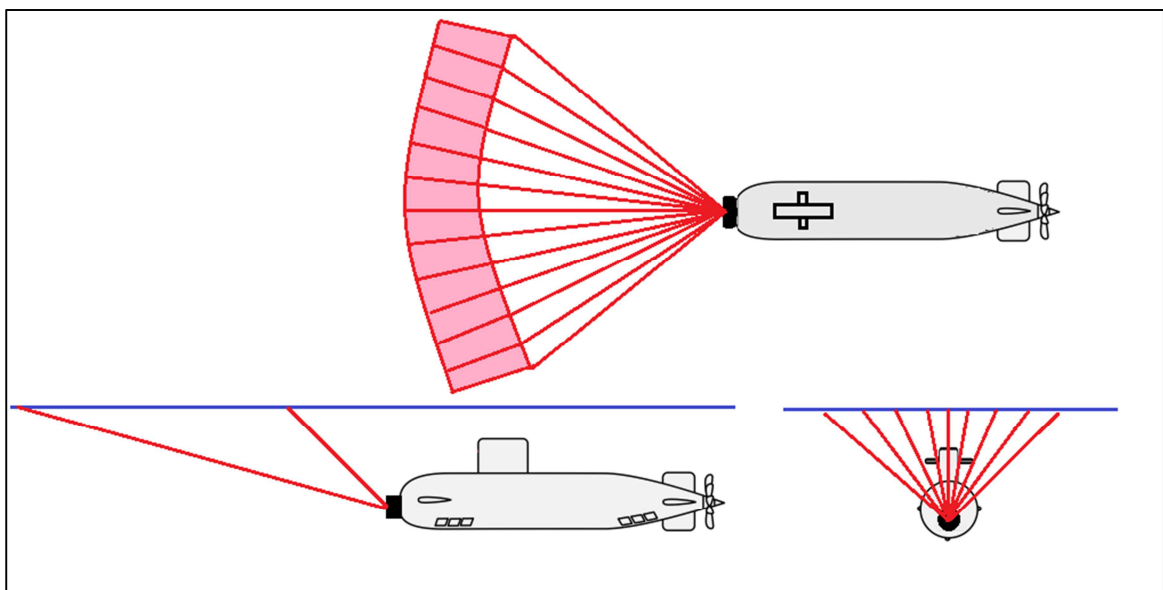


Figure 26 - Volume balayé par l'ULS en configuration horizontale

Cette étude a été faite en 2014/2015 en collaboration avec une ingénieure de recherche, Imen Karoui, après une première étude en 2012/2013 pour la société iXblue dans le cadre du projet MUSE (cf. §2.6.2). Une partie des résultats de ces travaux ont été publiés dans les actes de la conférence Underwater Acoustics et surtout une revue IEEE dont l'article est fourni annexe 6 page 125 ([KAR14] et [KAR15b] pour le projet MUSE, [KAR15a] pour le contrat avec DGA TN).

La littérature étant maigre sur le sujet, de même que les exemples d'images sub-surfaces par moyen acoustique, il nous a tout d'abord fallu emprunter le sonar d'évitement Reson Seabat 8101 de DGA TN Brest et faire des essais sur perche à l'aide de la vedette hydrographique de l'ENSTA Bretagne en 2013. Cet ensemble de données a été plus largement complété en 2015 lors d'essais menés par DGA TN avec l'AUV DAURADE équipé du sonar BlueView P900-90.

Deux méthodes ont été développées :

- Une méthode basée objet qui consiste à détecter un obstacle potentiel et à le pistier au cours de la séquence [KAR15a, KAR15b],
- Une méthode basée sur les grilles d'occupation qui permettent de capitaliser les informations de la surface de mer perçues par le sonar au fil du temps et de garantir une remontée sécurisée dans une certaine fenêtre.

Contrairement à la colonne d'eau, le fouillis de mer (nuages de bulles, vagues...) complique la détection sub-surface. Par ailleurs, les signatures acoustiques rencontrées sont très variables. Pour détecter les différentes signatures et pistier les obstacles en surface, une méthode hiérarchique a été proposée [KAR15a]. Brièvement, elle consiste tout d'abord à détecter et gommer les signatures les plus énergétiques liées aux navires en transit (détection des voies saturées par le bruit de cavitation des navires), puis appliquer un détecteur, réputé efficace dans le cas de fouillis de mer, l'algorithme adaptatif à taux de fausses alarmes constant (TFAC) et largement inspiré de [RAV92]. La recherche locale de sillage autour des différentes détections permet enfin d'écarter les fausses alarmes dues aux sillages et de ne garder que les positions potentielles des navires. Comme les obstacles ont une signature étendue avec une certaine structure, plusieurs détections correspondent souvent à un même objet. Une étape de regroupement est alors nécessaire et mène grossièrement à une segmentation des obstacles. La méthode retenue pour sa robustesse et sa faible exigence en termes de paramètres d'entrée s'appelle ASCDT pour Adaptive Spatial Clustering based on Delaunay Triangulation [DEN11] et est illustrée Figure 27. Une fois l'obstacle segmenté, sa position est extraite pour permettre son suivi au cours du temps (à noter que cette position est calculée de manière différente en fonction de la caractéristique de l'obstacle : avec ou sans sillage). D'une manière différente au pistage dans la colonne d'eau, l'enjeu ici est de permettre la reprise de vue dans un certain périmètre de la surface de l'océan. Le pistage s'effectue donc dans le repère absolu et le vecteur d'état contient la position et la vitesse de l'obstacle sur la surface. A nouveau, un filtre de Kalman sans parfum est mis en œuvre de façon à gérer la non linéarité du modèle. La mise à jour des pistes est enfin réalisée par combinaison probabiliste des mesures possibles selon la méthode JPDAF (Joint Probabilistic Data Association Filter). C'est une approche plus complexe et plus lente que l'approche du plus proche voisin, mais qui évite une erreur d'association dans un environnement qui peut être chargé en surface comparé à la colonne d'eau. Cette première méthode basée objet a l'avantage de détecter et pistier les obstacles fixes et mobiles. Néanmoins, elle est sensible aux fausses alarmes et exige d'avoir l'objet en totalité dans le champ de vision du sonar.

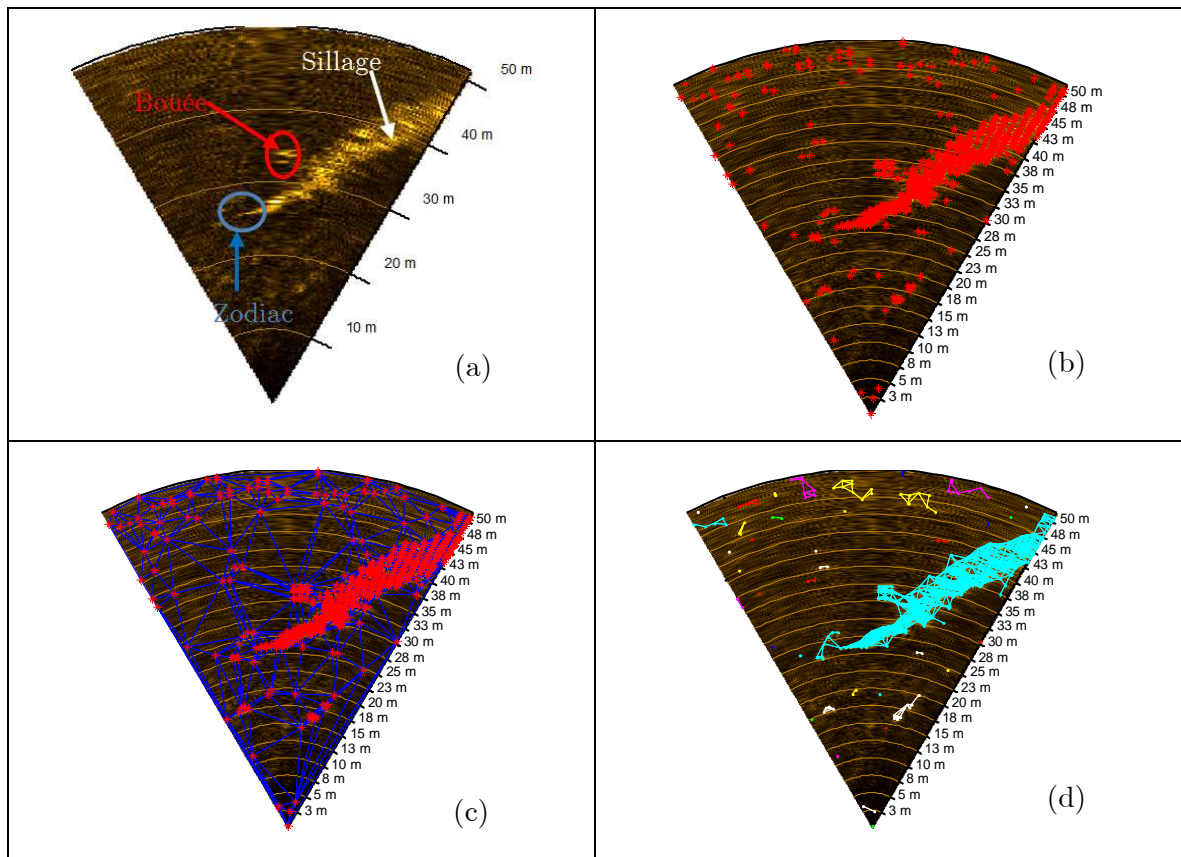


Figure 27 – Traitement d’une image de la surface acquise par le Reson Seabat 8101 (a): Image sonar avec l’étiquetage des signatures présentes, (b): résultat de la détection par le TFAC (points en rouge), (c) : après triangulation de Delaunay (en bleu), (d) : Sous-graphes restants après tri.

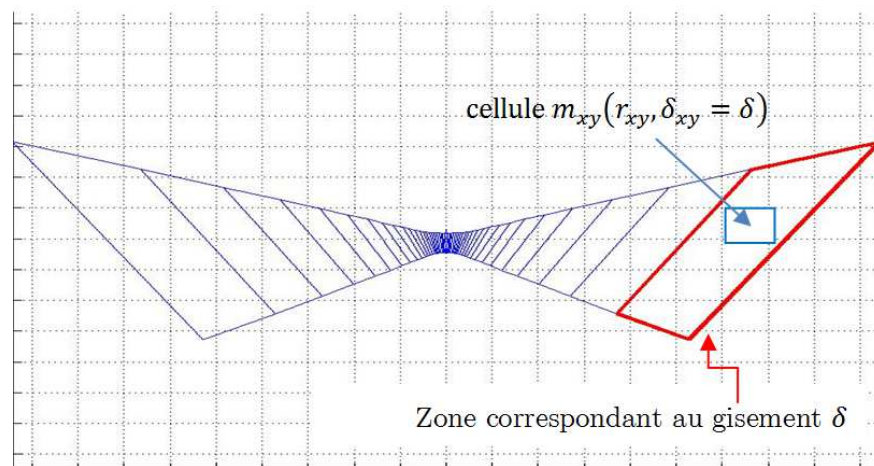


Figure 28 – Exemple de projection du champ de vision sonar sur la surface de mer quadrillée dans l’axe du sonar : en bleu foncé, superposition des pixels sonar projetés à la grille d’occupation de la surface de mer pour une ligne distance ; en rouge, surface projetée du pixel de la voie tribord extrême ; en bleu clair, cellule m_{xy} de la grille de coordonnées polaires (distance, gisement).

Une méthode alternative basée sur les grilles d'occupation a ainsi été envisagée mais seulement pour les obstacles fixes pour le moment. La philosophie est différente dans le sens où elle vise à garantir une fenêtre de relative sécurité pour la reprise de vue. La grille d'occupation dans le repère absolu correspond à un découpage de la surface de mer visible du sonar en un ensemble de cellules. Le but est d'estimer pour chaque cellule, la probabilité qu'elle soit occupée (par un obstacle ou une partie seulement). Pour un ping donné, l'ensemble des détections relatives au même obstacle permet de délimiter un volume 3D qui est ensuite projeté sur la grille d'occupation située à la surface. L'estimation de la probabilité d'une cellule se fait au fur et à mesure des détections retenues pour chaque ping selon une méthode décrite dans [GAN15] (méthode proposée pour détecter des obstacles à l'aide d'un sonar frontal mais sur le fond). Comme le montre la Figure 28, une bonne connaissance du capteur et de sa géométrie est exigée pour bien gérer occupation et intersection de l'obstacle avec les cellules de la grille. Une grille d'occupation 3D est également envisageable si l'on dispose de sonars frontaux en croix [CHE09].

3.3.3 Suivi de fond

Le suivi de fond vise à prévenir le véhicule d'une remontée de fond brutale ou encore à estimer grossièrement la bathymétrie dans le sens d'avance du véhicule. L'OAS adapté ici est un sondeur multifaisceau (SMF) utilisé comme sonar frontal en mode horizontal c'est-à-dire incliné de quelques degrés vers le bas par rapport au plan horizontal (voir Figure 29) que nous désignerons sous le terme Downward Looking Sonar ou, de manière plus synthétique, DLS. On utilise les données SMF brutes, *i.e.* en niveaux de gris, et non pas traduites en points de sonde comme d'usage en bathymétrie.

Le mode bathymétrique consiste en une faible ouverture en site permettant la visualisation d'une fine bande du fond éclairée, de même que des bandes secondaires si le contraste est assez bon (voir Figure 30). Cette bande principale et les bandes secondaires de part et d'autre correspondent à la rétrodiffusion du fond dans le lobe principal et les lobes secondaires du diagramme de directivité de l'antenne d'émission comme l'illustre schématiquement la Figure 31. La détection des limites de cette bande et son évolution ping à ping en termes de position et largeur permettent de suivre de manière grossière la bathymétrie (relief grossier et pente). Les performances de cette approche dépendent beaucoup de la précision de la détection des limites de la bande ou du premier trajet vers le fond. La facilité de détection est fonction du rapport signal à bruit qui est inversement proportionnel à la largeur de la bande balayée du fond. Cette largeur est liée à la distance moyenne au fond et à l'ouverture en site du sonar.

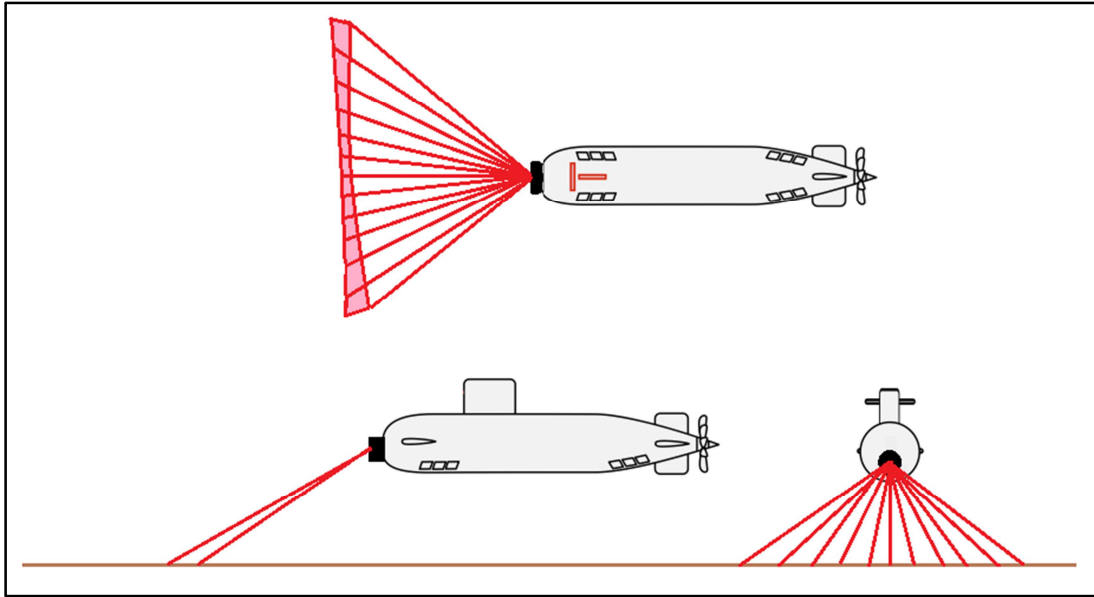


Figure 29 - Volume balayé par le DLS en configuration horizontale et mode bathymétrique

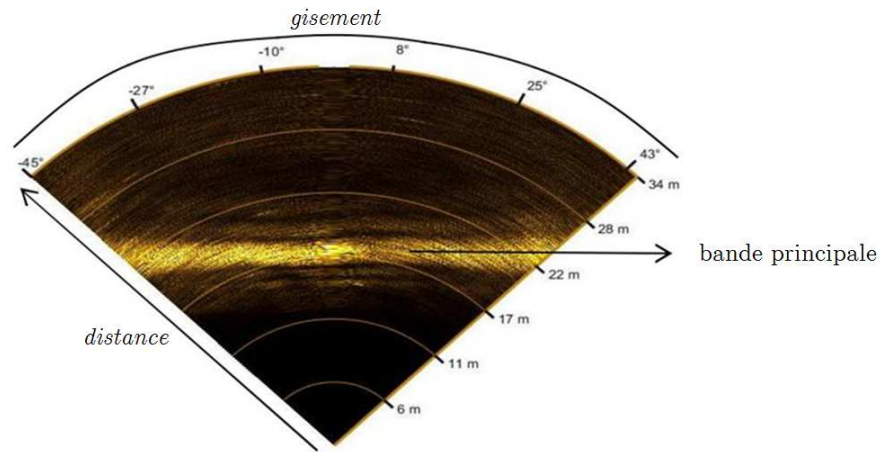


Figure 30 – Exemple d'image du fond obtenue par le sondeur R2020 de R2sonic

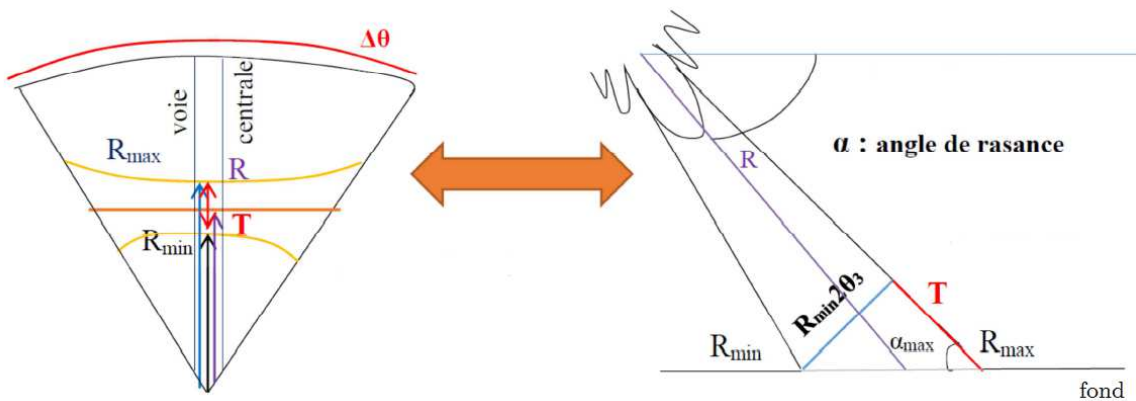


Figure 31 – Lien entre directivité de l'antenne d'ouverture 2θ , et largeur T de la bande principale sur le fond dans la voie centrale : à gauche, exemple schématique d'une image sectorielle du fond ; à droite, géométrie d'acquisition vue de profil

Cette partie de l'étude a été réalisée avec l'aide d'un stagiaire de Master Recherche « Signaux et Circuits » de l'Université de Bretagne Occidentale, Zied Hmissi, en 2016. L'étudiant n'étant pas formé aux sonars d'imagerie, le travail s'est effectué en deux temps. D'une part, un simulateur du masque de directivité sonar sur le fond marin (en fait seulement la trace du lobe principal) a été développé pour un fond de pente quelconque (problème direct). Cette première partie lui a permis de bien maîtriser la géométrie d'acquisition. D'autre part, une méthode d'estimation de la pente du fond basée sur la largeur de la bande segmentée à partir de l'image sonar brute a été proposée (problème inverse). Cette méthode a été validée sur des données réelles du sondeur multifaisceau R2020 de R2sonic en mode frontal incliné vers l'avant, loué pour les besoins de l'étude, et déployé, faute de véhicule sous-marin, depuis la vedette Panopée de l'ENSTA Bretagne fin novembre 2015 (voir Figure 30).

La pente du fond est une information importante mais partielle sur le fond. La remontée de fond par rétrécissement de la largeur de la bande du fond fut assez facile à détecter mais la caractérisation de macrostructures comme les piliers d'un ponton ou des monticules est plus délicats et doit faire l'objet de plus amples travaux.

3.3.4 Commentaires sur ces travaux consacrés à la détection d'obstacles

Ces travaux à fort caractère opérationnel n'ont pas amené de développements académiques importants mais ont exigé la mise en œuvre d'algorithmes robustes à la hauteur de données difficiles (dans le cas de l'ULS particulièrement) collectées par une plateforme mobile sous-marine. Par ailleurs, ils peuvent être mis à profit de la reconnaissance de mines à orin ou dérivantes, alors que la guerre des mines se concentre plus traditionnellement sur les mines de fond.

4 Projet de recherche

Les prochaines années vont me permettre de mettre à profit les compétences que j'ai acquises en chasse aux mines, d'abord dans l'industrie puis en recherche académique, dans le contexte actuel d'automatisation des tâches pour un portage sur véhicule sous-marin sans pilote et autonome. L'automatisation qui était auparavant une simple aide aux opérateurs passe ici un cap et l'enjeu est suffisamment important pour que recherches académique et industrielle se mobilisent avec un soutien et un encouragement de l'état (DGA) et plus récemment de la région Bretagne pour le développement d'une compétence drone maritime sur le territoire (programme SAMM – Systèmes Autonomes en Milieu Maritime). Depuis mon arrivée à l'ENSTA Bretagne, je travaille sur la perception des véhicules sous-marins autonomes avec la DGA pour la problématique d'évitement d'obstacles visant à garantir la sûreté du véhicule. La variété des situations (fond, colonne d'eau, sub-surface) et des obstacles pour cette seule tâche montre l'ampleur des travaux à mener avant d'atteindre des performances minimales pour une mission en autonomie. Le Lab-STICC a bien perçu ce besoin et proposé la création d'une nouvelle équipe au prochain quinquennal, création qui a été validée par l'HCERES au printemps dernier.

Mes projets à quatre ans se décomposent ainsi en deux parties : l'une consacrée à la naissance de cette nouvelle équipe du Lab-STICC dont je suis responsable, l'autre à l'évolution de mes travaux de recherche dans un contexte exigeant de portage des algorithmes sur véhicule autonome, et en particulier pour la tâche d'identification.

4.1 Animation équipe PRASYS

4.1.1 Historique et positionnement

La création d'une nouvelle équipe au sein du Lab-STICC répond à la volonté de valoriser les activités d'une partie des enseignants-chercheurs du pôle CID (Connaissance Information Décision) autour de la robotique mobile autonome de perception de l'environnement physique. Une quinzaine d'enseignants-chercheurs auparavant rattachés à d'autres équipes se retrouvent ainsi réunis pour développer, analyser et évaluer en conditions réelles les algorithmes nécessaires à la réalisation de systèmes intelligents autonomes distribués. Cette problématique récente est issue du domaine de la robotique dans laquelle l'interaction robot-environnement se fait via le capteur, ce qui nécessite l'étude des mécanismes perception-action. Les mots clefs Perception, Robotics et Autonomous SYStems ont donné le nom de cette nouvelle équipe : PRASYS. Le système

autonome (SA) désigne un seul robot, porteur d'un ensemble de capteurs ou un groupe de robots (notamment hétérogènes) permettant la distribution de ces mêmes capteurs. C'est un système bouclé de sorte que l'ensemble des capteurs du SA dédiés à l'acquisition automatique de données sur l'environnement à explorer (perception) aident réciproquement (interaction) le SA à adapter son comportement à l'environnement (action). En acceptant de mettre en avant cette thématique par la création d'une nouvelle équipe, le bureau du Lab-STICC fait de la robotique mobile autonome un axe fort de recherche de son laboratoire. L'idée de créer cette équipe a germé fin 2014 et de nombreuses discussions au sein du Lab-STICC ont permis de définir son projet scientifique en permettant également aux équipes SFIIS (qui est devenu IRIS) et IHSEV dont venaient les membres de PRASYS de recentrer leurs activités. A l'extérieur, la nouvelle équipe doit trouver sa place parmi les équipes nationales et internationales existantes de thématiques voisines. Sur le plan national, des équipes partagent quelques compétences avec PRASYS mais souvent avec un effectif et un périmètre plus étendus : équipe LAGADIC de de l'UMR IRISA (Institut de Recherche en Informatique et Systèmes Aléatoires), département Traitement de l'Information et Systèmes de l'ONERA (Office National d'Etudes et de Recherches Aérospatiales), équipe Robotics and InteractionS du LAAS-CNRS (Laboratoire d'analyse et d'architecture des systèmes), équipe EXPLORE de l'UMR LIRMM (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier). Néanmoins, PRASYS propose des outils mathématiques originaux, en particulier les méthodes ensemblistes, adaptés à la gestion des incertitudes dans un environnement non structuré, par exemple le milieu marin. Au plan international, on peut citer trois laboratoires d'activités proches de PRASYS : Robotics & Perception Group (Université de Zurich, Suisse), Autonomous Robotics & Perception Group (University of Colorado, US), Mobile Robotics Group (University of Oxford, UK) mais la robotique marine, familière des membres de PRASYS, y est peu représentée. La compétence robotique sous-marine n'est bien évidemment pas absente du paysage européen mais plutôt avec des aspects de design et vision dans l'équipe Girona Underwater Vision and Robotics (University of Girona, Espagne), design et coopération entre véhicules pour le Bioengineering and Robotics Research Center (University of Pisa, Italie), de contrôle pour Genoa Robotics and Automation Lab (University of Genova, Italie) et de perception (y compris distribuée) pour Ocean Systems Laboratory (Heriot-Watt University, Ecosse).

4.1.2 Contribution

En tant qu'animatrice, l'un des défis à relever pour la construction de cette équipe est de faire naître une synergie entre des chercheurs de domaines différents, le traitement de l'information pour les uns (axe perception), la robotique/automatique pour les autres

(axe action autonome), en les faisant communiquer (séminaires) et collaborer autour de projets mixtes à l'instar de ce qui se fait déjà (anti-collision, perception par meute...). Par ailleurs, ces chercheurs viennent de trois établissements différents en proportions inégales : l'ENSTA Bretagne en majorité, suivie de IMT Atlantique (site de Brest) et enfin l'UBO. La réunion de lancement de l'équipe de juin 2016 a permis de réunir chercheurs et doctorants et de rappeler les enjeux pour chacun, pour l'équipe et enfin le Lab-STICC.

En tant que chercheuse, mon travail de recherche se place plus naturellement dans l'axe 'perception' de l'équipe. En rupture avec mes activités en chasse aux mines chez Thales, je me suis intéressée dès l'année suivant mon arrivée à l'ENSTA Bretagne à la problématique de l'évitement d'obstacles pour véhicule sous-marin autonome ou plus simplement AUV (Autonomous Underwater Vehicle). Le GESMA (aujourd'hui DGA TN Brest) avait exprimé ce besoin pour ses démonstrateurs, l'AUV REDERMOR tout d'abord et l'AUV DAURADE ensuite [QUI07, QUI09, QUI10, QUI11]. Mon travail a consisté entre autres à extraire toutes les informations perçues par un capteur dédié à l'évitement d'obstacles pour induire une manœuvre d'évitement. Une fois détectée, la cible exige un suivi au cours du temps. En collaboration avec les roboticiens de mon équipe locale à l'ENSTA Bretagne, un pistage par filtrage de Kalman a été mis en place avec une prise en compte des données de navigation (vitesse, attitude) du véhicule [QUI12c]. Cet exemple montre bien que capteur et véhicule sont indissociables de sorte que la connaissance du comportement du véhicule évite un décrochage de l'algorithme ou une mauvaise interprétation de la donnée perçue. Réciproquement, les informations perçues permettent de définir la meilleure manœuvre d'évitement mais également l'acquisition d'informations complémentaires sous des géométries et/ou réglages différents. L'adaptation fait donc inévitablement partie de la mission sous peine d'arrêt brutal (collision) ou de performances médiocres. L'apprentissage au sens capitalisation de l'expérience est également une piste à explorer compte tenu du manque de données dans le domaine sous-marin. Plus généralement, alors que le bouclage perception-action en robotique autonome mobile est étudié très tôt en terrestre [RIV91], il est encore à formaliser en sous-marin du fait des spécificités de ce milieu. L'environnement marin est en effet peu structuré et mal connu. Pourtant, le robot doit s'appuyer sur cet environnement pour par exemple, corriger sa position ou ponctuellement agir sur la loi de commande. Par ailleurs, le robot mobile sous-marin a pour spécificité d'être toujours en mouvement ce qui induit un couplage direct entre la perception et l'action afin de prendre rapidement en compte la dynamique de l'environnement. Enfin, en raison d'un débit limité en communication acoustique, une partie des traitements et des décisions doivent se passer à bord ce qui exige de faire un choix entre ce que l'on traite à la volée et ce qu'on transmet ou stocke. Tous ces aspects propres à l'environnement marin

exigent la définition d'un formalisme spécifique de la boucle perception-action et une réflexion plus générale sur les évolutions possibles du système autonome sous-marin avec l'introduction de capteurs intelligents, la coopération avec d'autres robots...

Nous verrons au paragraphe suivant que les états-majors ont depuis plusieurs années compris l'intérêt du robot mobile sous-marin pour des missions à risque ou de discrétion, et investi pour son développement. Son usage civil d'abord cantonné à l'industrie pétrolière et gazière s'étend aujourd'hui, notamment dans le domaine de l'inspection d'ouvrages.

4.2 Emergence du robot mobile sous-marin pour le militaire et le civil

La tendance actuelle pour les missions à risque est de les confier à des robots, communément appelés drones, autonomes ou pas. Le milieu sous-marin ne fait pas exception et la mission de chasse aux mines sera déportée à moyen terme sur un AUV ou une meute d'AUVs, même s'il faut en parallèle convaincre des opérateurs pas toujours enclins à une confiance totale dans l'automatisation et l'autonomie de systèmes opérant hors de leur champ de vision. Par ailleurs, dès le début de mission, le véhicule autonome disparaît sous la surface de l'eau et par conséquent, du champ de vision ce qui nécessite de gagner la confiance des opérateurs. Les avantages en matière de coûts (matériels et humains), de sécurité et de discrétion sont évidents. Par ailleurs, la possibilité d'acquérir des données à altitude constante par rapport au fond et à partir d'une plateforme stable garantit des données de meilleure qualité. L'utilisation civile des AUVs a vu le jour à la fin des années 90 dans l'industrie pétrolière et gazière, suivie de son utilisation militaire par quelques nations de l'OTAN une dizaine d'années plus tard. En Europe, les développements et les alliances dans le domaine de la contremesure marine par plateformes autonomes concernent essentiellement la Norvège (FFI, Kongsberg Maritime), l'Allemagne (ATLAS ELEKTRONIK), le Royaume-Uni (MoD, BAE Systems,...) et la France (DGA, THALES, ECA) bien sûr. Initié en 2012 dans le cadre de la coopération franco-britannique, le programme Maritime Mine Counter Measures (MMCM) piloté par l'OCCAR (<http://www.occar.int/325>) vise le développement d'un prototype de système autonome destiné à la détection et à la neutralisation des mines marines et d'IEDs. Ce programme est une partie du programme français SLAMF, pour Système de Lutte Anti-mines Marines Futur, conduit par la DGA visant le renouvellement de la capacité de guerre des mines de la marine nationale qui s'appuie sur le concept de bateau-mère capable de mettre en œuvre un système de drones. Le PEA ESPADON (Evaluation incrémentale de Solutions Potentielles d'Automatisation de

Déminage pour les Opérations Navales), via le démonstrateur de drone de surface de la Figure 32 baptisé *Sterenn Du*, a déjà permis de démontrer la validité du concept de déploiement d'un ou plusieurs drones sous-marins à partir d'un USV (Unmanned Surface Vehicle), lui-même déployé depuis un bateau-mère. L'objectif est de réduire l'intervention humaine directe dans le processus de déminage. Cependant, l'USV associé à un (ou des) ROV ou AUV serait plutôt destiné aux dernières phases du processus de déminage, à savoir la réinspection (identification), l'intervention et la neutralisation (mission affichée de l'USV d'inspection *Inspector* du groupe ECA). Les phases de détection, classification et localisation (DCL) restent l'apanage d'un (ou d'une meute d') AUV afin de dégager rapidement (le véhicule peut être amené facilement sur place) une route en cas de conflit ou de nettoyer une large zone, comme le fait actuellement le bâtiment de surface mais avec une certaine lourdeur et lenteur d'intervention.



Figure 32 - Drone de surface (USV) Sterenn Du, réalisé dans le cadre du PEA ESPADON

Mes activités de recherche pour les prochaines années sont fortement liées à ce contexte mais aussi en continuité avec mes activités passées en prenant mieux en compte les contraintes et les avantages du portage sur AUV.

4.2.1 Reconnaissance de cibles (détection/classification)

Depuis quelques années, et en particulier sous l'impulsion du comité d'organisation du workshop iSAAM (International Workshop on Intersections in Signal Processing, Acoustics & ATR for Maritime Applications) dès 2012, une réflexion est menée au sein de la communauté de la reconnaissance automatique de cibles sous-marines pour pointer les limites et imaginer les solutions à venir pour ce qu'on résume plus simplement par ATR. Bien que l'accent ne soit pas mis sur les véhicules autonomes, les problématiques posées sont discutées à la lumière des développements technologiques, de leurs atouts et des contraintes associés. La première édition de 2012 a montré la nécessité de quantifier la performance de l'ATR en incluant les effets de l'environnement et du traitement du

signal. De ce fait, nous avons fait la demande en 2013 d'une bourse de thèse DGA consacrée à la qualification du fond marin à partir d'images sonar latéral. La prise en compte de l'environnement dans le processus ATR est en rupture avec ce qui se faisait jusqu'alors. Les travaux de thèse, réalisés par Laurent Picard de 2013 à 2016, ont permis, par le biais d'un niveau de confiance, de qualifier le fond en termes d'homogénéité, anisotropie et complexité (voir paragraphe 3.2.1.2.3). Pour un traitement embarqué, ces informations peuvent orienter le choix de tel ou tel algorithme plus ou moins élaboré, et optimiser la ressource disponible. On peut ainsi ne privilégier un algorithme complexe et consommateur que si nécessaire ou encore, changer de point de vue en déroutant le véhicule en cas de rides de sable pouvant occulter une cible. A cette qualification environnementale peut s'ajouter une capacité d'auto-évaluation de la qualité de la mesure qui peut mener l'AUV à diminuer son altitude par rapport au fond si la réverbération (estimée) est trop forte ou bien en cas de multi-trajets polluant la visibilité du fond... On retrouve ici encore un avantage de la chasse aux mines par AUVs, capables d'adapter leur navigation pour augmenter les performances des traitements en fonction de la qualité des données (voir par exemple dans [COU12], le gain de performance obtenu par une stratégie adaptative *in situ* en fonction de la portée effective du sonar ping à ping) ou du faible degré de confiance du classifieur qui exige un second point de vue sur l'objet (cf. classification multi-vue évoquée plus haut). D'un point de vue plus prospectif à présent, la modularité à venir du capteur peut être exploitée de différentes manières : faire un traitement SAS de gain juste nécessaire, choisir la sous-bande (d'un sonar large-bande) appropriée à un type de fond donné (plus basse fréquence pour les champs de rides ou pour la suspicion de mines enfouies, plus haute fréquence pour mieux classer sur ombre), ou encore exploiter davantage la possibilité de combiner les données acquises dans différentes sous-bandes (comme on l'a évoqué au paragraphe 3.2.2.3.2).

Outre la flexibilité apportée par l'AUV, il convient de maîtriser son usage et en particulier, formaliser la boucle perception-action dans le contexte sous-marin. Comme indiqué plus haut, l'AUV a pour obligation d'être perpétuellement en mouvement ce qui induit quelques actions réflexes comme l'évitement d'obstacles pour assurer sa sécurité ou induire une remontée d'urgence en cas de panne. En ce qui concerne la navigation et le suivi d'une trajectoire de consigne, les seules informations proprioceptives comme les données de la centrale inertielle subissent une inévitable dérive et l'AUV doit s'appuyer sur son environnement pour ainsi, corriger sa position ou ponctuellement agir sur la loi de commande. Or, dans le contexte sous-marin, l'environnement est peu structuré et mal connu. Pour s'en sortir, il devra s'appuyer sur des amers d'opportunité et gérer les incertitudes concernant leur localisation avec des outils mathématiques adaptés comme l'arithmétique des intervalles. Ainsi, informations proprioceptives et extéroceptives se complètent mais leur articulation doit être formalisée.

Enfin, l'avènement des capteurs intelligents est réellement prometteur pour les AUV. Pour un OAS, il pourrait par exemple auto-adapter son ouverture verticale s'il intercepte la surface de mer et que le taux de fausse alarme engendré est trop élevé. Pour la classification de cible sur le fond, le sonar pourrait décider de lui-même de la nécessité d'actionner sa mode multi-aspects ou pas, le SAS de monter en résolution ou pas en fonction du signal mesuré...

4.2.2 Identification de cibles et inspection d'ouvrage

La phase d'identification de cibles à partir d'un chasseur de mines français peut être réalisée, soit par deux plongeurs démineurs, soit à l'aide du ROV PAP (poisson auto-propulsé, voir Figure 33). Équipé d'une caméra et d'un projecteur, le PAP peut transmettre instantanément des images de l'objet suspect au centre opérationnel et éventuellement déposer sa charge explosive à proximité de celle-ci. La caméra du PAP est une caméra sous-marine classique dont la limite en portée est principalement liée à la turbidité de l'eau. Comme nous l'avons noté lors des travaux de thèse de Stéphane Bazeille, les caméras optiques classiques ont en effet comme principaux défauts une rapide atténuation en distance, un contraste faible et souvent un éclairage non uniforme (dû au spot du projecteur) [BAZ08a]. Ces travaux financés par la région Bretagne ont été réalisés en collaboration avec Thales Underwater Systems (TUS) dans le cadre du projet TOPVISION (Test Opérationnels de Vidéos Sous-marines pour l'Identification d'Objets Nuisibles) qui faisait partie d'un programme nommé Techno-Vision lancé par le Ministère de la Recherche et le Ministère de la Défense en 2005. Ce projet visait à réaliser une campagne d'évaluation d'algorithmes de traitements en vidéo sous-marine pour des fonctions de détection, localisation, caractérisation et identification automatiques d'objets nuisibles. Trois des quatre membres de son comité d'organisation appartenaient à la région Bretagne : TUS, l'ENSIETA (aujourd'hui ENSTA Bretagne) et le GESMA (aujourd'hui DGA TN Brest). Stéphane a développé une chaîne de pré-traitement des images optiques sous-marines qui corrige chacun des défauts énoncés précédemment [BAZ08b].

Néanmoins, des solutions technologiques permettant de s'affranchir de la turbidité existent [JAF13]. L'AUV norvégien HUGIN par exemple est équipé d'une caméra TileFish ou TileCam qui consiste en une caméra optique haute résolution synchronisée avec un stroboscope à LEDs (Figure 34), l'un et l'autre positionnés à une extrémité différente du porteur afin de réduire l'impact de la rétrodiffusion. Il en résulte des images d'une clarté étonnante comparées aux images obtenues par caméras optiques classiques (Figure 35) tout en permettant au véhicule de se maintenir à une altitude de sécurité par rapport au fond [MID11] et avec une consommation limitée grâce aux LEDs. Une option

concurrente est le *range-gated lidar* qui élimine toute mesure en dehors d'une plage distance comme par exemple le système LUCIE (Laser Underwater Camera Image Enhancer) développé initialement par le DRDC [JAF01]. Un laser pulsé émet un faisceau court de lumière qui va se réfléchir sur l'objet et la caméra n'enregistre la réponse que sur une fenêtre temporelle correspondant au temps de trajet aller-retour de l'impulsion sachant l'altitude du porteur [FOU93].



Figure 33 – Pilotage du PAP depuis le chasseur de mines Andromède (©Pierre Lansonneur)

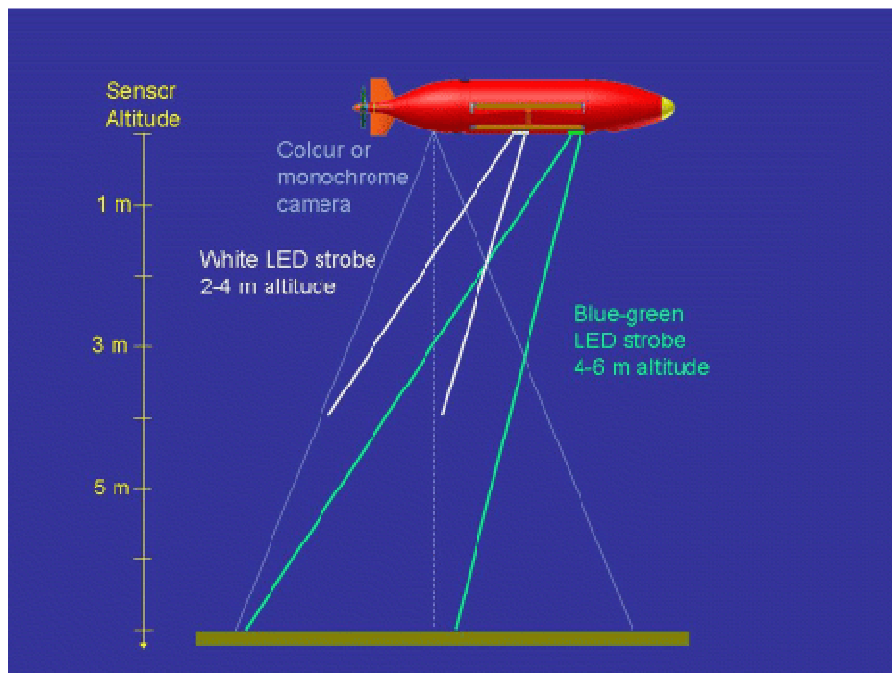


Figure 34 – Système TileFish monté sur Hugin : une caméra associée à un stroboscope à LEDs
(source : <http://www.neosubsea.com/products/tilefish/>)

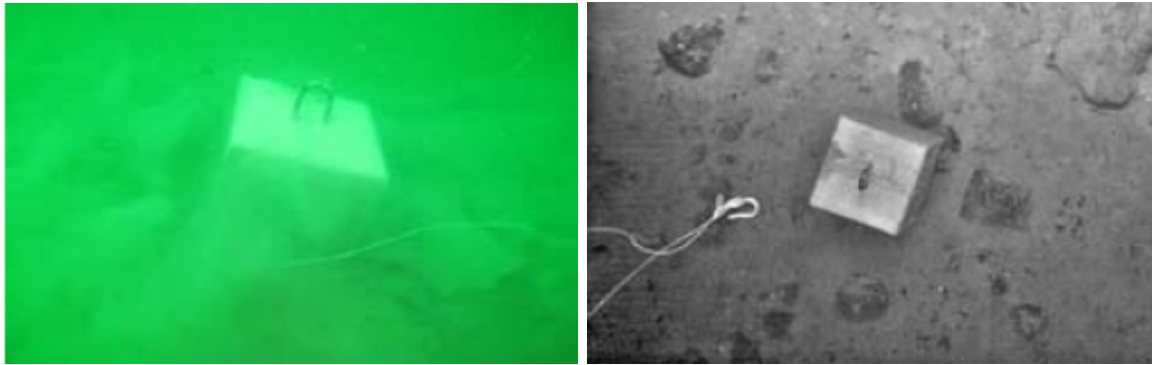


Figure 35 – Comparaison entre une caméra optique classique à gauche et la caméra TileCam à droite sur le même objet (extrait de [MID11])

Une autre option pour s'affranchir de la turbidité est l'utilisation d'une caméra acoustique, soit un réseau 2D de capteurs. A noter que le terme caméra acoustique est parfois abusivement employé pour désigner des sonars sectoriels haute résolution comme les sonars DIDSON (Dual-frequency IDentification SONar) qui utilisent une technologie à base de lentilles acoustiques développée initialement par l'Applied Physics Lab de l'Université de Washington pour permettre aux plongeurs d'identifier des mines dans des eaux turbides [BEL01]. Une formation de voies optique a quelques avantages par rapport à une formation de voies électronique. Une lentille acoustique peut se focaliser dès la distance d'un mètre. Par ailleurs, la formation de voies optique exige peu d'énergie (avantage certain pour un AUV) et diminue les problèmes de diaphonie (*cross-talk*) entre voies voisines. Une *vraie* caméra acoustique en revanche effectue une formation de voies électronique mais bi-dimensionnelle qui convertit les données sonar en éléments d'information volumique, les « voxels ». Une seule récurrence contient à la fois l'information en distance et transversale (voir par exemple, le système AMI développé par TUS Pty au début des années 2000 [MAG01]).

Une caméra acoustique ayant malheureusement un coût très élevé, un sondeur multifaisceau et à défaut un sonar sectoriel sont des options que nous allons retenir, le premier pour l'identification de mines dans le cadre d'un contrat de recherche pour le compte de DGA TN, le second pour l'inspection de barrage proposée au titre du projet pédagogique Guerledan.

Un sondeur multifaisceau dispose de deux antennes en croix dont l'une, dans le sens de déplacement du véhicule, émet et l'autre, multicapteur, dans le sens perpendiculaire, reçoit. Les voies formées à la réception permettent par intersection avec le faisceau formé à l'émission de balayer un grand nombre de cellules du fond de part et d'autre du porteur (Figure 36). DGA TN a choisi depuis quelques années d'étudier le potentiel de ce système pour identifier les mines à partir d'un AUV dédié à la phase d'identification. L'identification a dans un premier temps été envisagée de manière classique à partir des

points de sonde détectés par les algorithmes du constructeur du sonar en survolant la cible de très près c'est-à-dire à moins d'une dizaine de mètres de l'objet étudié (ce travail a été réalisé avec l'aide d'un étudiant de l'école, Olivier Clert, durant sa formation en alternance et surtout son stage de fin d'étude en 2014 [CLE14]). Cependant, les données brutes, données d'amplitude et phase dans la colonne d'eau, méritent une étude approfondie. En effet, les points de sonde détectés⁵ automatiquement le sont à des fins d'estimation de la bathymétrie du fond mais réduisent de fait drastiquement l'information et sous certaines hypothèses qui s'avèrent inexactes pour l'identification 3D des objets (en particulier les mines à orin). Or dans le cas de l'imagerie des objets complexes, on s'aperçoit que les données de colonne d'eau possèdent plusieurs échos par voie dont certains peuvent être pertinents pour remonter à la forme tridimensionnelle de l'objet. C'est un des aspects du contrat de recherche nommé Ident3D pour « Etudes de traitement de données 3D pour l'aide à l'identification de mines sous-marines » qui est en cours pour DGA TN. Deux pistes sont notamment possibles : retravailler la détection des points de sonde avec de nouvelles hypothèses et effectuer la reconnaissance de l'objet à partir du nuage de points, ou bien effectuer cette reconnaissance à partir des données brutes de colonne d'eau directement. Cette dernière piste est la plus ambitieuse.

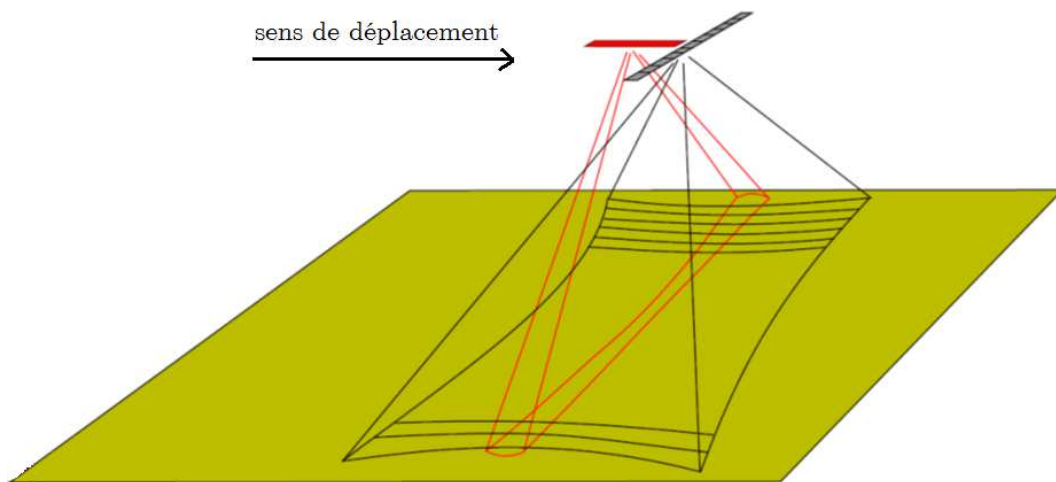


Figure 36 – Principe de balayage du sondeur multifaisceau

Les problématiques de l'identification de cibles et de l'inspection d'ouvrage sont assez voisines : nécessité de s'approcher de l'objet d'intérêt, d'y déceler de fins détails à l'aide de capteurs haute résolution pour reconstruire ou inspecter c'est-à-dire fournir une

⁵ La détection d'un point de sonde se fait classiquement pour chaque voie sonar par détection du maximum d'amplitude de la voie ou du passage par zéro d'une mesure de déphasage entre les deux extrémités de l'antenne

représentation à la fois interprétable et permettant l'extraction d'informations fiables. Nous nous sommes intéressées à l'inspection d'ouvrage dans le cadre du projet pédagogique Guerlédan porté par l'ENSTA Bretagne et ses partenaires (<http://hydrob.ensta-bretagne.fr/hydrographes-roboticiens-explorent-guerledan/les-partenaires/>). Ce projet vise à réunir les communautés de la robotique et de l'hydrographie au sein d'un stage pédagogique de terrain, situé au Lac de Guerlédan en centre Bretagne pour les années scolaires 2016/2017 et 2017/2018. Le sujet propose d'inspecter la partie immergée du barrage et en particulier les réparations qui avaient été faites lors de l'assec de 2015.

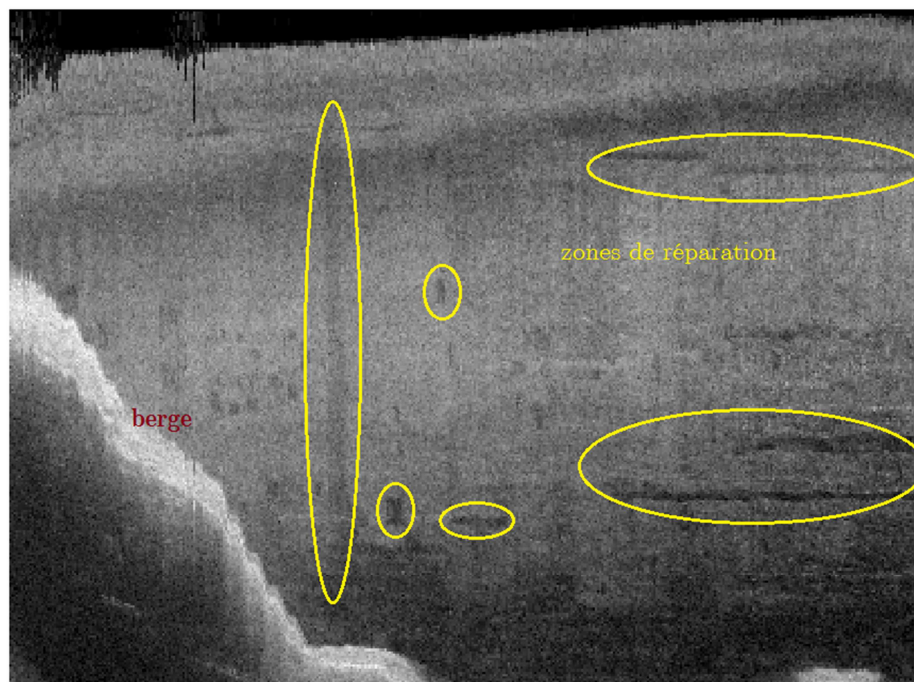


Figure 37 – Image sonar latéral (Tritech SeaKing Towfish) rectifiée du barrage de Guerlédan

Le monitoring de la paroi du barrage intéresse particulièrement EDF chargé de son exploitation et de son entretien. La réglementation en matière de sécurité des ouvrages hydrauliques a évolué depuis le décret de décembre 2007 et permet dans le cadre des revues de sûreté d'engager sans dérogation des examens des parties habituellement immergées sans vidange de la retenue. EDF voit dans l'inspection subaquatique un moyen de réduire les impacts négatifs liés à une vidange et réalise des études sur ce sujet depuis une bonne dizaine d'années [POU00] [ROC15]. Dans ce domaine aussi, la tendance est de limiter au maximum l'intervention humaine en milieu hyperbare et de privilégier une inspection à distance, par sécurité mais également par souci de productivité (plus grande accessibilité et durée d'intervention plus longue). Les sociétés SUB-C Marine (Feyzin - 69) ou encore Subsea Tech (Marseille - 13) proposent des ROV équipés de moyens optiques et acoustiques divers [GUI15]. Nous nous penchons également sur

l'usage de sonars frontaux haute résolution. Un sonar sectoriel n'étant pas disponible pour la première année d'expérimentation, un sonar latéral TriTech a été détourné de son usage classique (visualiser une bande latérale du fond marin) et fixé sur perche pour balayer la paroi du barrage à partir d'une prame. Malgré la qualité moyenne de ce sonar, l'image sonar rectifiée fait bien apparaître des traces de niveaux de gris plus foncés liées au niveau des reprises ponctuelles de l'enduit d'étanchéité (voir Figure 37). Ces résultats encourageants seront enrichis durant la nouvelle année scolaire 2017-2018 par l'inspection à l'aide des sonars BlueView M900 et M2250-130 dont les résolutions aideront non seulement à la détection mais également à l'analyse des réparations. A terme, lorsque ces capteurs seront placés sur un ROV, l'examen de la vantellerie immergée pourra être envisagé pour mesurer l'état d'envasement, détecter la présence d'embâcles ou encore faire du contrôle non destructif sur les organes de vantellerie. Notre contribution se place dans ce contexte à la fois au niveau du choix matériel (porteur, capteur) et algorithmique (stratégie d'exploration et analyse/traitement des données). A plus long terme, une approche par moyen acoustique suivie d'une inspection fine sur les derniers mètres voire centimètres en milieu très turbide par moyen optique est en particulier à étudier. Cette combinaison acoustique/optique est judicieuse tant au niveau du positionnement du véhicule que de la fusion des données avec un bouclage inévitable entre les deux. Le traitement de données sonar pourra s'inspirer de quelques travaux académiques récents et connexes comme la reconstruction de scènes sous-marines ou encore le mosaïquage [BRA14] [HUR14].

5 Conclusion

Au terme de ce manuscrit, le lecteur aura perçu le caractère très applicatif des travaux présentés. Ceci est le reflet de nombreuses opportunités de travail pour l'industrie de défense et la DGA (Direction Générale de l'Armement). L'ENSTA Bretagne sous tutelle du ministère de la défense est un cadre privilégié pour mener des études sur la contremesure dans le domaine marin. La DGA et l'école ont par ailleurs soutenu mon projet de recherche à l'étranger, me permettant d'élargir mon réseau à l'international et d'initier des collaborations pérennes. La communauté de la reconnaissance de cibles sous-marines est petite mais a la volonté de partager et d'échanger sur les challenges futurs comme l'attestent les éditions du workshop iSAAM et du colloque consacré à la détection et la classification de cibles sous-marines ainsi que nos diverses collaborations à l'international. Ces collaborations et les quelques thèses encadrées ont dégagé le temps nécessaire pour explorer des pistes académiques comme l'exploitation des données multi-passe ou encore la qualification de l'environnement proche de la cible pour améliorer sa reconnaissance. L'acquisition de données SAS lors de passes répétitives sur une même zone est opérationnellement exigeant mais à la portée d'un AUV destiné à des missions de défense et équipé pour cela de systèmes de navigation de précision suffisante. Nous avons vu que le traitement cohérent de ces données permet de détecter des changements aussi subtils qu'un simple mouvement de sédiments, ce qui peut survenir lors de l'enfouissement de dispositifs artisanaux qui constituent aujourd'hui une réelle menace. Une autre piste est l'amélioration du processus ATR qui, dans son approche classique d'extraction de l'information et classification supervisée, est vulnérable à la forte variabilité de l'environnement marin. Une meilleure compréhension des phénomènes environnementaux impactant l'ATR est la voie explorée dans ce manuscrit pour mieux répondre à cette problématique et maintenir un bon niveau de performance, en détection notamment.

Le système sonar ne peut être dissocié du porteur et de l'ensemble des autres capteurs. Les données de navigation sont aujourd'hui communément prises en compte pour rendre les traitements de données sonar plus robustes. En comparaison avec un navire de surface, la replanification est facilitée par la maniabilité d'un AUV. Par ailleurs, la stabilité des AUVs permet des gains élevés de traitement, pour le SAS par exemple... Si on ajoute la discrétion, ce sont autant d'atouts qui poussent les états-majors à plébisciter l'usage de ces véhicules tout en maintenant les opérateurs hors de danger. C'est dans ce contexte que nous nous sommes intéressés à des problématiques propres à la robotique mobile comme la détection d'obstacles et plus récemment l'enjeu de la boucle perception-

action pour une mission plus efficace, problématiques qui entrent parfaitement dans le périmètre des activités de la nouvelle équipe PRASYS du Lab-STICC.

Les contributions principales des ces dernières années de recherche ont été rappelées et complétées. Des limites et des perspectives ont été données : rôle de l'environnement dans le processus ATR, potentiel des données multi-passe, paramétrage automatique des fenêtres de calcul, enjeu du portage des algorithmes sur AUV...

Enfin, la double casquette de l'enseignant et du chercheur offre des opportunités de travail variées allant du contrat au projet pédagogique de terrain. La polyvalence de ce métier est une richesse et source de motivation. L'habilitation à diriger des recherches ajoutera un degré de liberté à mes activités dont j'espère faire profiter les étudiants qui souhaiteront avec moi améliorer notre perception de l'environnement marin.

6 Bibliographie

- [BAC04] S. T. Bachelor, R. L. Thompson and J. L. Seawall, “Acoustic Imaging with Time-Frequency Beamforming, Blazed Arrays, and Other Array Designs”, *US Patent* number US-4, 022,377, July 12, 2004.
- [BAZ08a] S. Bazeille, *Vision sous-marine monoculaire pour la reconnaissance d’objets*, Thèse de l’Université de Bretagne Occidentale, soutenue le 1^{er} octobre 2008.
- [BAZ08b] S. Bazeille, I. Quidu, L. Jaulin & J. P. Malkasse, « Une méthode de pré-traitement automatique pour le débruitage des images sous-marines », *Revue Traitement du Signal*, Vol. 25, No. 1 & 2, 2008/02-04.
- [BEL01] E. Belcher, B. Matsuyama and G. Trimble, “Object identification with acoustic lenses”, *OCEANS Conference*, Honolulu, HI, USA, Nov. 2001.
- [BEL02] A. Bellettini and M. A. Pinto, “Theoretical accuracy of synthetic aperture sonar micronavigation using a Displaced phase-center antenna”, *IEEE Journal of Oceanic Engineering*, 2002.
- [BRA14] N. Brahim, *Reconstruction tridimensionnelle de scènes sous-marines à partir de séquences d’images acquises par des caméras acoustiques*, Thèse de doctorat, Télécom Bretagne, Université de Rennes 1, 2014.
- [BRO14] D. C. Brown and A. P. Lyons, “Spatial Coherence Theory and Its Application to Synthetic Aperture Systems”, *Proceedings of the Institute of Acoustics*, Vol.36, Pt.I, 2014.
- [CHA94] M. J. Chantler, *The effect of variation in illuminant direction on texture classification*, Ph.D. thesis, Department of Computing and Electrical Engineering, Heriot-Watt University, August 1994.
- [CHA05] C. Chailloux and B. Zerr, « Non Symbolic Methods to register SONAR images », *Oceans 2005 – Europe*, 2005.
- [CHA07] C. Chailloux, *Recalage d’images sonar par appariement de régions. Application à la génération d’une mosaïque*, Thèse de Doctorat, ENST Bretagne, Brest, June 2007.
- [CHA08] P. Chapple, *Automated Detection and Classification in High-resolution Sonar Imagery for Autonomous Underwater Vehicle Operations*, DSTO-GD-0537 report, December 2008.
- [CHA11] C. Chailloux, J.-M. Le Caillec, D. Guériot and B. Zerr, “Intensity-Based Block Matching Algorithm for Mosaicing Sonar Images”, *IEEE Journal of Oceanic Engineering*, Vol. 36, No. 4, October 2011.

[CHA14] M. Chabah, N. Burtle, J.-P. Malkasse, G. Le bihan and B. Quéllec, « SAMDIS : a new SAS imaging system for AUV », in *proceedings of Quantitative Monitoring of the Underwater Environment, Results of the International Marine Science and technology Event MOQESM'14, Brest, France*, Springer, 2014.

[CHE09] N. A. McChesney, *Three-Dimensional Feature Reconstruction with Dual Forward Looking Sonars for Unmanned Underwater Vehicle Navigation*, Master's Thesis, Naval Postgraduate School, Monterey, CA, 2009.

[CLA92] S. J. Clarke, *The analysis and synthesis of texture in sidescan sonar data*, Ph.D. thesis, Dept. Electrical & Electronic Engineering, Heriot-Watt University, May 1992.

[CLE14] O. Clert, *Etude et optimisation pour les systèmes robotiques pour l'identification de mines sous-marines*, Mémoire de fin d'études (FIPA 2014), DGA Techniques Navales Brest, 14 août 2014.

[COI08] E. Coiras, J. Groen, B. Evans and M. Pinto, "Automatic change detection for the monitoring of cluttered underwater areas", *Proceedings of the 1st International Conference and Exhibition on Waterside Security (WSS 2008)*, Technical University of Denmark, Copenhagen, Denmark, 25-28 August 2008.

[COI09a] E. Coiras, J. Groen, "3D Target Shape from SAS Images Based on a Deformable Mesh", *3rd International Conference & Exhibition on Underwater Acoustic Measurements: Technologies & Results*, Nafplion, Greece, 2009.

[COI09b] E. Coiras, J. Groen, "Simulation and 3D Reconstruction of Side-looking Sonar Images", *Book chapter in "Advances in Sonar Technology"*, In-Tech Books, Vienna, Austria, pp.1- 14, 2009.

https://www.intechopen.com/books/advances_in_sonar_technology/simulation_and_3d_reconstruction_of_side-looking_sonar_images

[COO01] D. A. Cook, J. T. Christoff and J. E. Fernandez, "Broadbeam Multi-Aspect Synthetic Aperture Sonar", *OCEANS conference*, Honolulu, HI, USA, Nov. 2001

[COU12] M. Couillard and D. P. Williams, "Performance assessment of an adaptative AUV survey", in *Proc. of the International Conference on Underwater Remote Sensing (ICoURS) - Detection and Classification of Underwater Targets, Brest, France, 8-11 October 2012*, Cambridge Scholars publishing, 2014.

[DAN98] S. Daniel, *Fusion multisource appliquée à la reconnaissance d'objets dans le milieu sous-marin*, Thèse de doctorat, ENST Bretagne, 1998.

[DAN12] O. Daniell, Y. Petillot, and S. Reed, "Unsupervised sea-floor classification for automatic target recognition," in *Proc. of the International Conference on Underwater Remote Sensing (ICoURS) - Detection and Classification of Underwater Targets, Brest, France, 8-11 October 2012*, Cambridge Scholars publishing, 2014.

- [DAN14] O. Daniell, Y. Petillot, S. Reed, J. Vazquez, and A. Frau, "Reducing false alarms in automated target recognition using local sea-floor characteristics," in *Sensor Signal Processing for Defence (SSPD) conference*, 2014.
- [DEN11] M. Deng, Q. Liu, T. Cheng and Y. Shi, "An adaptive spatial clustering algorithm based on Delaunay triangulation", *Computers, Environment and Urban Systems* 35, pp. 320-332, 2011. [DIL14] J. Dillon and V. Myers, "Coherence estimation for repeat-pass interferometry", *Oceans conference*, St. John's, 2014.
- [DOB97] G. Dobeck, J. Hyland, L. Smedley, "Automated Detection/Classification of Seamines in Sonar Imagery," *Proc. of SPIE*, vol. 3079, Apr. 1997.
- [DOB05] G. Dobeck, "Image Normalization using the Serpentine Forward-Backward Filter: Application to high-resolution sonar imagery and its impact on mine detection and classification," *Proc. of SPIE 2005*, vol. 5794, Apr 2005.
- [DOB10] G. Dobeck, "Adaptive large-scale clutter removal from imagery with application to high-resolution sonar imagery," *Proc. of SPIE 2010*, vol. 7764, Apr. 2010.
- [DON00] P. Dong, "Test of a new lacunarity estimation method for image texture analysis", *International Journal of Remote Sensing*, Vol. 21, N° 17, pp. 3369 – 3373, 2000.
- [DUB08] J. Dubberley, M. Gendron and M. Lohrenz, "Automated change detection with area matching", *Acoustics'08 Paris*, June 29-July 4, 2008.
- [ELB12a] A. Elbergui, I. Quidu, B. Zerr and B. Solaiman, "Model Based Classification of Mine-Like Objects in Sidescan Sonar using the Highlight Information", *Proceedings of the 11th European Conference on Underwater Acoustics*, Edinburgh, Scotland, UK, 2-6 July 2012.
- [ELB12b] A. Elbergui, *Amélioration des techniques de reconnaissance automatique de mines marines par analyse de l'écho à partir d'images sonar haute résolution*, Thèse de Doctorat, Université de Bretagne Occidentale, ENSTA Bretagne, Brest, Décembre 2012.
- [FAK13] E. Fakiris, D. P. Williams, M. Couillard and W. L. J. Fox, "Sea-floor acoustic anisotropy and complexity assessment towards prediction of ATR performance", in *1st International Conference and Exhibition on Underwater Acoustics*, Corfu, Greece, June 2013, pp. 1277-1284.
- [FAW10] J. Fawcett, V. Myers, D. Hopkin, A. Crawford, M. Couillard and B. Zerr, "Multiaspect classification of sidescan sonar images: Four different approaches to fusing single-aspect information", *IEEE Journal of Oceanic Engineering*, Vol. 35, No. 4, pp. 863-876, October 2010.
- [FAW12] J. Fawcett, "Automatic target recognition methods for sidescan sonar images: the advances and the challenges", in *Proc. of the International Conference on*

Underwater Remote Sensing (ICoURS) - Detection and Classification of Underwater Targets, Brest, France, 8-11 October 2012, Cambridge Scholars publishing, 2014.

[FEL01] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, 2001.

[FER12] J. Ferrand and N. Mandelert, "Change detection for MCM survey mission," in *Proc. of the International Conference on Underwater Remote Sensing (ICoURS) - Detection and Classification of Underwater Targets, Brest, France, 8-11 October 2012*, Cambridge Scholars publishing, 2014.

[FIS58] W. D. Fisher, "On grouping for maximum homogeneity", *American Statistical Association Journal*, vol. 53, pp789-798, 1958.

[FOR08] A. Fortin, P. Simard and L. Simard, *An Automated Change Detection System for Imaging Sonars*, Contract Report, DRDC Atlantic CR 2008-279, December 2008.

[FOU93] G. R. Fournier, D. Bonnier, J. L. Forand and P. Pace, "Range-gated underwater laser imaging system", *Optical Engineering* 32(9), pp. 2185-2190, Sept. 1993.

[FUR06] T. H. Furukawa, *Autonomous Underwater Vehicle Utilizing a Forward Looking Sonar*, Master's Thesis, Naval Postgraduate School, Monterey, CA, 2006.

[GAN15] V. Ganesan, M. Chitre, E. Brekke, "Robust Underwater Obstacle Detection for Collision Avoidance", *Autonomous Robots*, pp 1-21, 29 December 2015.

[GEI14] M. Geilhufe and O. Midtgaard, "Quantifying the complexity in sonar images for MCM performance estimation," *Proc. of 2nd Underwater Acoustics Conference*, 22-27 June, 2014, Island of Rhodes, Greece.

[GMI10] T. G-Michael and J. D. Tucker, "Canonical correlation analysis for coherent change detection in synthetic aperture sonar imagery", *Proc of IEEE SAR/SAS 2010 Conference*, vol. 32, no. 4, pp. 117-122, Sep. 2010.

[GRO10] J. Groen, E. Coiras & D. P. Williams, "False-alarm reduction in mine classification using multiple looks from a synthetic aperture sonar", *Proc. of OCEANS'10 conference*, Sydney, 2010.

[GRO17] J. Groen, W. Jans, Y. Pailhas, V. Myers, "Multi-look processing for Coherent Change Detection with Synthetic Aperture Sonar", in *4th Underwater Acoustics Conference*, 3-8 September, 2017, Skiathos Island, Greece.

[GUE96] D. Gueriot, E. Maillard, and J.-P. Kernin, "Sonar image registration through symbolic matching: a fuzzy local transform approach using genetic algorithms," *Proc. of OCEANS'96, Prospects for the 21st Century*, vol. 3, Sep 1996.

[GUE07] D. Gueriot, C. Sintes and R. Garello, "Sonar Data Simulation based on Tube Tracing", *Proc. of OCEANS Europe*, June 18-21, Aberdeen, Scotland, UK, 2007.

- [GUI15] F. Guillot, « Quand la réglementation en matière de sécurité des ouvrages hydrauliques conduit à mettre en oeuvre des technologies innovantes », *Colloque Comité Français des Barrages et Réservoirs : « Vantellerie, contrôle-commande, télécom et alimentations électriques pour des barrages plus sûrs »*, Chambéry, 2-3 décembre 2015.
- [HAY09] M. P. Hayes and P. T. Gough, “Synthetic Aperture Sonar: A Review of Current Status”, *IEEE Journal of Oceanic Engineering*, Vol. 34, No. 3, July 2009.
- [HOE98] U. Hoelscher-Hoebing and D. Kraus, “Unsupervised image segmentation and image fusion for multi-beam/multi-aspect sidescan sonar images”, *Proc. OCEANS’98 conference*, Nice, France, 28 Sept-1 Oct 1998.
- [HUN09] A. J. Hunter and S. Dugelay, “Exploiting the coherence of seabed scattering for repeat-pass SAS micro-navigation”, *Proceedings of the Institute of Acoustics*, 2009.
- [HUN16] A. J. Hunter, S. Dugelay and W. L. J. Fox, “Repeat-pass synthetic aperture sonar micronavigation using redundant phase center arrays”, *IEEE Journal of Oceanic Engineering*, Vol. 41, No. 4, October 2016.
- [HUR14] N. Hurtós, *Forward-Looking Sonar Mosaicing for Underwater Environments*, PhD Thesis, University of Girona, 2014.
- [JAF01] J. S. Jaffe, K. D. Moore, J. McLean and M. P. Strand, “Underwater Optical Imaging: Status and Prospects”, in *Oceanography*, Vol. 14, No. 3, 2001.
- [JAF13] J. S. Jaffe, “A Contemporary Perspective on Underwater Optical Imaging”, *Proc. of OCEANS Europe*, Bergen, 2013.
- [JAN87] D. Jan, « Traitement d’images numérique appliqué aux sonars à haute résolution », *Colloque GRETSI 87*, Nice, France, pp. 357-160, 1987.
- [JI05] S. Ji, X. Liao and L. Carin, “Adaptive Multiaspect Target Classification and Detection with Hidden Markov Models”, *IEEE Sensors Journal*, Vol. 5, No. 5, October 2005.
- [KAR14] I. Karoui, I. Quidu and M. Legris, « Acoustic obstacle detection for safe AUV surfacing », *2nd Underwater Acoustics Conference*, Island of Rhodes, Greece, Sunday 22nd through Friday 27th June, 2014.
- [KAR15a] I. Karoui, I. Quidu and M. Legris, « Automatic spatial clustering and tracking of sea surface obstacles in Forward Looking Sonar sequences », *3rd Underwater Acoustics Conference*, Greek island of Crete, 21st – 26th June, 2015.
- [KAR15b] I. Karoui, I. Quidu and M. Legris, "Automatic sea surface obstacle detection and tracking in Forward Looking Sonar image sequences", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 53, no 8, August 2015.

- [KEL89] J. M. Keller, S. Chen, R. M. Crownover, "Texture description and segmentation through fractal geometry", *Computer vision, graphics, and image processing*, vol. 45, n° 2, Feb. 1989, p. 150-166.
- [KES02] R. T. Kessel, "Using sonar speckle to identify regions of interest and for mine detection," *proceedings of SPIE*, vol. 4742, pp. 440–451, Orlando, Florida, USA, April 2002.
- [KRU03] N. Krüger and M. Felsberg, "A continuous formulation of intrinsic dimension," in *British Machine Vision Conference* (R. Harvey and A. Bangham, eds.), 2003.
- [LEB06] I. Leblond, *Recalage à long terme d'images sonar par mise en correspondance de cartes de classification automatique des fonds*, Thèse de Doctorat, Université de Bretagne Occidentale, ENSIETA, Brest, February 2006.
- [LEB10] I. Leblond, I. Quidu et C. Berron, "Mise en œuvre de stratégies prédictives sur les vues à acquérir pour la classification multi-vues d'objets immergés à partir d'images SAS", *17ème conférence en Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, Caen, France, 19-22 Janvier 2010.
- [LER11] C. Leroy and B. Zerr, "Estimation of the "best angle for next view" in a sequential multiple aspect classification scheme", *4th international conference and exhibition on "underwater acoustic measurements: technologies & results"*, 2011.
- [LIN91] L.M. Linnett, *Multi-texture image segmentation*, Ph.D. thesis, Dept. Computing and Electrical Engineering, Heriot-Watt University, March 1991.
- [LOW99] D. Lowe, "Object recognition from local scale-invariant features", *7th IEEE International Conference on Computer Vision*, vol. 2, pp. 1150-1157, 1999.
- [LYO10] T. Lyons and D. Brown, "Temporal variability of seafloor roughness and its impact on coherent change detection," *Proceedings of the International Conference on SAS and SAR*, Lerici, Italy, 2010.
- [LYO13] A. P. Lyons and D. C. Brown, "The Impact of the Temporal Variability of Seafloor Roughness on Synthetic Aperture Sonar Repeat-Pass Interferometry", *IEEE Journal of Oceanic Engineering*, Vol. 38, No. 1, January 2013.
- [MAG01] A. Maguer, Y. Gao, R. Vesetas, M. Narozny and G. Manzie, "AMI: A 3D mine imaging sonar for mine identification in turbid waters", *Proc. UDT Europe*, 2001.
- [MAL00] J. P. Malkasse, « Identification tool by sonar means with synthetic aperture sonar and 3D reconstruction", *Proc. UDT Europe 2000*, London, June 2000.
- [MAN83] B. B. Mandelbrot, *The fractal geometry of nature*, W.H. Freeman, New York, 1983.
- [MAS08] T. D. Masek, *Acoustic image models for navigation with forward-looking sonars*, Master's Thesis, Naval Postgraduate School, December 2008.

- [MID11] Ø. Midtgaard, R. E. Hansen, T. O. Sæbø, V. Myers, J. R. Dubberley and I. Quidu, “Change Detection Using Synthetic Aperture Sonar: Preliminary Results from the Larvik Trial”, *OCEANS’2011 MTS/IEEE*, Kona, Hawai, Sept 19-22, 2011.
- [MID14] Ø. Midtgaard, I. Alm, T. O. Sæbø, M. Geilhufe, R. E. Hansen, “Performance assessment tool for AUV based mine hunting”, *Proc. Institute of Acoustics*, vol. 36, pp. 120-129, September 2014.
- [MIG98] M. Mignotte, *Segmentation d’images sonar par approche markovienne hiérarchique non supervisée et classification d’ombres portées par modèles statistiques*, Thèse de doctorat, Université de Bretagne Occidentale, 1998.
- [MIG05] P.Y. Mignotte, M. Lianantonakis, and Y. Petillot, "Unsupervised registration of textured images: applications to sidescan sonar", *Oceans 2005 – Europe*, 2005.
- [MYE02] V. L. Myers, “Decision trees for computer-aided detection and classification (CAD/CAC) of mines in sidescan sonar imagery”, *Technical Memorandum*, DRDC Atlantic TM 2002-144, December 2002.
- [MYE09a] V. Myers, *A data set for change detection in port environments using sidescan sonar*, Technical memorandum, DRDC Atlantic TM 2008-247, 2009.
- [MYE09b] V. Myers, A. Fortin and P. Simard, “An Automated Method for Change Detection in Areas of High Clutter Density Using Sonar Imagery”, *Underwater Acoustics Conference 2009*, Nafplion, Greece, June 2009.
- [MYE12] V. Myers and D. P. Williams, “Adaptive Multiview Target Classification in Synthetic Aperture Sonar Images Using a Partially Observable Markov Decision Process”, *IEEE Journal of Oceanic Engineering*, Vol. 37, No. 1, January 2012.
- [MYE13] V. Myers, I. Quidu, T. Sæbø, Ø. Midtgaard and R. Hansen, “Results and analysis of coherent change detection experiments using repeat-pass synthetic aperture sonar images”, *Underwater Acoustics Conference*, Corfu, Greece, June 2013.
- [MYE17a] V. Myers, I. Quidu, T.O. Sæbø, R.E. Hansen and B. Zerr, "Computing navigation corrections for co-registration of repeat-pass synthetic aperture sonar images", *Acoustics'17 Boston, Journal of the Acoustical Society of America*, 141:5, 4042-4042, 2017.
- [MYE17b] V. Myers, J. Groen, H. Schmaljohann, I. Quidu and B. Zerr, “Multi-look processing for Coherent Change Detection with Synthetic Aperture Sonar”, *Underwater Acoustics Conference*, Skiathos, Greece, 2017.
- [NEL12] J. D. B. Nelson and N. G. Kingsbury, “Fractal dimension, wavelet shrinkage and anomaly detection for mine hunting”, *IET Signal Processing*, vol. 6, pp. 484-493, 2012. ISSN 1751-9675.

- [NIC16a] F. Nicolas, A. Arnold-Bos, I. Quidu and B. Zerr, “Rigid Sonar Tracks Registration For MCM Survey Missions”, *proceedings of Undersea Defence Technology conference*, Oslo, Norway, June 1-3, 2016.
- [NIC16b] F. Nicolas, A. Arnold-Bos, I. Quidu and B. Zerr, “Fast Fourier-Based Block-Matching Algorithm for Sonar Tracks Registration in a Multi-Resolution Framework”, in *MOQESM’16 conference*, Brest, October 2016.
- [PAD12] D. Padfield, “Masked object registration in the fourier domain,” *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2706–2718, May 2012.
- [PAI10] Y. Pailhas, Y. Petillot and C. Capus, “High-Resolution Sonars: What Resolution Do We Need for Target Recognition?”, *EURASIP Journal on Advances in Signal Processing*, 2010.
- [PER00] S. W. Perry, *Applications of Image Processing to Mine Warfare Sonar*, DSTO-GD-0237 report, February 2000.
- [PET06] Y. Petillot, S. Reed, E. Coiras, J. Bell, "A framework for evaluating underwater Mine detection and classification algorithms using augmented reality", *IEEE journal of Oceanic Engineering*, 2006.
- [PIC15a] L. Picard, I. Quidu, M. Legris, “Rectilinear paths detection of target in the water column by using a three-dimensional Hough transform of forward looking sonar data”, 3rd *Underwater Acoustics Conference*, Greek island of Crete, 21st – 26th June, 2015.
- [PIC15b] L. Picard, A. Baussard, G. Le Chenadec, and I. Quidu, “Potential of the intrinsic dimensionality for characterizing the seabed in the ATR context,” in *IEEE/OES & MTS OCEANS conference*, Genoa, Italy, May 2015.
- [PIC16] L. Picard, A. Baussard, G. Le Chenadec, and I. Quidu, “Seafloor characterization for ATR applications using the monogenic signal and the intrinsic dimensionality”, in *IEEE/OES & MTS OCEANS conference*, Monterey, CA, USA, September 2016.
- [PIC17] L. Picard, *Prise en compte de l’environnement marin dans le processus de reconnaissance automatique de cibles sous-marines*, thèse de l’Université de Bretagne Occidentale, soutenue le 18 mai 2017.
- [POU00] M. Poupert, P. Bénéfice and M. Plutarque, « Subaquatic Inspections of EDF (Electricité de France) Dams », *Proc. OCEANS’2000 MTS/IEEE*, Providence, Rhode Island, Sept 2000.
- [PRE06] M. Preiss and N. Stacy, *Coherent Change Detection: Theoretical Description and Experimental Results*, technical report DSTO-TR-1851, August, 2006.
- [PUY00] P. Puyhabilier, « La modernisation des chasseurs de mines tripartites », *L’Armement*, n° 71, octobre 2000.

- [QUI00] I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé, « Mine Classification based on raw sonar data: an approach combining Fourier Descriptors, Statistical Models and Genetic Algorithms », *Proc. OCEANS'2000 MTS/IEEE*, Providence, Rhode Island, Sept 2000.
- [QUI01a] I. Quidu, *Classification multi-vues d'un objet immergé à partir d'images sonar et de son ombre portée sur le fond*, Thèse de doctorat, Université de Bretagne Occidentale, Brest, soutenue le 11 décembre 2001.
- [QUI01b] I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé, « A 2-D Filter Specification for Sonar Image Thresholding », *Advanced Concepts for Intelligent Vision Systems (ACIVS'2001) conference*, Baden-Baden, Germany, July 30 – August 3, 2001.
- [QUI03] I. Quidu, J.P. Malkasse et G. Burel, « Classification de mines sous-marines à partir de l'image sonar brute : caractérisation du contour de l'ombre portée par algorithme génétique », *Revue Traitement du Signal*, Vol. 20, No 1, 2003.
- [QUI05] I. Quidu, N. Burette, J-P. Malkasse et F. Florin, « Automatic classification for MCM systems », *IEEE Conf. Oceans'05 Europe*, Brest, 20-23 juin 2005.
- [QUI07] I. Quidu, A. Héret, Y. Dupas and S. Lefèvre, « AUV (REDERMOR) obstacle detection and avoidance experimental evaluation », *IEEE OCEANS'07 conference*, Aberdeen (Scotland), 18-21 June 2007.
- [QUI09] I. Quidu, Y. Dupas & L. Jaulin, “Theoretical optical flow for target position prediction on FLS images”, *Underwater Acoustic Measurement (UAM'09)*, Nafplion, Greece, 21-26 June 2009.
- [QUI10] I. Quidu, A. Bertholom and Y. Dupas, “Ground obstacle tracking on forward looking sonar images”, *European Conference on Underwater Acoustics (ECUA) 2010*, Istanbul, Turkey, July 2010.
- [QUI11] I. Quidu, “Incoherent change detection using amplitude sidescan sonar images”, *Proceedings of the 11th European Conference on Underwater Acoustics*, Edinburgh, Scotland, UK, 2-6 July 2012.
- [QUI12a] I. Quidu, “Incoherent change detection using amplitude sidescan sonar images”, *Proceedings of Meetings on Acoustics (POMA)*, 2012.
- [QUI12b] I. Quidu, V. Myers, Ø. Midtgaard and R. E. Hansen, “Subpixel image registration for coherent change detection between two high resolution sonar passes”, in *Proc. of the International Conference on Underwater Remote Sensing (ICoURS) - Detection and Classification of Underwater Targets*, Brest, France, 8-11 October 2012, Cambridge Scholars publishing, 2014.

- [QUI12c] I. Quidu, L. Jaulin, A. Bertholom and Y. Dupas, « Robust Multitarget Tracking in Forward Looking Sonar Image Sequences using Navigational Data », *IEEE Journal of Oceanic Engineering*, Vol. 37, No 3, July 2012.
- [RAV92] R. Ravid, N. Levanon, “Maximum-Likelihood CFAR for Weibull background”, *IEE Proceedings of Radar and Signal Processing*, vol. 139, pp. 256 -264, 1992.
- [REE00] S. Reed, J. Bell and Y. Petillot, “Unsupervised segmentation of object shadow and highlight using statistical snakes”, *Proceedings of GOATS 2000 conference*, NURC, La Spezia, Italy, August 2000.
- [REE04a] S. Reed, Y. Petillot and J. Bell, “Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information”, *IEE Proc.-Radar Sonar Navig.*, Vol. 151, No. 1, February 2004.
- [REE04b] S. Reed, Y. Petillot and J. Bell, “Model-based approach to the detection and classification of mines in sidescan sonar”, in *Applied optics*, 43(2) : 237–246, 2004.
- [RIG93] E. Rignot and J. Van Zyl, “Change Detection Techniques for ERS-1 SAR Data”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, No. 4, July 1993.
- [RIV91] P. Rives, *La relation Perception/Action en robotique : vers une approche « commande en boucle fermée sur l’environnement »*, Rapport d’Habilitation à Diriger les recherches, INRIA, centre de Sophia-antipolis, 1991.
- [ROC15] J.-P. Roche, M. Tron et F. Mauris, «Maîtrise de la sûreté des ouvrages, retour d’expérience sur l’utilisation des drones pour l’inspection des ouvrages hydrauliques d’EDF », *Congrès Société Hydrotechnique de France : «Drones et hydraulique»*, Paris, avril 2015.
- [SAE11a] T. O. Sæbø, R. E. Hansen, H. J. Callow, and S. A. V. Synnes, “Coregistration of synthetic aperture sonar images from repeated passes,”in *4th international conference and exhibition on "underwater acoustic measurements: technologies & results"*, 2011.
- [SAE11b] T. O. Sæbø, *Seafloor Depth Estimation by means of Interferometric Synthetic Aperture Sonar*, PhD thesis of the University of Tromso, Norvège, September 2010.
- [SAE13] T. O. Sæbø, S. A. V. Synnes, and R. E. Hansen, “Wideband interferometry in synthetic aperture sonar,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 51, no. 8, pp. 4450–4459, Aug. 2013.
- [SAW10] J. Sawas, Y. Petillot, and Y. Pailhas, “Cascade of boosted classifiers for rapid detection of underwater objects,” in *Proceedings of the European Conference on Underwater Acoustics*, Istanbul, Turkey, 2010.
- [SCH96] F. Schmitt, L. Bonnaud and C. Collet, “Contrast control for sonar pictures“, *Signal and Image Processing, SPIE’96 – Technical Conference on Application of Digital Image Processing XIX*, 2847:70-82, August 1996.

- [SIN89] A. Singh, "Digital change detection techniques using remotely sensed data," *Int. J. Remote Sensing*, vol. 10, pp. 989-1003, 1989.
- [STA84] T. Stanton, « Sonar estimates of seafloor microroughness », *Journal of the Acoustical Society of America*, 75, pp. 809-818, 1984.
- [STA96] B. Stage and B. Zerr, « Application of the Dempster-Shafer theory to combination of multiple view sonar images », *UDT'96 conference*, 1996.
- [STE09] D. D. Sternlicht, J. K. Harbaugh, and M. A. Nelson, "Experiments in coherent change detection for synthetic aperture sonar," *OCEANS 2009*, Biloxi, MI, 2009.
- [SYN10] S. A. V. Synnes, H. J. Callow, T. O. Sæbø, and R. E. Hansen, "Multipass coherence processing on synthetic aperture sonar data," in *Proceedings of European Conference on Underwater Acoustics (ECUA 2010)*, Istanbul, Turkey, July 2010.
- [THO96] P. Thourel, *Segmentation d'images sonar par modélisation markovienne hiérarchique et analyse multirésolution*, Thèse de doctorat, Université de Bretagne Occidentale, 1996.
- [THR01] S. Thrun, "Learning occupancy grids with forward models", *Proc. of the IEEE-RSJ Int. Conf. on Intelligent Robots and Systems*, Hawaii, HI (US), oct-nov 2001.
- [TIZ98] H.R. Tizhoosh, «Fuzzy Image Processing : Potentials and State of the Art», *Methodologies for the Conception, Design and Application of Soft Computing, proc. of IIZUKA'98*, pp. 321-324, 1998.
- [TOU99] R. Touzi, A. Lopes, J. Bruniquel and P. W. Vachon, "Coherence Estimation for SAR Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, No. 1, January 1999.
- [VAN11] P. Vandrish, A. Vardy, D. Walker, and O. Dobre, "Side-scan sonar image registration for auv navigation," in *IEEE Symposium on Underwater Technology (UT) and Workshop on Scientific Use of Submarine Cables and Related Technologies (SSC)*, April 2011, pp. 1–7.
- [WEI09] S. Wei, H. Leung and V. Myers, "An automated change detection approach for mine recognition using sidescan sonar data ", *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics*, San Antonio, TX, USA, 11-14 Oct. 2009.
- [WIL09] D. Williams and J. Groen, "Multi-View Target Classification in Synthetic Aperture Sonar Imagery", *3rd International Conference & Exhibition on Underwater Acoustic Measurements: Technologies & Results*, Nafplion, Greece, 2009.
- [WIL10a] D. P. Williams, "On Adaptive Underwater Object Detection", *International Conference on Intelligent Robots and Systems (IROS)*, 2011.

- [WIL10b] D. P. Williams and E. Coiras, "On Sand Ripple Detection in Synthetic Aperture Sonar Imagery," *Proceedings of the 35th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, Texas, pp. 1074-1077, March 2010.
- [WIL11] D. P. Williams and J. Groen, "A fast physics-based, environmentally adaptive underwater object detection algorithm", *Proc. of OCEANS 2011*, Spain, 2011.
- [WIL12] D. P. Williams, "AUV-enabled adaptive underwater surveying for optimal data collection," *Intelligent Service Robotics*, vol. 5, no. 1, pp. 33–54, 2012.
- [WIL14] D. Williams and E. Fakiris, "Exploiting environmental information for improved underwater target classification in sonar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6284–6297, 2014.
- [WIL15] D. P. Williams, "Fast Unsupervised Seafloor Characterization in Sonar Imagery Using Lacunarity", in *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 53, No. 11, November 2015.
- [ZER96] B. Zerr and B. Stage, « Three-dimensional reconstruction of underwater objects from a sequence of sonar images », *International conference on image processing (ICIP)*, Vol. 3, 1996.
- [ZER05] B. Zerr, G. Mailfert, A. Bertholom, and H. Ayreault, "Sidescan sonar image processing for AUV navigation," in *Oceans 2005 - Europe*, vol. 1, June 2005, pp. 124–130.
- [ZER09] B. Zerr, J. Fawcett and D. Hopkin, "Adaptive algorithm for sea mine classification", *Underwater Acoustic Measurements (UAM) Conference Proceedings*, 2009.

Annexe 1: segmentation orientée classification

Cette annexe contient l'article tiré de mes travaux de thèse [QUI01a] et publiés en 2003 dans la revue Traitement du Signal [QUI03]. Afin d'optimiser la procédure de classification en se concentrant sur l'objectif ultime de la chaîne de traitement, nous avons mis en œuvre un processus dynamique pour caractériser le contour de l'ombre à partir de l'image sonar brute en offrant en outre la possibilité de classifier l'objet détecté. Cette approche fait appel aux notions de modèles déformables, modèles statistiques et algorithmes génétiques.

Classification de mines sous-marines à partir de l'image sonar brute : caractérisation du contour de l'ombre portée par algorithme génétique

Mine classification based on raw sonar data: shadow contour characterization using a genetic algorithm

par I. QUIDU*, J. PH. MALKASSE*, G. BUREL**

(*) Thales Underwater Systems, Route de Sainte Anne du Portzic, CS 43614, 29238 BREST cédex 3, France

Tél : 02 98 31 39 95, 02 98 31 35 54, Fax : 02 98 31 35 50

E-mail: isabelle.quidu@fr.thalesgroup.com, jean-philippe.malkasse@fr.thalesgroup.com

(**) L.E.S.T. - UMR CNRS 6165, 6 avenue Le Gorgeu, BP 809, 29285 BREST cédex, France

Tél : 02 98 01 62 46, Fax : 02 98 01 63 95

E-mail: Gilles.Burel@univ-brest.fr

résumé et mots clés

Dans le domaine de la chasse aux mines sous-marines, l'objet détecté peut être caractérisé par son ombre portée sur le fond. L'approche classique est séquentielle : l'image sonar est tout d'abord segmentée afin d'obtenir une image binaire partageant les pixels entre la zone d'ombre et la zone de réverbération de fond, puis des attributs caractéristiques sont extraits de la silhouette 2-D correspondant à l'ombre segmentée lesquels servent à classifier l'objet en fin de traitement. À chacune des étapes sont généralement associés des pré- et/ou post-traitements visant à éviter qu'une erreur intervenant à un instant donné de la chaîne de traitement se répercute jusqu'au résultat final. Afin d'optimiser la procédure de classification en se concentrant sur l'objectif ultime de la chaîne de traitement, nous avons mis en œuvre un processus dynamique pour caractériser le contour de l'ombre à partir de l'image sonar brute en offrant en outre la possibilité de classifier l'objet détecté. Cette approche innovante fait appel aux notions de modèles déformables, modèles statistiques et algorithmes génétiques.

Image sonar, classification de mine, algorithme génétique, discrimination de contours, descripteurs de Fourier, modèles statistiques.

abstract and key words

In the context of mine warfare, detected objects can be classified from their cast shadow. A standard solution consists in segmenting the image at first (we obtain binary from grey-level image giving the label zero for pixels belonging to the shadow and the label one elsewhere), and then in performing classification from features extracted from the 2D-shape of the segmented shadow. Other pre- or post-processings are generally used to make each step more robust by avoiding a mistake to be propagated through the following steps. In this paper, to focus on the actual goal, we propose a novel approach where a dynamic segmentation scheme is fully classification-oriented. Actually, classification is performed directly from raw image data. The approach is based on the combination of deformable models, genetic algorithms, and statistical image models.

Sonar image, mine classification, genetic algorithm, contour discrimination, Fourier descriptors, statistical models.

1. introduction

– La chasse aux mines

Les systèmes de chasse aux mines utilisent des sonars de deux types. Caractérisé par une portée relativement grande et une résolution faible, le sonar détecteur permet de repérer l'objet par son écho. Grâce à une meilleure résolution et à une plus courte portée, le sonar classificateur fournit une image interprétable sur laquelle l'ombre portée de l'objet se détache.

La faiblesse des moyens militaires face aux mines marines de nouvelle génération apparut au début des années 90 lors des événements intervenus dans le golfe Persique : une modernisation des équipements vise depuis à faire face à ces mines dites

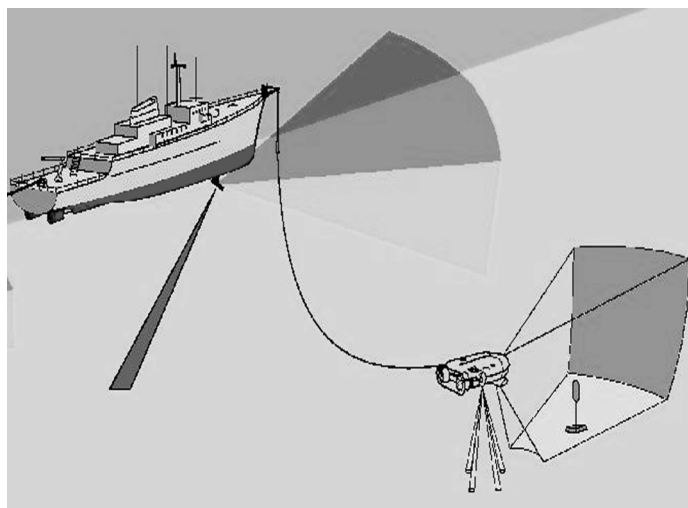


Figure 1. – Sonar de coque et sonar propulsé de l'avant.

furtives (petite taille, géométrie et structure étudiées présentant un index de cible réduit) [PUYH-00]. Les progrès technologiques de ces dix dernières années permettent de faire de l'imagerie des fonds marins avec une très haute résolution. La guerre des mines a tout naturellement exploité les performances des sonars classificateurs haute fréquence pour saisir les petits détails et finement caractériser les contours des ombres portées des objets immergés. Sur l'image obtenue apparaît l'ombre acoustique caractérisée par l'absence de réverbération de fond du fait de la présence de l'objet. La figure 2 illustre la similitude des ombres obtenues par voie acoustique et par voie optique (vue de dessus en éclairage rasant).

– La classification sur ombre classique

La classification sur ombre est un problème classique de reconnaissance de formes. La notion de forme englobe la région intérieure de l'ombre segmentée et son contour. Le processus classique de classification sur ombre d'un objet immergé consiste en trois étapes successives :

segmentation → **extraction d'attributs** → **classification**

La segmentation extrait les régions d'intérêt dans l'image, *i.e.*, dans le cas de la classification sur ombre d'images sonar, partage les pixels de l'image en deux classes : la classe « ombre » comprenant les pixels de la zone d'ombre et la classe « fond » comprenant à la fois les pixels de la zone de réverbération de fond et de l'écho. À partir de l'image binaire obtenue, des paramètres discriminants ou attributs caractéristiques sont calculés sur la forme de l'ombre ou uniquement son contour, et regroupés dans un vecteur caractéristique. L'information disponible est ainsi réduite à un nombre restreint de paramètres discriminants. Enfin, le vecteur caractéristique introduit dans le module de classification doit permettre d'affecter en finalité l'objet à une classe d'objets. La classification est faite de manière automatique, *i.e.* sans intervention de l'opérateur, et de manière supervisée nécessitant un nombre conséquent d'éléments représentatifs de chacune des classes dont on peut rapprocher l'élément à

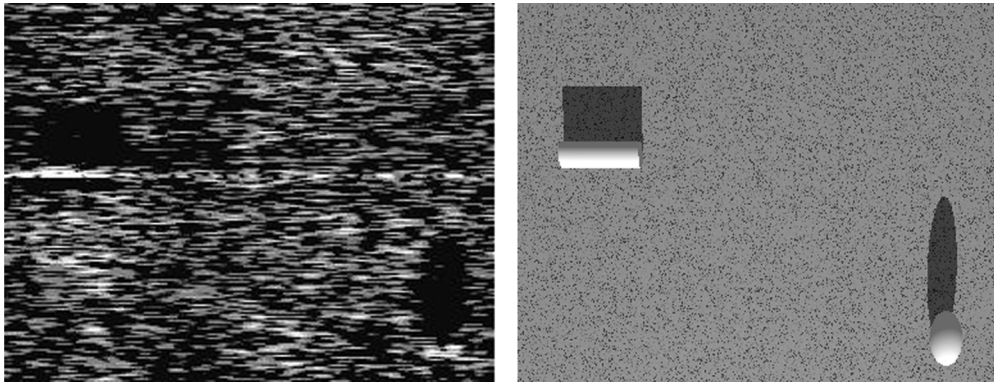


Figure 2. – Image sonar haute résolution d'une mine cylindrique et d'une mine sphérique et image vidéo de synthèse de la même scène.

classifier. À chacune des étapes du processus séquentiel ci-dessus sont associés des pré- et/ou post-traitements superflus au regard de l'objectif final de classification. Il s'agit par exemple du détail exigé sur le contour à l'étape de segmentation. Il s'agit aussi de préparer l'image pour l'étape qui suit et d'éviter qu'une erreur intervenant à un instant donné ne se répercute jusqu'au résultat final sans rétroaction possible.

– *Approche proposée*

Nous proposons une alternative novatrice qui consiste en un processus dynamique de « **segmentation orientée classification** » coordonné par un algorithme génétique. En se concentrant sur l'objectif ultime de la chaîne de traitement, une segmentation discriminante du contour est effectuée à partir des données sonar brutes. Outre l'économie de traitements superflus au vu de l'objectif visé, ce traitement « orienté solution » évite la propagation d'erreurs préjudiciables. Le résultat est double : tout en caractérisant le contour de l'ombre, on offre la possibilité de classifier l'objet détecté à partir d'une population réduite d'individus. L'exploration de l'espace des solutions est menée par un algorithme génétique. Cet algorithme manipule des solutions caractérisées par un ensemble de descripteurs de Fourier décrivant chacune un contour, sur lesquelles agissent des opérateurs génétiques classiques et nouveaux. L'évolution est réglée par un paramétrage adéquat des opérations génétiques et de l'opération de sélection des solutions. La sélection est basée sur l'adaptation du contour dans l'image au sens d'un critère statistique basé sur la modélisation du phénomène de réverbération. Il en résulte un partage des données sonar au sens du maximum de vraisemblance.

Le paragraphe suivant présente le problème et la spécificité de l'image sonar. Le troisième paragraphe explicite les contraintes sur l'objectif à atteindre et justifie le choix de l'algorithme d'optimisation. On rappelle le principe des algorithmes génétiques au paragraphe 4. La mise en œuvre de la segmentation orientée classification est détaillée au paragraphe 5 puis illustrée par quelques résultats sur images réelles et simulées au paragraphe 6.

2. problème et spécificités associées

2.1. problème posé : classification de mines marines à partir d'images sonar

Les sonars actifs émettent un signal sur la zone de fond à prospecter. La réception s'effectue sur une antenne et les signaux captés passent dans la chaîne de traitement sonar pour former l'image sonar. De par leur conception, les antennes favorisent certaines directions d'arrivée du signal et par extension, localisent précisément la cible. Dans le cas particulier de la classification sur ombre, le système sonar équipant les navires de chasse aux mines est composé d'un sonar détecteur et d'un sonar classificateur. Dans un premier temps, la recherche d'objets sur le fond ou dans le volume d'eau est effectuée par le sonar détecteur. Dans un second temps, le sonar classificateur fournit une image suffisamment résolue permettant de classer l'objet à partir de l'ombre acoustique portée sur le fond.

La caractérisation d'objets posés sur le fond marin est faite au moyen de **sonars classificateurs à haute résolution**, c'est-à-dire dont les dimensions de la cellule de résolution en distance et en gisement sont inférieures aux dimensions des objets recherchés (cf. paragraphe 2.2.2). Sur l'image obtenue apparaît **l'ombre acoustique** que porte l'objet sur le fond, laquelle est souvent plus révélatrice de sa géométrie que son écho du fait de la complexité des lois de la diffusion acoustique qui régissent ce dernier et du revêtement anéchoïque de certaines mines furtives. La présence ou l'absence de réverbération de fond sert à délimiter la zone d'ombre associée à l'objet ciblé. Ce principe de classification sur ombre est illustré figure 3 où l'on visualise l'évolution de l'amplitude du signal reçu dans une direction donnée (voie formée par le sonar) en fonction du temps [BOUV-92].

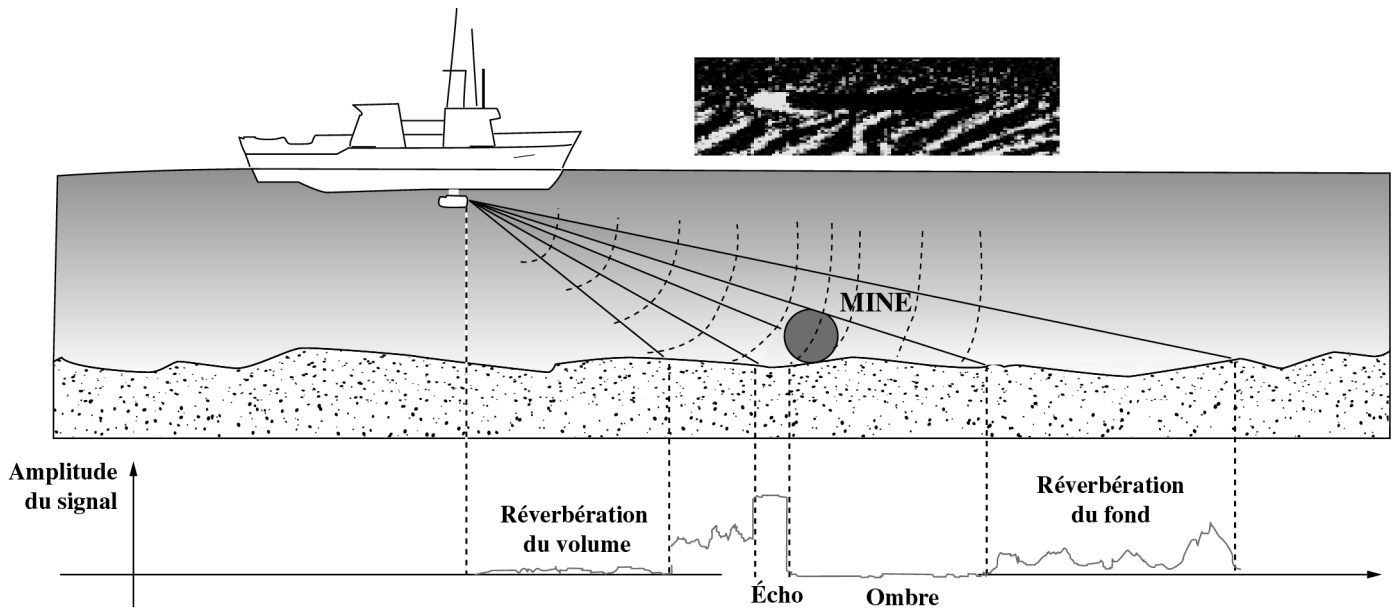


Figure 3. – Principe de classification sur ombre.

2.2. spécificités statistique et géométrique de l'image sonar

2.2.1. modélisation du bruit

La réverbération est source d'information pour la classification sur ombre. Ce phénomène propre à tous les systèmes d'imagerie cohérente apparaît lorsque la longueur d'onde incidente est grande devant la dimension des aspérités de la surface illuminée [GOOD-76]. De ce fait, l'onde réverbérée par le milieu (eau de mer ou fond) est le résultat de la combinaison tantôt destructive tantôt constructive des ondes rétrodiffusées mais déphasées aléatoirement. L'amplitude résultante de la somme de ces ondes

complexes reçues par un capteur suit alors une loi de Rayleigh [PAPO-84] caractéristique d'un bruit multiplicatif appelé chatoiement ou speckle [MAIT-01]. Deux types de réverbération nous intéressent : la réverbération de volume et la réverbération de fond générées par la somme des multiples rétrodiffusions dues, respectivement, aux inhomogénéités de l'eau (particules, bulles) et aux irrégularités du fond (micro-relief). Un chatoiement de fort niveau moyen constitue ainsi le signal utile pour la zone du fond et contraste avec un chatoiement de faible niveau moyen pour la réverbération de volume prépondérante dans la zone d'ombre. Dans la présente étude, nous négligeons une dégradation parfois observée à la frontière de la zone d'ombre. Celle-ci est causée par la remontée du niveau de bruit dû à la présence de lobes secondaires associés à la fonction de directivité de l'antenne.

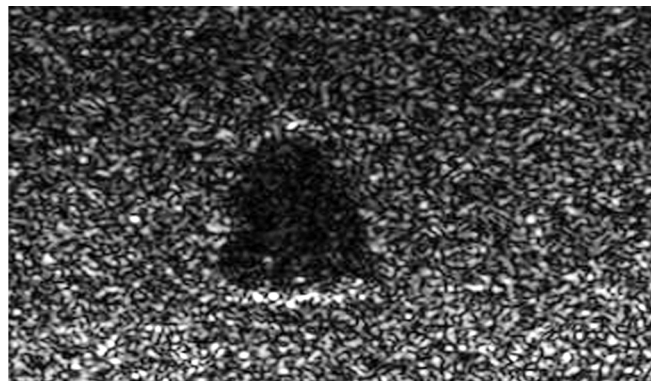
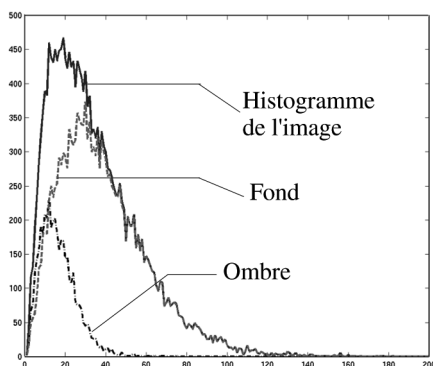


Figure 4. – Mélange des deux distributions de Rayleigh associées aux zones « ombre » et « fond ».

2.2.2. caractéristiques géométriques des images sonar

La **formation de l'image sonar** consiste à enregistrer et représenter les données sonar (ondes acoustiques réfléchies) dans le plan oblique. La dimension effective du pixel dépend à la fois des paramètres sonar et du traitement des données. D'une part, les paramètres sonar, caractéristiques de l'antenne et du signal émis, fixent la dimension de la **cellule de résolution**, c'est-à-dire l'écart minimal suivant l'axe de propagation de l'onde (distance) et angulairement (gisement) devant séparer deux points pour qu'on puisse les distinguer [SOME-93] [BLON-97]. Le traitement des données permet de passer des données sonar brutes à l'image en niveaux de gris illustrant la variation de la pression acoustique reçue par les capteurs de l'antenne et proportionnelle à la puissance du signal rétrodiffusé (c'est-à-dire en direction du sonar). Le niveau de gris d'un pixel de l'image représente ainsi l'amplitude de la somme des ondes réfléchies vers le sonar par les points du fond (diffuseurs) appartenant à la cellule de résolution. La juxtaposition au fil du temps des enregistrements obtenus dans chacune des voies pointées permet de former une carte bidimensionnelle de la zone insonifiée. En général, la dimension du pixel est plus petite que celle de la cellule de résolution (dans le cas contraire, on perd de l'information). Elle dépend de plusieurs paramètres tels que la fréquence d'échantillonnage temporel, l'échantillonnage spatial effectué lors de l'opération de formation de voies, la pondération de l'antenne, ... Ces opérations induisent un suréchantillonnage plus ou moins important de la zone insonifiée qui n'apporte pas d'information supplémentaire mais améliore la visualisation.

Par ailleurs, une distorsion géométrique classiquement observée en imagerie sonar est **l'élongation des ombres en fonction de la rasance**. Cette dernière désigne l'angle formé entre le plan oblique, *i.e.* le plan de collectage des données sonar, et le fond (*cf.* figure 3).

3. justification de l'approche proposée

Le problème de la classification sur ombre est un problème classique de reconnaissance de formes qui nécessite tout d'abord la modélisation du contour solution. Le problème peut alors être considéré comme un problème inverse résolu par l'optimisation d'un critère basé sur les caractéristiques statistiques des données sonar partagées par ce contour. Outre la spécificité statistique, la prise en compte de la spécificité géométrique de l'imagerie sonar apporte de la robustesse au traitement. La solution optimale est alors obtenue par un algorithme itératif opérant sur un espace de recherche étendu de solutions potentielles dont les contours évoluent dans l'image sonar brute sans segmentation préalable.

3.1. modélisation du contour [WALL-80] [MITC-83] [ARBT-89] [HENO-93]

Parmi les multiples procédés utilisés par les traiteurs d'images pour décrire un contour, nous avons retenu la décomposition en série de Fourier pour trois raisons :

1. Les descripteurs de Fourier ont l'avantage de décrire le contour de manière synthétique. On dispose d'une représentation multirésolution offrant la possibilité de lisser un contour bruité par transformation inverse limitée aux descripteurs d'indice inférieur ou égal à K . En ce sens, la décomposition en série de Fourier s'apparente à la décomposition multirésolution par ondelettes de l'image observée appliquée aux images sonar par Collet *et al.* : cette dernière permet de s'affranchir en partie du bruit de speckle en associant opération de décimation spatiale et filtrage [COLL-98].
2. La normalisation des descripteurs par rapport au point de départ sur le contour garantit une caractérisation unique des contours offrant la possibilité de discriminer les formes 2-D correspondantes.

Les solutions recherchées sont caractérisées par les descripteurs de Fourier qui décrivent un contour. X_k et Y_k désignent les transformées de Fourier effectuées distinctement sur les abscisses x_i et les ordonnées y_i des N points de contour équirépartis. En notant $z_i = x_i + jy_i$, on a la relation $C_k = X_k + jY_k$, où C_k est la transformée de Fourier effectuée sur les coordonnées complexes. Les descripteurs normalisés pour les deux coordonnées des points de contour sont obtenus en suivant un procédé introduit par Arbter [ARBT-89] (*cf.* annexe 1). Pour un degré de décomposition donné K , le vecteur décrivant le contour est noté $\theta_{(K)} = (\tilde{X}_k, \tilde{Y}_k)_{k \in \{1, \dots, K\}}^T$ où :

$$\begin{cases} \tilde{X}_k &= X_k \frac{C_1^{-k}}{|C_1|^{-k}} = X_k \frac{(X_1 + jY_1)^{-k}}{|X_1 + jY_1|^{-k}} \\ \tilde{Y}_k &= Y_k \frac{C_1^{-k}}{|C_1|^{-k}} = Y_k \frac{(X_1 + jY_1)^{-k}}{|X_1 + jY_1|^{-k}} \end{cases}$$

3. Une homothétie des descripteurs de Fourier se concrétise par une homothétie des points de contour. Cette propriété nous permet d'adapter la taille des contours manipulés par simple factorisation des descripteurs de Fourier par un facteur d'échelle déduit des paramètres sonar et de sa rasance (*cf.* la *taille équivalente du pixel* définie au paragraphe 3.3). En effet, les solutions potentielles proviennent d'images simulées sous des conditions (paramètres sonar, position du sonar) propres qui peuvent différer de celles de l'image à segmenter.

3.2. critère à optimiser et contrainte

La recherche du contour de l'ombre à partir des données observées est guidée par l'optimisation d'un critère statistique basé sur la modélisation du phénomène de réverbération. Cette optimisation s'accompagne d'une contrainte sur la structure du contour.

3.2.1. critère à optimiser (attache aux données)

Deux hypothèses importantes concernent la distribution des pixels [FIGU-97] :

- les niveaux de gris des pixels de l'image sont indépendamment distribués
- les deux régions, l'une extérieure au contour recherché, *i.e.* la réverbération de fond, et l'autre intérieure, *i.e.* l'ombre, sont homogènes ce qui implique que la probabilité conditionnelle associée à chacun des pixels dépend uniquement de sa localisation dans l'une ou l'autre des régions

En d'autres termes, tous les pixels appartenant à la zone d'ombre (*resp.* de réverbération de fond) suivent une distribution commune de paramètre ϕ_{in} (*resp.* ϕ_{out}). La fonction de vraisemblance à maximiser peut donc s'écrire, avec $\phi = [\phi_{in}, \phi_{out}]$ [FIGU-97][CHES-99] :

$$p(I/\phi) = \left(\prod_{(i,j) \in O} p(I(i,j)/\phi_{in}) \right) \left(\prod_{(i,j) \in R} p(I(i,j)/\phi_{out}) \right),$$

$I(i,j)$ désigne l'intensité du pixel (i,j) , O (*resp.* R) la zone d'ombre (*resp.* de réverbération de fond). La frontière (contour) entre ces deux zones est caractérisée par le vecteur $\theta_{(K)}$ des descripteurs de Fourier d'indice inférieur ou égal à K fixé.

L'objectif visé est celui d'une segmentation **non supervisée** c'est-à-dire l'estimation, sans connaissance *a priori*, pour K fixé, de $\theta_{(K)}$ à partir de l'image observée seule en maximisant la fonction de vraisemblance ci-dessus. Dans le cas des images sonar, la distribution des niveaux de gris des pixels est généralement modélisée par une loi de Rayleigh. On a donc pour chacune des régions O ou R :

$$p(I(i,j)/\phi_{r\acute{e}gion}) = \frac{I_{r\acute{e}gion}(i,j)}{\alpha_{r\acute{e}gion}^2} \exp \left\{ -\frac{[I_{r\acute{e}gion}(i,j)]^2}{2\alpha_{r\acute{e}gion}^2} \right\}$$

avec $\phi_{r\acute{e}gion} = [\alpha_{r\acute{e}gion}^2]$ où $\alpha_{r\acute{e}gion}^2$ est appelé facteur d'échelle.

L'estimée au sens du maximum de vraisemblance est [EVAN-93] :

$$\hat{\alpha}_{r\acute{e}gion}^2 = \frac{\sum_{(i,j) \in r\acute{e}gion} [I_{r\acute{e}gion}(i,j)]^2}{2 \cdot N_{r\acute{e}gion}}$$

avec $N_{r\acute{e}gion}$ le nombre de pixels appartenant à la région.

On cherchera à maximiser la fonction log-vraisemblance :

$$\begin{aligned} \ln(P(I/\phi)) &= \\ & \sum_{I(i,j) \in O} \ln I(i,j) - \sum_{I(i,j) \in O} \ln \alpha_O^2 - \sum_{I(i,j) \in O} \frac{I^2(i,j)}{2\alpha_O^2} \\ & + \sum_{I(i,j) \in R} \ln I(i,j) - \sum_{I(i,j) \in R} \ln \alpha_R^2 - \sum_{I(i,j) \in R} \frac{I^2(i,j)}{2\alpha_R^2} \\ & = \sum_{I(i,j) \in \text{Image}} \ln I(i,j) - N_O \ln \alpha_O^2 - N_O - N_R \ln \alpha_R^2 - N_R \end{aligned}$$

soit à maximiser,

$$E = -N_O \ln \alpha_O^2 - N_R \ln \alpha_R^2$$

3.2.2. régularisation

La recherche d'informations à partir des données observées ne peut généralement se faire sans l'apport de contraintes de régularité (problème inverse *mal-posé* au sens d'Hadamard [HADA-23]). La régularisation classique consiste à imposer des contraintes sur la solution recherchée sous la forme d'un terme additionnel pondéré. Dans notre cas, il n'y a pas de terme additionnel de régularisation car les déformations du contour sont implicitement pénalisées par limitation du degré K de décomposition en série de Fourier. Il en résulte une estimation au sens du maximum de vraisemblance opposée à une estimation au sens du maximum *a posteriori* qui consiste à minimiser une énergie composée de deux termes : terme de régularisation et terme d'attache aux données (modèle de bruit).

3.3. prise en compte de la spécificité géométrique des images sonar

Du fait du principe de formation des images sonar explicité au paragraphe 2.2.2, une normalisation des images est nécessaire de manière à pallier la faible rasance et la faible précision en gisement entraînant des ombres généralement disproportionnées. Elle opère donc sur deux fronts conjoints :

- rendre les pixels de l'image représentatifs de la surface réelle du sol observé :

Il existe une distorsion entre la taille du pixel (habituellement carré) de l'image et l'aire effective qu'il représente sur le sol. Ce phénomène provient du principe d'acquisition des systèmes sonar qui, pour des raisons technologiques, amènent généralement à une plus grande précision en distance qu'en gisement.

- normalisation en rasant de façon à ce que la longueur des ombres soit indépendante de celle-ci :

L'angle choisi pour la normalisation en rasant doit permettre de comparer facilement des formes géométriques associées au même type d'objet en apportant de la robustesse à la classification. La rasant choisie de 45 degrés entraîne théoriquement une longueur de l'ombre visible sur l'image sonar (suivant la distance oblique) égale à la hauteur de l'objet multipliée par $\sqrt{2}$.

La résolution du pixel dépend initialement des paramètres sonar mais est modifiée à la présente étape de normalisation. La taille équivalente du pixel finalement imposée, identique suivant les deux axes, dépend ainsi à la fois des paramètres sonar et de la normalisation de l'image.

Nota Bene : cette manipulation des données sonar est transparente pour l'évaluation des grandeurs statistiques qui s'effectue sur les données sonar initiales préalablement stockées.

3.4. choix de l'algorithme d'optimisation

Les modèles déformables ont été diversement appliqués à la reconnaissance de formes, la segmentation, le suivi d'objet en mouvement,... [COHE-93] [KERV-94] [JAIN-96] [BOSS-97] [ZHON-98] [LAM-98] [CHES-99] [MIGN-00]. Ce procédé est utilisé dans deux approches largement discutées en traitement d'images : les modèles déformables paramétriques et les contours actifs.

D'une part, le principe des **modèles déformables paramétriques** est de disposer d'une base de *modèles prototypes* ou *gabarits* (« *templates* ») décrits par un nombre donné de paramètres auxquels on associe un *ensemble de déformations*. Le choix de tels modèles est donc fortement conditionné par les connaissances *a priori* sur les objets susceptibles d'apparaître. Le modèle linéaire interagit de manière dynamique avec les particularités de l'image en faisant varier ses paramètres : son degré de liberté correspond au nombre de paramètres utilisés. Leur estimation au sens du maximum *a posteriori* est obtenue en associant au modèle une densité de probabilité *a posteriori* conditionnellement aux observations (en fait la carte de segmentation en deux classes *ombre/réverbération de fond*). La règle de Bayes nous permet de distinguer deux termes : le premier intègre les connaissances *a priori* qui comprennent la description du modèle et l'ensemble des déformations admissibles (densité de probabilité *a priori*) tandis que le second a pour effet d'attirer le modèle vers les traits remarquables de l'image (vraisemblance des observations conditionnellement aux déformations).

Le **contour actif** est d'autre part un modèle élastique dont les contraintes *a priori* de régularité sont basées sur la *théorie de l'élasticité*. Le contour de l'objet est alors représenté par une courbe ayant des propriétés mécaniques de type membrane/plaque mince. Le modèle général de contours actifs se présente donc sous la forme d'une courbe (fermée ou non), dont l'initialisation est située à proximité du contour recherché et dont l'évolution s'effectue selon un processus itératif de déformation contrôlé par un test de convergence. Le critère selon lequel le processus de déformation fait évoluer le contour a été initialement associé à la recherche d'une minimisation d'une fonctionnelle [KASS-88]. Celle-ci est généralement constituée de deux termes d'énergie qui doivent s'équilibrer : l'un est un terme de régularisation caractérisant le comportement mécanique de la courbe, soit la rigidité et la tension (énergie interne), l'autre attire la courbe vers les structures pertinentes de l'image que l'on recherche (énergie externe). Jain *et al.* montrèrent que, d'un point de vue bayésien, les énergies interne et externe peuvent respectivement s'apparenter à une densité de probabilité *a priori* et à une vraisemblance comme dans le cas des modèles déformables paramétriques [JAIN-98]. La méthode proposée tire profit des avantages des deux approches décrites ci-dessus. Les modèles déformables, d'une part, permettent de classifier la forme segmentée en l'approchant d'un prototype connu *a priori*. Les contours actifs, d'autre part, visent à segmenter la forme en recherchant son contour à partir de l'image brute en niveaux de gris. Nous proposons une technique originale utilisant la déformation simultanée de plusieurs contours pour effectuer **une segmentation orientée classification**. En d'autres termes, le domaine des solutions possibles est limité mais supposé représentatif des cas susceptibles de se présenter, moyennant des déformations du même type que celles rencontrées pour les modèles déformables. L'exploration de ce domaine a pour finalité de proposer la solution la mieux adaptée aux données sonar brutes. Une stratégie d'optimisation appropriée à ce problème est le développement d'un **algorithme génétique** qui exploite les données brutes **sans devoir passer par une étape préliminaire et irréversible de segmentation**. Cette approche s'oppose à l'approche classique séquentielle « *segmentation* → *extraction des attributs* → *classification* » comme l'explique la figure 5. L'ultime étape de classification intervient dans le rapprochement de la solution avec une version étendue du domaine des solutions possibles initial.

Le tableau 1 synthétise les caractéristiques des trois approches énoncées. Il apparaît bien que notre approche est **intermédiaire** entre les modèles déformables paramétriques et les contours actifs. Sans idée précise de la nature de l'objet ciblé (comme pour les contours actifs), on fait l'hypothèse qu'il appartient à un ensemble connu d'objets (équivalents des prototypes énoncés utilisés pour les modèles déformables paramétriques) dont on a établi un panorama des contours des ombres portées pour différents points de vue soient un total de N_p solutions ou individus. De plus, comme dans le cas des contours actifs, on travaille directement avec les données brutes.

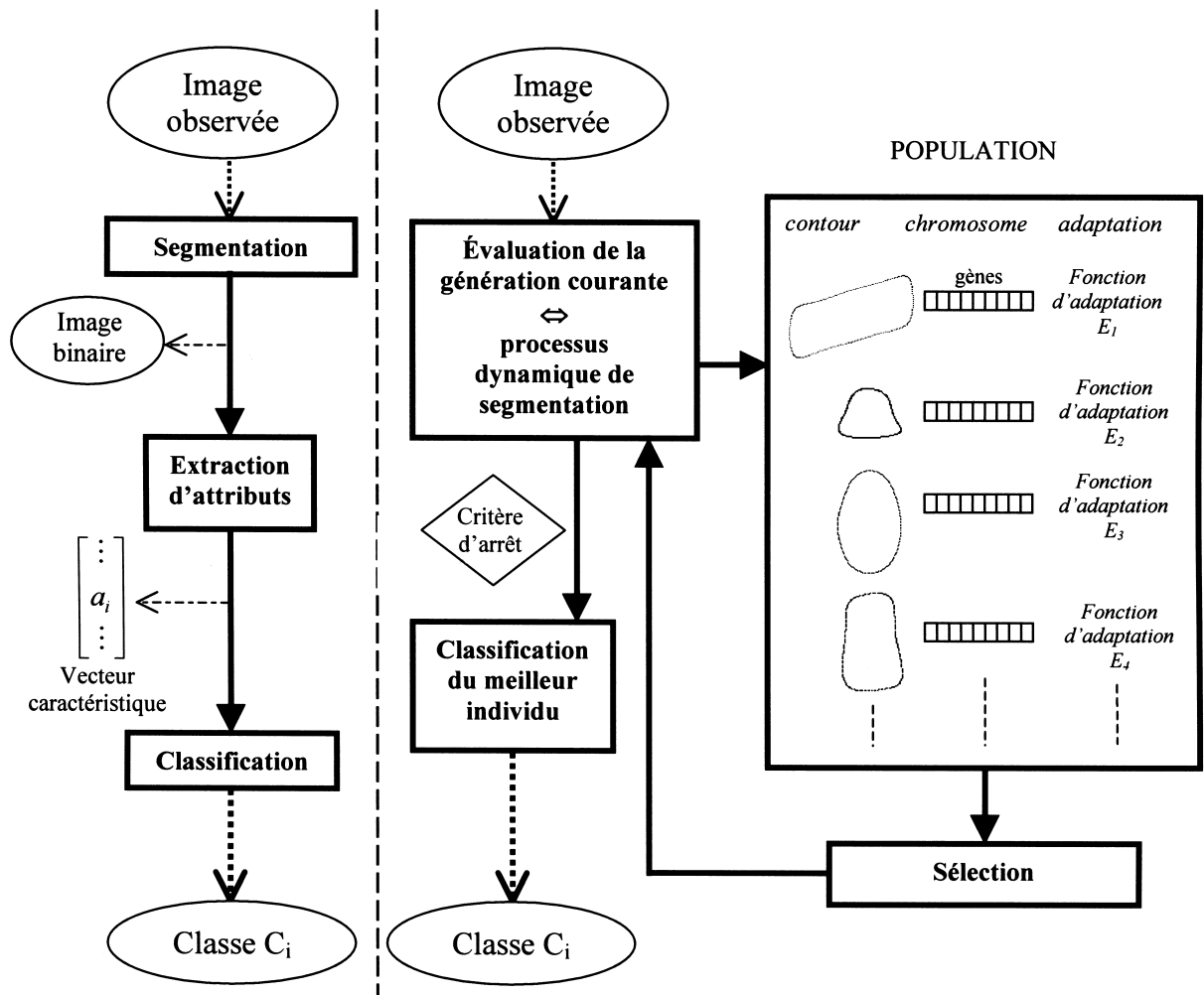


Figure 5. – Segmentation orientée classification opposée au processus de segmentation classique séquentiel.

Tableau 1. –

Type de traitement	Modèles déformables	Approche proposée	Contours actifs
Données	image segmentée en deux classes : ombre/ réverbération de fond	image brute	image brute
Espace des solutions possibles	un prototype	N_p individus	espace infini
Critère d'optimisation	fonction objective bayésienne	fonction d'adaptation	fonctionnelle
Objectif visé	classification	segmentation orientée classification	segmentation

4. algorithmes génétiques

4.1. généralités

Les algorithmes génétiques sont des méthodes d'exploration stochastique fondées sur les mécanismes de la sélection naturelle et de la génétique. Elles en tirent leurs principales caractéristiques : *principe de survie* des individus les mieux adaptés (par évaluation de la « performance » d'un descendant donné) et *échanges d'informations pseudo-aléatoires* entre individus (par opérations génétiques). Ainsi, à chaque génération, de nouveaux individus sont créés par le processus de sélection des parents compte tenu de leur adéquation au problème et de leur reproduction selon les principes généraux de la génétique que sont les croisements et les mutations.

4.2. terminologie

Les algorithmes génétiques procèdent de manière itérative à travers l'évolution d'une population de génération en génération. Chaque *individu* symbolise une solution de l'espace de recherche. Il est caractérisé par son *chromosome* constitué d'une liste finie de *gènes* (paramètres) pouvant prendre des valeurs dans un alphabet donné. À chaque génération (itération), une *fonction d'adaptation* juge de l'adéquation de chaque individu aux données. Cette fonction constitue le critère de *sélection* pour la reproduction. Les individus les mieux adaptés sont soumis à deux types d'opérations génétiques : les *croisements* et les *mutations*. Dans l'opération élémentaire de croisement, les chromosomes sont regroupés par paire et les chaînes ainsi appariées sont recombinaées. La mutation, accidentelle mais nécessaire afin de diversifier le matériel génétique, résulte classiquement d'une modification aléatoire de faible probabilité de la valeur d'un caractère de la chaîne. Dans notre cas, il nous a paru utile d'introduire en outre une mutation opérant une transformation globale du chromosome (*cf.* § 5.1.5). La figure 6 présente le principe général des algorithmes génétiques.

En somme, la mise en œuvre d'un algorithme génétique nécessite la détermination de *sept points fondamentaux* que nous détaillerons au §5 :

- la représentation chromosomique
- la création de la population initiale
- la définition de la fonction d'adaptation
- le critère de sélection
- les opérateurs génétiques réglant la reproduction
- la réinsertion
- le critère d'arrêt

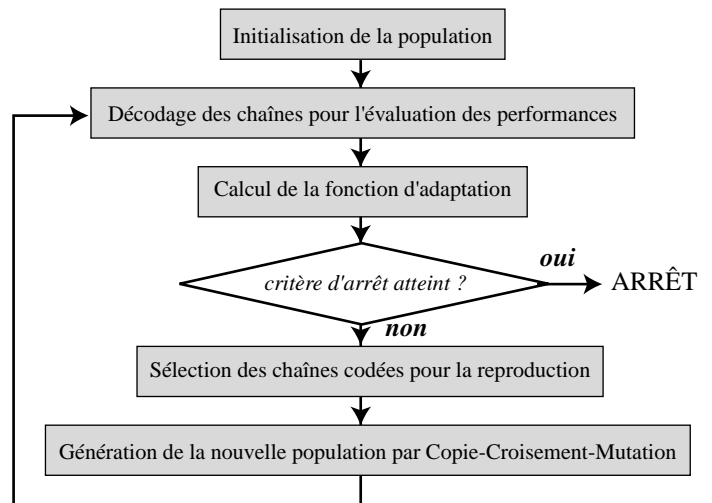


Figure 6. – Principe général des algorithmes génétiques.

4.3. motivations

4.3.1. propriétés théoriques

Une stratégie d'optimisation appropriée à notre problème est le développement d'un algorithme génétique qui possède les propriétés suivantes [DAVI-91] [PAL-96] [GUER-98] :

- *Parallélisme* : cette technique d'exploration stochastique (indépendante de l'initialisation) a l'avantage d'optimiser massivement plusieurs chaînes de paramètres (individus) à la différence de la technique du recuit simulé qui optimise une seule chaîne en évaluant ses configurations voisines. Par ailleurs, à une génération donnée, le potentiel de la population ne se résume pas à la solution représentée par chaque individu : leur combinaison par opérations génétiques est source de multiples autres solutions sous-jacentes.
- par un paramétrage des opérations de sélection et des opérations génétiques, un *compromis* est réalisé entre une exploitation des meilleures solutions et une exploration aléatoire de l'espace de recherche [MICH-94].
- *Rapidité* :
 - les descripteurs sont optimisés conjointement (contrairement au recuit simulé qui optimiserait un seul descripteur)
 - le nombre d'itérations est fonction du problème à optimiser (cas où le critère d'arrêt est fonction de l'optimisation) et non d'un paramètre fixé *a priori* tel la température pour le recuit simulé [DAVI-87]
 - l'élaboration d'opérateurs génétiques spécifiques à l'information portée par les chromosomes augmente la rapidité de convergence par une prise en compte de la spécificité des données traitées. (Dans notre application, il s'agit des distorsions caractéristiques des ombres acoustiques, de la limitation des possibilités de croisements entre chaînes chromosomiques pour éviter la prolifération d'individus non viables...)

4.3.2. travaux antérieurs

Les algorithmes génétiques ont déjà été utilisés pour diviser l'image en plusieurs régions [SEET-91] [HASE-99]. À cette segmentation par région, nous préférons une approche par contour. Dans notre cas en effet, la segmentation, limitée à deux régions, n'est pas une fin en soi mais doit permettre *in fine* de classifier l'objet ciblé par une analyse discriminante de la frontière entre ces régions. Cette frontière correspond au contour de l'ombre acoustique portée et peut être décrite de manière unique par des descripteurs de Fourier normalisés estimés par l'algorithme génétique. L'information région constitue l'attache aux données prise en compte par l'intermédiaire du critère d'adaptation fondé sur la statistique des pixels de l'image. À la différence de la classification des ombres portées en imagerie sonar également proposée par Mignotte *et al.*, l'optimisation vise ici à estimer le contour qui se rapporte à un type de mine parmi les cinq types possibles et non estimer les seules déformations qui affectent un type de mine donné [MIGN-00].

4.4. environnement de mise en œuvre

Le développement algorithmique a été facilité par l'utilisation de la boîte à outils nommée *GAOT* pour *Genetic Algorithms for Optimization Toolbox* codée en langage Matlab [HOUC-95]. Grâce à sa modularité, les modules de bases ont été enrichis de nouvelles routines spécifiques à notre problème.

5. segmentation orientée classification

5.1. segmentation par optimisation génétique

Conformément au principe des algorithmes génétiques, nous partons d'un ensemble d'individus ou *contours prototypes* constituant la population initiale. Chaque individu (ou contour) est caractérisé par $2K$ descripteurs de Fourier normalisés qui constituent son chromosome.

Grâce à un critère énergétique décrit au paragraphe 3.2.1, l'estimation du contour peut s'effectuer sur les données de l'image sonar brute (sans traitement préliminaire) présentant l'ombre portée d'un objet immergé. L'algorithme génétique vise à optimiser cette énergie (fonction d'adaptation), c'est-à-dire estimer les descripteurs de Fourier associés au contour qui correspond à une segmentation optimale.

5.1.1. représentation chromosomique

5.1.1.1. généalogie

Outre la caractérisation intrinsèque de l'individu au moyen des descripteurs de Fourier qui lui sont associés, une information extrinsèque relative à l'arbre généalogique de l'individu peut être résumée dans C gènes particuliers relatifs aux C classes considérées. En effet, les descendants héritent du patrimoine génétique des parents lesquels sont eux-mêmes issus de la génération précédente. Alors que nous connaissons *a priori* la nature de chacun des individus de la population initiale, cette information n'est plus évidente aux générations suivantes. Néanmoins, elle peut être récupérée à chaque génération par incrémentation d'un compteur $N_{c,c=1\dots C}$ pour chacune des classes (*cf.* § 5.1.5).

5.1.1.2. composition du chromosome

La figure 7 montre un chromosome constitué de la manière suivante (coefficients pour $k \in \{1, \dots, K\}$) :

- $2 \times K$ gènes représentant $2 \times K$ réels (parties réelle et imaginaire) pour K descripteurs de Fourier complexes \tilde{X}_k résultant de la transformée de Fourier appliquée aux abscisses des pixels du contour
- $2 \times K$ gènes représentant $2 \times K$ réels (parties réelle et imaginaire) pour K descripteurs de Fourier complexes \tilde{Y}_k résultant de la transformée de Fourier appliquée aux ordonnées des pixels du contour
- 2 gènes permettant de localiser le contour, c'est-à-dire les réels X_0 et Y_0
- C gènes (pour C classes d'objets) relatifs à la généalogie de l'individu

Remarque (*cf.* § 3.2.1) : l'image observée est divisée en deux zones de frontière caractérisée par le vecteur $\theta_{(K)} = (\tilde{X}_k, \tilde{Y}_k)_{k \in \{1, \dots, K\}}^T$

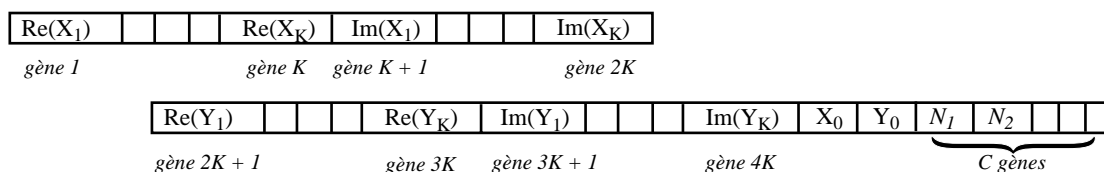


Figure 7. – Définition du chromosome.

5.1.2. population initiale

La population initiale de taille N_p doit couvrir l'ensemble des classes. Le choix des classes tient compte de l'évolution de la menace en guerre des mines. Alors que la plupart des mines à orin sont sphériques ou oblongues (cylindre vertical), les mines de fond ont des formes variées et complexes. À la classique mine cylindrique s'ajoutent de nouveaux types de mines dites furtives du fait de leur revêtement anéchoïque, de petites tailles et de formes particulières.

Cinq classes sont donc considérées : cylindres (classe C_1), sphères (classe C_2), mines Manta (classe C_3), mines Sigeel (classe C_4) et mines Rockan (classe C_5). Les mines Manta et Sigeel sont toutes deux tronconiques mais de dimensions différentes, tandis que la mine Rockan présente un profil particulièrement plat [JANE].

La population initiale doit être représentative des différentes configurations susceptibles de se présenter. En particulier, du fait de la géométrie complexe des mines Rockan et cylindriques, la forme associée à l'ombre portée dépend du point de vue considéré. En outre, si les mines tronconiques et Rockan ont des dimensions fixes, il n'en est pas de même pour sphères et cylindres.

Ces remarques montrent que la population initiale doit comporter un nombre plus important d'exemplaires pour les objets complexes et/ou de dimensions multiples. Notons en outre que cette population peut à tout moment être enrichie de nouveaux prototypes pour anticiper les menaces futures.

La figure 8 montre quelques exemplaires « prototypes » reconstruits à partir de leurs descripteurs de Fourier issus de cette population initiale qui correspondent aux mines illustrées figure 9.

5.1.3. fonction d'adaptation

Le terme E défini au paragraphe 3.2.1 correspond à une énergie (désignée par énergie externe dans la théorie des contours actifs) et constitue pour notre application de l'algorithme génétique la **fonction d'évaluation** ou **d'adaptation**.

Pour calculer cette énergie, le traitement appliqué à chaque individu de la population à chaque génération est le suivant :

1) *positionnement du marqueur* (barycentre approché de l'ombre portée) sur l'image observée

La reconstruction du contour s'effectue autour du pixel dont les coordonnées sont les coordonnées moyennes des points de contour de l'ombre détectée c'est-à-dire (X_0, Y_0) . La recherche de ce point ou **marqueur** n'est pas triviale dans le sens où nous ne connaissons justement pas le contour exact de

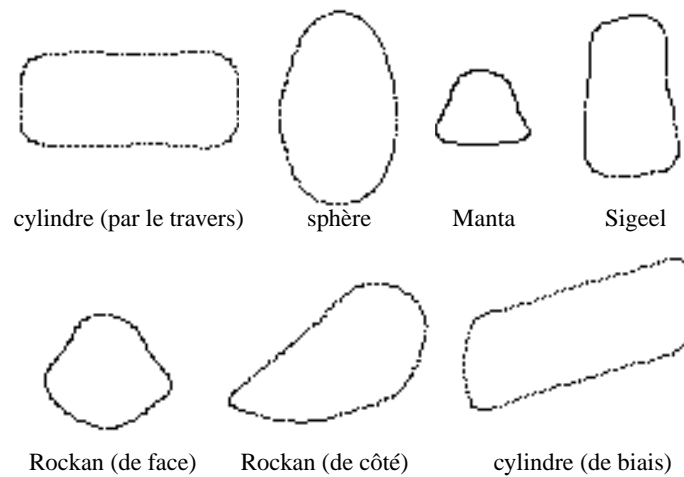


Figure 8. – Exemples d'individus de la population initiale : contours d'ombres acoustiques segmentées.

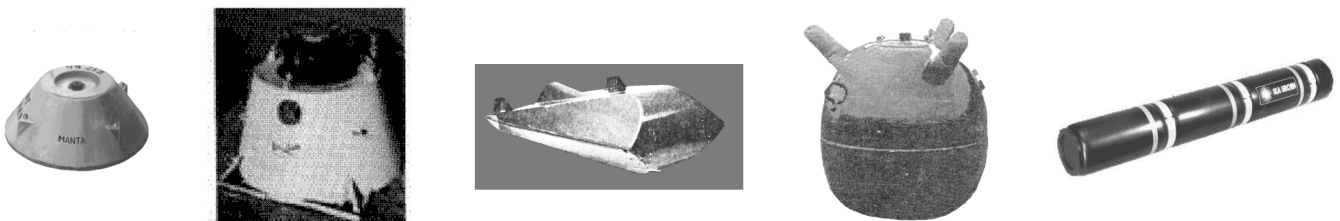


Figure 9. – Mines Manta, Sigeel et Rockan, sphère et cylindre.

cette ombre. Pratiquement, la recherche du marqueur est effectuée au voisinage (limité à plus ou moins dix pixels dans chacune des directions) d'un point initial et guidée par la valeur de l'énergie obtenue pour chacune des positions du contour dans l'image. Les coordonnées du marqueur correspondent à une localisation optimale du contour au sens du critère statistique choisi (cf. paragraphe 3.2.1) [GUER-99]. À la première génération, les coordonnées du point initial sont fixées par l'utilisateur. La position optimale du marqueur (au sens de l'évaluation choisie) est ensuite transmise et peut être modifiée de génération en génération comme les autres gènes mais suivant un processus d'évolution notablement différent. L'individu a la capacité de s'auto-adapter : avant un calcul définitif de la fonction d'adaptation, il augmente ses chances de survie en jouant sur deux gènes portant les coordonnées du marqueur. En quelque sorte, ces gènes particuliers se réfèrent à l'acquis par opposition à l'inné contenu dans les autres gènes.

- 2) lecture des K coefficients de Fourier, capital génétique de l'individu (cf. § 5.1.1)
- 3) reconstruction de la silhouette associée à l'individu par Transformée de Fourier Inverse autour de son barycentre telle que

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \sum_{\substack{k=-K \\ k \neq 0}}^K \begin{bmatrix} X_k \\ Y_k \end{bmatrix} \exp\left(2j\pi \frac{ki}{N}\right)$$

où $X_{-k} = X_k^*$, $Y_{-k} = Y_k^*$, et N désigne le nombre de points de contour.

- 4) calcul de la fonction d'adaptation E (cf. § 3.2.1)
 - si toutes les coordonnées des pixels du contour reconstruit sont inférieures aux dimensions de l'image :

$$E = -N_0 \ln \alpha_0^2 - N_R \ln \alpha_R^2,$$

– sinon, on pénalise fortement un contour dont une partie sort du cadre de l'image en considérant l'image comme une seule région *i.e.* :

$$E = -N_{image} \ln \alpha_{image}^2$$

Remarque : afin de diminuer le temps de calcul, nous utilisons l'histogramme de chaque région plutôt que l'ensemble des pixels de l'image. En effet, la contribution d'un pixel à la fonction d'adaptation dépend uniquement de son niveau et de la région à laquelle il appartient (et non de ses coordonnées). On passe ainsi d'une double somme sur les lignes et les colonnes à une simple somme sur les 256 niveaux de gris.

5.1.4. critère de sélection

L'étape de sélection consiste à choisir les parents de la génération suivante et se situe donc avant les opérations génétiques. Classiquement, une probabilité de sélection P_i est attribuée à chaque individu i , reflétant son adaptation au milieu. Supposons que l'on souhaite sélectionner N_a individus, N_a nombres aléa-

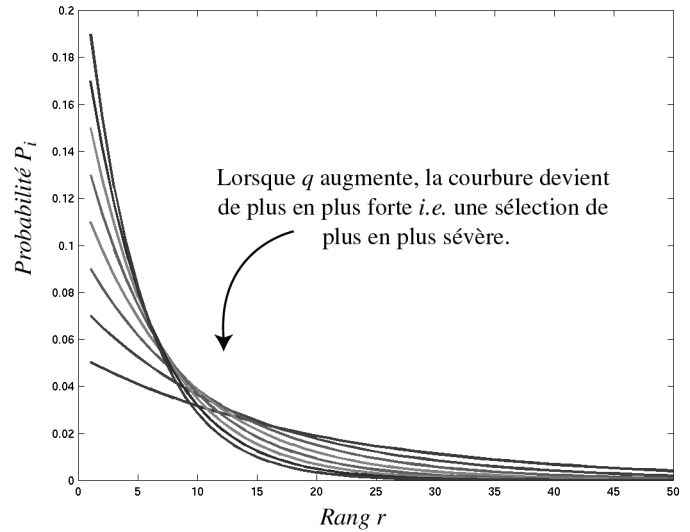


Figure 10. – Allures $P_i = f(r)$ pour différentes valeurs de q .

toires $u_{n,n=1\dots N_a} \in [0, 1]$ sont alors générés et comparés à la

somme cumulée des probabilités $C_i = \sum_{j=1}^i P_j$, $i \in [1, N_a]$. Pour

chaque entier n de 1 à N_a , on sélectionne l'individu i tel que $C_{i-1} < u_n \leq C_i$.

Diverses réalisations de P_i sont possibles. Une des plus répandues, dite « **roue de loterie** », attribue à l'individu une probabilité proportionnelle à son adaptation au milieu. Son principal défaut est de supprimer à coup sûr les individus les plus faibles au risque de converger vers une solution sous-optimale. Par contre, les méthodes de tri et de sélection en fonction du rang, dites « **ranking methods** », sont plus souples et plus robustes [GOLD-91]. Ces méthodes attribuent une probabilité fonction du rang $r(i)$ de l'individu i lorsque toutes les adaptations ont été classées par ordre décroissant. Houck et Joines ont ainsi développé un algorithme appelé « **Normalized Geometric Ranking** » dans lequel P_i est défini par [HOUC-95] :

$$P_i = \frac{q}{1 - (1 - q)^{N_p}} (1 - q)^{r(i)-1}$$

Avec $r(i)$ le rang de l'individu où 1 désigne le meilleur, N_p la taille de la population, q la probabilité de sélection du meilleur individu.

Pour une population de 100 individus par exemple, nous visualisons les probabilités P_i obtenues en fonction du rang pour q allant de 0,05 à 0,2 par pas de 0,02, *i.e.* 8 courbes. La figure 10 met en évidence une courbure de plus en plus forte *i.e.* une sélection de plus en plus sévère lorsque q augmente. Il s'agit donc d'effectuer le compromis entre une sélection trop sévère (q trop important) qui mettrait en danger la diversité génétique et une sélection trop large (q trop faible) qui augmenterait les temps de calcul.

5.1.5. opérateurs génétiques

5.1.5.1. définitions

Les parents sélectionnés vont se reproduire par croisement et subir des mutations. Nous avons introduit des routines originales qui exploitent la nature des descripteurs de Fourier en utilisant leurs propriétés et l'interprétation qui peut en être faite.

On note c_i (resp. p_i) les chromosomes des descendants obtenus (resp. des parents).

– *croisement (entre deux parents) :*

Par principe, l'individu issu du croisement génétique doit « ressembler » à ses parents. En terme de contour, on doit donc visuellement rapprocher le nouveau contour créé des contours parentaux. Ceci impose des contraintes qui limitent les opérations admissibles entre les chaînes chromosomiques. Les trois routines suivantes sont basées sur la moyenne des descripteurs des deux parents.

Routine crois1

La moyenne concerne tous les descripteurs, le barycentre de la silhouette inclus, telle que

$$c_1(1, \dots, 4K + 2) = \frac{p_1(1, \dots, 4K + 2) + p_2(1, \dots, 4K + 2)}{2}$$

parallèlement, l'information généalogique est transmise par

$$c_1(4K + 3, \dots, 4K + 2 + C) =$$

$$p_1(4K + 3, \dots, 4K + 2 + C) + p_2(4K + 3, \dots, 4K + 2 + C)$$

Afin de maintenir le même nombre d'individus à chaque génération, un deuxième descendant, copie conforme du parent le mieux adapté, est formé.

Routine crois2

Une moyenne pondérée est effectuée de manière à ce que les descendants ressemblent plutôt à l'un ou l'autre des parents *i.e.*

$$\left\{ \begin{aligned} c_1(1, \dots, 4K + 2) &= \frac{p_1(1, \dots, 4K + 2) + 0,5 \times p_2(1, \dots, 4K + 2)}{1,5} \\ c_2(1, \dots, 4K + 2) &= \frac{p_2(1, \dots, 4K + 2) + 0,5 \times p_1(1, \dots, 4K + 2)}{1,5} \end{aligned} \right.$$

et parallèlement, l'information généalogique conformément au « poids héréditaire » est transmise par

$$\left\{ \begin{aligned} c_1(4K + 3, \dots, 4K + 2 + C) &= p_1(4K + 3, \dots, 4K + 2 + C) \\ &+ 0,5 \times p_2(4K + 3, \dots, 4K + 2 + C) \\ c_2(4K + 3, \dots, 4K + 2 + C) &= p_2(4K + 3, \dots, 4K + 2 + C) \\ &+ 0,5 \times p_1(4K + 3, \dots, 4K + 2 + C) \end{aligned} \right.$$

Routine crois3

Les seuls coefficients de degré 1 sont pris en compte. Ils caractérisent les axes principaux de l'ellipse approchant au mieux le contour. c_1 est obtenu suivant la même opération de moyenne que pour *crois1* mais limitée à ce seul coefficient. Les autres coefficients conservent leur valeur. c_2 est une copie conforme du parent le mieux adapté.

Routine crois4

Pour garantir la diversité, une quatrième routine *crois4* permet de créer deux individus complémentaires dont les coefficients de Fourier sont choisis par tirage aléatoire entre les chaînes chromosomiques des deux parents.

Prenons le cas $K = 5$. Un exemple est illustré figure 11 pour lequel le tirage aléatoire attribue les coefficients $k = 2, 3, 5$ de p_1 au descendant c_1 et $k = 1, 4$ au descendant c_2 et inversement concernant p_2 .

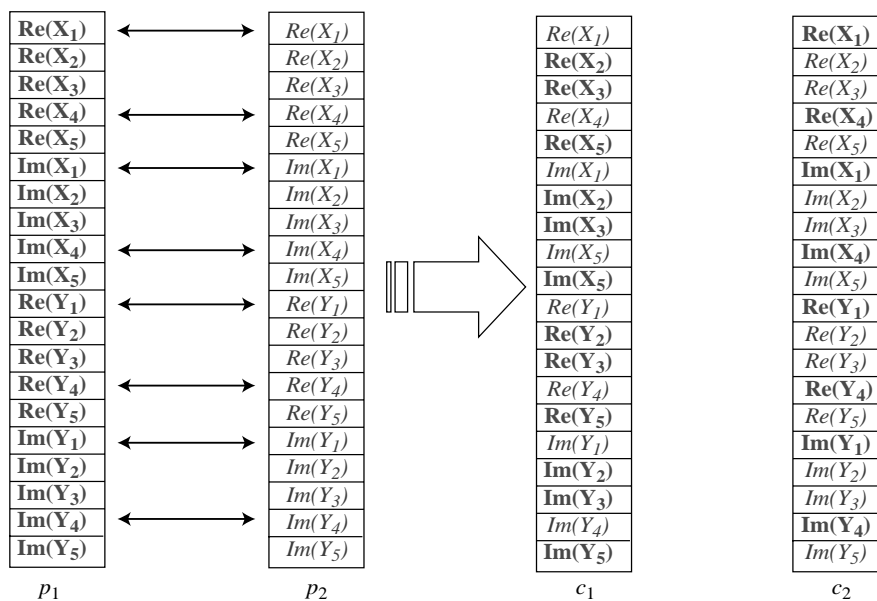


Figure 11. – Échanges aléatoires entre deux chaînes.

Les points moyens retournés sont identiques pour c_1 et c_2 i.e.

$$c_1(4K + 1, 4K + 2) = \frac{p_1(4K + 1, 4K + 2) + p_2(4K + 1, 4K + 2)}{2}$$

et parallèlement, l'information généalogique est transmise par

$$c_1(4K + 3, \dots, 4K + 2 + C) = \pi_1^i \times p_1(4K + 2, \dots, 4K + 2 + C) + \pi_2^i \times p_2(4K + 3, \dots, 4K + 2 + C)$$

où π_1^i (resp. π_2^i) représente la proportion de gènes reçus de p_1 (resp. p_2) par le descendant c_i .

– *mutation (s'applique à un seul parent)* :

Remarque : ces opérations ne concernent que les 4K premiers gènes c'est-à-dire les descripteurs de Fourier.

Routine mut1

Un bruit blanc gaussien est ajouté à l'un des descripteurs de Fourier. Ce bruit est de moyenne nulle et d'écart-type égal à la valeur maximale admissible pour ce descripteur divisée par 10.

Routine mut2

Cette routine utilise la propriété de linéarité de la décomposition en série de Fourier calculée distinctement sur les abscisses et les ordonnées des points de contour.

Soit la transformation géométrique suivante :

$$\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = M \times \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$$

en passant dans le domaine de Fourier, on a

$$\begin{bmatrix} U_k \\ V_k \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} X_k \\ Y_k \end{bmatrix} = M \times \begin{bmatrix} X_k \\ Y_k \end{bmatrix}$$

Cette décomposition en série de Fourier du contour est donc *linéaire*.

Réciproquement, une transformation géométrique des descripteurs se traduit par une transformation affine des points de contour. On dispose donc d'un modèle paramétrique linéaire de degré 4 correspondant aux quatre paramètres de la matrice M .

Par mutation, on peut donc modifier le contour de manière globale en lui affectant une transformation affine aléatoire (la translation exceptée). Grâce à cette opération, un seul individu possède un patrimoine génétique qui peut servir à constituer d'autres solutions.

Six déformations affines de matrice $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ sont possibles :

- l'homothétie qui agit sur a et d telle que

$$M = \begin{bmatrix} 1 + \alpha & 0 \\ 0 & 1 + \alpha \end{bmatrix}, \text{ avec } \alpha \in [-0, 4; +0, 4] \setminus \{0\} \text{ par tirage aléatoire}$$

- l'élongation sur les lignes qui agit sur a telle que

$$M = \begin{bmatrix} 1 + \alpha & 0 \\ 0 & 1 \end{bmatrix}, \text{ avec } \alpha \in [-0, 4; +0, 4] \setminus \{0\} \text{ par tirage aléatoire}$$

- l'élongation sur les colonnes qui agit sur d telle que

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 1 + \alpha \end{bmatrix}, \text{ avec } \alpha \in [-0, 4; +0, 4] \setminus \{0\} \text{ par tirage aléatoire}$$

- la rotation qui agit sur toutes les composantes telle que

$$M = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix},$$

avec $\alpha \in \frac{1}{iter}[-9^\circ; +9^\circ] \setminus \{0\}$ par tirage aléatoire

- le biais sur les lignes qui agit sur b tel que

$$M = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}, \text{ avec } \alpha \in \frac{1}{iter}[-0, 1; +0, 1] \setminus \{0\} \text{ par tirage aléatoire}$$

- le biais sur les colonnes qui agit sur d tel que

$$M = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix}, \text{ avec } \alpha \in \frac{1}{iter}[-0, 1; +0, 1] \setminus \{0\} \text{ par tirage aléatoire}$$

Pour les trois dernières possibilités, *iter* désigne l'itération de sorte que l'amplitude admissible des rotations et des biais diminue au fil des générations. L'amplitude du paramètre α a été fixée de manière empirique de manière à éviter des transformations trop importantes qui créeraient des individus non viables.

En modifiant ainsi la forme de manière globale (*mut2*), on simule une déformation affine qui peut être due à l'angle de prise de vue, à la différence de taille entre la mine ciblée et la mine prototype, voire aux artefacts de traitements préliminaires de formation et de normalisation en rasant de l'image sonar. La nature de la silhouette n'est donc pas altérée. À l'opposé, une modification partielle voire ponctuelle de la chaîne des descripteurs (*mut1*) modifie intrinsèquement la silhouette en modifiant sa nature.

Les deux séries d'opérations génétiques sont complémentaires : les croisements exploitent les informations existantes en les recombinaison, tandis que les mutations explorent plus largement l'espace des solutions. Ces opérations et les paramètres associés ont été choisis pour leur signification physique, en dépit du fait que les algorithmes génétiques bénéficient d'une robustesse inhérente vis-à-vis du paramétrage. On épargne ainsi du temps de calcul en évitant de générer des contours de forme complètement inadaptée au problème.

5.1.5.2. taux d'occurrence des opérations génétiques [PAL-96]

Le choix des *taux de croisements et de mutations* ainsi que celui de la *taille de la population initiale* n'est pas trivial :

- accroître le taux de croisements revient à augmenter le nombre de recombinaisons mais dissocie en contrepartie les bons chromosomes,
- accroître le taux de mutations risque de transformer la recherche en une exploration aléatoire de l'espace des solutions, mais permet en outre de réintroduire une partie du maté-

riel génétique perdu (par *mut1*) ou de diversifier la population (grâce à *mut2*),

- augmenter la taille de la population accroît sa diversité mais augmente les temps de calcul.

Les taux que nous avons choisis dépendent des opérateurs (cf. § 6.1) et de la taille de la population. Ainsi, pour une petite population, on est amené à autoriser davantage d'opérations génétiques de manière à pallier un manque de diversité. Ceci justifie un fort taux d'occurrence pour les mutations par transformation affine qui ne vise pas à réintroduire de l'information perdue (contrairement à la mutation accidentelle d'un gène) mais vise à augmenter la diversité d'une population de petite taille. Par ailleurs, la taille de notre population initiale est fortement conditionnée par l'objectif visé qui consiste à segmenter des ombres relatives à des objets connus que l'on souhaite classifier. On souhaite ainsi rapprocher l'ombre détectée d'un ensemble de cas répertoriés et préalablement segmentés (cf. § 5.1.2).

5.1.6. réinsertion

L'étape de sélection réarrange la population de sorte que les individus les plus faibles sont éliminés au profit d'individus plus forts ayant une probabilité de sélection plus grande. On maintient à chaque génération le même nombre d'individus de sorte que $N_a = N_p$ (cf. § 5.1.4). Ensuite, les opérations de croisements prennent chaque fois deux parents pour renvoyer deux descendants. Les mutations, d'un autre côté, ne concernent qu'un individu à la fois, remis en place après transformation. Quels que soient le choix des routines et leur occurrence (fonction des taux de croisement et de mutation choisis), à l'issue des opérations génétiques, N_p descendants doivent être transmis. Deux cas peuvent se présenter :

- si le nombre de descendants est supérieur à N_p , on ne réinsère que les individus les mieux adaptés,
- sinon, les descendants remplacent les individus de la génération précédente les moins bien adaptés.

5.1.7. critère d'arrêt

La fonction de terminaison est appelée à la fin de chaque itération c'est-à-dire lorsque la nouvelle génération est obtenue. Nous choisissons d'arrêter l'optimisation lorsque l'un des cas suivants est rencontré :

- la population des descendants est homogène c'est-à-dire l'écart entre les évaluations des individus le mieux et le moins bien adaptés est inférieur à une petite valeur,
- il n'y a pas eu d'amélioration pour trois itérations consécutives,
- le nombre maximum de générations admissible est atteint.

5.2. module de classification

Estimation faite des descripteurs de Fourier associés au contour le mieux adapté à l'image observée, il s'agit de déterminer la classe correspondante à cet « individu ». Pour ce faire, nous

décidons de rapprocher les descripteurs estimés avec les descripteurs caractérisant les N_p^c exemplaires associés à chacune des $C = 5$ classes d'une population de référence, représentative des cas susceptibles d'apparaître, et plus étendue (c'est-à-dire plus diverse) que la population initiale (i.e. $N_p^c > N_p$).

Compte tenu de l'unicité des descripteurs de Fourier normalisés, nous introduisons un critère de classification fondé sur le principe des *k*-plus-proches-voisins (*k*-ppv) qui permet d'attribuer à un individu la classe majoritaire parmi les *k* voisins les plus proches au sens d'une distance définie ci-dessous.

La distance d_j peut être calculée à partir des gènes de l'individu à classifier et de ceux du $j^{\text{ième}}$ individu de la population de référence telle que :

$$d_j = \left(\sum_{i=1}^{4K} (g_i^j - \tilde{g}_i)^2 \right)^{\frac{1}{2}}$$

où

$(\tilde{g}_i)_{i=1...4K}$ sont les gènes estimés

$(g_i^j)_{i=1...4K}$, les gènes du chromosome du $j^{\text{ième}}$ individu de la population de référence

Chaque indice j réfère à un prototype spécifique dont on connaît la classe. L'indice $j = \text{Arg} \min_{j \in \{1...N_p^c\}} d_j$ est en particulier relatif au plus proche prototype.

Les *k*-ppv consistent à trouver les *k* distances les plus faibles puis à déterminer la classe associée aux individus correspondants.

L'algorithme de classification est décrit sur la figure 12. Deux cas se présentent :

- plusieurs classes sont concurrentes,
- une classe seule est majoritairement représentée parmi les *k* voisins.

- Trouver les *k* indices j pour lesquels les distances d_j sont les plus faibles
- Répartir les individus correspondants par classe $C_{c,c=1...C}$, on a $N_{c,c=1...C}$ individus par classe tels que $\sum_{c=1...C} N_c = k$
- Calculer la distance moyenne associée à la classe $C_{c,c=1...C}$:

$$\bar{d}_c = \frac{1}{N_c} \sum_{j \in \{\text{indices des individus de } C_c\}} d_j$$

Deux cas possibles :

- Si plusieurs classes sont concurrentes, c'est-à-dire si elles bénéficient chacune du nombre maximal de représentants $\max_c N_{c,c=1...C}$ parmi les voisins la classe attribuée est celle dont la distance moyenne est la plus faible, i.e. $c = \text{Arg} \min_{c \in \{1...C\}} \bar{d}_c$
- Sinon, $c = \text{Arg} \max_{c \in \{1...C\}} N_c$, classe majoritaire parmi les *k* voisins les plus proches

Figure 12. – Algorithme de classification.

Prise en compte des gènes relatifs à la généalogie :

Cet algorithme des k -ppv est enrichi de l'hérédité portée par l'individu. Sans la considérer comme un moyen à part entière de le classifier, elle constitue une information supplémentaire qui permet à l'opérateur d'accorder plus ou moins de confiance en pondérant le résultat.

6. expérimentations

6.1. paramétrage

Une reconstruction lissée du contour peut être obtenue en fixant **K à 5** tout en préservant la forme globale de la silhouette qui permet donc de la discriminer.

La population initiale doit couvrir l'ensemble des $C = 5$ classes (cf. § 5.1.2). La proportion des exemplaires par classe dépend des propriétés de symétrie et de la taille (fixe ou multiple) des objets :

- les cylindres et les sphères ont des dimensions variées et doivent donc être représentés par un nombre d'exemplaires supérieur à celui des mines à taille fixe
- du fait de la géométrie des mines Rockan et cylindriques, la forme associée à l'ombre portée dépend du point de vue considéré : ces deux classes doivent comporter l'ensemble des configurations possibles.

Par simulation, différents points de vue sont acquis le long d'une trajectoire semi-circulaire du sonar autour de l'objet. La population initiale est de $N_p = 41$ individus répartis entre les cinq classes en fonction de leur dispersion (diversité des tailles et des formes possibles de l'ombre) comme suit :

- pour la classe C_1 , $N_{cyl} = 19$ en prenant 3 tailles de cylindres et 19 points de vue différents,
- pour la classe C_2 , $N_{sph} = 10$ en prenant 3 tailles de sphères,
- pour les classes C_3 et C_4 , $N_{man} = N_{sig} = 3$, car les mines tronconiques sont de taille constante et ont la propriété de symétrie radiale,
- pour la classe C_5 , $N_{rock} = 6$ en prenant 6 points de vue représentatifs.

La taille de la population initiale nous amène à fixer q à **0.25** pour la sélection des individus. Selon la formule et les considérations du paragraphe 5.1.4, la probabilité de sélection P_i est alors inférieure à 0,001 (ce qui équivaut à une sélection improbable) pour la moitié des individus.

La population de référence pour la classification est environ trois fois plus importante pour pouvoir fixer **le nombre de voisins à $k = 3$** dans le module de classification.

Le **taux de croisements** est fixé à **10 %** pour chaque opérateur de moyenne et à **20 %** pour l'échange aléatoire entre chaînes chromosomiques. Pour la mutation par modification ponctuelle d'un gène, on effectue une seule opération par génération. Par contre, un **taux de 50 %** a été choisi pour la **mutation par**

transformation affine afin de multiplier les différentes configurations en termes de format et de point de vue des ombres potentielles associées aux individus sélectionnés. Ce taux diminue au fil des générations de sorte que pour une génération ou itération donnée $iter$, le taux est de $50/iter$ % (cf. § 5.1.5).

Enfin, le **nombre maximal d'itérations** est **20**.

6.2. illustrations

Les figures 13 à 17 illustrent les résultats obtenus sur deux images réelles acquises par un traitement « antenne synthétique » et sur trois images simulées [BILL-98].

Pour chaque image testée, on affiche la classe C_c attribuée selon l'algorithme de la figure 12 et le taux N_c/k associé. L'information apportée par les C gènes relatifs à la généalogie de l'individu solution de l'algorithme d'optimisation est une aide à l'opérateur. Il accordera d'autant plus de confiance au résultat issu du module de classification que la proportion d'individus de même classe que la classe attribuée sera élevée dans l'hérédité.

Sur les figures 18 et 19, à l'histogramme de l'image observée, on superpose les histogrammes estimés des zones « ombre » et « réverbération de fond » résultant de la segmentation de l'image.

Nota Bene : pour améliorer la visibilité et faciliter la comparaison, nous avons imposé le même niveau de gris moyen des pixels pour les images réelles et les images simulées.

6.3. performances

Nous avons évalué notre algorithme sur une base d'images simulées composée des 5 types de mines précédemment décrites. Le nombre d'images par classe dépend des propriétés de symétrie et de la taille (fixe ou multiple) des objets. Ainsi, les objets à symétrie complexe (cylindres et mines Rockan) sont davantage représentés du fait de la variété des configurations possibles.

La base est constituée comme suit :

- 30 séquences de 19 images pour 3 tailles de cylindres *i.e.* 1710 images
- 30 séquences de 10 images pour 3 tailles de sphères *i.e.* 900 images
- 30 séquences de 10 images pour la mine Manta *i.e.* 300 images
- 30 séquences de 10 images pour la mine Sigeel *i.e.* 300 images
- 30 séquences de 19 images pour la mine Rockan *i.e.* 570 images

Le Tableau 2 donne la matrice de confusion obtenue.

La plus forte confusion intervient pour la mine Rockan dont l'un des points de vue est systématiquement associé à la classe des mines Manta de par la forme trapézoïdale de l'ombre portée (cf. figure 20).



Figure 13. – Résultat sur image réelle (ombre portée d'une mine Manta).

Objet classé « mine Manta » par k -ppv
avec un taux de 1

Généalogie :

0 % cylindre

0 % sphère

100 % mine Manta

0 % mine Sigeel

0 % mine Rockan

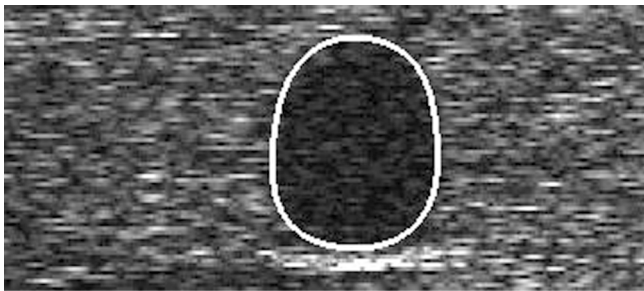


Figure 14. – Résultat sur image réelle (ombre portée d'une mine sphérique).

Objet classé « sphère » par k -ppv
avec un taux de 1

Généalogie :

0% cylindre

100% sphère

0% mine Manta

0% mine Sigeel

0% mine Rockan

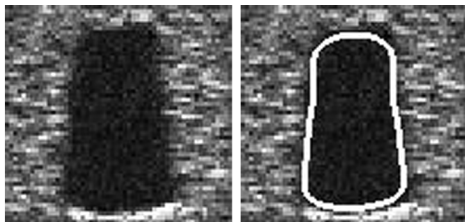


Figure 15. – Résultat sur image simulée (ombre portée d'une mine Sigeel).

Objet classé « mine Sigeel » par k -ppv
avec un taux de 1

Généalogie :

0% cylindre

0% sphère

0% mine Manta

100% mine Sigeel

0% mine Rockan

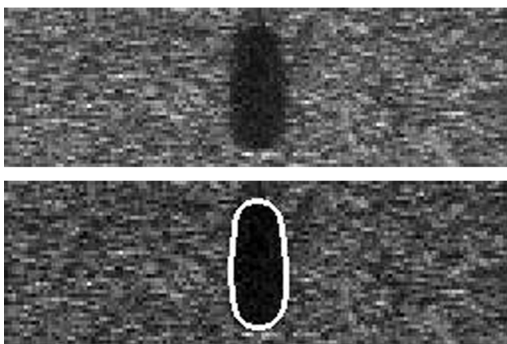


Figure 16. – Résultat sur image simulée (ombre portée d'une mine cylindrique vue par le bout).

Objet classé « cylindre » par k -ppv
avec un taux de 0,67

Généalogie :

64,63% cylindre

35,37% sphère

0 % mine Manta

0 % mine Sigeel

0 % mine Rockan

Classification de mines sous-marines à partir de l'image sonar brute

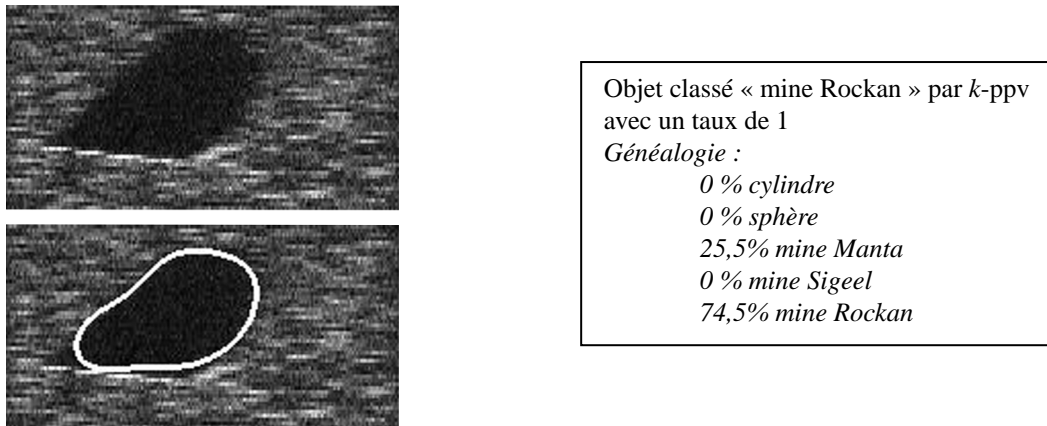


Figure 17. – Résultat sur image simulée (ombre portée d'une mine Rockan).

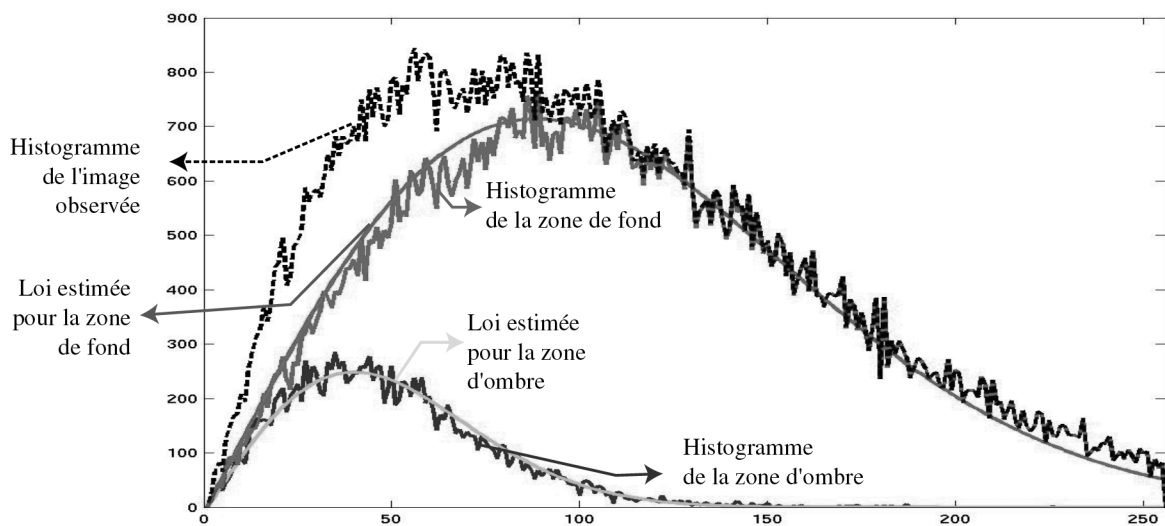
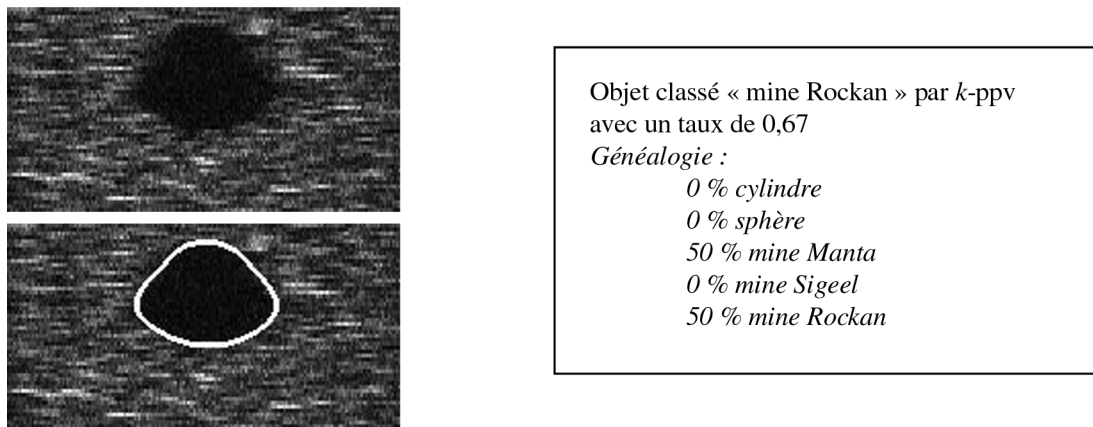
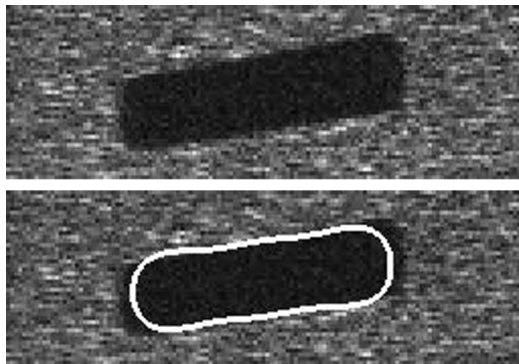


Figure 18. – Résultat de l'optimisation sur image simulée (ombre portée d'une mine Rockan) et le résultat de l'estimation des lois de répartition des niveaux de gris des pixels.



Objet classé « cylindre » par k -ppv
avec un taux de 1
Généalogie :
100 % cylindre
0 % sphère
0 % mine Manta
0 % mine Sigeeel
0 % mine Rockan

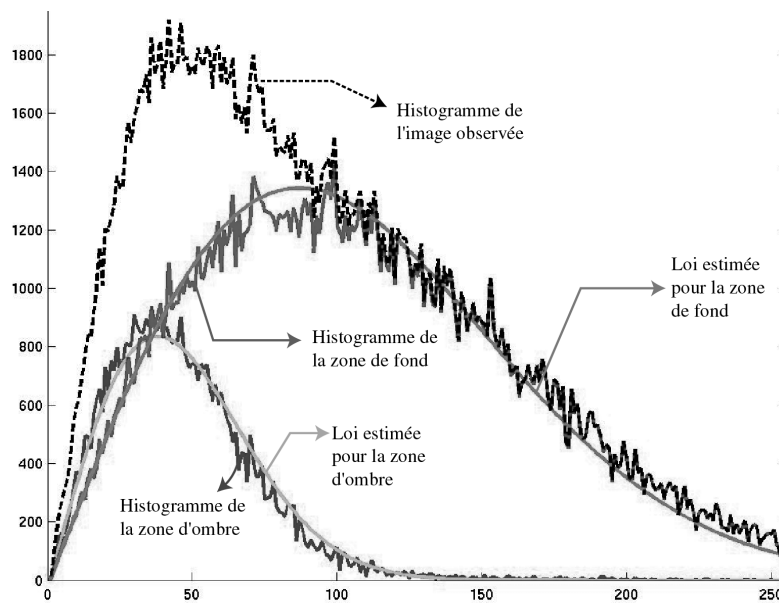


Figure 19. – Résultat de l’optimisation sur image simulée (ombre portée d’une mine cylindrique) et le résultat de l’estimation des lois de répartition des niveaux de gris des pixels.

Tableau 2 – Matrice de confusion (taux moyen de bonne classification de 97,4%).

	C_1	C_2	C_3	C_4	C_5
Cylindres	97,4	2,6		0	< 0,1 %
Sphères	0	100	0	0	0
Mines Manta	0	0	100	0	0
Mines Sigeeel	0	1,7	0	98,3	0
Mines Rockan	0	0	8,9	0	91,1

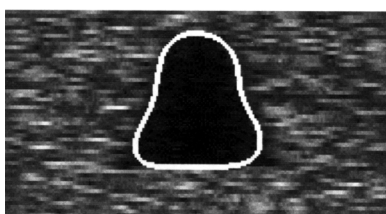


Figure 20. – Point de vue ambigu de la mine Rockan.

7. conclusion

Dans le contexte de la chasse aux mines, les informations que nous possédons sur les différents objets susceptibles de représenter une menace nous permettent d’accélérer les traitements en travaillant initialement avec un nombre réduit de solutions potentielles. La technique proposée consiste à explorer le domaine des solutions possibles par algorithme génétique. L’initialisation permet de se positionner dès le départ dans un contexte favorable, ensemble d’individus dont les caractéristiques sont parfaitement connues. Au fil des générations, ces individus montrent une forte capacité d’adaptation à l’image observée en convergeant vers une solution optimale au sens d’un critère statistique d’évaluation. Des opérateurs génétiques ont été spécifiquement développés pour tirer profit des propriétés des descripteurs de Fourier codant les individus. Les tests effectués sur images réelles et simulées ont montré que, par sa flexi-

bilité, l'algorithme génétique proposé estimait correctement les descripteurs de Fourier décrivant au mieux le contour de l'ombre portée. En outre, on classifie aisément l'objet ciblé en rapprochant la chaîne chromosomique estimée des chaînes des individus d'une population de référence.

En travaillant directement à partir des données sonar brutes, nous nous affranchissons des étapes intermédiaires (segmentation et extraction d'attributs caractéristiques) intervenant dans le traitement séquentiel classiquement mis en œuvre pour la reconnaissance de formes. À travers les générations successives, l'ensemble des contours codés dans les chromosomes des individus évoluent suivant les lois de la génétique. Il en résulte un processus dynamique de classification de l'ombre. Remarquons enfin que la mise en œuvre de cet algorithme ne nécessite pas un apprentissage sur une base de données étendue et s'adapte donc aisément à la menace par l'introduction de nouveaux individus dans la population initiale.

Annexe 1 : description du contour

Étant donné le contour d'une silhouette sous forme d'une liste de coordonnées de pixels, on rééchantillonne le contour de manière à construire un contour de vélocité uniforme de N points équidistants. Une transformée de Fourier est alors effectuée distinctement sur les abscisses x_i et les ordonnées y_i des N points de contour équirépartis :

$$\begin{cases} X_k &= \frac{1}{N} \sum_{i=0}^{N-1} x_i \exp\left(-j2\pi \frac{ki}{N}\right) \\ Y_k &= \frac{1}{N} \sum_{i=0}^{N-1} y_i \exp\left(-j2\pi \frac{ki}{N}\right) \end{cases}$$

En notant $z_i = x_i + jy_i$, une description plus compacte du contour est donnée par :

$$C_k = \frac{1}{N} \sum_{i=0}^{N-1} z_i \exp\left(-j2\pi \frac{ki}{N}\right) = X_k + jY_k$$

L'information « contour » de la silhouette est réellement contenue dans les descripteurs d'indice non nul. Le point (X_0, Y_0) permet uniquement de localiser la silhouette, il s'agit de son barycentre. La reconstruction s'effectue alors autour de ce point. Les abscisses et les ordonnées étant réelles, leur transformée de Fourier est nécessairement symétrique *i.e.* $X_{-k} = X_k^*$ et $Y_{-k} = Y_k^*$. La seule connaissance des coefficients d'indices positifs suffit donc à définir complètement la transformée.

Afin de caractériser le contour de manière unique, il convient d'effectuer une normalisation des descripteurs de Fourier pour assurer l'invariance par rapport au point de départ sur le contour. En effet, la paramétrisation du contour conditionne le résultat de la transformée de Fourier : soient C_k^0 et C_k , deux descriptions du même contour décrit par N points équirépartis, le décalage l du

point de départ sur le contour agit sur les descripteurs comme suit :

$$C_k = C_k^0 \tau^{-k} \quad \text{avec} \quad \tau = \exp\left(j2\pi \frac{l}{N}\right)$$

On peut s'affranchir de la dépendance en τ en s'inspirant des travaux de Arbter, *i.e.* en formant les invariants suivants [ARBT-89] :

$$\tilde{C}_k = |C_k| \Phi_k^{\lambda_0} \prod_{i=1}^n \Phi_{k_i}^{k_i \lambda_i}, \quad k_i = \pm 1, \pm 2, \dots$$

avec $\Phi_k = \frac{C_k}{|C_k|}$, $\Phi_{k_i} = \frac{C_{k_i}}{|C_{k_i}|}$ et $\lambda_0 + \sum_{i=1}^n k_i \lambda_i = 0$, $\lambda_i \in \mathbb{Z}$.

Pour $n = 1$, $\tilde{C}_k = |C_k| \Phi_k^{\lambda_0} \Phi_{k_i}^{k_i \lambda_i}$. Dans le cas le plus simple, $\lambda_0 = 1$ et $1 + k_1 \lambda_1 = 0$, soit par exemple, $k_1 = 1$ et $\lambda_1 = -1$. D'où,

$$\tilde{C}_k = |C_k| \Phi_k \Phi_1^{-k} = |C_k| \frac{C_k}{|C_k|} \frac{C_1^{-k}}{|C_1|^{-k}}$$

i.e.
$$\tilde{C}_k = C_k \frac{C_1^{-k}}{|C_1|^{-k}}$$

Pour une décomposition suivant lignes et colonnes, sachant $C_k = X_k + jY_k$, on a donc les descripteurs normalisés pour les deux coordonnées :

$$\begin{cases} \tilde{X}_k &= X_k \frac{C_1^{-k}}{|C_1|^{-k}} = X_k \frac{(X_1 + jY_1)^{-k}}{|X_1 + jY_1|^{-k}} \\ \tilde{Y}_k &= Y_k \frac{C_1^{-k}}{|C_1|^{-k}} = Y_k \frac{(X_1 + jY_1)^{-k}}{|X_1 + jY_1|^{-k}} \end{cases}$$

Remarque : $\tilde{X}_0 = X_0$ et $\tilde{Y}_0 = Y_0$. Il n'y a donc aucune modification du barycentre de la silhouette.

BIBLIOGRAPHIE

[ARBT-89] K. Arbter, « Affine-Invariant Fourier Descriptors », *From Pixels to Features*. Amsterdam, The Netherlands : Elsevier Science, pp. 153-164, 1989.

[BILL-98] D. Billon, F. Fohanno, « Theoretical performance and experimental results for synthetic aperture sonar self-calibration », *Proc. OCEANS'98 MTS/IEEE*, Nice, France, 1998, pp. 965-970.

[BLON-97] P. Blondel, and B.J. Murton, *Handbook of Seafloor Sonar Imagery*, Chichester: John Wiley and Sons, 1997.

[BOSS-97] P.L. Bossart, D. David, J.M. Dinten et J.M. Chassery, « Détection de contours réguliers dans des images bruitées et texturées : une approche par contours actifs multiéchelle », *Traitement du Signal*, vol. 14, n° 2, 1997.

[BOUV-92] M. Bouvet, *Traitements des signaux pour les systèmes sonar*, Ed. Masson, Coll. Technique et scientifique des télécommunications, 1992.

- [CHES-99] C. Chesnaud, P. Réfrégier, « Statistical region snake-based segmentation adapted to different physical noise models », *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no 11, pp.1145-1157, 1999.
- [COCQ-95] J.-P. Cocquerez, S. Philipp, *Analyse d'images : filtrage et segmentation*, Ed. Masson, 1995.
- [COHE-93] L.D. Cohen and I. Cohen, « Finite-Element Methods for Active Contour Models and Balloons for 2-D and 3-D Images », *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, n° 11, nov. 1993.
- [COLL-98] C. Collet, P. Thourel, M. Mignotte, P. Pérez et P. Bouthemy, « Segmentation markovienne hiérarchique multimodèle d'images sonar haute résolution », *Traitement du Signal*, vol. 15, n° 3, pp. 231-250, 1998.
- [DAVI-87] L. Davis, (Editor), *Genetic Algorithms and Simulated Annealing*, Morgan Kaufmann Publishers, Los Altos, CA, 1987.
- [DAVI-91] L. Davis, (Editor), *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New-York, 1991.
- [EVAN-93] M. Evans, N. Hastings And B. Peacock, *Statistical Distributions*, Second Edition - New-York: Chichester: Toronto: Singapore: John Wiley & Sons, Inc., 1993.
- [FIGU-97] M.A.T. Figueiredo, J.M.N. Leitão, A.K. Jain, « Adaptive parametrically deformable contours », *Proceedings of Computer Vision and Pattern Recognition*, pp.35-50, 1997.
- [GOLD-91] D.E. Goldberg, K. Deb, « A Comparative Analysis of Selection Schemes Used in Genetic Algorithms », *Foundations of Genetic Algorithms*, ed. Gregory J. E. RAWLINS, San Mateo, California, USA : Morgan Kaufmann Publishers, pp. 69-93, 1991.
- [GOOD-76] J.W. Goodman, « Some fundamental properties of speckle », *J. Opt. Soc. Am.*, vol. 66, n° 11, pp. 1145-1150, 1976.
- [GUER-98] D. Guériot, « Utilisation des algorithmes génétiques pour des problèmes d'optimisation spécifiques : application aux réseaux de neurones et au traitement d'images sonar », *Thèse de doctorat*, Université de Haute Alsace, décembre 1998.
- [GUER-99] F. Guérault, « Techniques statistiques pour l'estimation de la position d'un objet dans des images bruitées », *Thèse de doctorat*, Université d'Aix-Marseille, janvier 1999.
- [HADA-23] J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, 1923.
- [HASE-99] M. Haseyama, M. Kumagai, H. Kitajima, « A genetic algorithm based image segmentation for image analysis », *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, Phoenix, Arizona, 1999.
- [HENO-93] H. Henocq, G. Burel, « Reconnaissance d'objets 3D par analyse de Fourier de la silhouette », *Revue Technique Thomson-CSF*, vol. 25, n° 1, 1993.
- [HOUC-95] C.R. Houck, J.A. Joines, M.G. Kay, « a genetic algorithm for function optimization : a Matlab implementation », *NCSU-IE Technical Report* 95-09, 1995.
- [KERV-94] C. Kervrann and F. Heitz, « A hierarchical statistical framework for the segmentation of deformable objects in image sequences », *IEEE Conf. Computer Vision Pattern Recognition*, Seattle, pp. 724-728, june 1994.
- [JAIN-96] A. K. Jain, Y. Zhong and S. Lakshmanan, « Object Matching Using Deformable Templates », *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, n° 3, pp.267-278, 1996.
- [JAIN-98] A. K. Jain, Y. Zhong and M.P. Dubuisson-Jolly, « Deformable template models : a review », *Signal Processing*, vol. 71, pp. 109-129, 1998.
- [JANE] *Jane's Underwater Warfare Systems*, Ed. by Anthony J. Watts, 10th Edition 1998-1999.
- [KASS-88] M. Kass, A. Witkin, D. Terzopoulos, « Snakes : Active Contour Models », *International Journal of Computer Vision*, vol. 1, pp.321-331, 1988.
- [LAM-98] C.L. Lam and S.Y. Yuen, « An unbiased active contour algorithm for object tracking », *Pattern Recognition Letters*, vol. 19, pp. 491-498, 1998.
- [MAIT-01] H. Maître, *Traitement des images de RSO*, Hermès, 2001.
- [MICH-94] *Genetic Algorithms + Data Structures = Evolution Programs*, AI Series, Springer-Verlag, New-York, 1994.
- [MIGN-00] M. Mignotte, C. Collet, P. Perez, P. Bouthemy, « Hybrid Genetic Optimization and Statistical Model-Based Approach for the Classification of Shadow Shapes in Sonar Imagery », *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, n° 2, 2000.
- [MITC-83] O.R. Mitchell and T.A. Grogen, « Evaluation of Fourier Descriptors for Target Recognition in Digital Imagery », *Final Technical Report RADC-TR-83-33*, Purdue Univ., School of Electrical Engineering, Lafayette, IN, february 1983.
- [PAL-96] S.K. Pal and P.P. Wang, *Genetic Algorithms for Pattern Recognition*, CRC Press, Boca Raton, FL, 1996.
- [PAPO-84] A. Papoulis, *Probability Random, Variables, and Stochastic Processes*, Mc Graw-Hill, 1984.
- [PUYH-00] P. Puyhabilier, « La modernisation des chasseurs de mines tripartites », *L'Armement*, n° 71, octobre 2000.
- [SEET-91] G. Seetharaman, A. Narasimhan, A. Sathe and L. Storck, « Image Segmentation with Genetic Algorithms : a Formulation and Implementation », *SPIE, vol. 1569, Stochastic and Neural Methods in Signal Processing, and Computer Vision*, pp. 269-273, 1991.
- [SOME-93] M.L. Somers, « Sonar Imaging of the Seabed: Techniques, Performances, Applications », in *Acoustic Signal Processing for Ocean Exploration*, J.M.F. Moura and I.M.G. Lourtie (eds.), Canadian Government (pp. 355-369), 1993.
- [WALL-80] T.P. Wallace, P. A. Wintz, « An efficient three-dimensional aircraft recognition using normalized Fourier descriptors », *Comput. Graphics Image Processing*, vol. 13, pp. 96-126, 1980.
- [ZHON-98] Y. Zhong, A.K. Jain and M.P. Dubuisson, « Object Tracking Using Deformable Templates », *Int. Conf. On Computer Vision*, Bombay, India, pp. 440-445, 1998.

Manuscrit reçu le 31 octobre 2001

Classification de mines sous-marines à partir de l'image sonar brute

LES AUTEURS

Isabelle QUIDU



Née à Pontivy en 1974, Isabelle QUIDU est Docteur de l'Université de Bretagne Occidentale en électronique (décembre 2001). Elle a préparé sa thèse sous convention CIFRE au sein du Service Études Générales Sonar de Thales Underwater Systems à Brest en liaison avec le Laboratoire d'Électronique et Systèmes de Télécommunications (UMR CNRS 6165). Ses travaux portaient sur la classification multi-vues de mines marines à partir d'images sonar haute

résolution. Elle occupe aujourd'hui le poste d'ingénieur traitement du signal. Ses thèmes de recherche concernent l'amélioration des traitements du signal et de l'image équipant les systèmes de chasse aux mines et aéroportés.

Jean-Philippe MALKASSE



Jean-Philippe MALKASSE est ingénieur Supélec, promotion 93 (Radiocommunications, Option Recherche) et mastère en Image et Systèmes de Télévision de Télécom Paris, promotion 95. Il est responsable du Laboratoire Traitement du Signal et de l'Information du Service Études Générales Sonar de THALES UNDERWATER SYSTEMS à Brest. Ce laboratoire mène les études amont en détection, classification, identification et aides opérateurs appliquées aux systèmes

sonar des lignes de produits Guerre des Mines (sonar d'imagerie) et Aéroporté (systèmes de bouées, sonar héliporté).

Gilles BUREL



Gilles BUREL a obtenu le diplôme d'ingénieur de l'École Supérieure d'Électricité (Supélec) en 1988, le doctorat de l'Université de Bretagne Occidentale en décembre 1991 et l'Habilitation à Diriger des Recherches en avril 1996. Il est auteur de 19 brevets et de 80 articles. Après avoir débuté sa carrière à Thomson CSF en 1988, dans le domaine du traitement d'images, il a ensuite rejoint Thomson Multimédia. Depuis septembre 1997, il est Professeur à l'Université de Bretagne Occidentale, où il assure la responsabilité de l'une des trois équipes de recherche du Laboratoire d'Électronique et Systèmes de Télécommunications (UMR CNRS 6165), l'équipe « Traitement du Signal pour les Télécommunications ».

Ses nouvelles activités de recherche se situent essentiellement dans le domaine des communications numériques (interception et analyse, systèmes MIMO, transmissions furtives).

Annexe 2 : processus ATR développé chez Thales en 2005

Cette annexe est une copie de l'acte [QUI05] du colloque IEEE OCEANS en 2005 présentant une chaîne ATR développée avec mes collègues du service Etudes Générales Sonar (EGS) de Thales Underwater Systems Brest. Les deux modes sonar, les modes détecteur et classificateur, sont successivement exploités pour détecter sur écho un objet potentiellement dangereux puis le classifier en mine ou non-mine à partir de son ombre portée sur le fond.

Automatic Classification for MCM Systems

I. Quidu, N. Burlet, J.-P. Malkasse and F. Florin
ThalesUnderwater Systems, Route de sainte Anne du Portzic, CS 43814,
29238 Brest cedex 3, France
isabelle.quidu@fr.thalesgroup.com, nicolas.burlet@fr.thalesgroup.com

Abstract - Sonar systems, designed to detect stealthy mines, will now also hand over to the classification process many more small non-mine bottom objects (NOMBO), with mine-like target strength. In these conditions the global search effort is transferred to the classification function by reducing the number of pointless false detection alarms, i.e. by classifying a NOMBO as a non mine like contact (NON MILCO). This paper then aims at giving demonstrative results and means to perform automatic classification while keeping low false classification rate. The considered method is divided into two steps : a first coarse long range pre-classification only provides relevant mine like echoes (MILEC) to be then processed by a shorter range classification. The false alarm reduction of the first step is performed with a long range detection sonar by analysing the echo signal structure and its evolution given the sonar motion. Then, the second step is image based classification and needs a high frequency classification sonar to provide a more detailed acoustic image of the object (especially concerning its cast shadow shape).

I. INTRODUCTION

Modern mines have become very stealthy. High resolution sonars have been designed to counter this threat by performing improved probability of detection at the first step of the mine hunting operation. Nevertheless we are now facing an increasing number of data passed through the classification module which has to discriminate Mines from Non Mine Bottom Objects (NOMBO). The need for automatic enhanced classification becomes of high necessity.

Mine hunting is classically considered as a three steps process: (1) detection, (2) classification and (3) identification (possibly followed by mine disposal). Detection aims to find objects in the water column or on the sea floor which feature a high acoustic return i.e. mine like echoes (MILEC), and aims to eliminate most of non-mine returns, i.e. false alarm reduction. Classification uses the measured acoustical characteristics of MILEC (echoes strength and signature, dimensions, shadows) to eliminate most of remaining false alarms and keep mine like contacts (MILCO). The final stage, identification, confirms classification by a formal recognition of the MILCO as a mine, and collects a maximum of information on the mine (type of mine, country of origin, settings, etc.). The detection is classically based on the simple echo level compared to the background noise (Detection Threshold), and discriminates between MILEC and NON-MILEC objects. At the detection stage, any object having an echo level similar to those of the mines can be considered as a

mine and is most often pointed out as MILEC. When mines still used to be high echo level objects, the number of non-mine objects passed on to the classification step was relatively small and, as a result, the overall mine hunting effort was mostly dimensioned by the detection process. This has rather drastically changed now that modern mines have become very stealthy. Sonar systems, designed to detect stealthy mines, will now also hand over to the classification process many more small non-mine bottom objects, with mine-like echo levels. In these conditions the global search effort is transferred to the classification function ([8], [9], [10]).

We then propose a two-steps classification for Autonomous Underwater Vehicles (AUV). The first coarse long range pre-classification (paragraph II) consists in a high resolution analysis of the echo along range axis. The second step is high resolution sonar image based classification while reducing pointless false detection alarms (paragraph III).

II. Echo-based classification

The long-range echo pre-classification is divided into three main steps:

- multi-levels threshold
- temporal analysis of the echoes
- automatic pre-classification

The final product of the processing is a FAR (False Alarm Reduction) image, where only relevant echoes are visible.

A. Detection image

The image obtained in this first step of pre-classification is the result of a special normalization of the sonar data. The result depends only on the physical features of the targets, the grazing angle and the nature of the bottom. Then, a multi-levels threshold, corresponding to different possible threats, is applied.

B. Temporal analysis

The second step of echo pre-classification is based on the temporal analysis of a sequence of images (assuming that the sonar is moving).

This processing extracts different features for each persistent echo detected during the first step and analyzes the fluctuation of the echo structure with time due to variations of grazing and view angles.

C. Automatic Pre-Classification

At last, an automatic classification algorithm (based on maximum likelihood with models stored in a database) is performed. The result is the FAR image.

Although a completely automatic processing is made, an OAM (Operator Aided Mode) is proposed too. It uses the resulting by-products of the entire pre-classification processing (visualization of the detection image, for example).

D. Results

Here is a result of echo pre-classification, performed on real data.

The following figure is a visual comparison between the sonar image and the final FAR image. The pre-classification process reduces significantly the number of MILEC (Fig. 1). On the FAR image, four echoes are detected and are surrounded with a blue circle. Three of them correspond to mines placed for the sea trial.

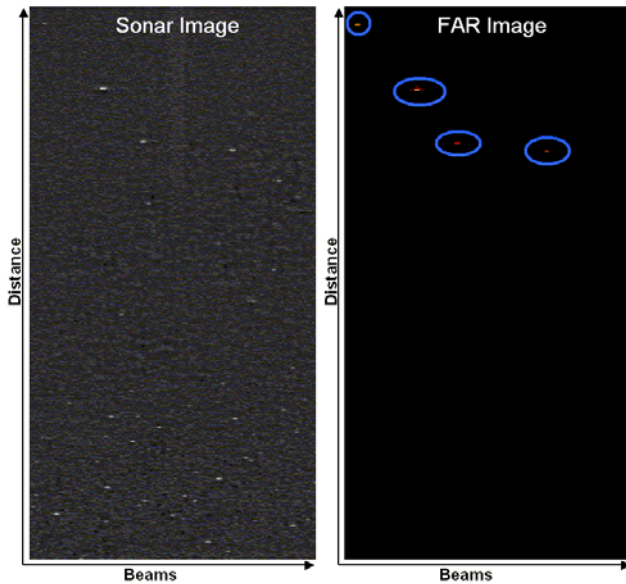


Fig. 1. Comparison between sonar and FAR images

III. Shadow-based classification

A. Flowchart

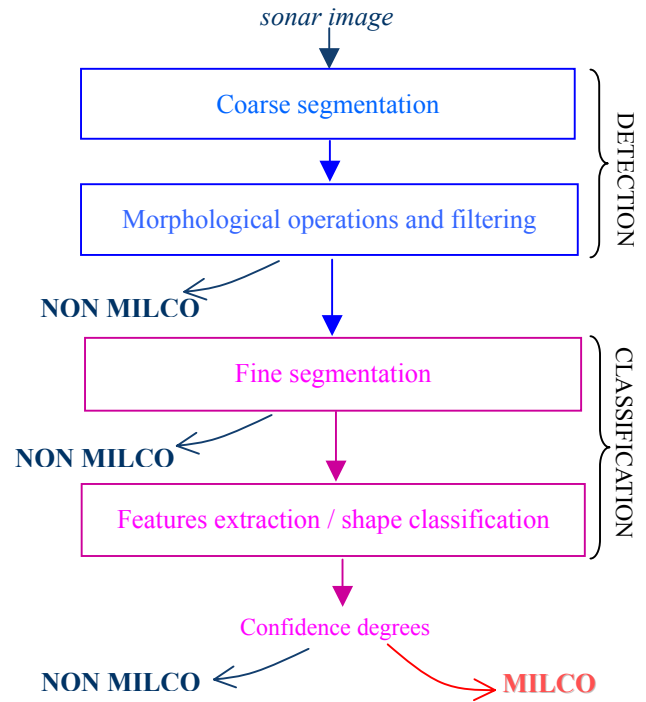


Fig. 2. Flowchart of shadow-based classification

This flowchart consists of successive operations on sonar image designed to classify a NOMBO as a NON MILCO. We can distinguish a shadow-based detection phase (blue boxes) from the classification phase (pink boxes).

B. Shadow-based detection phase

a) Coarse segmentation of shadow areas

This is a three-steps technique :

- Image enhancement: the goal is to mean bottom reverberation grey-levels while expanding echo and shadow grey-levels. As a consequence, shadow areas are easier to detect by thresholding. At this step we memorize echo pixels (maximal grey-level).
- Filtering: it increases bottom reverberation and shadow grey-levels smoothing.
- Thresholding between areas of interest easily obtained thanks to the first step.

b) Morphological operations and filtering: we aim at discarding 2D coarse shadows that don't respect some geometrical conditions. By performing erosion, we only keep the objects of a minimal size depending on the researched targets. Then, shadows shapes must have some typical global properties and must contrast with seafloor reverberation.

C. Shadow-based classification phase

- a) Fine segmentation by means of statistical snakes [1]. A quick comparison between the coarse and the fine segmentation allows us to discard shadows rather related to bottom structures.
- b) Features extraction: in [3], we have pointed out the advantage of using a hybrid set of features. In

addition to the common used topological parameters, we were interested in features derived from statistical moments of 2D-shapes [5]. Some particular functions of moments have the useful property of affine invariance [4]. So they are well-suited for cast shadows viewed under weak perspective conditions. Thanks to their different origins, these features appeared complementary for classifying both classical and stealthy mines.

- c) Multi-class classification by means of a Multi-Layer Perceptron (MLP) and introduction of a degree of confidence [6]. The number of input layer nodes equals to the number of features used. The output one is made of as many nodes as the number of types of mines in the database.

Given the output O_j of each output node j , we derive a degree of confidence d_{conf} from the elected class and rival classes.

If d_{conf} is insufficient, then the object is classified as a NON MILCO, otherwise it is classified as

MILCO of type $j = \text{Arg} \max_{j \in \{1, \dots, 5\}} O_j$ with a degree of

confidence d_{conf} .

D. Experiments

This flowchart has been tested on a real Synthetic-Aperture Sonar (SAS) image where four mines lie [10], [7]. Hereafter are given results of each step (see Fig. 3 and following ones).

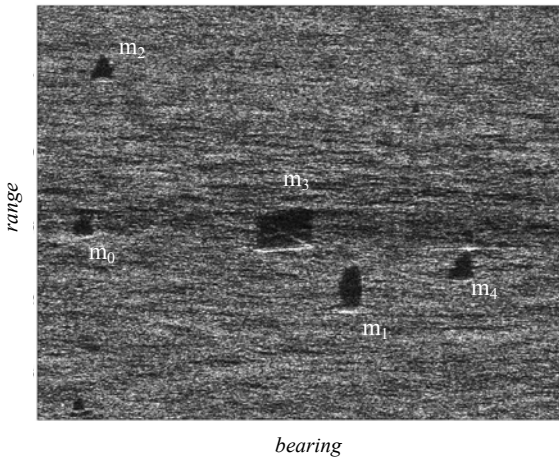


Fig. 3. SAS image



Fig. 4. After enhancement

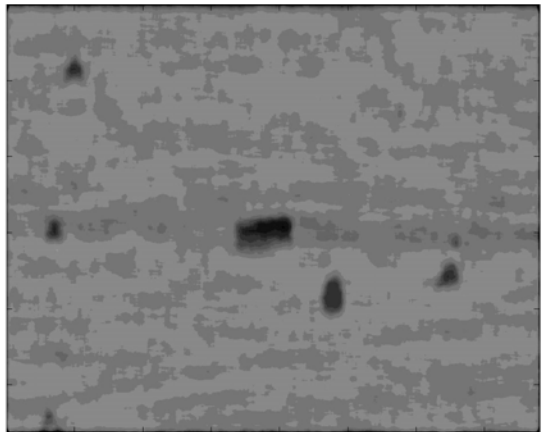


Fig. 5. After filtering

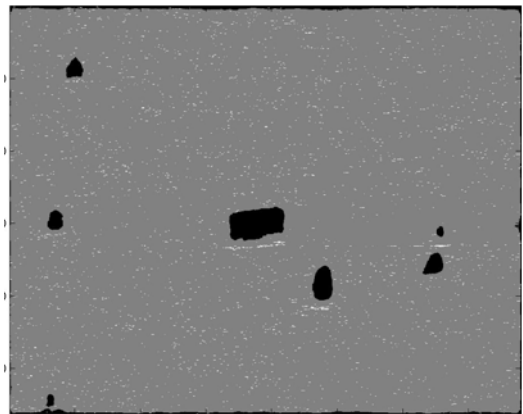


Fig. 6. After shadow/bottom thresholding and echo pixels retrieval

The coarse shadow shapes obtained on Fig. 3 which respect geometrical conditions in terms of size and topology are selected for the classification phase which begins by a finer segmentation (see Fig. 7).

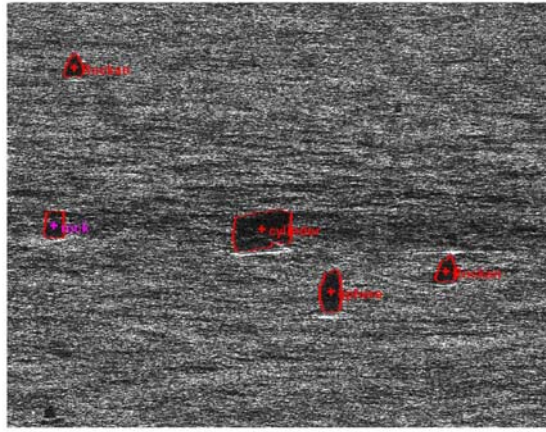
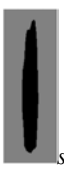





Fig. 7. Fine segmentation and classification results

Fine segmentation allows us to discard the m_0 object, which is a rock, regarding its difference with coarse segmentation result. The other objects which are all mines and classified as such (see Table I).

Object (m_x on Fig. 3)	Class	d_{conf}	Rival class
 m_1 sphere ($\varnothing=0.8m$)	sphere	0.72	Manta mine
 m_2 Rockan mine	Rockan Mine	0.99	Manta mine
 m_3 Cylindre ($L=2m$ et $\varnothing=0.5m$)	cylindre	1	-
 m_4 Manta Mine	Rockan Mine	0.84	cylinder

IV. Conclusion

The automatic classification for MCM systems proposed in this article is based on a two-steps processing after the classical detection step of the mine hunting process. A long-range preclassification, associated with an efficient automatic shadow classification algorithm, allows to better readjust the global search effort of mines between the different processings. This two-steps classification fully exploits sonar capabilities by performing :

- signal-based analysis of the echo thanks to high

range resolution,

- image-based analysis of the acoustic shadow thanks to SAS high cross-range resolution.

Moreover, combining echo and shadow information is more robust. This is especially of high interest for AUV missions.

Acknowledgments

The authors thank very much DGA/SPN (Service des Programmes Navals) that supported part of preliminary studies and DGA/GESMA (Groupe d'Etudes Sous-Marines de l'Atlantique) that provided the associated sonar data.

REFERENCES

- [1] C. Chesnaud, P. Réfrégier, "Statistical region snake-based segmentation adapted to different physical noise models", in *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 11, November 1999, pp.1145-1157.
- [2] I. Quidu, "Classification multi-vues d'un objet immergé à partir d'images sonar et de son ombre portée", *Ph.D. thesis*, Université de Bretagne Occidentale, décembre 2001.
- [3] I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé, "Mine classification using a hybrid set of descriptors", *Proceedings of OCEANS'2000 MTS/IEEE*, Providence, Rhode Island, 2000.
- [4] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants", *Pattern Recognition*, Vol.26, No.1, pp.167-174, 1993.
- [5] R.G. Prokop and A.P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP: Graphical Models and Image Processing*, vol. 54, No. 2, pp. 438-460, September 1992.
- [6] J. Hertz, A. Krogh and R. G. Palmer, *Introduction to the theory of neural computation*, Vol. I, Santa Fe Institute, Addison-Wesley Publishing Company, 1991.
- [7] Billon, D. and F. Fohanno, "Theoretical performance and experimental results for synthetic aperture sonar self-calibration", *OCEANS'98 MTS/IEEE*, pp. 965-970, 1998.
- [8] F. Florin, F. Van Zeebroeck, I. Quidu and N. Le Bouffant, « Classification performances of Mine Hunting Sonar : Theory, practical results and operational applications », *Undersea Defence Technology (UDT) Europe 2003*, Malmö, Sweden, 24-26 June 2003.
- [9] F. Florin, I. Quidu et N. Le Bouffant, « Shadow classification performances evaluation of Mine Hunting Sonar », *Undersea Defence Technology (UDT) Asia 2003*, Singapore, 12-14 novembre 2003.
- [10] F. Florin, F. Fohanno, I. Quidu et J.-P. Malkasse, « Synthetic Aperture and 3D Imaging for Mine Hunting Sonar », *Undersea Defence Technology (UDT) Europe 2004*, Nice, 22-25 juin 2004.

Annexe 3 : Détection incohérente de changement

Cette annexe est une copie de l'acte [QUI12a] de la conférence *European Conference on Underwater Acoustics (ECUA) 2012*, publié dans les actes de *Meetings of Acoustics (POMA)*, décrivant la chaîne de traitement, recalage puis détection de changement incohérente, appliquée aux données canadiennes et développée dans le cadre de l'arrangement spécifique n°42 franco-canadien.

INCOHERENT CHANGE DETECTION USING AMPLITUDE SIDESCAN SONAR IMAGES

I. Quidu ENSTA Bretagne, STIC/OSM, Lab-STICC UMR CNRS 6285, 2 rue François Verny,
29806 Brest Cedex 9, France

1 INTRODUCTION

The automatic detection of temporal changes in sonar images is of high interest for monitoring critical areas as ports or channels used by submarines for instance. This problem is addressed here as incoherent change detection between two sonar passes.

To be able to compare two surveys, sonar images have first to be registered to each other. A global translation between the master image and the slave image has been estimated here at a coarse-to-fine image resolution by minimizing the Kullback-Leibler divergence between the first and the second sets of amplitude distributions related to homogeneous areas of the master image. This is detailed in section 2.

Then two different methods for the detection of changes in terms of contacts in the common area of the registered images are proposed in section 3. The first method is one of the well-known techniques used for synthetic aperture radar imagery: the log ratio of images. The second one is made of two stages: firstly a goodness-of-fit test is applied to every image divided into small overlapped snippets in order to detect statistical deviations due to possible contacts; secondly a pixelwise difference is simply performed between two images whose detected snippets have been put in a non-zero pixel value. For both methods morphological operations and local correlations have been performed in order to discard false alarms.

Advantages and drawbacks of the two proposed methods are discussed in section 4 with results on a set of data collected by Defence R&D Canada Atlantic with a high-frequency sidescan sonar a month apart in the winter of 2008 in a port area.

2 RIGID-BODY REGISTRATION FOR SONAR IMAGES

2.1 Previous work

Sonar images registration has been widely studied in the 2000s. Proposed methods can be divided into two classes: symbolic methods and non symbolic or iconic methods [1]. The first ones are based on segmented images and/or extracted features or salient points (following a SIFT or SURF algorithm for instance) [2,3,4,5,6]. The second ones are only based on pixel intensities and aim to establish a global relationship between the pixels of the two images using their gray levels only [7]. We proposed here a method of the second class that uses the Kullback-Leibler divergence to measure the similarity between the master and the slave sets of amplitude distributions of homogeneous areas.

2.2 Proposed method

Here the two images to be registered are geo-referenced and the basic linear transform to be estimated is a translation. Moreover the registration procedure aims to provide overlapped areas where changes in terms of contacts have to be found. In such a context, registration cannot be based on the contacts themselves but rather on pixel distributions of homogeneous areas. Hence we were interested in developing an iconic method for estimating the global translation that provides the best image matching in terms of homogeneous areas. For this the similarity between the master and the slave sets of amplitude distributions is computed using a global Kullback-Leibler divergence. These sets are carried out into two steps: two 3-staged image pyramids are first built. The first pyramid decomposition is needed to speed up the registration procedure and only consists in subsampling the image by a factor of 2 along each spatial dimension. Between stages of the second pyramid, besides subsampling, a median filtering is performed in order to denoise the

images. This allows performing a segmentation of the master image at the coarser stage into six homogeneous areas by applying the Fisher method which consists in minimizing the variance within classes [8]. The resulting segmented image is used in the remaining as a mask providing the six homogeneous areas. Note that even at the coarser stage the image size is sufficient for allowing a good estimate of the six distributions of pixel level.

At each stage of the first pyramid this mask is used to collect six pairs of pixel level distributions for a given translation of the slave image compared to the master image. Then a symmetric Kullback-Leibler divergence D_{KL} is computed for each pair of discrete distributions P and Q [9]:

$$D_{KL}(P||Q) = \left[\sum_i P(i) \log \left(\frac{P(i)}{Q(i)} \right) + \sum_i Q(i) \log \left(\frac{Q(i)}{P(i)} \right) \right] / 2 \text{ where } i \in [0, 255]$$

The global Kullback-Leibler divergence for a given translation and for a given stage is the sum of the six symmetric Kullback-Leibler divergences. The estimated translation is the one that gives the lower global Kullback-Leibler divergence. At the coarser stage, the translation interval is bounded by the georeferencing accuracy. The estimated translation is used at the upper stage to initialize the new translation estimation bounded by only two pixels this time: this indeed corresponds to the subsampling factor between stages in rows and in columns. The global translation used for registering images is obtained by combining the translations estimated at every stage.

3 INCOHERENT CHANGE DETECTION

3.1 Log-ratio image

In the synthetic aperture radar (SAR) domain, image rationing pixel by pixel is performed for decades [10]. Ratio of images is preferred to difference because the ratio value only depends on relative change in amplitude between the two dates so that the detection of changes does not depend on the level of the pixels [11]. It is usually expressed in decibels (that is, by taking ten times the logarithm in base ten of the ratio of the intensities). This method aims to identify changes in the mean backscatter power of a scene. An estimate of the mean backscatter power is given by the pixel intensity. Moreover as the image is corrupted by speckle an averaging is required in order to ensure a good estimate [12]. Whereas the averaging is performed using N looks of the SAR data, we have to manage here with a single sonar look. Hence the two images are first processed by an average filtering using a sliding window. Moreover we deal with amplitude data, *i.e.* the square root of the intensity data: that induces a multiplication of the above ratio by two. Finally the log ratio of sonar images is given by:

$$I_{div} = 20 \log_{10} \frac{\langle I_m \rangle}{\langle I_s \rangle}, \text{ where } \langle I_m \rangle \text{ (resp. } \langle I_s \rangle) \text{ is the averaged master (resp. slave) image.}$$

In order to detect any change whatever its origin we take the absolute value, *i.e.* $|I_{div}|$. In this latter, pixel values over a given threshold are supposed to show a change in the backscattered energy. Here this threshold is fixed to the 99th percentile. In order to discard remaining false alarms in the thresholded image I_b , a binary morphological erosion is performed so that to keep objects as least as large than an object of size 30cm in azimuth and 50cm in range, this provides the image I_{bd} . Remaining detections are then processed individually: if the detected surface is not compact enough (proportion of the pixels in the convex hull that are also in the region must be over 0.7), it is rejected; otherwise we first perform a new but finer registration by cross-correlating the snippets corresponding to the area surrounding the detection in the master and the slave images, and then the new resulted log-ratio area is thresholded as above and the result compared to the previous one: no common pixel leads to a rejection.

3.2 Detection difference image

High-frequency sonar images suffer from a well-known multiplicative noise process that is commonly called *speckle* [13, 14]. When the number of scatterers in a resolution cell is sufficiently high, the pixel gray levels follow a Rayleigh law. Under this hypothesis, the ratio value between the standard deviation and the mean of pixels, called the coefficient of variation c_v , is constant and

equal to

$$c_v^{th} = \sqrt{\frac{4}{\pi}} - 1 = 0.52$$

When this coefficient locally differs from the above theoretical value, it is probably due to an object that locally disturbs the distribution.

The idea is to locally estimate this coefficient in the master and slave images and to compare the two processed images. Practically, each sonar image is first divided into snippets that consist of rectangular patches corresponding to an area of 3m by 3m with an overlap of 75%. For each snippet, the coefficient of variation c_v is computed and all the pixel levels of the corresponding snippet are set to the coefficient value or, for overlapped snippets, the average coefficient of variation. Let us call I_{cv_m} and I_{cv_s} the two processed images.

Before differencing, we only keep pixels whose level is at least equal to $m_{I_{cvx}} + 2\sigma_{I_{cvx}}$ ($x=m$ or s) where $m_{I_{cvx}}$ (*resp.* $\sigma_{I_{cvx}}$) is the average level (*resp.* the standard deviation) of the pixels of the entire processed image. Let us call I_{t_m} and I_{t_s} the two new images. The difference image $I_{td} = |I_{t_m} - I_{t_s}|$ is computed and normalised to one. In this image, we can see the main changes but some are false alarms. To discard them, a grey-level morphological opening with a structural element equivalent to an area of 1m in azimuth and 2m in range is performed: the resulting I_{tdo} image is normalised to one as well. A binary image I_{tdob} is then obtained by thresholding I_{tdo} to 0.7: remaining areas are detections. Each detection is processed individually: we first perform a new but finer registration by cross-correlating the snippets corresponding to the area surrounding the detection in the master and the slave images, and then the new difference area is thresholded as above and the result compared to the previous one: no common pixel leads to a rejection. A last test is performed in order to discard "ring-shaped" false alarms due to scale difference between detections: this test consists in verifying the central position of the hole in the ring.

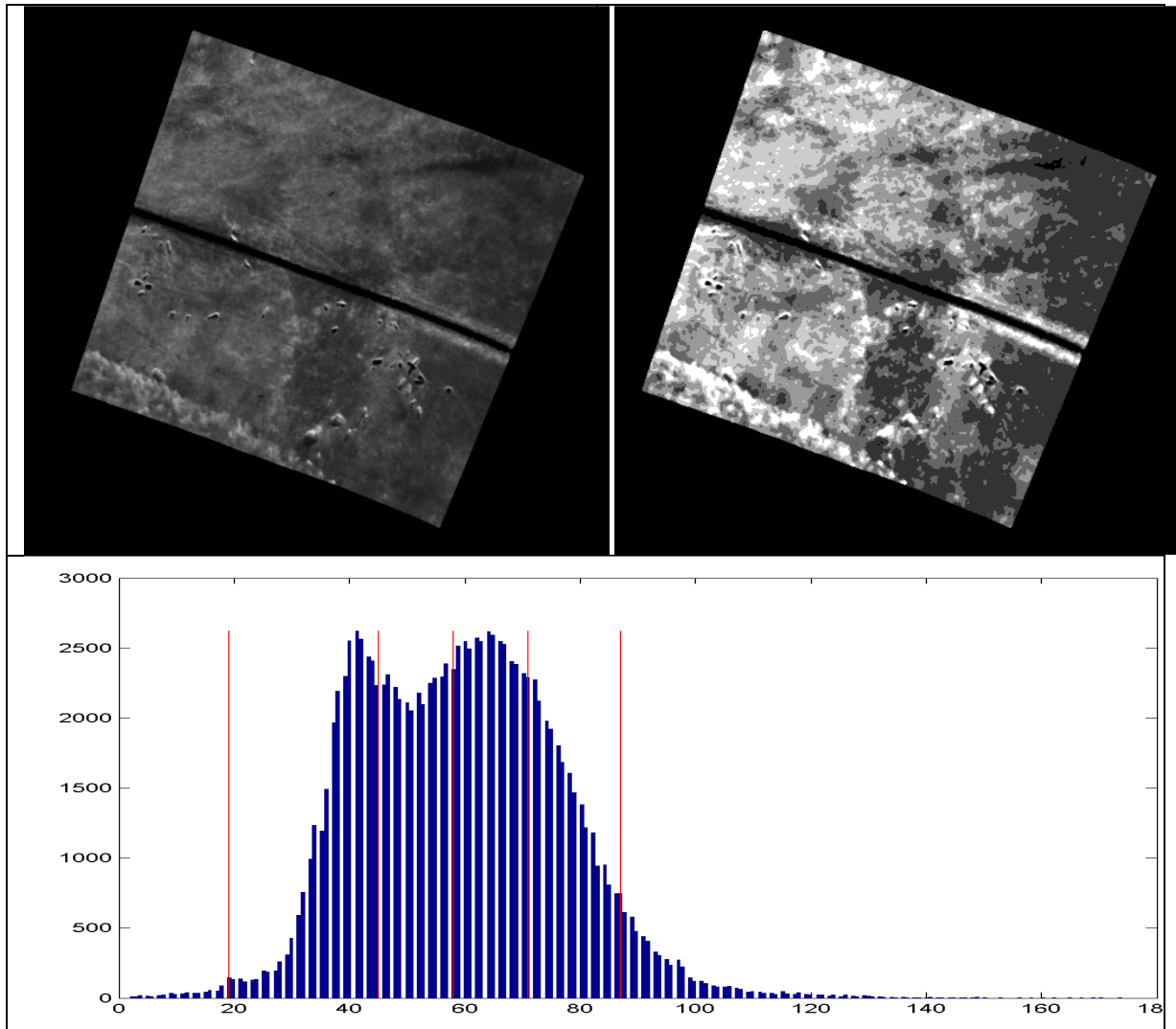


Figure 1 – First pair results. Left upper part: Filtered master image at the coarser resolution; Right upper part: segmented (6 classes) master image = mask; lower part: histogram of the filtered master image and the five thresholds (red lines) given by the Fisher method.

4 RESULTS

4.1 Data description

The data set was gathered with the Canadian Interim Remote Minehunting and Disposal System (IRMDS) which is a semi-submersible drone towing a Klein 5500 sidescan sonar operating at a centre frequency of 455 kHz and a bandwidth of 20 kHz. Two surveys were undertaken roughly one month apart during the winter of 2008 in a port area. Images were also normalised using a simple technique to remove effects of beam pattern, grazing angle and propagation loss and to create an image which is constant (in amplitude) with range. Finally, images were georeferenced on a grid with a resolution of 0.11m x 0.11m. Data samples falling into the same grid cell were averaged. See [15] for details. Here we will use only two image pairs among the five available ones:

- The first pair with many objects on the seafloor, some texture on the edges of the insonified area
- The second pair acquired during the vehicle turn (inducing distortion in the images), many objects and debris scattered on the seafloor with characteristic shapes.

4.2 Rigid-body registration

Intermediate results concerning the segmentation step are only given for the first pair in Figure 1. Figure 2 and Figure 3 give the registration results for the first and the second image pair.

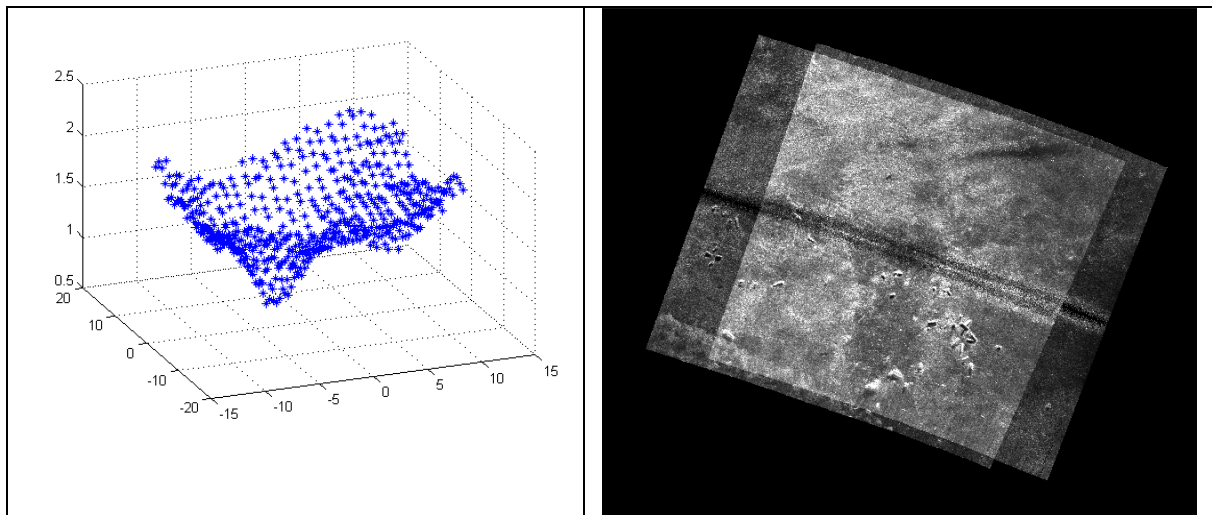


Figure 2 – Registration results for the first image pair: on the left, global Kullback-Leibler divergence vs. xy-translation at the coarser stage (the lower divergence indicates the good translation); on the right, superimposed registered images.

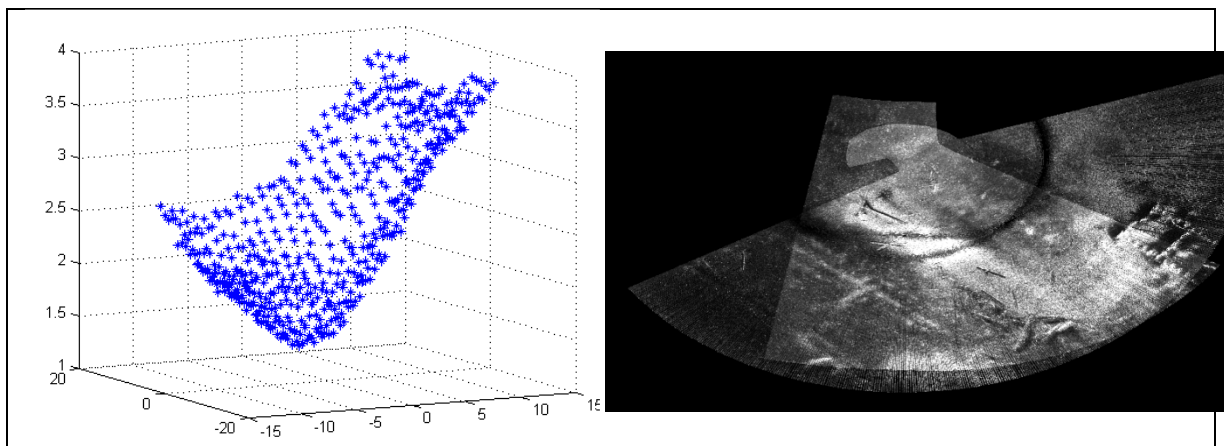


Figure 3 – Registration results for the second image pair: on the left, global Kullback-Leibler divergence vs. xy-translation at the coarser stage (the lower divergence indicates the good translation); on the right, superimposed registered images.

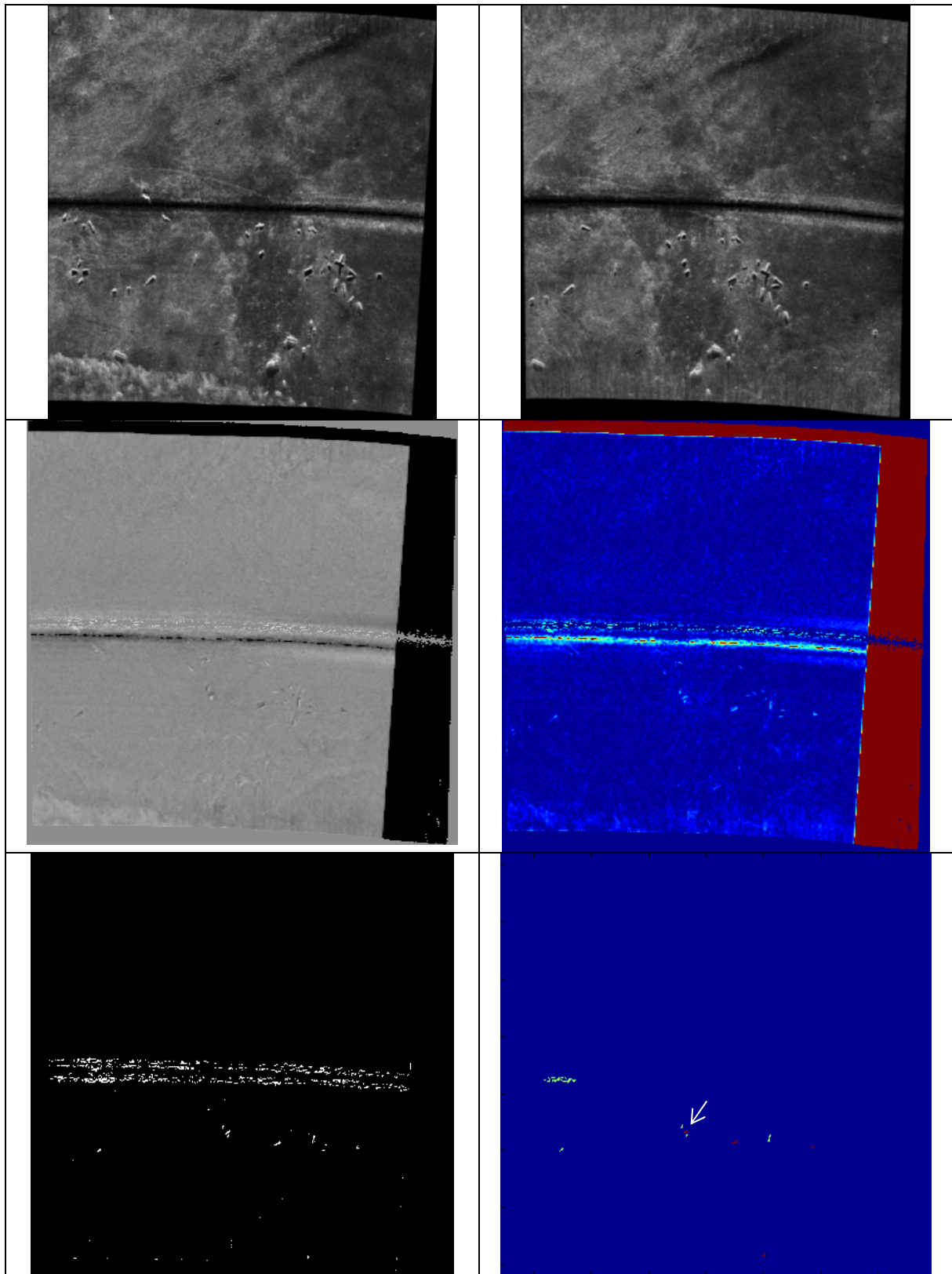


Figure 4 – First pair results for the log-ratio image method. Upper part: averaged images $\langle I_m \rangle$ and $\langle I_s \rangle$; Middle part : left figure I_{div} , right figure $|I_{div}|$; Lower part: left figure shows the thresholded image I_b and the right figure shows the four final detections in red (the added object is indicated by an white arrow) superimposed to the remaining detected objects in green after morphological operation I_{bd} .

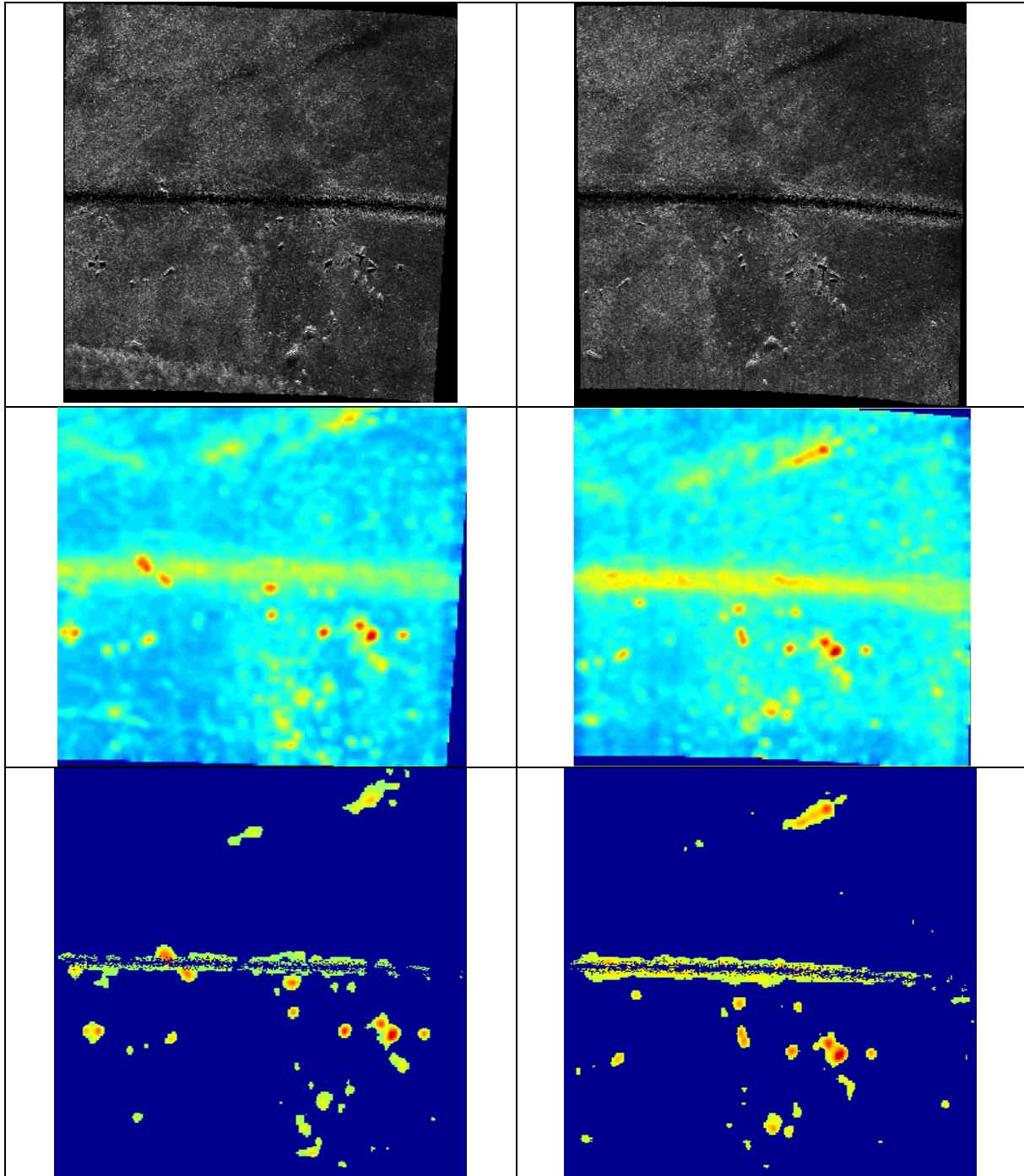


Figure 5 – First pair results. Upper part: sonar images (master one I_m on the left, slave one I_s on the right); Middle part: the processed image I_{cv_x} ($x=m$ or s) by computing the coefficient of variation; Lower part: I_{t_x} after having kept pixels whose level is at least equal to $m_{I_{cvx}} + 2\sigma_{I_{cvx}}$.

4.3 Incoherent change detection

Both methods are assessed on the first image pair where several objects lay on the seafloor. An object has been added manually in the slave image in order to provide some ground truth. From now, the images are rotated so as to be able to adapt morphological structural elements in range and in azimuth.

4.3.1 Log-ratio image

Figure 4 shows successively the averaged images $\langle I_m \rangle$ and $\langle I_s \rangle$, the log-image ratio I_{div} , its absolute $|I_{div}|$, the thresholded image I_b and then, selected detections given the procedure detailed in section 3.1. The added object is well detected with 3 false alarms mainly due to broken objects.

4.3.2 Detection difference image

Figure 5 shows columnwise: the sonar image (master one I_m on the left, slave one I_s on the right), the resulted image after the computation of the coefficient of variation I_{cv_x} ($x=m$ or s), and then I_{t_x} after having kept pixels whose level is at least equal to $m_{I_{cv_x}} + 2\sigma_{I_{cv_x}}$.

Figure 6 shows in the upper part the difference image before (I_{td} on the left) and after (I_{tdo} on the right) morphological opening. In order to compare these two images, they have been binarised by setting all non-zero pixels to one and then superimposed. The resulting image is showed in the left lower part of the figure whereas the right lower part shows the remaining detections following the procedure detailed in section 3.2. The added object is well detected with 3 false alarms due to image frontier or contrast change.

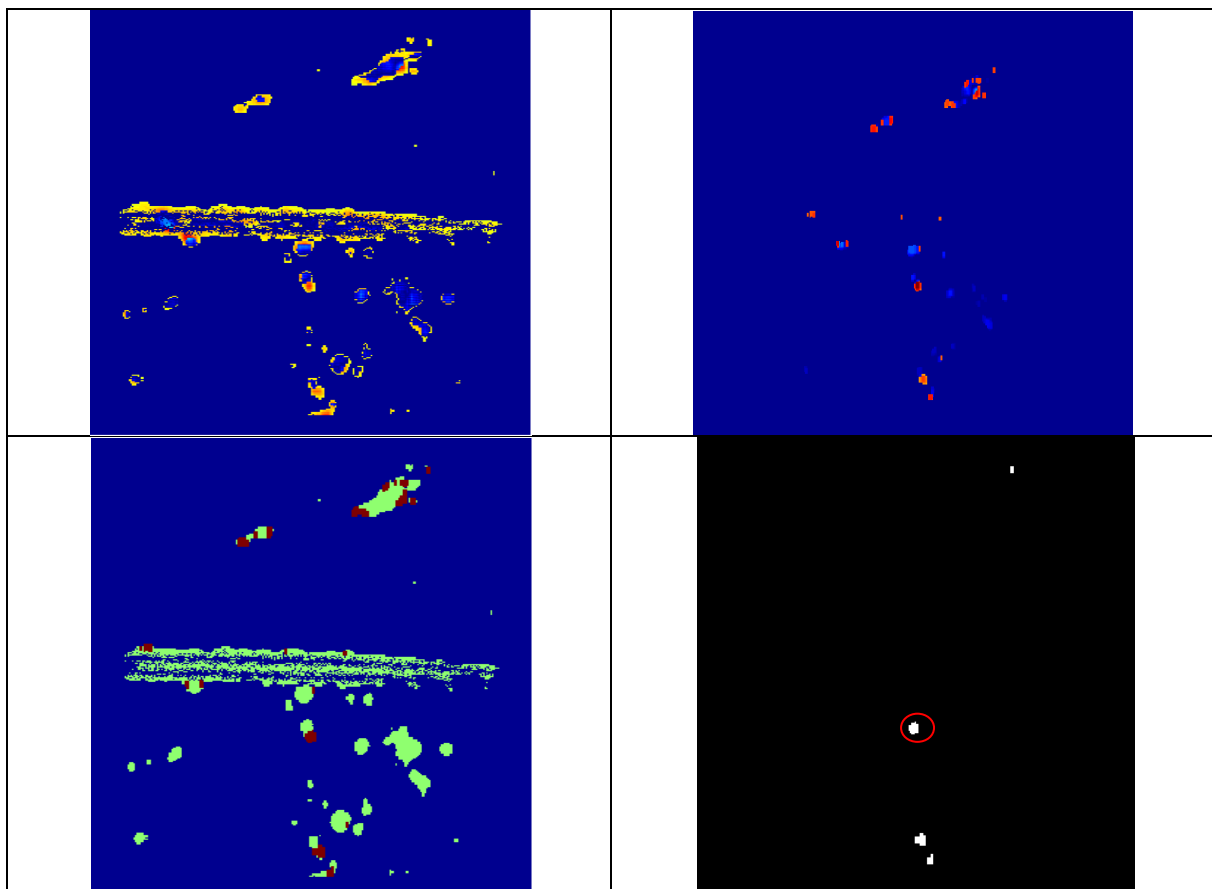


Figure 6 – First pair results for the detection difference image method. Upper part: difference image I_{td} on the left and after morphological grey-level opening I_{tdo} on the right ; Lower part : on the left the superimposition of the images I_{td} (green) and I_{tdo} (red) after a binarization and on the right the 4 remaining detected objects where the added object is encircled.

5 CONCLUSION

In this paper, the problem of incoherent change detection between two sonar passes has been addressed. Two methods have been proposed and applied to a set of amplitude sonar data gathered in a port area. Both proposed methods have similar performances in terms of change detection with a low false alarm rate. However, the log-image ratio is more sensitive to small changes as shape distortion than the second method based on the difference of values corresponding to the coefficient of variation C_v . Indeed whereas the first is a low level processing that compares (averaged) images pixel by pixel, the second compares the high level information returned by a C_v computer.

ACKNOWLEDGEMENTS

The author thanks DRDC Atlantic for the release of KLEIN 5500 Change Detection dataset and French DGA/DS for having supported this work.

REFERENCES

1. C. Chailloux, *Recalage d'images sonar par appariement de régions. Application à la génération d'une mosaïque*, PHD thesis, ENST Bretagne, Brest, June 2007.
2. D. Gueriot, E. Maillard, and J.-P. Kernin, "Sonar image registration through symbolic matching: A fuzzy local transform approach using genetic algorithms," in Proc. *MTS/IEEE OCEANS Conf.*, Sep. 1996, vol. 3, pp. 1324–1329.
3. S. Daniel, F. Le Leannec, C. Roux, B. Solaiman, and E. Maillard, "Side-scan sonar image matching," *IEEE J. Ocean. Eng.*, vol. 23, no. 3, pp. 245–259, Jul. 1998.
4. P.Y. Mignotte, M. Lianantonakis, and Y. Petillot, "Unsupervised registration of textured images: applications to sidescan sonar", *Oceans 2005 – Europe*, 2005.
5. I. Leblond, *Recalage à long terme d'images sonar par mise en correspondance de cartes de classification automatique des fonds*, PHD thesis, Université de Bretagne Occidentale, ENSIETA, Brest, February 2006.
6. S. Daniel and M. Bach, *Automatic Image Co-registration Study for Sidescan Sonar*, Contract Report, DRDC Atlantic CR 2010-022, January 2010.
7. C. Chailloux and B. Zerr, « Non Symbolic Methods to register SONAR images », *Oceans 2005 – Europe*, 2005.
8. W.D. Fisher, "On grouping for maximum homogeneity", *American Statistical Association Journal*, vol. 53, pp789-798, 1958.
9. S. Kullback, *Information theory and statistics*, John Wiley and Sons, NY, 1959.
10. A. Singh, "Digital change detection techniques using remotely sensed data," *Int. J. Remote Sensing*, vol. 10, pp. 989-1003, 1989.
11. E. Rignot and J. Van Zyl, "Change Detection Techniques for ERS-1 SAR Data", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, No. 4, July 1993.
12. M. Preiss and N. Stacy, *Coherent Change Detection: Theoretical Description and Experimental Results*, technical report DSTO-TR-1851, August, 2006.
13. J. S. Lee, "Statistical modeling and suppression of speckle in synthetic aperture radar images", in Proc. of *IGARSS'87*, pages 1331-1339, Ann Arbor, 18-21 May 1987.
14. F. Schmitt, L. Bonnaud and C. Collet, "Contrast control for sonar pictures", *Signal and Image Processing, SPIE'96 – Technical Conference on Application of Digital Image Processing XIX*, vol. 2847, pp. 70-82, Aug. 1996.
15. V. Myers, *A data set for change detection in port environments using sidescan sonar*, Technical memorandum, DRDC Atlantic TM 2008-247, 2009.

Annexe 4 : Détection cohérente de changement

Cette annexe est une copie de l'acte [QUI12b] de la conférence *Detection and Classification of Underwater Targets (DCUT) 2012* co-écrit par les collègues canadiens et norvégiens sur le recalage élastique et la détection cohérente de changement appliqués aux données norvégiennes. C'est une synthèse des travaux réalisés sur ce sujet dans le cadre de l'arrangement spécifique n°42 franco-canadien et d'un accord ENSTA Bretagne – FFI de collaboration autour des données de la campagne norvégienne d'avril 2011. Cette campagne a eu lieu dans le cadre d'un consortium entre le Norwegian Defence Research Establishment (FFI), Defence Research and Development Canada (DRDC) et Naval Research Laboratory (NRL), US mais un accord franco-norvégien a ensuite été signé pour permettre une exploitation commune des données de cette campagne.

Subpixel image registration for coherent change detection between two high resolution sonar passes

I. Quidu⁽¹⁾, V. Myers⁽²⁾, Ø. Midtgaard⁽³⁾ and R. E. Hansen⁽³⁾

(1) ENSTA Bretagne, Ocean Sensing and Mapping Team, Lab-STICC UMR CNRS 6285, 2 rue François Verny, 29806 Brest Cedex 9, France

Email: isabelle.quidu@ensta-bretagne.fr

(2) Defence R&D Canada, 9 Grove St, Dartmouth, Nova Scotia, B2Y 3Z7, Canada

Email: vincent.myers@drdc-rddc.gc.ca

(3) Norwegian Defence Research Establishment (FFI), P.O. Box 25, N-2027 Kjeller, Norway

Email: [\[Oivind.Midtgaard,Roy-Edgar.Hansen\]@ffi.no](mailto:{Oivind.Midtgaard,Roy-Edgar.Hansen}@ffi.no)

ABSTRACT: *The use of phase coherence between two pings is used in many underwater imaging applications, in particular the micronavigation algorithms used in correcting receiver positions when focussing a synthetic aperture sonar (SAS) array. Recently, interest has turned to using the phase coherence between two sonar images that consist of two repeated passes over the same area in order to co-register the images and perform coherent change detection, meaning the detection of differences in two areas that causes a decorrelation between the two images. Such techniques are used in synthetic aperture radar (SAR) applications, and can detect subtle changes that are not immediately detectable by visual inspection or other traditional methods. Before computing the coherence the two images have to be accurately co-registered. Indeed remaining misregistrations after a 2D rigid translation lead to a loss of phase coherence. This can be due to navigational disturbances or data processing defects. The problem of non-rigid and subpixel sonar image registration is achieved here in two main steps. The first step aims to estimate the displacement field from which a warping function is derived; then, the repeat-pass image is resampled using the warping function to co-register it to primary image.*

Keywords: *elastic registration, change detection, coherence, SAS image*

1 INTRODUCTION

Change detection between two repeated sonar passes over the same area is a well-known tactic for detecting targets on the seafloor, and is the most reliable method for detecting objects of an improvised nature meaning that little or no prior knowledge of size and shape is available or that it has been altered or disguised in some way. Two modes of change detection can be considered: incoherent and coherent change detection [1]. Incoherent change detection methods aim to identify changes in the mean backscatter power that can be estimated by the pixel intensity. This method has been widely used in SAR imagery [2,3] and recently applied to high-frequency sonar imagery [4,5]. Coherent change detection, on the other hand, identifies changes in both the amplitude and phase of the image data. By using the information that is provided by the phase, it is possible to detect very subtle changes, even those not visually detectable. However, preserving coherence is made difficult by the natural temporal decorrelation and by the stringent navigational accuracy requirements for minimizing

differences in the imaging geometry between the primary and repeat pass [1,6]. In this paper, these issues are examined using a data set collected in April 2011 by the Norwegian Defence Research Establishment (FFI) in an area outside of Larvik, Norway with the HUGIN 1000-HUS AUV equipped with a HISAS 1030 interferometric SAS, deployed from the research vessel HU Sverdrup II [4,7]. Coherent change detection also requires that the primary and repeat pass images are registered to sub-resolution accuracy, typically a tenth of a resolution cell [1]. This requirement, together with the problem of local misregistration, is addressed in two steps: A large scale translation is first applied to the data, which is presented in Section 2, along with a method for estimating the subpixel displacement field; then, two warping procedures are proposed in Section 3 for performing an elastic co-registration. Section 4 shows experimental results on the HiSAS data set, which contains different objects on different types of seabeds.

2 DISPLACEMENT FIELD ESTIMATION

Registration is the process of aligning two images: the primary image I_p and the repeat pass image I_{rp} . Finding the spatial transformation to apply to I_{rp} is generally the key to any registration problem. Sometimes a rigid transformation is not sufficient and local misregistrations remain. In SAS images, it is usually due to a change in the sensor position or viewpoint. In that case, an elastic registration following a rigid one must be performed. The displacement field which gives a feature-to-feature correspondence between the primary image and the repeat pass image is estimated by finding matching features extracted in both images. Many matching techniques are available and recent techniques have focussed on extracting local invariant features such as the Scale Invariant Feature Transform (SIFT) method and Speeded Up Robust Features (SURF) [8,9]. However these techniques do not use the phase of the image, which is a key component when using coherence to detect changes. Moreover the sonar image formation process leads to a different sampling of a given scene at different times. This is the source of the motivation for evaluating subpixel correlation-based matching methods. Two methods are presented here. For both, in order to speed up the cross-correlation computation over the area to register, the repeat pass $M \times N$ image is divided into a number of non-overlapping chips $C_{rp} = I_{rp}(x, y, m, n)$ of size $m \times n$ centred at the image indices x, y where $m \ll M$ and $n \ll N$.

2.1 Coarse shift and re-navigation

The first step in the displacement field estimation is a compensation for global translation, as described in [4, 7] and shown in Figure 1.

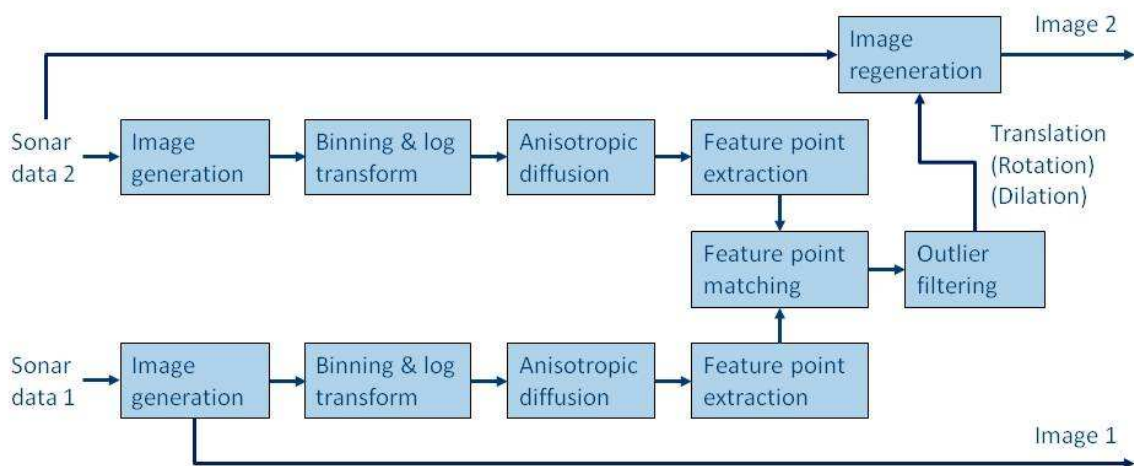


Figure 1 - Data flow of the coarse shift image removal. After the features are extracted and matched, the SAS imaging process is repeated with the corrected navigation solution.

Corresponding SAS images are produced from the primary and repeat pass sonar data, using the vehicle navigation solutions. As preprocessing for feature point extraction, these images are down sampled by a factor of two in both dimensions using the mean operator, and the resulting magnitude values are transformed onto a logarithmic scale. Anisotropic diffusion is applied to reduce the image speckle, thus creating more uniform magnitude regions while preserving region edges. Feature points are then extracted from both images by the SURF algorithm and matched using a similarity score. The SURF feature points typically correspond to image

high-light and shadow “blobs” of various sizes. An outlier filtering stage has been included, as false matches sometimes occur. The filter only accepts matches that are part of the main match cluster in the parameter space. The parameters for a rigid spatial transformation (translation, rotation and dilation) of the images can then be found from the set of resulting point pairs.

The estimated translation between the two images are then applied to the navigation solution of the repeat pass data to regenerate that image (Note that even though the method also estimates rotation and dilation, only the translation was used in this study). The method managed to bring the global relative position error down from around five meters to a few centimeters.

2.2 Subpixel complex correlation

As proposed in [10], the cross-correlation of the two small chips $C_p = I_p(x, y, m, n)$ and $C_{rp} = I_{rp}(x, y, m, n)$ of the same size is achieved into two steps. It starts with an initial estimate of the location of the cross-correlation peak obtained to within a fraction $\frac{1}{2}$ of a pixel by the usual fast Fourier transform (FFT) approach. This approach consists in (X^* represents the complex conjugate of X): (i) computing the fast Fourier transforms of each chips $FC_p = FFT(C_p)$ and $FC_{rp} = FFT(C_{rp})$; (ii) embedding the element-wise product $FC_p \times FC_{rp}^*$ in a larger array of zeros of dimension $(2m \times 2n)$; (iii) computing an inverse FFT to obtain an upsampled (by a factor of 2) cross correlation; and (iv) locating its peak. This peak is a starting point, the initial estimate, around which a new upsampled cross-correlation (by a factor k) is computed in a 1.5×1.5 pixel neighbourhood. This results from the following property:

$$IFFT(FC_p \times FC_{rp}^*) = FFT^*(FC_p^* \times FC_{rp}) = FFT^*(h) = H^*,$$

where H can be computed using a Matrix Fourier Transform such that [11] (X^T represents X transpose):

$$H(U, V) = e^{-2\pi jUX^T} \cdot h(X, Y) \cdot e^{-2\pi jYV^T}$$

$$\text{where } \begin{cases} U = V = \frac{[0 \dots \text{ceil}(1.5k) - 1]^T}{k} \text{ around the initial peak} \\ X = \frac{[0 \dots m - 1]^T}{m} - \frac{1}{2}, \quad Y = \frac{[0 \dots n - 1]^T}{n} - \frac{1}{2} \end{cases}$$

The output H is a $1.5k \times 1.5k$ array in units of upsampled pixels. The searched subpixel displacement $(\Delta x, \Delta y)$ is given by the location of the output peak in the complex modulus $|H|$, then in the phase angle $\arg(H^*)$, if necessary.

Possible outliers or erroneous displacement estimates can simply be suppressed by thresholding the value of the output peak. From the remaining estimates two sets of n_{cp} corresponding feature points $\mathbf{x}_p = [x_k], \mathbf{y}_p = [y_k]$ on the primary image and $\mathbf{x}_s = [u'_k], \mathbf{y}_s = [z'_k]$ on the repeat pass image with $k = [1, \dots, n_{cp}]$ can be derived where: $u'_k = x_k + \Delta x_k$ and $z'_k = y_k + \Delta y_k$.

2.3 Amplitude correlation with parabolic interpolation

Unlike the previous one, this method takes as inputs the amplitude of the sonar images. The correlation procedure works as follows: a normalized cross-correlation is performed in a chip, centred at the same image indices x, y with an area in the primary image:

$$\gamma(x, y) = \frac{1}{(mn) - 1} \sum \frac{\bar{C}_p \bar{C}_{rp}}{\sigma_p \sigma_{rp}},$$

where $\bar{C}_p = C_p - \mu_p$ is the mean subtracted primary chip and σ_p is the standard deviation of \bar{C}_p . The index $\gamma(x, y)$ of the maximum correlation over the central part (that is, where the cross-correlation is valid) of the

primary chip is found. Then, a parabolic interpolation is used to estimate the sub-pixel position of the peak independently in the u and z dimensions [1, 12]:

$$u' = u + 0.5 \frac{(u-v)^2 |\gamma(u, z) - \gamma(w, z)| - (u-w)^2 |\gamma(u, z) - \gamma(v, z)|}{(u-v) |\gamma(u, z) - \gamma(w, z)| - (u-w) |\gamma(u, z) - \gamma(v, z)|}$$

where $v = u - 1$ and $w = u + 1$. An analogous method is used to find z' . The maxima $\gamma(u, z)$ can also be filtered using a threshold τ . What remains are two sets of n_{cp} corresponding feature points $\mathbf{x}_p = [x_k], \mathbf{y}_p = [y_k]$ on the primary image and $\mathbf{x}_s = [u'_k], \mathbf{y}_s = [z'_k]$ on the repeat pass image, with $k = [1, \dots, n_{cp}]$.

3 WARPING PROCEDURE

From the displacement field a warping procedure has to be derived. The warping function specifies the pixel-to-pixel correspondence, namely the spatial mapping, between the primary image and the repeat pass image. Registration is performed by resampling the repeat pass image onto the primary image by applying the spatial mapping along with an interpolation technique. The spatial mapping is frequently expressed parametrically as two approximation functions f_x and f_y so as to compute the transformed image I_{rpw} [13]:

$$I_{rpw}(f_x(x, y), f_y(x, y)) = I_{rp}(x, y)$$

The two approximation functions f_x and f_y must be computed from the estimated displacement field. On the one hand, the displacement field is a set of n feature points $(x_i, y_i)_{i=1 \dots n}$ in the repeat pass image to be displaced by some value $(dx_i, dy_i)_{i=1 \dots n}$. On the other hand, the warp function relates the (x, y) points in the original image to their counterparts $(x', y') = (f_x(x, y), f_y(x, y))$ in the warped image. The warping procedure tries to estimate two smooth surfaces from two sets of scattered data: X that passes through $(x_i, y_i, dx_i)_{i=1 \dots n}$ and Y that passes through $(x_i, y_i, dy_i)_{i=1 \dots n}$. The two approximation functions provide a displacement to every pixel of the repeat pass image:

$$f_x(x, y) = (x, y) + X(x, y),$$

and

$$f_y(x, y) = (x, y) + Y(x, y).$$

After the feature matching procedure from Section 2, the set $(dx_k, dy_k)_{k=1 \dots n_{cp}} = (x_k - u'_k, y_k - z'_k)_{k=1 \dots n_{cp}}$ can be used to estimate f_x and f_y .

3.1 Multilevel B-spline approximation

Cubic B-splines have many properties, such as local control, smoothness and computational efficiency that make them widely used in surface modeling [14]. Moreover, in order to deal irregularly spaced feature points, multilevel B-splines, proposed by Lee *et al* for generating a smooth surface by interpolating the scattered data points [15], seemed to be promising. This approach was chosen over thin-plate splines, introduced by Bookstein [16], as they have two major drawbacks in the present application: firstly, the estimation requires the inversion of a matrix made of data point coordinates and can be computationally expensive; secondly it is based on radial basis functions that cannot adapt themselves to the data point density.

The basic idea is to estimate the approximation function f as a uniform bicubic B-spline function by minimizing (in the least-squared sense) the error between this function and the feature points. The function f is defined in terms of sixteen control points ϕ_{ij} weighted by a product of uniform cubic B-spline basis functions $B_k, k = 0 \dots 3$:

$$f(x, y) = \sum_{k=0}^3 \sum_{l=0}^3 B_k(s) B_l(t) \phi_{i+k, j+l}, i = \lfloor x \rfloor - 1, j = \lfloor y \rfloor - 1, s = x - \lfloor x \rfloor, t = y - \lfloor y \rfloor,$$

with

$$B_0(t) = \frac{(1-t)^3}{6}, B_1(t) = \frac{(3t^3 - 6t^2 + 4)}{6}, B_2(t) = \frac{(-3t^3 + 3t^2 + 3t + 1)}{6}, B_3(t) = \frac{t^3}{6}, 0 \leq t < 1.$$

Let Φ be the control lattice that spans the integer grid in a rectangular domain Ω of the xy-plane and the set of scattered points $P = \{(x_c, y_c, z_c)\}$, where (x_c, y_c) is a point of Ω as shown in Figure 2. The objective is to determine the control points of Φ that provide the value of the approximation function f at every (x_c, y_c) the closest of z_c . This is found by solving the minimization problem ($\lfloor x \rfloor$ represents the floor function):

$$\min_{\phi_{ij}} \sum_c (z_c - f(x_c, y_c))^2 + \sum_{ij} \phi_{ij}^2,$$

where $z_c = \sum_{k=0}^3 \sum_{l=0}^3 B_k(s) B_l(t) \phi_{i+k, j+l}^c$, $i = \lfloor x \rfloor - 1$, $j = \lfloor y \rfloor - 1$, $s = x - \lfloor x \rfloor$, $t = y - \lfloor y \rfloor$ and \sum_c is the sum of the feature points (x_c, y_c, z_c) that influence the value of ϕ_{ij} .

The solution is given by:

$$\phi_{i+k, j+l} = \frac{\sum_c w_{k,l}^2 \phi_{k,l}^c}{\sum_c w_{k,l}^2}$$

where $\phi_{k,l}^c$ is the contribution of the control point $\phi_{i+k, j+l}$ to the feature points (x_c, y_c, z_c) that are within a 4×4 neighborhood. $\phi_{k,l}^c$ is found by using pseudoinverse method:

$$\phi_{kl}^c = \frac{w_{kl} z_c}{\sum_{a=0}^3 \sum_{b=0}^3 w_{ab}^2}$$

with $w_{kl} = B_k(s) B_l(t)$, $k, l = 0 \dots 3$, and $k = i + 1 - \lfloor x_c \rfloor$, $l = j + 1 - \lfloor y_c \rfloor$, $s = x_c - \lfloor x_c \rfloor$, $t = y_c - \lfloor y_c \rfloor$

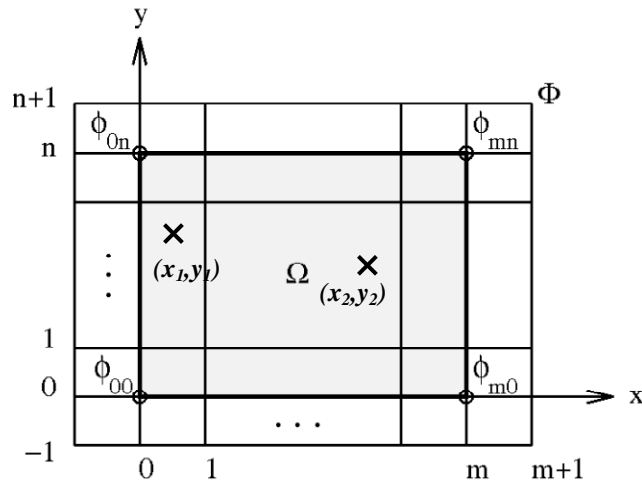


Figure 2 – Control lattice Φ overlaid on domain Ω [15] and two features points (x_c, y_c)

This B-spline approximation is parametrized by the size of the control lattice. A coarse lattice leads to a smooth but inaccurate approximation. Conversely with a fine lattice, local peaks near feature points appear. In order to circumvent this tradeoff between smoothness and accuracy a multi-level B-spline approximation has been proposed by Lee *et al.* [15]. The method is summarized in Algorithm 1, below. The refinement consists in deriving from Ψ a larger control lattice Ψ' so that they both generate the same B-spline functions. It is less

computationally expensive than keeping the approximation function of all levels and adding them at the end.

Algorithm 1: Multi-level B-spline pseudo-code.

$h \leftarrow$ the number of levels
 $P \leftarrow \{(x_c, y_c, z_c)\}$ \Rightarrow The set of scattered points
 $f_0 \leftarrow$ result of applying B-spline approximation to P with coarsest control lattice Φ_0 .
for each feature point
 $\Delta^1 z_c \leftarrow z_c - f_0(x_c, y_c)$
endfor
 $\Psi' \leftarrow 0$ \Rightarrow Initialize the control lattice
for $k = 1$ to h
 Compute Φ_k using $P_k = \{(x_c, y_c, \Delta^k z_c)\}$ $\Rightarrow \Phi_k$ is twice as fine as Φ_{k-1}
for each feature point
 $\Delta^{k+1} z_c \leftarrow \Delta^k z_c - f_k(x_c, y_c)$
end for
 $\Psi = \Psi' + \Phi_k$ \Rightarrow Update the control lattice
 $\Psi' \leftarrow \Psi$, where the control point spacing is half as large as Ψ \Rightarrow Refinement
endfor

3.2 Polynomial warp

Another method for computing the warping function is to fit a polynomial of the points (x_p, y_p) and (x_s, y_s) .

For a 2D polynomial fit of order p , the Vandermonde matrix V must be constructed, where:

$$V = \begin{bmatrix} 1 & u'_1 & z'_1 & u'_1 z'_1 & u_1'^2 & \dots & z_1'^p \\ 1 & u'_2 & z'_2 & u'_2 z'_2 & u_2'^2 & \dots & z_2'^p \\ 1 & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u'_{n_{cp}} & z'_{n_{cp}} & u'_{n_{cp}} z'_{n_{cp}} & u_{n_{cp}}'^2 & \dots & z_{n_{cp}}'^p \end{bmatrix}$$

which is an $n_{cp} \times (p+1)(p+2)/2$ matrix. The coefficients of the least-square fitted polynomial can then be computed as

$$a = V^+ \mathbf{x}_p,$$

and

$$b = V^+ \mathbf{y}_p,$$

where V^+ is the Moore-Penrose pseudo-inverse of V . The polynomial defined by the coefficients in a and b can be used to compute the required warping to interpolate and co-register the repeat-pass image with the primary image.

3.3 Interpolation

The previous warping function gives a shift for every pixel of the image to be co-registered. This warped grid must be interpolated between points of the mesh grid that overlay the expected registered image. Here, a 2D cubic data interpolation is performed, which avoids phase jumps that can happen by a nearest neighbor interpolation.

4 RESULTS

4.1 Data set

The data set used in this experiment was gathered in April 2011 in an area near Larvik, Norway using the by the HUGIN 1000-HUS Autonomous Underwater Vehicle, equipped with the HISAS 1030 Synthetic Aperture Sonar, which operates in the 60-120 kHz frequency range with up to 50 kHz in bandwidth.

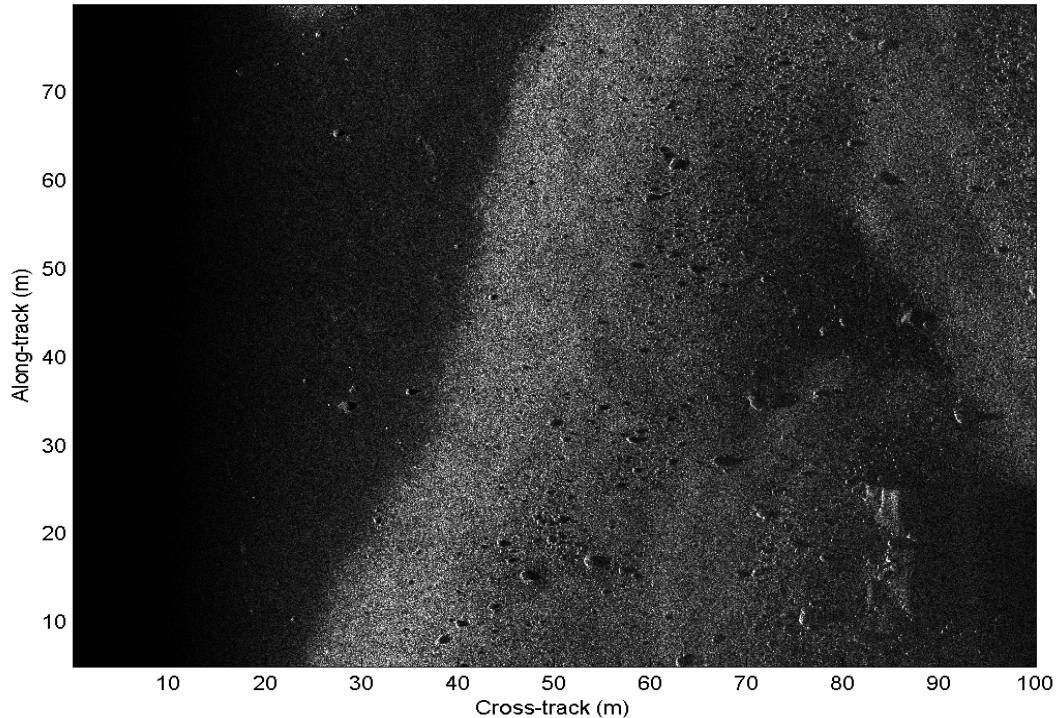


Figure 3 – Primary SAS image of Area 2. The image size is 100m (range) by 75m (azimuth).

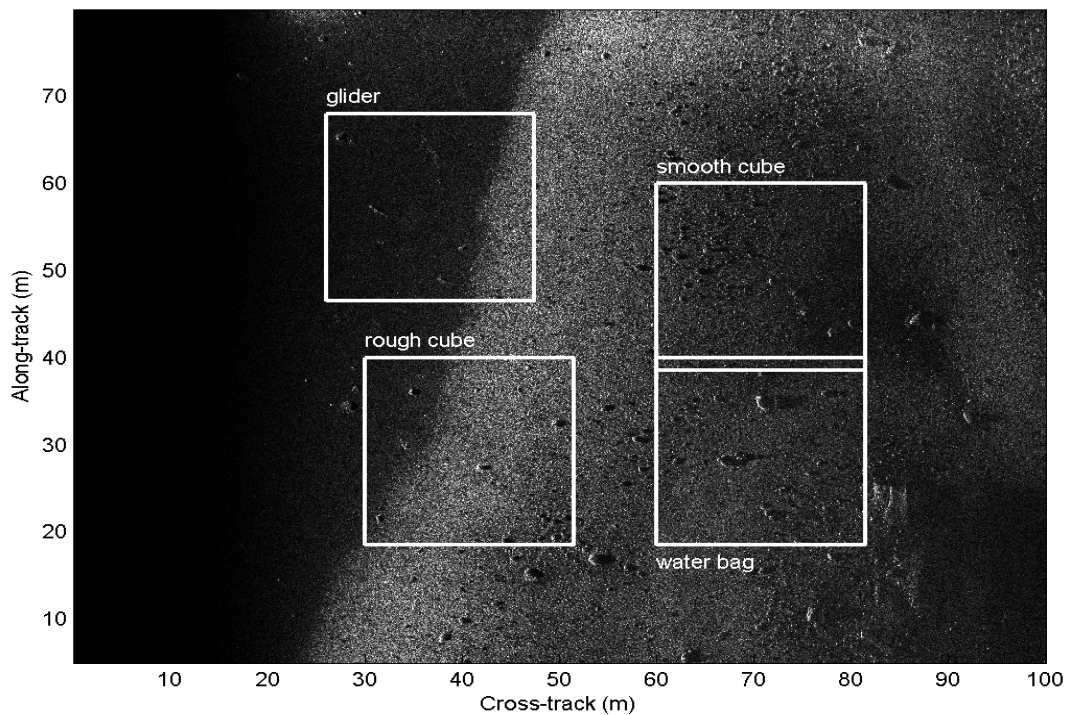


Figure 4 – Repeat pass SAS image of Area 2. The image size is 100m (range) by 75m (azimuth). The four studied regions are bounded by a rectangular box with the named object to be detected.

The resolution of the beamformed image is better than 5×5 cm and the pixel size is 2×2 cm. An area of high clutter was chosen specifically for these experiments: the area was first surveyed without any objects on April 8th, after which four targets were deployed and the area resurveyed on April 10th. The objects were: 2 concrete cubes of $40 \times 40 \times 40$ cm (one rough, one smooth), an underwater glider, and a large water-filled “bag”. Figure 3 and Figure 4 show the primary and the repeat pass areas respectively. The vehicle track is along the y-axis, and sonar range is along the x-axis. On the repeat pass area, four boxes surrounding the four deployed objects give the specific regions on which the proposed algorithms are assessed.

4.2 Displacement field estimation

Cross-correlation is performed using a squared chip of $m=51$ by $n=51$ pixels. The primary and repeat pass images are $M=1024$ by $N=1024$ pixels in size. Local displacements estimated by both methods of section 2 (see 2.2 and 2.3) were fused so as to provide a better estimate of the field. The cross-correlation of complex-valued chips ensures that the phase component also contribute to the alignment process, however where coherence is too weak, displacement can be estimated by correlating the amplitude. For complex subpixel correlation, the upsampling factor k is set to 10 and only displacements for which the peak is over the maximum less 0.2 are kept. For amplitude correlation, the threshold is fixed to 0.5.

Figure 5 shows the result of the displacement field estimation for the region surrounding the water bag. Extracted feature points have been plotted on the primary (left) and the repeat pass (right) images. At the bottom left of the figure, the estimated displacements can be seen. The amount of displacement has been exaggerated in order to be visible. Blue arrows are displacements estimated by both methods but for which the amplitude correlation has been kept. Magenta arrows are displacements estimated by the complex cross-correlation only. And the red arrows show displacements estimated by the amplitude cross-correlation only. As expected, these last are found even if the coherence is very weak (blue areas of the image at the bottom right of the figure).

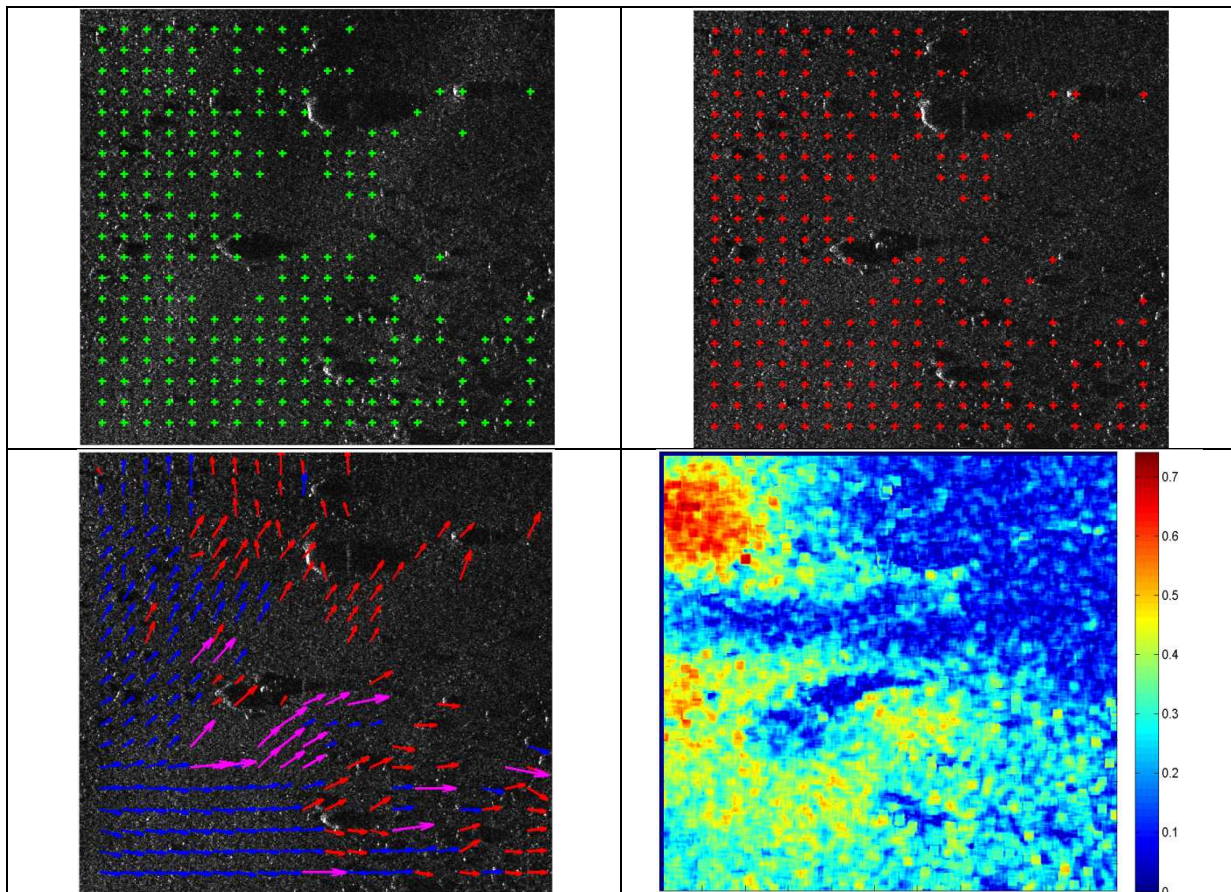


Figure 5 – Results for the “water bag” region. In the first row, feature points on the primary (on left) and the repeat pass (on right) images are plotted. In the second row, on the left the estimated displacements are shown in arrows (length exaggerated to be visible); the coherence between the two images is shown on right

4.3 Warping and interpolation

In the multilevel B-spline approximation algorithm, the size of the finest control lattice is determined by the size of the image to register. In other words, the last level h is reached when the control lattice size cannot be refined without oversizing the input image.

Note that in order to avoid aliasing problem while interpolating, the image phase was modified such that the fast fluctuation of the phase in the range direction was corrected by subtracting the fluctuation due to the carrier frequency. This gives a slower fluctuation due to the along-track sonar resolution.

Figure 6 shows the two surfaces (one for estimated displacements along the x -axis, one for estimated displacements along the y -axis) that are estimated by the two warping methods given the feature points extracted in the “water bag” region that were shown in Figure 5.

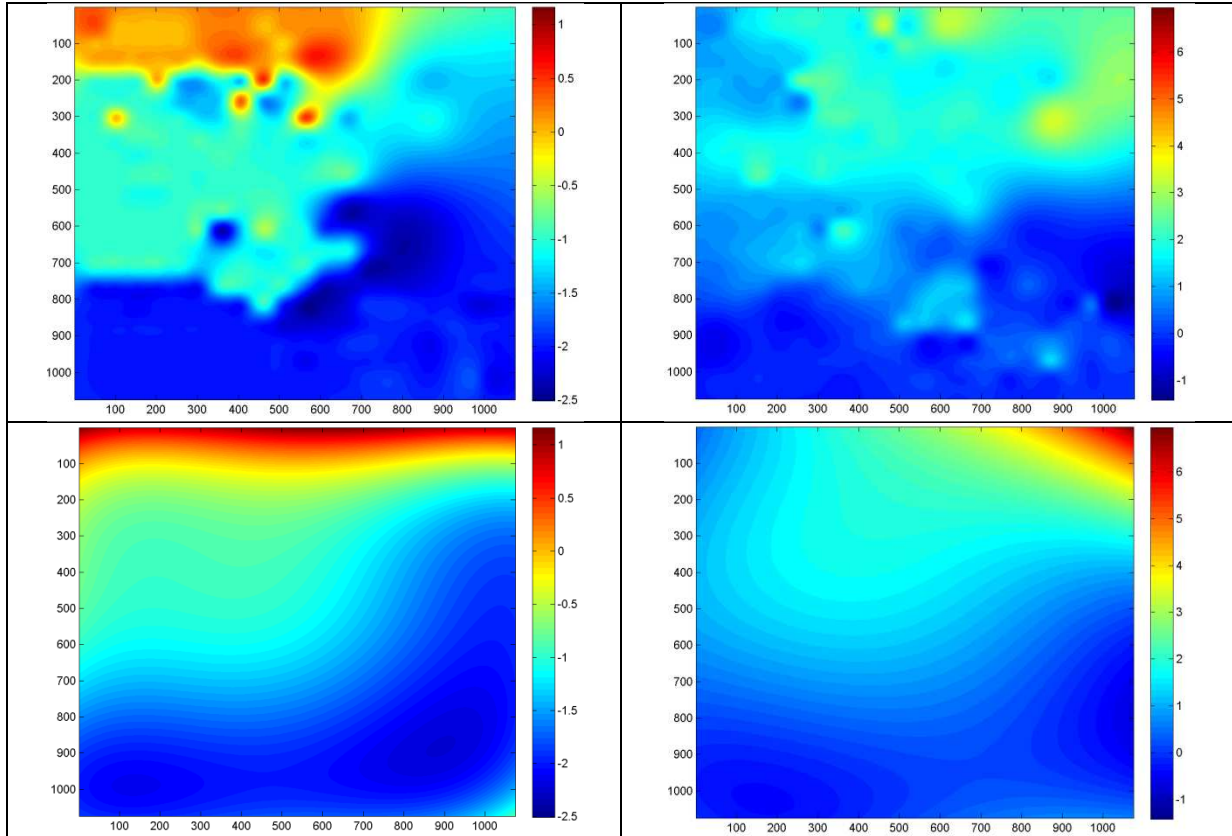


Figure 6 - Estimated displacements along the x - and y -axes for the “water bag” region. Above is the field given by the multilevel B-spline warping and below is that computed by polynomial warping.

In figures 7, 8, 9 and 10, the two warping methods are compared in terms of phase registration, amplitude registration and coherence. The first column gives results after only the rigid (coarse) registration. The second column shows results after elastic registration based on polynomial warp. The third column shows results with elastic registration based on the multilevel B-spline approximation.

For the phase registration, the phase difference (pixel by pixel) is shown.

For the amplitude registration, the log-ratio of averaged images (by an average filter of size 20 by 20 pixels) before and after elastic registration [5] are compared. The change-detected object(s) should be visible in the log-ratio image.

Let us recall the definition of the coherence used to detect changes between the two sonar passes. First, the cross-correlation between two time series $s_1(t)$ and $s_2(t)$ is defined as:

$$\rho(s_1, s_2, \tau) = \int_{-\infty}^{+\infty} s_1^*(t) s_2(t + \tau) dt,$$

where $s^*(t)$ represents the complex conjugate of $s(t)$. The complex coherence function between $s_1(t)$ and $s_2(t)$ is then defined as:

$$\Gamma(s_1, s_2, \tau) = \frac{\rho(s_1, s_2, \tau)}{\sqrt{\int_{-\infty}^{+\infty} |s_1(t)|^2 dt} \sqrt{\int_{-\infty}^{+\infty} |s_2(t)|^2 dt}}$$

The absolute value of $\Gamma(s_1, s_2, \tau)$ is the maximum likelihood estimator for coherence. It is here computed using a sliding window of 21×21 pixels.

The ‘water-bag’ case (Figure 7) shows the potential of using change detection to detect objects without *a priori* knowledge of size and shape. The water bag target is quite stealthy by nature and had been deployed very near to a rock, which is a very difficult scenario in real-world applications. In spite of these difficulties, it can be detected here by taking the log-ratio of the two images. In this case, the elastic registration step did not provide a noticeably better log-ratio image.

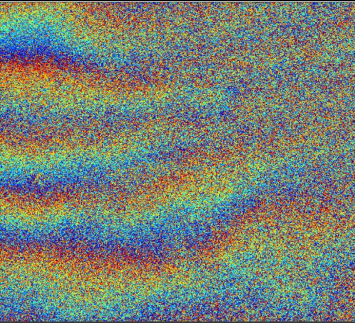
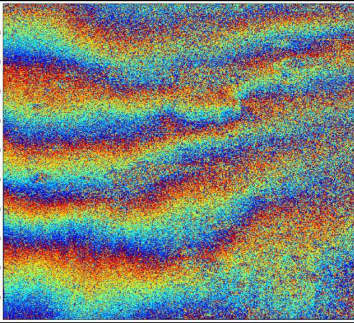
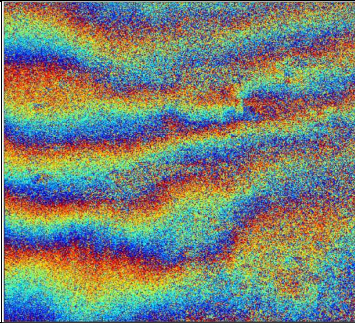
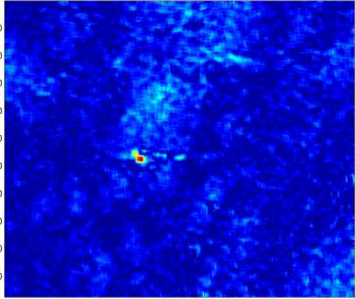
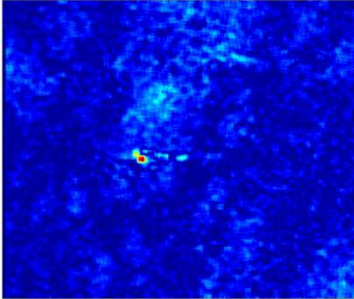
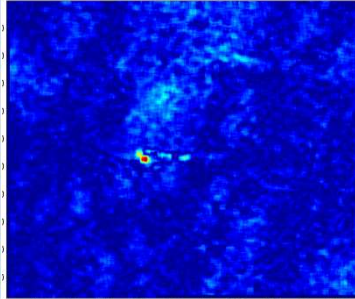
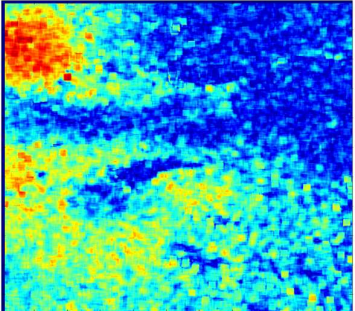
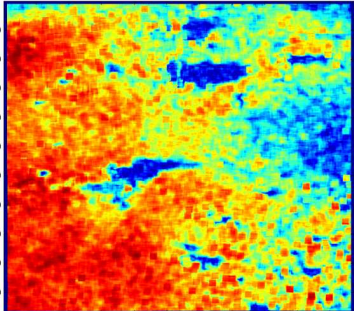
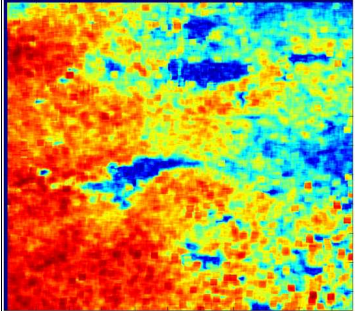
	After rigid registration	After elastic registration / polynomial warp	After elastic registration / multilevel B-spline approximation
Phase difference			
Log-ratio image			
Coherence			

Figure 7 –“Water bag” region: registration results for coarse registration, polynomial warping and B-spline approximation. The resulting phase difference, log-ratio and coherence for the three methods are shown.

Figure 8 shows the case of the area surrounding the smooth cube. In this case, the elastic registration provides a significant boost in coherence that is required to detect it by coherent change detection. The log-ratio image is also improved by the elastic registration step; however the target is easily detectable in all three cases.

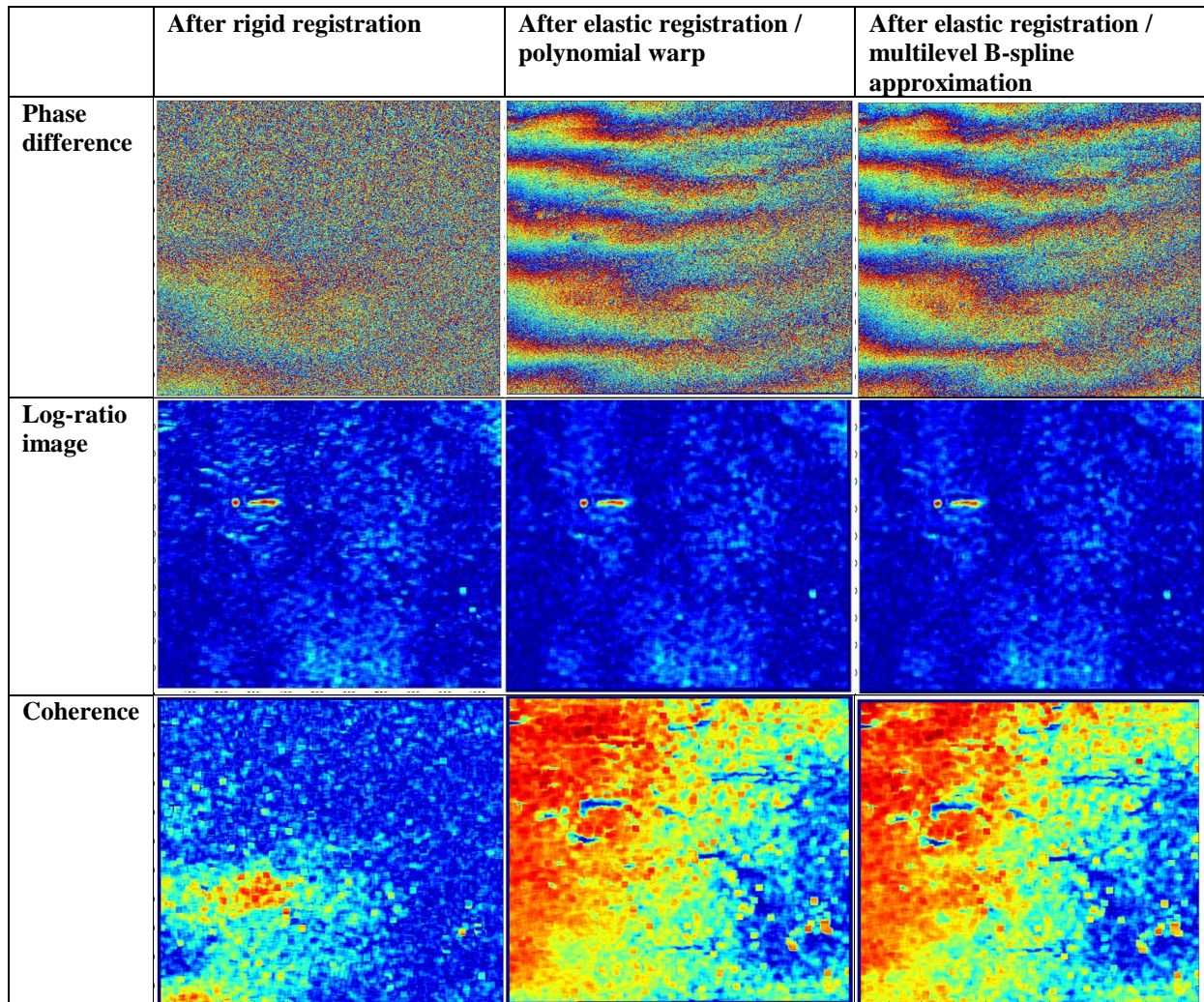


Figure 8 – “Smooth cube” region: registration results for coarse registration, polynomial warping and B-spline approximation. The resulting phase difference, log-ratio and coherence for the three methods are shown.

In Figure 9, more dramatic improvements in coherence can be observed in the region surrounding the “rough cube”. As was the case for the smooth cube, the rough cube that was invisible in the coherence image after the rigid registration appear clearly with a very weak coherence surrounded by high coherence after the elastic registration. Moreover the remaining false alarms after the rigid registration disappear after the elastic registration. Looking at the phase difference image, the phase rotation is well retrieved, or denoised, over the whole region. Note that the vertical stripes in the coherence and phase difference are probably due to an incorrect wavenumber shift in the image formation and can be corrected.

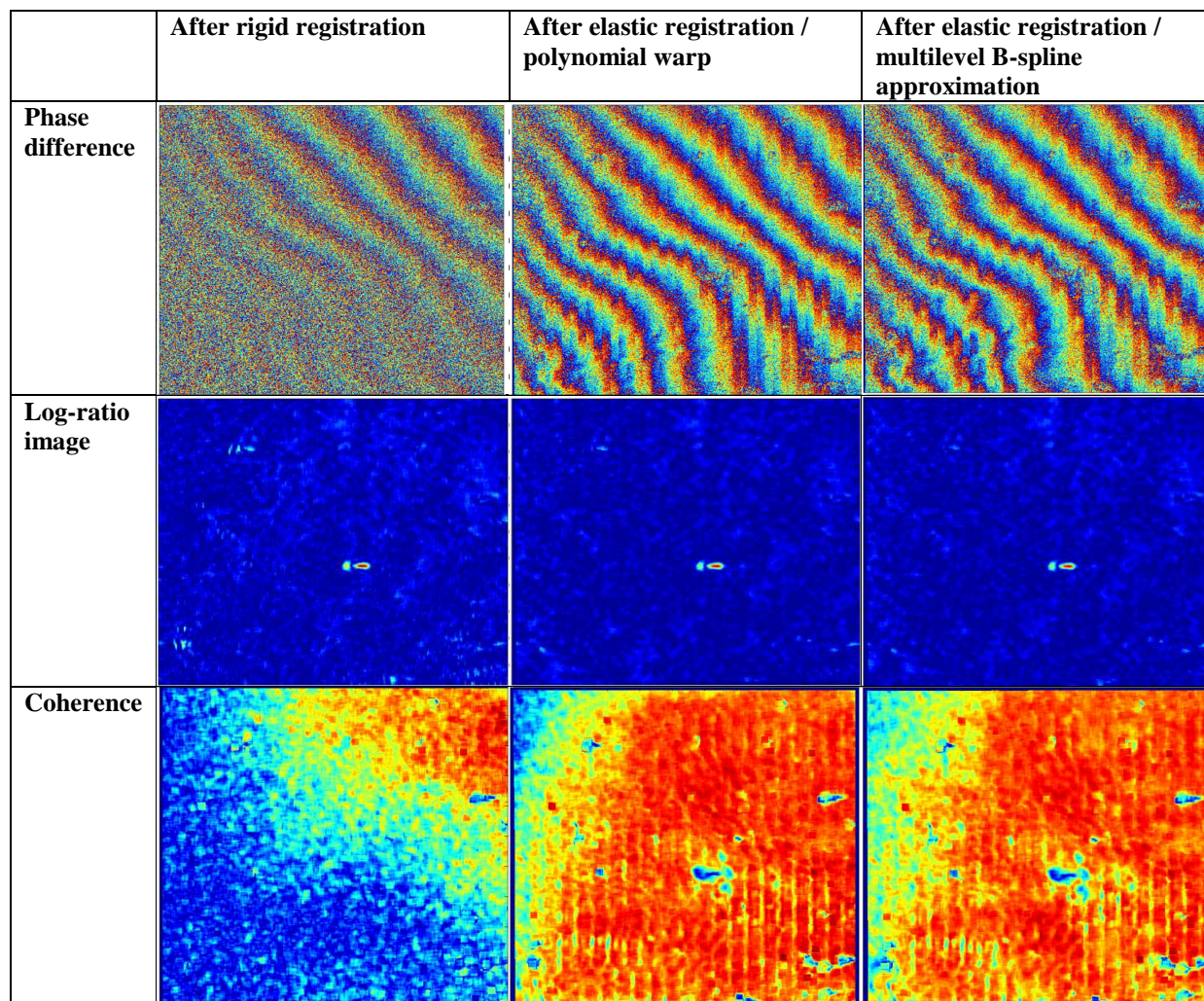


Figure 9 –“Rough cube” region: registration results for coarse registration, polynomial warping and B-spline approximation. The resulting phase difference, log-ratio and coherence for the three methods are shown.

Finally, Figure 10 shows a significant amount of coherence gain the “glider” region, however the target, along with a smaller object just below it (which was an anchor used for keeping the glider in place) are in an area where coherence was not maintained, resulting in not a great deal of benefit for coherent change detection. Moreover, we can notice here a border defect of the polynomial warp that appears in the left part of the image leading to a loss of coherence conjugated with a noisy phase difference. The objects are visible in the log-ratio image.

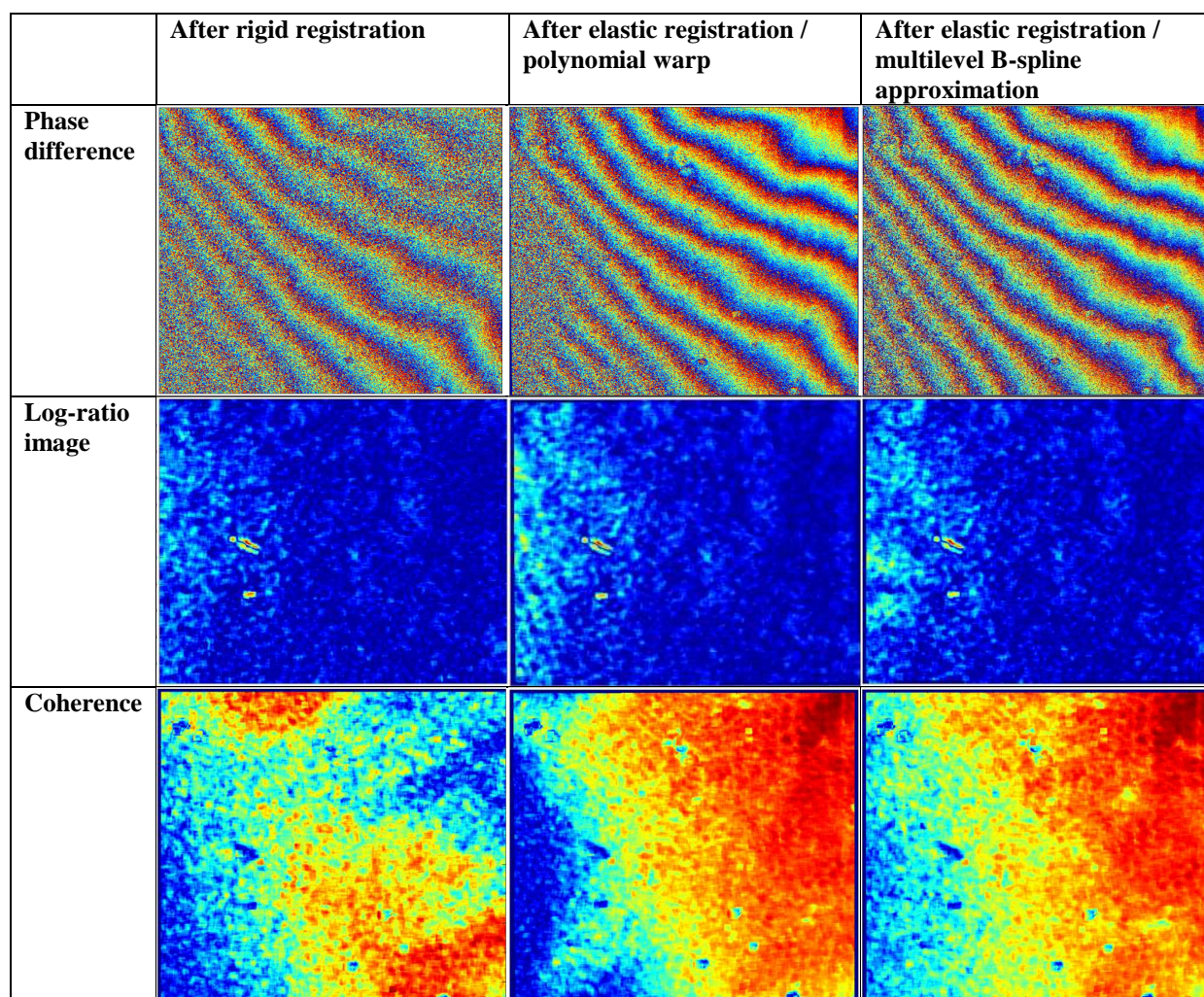


Figure 10 –“Glider region”: registration results for coarse registration, polynomial warping and B-spline approximation. The resulting phase difference, log-ratio and coherence for the three methods are shown.

5 CONCLUSION AND FUTURE WORK

This paper examined a number of methods for feature matching, estimating the displacement field and warping to coregister two synthetic aperture sonar images over the same area. The methods were tested on image gathered by the HISAS-equipped HUGIN AUV during experiments at sea in an area near Larvik, Norway, in 2011. In all the cases, the elastic registration provided an increase in coherence between the two images, and in a few cases were need to coherently detect the deployed objects. However, it was also possible to detect the objects using purely amplitude-based methods. This is to be expected, since the true strength of coherent change detection is to be able to identify very small changes those not immediately perceptible by visual inspection, and in the case of the targets deployed in the Larvik trial, a combination of object size and good sensor resolution results in the lack of coherence not being strictly necessary to detect the targets. However, more importantly, it was found to be in fact possible to maintain some amount phase coherence over the seafloor after two days. This suggests that a study on the coherence would be of significant value. In fact, the data collected during the Larvik trial could be used for this purpose as sediment samples were taken and analyzed, and areas were surveyed in such a way as to provide time intervals of one day to over five days between surveys. The temporal coherence of the seabed determines the practical limitations of using coherent change detection (or any other kind of multi-pass coherent processing) and thus needs to be better understood. Also, the phase coherence allows one to envision the use of multiple passes over an area to create very high precision bathymetric maps using interferometry from two passes. Both are suggested as avenues for future research.

6 REFERENCES

- [1] M. Preiss and N. Stacy, *Coherent Change Detection: Theoretical Description and Experimental Results*, technical report DSTO-TR-1851, August, 2006.
- [2] A. Singh, "Digital change detection techniques using remotely sensed data," *Int. J. Remote Sensing*, vol. 10, pp. 989-1003, 1989.
- [3] E. Rignot and J. Van Zyl, "Change Detection Techniques for ERS-1 SAR Data", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, No. 4, July 1993.
- [4] Ø. Midtgaard, T. O. Sæbø, R. Hansen, V. Myers, J. R. Dubberley and I. Quidu, "Change Detection Using Synthetic Aperture Sonar: Preliminary Results from the Larvik Trial", *OCEANS'2011 MTS/IEEE*, Kona, Hawaiï, Sept 19-22, 2011.
- [5] I. Quidu, "Incoherent change detection using amplitude sidescan sonar images", *11th European Conference on Underwater Acoustics (ECUA)*, Edinburgh, Scotland, 2nd and 6th July, 2012.
- [6] T. Lyons and D. Brown, "Temporal variability of seafloor roughness and its impact on coherent change detection," *Proceedings of the International Conference on SAS and SAR*, Lercici, Italy, 2010.
- [7] T. O. Sæbø, R. E. Hansen, H. J. Callow, S. A. V. Synnes, "Coregistration of synthetic aperture sonar coregistration of synthetic aperture sonar Images from repeated passes", *proceedings of the 4th International Conference on "Underwater Acoustic Measurements: Technologies & Results"*, Kos Island, Greece, 2011.
- [8] D.G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, 60(2), pp. 91-110, 2004.
- [9] H. Bay, T. Tuytelaars, L.V. Gool, "Surf: Speeded Up Robust Features", *In: 9th European Conference on Computer Vision (ECCV'06)*, May 2006.
- [10] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Optics Letters*, Vol. 33, pp. 156-158, 2008.
- [11] J. O. Smith, *Mathematics of the Discrete Fourier Transform (DFT)* (W3K Publishing, <http://www.w3k.org/books/>), 2007.
- [12] W.H. Press, B.P. Flannery, W.A. Teukolsky, W.T Vetterling, *Numerical Recipes in C*, Cambridge University Press, pp. 299, 1988.
- [13] L. G. Brown, "A survey of image registration techniques", *ACM computing Surveys*, vol. 24, No 4, pp. 325-376, 1992.
- [14] Z. Xie and G. Farin, "Image registration using hierarchical B-splines", *IEEE Transaction on Visualization and Computer Graphics*, 10(1), pp. 85-94, 2004.
- [15] S. Lee, G. Wolberg and S. Y. Shin, "Scattered data interpolation with multilevel B-splines", *IEEE transactions on visualization and computer graphics*, Vol. 3, No; 3, July-September 1997.
- [16] F.L. Bookstein, "Principal warps: thin-plate splines and the decomposition of deformations", *IEEE Trans Pattern Anal Mach Intell*, vol. 11, no 6, pp. 567-585, 1989.

Annexe 5 : Détection et suivi d'obstacles dans la colonne d'eau

Cette annexe est une copie de l'article [QUI12c] publié dans *IEEE Journal of Oceanic Engineering* en 2012 dans le cadre d'un contrat de recherche pour DGA TN Brest. Cet article décrit la détection d'obstacles dans la colonne d'eau, le modèle d'état utilisé dans le filtre de Kalman sans parfum mis en œuvre pour le pistage des obstacles au cours de la séquence, et l'estimation hors ligne de l'altitude de l'obstacle. Ce dernier point a fait l'objet d'une amélioration lors du second contrat consacré à ce sujet qui est détaillée au paragraphe 3.3.1.

Peer-Reviewed Technical Communication

Robust Multitarget Tracking in Forward-Looking Sonar Image Sequences Using Navigational Data

Isabelle Quidu, Luc Jaulin, Alain Bertholom, and Yann Dupas

Abstract—This paper presents a new approach to the problem of tracking objects in sequences of forward-looking sonar images. Unlike previous work, navigational data are taken as inputs to the state model of the Kalman filter used for tracking fixed obstacles. This model allows a robust prediction of their apparent motion in relation to the position of the sonar. A complete framework is presented where detection and data association issues are also discussed. An assessment of the proposed method has been carried out on real data from two different systems. Moreover, whereas the state model was first derived for a ground obstacle, a modified state model is proposed to estimate the altitude of the obstacle in relation to the sonar position using a number of successive pings.

Index Terms—Autonomous underwater vehicles (AUVs), multitarget tracking, obstacle avoidance, process modeling.

I. INTRODUCTION

AUTONOMOUS UNDERWATER VEHICLES (AUVs) are expected to perform survey missions in both known and unknown environments. While the primary mission of an AUV is usually data collection, generally achieved by means of a sidescan sonar or multibeam echosounder, another key task is to ensure its safety. For this, the AUV can be equipped with a forward-looking sonar (FLS) by which it is able to sense the environment at a certain distance in the direction of travel. The main objective is to determine whether it is faced with an obstacle, to avoid a collision that could damage the vehicle. After an initial detection, the AUV first has to track the object to confirm the existence of an obstacle and obtain information

about it. This must be done sufficiently early, i.e., at a safe distance between the AUV and the obstacle, to perform an avoidance maneuver if necessary.

This paper focuses on the tracking of multiple fixed objects through a sequence of FLS images. Robustness relies on the integration in the process of accurate navigational data obtained via an iXSea (iXBlue company, Marly Le Roi, France) Inertial Navigation Unit (INU) and an RD Instruments (RDI; Poway, CA) Doppler velocity log (DVL).

In previous work dealing with FLS data, object-based tracking between successive frames is generally based solely on the incoming sonar data. In addition, most of them focus on moving obstacles, although Williams *et al.* developed a Kalman filter for the localization of stationary obstacle between successive frames [1]. In a broad sense, an object appears to be moving due to either platform motion (the present case) or its own movement (the more general case). Cobb *et al.* proposed a low-power tracking and homing algorithm for a stationary floating object [2]. Lane *et al.* performed optical flow calculations on moving objects and used the motion estimates to match them with new returns (i.e., object detections) with a metric called the compatibility value [3]. A tracking tree then allows keeping several possible tracks for a deferred decision. More recently, Trucco *et al.* followed by Petillot *et al.* used Kalman filters as the core element of their tracking schemes with an extended state vector composed of position and area and their associated first and second derivatives [4], [5]. Clark *et al.* then compared this traditional approach including a Kalman filter for each target and a nearest neighbor algorithm for data association with a new approach using the developed particle implementation of the multiple-target probability hypothesis density (PHD) filter proposed by Mahler and a target state estimate-to-track data association technique [6], [7]. Finally, a multitarget tracking method using a Dual frequency Identification SONar (DIDSON) sonar to monitor fish movement in trawls was presented in [8]. Tracking is performed by a simple alpha-beta fixed gain filter and a global nearest neighbor method [9].

In this paper, we aim to perform real-time tracking by means of Kalman filtering that takes navigational data as inputs. Section II describes the detector algorithm that provides measurements for track initiation and update. Unlike previous work, a new state model, described in Section III, has been derived from the AUV process model. This allows the prediction of the polar coordinates of an obstacle under the assumptions that the obstacle is fixed and lies on the seabed. This process model

Manuscript received February 21, 2011; revised November 15, 2011; accepted March 05, 2012. Date of publication May 16, 2012; date of current version July 10, 2012. This work was supported by the Groupe d'Etudes Sous-Marines de l'Atlantique (GESMA), a division of the French Ministry of Defense (DGA).

Associate Editor: D. A. Abraham.

I. Quidu and L. Jaulin are with the Lab-STICC UMR CNRS 6285, Ecole Nationale de Techniques Avancées (ENSTA) Bretagne, Brest Cedex 9 29806, France (e-mail: isabelle.quidu@ensta-bretagne.fr; luc.jaulin@ensta-bretagne.fr).

A. Bertholom is with the Groupe d'Etudes Sous-Marines de l'Atlantique (GESMA), DGA/DET/GESMA, Brest Armées 29240, France (e-mail: alain.bertholom@dga.defense.gouv.fr).

Y. Dupas is with the Service Hydrographique et Océanographique de la Marine (SHOM), SHOM/DO/HOM/DEV, Brest 29200, France (e-mail: yann.dupas@shom.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

This paper has supplementary downloadable materials available at <http://ieeexplore.ieee.org>, provided by the author. Contact isabelle.quidu@ensta-bretagne.fr for further questions about this work.

Digital Object Identifier 10.1109/JOE.2012.2190572

TABLE I
GENERAL PARAMETERS

General parameters	
$\mathbf{x} = (d \ \delta)^T$	Polar coordinates (meters, radians) of the target in the sonar frame
$\mathbf{y} = (y^i \ y^j)^T$	Screen coordinates of the target in pixels
Δd	Along-track sampling rate in meters
$\Delta \delta$	Across-track sampling rate in degrees
$\mathbf{p}_a = (p_a^x \ p_a^y \ p_a^z)^T$	Coordinate vector of the AUV in the world reference frame in meters
$\mathbf{m}_a = (m_a^x \ m_a^y \ m_a^z)^T$	Coordinate vector of a proud seafloor object in the world reference frame in meters
$\mathbf{p}_r = (p_r^x \ p_r^y \ p_r^z)^T$	Coordinate vector of the AUV in the relative reference frame in meters
$\mathbf{m}_r = (m_r^x \ m_r^y \ m_r^z)^T$	Coordinate vector of a proud seafloor object in the relative reference frame in meters
$\mathbf{V}_a = (v_a^x \ v_a^y \ v_a^z)^T$	Velocity vector of the AUV in the world reference frame in meters per second
$\mathbf{V}_r = (v_r^x \ v_r^y \ v_r^z)^T$	Velocity vector of the AUV in the relative reference frame in meters per second
h	AUV altitude in meters
d_{st}	Distance between the sonar and the target along the z-axis
(φ, θ, ψ)	Euler angles ¹ (roll, pitch, yaw) in radians
R_ψ	Yaw rotation matrix
R_θ	Pitch rotation matrix
R_φ	Roll rotation matrix
$R_{euler}(\varphi, \theta, \psi)$	Euler rotation matrix, <i>i.e.</i> three composed rotations $R_\psi, R_\theta, R_\varphi$ that move the relative reference frame to the world reference frame. Its arguments are the Euler angles.
f_a, f_b, f_c	Intermediate functions involved in the process modeling
f	Process model function
c_v	Coefficient of variation

¹ The Euler angles are three angles introduced by Euler (1770) to describe the orientation of a rigid body or equivalently a rotation matrix. To describe such an orientation three parameters are required. These parameters make possible to describe a motion from a reference frame to a given referred frame. This motion can be described in various ways. One of them, called the ZYX or Yaw-Pitch-Roll formulation, is commonly used in underwater robotics.

is the state model of a Kalman filter described in Section IV and navigational data are used as inputs so as to allow robust tracking of an obstacle even if the vehicle changes its velocity and/or its attitude. In Section V, data association is performed to take advantage of the behavior of the Kalman filter via the innovation term. According to the assumed condition of a seafloor object, tracking of a tethered object in the water column is not performed especially well. To address this, a “batch” procedure which requires some number of consecutive pings to function has been proposed: it performs a deferred estimation of the tether length by minimizing the innovation value through the tracking procedure. This procedure is explained in Section VI. Section VII gives several results on real data obtained by two different FLS, after which some concluding remarks are made in Section VIII.

Notations for the main parameters used in this paper are listed in Tables I and II. Some parameters are to be used in the world (absolute) reference frame ($O_a, \mathbf{x}_a, \mathbf{y}_a, \mathbf{z}_a$) (O_a : center of gravity of Earth, \mathbf{x}_a : geographical north direction, \mathbf{y}_a : east direction, \mathbf{z}_a : gravity direction) on the one hand, and in the relative reference frame ($O_r, \mathbf{x}_r, \mathbf{y}_r, \mathbf{z}_r$) (O_r : center of gravity of the vehicle, \mathbf{x}_r : the direction of movement of the vehicle, \mathbf{y}_r : toward its right, \mathbf{z}_r : down) on the other hand.

II. OBJECT DETECTION

To track an obstacle in a series of sonar images, it must first be discriminated from the background signal by employing some kind of a detection algorithm. In imaging sonar, detection is sometimes performed by segmentation and can be achieved by a double thresholding (one for echoes from the object, one for the acoustic shadows cast on the seafloor) with or without pre-filtering the image [5], [10].

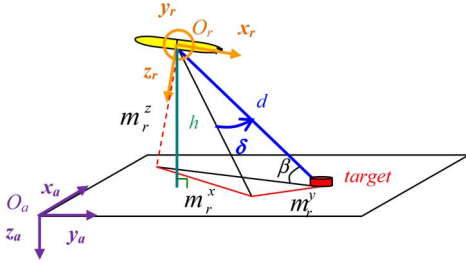
Here, a simple goodness-of-fit test is performed to detect statistical deviations from which detected points can be derived. This test exploits the fact that high-frequency sonar images suffer from a well-known multiplicative noise process that is commonly called the *speckle* [11], [12].

When the number of scatterers in a resolution cell is sufficiently high, the pixel gray levels follow a Rayleigh law. Under this hypothesis, we have derived a simple goodness-of-fit test that consists in verifying the ratio value between the mean and the standard deviation of pixels, called the coefficient of variation c_v [13]. This coefficient is constant for a Rayleigh-distributed variable and equal to

$$c_v^{\text{th}} = \sqrt{\frac{4}{\pi} - 1} = 0.52. \quad (1)$$

TABLE II
 KALMAN FILTER PARAMETERS

Kalman filter parameters	
k	Step number
\mathbf{x}_k	State vector at step k
$\mathbf{x}_{k k-1}$	State vector at step k given data from step $k-1$
\mathbf{y}_k	Measurement vector at step k
\mathbf{u}_k	Input vector at step k
\mathbf{v}_k	Gaussian state noise at step k
\mathbf{w}_k	Gaussian measurement noise at step k
\mathcal{Q}_k	Noise covariance matrix at step k
\mathcal{R}_k	Measurement covariance matrix at step k
d_R	Scaling scalar in meters for \mathcal{R}_k
P_k	State covariance matrix at step k
$\mathbf{v}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k k-1}$	Innovation term at step k
\mathcal{R}_k^v	Innovation covariance matrix at step k
\mathcal{X}_k^i	Sigma point number i for unscented Kalman filtering at step k
W_i	Weight parameter of the Sigma point number i for unscented Kalman filtering


 Fig. 1. Operational configuration (O_a supposed to be the center of gravity of Earth).

The practical implementation of this test is given in Section VII-B.

The simple ratio test above has been compared to the non-parametric Kolmogorov–Smirnov statistical test for equal distributions, which gives similar results but requires a higher computation time.

III. PROCESS MODEL

To use the Kalman filter to track obstacles given noisy detection points, we first have to define the state model. This is the process model that provides the sonar coordinates (d, δ) , i.e., the polar coordinates (meters, radians) of the target in the sonar frame, of a detected object given the AUV motion [15]. It is obtained as follows.

In the relative reference frame, a sea floor target is located by means of the following equations (see Fig. 1):

$$\begin{cases} (m_r^x)^2 + (m_r^y)^2 + (m_r^z)^2 = d^2 \\ m_r^y = d \sin(\delta) \\ -\sin(\theta) m_r^x + \cos(\theta) \sin(\varphi) m_r^y + \cos(\theta) \cos(\varphi) m_r^z = h \end{cases} \quad (2)$$

where $\mathbf{m}_r = (m_r^x \ m_r^y \ m_r^z)^T$ is the coordinate vector of the proud seafloor target, h is the AUV altitude in meters, and (φ, θ, ψ) are the Euler angles that define the attitude of the vehicle.

From these equations, it can be shown that $a_1 (m_r^x)^2 + b_1 m_r^x + c_1 = 0$, when

$$a_1 = \left[1 + \left[\frac{\sin(\theta)}{\cos(\theta) \cos(\varphi)} \right]^2 \right]$$

$$b_1 = 2 \frac{\sin(\theta)}{\cos^2(\theta) \cos^2(\varphi)} [h - d \cos(\theta) \sin(\varphi) \sin(\delta)]$$

$$c_1 = d^2 [\sin^2(\delta) - 1] + \left[\frac{h - d \cos(\theta) \sin(\varphi) \sin(\delta)}{\cos(\theta) \cos(\varphi)} \right]^2$$

so that

$$\begin{pmatrix} m_r^x \\ m_r^y \\ m_r^z \end{pmatrix} = \begin{pmatrix} \frac{-b_1 + \sqrt{\Delta_1}}{2a_1} \\ d \sin(\delta) \\ \frac{h + \sin(\theta) \left(\frac{-b_1 + \sqrt{\Delta_1}}{2a_1} \right) - d \cos(\theta) \sin(\varphi) \sin(\delta)}{\cos(\theta) \cos(\varphi)} \end{pmatrix}$$

where $\Delta_1 = b_1^2 - 4a_1c_1$. That is

$$\begin{pmatrix} m_r^x \\ m_r^y \\ m_r^z \end{pmatrix} = f_a(d, \delta). \quad (3)$$

Let $(\dot{d}, \dot{\delta}) = (\partial d/\partial t, \partial \delta/\partial t)$ be the derivative of the state vector with respect to time. It can be derived from the two first equations of (2), such that

$$\begin{aligned} \dot{d} &= \frac{\dot{m}_r^x m_r^x + \dot{m}_r^y m_r^y + \dot{m}_r^z m_r^z}{d} \\ \dot{\delta} &= \frac{\dot{m}_r^y - \frac{\sin(\delta)}{d} (\dot{m}_r^x m_r^x + \dot{m}_r^y m_r^y + \dot{m}_r^z m_r^z)}{d \cos(\delta)}, \\ \text{i.e., } (\dot{d}, \dot{\delta}) &= f_c(\dot{m}_r^x, \dot{m}_r^y, \dot{m}_r^z, d, \delta). \end{aligned} \quad (4)$$

By introducing the Euler rotation matrix R_{euler} defined as

$$\begin{aligned} R_{\text{euler}}(\varphi, \theta, \psi) &= R_\psi R_\theta R_\varphi = \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &\times \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix} \end{aligned}$$

we can write

$$\mathbf{p}_a - \mathbf{m}_a = -R_{\text{euler}}(\varphi, \theta, \psi) \cdot \mathbf{m}_r \quad (5)$$

where $\mathbf{p}_a = (p_a^x \ p_a^y \ p_a^z)^T$ is the coordinate vector of the AUV in the world reference frame in meters (we assume that its location is coincident with all the other sensors) and $\mathbf{m}_a = (m_a^x \ m_a^y \ m_a^z)^T$ is the coordinate vector of the proud seafloor object in the world reference frame in meters. Also, the Euler angles (φ, θ, ψ) for roll, pitch, and yaw are defined in radians.

By differentiating (5), we obtain

$$\mathbf{V}_a = -\dot{R}_{\text{euler}} \cdot \mathbf{m}_r - R_{\text{euler}} \cdot \dot{\mathbf{m}}_r$$

with $\mathbf{V}_a = R_{\text{euler}}(\varphi, \theta, \psi) \cdot \mathbf{V}_r$ being the velocity of the AUV in the world reference frame in meters per second, so that

$$\begin{aligned} \dot{\mathbf{m}}_r &= -R_{\text{euler}}^T(\varphi, \theta, \psi) \dot{R}_{\text{euler}}(\dot{\varphi}, \dot{\theta}, \dot{\psi}) \mathbf{m}_r - \mathbf{V}_r, \\ &= f_b(m_r^x, m_r^y, m_r^z, \mathbf{V}_r, \varphi, \theta, \psi, \dot{\varphi}, \dot{\theta}, \dot{\psi}) \end{aligned} \quad (6)$$

by exploiting the unitary nature of the Euler matrix and where $\mathbf{V}_r = (v_r^x \ v_r^y \ v_r^z)^T$ is the velocity vector of the AUV in the relative reference frame, and $\dot{R}_{\text{euler}}(\varphi, \theta, \psi, \dot{\varphi}, \dot{\theta}, \dot{\psi})$ is defined as

$$\begin{aligned} \dot{R}_{\text{euler}}(\varphi, \theta, \psi, \dot{\varphi}, \dot{\theta}, \dot{\psi}) &= \dot{\psi} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} R_\psi R_\theta R_\varphi + \dot{\theta} R_\psi \\ &\times \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} R_\theta R_\varphi + \dot{\varphi} R_\psi R_\theta \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} R_\varphi. \end{aligned}$$

We can now derive the expression for a moving object as a composite function of its initial position and navigational data

$$(\dot{d}, \dot{\delta}) = (f_c \circ f_b \circ f_a)(d, \delta, \mathbf{V}_r, \varphi, \theta, \psi, \dot{\varphi}, \dot{\theta}, \dot{\psi}) \quad (7)$$

where $(g \circ h)(\cdot) = g(h(\cdot))$ stands for a composite function.

IV. KALMAN FILTERING

A. State Equation

The state vector is composed of the sonar coordinates, i.e., $\mathbf{x} = (d \ \delta)^T$. Considering the previous part and (7), the state equation can be written in the discrete domain according to the Euler method with a step number k

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{v}_{k-1} \quad (8)$$

where the input vector \mathbf{u}_{k-1} is derived from navigational data such that $\mathbf{u}_{k-1} = (\mathbf{V}_r \ \varphi \ \theta \ \psi \ \dot{\varphi} \ \dot{\theta} \ \dot{\psi})_{k-1}^T$, \mathbf{v}_{k-1} stands for the state noise vector whose covariance matrix is Q_{k-1} (explained in Section IV-C) and $f = f_c \circ f_b \circ f_a$ is a nonlinear state function determined in Section III.

B. Measurement Equation

The measurement vector $\mathbf{y} = (y^i \ y^j)^T$ comprises obstacle coordinates on the screen in pixels, as determined by the detection step from Section VII-B. The measurement equation is then

$$\mathbf{y}_k = H_k \mathbf{x}_k + \mathbf{w}_{k-1} \quad (9)$$

where \mathbf{w}_{k-1} stands for the measurement noise vector whose covariance matrix is R_{k-1} (see Section IV-C), and

$$H_k = H = \begin{pmatrix} 1 & 0 \\ \Delta d & 0 \\ 0 & 1 \\ & \Delta \delta \end{pmatrix} \quad \forall k$$

is the measurement matrix, where Δd (resp., $\Delta \delta$) stands for the along-track (resp., across-track) sampling rate in meters (resp., in degrees).

C. Covariance Matrices

For an optimal Kalman filtering, state and measurement noises as well as the initial state vector have to be Gaussian and mutually independent. However, the Kalman filter is known to be robust enough to work even in cases of noises that differ slightly from these hypotheses, in particular, for noises that are not exactly Gaussian.

The state variances were estimated by a heuristic procedure. The worst precision bounds of navigational parameters between the two AUVs are taken into account so as to simulate inputs of the state equation as uniformly distributed. Then, more than 8000 possible values of sonar coordinates were propagated in the process model, typically a range from 20 to 100 m by step of 1 m for d and a range from -0.5 to 0.5 rad by step of 0.01 rad for δ . This operation was repeated a hundred times and the average values stand as the estimated elements of the state covariance matrix, where state noises appear decoupled (actually with an absolute covariance less than 10^{-8})

$$Q_k = Q = 10^{-6} \begin{bmatrix} 0.48 & 0 \\ 0 & 0.024 \end{bmatrix} \quad \forall k.$$

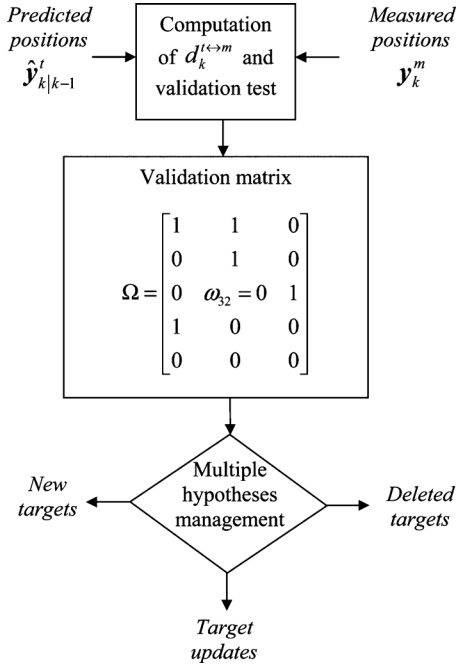


Fig. 2. Principle of data association procedure: example in the case of $M = 5$ and $T = 3$.

We assume that the measurement accuracy is about d_R meters in both the along-track and across-track directions. The measurement noises are then modeled as decoupled Gaussian noises with a standard deviation given in pixels but equivalent to d_R meters, which leads to the following measurement covariance matrix by introducing the sonar parameters:

$$R_k = \begin{bmatrix} [\text{int}(d_R/\Delta d)]^2 & 0 \\ 0 & [\text{int}(d_R/(\Delta\delta\Delta dy_k^i))]^2 \end{bmatrix}$$

where y_k^i stands for the first component of the measurement vector y_k such that Δdy_k^i determines the distance between the sonar and the detected obstacle. The $\text{int}()$ function returns the closest integer. This latter function is coherent with the measurement vector components that are the screen coordinates of the target and hence integer values.

In this paper, the possible values of d_R range from 2 to 15 m. On the one hand, a high value is better to track large obstacles such as a wreck to avoid multiple tracks simultaneously on this specific obstacle. On the other hand, a small value is necessary to properly distinguish two smaller obstacles within close proximity of each other. In other words, for smaller values of d_R , the probability of detection as well as the probability of false alarm is higher.

D. Initialization

A Kalman filter is initialized for each new detected target. To avoid several Kalman filters on the same target, detection points that are too close to each other are fused by averaging their coordinates. This vicinity is given by a gate defined by the uncertainties of the measurement covariance matrix. This point will be discussed again in Section V.

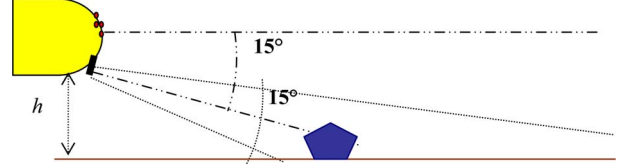


Fig. 3. Configuration of the *Reson Seabat 8101* FLS on *Redermor* AUV.

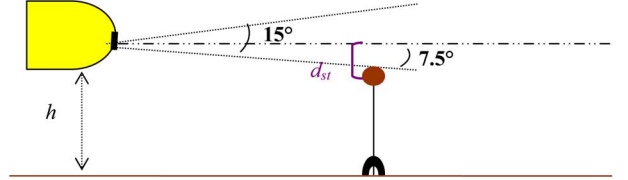


Fig. 4. Configuration of the *Blueview P450* FLS on *Daurade* AUV.

The initial state $\mathbf{x}_{0|0}$ consists of the sonar coordinates corresponding to the detected obstacle, i.e.,

$$\hat{\mathbf{x}}_{0|0} = \mathbf{x}_0 = (d_0 \quad \delta_0)^T = (y^i \Delta d \quad y^j \Delta \delta)^T$$

where y^i and y^j are the coordinates (in pixels) of the detected point on the screen for this obstacle.

To not be too confident in the first state, we set large uncertainties of about 5 m in range and 5° in azimuth. Indeed, these large values allow the Kalman filtering not to be too confident in the first measurement. Given the sonar parameters, we can derive the initial covariance matrix of

$$P_{0|0} = \begin{bmatrix} (100\Delta d)^2 & 0 \\ 0 & (3\Delta\delta)^2 \end{bmatrix}$$

for the *Reson Seabat 8101* FLS and

$$P_{0|0} = \begin{bmatrix} (30\Delta d)^2 & 0 \\ 0 & (30\Delta\delta)^2 \end{bmatrix}$$

for the *Blueview P450* FLS.

E. Prediction Stage

The prediction stage computes the new state $\hat{\mathbf{x}}_{k|k-1}$ given the previous one $\hat{\mathbf{x}}_{k-1|k-1}$, and is carried out by performing an unscented transform of $\hat{\mathbf{x}}_{k-1|k-1}$ because of the strong nonlinearity of the state function f . Details on the unscented Kalman filter can be found in [16]. The unscented transform is described hereafter.

We begin by creating a number of Sigma points $\mathbf{X}_{k-1|k-1}^i$, for $i = 1$ to $2n + 1$ with respective weight W_i , where n is the dimension of $\hat{\mathbf{x}}_{k-1|k-1}$. These points are uniformly distributed on an ellipsoid such that their mean and covariance are $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$ [17]. Practically, for $j \in \{1 \dots n\}$, we compute

$$\begin{aligned} \mathbf{X}_{k-1|k-1}^0 &= \hat{\mathbf{x}}_{k-1|k-1}, & W_0 &= \frac{\kappa}{n + \kappa} \\ \mathbf{X}_{k-1|k-1}^j &= \hat{\mathbf{x}}_{k-1|k-1} + \left(\sqrt{(n + \kappa)\mathbf{P}_{k-1|k-1}} \right)_j \\ W_j &= \frac{1}{2(n + \kappa)} \end{aligned}$$

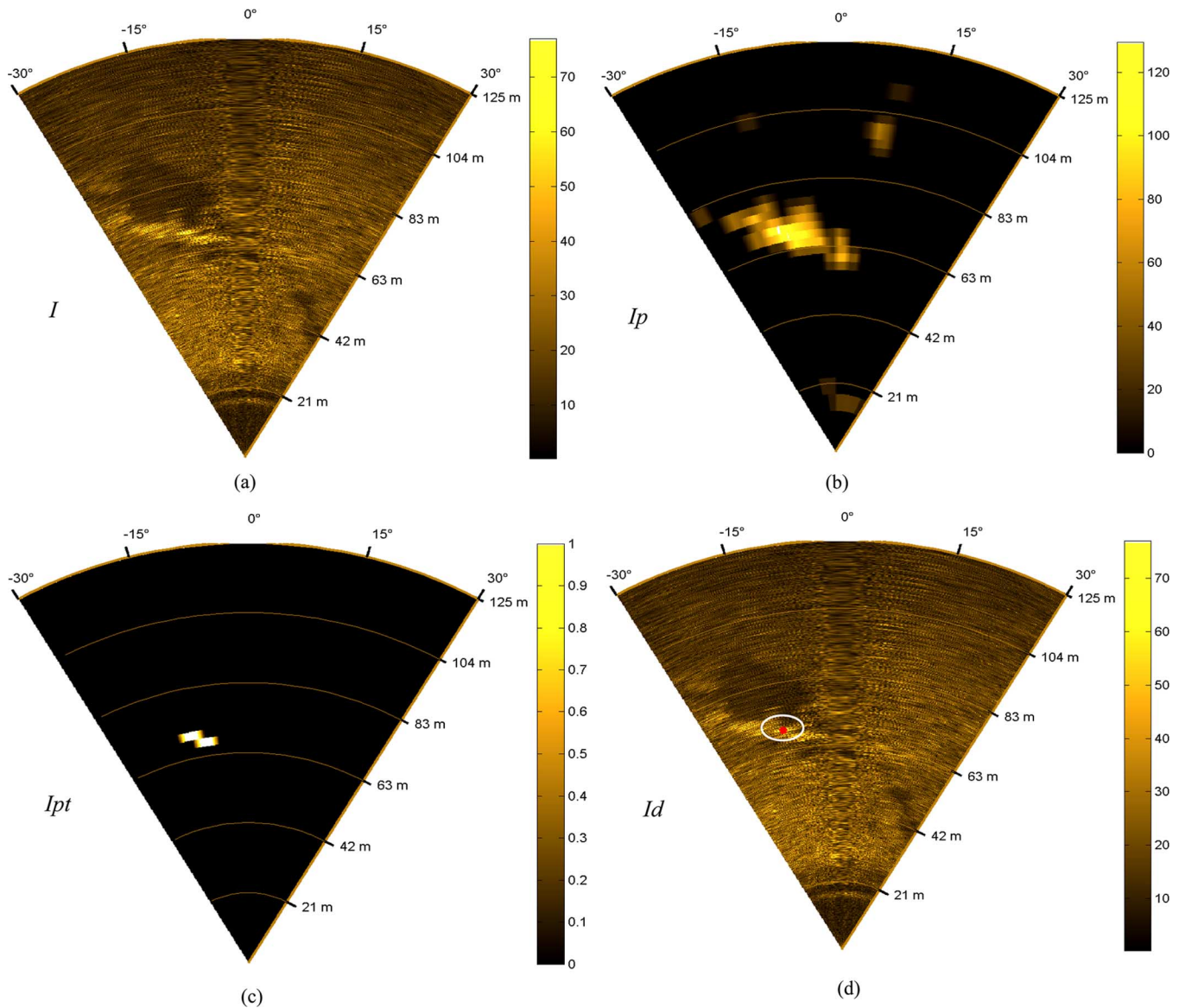


Fig. 5. Example of successive detection stages for a real sector-scan image (data of the 20th ping of the *Reson Seabat 8101* FLS sequence) where a shipwreck stands: (a) raw image I , (b) image I_p of detected areas, (c) thresholded image I_{pt} , (d) raw image I_d with the final detected point in red (encircled in white).

$$\mathbf{X}_{k-1|k-1}^{j+n} = \hat{\mathbf{x}}_{k-1|k-1} - \left(\sqrt{(n+\kappa)P_{k-1|k-1}} \right)_j$$

$$W_{j+n} = \frac{1}{2(n+\kappa)}$$

where κ is usually set to $3-n$ [16] and $(\sqrt{(n+\kappa)P_{k-1|k-1}})_j$ stands for the j th column of the square root of the matrix $(n+\kappa)P_{k-1|k-1}$ obtained by a Cholesky decomposition [18].

Next, we compute the predicted values of these points given by the nonlinear function f

$$\mathbf{X}_{k|k-1}^i = f\left(\mathbf{X}_{k-1|k-1}^i, \mathbf{u}_{k-1}\right), \quad \text{for } i = 1 \text{ to } 2n+1.$$

The predicted state can now be derived by the following weighted sum:

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=0}^{2n} W_i \mathbf{X}_{k|k-1}^i$$

as well as the associate covariance matrix

$$P_{k|k-1} = \sum_{i=0}^{2n} W_i \left[\mathbf{X}_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1} \right] \left[\mathbf{X}_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1} \right]^T.$$

F. Correction Stage

This final stage calculates the corrected sonar coordinates $\hat{\mathbf{x}}_{k|k}$ given the predicted one $\hat{\mathbf{x}}_{k|k-1}$ and the new measurement \mathbf{y}_k which is a new detection at the k th filtering step. After computing $\hat{\mathbf{y}}_{k|k-1} = H\hat{\mathbf{x}}_{k|k-1}$ given the predicted state $\hat{\mathbf{x}}_{k|k-1}$, we can correct the state by applying the Kalman equations in the linear case

$$\begin{cases} \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \\ K_k = P_{k|k-1}H_k^T (H_kP_{k|k-1}H_k^T + R_k)^{-1} \\ = P_{k|k-1}H_k^T (R_k^v)^{-1} \\ P_{k|k} = P_{k|k-1} - K_kR_k^v(K_k)^T. \end{cases} \quad (10)$$

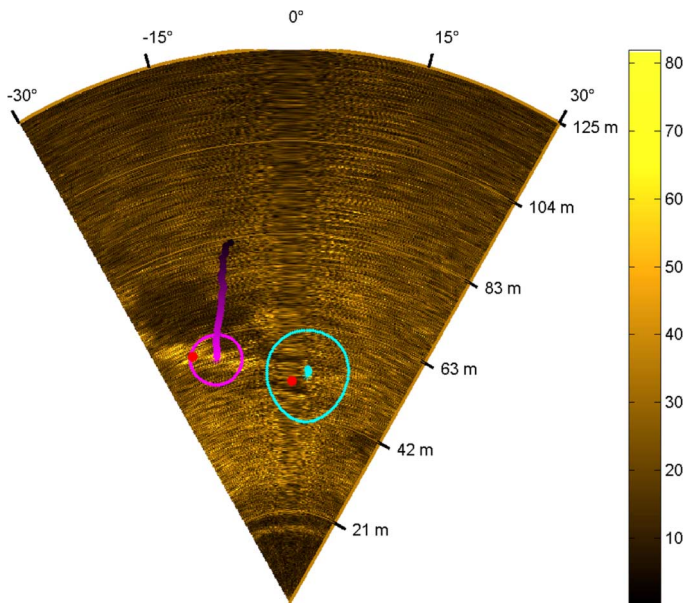


Fig. 6. Data of the 33rd ping of the *Reson Seabat 8101* FLS sequence: tracking of a shipwreck where $d_R = 15$ m: measurements in red, new filtered positions (at the center of the confidence ellipses) and trailing history of 32 previous positions in purple (vanishing colors for the oldest positions) and a new track with a large ellipse initialized by a false alarm in cyan.

The difference $\mathbf{v}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$ is called the innovation term and the associated covariance matrix is R_k^v . The impact of this term is controlled by the Kalman filter gain K_k .

If no detection occurs at this step, we keep the previous detection and the corresponding covariance matrix, i.e., $\mathbf{y}_k = \mathbf{y}_{k-1}$ and $R_k = R_{k-1}$. Another possibility would be to put large values in the diagonal matrix measurement R_k . However, this could lead to errors for the data association procedure described in the next section.

V. DATA ASSOCIATION

Another crucial point of this multitarget tracking algorithm is the association of new measurements to existing tracks at each ping. Data association techniques can be divided into the following two categories [19]: 1) approaches that focus primarily on selecting target measurements that only fall within “validation gates” generated by existing tracks; or 2) approaches that focus on measurements and that try, for each of them, to find an existing track in their vicinity and create a new track if not.

The first approach is more suitable for tracking based on Kalman filtering that requires one or more measurements for use during the correction stage. This approach was chosen in [4] and [5], where data association was based on a nearest neighbor algorithm using the Euclidian distance between all the observations (i.e., measurements) and the predicted positions of tracked objects. As mentioned in [9], the global nearest neighbor method is indeed one of the simplest methods that can be used for data association. It also includes the formation of an observation-to-track assignment matrix to deal with conflict situations. In a similar way, we developed a process guided by previous works such as [20] and [21]. This process consists of two stages: first, we compute a validation matrix to visualize all possible matches between observations and existing tracks, and

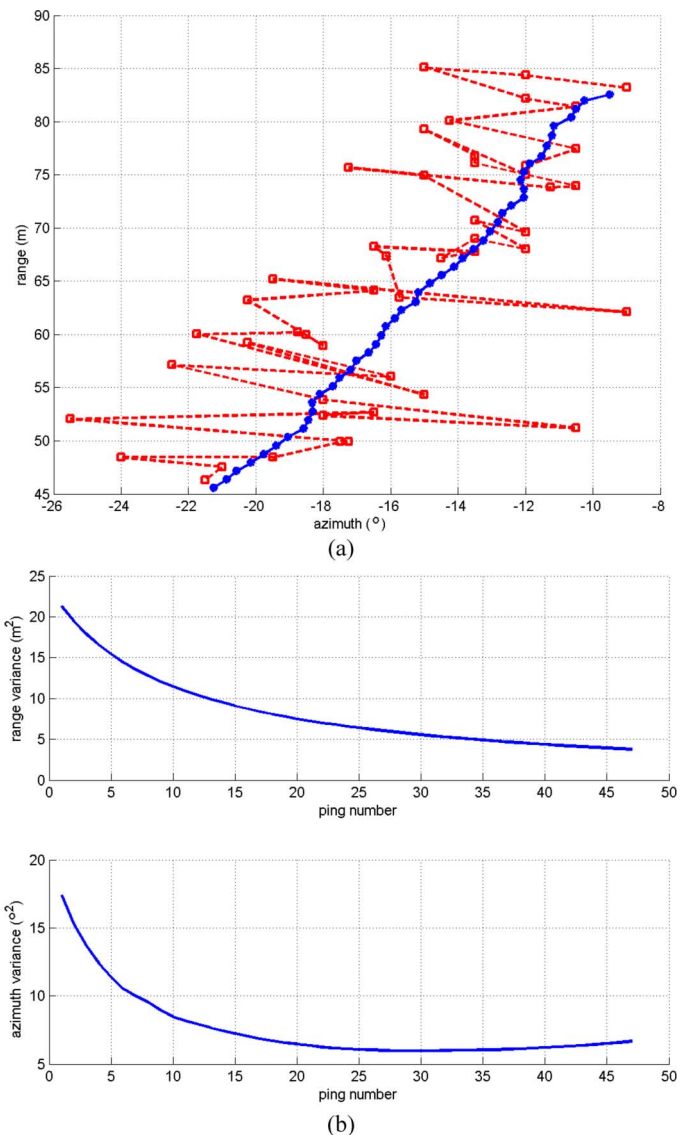


Fig. 7. (a) Filtering results on shipwreck sequence (dashed red line: measurements; blue line: estimated trajectory). (b) *A posteriori* variances evolution along the sequence.

then, we perform multiple-hypotheses management to decide on the final associations. Fig. 2 gives a diagram of the overall process.

In the first stage, for a given step $k > 0$, we need to deal with T existing tracks and M measurements. The validation matrix Ω of size $M \times T$ is defined with binary elements ω_{mt} to indicate if measurement $m \in \{1, \dots, M\}$ lies in the validation gate for the target $t \in \{1, \dots, T\}$. A validation gate is an ellipse centered on the predicted position $\hat{\mathbf{y}}_{k|k-1}^t$ whose axes are defined by the diagonal elements of the covariance matrix R_k^v of the innovation term $\mathbf{v}_k = \mathbf{y}_k^m - \hat{\mathbf{y}}_{k|k-1}^t$ (assumed to be normally distributed with zero mean). The Mahalanobis distance $d_k^{t \leftrightarrow m} = \mathbf{v}_k^T (R_k^v)^{-1} \mathbf{v}_k$ has a χ^2 distribution with r (equal to the number of components of the vector \mathbf{y}_k) degrees of freedom. Let χ_r^2 be the χ^2 variable with r degrees of freedom and $\chi_{r,1-\alpha}^2$ the critical value of this variable (tabulated) at a level of significance α . The second stage performs a validation test as follows: we accept the measurement at step k if $d_k^{t \leftrightarrow m} \leq \chi_{r,1-\alpha}^2$ where

$\chi_{r,1-\alpha}^2 = 5.991$ for $r = 2$ and a significance level $\alpha = 0.05$ [22]. This validation test based on statistical properties of the innovation term of the Kalman filter was preferred to a matching function found in [1]. This matching function to be minimized is designed to check possible matches between new detections and existing tracks and takes into account Euclidian distances between elements being matched and their covariances.

The management of multiple hypotheses consists in deciding if a given track has to be initialized, can be updated, or has to be deleted. Of course, a measurement can be associated to only one target.

For a given track, i.e., for a column of the validation matrix, this decision depends on the number of binary elements set to one, and can be one of four cases.

- 1) If there is only one binary element set to one, the corresponding measurement is associated to this track for the correction stage of the Kalman filtering (if this is really the closest possible track).
- 2) If there are more than one binary elements set to one, we first consider the closest measurement m_{\min} (the lowest corresponding distance $d_k^{l \leftrightarrow m_{\min}}$) and a validate gate defined by the uncertainties of the measurement covariance matrix. If other measurements “fall” in the ellipse, the effective measurement is the average of these measurements.
- 1) If no binary element is set to one, a copy of the previous measurement and of the corresponding covariance matrix is used for the correction stage.
- 2) Finally, for a row of the validation matrix with only zeros, a new track is created and a Kalman filter initialized.

Another point concerns track interruption. Hypothesis tests, like the Wald sequential probability ratio test or statistical tests of Mehra and Peschon, appear interesting because they are based on the innovation values [9], [23]. Unfortunately these tests were unusable here because of the difficulty in determining appropriate parameter values for the first one and because of lack of statistical samples (available innovation values) for the others. For these reasons, we chose a much simpler rule, besides the one used by Cobb *et al.*, which states that if no measurement is associated to a given track in three successive pings, then that track is interrupted [2].

VI. CASE OF AN OBSTACLE IN THE AUV PATH

In the previously developed method, the obstacle is assumed to be lying on the seafloor. However, the case of an obstacle that is in the water column, such as a tethered mine, also has to be considered. In this case, the Kalman filtering leads to a biased estimation of successive positions of the floating target. Indeed it will appear to be moving faster than the predicted positions according to the process model. Starting with this observation, a new algorithm is proposed that allows a deferred estimation of the z -coordinate (along the absolute z -axis) of the obstacle related to the vehicle using the information given by successive pings. This piece of information is of particular interest for obstacle avoidance but generally unavailable because of little or no sampling of the sonar data along the vertical axis for typical FLS with a single linear or cylindrical array. The only alternative is to use a more expensive and more complex sonar system (an interferometric sonar, a steerable sonar system, or a 3-D sonar

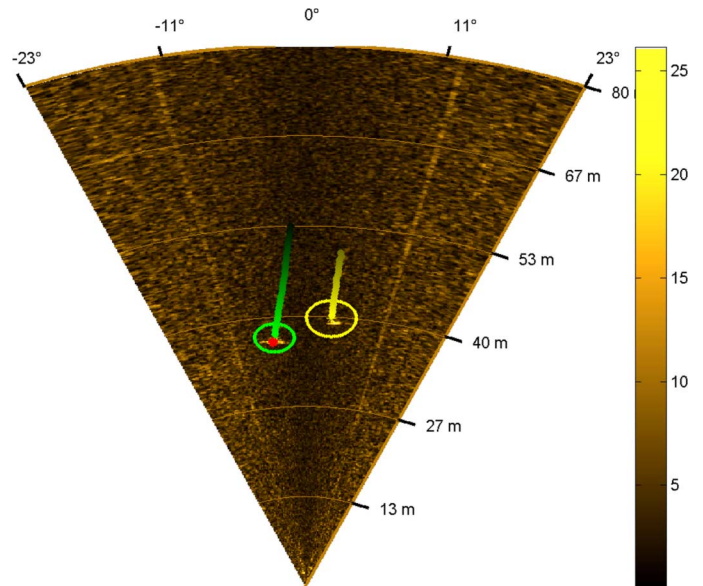


Fig. 8. Data of the 32nd ping of the first *Blueview P450* FLS sequence: tracking of two fake tethered mines in sight where $d_R = 2$ m: measurements in red, new filtered positions (at the center of the confidence ellipses) and trailing history of 31 previous positions in green (the steel mine) and of 19 previous positions in yellow (the rubber mine).

system), which should commonly be avoided for the obstacle avoidance task of the AUV [24].

An estimation of the z -coordinate of the obstacle can be achieved by minimizing, at each step of the sequence, the root mean squared deviation (RMSD) between measured sonar positions (deduced from measurements \mathbf{y}_k) and predicted positions given by a new process model derived from the previous one with an additional parameter related to the z -coordinate. These predicted points only depend on the first measurement \mathbf{y}_0 and successive predicted displacements according to vehicle motion and assumed obstacle altitude.

This new model is achieved by modifying the third equation of (2) to put the distance d_{st} in meters between the sonar and the target along the z -axis in the world reference frame in the place of the altitude h of the sonar (as seen in Fig. 4). This leads to

$$-\sin(\theta)m_r^x + \cos(\theta)\sin(\varphi)m_r^y + \cos(\theta)\cos(\varphi)m_r^z = d_{st} \quad (11)$$

and the previous f function becomes a new function \tilde{f} . For different possible values of $d_{st} \geq 0$, we compute all of the predicted positions as follows.

- In the first step, the predicted sonar position is computed by taking into account the first measurement \mathbf{y}_0 such that $\tilde{\mathbf{x}}_0 = \tilde{H} \mathbf{y}_0$ where $[\tilde{H}] = \begin{bmatrix} \Delta d & 0 \\ 0 & \Delta \delta \end{bmatrix}$.
- For steps k greater than 0, we compute predicted sonar positions $\tilde{\mathbf{x}}_k = \tilde{f}(\tilde{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}, d_{st})$ given the previous one $\tilde{\mathbf{x}}_{k-1}$.

Finally, we estimate the distance from the sonar to the target as

$$\hat{d}_{st} = \arg \min_{d_{st} \in [0, 1.2 \times h]} \sqrt{\sum_{k=0}^N \|\mathbf{x}_k - \tilde{\mathbf{x}}_k(d_{st})\|^2 / (N + 1)}$$

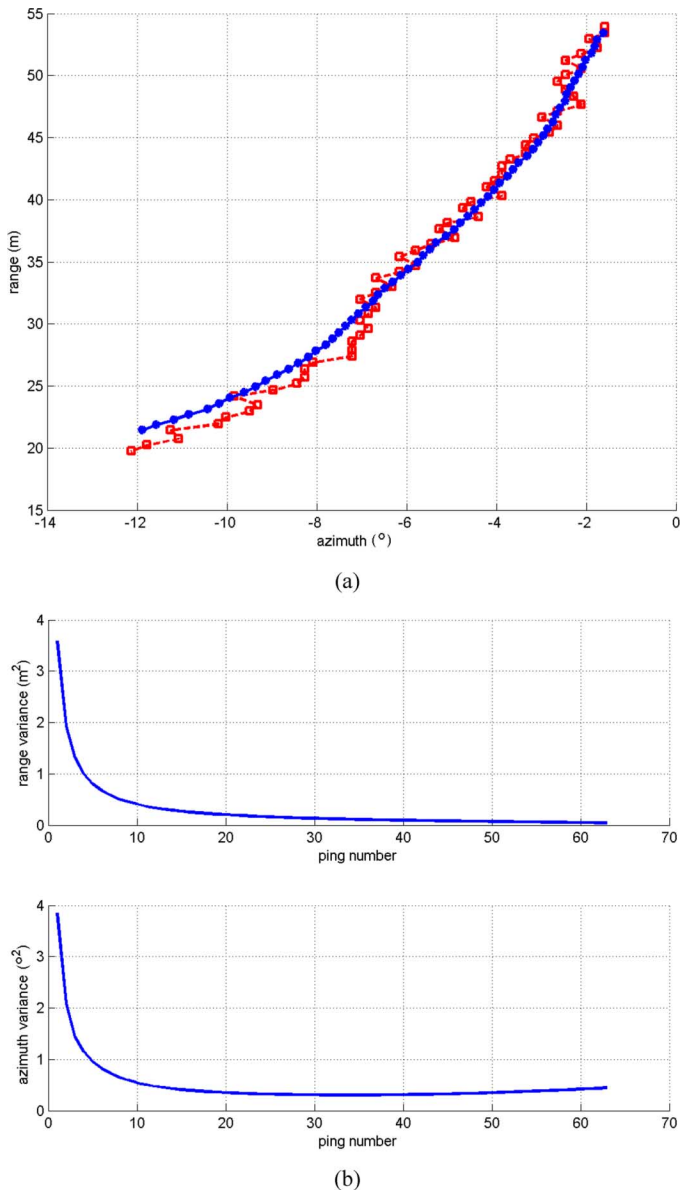


Fig. 9. (a) Filtering results for the longer track on the first *Blueview P450* FLS sequence (dashed red line: measurements; blue line: estimated trajectory). (b) A posteriori variances evolution along the sequence.

where $N + 1$ is the number of pings used, $\mathbf{x}_k = \tilde{H} \mathbf{y}_k$, and h is the altitude of the sonar at the first ping used. If the obstacle is below the vehicle, the maximum possible value of the distance from the sonar to the target is theoretically equal to the altitude of the sonar (case of a bottom object), but to prevent a wrong measured value, we choose a larger value, i.e., $1.2h$. If the object is not below but above the vehicle, this value is large enough to be considered as a safe distance.

Note that $\mathbf{x}_k - \hat{\mathbf{x}}_k = \tilde{H} (\mathbf{y}_k - \tilde{\mathbf{y}}_k) = \tilde{H} \tilde{\mathbf{v}}_k$, therefore we can say that the estimation consists in finding \hat{d}_{st} that minimizes the new innovation term $\tilde{\mathbf{v}}_k$.

Here the RMSD is computed from two sets of polar coordinates whose components are not comparable as such. As a consequence, to make the RMSD similar to a distance measure between the two tracks, the azimuth component of sonar positions has to be multiplied by the corresponding range component so as to have the arc length instead.

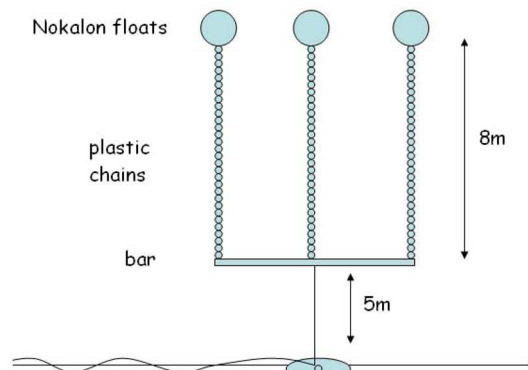


Fig. 10. Obstacle of the third sequence: a large thin obstacle made of a bar supporting chains and floats.

VII. EXPERIMENTAL RESULTS

A. Data Description

1) Obstacle Avoidance Sonar Data:

a) *Redermor AUV*: The *Redermor* vehicle is an experimental platform developed by the GESMA (Groupe d'Etudes Sous-Marines de l'Atlantique). In its last release, the vehicle is equipped with an obstacle avoidance sensor that consists of a 240-kHz *Reson Seabat 8101* FLS [25]. The system can obtain a beamformed image over a 15° (vertical) \times 60° (horizontal) sector with an across-track sampling rate $\Delta\delta$ equal to 1.5° and an along-track sampling rate Δd equal to 0.05 m. The sonar has been oriented 15° down from the horizontal plane (shown in Fig. 3).

To test the capability of the *Redermor* vehicle to react when obstacles are encountered on its way, GESMA organized experimental trials in April 2006 in Douarnenez Bay, France, named DETection et EVitement d'OBstacles (DEVITOBS).

b) *Daurade AUV*: The *Daurade* vehicle is a multipurpose experimental AUV for rapid environment assessment (REA) applications. It has been developed in the context of a project conducted by the Service Hydrographique et Océanographique de la Marine (SHOM) in cooperation with GESMA. The obstacle avoidance system is a *Blueview P450* FLS [26]. The system can obtain a beamformed image over a 15° (vertical) \times 45° (horizontal) sector with an across-track sampling rate $\Delta\delta$ equal to 0.18° and an along-track sampling rate Δd equal to 0.17 m. The sensor looks straight ahead (see Fig. 4).

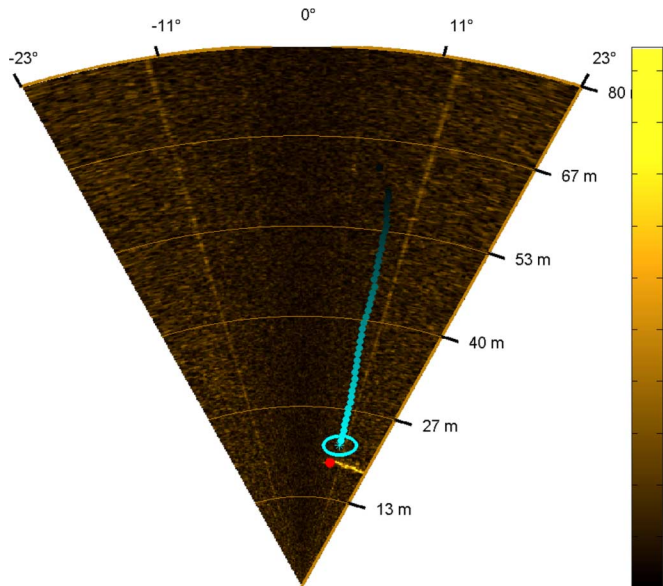


Fig. 11. Data of the 46th ping of the second *Blueview P450* FLS sequence: tracking of a large floating obstacle where $d_R = 5$ m: measurement in red, new filtered position (at the center of the confidence ellipse) and trailing history of 45 previous positions in cyan.

Data used in this paper have been collected by GESMA during two different sea trials near Brest, France, in March 2009 and April 2010.

2) *Navigational Data*: The two previous vehicles are fitted with different navigation systems that give us the following information:

- the vehicle velocity in relation to the seafloor, i.e., $\mathbf{V}_r = (v_r^x \ v_r^y \ v_r^z)^T$;
- the vehicle orientation (roll, pitch, and yaw) and its acceleration in relation to the world reference frame, i.e., the Euler angles (φ, θ, ψ) and their derivatives;
- the latitude and longitude coordinates of the vehicle;
- the altitude h of the vehicle from seafloor and also its depth from sea surface.

a) *Redermor AUV*: The Redermor AUV equipped with the *Reson Seabat 8101* uses a dead reckoning algorithm to compute the geographical position of the vehicle. The bottom velocity was obtained by a 300-kHz RDI DVL, and the vehicle orientation in the world reference frame was delivered by the OCTANS Attitude and Heading Reference System from iXSea (France). The dead reckoning process is initialized with a Wide Area Augmentation System (WAAS) Global Positioning System (GPS) receiver position.

b) *Daurade AUV*: The *Daurade* AUV equipped with the *Blueview P450* is built by ECA Robotics (Toulon, France) and uses the iXSea PHINS Inertial Navigation System. The PHINS system is directly fed by other navigation sensors, namely, a 300-kHz RDI DVL, a Paroscientific Depth Sensor, and a WAAS GPS receiver, and computes geographical position and velocity of the vehicle using Kalman filtering.

B. Obstacle Detection

As explained in Section II, an obstacle leads to a disturbance of the local statistics of the sonar image. To localize it, the sonar

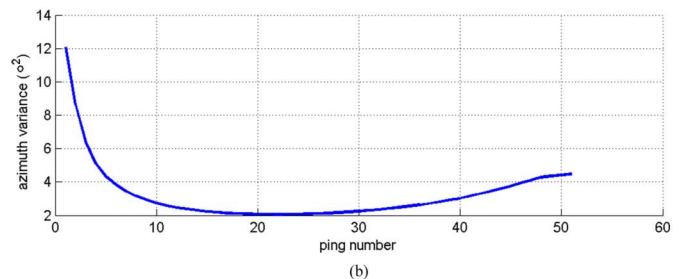
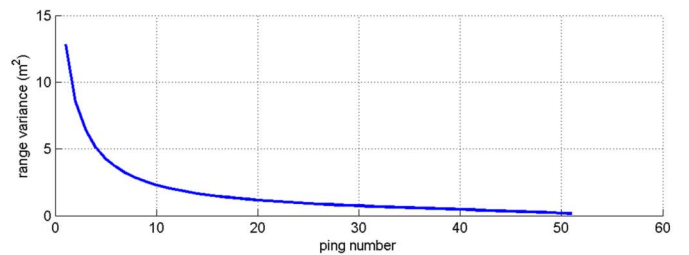
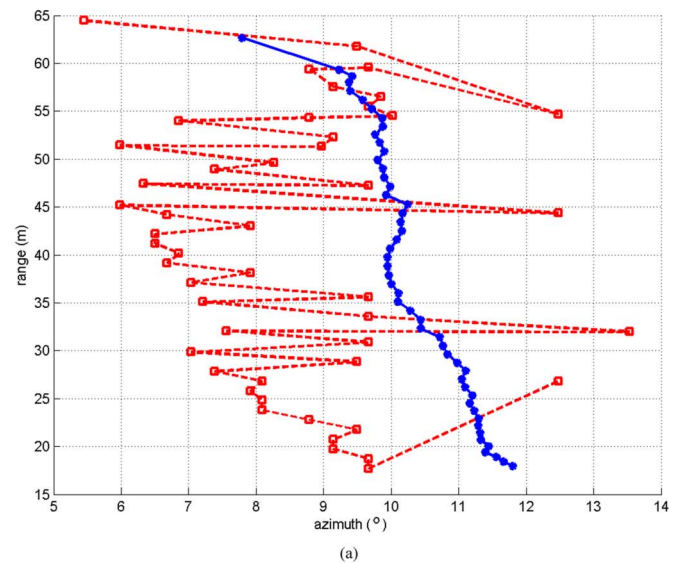


Fig. 12. (a) Filtering results on the second *Blueview P450* FLS sequence (dashed red line: measurements; blue line: estimated trajectory). (b) *A posteriori* variances evolution along the sequence.

image is first divided into snippets that consist of rectangular patches corresponding to an area of 5 m by 5 m with an overlap of 50%. For each snippet, we first compute the coefficient of variation c_v . For a Rayleigh-distributed variable, this ratio is constant and equal to 0.52 (see Section II) but this value can differ slightly for some systems like the *Reson Seabat 8101* FLS used in this paper, for which the revised ratio was estimated at 0.55. This difference can be explained by specific operational conditions that induce a better modeling of speckle by a Weibull law instead of a Rayleigh law [14]. Let c_v^{th} denote this theoretical value.

The test consists in comparing the local value of c_v to the theoretical value c_v^{th} . In other words, we consider that an object is present in the area if: $|c_v - c_v^{\text{th}}| > 0.1$. In that case of detection, the value of the pixels of the corresponding snippet is set to the mean pixel value (of the snippet). Otherwise, the pixels of the corresponding snippet are set to zero. By thresholding this image of overlapped snippets with a level of 80% of the maximum pixel value in this image, we can see areas of the image where pixels do not follow a Rayleigh distribution,

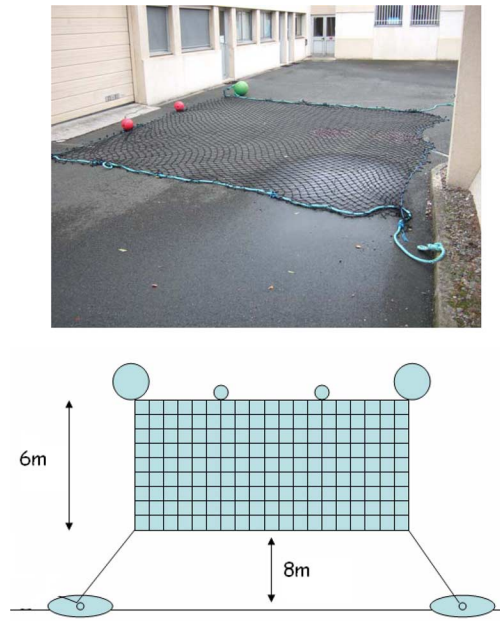


Fig. 13. Obstacle of the last sequence: a net with a large mesh.

i.e., where an object has potentially been detected. The value of the previous threshold is a compromise between a lower value that would lead to an aggregation of detections and false alarms in the same area and a higher value that would result in lower detection rates and missing of faint objects. Last, for each of these detected areas, the pixel that had the maximum level in the original image is defined as the detection point (or measurement) and can initiate a Kalman filter. Fig. 5 gives an example of this process on a *Reson Seabat 8101* FLS image.

C. Ground Obstacle Avoidance

Figs. 6, 8, 11, and 14 are screenshots of a single ping where we can see one or two tracks with a trailing history from the previous filtered positions. Around the current predicted position, the 95% confidence ellipse that is a function of the diagonal elements of the state covariance matrix $P_{k/k}$ [27] is plotted. Whereas the target localization is more accurate, the length of the ellipse axes decreases; see Figs. 6, 8, 11, and 14 or videos downloadable from www.ieeexplore.ieee.org for different sizes of ellipse. Also during an image sequence some tracks can be initiated but interrupted later and hence, are not visible on a single ping.

The first sequence was achieved by *Reson Seabat 8101* FLS where a shipwreck lies on the seafloor (see screenshot shown in Fig. 6, where a large echo area is followed by a large shadow area). As explained in Section IV-C, a large value fixed at 15 m for d_R is necessary to fuse all possible detected points relative to the same obstacle that is the large echo of the wreck. Because of that, a single track is initiated and updated with an average measurement. Fig. 7 shows a good filtering despite some noisy measurements at the end of the sequence. Variances, i.e., the diagonal elements of the state covariance matrix $P_{k/k}$, decrease monotonically toward a small value according to the accuracy of navigational data. Nevertheless, we can notice at the end a slight drawback for the azimuth variance that is probably due to a less accurate localization of the target in azimuth.

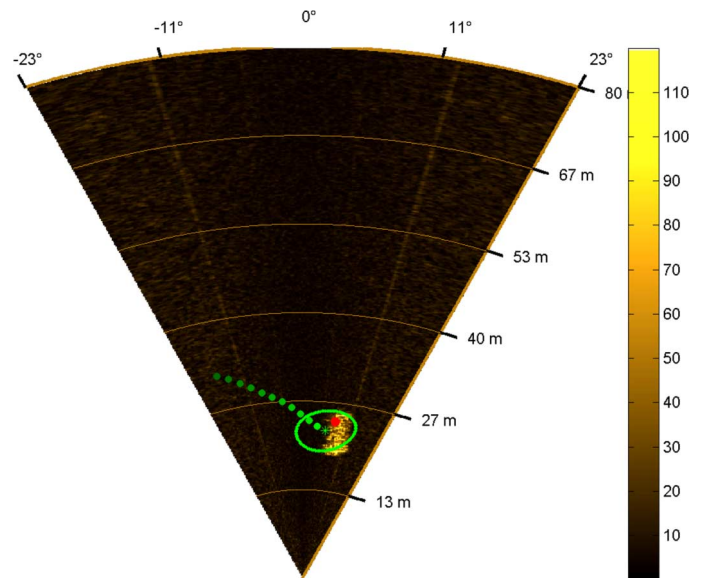


Fig. 14. Data of the 19th ping of the last *Blueview P450* FLS sequence: tracking of the net where $d_R = 2$ m: measurement in red, new filtered position (at the center of the confidence ellipse) and trailing history of 11 previous positions in green.

The second sequence is achieved by *Blueview P450* FLS where two fake tethered mines (see screenshot shown in Fig. 8) are in sight: a rubber one (yellow track in Fig. 8) was floating at about 15 m over the seafloor and had a diameter equal to 1 m (but probably smaller due to the pressure), another steel one (green track in Fig. 8) was floating at about 13 m and had a diameter equal to 0.75 m. These data were gathered in April 2010 near Brest, France. Sonar altitude was about 15 m above the seabed. The rubber mine is not tracked very far because of weaker echo level compared to the other. This is a consequence of the threshold based on the maximum level of the image of detected areas during the detection step (recall Section VII-B). The evolution of the variances is showed in Fig. 9(b). As the obstacle approaches (not visible at the step shown in Fig. 8, but can be seen in the corresponding downloadable video file), we can observe for this specific case of a tethered mine that Kalman filtering does not stop, however, it does underestimate the obstacle velocity. This is a consequence of the assumption of a ground obstacle. This case is discussed in Section VII-D.

The third sequence is also achieved by *Blueview P450* FLS where a large thin obstacle made of a bar supporting chains and floats is floating over the seafloor shown in the top part of Fig. 10. These data were gathered in March 2009, also near Brest, France. Sonar altitude was about 15 m above the seafloor. A track is shown in Fig. 11 with d_R of 5 m according to the size of the obstacle. A smaller value of d_R , such as 2 m for instance, would induce more than a single track on this obstacle. According to the complexity of this obstacle, it is impossible to know which part of it is detected at a given ping but, by considering the measurement variation in Fig. 12 (also visible in the video downloadable from www.ieeexplore.ieee.org), it is clear that it is not always the same part. The evolution of the variances shown in Fig. 12 is approximately the same as that of the shipwreck sequence with a more difficult convergence for the azimuth prediction. As in previous cases, Kalman filtering underestimates the obstacle velocity.

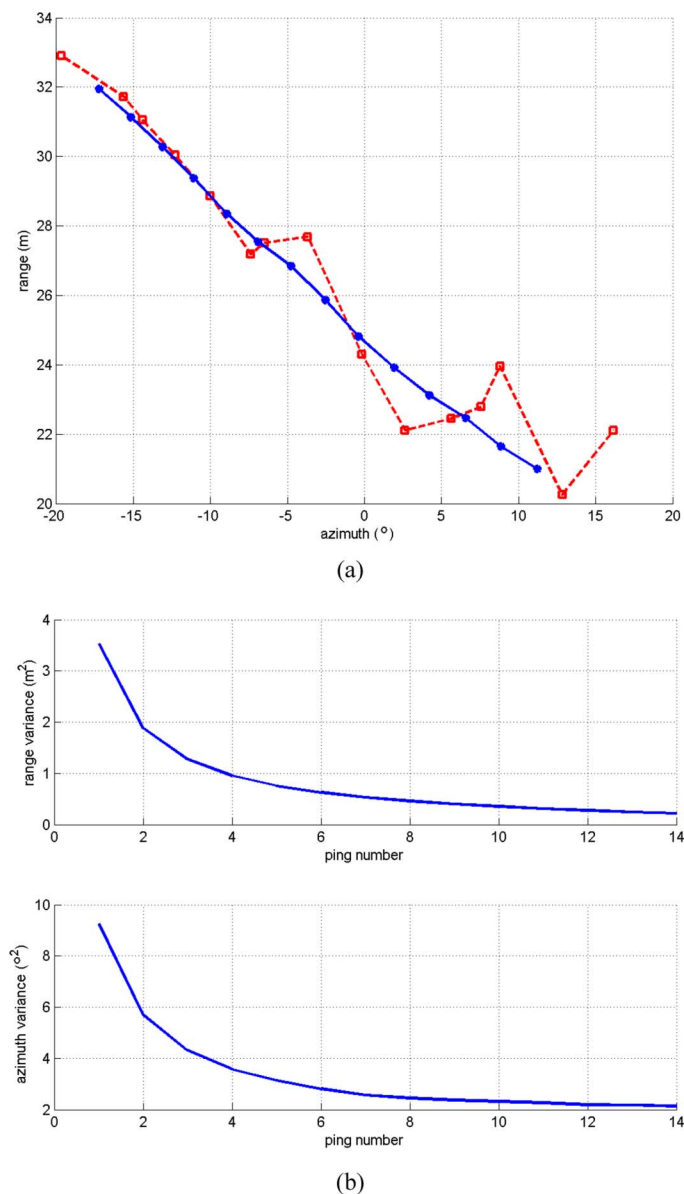


Fig. 15. (a) Filtering results on the last *Blueview P450* FLS sequence (dashed red line: measurements; blue line: estimated trajectory). (b) *A posteriori* variances evolution along the sequence.

The last sequence is obtained using the *Blueview P450* FLS, and shows images acquired along a turn of the vehicle, i.e., with large and fast attitude parameters variation. That is the reason that the obstacle is crossing from the left to the right of the screen. This obstacle, a net with a large mesh, is shown on Fig. 13. As was the case for the previous obstacle, these data were gathered in March 2009 near Brest, France. The track shown in Fig. 14 is obtained with d_R of 2 m. Both variances in Fig. 15 decrease monotonically, which proves a good performance of the Kalman filter even in this special case.

D. Relative Obstacle Altitude Estimation

In the following, we will call the experimental measurements as those given by the goodness-of-fit test described in Section II, and the manual measurements as those we estimated visually, i.e., by hand on each frame.

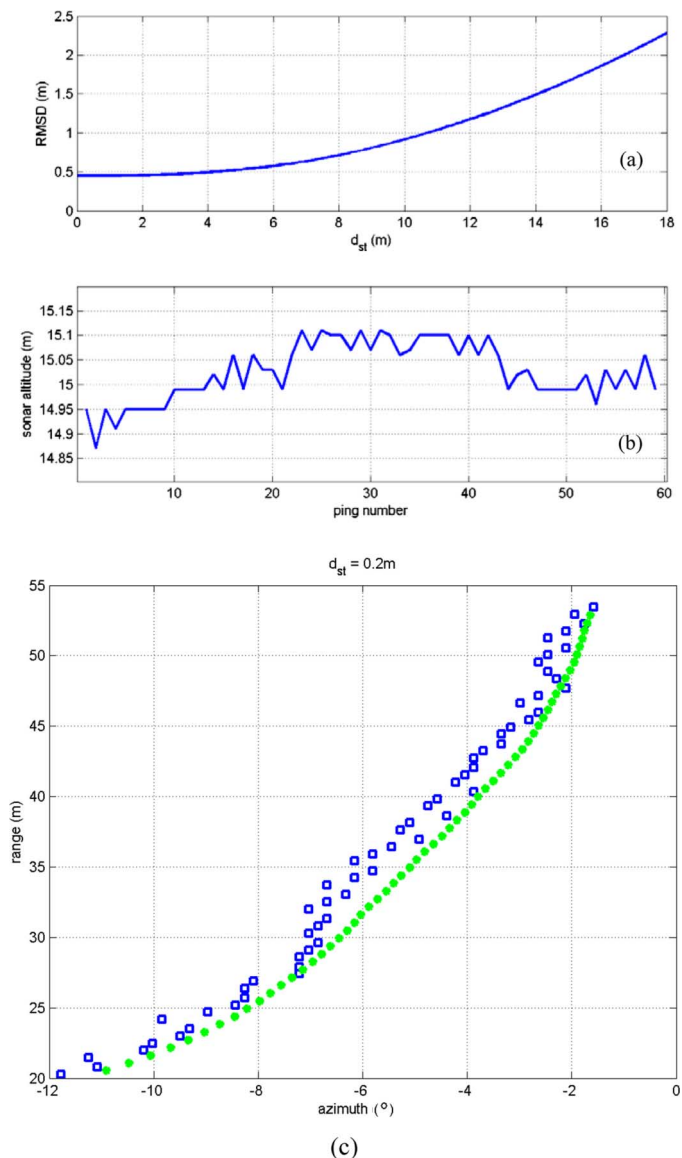
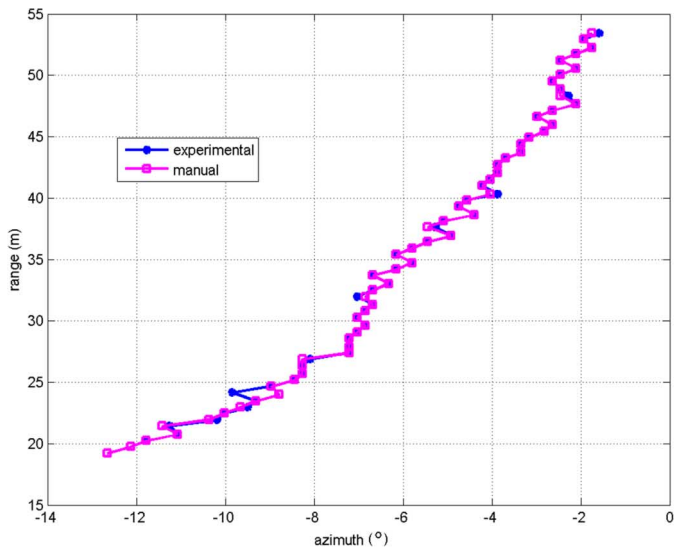
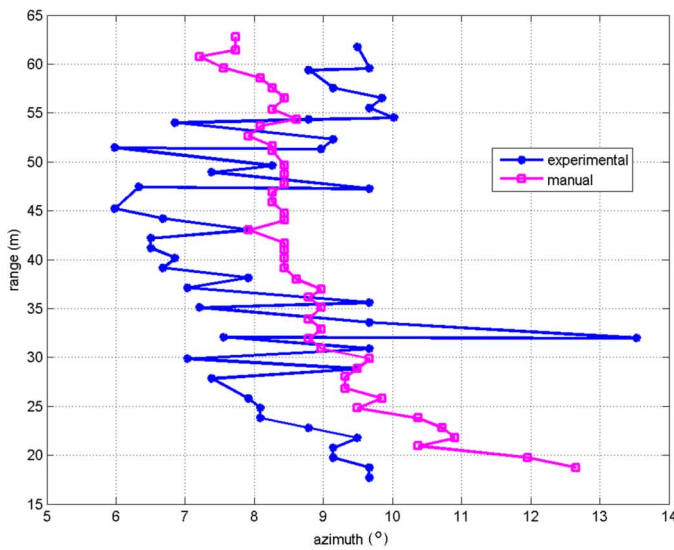


Fig. 16. Relative obstacle altitude estimation for track seen in Fig. 9: (a) RMSD versus d_{st} ; (b) sonar altitude versus ping number; (c) predicted points (green data points) and measurements (blue squared data points) for $d_{st} = 0.2$ m.

The *Daurade* sea trials in April 2010 were designed to assess the performance of the estimation method of the length of the tether of a moored mine proposed in Section VI. Two tethered spherical targets were deployed: a rubber one with a tether of 15 m and a steel one with a tether of 13 m. Fig. 16 shows the results concerning the second target. Fig. 16(a) and (b) shows a sonar altitude about 15 m and the RMSD minimization that gives an estimated distance \hat{d}_{st} of 0.2 m. That was approximately the case (expected value about 1.5 m): GESMA was testing avoidance capability in the case of an obstacle in the direction of travel of the vehicle. Predicted points given by the estimation method and experimental measurements points are plotted in Fig. 16(c) in case of $d_{st} = 0.2$ m, that is to say, when these two sets of points are the closest to each other. In this case, as the target is punctual, the set of experimental measurements is very similar to the set of manual measurements, as shown in Fig. 17(a). That is why the estimation leads to the same estimated value for d_{st} , whatever the set of measurements is.



(a)



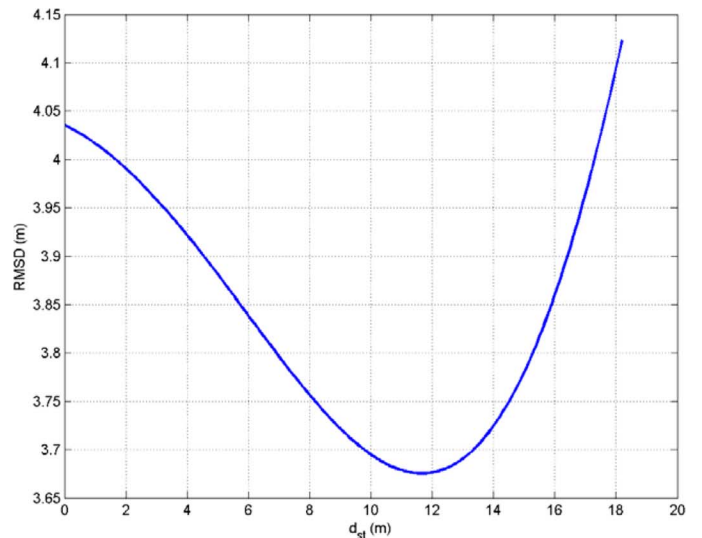
(b)

Fig. 17. Comparison between the set of experimental measurements (blue line) and the set of manual measurements (magenta line with squared data points) on *Blueview P450* FLS sequences: (a) data gathered in April 2010, (b) data gathered in March 2009.

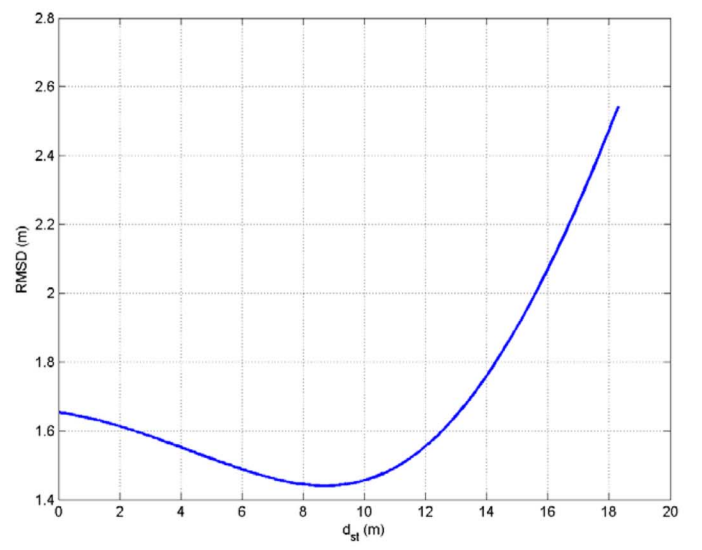
This task was more difficult for data gathered in March 2009 on a large obstacle seen in Fig. 10. Indeed, we can see in Fig. 18 that the estimated distance \hat{d}_{st} (11.7 m) based on experimental measurements differs from the one (8.7 m) based on manual measurements. This is probably due to noisy experimental measurements compared to manual measurements, as shown in Fig. 17(b). Moreover, as the three Nokalon floats are not clearly visible, we can assume that the obstacle was pushed down by current forces. Under this assumption, the detected echo area does not represent the Nokalon floats or the chains or the bar but probably a mix of them that leads to a value of 8.7 m between the obstacle and the sonar along the z -axis.

VIII. CONCLUSION

This paper has presented a framework for performing robust multitarget tracking, using navigational data. The main contribu-



(a)



(b)

Fig. 18. Relative obstacle altitude estimation for track seen in Fig. 11: (a) RMSD versus d_{st} based on experimental measurements; (b) RMSD versus d_{st} based on manual measurements.

tion of this work is the process modeling on which the Kalman filter is based. A Kalman filter is initialized every time a new target appears. The predicted position of the target is carried out by the state equation derived from the vehicle model and that takes navigational data as inputs. The nonlinearity of the state model was circumvented by performing an unscented transform. When a new measurement is available, it is put in the correction phase to adjust the target position. Multitarget tracking is achieved by a data association procedure that includes a validation test and a validation gate based on the innovation term of the Kalman filter. The robustness of the tracking has been shown on real data from two different FLSs as well as various obstacles. An interesting related contribution is the deferred estimation of the relative altitude of the obstacle based on the revised process model. This parameter is often an unavailable but important piece of information for obstacle avoidance purposes.

Future work will focus on obstacle characterization and classification. The objective is twofold: first, to discard benign obstacles such as a harmless school of fish, and second, to induce an appropriate avoidance maneuver. In addition, we aim to consider extending this work to the case of moving obstacles.

REFERENCES

- [1] G. N. Williams, G. E. Lagace, and A. Woodfin, "A collision avoidance controller for autonomous underwater vehicles," in *Proc. IEEE Symp. Autom. Underwater Veh. Technol.*, 1990, pp. 206–212.
- [2] J. T. Cobb, B. Schulz, and G. Dobeck, "Forward-looking sonar target tracking and homing from a small AUV," in *Proc. MTS/IEEE OCEANS Conf.*, Washington, DC, Sep. 17–23, 2005, DOI: 10.1109/OCEANS.2005.1639948.
- [3] D. M. Lane, M. J. Chantler, and D. Y. Dai, "Robust tracking of multiple objects in sector-scan sonar image sequences using optical flow motion estimation," *IEEE J. Ocean. Eng.*, vol. 23, no. 1, pp. 31–46, Jan. 1998.
- [4] E. Trucco, Y. R. Petillot, I. T. Ruiz, K. Plakas, and D. M. Lane, "Feature tracking in video and sonar subsea sequences with applications," *Comput. Vis. Image Understand.*, vol. 79, pp. 92–122, 2000.
- [5] Y. Petillot, I. T. Ruiz, and D. M. Lane, "Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar," *IEEE J. Ocean. Eng.*, vol. 26, no. 2, pp. 240–251, Apr. 2001.
- [6] D. Clark, I. T. Ruiz, Y. Petillot, and J. Bell, "Particle PHD filter multiple target tracking in sonar image," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 409–416, Jan. 2007.
- [7] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.
- [8] N. O. Handegard and K. Williams, "Automated tracking of fish in trawls using the DIDSON (Dual frequency Identification SONar)," *ICES J. Mar. Sci.*, vol. 65, pp. 636–644, 2008.
- [9] S. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Reading, MA: Artech House, 1999, ch. 6 and 13.
- [10] L. Henriksen, "Real-time underwater object detection based on an electrically scanned high-resolution sonar," in *Proc. Symp. Autom. Underwater Veh. Technol.*, Cambridge, MA, Jul. 19–20, 1994, pp. 99–104.
- [11] J. S. Lee, "Statistical modeling and suppression of speckle in synthetic aperture radar images," in *Proc. Int. Geophys. Remote Sens. Symp.*, Ann Arbor, MI, May 18–21, 1987, pp. 1331–1339.
- [12] F. Schmitt, L. Bonnaud, and C. Collet, "Contrast control for sonar pictures," *Proc. SPIE—Signal Image Process.*, vol. 2847, pp. 70–82, Aug. 1996.
- [13] R. Caulett, *Statistics in Research and Development*. London, U.K.: Chapman & Hall, 1991, p. 19.
- [14] F. Maussang, "Traitement d'images et fusion de données pour la détection d'objets enfous en acoustique sous-marine," Thèse, de l'Université Joseph Fourier—Grenoble 1, Grenoble, France, Nov. 2005.
- [15] I. Quidu, Y. Dupas, and L. Jaulin, "Theoretical optical flow for target position prediction on FLS images," in *Proc. Underwater Acoust. Meas. Conf.*, Nafplion, Greece, Jun. 21–26, 2009 [Online]. Available: <http://www.uam-conferences.org/index.php/past-proceedings>
- [16] S. J. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proc. 11th Int. Symp. Aerosp./Defense Sens. Simul. Controls*, 1997, Multi Sensor Fusion, Tracking and Resource Management II, pp. 182–193.
- [17] R. Van Der Merwe and E. Wan, "Sigma-point Kalman filter for probabilistic inference in dynamic state-space models," in *Proc. Workshop Adv. Mach. Learn.*, Montreal, QC, Canada, Jun. 2003 [Online]. Available: <http://cslu.cse.ogi.edu/publications/ps/merwe03a.ps.gz>
- [18] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992, p. 350.
- [19] Y. Bar-Shalom and T. E. Fortmann, "Mathematics in science and engineering," in *Tracking and Data Association*. San Diego, CA: Academic, 1988, vol. 179, ch. 1.
- [20] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE J. Ocean. Eng.*, vol. OE-8, no. 3, pp. 173–184, Jul. 1983.
- [21] B. Jida, R. Lherbier, M. Wahl, and D. Hamad, "Détection automatique dans les données télémétriques et association spatio-temporelle par réseaux bayésiens—Application au suivi de véhicules," presented at the MajecSTIC Conf., Lorient, France, Nov. 22–24, 2006.
- [22] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, pp. 451–460, 1975.

- [23] R. K. Mehra and J. Peschon, "An innovations approach to fault detection and diagnosis in dynamic systems," *Automatica*, vol. 7, pp. 637–640, 1971.
- [24] C. D. Loggins, "A comparison of forward-looking sonar design alternatives," in *Proc. MTS/IEEE OCEANS Conf.*, Honolulu, HI, 2001, pp. 1536–1545.
- [25] Reson, "SeaBat," [Online]. Available: www.reson.com/products/seabat/
- [26] BlueView Technologies, "2D multibeam imaging sonar," [Online]. Available: www.blueview.com/2d-Imaging-Sonar.html
- [27] N. R. Draper and H. Smith, *Applied Regression Analysis*, 2nd ed. New York: Wiley, 1981, pp. 108–109, ISBN: 0471029955.



Isabelle Quidu was born in Pontivy, France, in 1974. She received the Ph.D. degree in electronics from the University de Bretagne Occidentale, Brest, France, in 2001.

From 2002 to 2005, she was a Research Engineer at Thales Underwater Systems in the field of mine warfare. Since 2005, she has been an Associate Professor at the Ecole Nationale Supérieure de Techniques Avancées (ENSTA) Bretagne (formerly ENSIETA), Brest, France, and a member of the Ocean Sensing and Mapping (OSM) team. She is also a member of Lab-STICC UMR CNRS 6285. Her research interests are in detection and classification of underwater targets in sonar images.



Luc Jaulin was born in Nevers, France, in 1967. He received the Ph.D. degree in automatic control from the University of Orsay, Orsay, France, in 1993.

He has been a Professor of Robotics at the Ecole Nationale Supérieure de Techniques Avancées (ENSTA) Bretagne, Brest, France, since 2004. He is part of the Ocean Sensing and Mapping (OSM) team and his research is in underwater robotics using interval methods and constraint propagation. He is a member of Lab-STICC UMR CNRS 6285.



Alain Bertholom was born in Quimper, France, in 1971. He graduated in electronics from the French Department of Defense in 1993.

He joined the Naval Robotics Department, Groupe d'Etudes Sous-Marines de l'Atlantique (GESMA), Brest, France, in 1998. He was involved in real-time embedded software development and sea trials for the *Redermor I* and *II* autonomous underwater vehicles (AUVs). He also worked on navigation purposes. Since September 2009, he has been the Technical Manager of the Covert REA *Daurade*

AUV.



Yann Dupas was born in Montereau, France, in 1981. After finishing his first studies at the Autun Military College, Autun, France, he became a military student at the the Ecole Nationale Supérieure de Techniques Avancées (ENSTA) Bretagne, Brest, France, in September 2001.

He became a Military Engineer in September 2005. He joined the Naval Robotics Department, Groupe d'Etudes Sous-Marines de l'Atlantique (GESMA), Brest, France, and started working on obstacle detection and avoidance for autonomous underwater vehicles (AUVs), mission planification, and sea trials organization with the GESMA AUV *Redermor II*. In 2008, he became a Technical Manager for the new REA project and achieved acquisition of the new covert REA AUV *Daurade*. In September 2009, he became a Project Manager for the covert REA project at the Service Hydrographique et Océanographique de la Marine (SHOM), Brest, France.

Annexe 6 : Aide à la reprise de vue des AUV

Cette annexe est une copie de l'article [KAR15b] publié dans la revue *IEEE Transactions on Geoscience and Remote Sensing* en 2015 suite au projet MUSE avec la société iXblue (cf. paragraphe 2.6.2).

Automatic Sea-Surface Obstacle Detection and Tracking in Forward-Looking Sonar Image Sequences

Imen Karoui, Isabelle Quidu, and Michel Legris

Abstract—Automatic sea-surface object detection and tracking for safe autonomous underwater vehicle and submarine surfacing is a critical issue in relation to the accidents reported in the last decades. Here, we propose an efficient tool to detect and track sea-surface obstacles by processing forward-looking sonar images. The proposed method can detect either still or moving objects with and without wake. For each image sequence, a sequential procedure is proposed to detect various obstacle signatures. Then, target positions and velocities are estimated in Cartesian coordinates using the debiased converted measurement Kalman filter and the joint probabilistic data association filter. Detection and tracking stages exchange information in order to reduce the number of false alarms. Promising results are obtained using real data collected at sea with various objects and scenarios.

Index Terms—Multitarget tracking, obstacle avoidance, sea-surface obstacle detection.

I. INTRODUCTION

OBSTACLE avoidance sonar systems are usually used in a forward-looking mode to detect and avoid obstacles either floating in the sea column or lying on the sea bottom [1], [2]. Automatic detection and tracking of sea-surface obstacles from forward sonar images is undoubtedly an interesting application that may help enable safe submarine and autonomous underwater vehicle (AUV) surfacing. In fact, in the last decades, many accidents have occurred. A famous accident was the collision between the Japanese vessel *Ehime Maru* and the submarine USS *Greeneville* in 2001. Within minutes of the collision, *Ehime Maru* sank. Nine of her crewmembers were killed, including four high school students.

To our knowledge, no former work has been published on safe AUV surfacing, at least in open literature. However, some surface object detection and classification methods exist for harbor surveillance. However, most of these methods are based on passive acoustic detection [3], [4]. Active sonar systems dedicated to the ocean surface characterization mainly focus

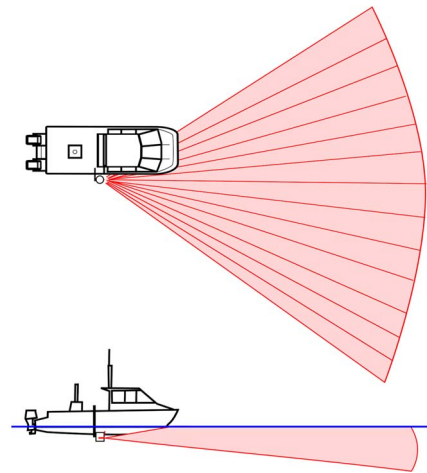


Fig. 1. Sonar configuration.

on experimental and theoretical analysis of ship wake structure through the study of bubble dynamics and backscattering [5]–[7], and only a few recent methods have been proposed for automatic ship detection using active sonar images [8]–[10]. These methods are mainly suited to a static configuration for harbor surveillance purposes, and they concern single ship detection.

Here, we propose a pattern analysis procedure for automatic detection and tracking of multiple surface objects from forward-looking sonar images. The proposed method can detect still and moving vessels with or without self-noise and wake.

This paper is organized as follows. In Section II, the experiment setup is described. In Section III, the different target signatures and the associated detection methods are detailed. In Section IV, the association and tracking procedure are depicted. In Section V, some results are given using real data gathered at sea. In Section VI, the conclusion and future work are outlined.

II. EXPERIMENTAL SETUP

Obstacle detection is usually accomplished using forward-looking sonar. Such a system is an active sonar that emits a short acoustic pulse forward on a large horizontal sector (of about 120°) and a medium vertical sector (from 15° to 20°) (see Fig. 1). Acoustic echoes are recorded on a horizontal linear array of transducers. A beamforming process [11] (consisting in phase shifts and amplitude summation) is applied on the recorded data to sort the acoustic energy in time (related to the

Manuscript received March 12, 2014; revised July 16, 2014 and November 21, 2014; accepted January 23, 2015. This work was supported by Direction Générale de l'Armement through the "Rapid" Sponsorships.

The authors are with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Lab-STICC UMR CNRS 6285, 29806 Brest, France, and also with SPARTE iXBlue & ENSTA Bretagne Joint Laboratory, 29806 Brest, France (e-mail: imen.karoui@ensta-bretagne.fr; isabelle.quidu@ensta-bretagne.fr; michel.legris@ensta-bretagne.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2015.2405672



Fig. 2. (a) ENSTA survey boat and (b) inflatable boat involved in the experiments.

distance of the echoes) and angle of arrival (named acoustic beams). The envelope of the resulting signals is then displayed as an image (see Fig. 3). That image could be represented in polar coordinates (range and bearing) or Cartesian coordinates (the acoustic image then appears in a fan shape).

Here, we use the RESON 8101 sonar with the image sector mode. Experiments were conducted by ENSTA Bretagne in Brest, Marina Du Château, in June 2013. The RESON 8101 forward-looking sonar was mounted on the ENSTA Bretagne survey boat at roughly 2 m below the sea surface. The sonar looks forward but intercepts the sea surface (see Fig. 1).

This sonar works at a central frequency of 240 kHz and a bandwidth of approximately 15 kHz. The horizontal transmitting aperture is wide and covers 150° horizontally. The vertical aperture is about 15° . However, the recorded data only covers a 60° -wide sector in the broadside direction due to electronic bandwidth constraints. The beamwidth for each of the received beams is 1.5° , and the range resolution is about 0.05 m.

Our survey boat was moving slowly in the zone under investigation, and data were recorded according to various scenarios with different sea-surface objects: static man-made objects laid on the sea surface, and moving ferries and sailboats arriving at or leaving the military harbor nearby. Some scenarios include a cooperative inflatable boat (see Fig. 2).

III. SEA-SURFACE TARGET DETECTION

In this paper, the sea-surface obstacles considered are man-made objects: buoys, boats (motorboats or sailboats), containers, etc. According to the object type and state (still or moving), the surface acoustic signature can be as follows:

- strong-intensity beam(s) in the sonar image, (radial strips in Cartesian representation of sonar data) due to the stationary ship self-noise. This kind of signature indicates the ship bearing (it is the strongest among all surface object signatures and the easiest to detect, see Fig. 3);
- a high-contrast-intensity feature in the case of noise-free objects such as buoys, sailboats, and static ships such as fishing vessels (see Fig. 4);
- some high-intensity lines due to the wake in the track of moving vehicles (see Fig. 4). The wake echo level strongly depends on parameters such as the sensor distance to the ship, the aspect angle, and the center frequency of the transmitted pulse [7].

A detection and tracking procedure is proposed whatever the signature. The proposed approach is depicted in the flowchart of Fig. 5. It is a hierarchical detection procedure in order to deal

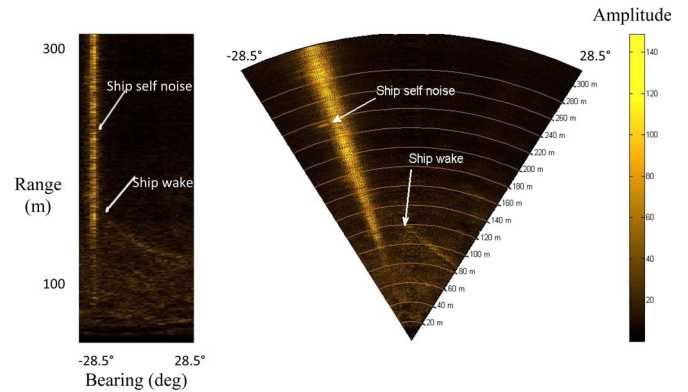


Fig. 3. Stationary-ship self-noise signature. (Left) Polar coordinates (x -axis corresponds to bearing and y -axis to range values). (Right) Cartesian coordinates.

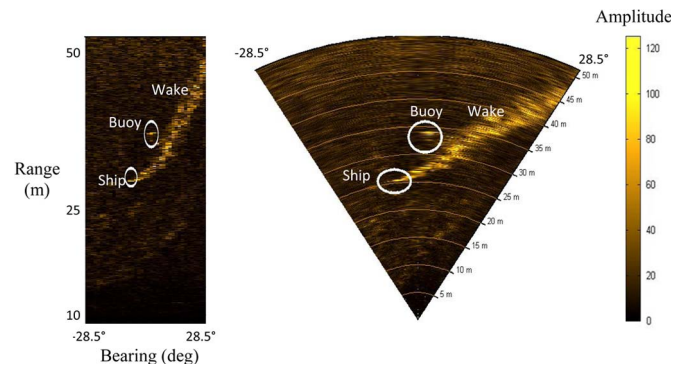


Fig. 4. Stationary-ship self-noise signature. (Left) Polar coordinates (x -axis corresponds to bearing and y -axis to range values). (Right) Cartesian coordinates.

with the various target signatures: stationary-ship self-noise, noise-free obstacles, and finally, wakes if they exist. In addition, as will be explained later on, object detection and tracking are carried out together: Information about the scene extracted from the last recurrences is taken into account to improve the detection rate for the image under process.

The proposed method is organized as follows: As the stationary-ship self-noise is strong and is spread on one or several adjacent beams, it could mislead the detection of other features. Therefore, we first try to detect this strongest signature. In the case of detection, the strong intensity beams corresponding to the ship bearing are ignored during the second step that consists in detecting the other types of obstacles. The third and final detection step consists in analyzing the area around every detected position in order to state if it corresponds to a reliable obstacle or to wakes only.

A. Detection of Ships With Self-Noise

The detection of the stationary-ship noise is often based on the evaluation of the mean echo level along each image beam. The beam of the highest energy is selected to give ship bearing [8]. As we assume neither the existence nor the uniqueness of vessel, we search for the more energetic beam(s) with respect to their neighbors. Instead of using the average of the beam amplitude as an indicator of ship self-noise presence,

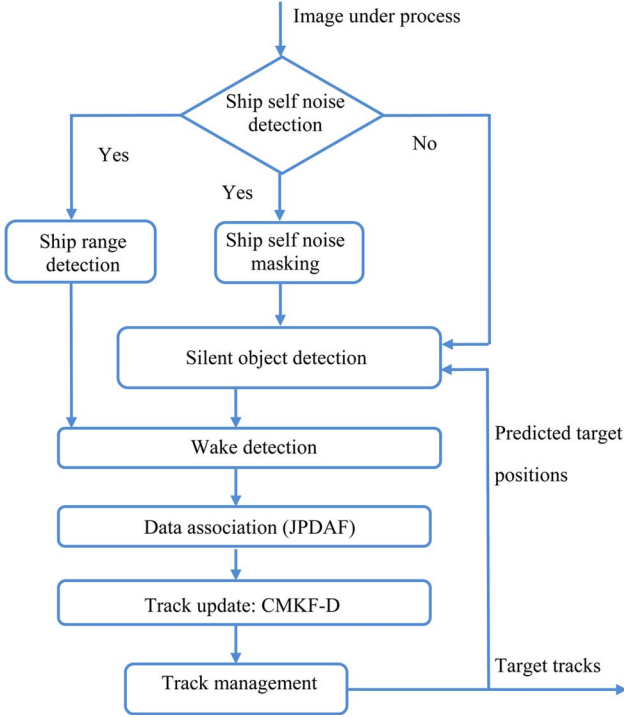


Fig. 5. Flowchart of the overall method.

we use the median amplitude value of each image beam that is more suitable for the multiple-target case. Indeed, the average of the high amplitude of a given beam may correspond to several targets without stationary noise and not to a single ship self-noise.

Formally, the indexes of the bearing measurements associated to ship self-noise at each image sequence \mathbf{I}_k are given by

$$\{\tilde{n}_k\} = \left\{ j \mid \mathbf{d}(j) > \frac{\xi}{2(n-n_g)} \left[\sum_{k=-n}^n \mathbf{d}(j-k) - \sum_{k=-n_g}^{n_g} \mathbf{d}(j-k) \right] \right\} \quad (1)$$

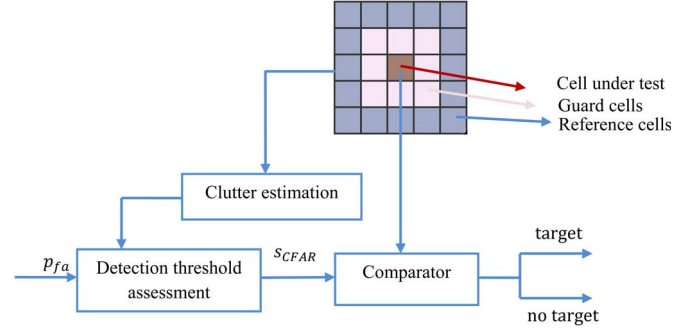
where $\mathbf{d}(j)$ is the median of the image beam j . ξ is a preselected threshold higher than one, which is here set to 1.5.

$2(n-n_g)$ stands for the number of adjacent beams used to estimate the detection threshold. The $2n_g$ closest neighbors are excluded from the threshold assessment because the signal is spread on several neighboring beams. This spreading is the result of two factors. First, no beamforming can achieve an ideal directivity pattern: a small part of the acoustic energy leaks on adjacent beams (sidelobe effects [11]). Second, the sampling interval is generally smaller than the beam resolution. For our sonar, we observed that a good value of n_g is

$$n_g = 3 \left\lfloor \frac{\Delta_\varphi}{\Delta\delta} \right\rfloor \quad (2)$$

where $\lfloor \cdot \rfloor$ is the smaller integer inside the brackets, Δ_φ stands for the beamwidth, and $\Delta\delta$ for the angular sampling rate (for the RESON 8101, $\Delta_\varphi = \Delta\delta = 1.5^\circ$).

Ship echoes are usually strong, and thus, ship range detection consists in finding the highest intensity pixels in the detected beams. Because of the time-varying gain or the automatic gain amplifier usually applied to sonar images, the ship noise


 Fig. 6. CFAR detector: p_{fa} is the probability of false alarm, and s_{CFAR} is the detection threshold.

level increases with range, and the target could be masked particularly if it occurs at far ranges. To reduce this effect, as in [8], we normalize the detected image beams by a moving window median filter as follows:

$$\begin{aligned} \tilde{\mathbf{I}}_k(m, n) &= \frac{\mathbf{I}_k(m, n)}{\text{med}(\mathbf{I}_k((m-n_b), n), \dots, \mathbf{I}_k((m+n_b), n))} \quad \forall n \in \{\tilde{n}_k\} \end{aligned} \quad (3)$$

where $\mathbf{I}_k(m, n) \equiv \mathbf{I}_k(m\Delta r, n\Delta\delta)$ stands for the image amplitude at polar coordinates $(m\Delta r, n\Delta\delta)$; Δr and $\Delta\delta$ are, the range and bearing sampling rates, respectively; $\text{med}(\cdot)$ stands for the median value of the quantity inside the brackets; and $2n_b + 1$ is the length of the moving window median filter. For our experiments, we set $n_b = 50$ (about 2.5 m).

We then pick the first ten strongest peaks of each normalized detected beam. If ship wakes are visible, their positions indicate ship ranges (see Fig. 3). In such cases, only the range value selected after the wake detection stage is kept (see Section III-C).

B. Detection of Noise-Free Targets

To our knowledge, there is no work dealing with the detection of ships without self-noise for active sonar images. In SAR imagery, the topic of ship detection has been extensively studied [12]. The most powerful detector widely used in SAR imagery for ship detection is the constant false-alarm rate (CFAR) algorithm with an adaptive threshold [13]. This threshold is computed for each image cell by taking into account the clutter power estimated from a set of reference cells surrounding the cell under investigation, as shown in Fig. 6.

Guard cells are the nearest neighbors of the cell under investigation and are not used in the local clutter estimation in order to ensure that pixels belonging to an extended target cannot be included in the clutter statistics estimation.

Adaptive radar detection for several clutter probability density functions (pdfs) (Gaussian, Weibull, K distribution, Pearson, etc.) have been widely studied in the literature [14]–[16].

Here, we deal with amplitude sonar images, and the distribution of pixel levels is usually modeled by a Rayleigh distribution [17]: $f(x) = (2x/b^2) \exp(-(x^2/b^2))$. We assume that clutter samples are independent and identically distributed, and we estimate the parameter b according to the maximum-likelihood method [15]: $\hat{b} = ((1/J) \sum_{j=1}^J x_j^2)^{1/2}$,

where J is the number of reference cells, and x_j the j th cell amplitude. The CFAR detector is then based on the following threshold [14]:

$$s_{\text{CFAR}} = \left[\left(p_{\text{fa}}^{-\frac{1}{J}} - 1 \right) \sum_{j=1}^J x_j^2 \right]^{\frac{1}{2}}. \quad (4)$$

For our data, two events can increase the false-alarm rate or induce target missing: the presence of wake and other targets in the immediate vicinity of pixels under test. Indeed, every interfering target or wake that appears in the reference window associated to the target under investigation may raise the CFAR threshold extemporaneously. To avoid a wrong estimation of the detection threshold, we allow information exchanges between the tracker and the different detectors. Practically, at the k th image of the sequence, we exclude every cell from the reference cells, which are used in the clutter parameter estimation (\hat{b}), belonging to the following:

- detected beams corresponding to ship self-noise if they exist;
- previously detected wakes;
- target positions predicted by the Kalman filter.

C. Wake Detection

Moving ships generate bubbly turbulent wake in their track. The turbulence can persist for several minutes and extend over hundreds of meters [9]. Commonly, wake width is five to ten times the ship's width, and it increases slowly with the distance in relation to the ship. The exact wake structure depends, however, on the ship propeller, hull design, ship speed, and maneuver [18].

When wake cell amplitudes are high enough to be detected by the CFAR detector, detections associated to the wake drastically increase the amount of false alarms and hence mislead ship localization and tracking.

Here, we do not need an exact wake detection and characterization. Our goal is only to detect the reliable position of the ship among all the detections generated by wakes. To do so, we first search for wake ends, and we assume that ship position is given by the wake endpoint with the highest echo level. However, for our experiments, as we work at quite short maximum ranges (50 m for a typical sequence for which the wake is visible), we notice that, for some cases, ship echo is the wake endpoint with the lowest energy. A pragmatic solution is to keep both wake ends if both of them are detected, and the spurious end could easily be removed after tracking.

The proposed wake end detection consists in analyzing the area around each detection within oriented strips as shown in Fig. 7. We assume that if the amount of detections for one orientation is much higher than the detection number in the other orientations, the detection is likely to be a wake end. In order to reduce the computation time, the search for wake ends is not carried out for each detection but only around cluster centers: closely spaced detections are gathered into clusters, and each cluster center is given by the averaged position of all detections belonging to the cluster.

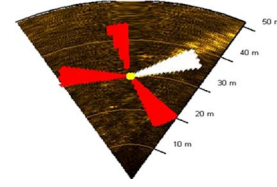


Fig. 7. Wake detection procedure for (yellow) a given detection. (Red) Three strips of 15-m length and about 20° width for three different orientations. (White) detected strip.

Formally, we jointly use two strip lengths l_1 and $l_2 = l_1/2$. We consider that a given detection could correspond to a wake end if these two conditions are satisfied.

- 1) The number of detections in one of its associated strip of length l_1 for a given orientation is higher than five times the mean detection number of all different strips of length l_1 , and this number exceeds n_w detections.
- 2) The number of detections in its associated strip of length l_2 in the same orientation is higher than five times the mean detection number of all different strips of length l_2 , and this number exceeds $2n_w/3$ detections.

Condition 2 avoids the detection of isolated detections near the wake by taking into account a high number of detections near the ship. The minimum detection number n_w ensures that the wake is dense. We set $n_w = 50$.

For our data, we noticed that turbulent wakes have a maximum width of 20° ; hence, we use strips of 20° in width, and we set $l_1 = 15$ m.

Many detections belonging to the same wake can satisfy these two conditions. The clustering of wake detections inside the same wake is carried out as follows: All pixels belonging to overlapped strips and having similar detected orientations (modulo 2π because we keep both wake ends) are clustered, and only the end position is kept.

All the other detections inside the detected wakes are removed from the measurement list.

IV. OBJECT ASSOCIATION AND TRACKING

A. Target Tracking

Once the detection is carried out and potential target coordinates are estimated, we aim at finding reliable targets and tracking them. Target tracking is performed in the Cartesian frame according to the classic near constant velocity dynamical model coded in matrices \mathbf{F} and \mathbf{G} of the following state equation:

$$\mathbf{x}_k^t = \mathbf{F}\mathbf{x}_{k-1}^t + \mathbf{G}\mathbf{v}_{k-1} \quad (5)$$

where

- $\mathbf{x}_k^t = [p_x, p_y, \dot{p}_x, \dot{p}_y], [p_x, p_y]$, and $[\dot{p}_x, \dot{p}_y]$ are the position and the velocity vectors in the xy plane relative to the sonar, respectively;

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} (\Delta t^2/2) & 0 \\ 0 & (\Delta t^2/2) \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}$$

and Δt refers to the time sampling;

— \mathbf{v}_{k-1} is a 2-D zero-mean independent and identically distributed white Gaussian noise with standard deviation σ_v for each component.

At ping k , measurements that originated from target t are related to the state vector \mathbf{x}_k^t according to the following measurement equation:

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k^t + \mathbf{w}_k \quad (6)$$

where \mathbf{H} is a known matrix, and \mathbf{w}_k is a known measurement noise.

Target positions are reported in polar coordinates (range and bearing) with respect to the sonar position. In order to deal with a linear measurement equation (6), we convert polar coordinates to Cartesian coordinates, and the bias due to the nonlinearity of measurement conversion is tackled by using the converted measurement Kalman Filter with debiased conversion (CMKF-D) [19].

Let us consider $\{(d_m, \delta_m)\}$ the polar measurements corrupted by a white and Gaussian noise with covariance matrix

$$\mathbf{R}_{\text{polar}} = \begin{bmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_\delta^2 \end{bmatrix}$$

where σ_d and σ_δ are the error standard deviations in range and bearing, respectively.

The conversion of the polar measurement to debiased Cartesian measurement is as follows:

$$\mathbf{y}_k^c = \begin{bmatrix} p_x^c \\ p_y^c \end{bmatrix} = \begin{bmatrix} d_m \cos(\delta_m) \\ d_m \sin(\delta_m) \end{bmatrix} - \boldsymbol{\mu}_c \quad (7)$$

where $\boldsymbol{\mu}_c$ is the bias compensation term given by (see [19] for details)

$$\boldsymbol{\mu}_c = \begin{bmatrix} d_m \cos(\delta_m) \left(\exp^{-\sigma_\delta^2} - \exp^{-\sigma_\delta^2/2} \right) \\ d_m \sin(\delta_m) \left(\exp^{-\sigma_\delta^2} - \exp^{-\sigma_\delta^2/2} \right) \end{bmatrix}. \quad (8)$$

The converted measurement coordinates are not mutually independent, and their error covariance matrix $\mathbf{R}_c = \begin{bmatrix} r_c^{11} & r_c^{12} \\ r_c^{21} & r_c^{22} \end{bmatrix}$ depends on the target bearing and range and on their error standard deviations [19], i.e.,

$$\begin{aligned} r_c^{11} = & d_m^2 \exp^{-2\sigma_\delta^2} \left[\cos^2(\delta_m) (\cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \right. \\ & \left. + \sin^2(\delta_m) (\sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2)) \right] \\ & + \sigma_d^2 \exp^{-2\sigma_\delta^2} \left[\cos^2(\delta_m) (2 \cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \right. \\ & \left. + \sin^2(\delta_m) (2 \sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2)) \right] \end{aligned} \quad (9)$$

$$\begin{aligned} r_c^{22} = & d_m^2 \exp^{-2\sigma_\delta^2} \left[\sin^2(\delta_m) (\cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \right. \\ & \left. + \cos^2(\delta_m) (\sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2)) \right] \\ & + \sigma_d^2 \exp^{-2\sigma_\delta^2} \left[\sin^2(\delta_m) (2 \cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \right. \\ & \left. + \cos^2(\delta_m) (2 \sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2)) \right] \end{aligned} \quad (10)$$

$$\begin{aligned} r_c^{12} = & r_c^{21} = \sin(\delta_m) \cos(\delta_m) \exp^{-4\sigma_\delta^2} \\ & \times \left[\sigma_d^2 + (d_m^2 + \sigma_d^2) \left(1 - \exp^{-\sigma_\delta^2} \right) \right] \end{aligned} \quad (11)$$

We notice that under some conditions ($d\sigma_\delta^2/\sigma_d < 0.4$ and $\sigma_\delta < 23^\circ$) a simpler covariance form, obtained through the linearization of the Cartesian coordinate errors, could be used [19].

The CMKF-D consists in iterating the two following standard Kalman filter steps for each target t .

Prediction step:

— The predicted state

$$\hat{\mathbf{x}}_{k|k-1}^t = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}^t. \quad (12)$$

— The predicted state error covariance

$$\mathbf{P}_{k|k-1}^t = \mathbf{F}\mathbf{P}_{k-1|k-1}^t\mathbf{F}^T + \mathbf{Q}. \quad (13)$$

— The predicted measurement

$$\hat{\mathbf{y}}_{k|k-1}^t = \mathbf{H}\hat{\mathbf{x}}_{k|k-1}^t. \quad (14)$$

— The predicted measurement covariance matrix

$$\mathbf{S}_k^t = \mathbf{H}\mathbf{P}_{k|k-1}^t\mathbf{H}^T + \mathbf{R}_c^t \quad (15)$$

where $\mathbf{Q} = \sigma_v^2 \mathbf{G}\mathbf{G}^T$ is the process noise covariance, T stands for the matrix transpose, and \mathbf{R}_c^t is the debiased converted measurement error covariance for target t . The components of \mathbf{R}_c^t are computed according to (9), (10), and (11).

Update step: This step is carried out for each existing target track using its associated measurement(s) and is done according to the following standard manner:

$$\hat{\mathbf{x}}_{k|k-1}^t = \hat{\mathbf{x}}_{k|k-1}^t + \mathbf{K}_k^t \boldsymbol{\vartheta}_k^t \quad (16)$$

$$\mathbf{P}_{k|k}^t = \mathbf{P}_{k|k-1}^t + \mathbf{K}_k^t \mathbf{S}_k^t (\mathbf{K}_k^t)^T \quad (17)$$

where $\boldsymbol{\vartheta}_k^t$ is the innovation vector, \mathbf{S}_k^t is the associated covariance matrix, and \mathbf{K}_k^t is the conventional Kalman filter gain

$$\mathbf{K}_k^t = \mathbf{P}_{k|k-1}^t \mathbf{H}^T (\mathbf{S}_k^t)^{-1}. \quad (18)$$

B. Data Association

A specific procedure is needed to assign each target to its associated measurement(s). This procedure is designated as data association. Many data association algorithms have been developed [20]–[22]. Among them, the global nearest neighbor, the joint probabilistic data association (JPDA), and multiple hypothesis tracking algorithms are considered the most common and the most efficient techniques. Here, we chose to use the JPDA filter because of its relatively low computational complexity, its suitability to multitarget tracking, and its robustness to noisy environments. In JPDA-filter-based association, all measurements lying in the target validation gate are combined to assess the innovation used to update its state (16).

Formally, a validation gate for a target t is an ellipsoid defined as $\{\mathbf{y}_i | (\mathbf{y}_i - \hat{\mathbf{y}}_{k|k-1}^t)(\mathbf{S}_k^t)^{-1}(\mathbf{y}_i - \hat{\mathbf{y}}_{k|k-1}^t)^T \leq \gamma\}$. γ ensures that the correct measurements will lie within the gate with a specified probability, and it is determined from the Chi-square distribution [20]. This region is computed at every sampling time and is used to select believable correct measurements.

The combined (weighted) innovation vector is given by

$$\boldsymbol{\vartheta}_k^t = \sum_{j=1}^{m_k} \beta_{jt}^k \left(\mathbf{y}_k^j - \hat{\mathbf{y}}_{k|k-1}^t \right) \quad (19)$$

where

- \mathbf{y}_k^j is the measurement j , $j = 1, \dots, m_k$. m_k is the measurement number at ping k ;
- β_{jt}^k is the posterior probability that the measurement \mathbf{y}_k^j truly originated from target t , $t = 0, \dots, t_k$. $t = 0$ stands for false alarm, and t_k is the target number at ping k .

The probability β_{jt}^k of the event “measurement j originated from target t ” is obtained by evaluating the probabilities of the joint events conditioned on all measurements assuming that a measurement cannot be generated by more than one target. Formally, if we denote these joint events $\theta_k = \bigcap_{j=1}^{m_k} \theta_{jt}$, where θ_{jt} is the event that the measurement \mathbf{y}_k^j is originated from target t_j , we have

$$\beta_{jt}^k \triangleq p(\theta_{jt} | \mathbf{y}_{1, \dots, k}) = \sum_{\theta_k | \theta_{jt} \in \theta_k} p(\theta_k | \mathbf{y}_{1, \dots, k}) \quad (20)$$

$$\beta_{0t}^k \triangleq 1 - \sum_{j=1}^{m_k} \beta_{jt}^k \quad (21)$$

where

- $\mathbf{y}_{1, \dots, k} = \{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \dots, \mathbf{y}_k\}$ is the measurement set from time one up to time k and $\mathbf{y}_k = \{\mathbf{y}_k^j\}_{j=1, \dots, m_k}$ the measurement set at ping k .
- β_{0t}^k is the probability that none of the measurements is originated from target t .

The evaluation of the conditional joint event probabilities $p(\theta_k | \mathbf{y}_{1, \dots, k})$ is detailed in [20]. For a parametric JPDA with Poisson distributed false alarms, we obtain the following:

$$p(\theta_k | \mathbf{y}_{1, \dots, k}) = \frac{1}{c} \prod_{j=1}^{m_k} \left\{ \lambda^{-1} N \left(\hat{\mathbf{y}}_{k|k-1}^{t_j}, \mathbf{S}_k^{t_j} \right) \left(\mathbf{y}_k^j \right) \right\}^{\tau_j(\theta_k)} \times \prod_{t=1}^{t_k} \left(p_d^t \right)^{\delta_t(\theta_k)} \left(p_d^t \right)^{1 - \delta_t(\theta_k)} \quad (22)$$

where

- c is a normalization term;
- N is the normal distribution;
- λ is the expected density of false measurement;
- $\tau_j(\theta_k)$ is a binary measurement association indicator equal to one if the measurement is associated to a target (not considered as a false alarm) in the joint event θ_k and zero otherwise;
- $\vartheta(\theta_k) = \sum_{j=1}^{m_k} [1 - \tau_j(\theta_k)]$ is the number of false measurements in event θ_k ;
- p_d^t stands for the detection probability for target t ;
- δ_t is the target detection indicator equal to zero if no measurement is assigned to target t and to one, otherwise.

The state error covariance matrix update (17) expressed according to $\{\beta_{jt}^k\}$ is

$$\mathbf{P}_{k|k}^t = \beta_{0t}^k \mathbf{P}_{k|k-1}^t + (1 - \beta_{0t}^k) \mathbf{P}_{k|k}^{t,c} + \tilde{p}_k \quad (23)$$

where

$$\mathbf{P}_{k|k}^{t,c} = \mathbf{P}_{k|k-1}^t - \mathbf{K}_k^t \mathbf{S}_k^t \mathbf{K}_k^{t,T} \\ \tilde{p}_k = \mathbf{K}_k^t \left(\sum_{j=1}^{m_k} \beta_{jt}^k \boldsymbol{\vartheta}_k^{t,j} (\boldsymbol{\vartheta}_k^{t,j})^T - \boldsymbol{\vartheta}_k^t (\boldsymbol{\vartheta}_k^t)^T \right) (\mathbf{K}_k^t)^T.$$

B. Track Management

JPDAF does not allow for automatic track initiation and termination. In our approach, tracks are managed according to the following rules.

- If there are n_1 detections falling outside all already existing target gates in the last m_1 recurrences below a given distance from the most recent detection (difference between x -coordinates below x_{\max} and between y -coordinates below y_{\max}), a new track is initiated at that position.
- If an existing track is not updated at least n_2 times during the m_2 later recurrences, it is interrupted.
- In order to decrease the processing time and to avoid likely false track maintenance, we remove tracks with a position error standard deviation exceeding a given value (20 m for our data).

V. RESULTS

The proposed approach is applied to several sequences with still and moving objects for mainly two maximum range values: 200 and 50 m. In this section, we show detections and tracking results for four recorded sequences.

- Sequence I: 600 images of 4000 pixels in range and 39 beams (equivalent to 5 min) recorded as our survey boat was slowly approaching a large buoy (of nearly 6 m diameter), two smaller buoys behind it (of nearly 3 m diameter), and concrete blocks near the pier (see Fig. 8);
- Sequence II: 280 images of 1000 pixels in range and 39 beams (equivalent to about 2.5 min) where the inflatable boat was running on its own momentum crossing behind an emerged buoy (see Fig. 9);
- Sequence III: 480 images of 1000 pixels in range and 39 beams recorded while the inflatable boat was running on its own momentum and crossing very close to a sailboat (see Fig. 10);
- Sequence IV: very short sequence of 20 images of 6000 pixels in range and 39 beams, which corresponds to the crossing of a passenger boat with a self-noise visible in the acoustic images (see Fig. 3). Given our sensor frequency (240 kHz), the boat noise specifically corresponds to the cavitation noise.

Standard deviation of the polar measurement error σ_d and σ_δ are set according to the sonar range and bearing sampling rates: $\sigma_d = \Delta_d = 0.05$ m, $\sigma_\delta = \Delta_\delta = 1.5^\circ$, and sampling time $\Delta t = 0.5$ s. The noise process standard deviation is set low but nonzero: $\sigma_v = 10^{-3}$ m/s². Target detection probability and clutter density parameters are set to $p_d^t = 0.9 \forall t$ and $\lambda = p_{\text{fa}}$. Tracks are initiated if a detection is observed at least 16 times during the last 20 pings within a region of 5 m by 5 m for sequences of maximum range of 50 and 10 m by 10 m for

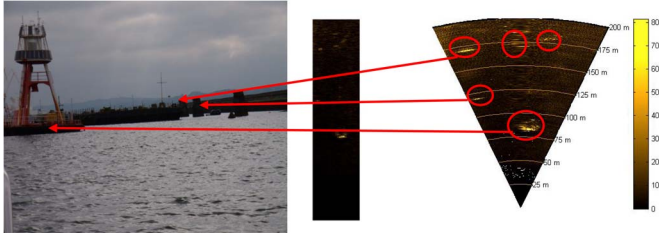


Fig. 8. Sequence I, known surface objects: (from left to right) a photograph of the surface objects, sonar data at ping 550 in polar coordinates, and sonar data at ping 550 in Cartesian coordinates.

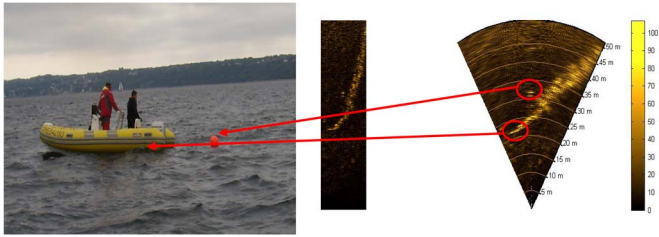


Fig. 9. Sequence II, known surface objects: (from left to right) a photograph of the surface objects, sonar data at ping 250 in polar coordinates, and sonar data at ping 250 in Cartesian coordinates.



Fig. 10. Sequence III: the known surface objects: (from left to right) a photograph of the surface objects, sonar data at ping 450 in polar coordinates, and sonar data at ping 450 in Cartesian coordinates.

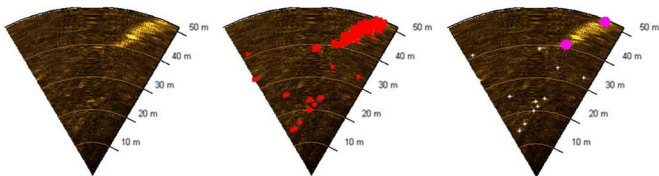


Fig. 11. Sequence II: Detection results at ping 47: (from left to right) raw data in Cartesian coordinates, all detected positions ($p_{fa} = 10^{-6}$), remaining measurements after clustering, and wake detection removal (in white are the cluster center positions and in magenta the detected wake ends).

sequences of maximum range of 200 m. A track is terminated if no measurements are associated to it at least eight times during the last 20 pings ($m_1 = m_2 = 20$, $n_1 = 2n_2 = 16$), $x_{\max} = y_{\max} = 5$ m for sequences of maximum range of 50 m, and $x_{\max} = y_{\max} = 10$ m for sequences with maximum range of 200 m.

Initial state estimate error covariance $\mathbf{P}_{0|0}^t$ for a track associated to a target t is set as follows:

$$\mathbf{P}_{0|0}^t = 5 \begin{bmatrix} \mathbf{R}_0^t & 0 & 0 \\ 0 & (x_{\max}/m_1 \Delta t)^2 & 0 \\ 0 & 0 & (y_{\max}/m_1 \Delta t)^2 \end{bmatrix}$$

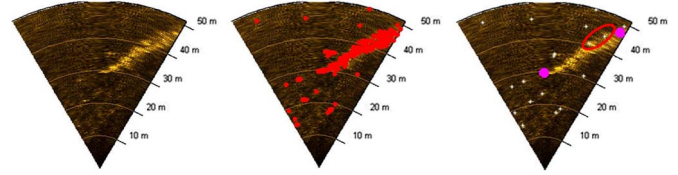


Fig. 12. Sequence II: Detection results at ping 161: (from left to right) data in Cartesian coordinates, all detections ($p_{fa} = 10^{-6}$), remaining measurements after wake detection and detection clustering (in white are the cluster centers, and in magenta are the detected wake ends; the red ellipse shows two detections inside the wake that are not removed by the proposed method).

where \mathbf{R}_0^t is the measurement error covariance matrix at the initial target position computed according to (9), (10), and (11).

The proposed method was tested for various p_{fa} values and various reference cell numbers J . Detection capabilities were assessed visually by the analysis of the detection results and according to our knowledge of the sea-surface environment during the sequence acquisition. To allow this analysis, we estimated the theoretic average false-alarm number as the product of the theoretical p_{fa} by the number of image pixels. The number of false alarms is computed as the difference between the total detection number and ‘true’ detections that are visually identified. We noticed that, in the absence of wakes, we globally obtain false-alarm numbers consistent with the theoretical p_{fa} for a number of reference cells below 50.

In the presence of wake, the number of false alarms can be very high compared to the theoretical false-alarm number. For example at ping 47 of Sequence II, even for low p_{fa} values ($p_{fa} = 10^{-6}$), the number of false alarms is high (about 300 false alarms). Most of these false alarms are related to the inflatable boat wake (see Fig. 11). The amount of detections drastically increases as the ship wake becomes longer (see Fig. 12). For such cases, removal of wake cells is essential. Figs. 11 and 12 show the detection results, respectively, for ping 47 and ping 161 of Sequence II with p_{fa} equal to 10^{-6} . In the middle, we show all detections using the CFAR detector (4). On the right, we draw the measurements used for tracking: Most of detections belonging to the wake are removed, and the number of detections drops from 350 to 15 after clustering (detections within 2 m are replaced by their average position) and wake pixel removal.

Wake detection works well in most of cases and particularly in the case of a strong wake behind the ship. However, in some cases, as the inflatable boat changed its heading frequently, the wake signature could appear distorted, and the method did not succeed in removing all the detections originated from the wake (see Fig. 12). Despite these failures, prior wake detection has the advantage of reducing the overall method computation time while maintaining robust ship tracking (see Figs. 16 and 17). Cavitation noise is easily detected in all pings of Sequence IV, and the ship range is well deduced from the wake detection step (see Fig. 13).

Tracking results can only be assessed in a qualitative way because our targets are not equipped with Global Positioning System. We notice that the proposed procedure succeeds in tracking the known targets. Figs. 14, 16, and 17 show the tracks associated to the known targets in Sequences I, II and III, respectively, for three different pings. We superimpose a confidence ellipse on each target track that describes the Kalman

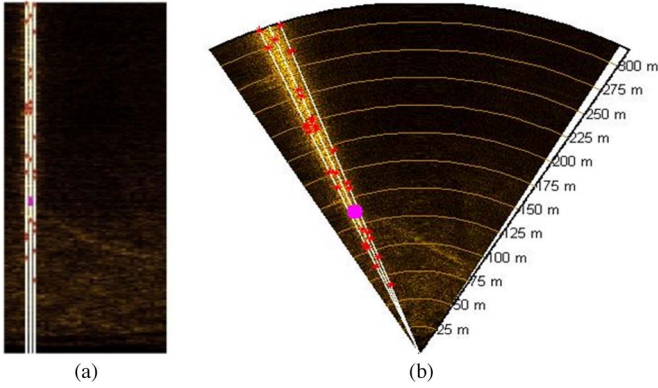


Fig. 13. Sequence IV: Detection results: (white lines) cavitation noise detection, (red) candidate range detections, and (magenta) the final range deduced from wake detection ($p_{fa} = 10^{-6}$), in (a) polar coordinates and in (b) Cartesian coordinates.

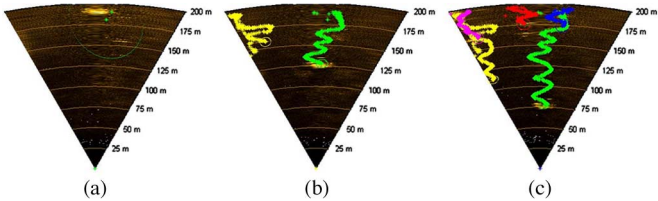


Fig. 14. Sequence I: The known object tracks and their corresponding confidence ellipse up to ping: (a) $k = 27$, (b) $k = 350$, and (c) $k = 590$. In green is the track associated to the large buoy, yellow and magenta tracks correspond to the two concrete blocks, and in red and in blue are the tracks associated to the smaller buoys ($p_{fa} = 10^{-5}$, $\gamma = 9$).

filter own assessment of tracking accuracy, expressed by the covariance matrix $\mathbf{P}_{k|k}^t$. This ellipse is only shown for the predicted target position and it is defined as follows:

$$\left\{ \mathbf{y} = [p_x, p_y] \mid \left(\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_{k|k}^t \right) \left(\mathbf{L}\mathbf{P}_{k|k}^t\mathbf{L}^T \right)^{-1} \left(\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_{k|k}^t \right)^T \leq \lambda \right\} \quad (24)$$

where $\mathbf{L} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$.

At the beginning of Sequence I, only the largest buoy is in the field of view of the sonar, and the concrete blocks and the two smaller buoys are seen later. The proposed method succeeds in automatically initiating every target track since they appear in the field of view of the sonar. However, the first concrete block was initiated by a ‘false’ track (the track in yellow in Figs. 14 and 15). Indeed, for this sequence, despite using low p_{fa} value ($p_{fa} = 10^{-5}$), ‘false’ tracks were created. These false tracks were related to pier pillars present at the edge of the image but not ‘seen’ by the sonar at every ping because of the survey boat yaw (see Fig. 15).

For Sequence I, the targets are still. The shape of the tracks (see Fig. 14) was due to the survey boat yaw. In an operational system, the sonar image will be stabilized according to carrier movement, and uniform target trajectory will appear as straight lines; hence, we expect better performances.

For Sequences II and III, we show three tracks in Figs. 16 (Sequence II) and 17 (Sequence III): one associated to the fixed buoy and two tracks associated to both wake

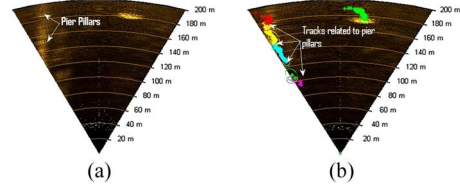


Fig. 15. Sequence I: (a) Image at ping: $k = 77$. (b) Existing tracks at ping $k = 111$: All tracks on the left are initiated by pier pillar associated detections.

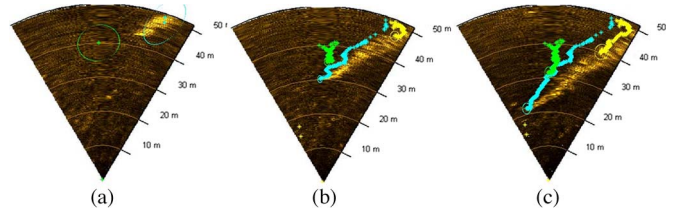


Fig. 16. Sequence II—The known object tracks and their confidence ellipses up to ping (a) $k = 22$, (b) $k = 150$, and (c) $k = 275$. In green is the track associated to the buoy, in cyan is the track associated to the inflatable boat, and in yellow is the track associated to the second wake end ($p_{fa} = 10^{-6}$, $\gamma = 4$).

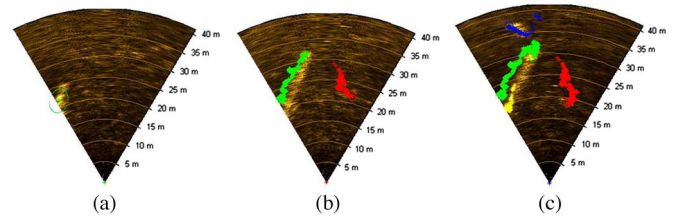


Fig. 17. Sequence III—The known object tracks and their confidence ellipses up to (a) ping $k = 21$, (b) $k = 330$, and (c) $k = 460$. In green is the track associated to the inflatable boat, in yellow is the track associated to the second wake endpoint, in red is the track associated to the fixed buoy, and in blue is the track related to a sailboat ($p_{fa} = 10^{-6}$, $\gamma = 4$).

endpoints. At the end of Sequence III, we can see the track related to the sailboat.

In Sequence II, when the zodiac approaches the buoy then goes past it, the algorithm succeeds in maintaining both tracks (see Fig. 16).

VI. CONCLUSIONS AND DISCUSSIONS

Sea-surface detection and tracking by means of forward-looking sonar is a useful application for safe AUV and submarine surfacing. In this paper, we have proposed an automatic method for detecting and tracking various sea-surface targets: Still and moving objects with or without wake and cavitation noise. Promising results have been obtained with real data gathered at sea.

Future improvements will first concern the integration of additional features in the target state vector, e.g., the object size and the wake orientation. Indeed, in the proposed method, wake detection was mainly used to improve the CFAR detector performance and to reduce the false-alarm rate due to the removal from the measurement list of unwanted detection inside the wakes. In addition, in the operational system, the data are stabilized from the sensor yaw, pitch, and roll; hence, target trajectories will appear straighter than the tracks presented here, and the wake orientations will be consistent with target

trajectories. For such case, wake orientation could be used as an additional target feature, which may improve target tracking. However, adding wake orientation to the target state vector will lead to a highly nonlinear measurement system, and the CMKF-D will no longer be suitable; for example, an extended Kalman filter or a JPDAF particle filter should be used.

The proposed method has been carried out on images without any preprocessing, and image denoising may also improve the detection performances.

REFERENCES

- [1] Y. Petillot, I. T. Ruiz, and D. M. Lane, "Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar," *IEEE J. Ocean. Eng.*, vol. 26, no. 2, pp. 1194–1199, Apr. 2001.
- [2] I. Quidu, A. Hétet, Y. Dupas, and S. Lefèvre, "AUV (REDERMOR) obstacle detection and avoidance experimental evaluation," in *Proc. IEEE OCEANS Conf.*, Aberdeen, Scotland, Jun. 2007, pp. 18–21.
- [3] L. Fillinger, A. Sutin, and A. Sedunoy, "Acoustic ship signature measurements by cross-correlation method," *J. Acoust. Soc. Amer.*, vol. 129, no. 2, pp. 774–778, 2011.
- [4] G. L. Ogden, L. M. Zurk, M. E. Jones, and M. E. Peterson, "Extraction of small boat harmonic signatures from passive sonar," *J. Acoust. Soc. Amer.*, vol. 129, no. 2, pp. 3768–3776, 2011.
- [5] M. V. Trevorrow, S. Vagle, and D. M. Farmer, "Acoustical measurements of microbubbles within ship wakes," *J. Acoust. Soc. Amer.*, vol. 95, no. 4, pp. 1922–1930, 1994.
- [6] S. Vagle and H. Burch, "Acoustic measurements of the sound-speed profile in the bubbly wake formed by a small motor boat," *J. Acoust. Soc. Amer.*, vol. 117, no. 1, pp. 153–163, 2005.
- [7] A. Sutin, A. Benilov, H. S. Roh, and Y. I. Nah, "Acoustic measurements of bubbles in the wake of ship model in tank," *J. Acoust. Soc. Amer.*, vol. 123, no. 5, pp. 8651–8656, 2008.
- [8] K. W. Lo and B. G. Ferguson, "Automatic detection and tracking of a small surface watercraft in shallow water using a high-frequency active sonar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 4, pp. 1377–1388, Oct. 2004.
- [9] Ç. Önur, "Acoustic tracking of ship wakes," M.S. thesis, Graduate School Natural Appl. Sci., Middle East Tech. Univ., Ankara, Turkey, 2013.
- [10] S. Jeong, S. W. Ban, S. Choi, and D. Lee, "Surface ship-wake detection using active sonar and one-class support vector machine," *IEEE J. Ocean. Eng.*, vol. 37, no. 3, pp. 456–466, Jul. 2012.
- [11] X. Lurton, *An Introduction to Underwater Acoustics: Principles and Applications*. New York, NY, USA: Springer-Verlag, 2002.
- [12] D. J. Crisp, "The state-of-the-art in ship detection in synthetic aperture radar imagery," *Intell., Surveillance Reconnaissance Div. Inf. Sci. Lab.*, Cambridge, MA, USA, DSTO-RR-0272, 2004.
- [13] H. M. Finn and R. S. Johnson, "Adaptive detection mode with threshold control as a function of spatially sampled clutter-level estimates," *RCA Rev.*, vol. 29, no. 3, pp. 414–464, 1968.
- [14] R. Ravid and N. Levanon, "Maximum-likelihood CFAR for Weibull background," *Proc. Inst. Elect. Eng. Radar Signal Process.*, vol. 139, no. 3, pp. 256–264, Jun. 1992.
- [15] P. J. Grieve, "The optimum constant false alarm probability detector for relatively coherent multichannel signals in Gaussian noise of unknown power," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 6, pp. 708–721, Nov. 1977.
- [16] A. Cetin, "CFAR Detection in k-distributed sea clutter," Ph.D. dissertation, Electron. Eng. Dept., Middle East Tech. Univ., Ankara, Turkey, 2008.
- [17] I. Quidu, L. Jaulin, A. Bertholom, and Y. Dupas, "Robust multitarget tracking in forward looking sonar image sequences using navigational data," *IEEE J. Ocean. Eng.*, vol. 37, no. 3, pp. 417–430, Jul. 2012.
- [18] M. Trevorrow, "High frequency acoustic scattering and absorption effects within ship wakes," Defence Res. Develop. Canada, Ottawa, ON, Canada, DRDC-ATLANTIC-SL-2005-125, Aug. 2005.
- [19] D. Lerro and Y. Bar-Shalom, "Tracking with debiased consistent converted measurements versus EKF," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 3, pp. 1015–1022, Jul. 1993.
- [20] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*. New York, NY, USA: Academic, 1988.
- [21] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "Tracking multiple moving targets with a mobile robot using particle filters and statistical data association," in *Proc. IEEE ICRA*, 2001, vol. 2, pp. 1665–1670.
- [22] I. Cox and S. Hingorani, "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 2, pp. 138–150, Feb. 1996.



Imen Karoui received the Engineer's degree from Ecole Polytechnique de Tunis, Tunis, Tunisia, in 2002 and the Ph.D. degree in signal processing and telecommunications from Telecom Bretagne, Plouzané, France, in 2007.

She is currently a Research Engineer with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Brest, France. Her main research interests include texture segmentation, detection, and tracking applied to sonar and synthetic aperture radar images.



Isabelle Quidu received the Ph.D. degree in electronics from Université de Bretagne Occidentale, Brest, France, in 2001.

From 2002 to 2005, she was a Research Engineer with Thales Underwater Systems in the field of mine warfare. Since 2005, she has been an Associate Professor with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Brest, and a member of the Ocean Sensing and Mapping Team. She is also a member of Lab-STICC UMR CNRS 6285. Her research interests include detection and classification

of underwater targets in sonar images.



Michel Legris received the Engineer's degree from Ecole Centrale de Paris, Paris, France, in 1985 and the Ph.D. degree in electrical engineering from the University of Grenoble, Grenoble, France, in 1996.

From 1986 to 2001, he was an engineer with the French Ministry of Defense, where he did research on low magnetic fields, hydroacoustics noise and high-frequency sonar technology. Since 2001, he has been an Associate Professor with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Brest, France. His current research interests include under-

water acoustics systems and innovative sonar.

Titre : Perception de l'environnement marin et de la menace à partir d'images sonar haute fréquence : du fond marin à la surface de mer

Mots clés : imagerie sonar, reconnaissance automatique de cible sous-marine, perception

Résumé : A l'instar des mammifères marins, l'homme a développé des systèmes sonar capables de percevoir l'environnement marin sur des distances plus ou moins longues et avec plus ou moins de détails en fonction de ses besoins. Le principal besoin traité dans ce travail préparé pour l'habilitation à diriger les recherches est la reconnaissance automatique de cibles sous-marines (plus connue sous l'acronyme ATR pour Automatic Target Recognition), qu'elles soient posées sur le fond ou flottantes entre deux eaux.

La chronologie de la sélection de travaux présentés suit le perfectionnement au fil des années des systèmes d'imagerie et des algorithmes de traitement associés pour répondre à l'évolution de la menace, depuis les mines furtives des années 90 aux engins explosifs improvisés d'aujourd'hui. Pour contrer ces menaces à partir de données sonar latéral ou sonar à antenne synthétique (dites données SAS pour Synthetic Aperture Sonar), deux points sont en particulier exposés : une nouvelle approche de l'ATR pour les missions de chasse aux mines qui passe par la qualification de l'environnement proche de la mine d'une part, la compréhension et l'exploitation de données SAS acquises lors de passes répétitives pour la mise en œuvre d'algorithmes de détection (incohérente puis cohérente) de changements dans un contexte de surveillance d'autre part. Outre la reconnaissance de cibles sur le fond

marin, la problématique de la détection d'obstacles s'est imposée plus récemment afin de permettre à des drones sous-marins d'autonomie croissante de mener leur mission en toute sécurité et ce, quelle que soit leur immersion. Dans ce manuscrit, la configuration et le mode des systèmes sonar d'évitement, de même que les algorithmes dédiés à chaque mission - suivi de fond, pistage d'obstacles dans la colonne d'eau, et enfin reprise de vue-, sont discutés.

La synthèse des travaux effectuée pour cette habilitation à diriger les recherches repose ainsi sur l'exploitation des systèmes d'imagerie sonar haute fréquence (HF), des sonars latéraux aux sonars frontaux, et est le plus souvent le fruit de diverses collaborations. Initialement destiné aux opérateurs, l'usage des images sonar HF est aujourd'hui à adapter face au flot croissant de données à traiter et à l'automatisation progressive des tâches à bord de drones sous-marins. La discrétion, la souplesse et la sécurité de ces véhicules ouvrent de nombreuses possibilités en termes de missions et d'utilisation de capteurs ad hoc d'imagerie et de navigation, y compris pour le domaine civil. Plus globalement, cette adaptation à la robotique mobile sous-marine exigera l'étude des mécanismes de perception-action dans un environnement incertain et peu structuré avec une communication limitée.

Title : Underwater environment and threat perception by high-frequency sonar images: from seabed to water surface

Keywords : sonar imaging, automatic target recognition, perception

Abstract : Similarly to marine mammals, mankind has developed sonar systems which are able to perceive underwater environments on a range of distances and with more or less detail, depending on its needs. The need addressed in this manuscript is automatic underwater target (ATR) recognition, whether the target is laying on the seafloor or floating in the water column.

The chronology of the work which is presented shows how sonar systems and underlying algorithms have progressively been improved in order to adapt to new threats, from the stealthy mines of the 90s to the improvised explosive devices of today. To counter these threats using side-scan or synthetic aperture sonar (SAS) data, two topics in particular are discussed: first, a new mine-hunting ATR approach which takes into consideration the environment in the vicinity of the target; second, the analysis and use of repeat-pass SAS data acquired in a monitoring context in order to perform change detection. Apart from the detection of targets lying on the seabed, obstacle detection issues have more recently been tackled in order to allow underwater drones to operate in

complete safety whatever the depth. For each type of mission –seabed tracking, obstacle detection and tracking in the water column, and vehicle surfacing–, the combination of configuration and mode of the obstacle-avoidance sonar systems, and dedicated algorithms are described.

The body of work presented is based on high-frequency sonar imaging, ranging from side-scan sonars to forward-looking sonars, and most of it is the result of various collaborations. Initially intended for operators, the use of sonar images needs to be adapted to an increasing quantity of data to be processed, and to the increasing automation of tasks aboard underwater drones. The stealth, flexibility and safety of these small-size vehicles open up new opportunities both of missions and of choice of relevant imaging and navigation sensors, not only in the military but also in the civilian domain. More generally, this adaptation process to underwater mobile robotics will require the analysis of perception-action mechanisms in an uncertain and unstructured environment, with limited communication.