



HAL
open science

Vision et reconstruction 3D : application à la robotique mobile

Rihab Hmida

► **To cite this version:**

Rihab Hmida. Vision et reconstruction 3D : application à la robotique mobile. Robotique [cs.RO]. Université Montpellier; Laboratoire d'électronique et de micro-électronique (Monastir, Tunisie), 2016. Français. NNT : 2016MONTT335 . tel-01817574

HAL Id: tel-01817574

<https://theses.hal.science/tel-01817574>

Submitted on 18 Jun 2018

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.



THESE EN COTUTELLE INTERNATIONALE



ENTRE
L'UNIVERSITE DE MONASTIR
Ecole Doctorale Sciences de l'Ingénieur et Technologique
ET
L'UNIVERSITE DE MONTPELLIER
Ecole Doctorale I2S

*_*_*_*_*_*_*_*_*_*_*_*_*_*_*_*

THESE

Présentée à

L'Ecole Nationale d'Ingénieurs de Monastir

Pour l'obtention du grade de

DOCTEUR D'UNIVERSITE

Discipline :

Génie Electrique
Systemes Automatiques et Microélectronique

Présentée par :

Rihab HMIDA

Vision et reconstruction 3D : Application à la robotique mobile.

Soutenue le 16/ 12/2016 devant la commission d'examen composée de :

Mr. Ali DOUIK	Professeur à l'université de Sousse	Rapporteur
Mr. Youssef MEZOUAR	Professeur à l'université de Clermont-Ferrand	Rapporteur
Mr. Mohamed ATRI	Maître de conférences à l'université de Monastir	Examineur
Mr. Frédéric COMBY	Maître de conférences à l'université de Montpellier	Examineur
Mr. René ZAPATA	Professeur à l'université de Montpellier	Directeur de thèse
Mr. Abdellatif MTIBAA	Professeur à l'université de Monastir	Directeur de thèse

Laboratoire Electroniques et microélectroniques (Code: LR99ES30)
Laboratoire Informatique, robotique et Microélectronique (Code: UMR5506)

Avant-propos

Tout d'abord, les travaux de cette thèse ont été effectués entre deux unités de recherches. La première équipe est celle du groupe CSR (Circuits, Systèmes Reconfigurables) faisant partie du Laboratoire Electronique et Micro-Electronique de l'Université de Monastir (Tunisie). C'est au sein de cette équipe que j'ai commencé mes premiers travaux de recherche de maîtrise et puis de doctorat et où j'ai pu rencontrer des membres aussi généreux que serviables. La deuxième consiste aux deux équipes ICAR (Département Image & Interaction) et EXPLORE (Département Robotique) au sein du Laboratoire Informatique, Robotique et Micro-Electronique de l'Université de Montpellier (France). Une équipe qui m'a accueillie avec beaucoup d'hospitalité et de gratitude tout le long de mes séjours en France.

Je remercie mes directeurs de thèse, Mr. René ZAPATA et Mr. Abdellatif MTIBAA de m'avoir suivie tout au long de cette thèse. Je tiens à remercier Mr. MTIBAA de m'avoir encadrée depuis mes études de maîtrise, pour sa confiance, sa gentillesse et pour la chance qu'il m'a donnée pour lancer cette thèse en cotutelle avec l'Université de Montpellier. Aussi, je remercie Mr. ZAPATA d'avoir accepté de diriger cette thèse au sein du laboratoire LIRMM et pour sa générosité et sa collaboration humaine et scientifique. Je me considère chanceuse de travailler avec tous les membres de toutes les équipes : tous dynamiques, disponibles et jamais à court d'idées. Vous m'avez toujours offert de bonnes conditions pour travailler et j'ai beaucoup appris de vous !

Je tiens à remercier sincèrement mes deux co-encadrants au LIRMM, Mr. Frédéric COMBY et Mr. Lionel LAPIERRE, qui m'ont suivi tout au long de cette thèse, pour leur gentillesse et leur disponibilité pour répondre à mes questions. Je vous remercie de m'avoir accompagné lors de mes séjours en France et de me faire part de votre générosité et savoirs.

Un grand merci à mon co-encadrant Mr. Abdeslem BEN ABDELLALI, membre du groupe CSR, pour sa grande écoute, disponibilité et son suivi sérieux et exigeant, qui m'ont été d'une aide précieuse. Merci pour vos encouragements et conseils lors des réunions régulières d'avancement de thèse.

Merci à Mr. Youssef MEZOUAR et Mr. Ali DOUIK d'avoir accepté d'évaluer et de juger ma thèse. Je leur remercie, également, pour l'intérêt qu'ils ont porté à mon travail et pour leurs remarques constructives.

Un grand merci à tous mes amis du groupe CSR et au LIRMM. Je leur exprime ma profonde gratitude et je leur souhaite beaucoup de réussite.

Enfin, j'ai une pensée toute particulière pour mes parents, ma sœur, mon frère, mes beaux-parents et mon cher époux. Je suis heureuse de leur dédier ce travail.

Table des matières

Introduction Générale	10
1. Contexte général	11
2. Problématique	12
3. Contributions	12
4. Plan du mémoire	13
5. Contributions scientifiques	14
Partie I. Etude Théorique : Etat de l'Art	16
Chapitre I. Etat de l'Art : Vision, Reconstruction 3D et Robotique Mobile	17
1. Vision : Généralités et Systèmes de vision	18
1.1 Généralités	18
1.1.1. Historique.....	18
1.1.2. Paradigme de Marr.....	19
1.1.3. Vision par ordinateur.....	19
1.2 Systèmes de vision.....	21
1.2.1 Systèmes passifs / actifs.....	22
1.2.2. Systèmes à large champ.....	24
1.3 Les systèmes catadioptriques	25
2. Reconstruction 3D : Systèmes et Méthodes	30
2.1. Les systèmes non-optiques.....	30
2.2. Les systèmes optiques actifs.....	31
2.3. Les systèmes optiques passifs	32
3. Robotique mobile : Généralités et Exemples d'applications	33
3.1. Généralités	33
3.2. Exemples d'applications	34
3.2.1. Exemples de robots terrestres.....	35
3.2.2. Exemples de robots sous-marins.....	37
4. Conclusion	41
Chapitre II. Taxonomies des Algorithmes Existants	43
1. Introduction	44
2. Etalonnage des systèmes omnidirectionnels	44
2.1. Classification des méthodes d'étalonnage.....	45
2.2. Approches d'étalonnage utilisant un modèle de test 2D.....	47
2.3. Etalonnage sous-l'eau.....	49

3. Détection et correspondance de points caractéristiques	50
3.1. Algorithmes de détection de spot(s) laser.....	50
3.2. Algorithmes de correspondances.....	52
4. Reconstruction 3D	53
4.1. Géométrie épipolaire.....	54
4.2. Triangulation.....	55
4.3. Géométrie du tracé de rayon.....	57
5. Conclusion	57

Partie II. Etude Expérimentale : Etude, Validation et Implémentation du système proposé.....60

Chapitre III. Etude et Validation du Système Proposé61

1. Introduction	62
2. Description des étapes de traitement	62
3. Etalonnage des systèmes stéréoscopiques catadioptriques	63
3.1. Description du système stéréoscopique.....	64
3.2. Démarche d'étalonnage du système.....	67
3.3. Validation de l'approche d'étalonnage.....	69
3.4. Etalonnage du système stéréoscopique réel.....	73
3.4.1. Etalonnage du système stéréoscopique réel dans l'eau.....	75
3.4.2. Nouvelle approche d'étalonnage dans l'air pour une utilisation sous l'eau.....	76
3.5. Estimation de l'indice de réfraction.....	78
4. Détection et correspondance de points caractéristiques	80
4.1. Détection des points laser.....	80
4.2. Correspondance des points laser stéréo.....	83
5. Reconstruction 3D sous-marine	84
5.1. Equations de reconstruction 3D dans l'eau avec étalonnage dans l'air.....	85
5.2. Equations de reconstruction 3D avec étalonnage dans l'eau.....	86
6. Expérimentations et interprétation des résultats	88
6.1. Conditions expérimentales.....	89
6.2. Evaluation de la précision des mesures sur la structure 3D.....	90
6.3. Projection des expérimentations pour application dans un environnement sous-marin réel.....	92
7. Conclusion	95

Chapitre IV. Architectures et Implémentation matérielle 96

1. Introduction	97
2. Environnement de développement et Plateforme FPGA	98

2.1. Environnement de développement	98
2.1.1. Flot de conception	99
2.1.2. Exploration architecturale et optimisation	100
a. Extraction de parallélisme	100
b. Optimisation algorithmique.....	101
2.2. Plateforme FPGA.....	103
3. Architecture matérielle proposée pour la reconstruction 3D.....	105
3.1. Bibliothèque Xilinx.....	105
3.2. Bibliothèque d'opérateurs personnalisés et optimisation.....	105
3.2.1. Fonctions personnalisées	106
a. Fonction de division	106
b. Fonction Racine Carrée	107
c. Fonctions trigonométriques	108
3.2.2. Comparaison entre le CORDIC et les algorithmes proposés après optimisation	109
3.3. Architecture proposée en XSG	109
3.4. Module de détection et extraction des spots laser.....	110
3.5. Module de reconstruction 3D depuis les projections images 2D	113
3.5.1. Module de détermination de la fonction d'approximation "Scaramuzza" :.....	115
3.5.2. Module de normalisation des rayons vecteurs correspondants "Normalisation" :.....	116
3.5.3. Module de détermination des angles de réfractions par la méthode de Snell "refraction_angle"	117
3.5.4. Module de détermination des coordonnées 3D du point réel "Distance measurement" :.....	118
4. Expérimentation et analyse de performances.....	119
4.1. Résultats de co_simulation.....	119
4.2. Résultats de synthèse.....	122
4.3. Analyse temporelle de la conception matérielle.....	124
4.3.1. Fréquence maximale	124
4.3.2. Analyse de performances	125
5. Conclusion	128
Conclusion Générale	130
Bibliographies.....	133

Liste des figures

Figure I.1. Photos de Alhazen, René Descartes, Isaac Newton, Thomas Young et Samir Zeki.....	18
Figure I. 2: Chaîne d'acquisition d'une scène réelle à l'image numérique	20
Figure I. 3: Exemples de systèmes de perception	21
Figure I. 4: Principe de stéréovision passive.....	22
Figure I. 5: Banc stéréoscopique pour endoscopie médicale	23
Figure I. 6: Exemples de systèmes actifs à différents motifs de projection laser.....	23
Figure I. 7: Exemples de système actif avec projecteur de lumière.....	24
Figure I. 8: Champ de vue des systèmes de vision : (a) Vision classique, (b) Vision panoramique, (c) Vision omnidirectionnelle.....	24
Figure I. 9: Systèmes à large champ de vue.....	25
Figure I. 10: Projection d'une image perspective sur le plan image d'un capteur catadioptrique	26
Figure I. 11: Exemple de projection sur un miroir : (a) parabolique et (b) hyperbolique	27
Figure I. 12: Illustration des repères utilisés.....	29
Figure I. 13: Taxonomie des méthodes d'acquisitions de formes	30
Figure I. 14: Reconstruction 3D avec un système à motifs lasers structurés [28]	31
Figure I. 15: Reconstruction à partir de plusieurs images.....	32
Figure I. 16: Principe de la photométrie [35]	33
Figure I. 17: Système stéréoscopique de capteurs catadioptriques: (a) Système passif, (b) Système actif.....	35
Figure I. 18: Système omnistéréo binoculaire-V à une seule caméra.....	36
Figure I. 19: Système catadioptrique à double miroir pour observation du conducteur	36
Figure I. 20: Système catadioptrique parabolique et conique avec système laser	37
Figure I. 21 : (a) Système d'inspection "Sabertooth" (b) Exemple d'image Sonar	38
Figure I. 22 : Système stéréoscopique pour application sous-marine.....	39
Figure I. 23 : Résultats de l'algorithme de correction sur un exemple d'image sous-marine.....	39
Figure I. 24 : (a) Système de navigation DEPTHX à faisceaux sonar, (b) Carte de profondeur résultante	39
Figure I. 25: Système stéréoscopique pour application sous-marine	40
Figure I. 26: Système stéréoscopique pour application sous-marine	40
Figure II. 1: Apprentissage et commande du robot [37].....	52
Figure II. 2: Système stéréo à deux caméras [79] Figure II. 3: Système stéréo à une seule caméra [80]	54
Figure II. 4: Reconstruction de formes 3D par géométrie épipolaire.....	55
Figure II. 5: Exemple de reconstruction 3D sous-marine [85].....	56
Figure II. 6: Reconstruction 3D sous-marine par triangulation [80].....	56
Figure II. 7: Principe de réfraction de rayon laser sous l'eau	57
Figure II. 8: Schéma de réfraction d'un rayon lumineux passant par différents milieux	58
Figure III. 1: Etapes de traitements d'images	63
Figure III. 2: Modélisation paramétrique du système catadioptrique	65
Figure III. 3: Modélisation du système stéréo actif (a) Système 3D (b) Image stéréo acquises.....	65
Figure III. 4: Différents périphériques du capteur catadioptrique	66
Figure III. 5: Vérification de la satisfaction de la contrainte PVU	66
Figure III. 6: Système stéréoscopique à pointeurs laser.....	67
Figure III. 7: Diagramme d'étalonnage stéréo du système catadioptrique et à lasers	69
Figure III. 8: Interface graphique de simulation	69

Figure III. 9: Illustration de la relation entre Baseline et les positions des lasers	70
Figure III. 10: Résultats de simulation: (a) Détection des coins (b) Détection du centre de l'image et optimisation par projection inverse (c) Projection des positions de la mire (d) Estimation de la Baseline.....	71
Figure III. 11: Comparaison entre points 3D réels (en vert) et points 3D reconstruits (en rouge).....	73
Figure III. 12: Résultat de comparaison : (a) Modèle 3D réel (b) Modèle reconstruit selon Scaramuzza (en rouge) et selon Mei (en bleu).....	73
Figure III. 13: Résultat de simulation du processus d'étalonnage.....	74
Figure III. 14: Image stéréo d'une même mire vue sous l'eau	76
Figure III. 15: Illustration géométrique du passage air/plexi/eau	77
Figure III. 16: Illustration géométrique de la nouvelle approche d'étalonnage.....	78
Figure III. 17: Systèmes de vision avec un système d'étanchéité	79
Figure III. 18: Projection du laser (a) environnement éclairé (b) environnement sombre.....	80
Figure III. 19: Illustration des régions de recherche.....	82
Figure III. 20: Résultat de segmentation de l'image couleur pour une scène (a) sombre et (b) éclairée.....	82
Figure III. 21: Résultats de simulation : (a) image couleur, (b) détection des points laser et (c) extraction des coordonnées 2D	83
Figure III. 22: Résultat de correspondance.....	83
Figure III. 23: Illustration des scénarios possibles.....	84
Figure III. 24: Géométrie de reconstruction 3D avec étalonnage dans l'air.....	85
Figure III. 25: Schéma d'approximation de l'intersection de deux vecteurs 3D dans l'espace.....	86
Figure III. 26: Schéma illustrant la reconstruction stéréo active d'un point 3D.....	88
Figure III. 27: Organigramme du processus de traitement	89
Figure III. 28: Environnement de test.....	90
Figure III. 29: Exemple d'images prises sous l'eau avec différentes orientations de la mire d'étalonnage.....	90
Figure III. 30: Mesures métriques sur deux différentes structures 3D sous l'eau.....	91
Figure III. 31: Mesures métriques de distances par rapport à la position du spot laser dans l'image	92
Figure III. 32: Effet de la variation d'indice de réfraction sur des mesures métriques	94
Figure III. 33: Illustration du phénomène de réflexion dans (a) l'air et (b) l'eau	94
Figure IV. 1 : Comparaison entre les technologies d'implémentation.....	97
Figure IV. 2 : Principe de la co_simulation logicielle-matérielle.....	99
Figure IV. 3 : Flot de conception proposé pour le prototypage rapide et efficace	100
Figure IV. 4 : Exemple d'extraction de parallélisme au niveau des instructions.....	101
Figure IV. 5 : Comparaison entre le même module avant et après optimisation	102
Figure IV. 6 : La carte de test à base de FPGA : plateforme ML605	103
Figure IV. 7 : Structure interne d'un FPGA	104
Figure IV. 8 : Organigramme de l'architecture proposée pour l'approximation de la division.....	107
Figure IV. 9 : Organigramme de l'architecture proposée pour l'approximation de la racine carrée	108
Figure IV. 10 : Organigramme de l'architecture proposée pour l'approximation de la fonction ArcSinus	108
Figure IV. 11 : Processus de détection de points laser et de reconstruction 3D	109
Figure IV. 12 : Module de segmentation des images RVB converties en HSV	110
Figure IV. 13 : Module d'extraction des coordonnées des points lasers.....	111
Figure IV. 14 : Sous-système 2: Elimination des fausses projections	112
Figure IV. 15 : Sous-système 3: Extraction parallèle des coordonnées des spots laser	113
Figure IV. 16 : Module de passage entre les deux phases : Détection et reconstruction	114
Figure IV. 17 : Module de reconstruction 3D.....	114
Figure IV. 18 : Module de Scaramuzza.....	115
Figure IV. 19 : Module d'approximation de la racine carrée d'un nombre réel.....	116

Figure IV. 20 : Module de normalisation	117
Figure IV. 21 : Module de détermination de l'angle de réfraction entre la surface du plexi et l'eau.....	117
Figure IV. 22 : Module de détermination des angles de réfraction à chaque surface séparatrice.....	118
Figure IV. 23 : Module de détermination des coordonnées 3D et de la mesure de distance	118
Figure IV. 24 : Module de détermination des coordonnées 3D et de la mesure de distance	119
Figure IV. 25 : Module de Co_simulation.....	120
Figure IV. 26 : Résultats de pré-traitement d'image (a) Segmentation par code Matlab, (b) Segmentation par co_simulation, (c) Identification des spots Laser par Hw_cosimulation, et (d) Extraction des centres des spots par code Matlab et blocs XSG.....	121
Figure IV. 27 : Résultats de co_simulation hardware sur FPGA	121
Figure IV. 28 : Schéma RTL du module de reconstruction 3D	123
Figure IV. 29 : Illustration du processus de contrôle du robot	124

Liste des tableaux

Tableau II. 1 : Classification des méthodes d'étalonnage de système omnidirectionnel	46
Tableau III. 1 : Paramètres intrinsèques de la caméra 1.....	72
Tableau III. 2 : Paramètres intrinsèques et extrinsèques dans un environnement aérien.....	74
Tableau III. 3 : Paramètres intrinsèques et extrinsèques dans un environnement aquatique	75
Tableau III. 4 : Exemple d'évaluation d'erreur.....	92
Tableau IV. 1 : Comparaison des ressources utilisées avant et après optimisation du bloc hardware	102
Tableau IV. 2 : Comparaison de performances entre CORDIC et le bloc XSG proposé.....	109
Tableau IV. 3 : Reliquat des ressources FPGA utilisées pour toute l'architecture	123
Tableau IV. 4 : Temps d'exécution des blocs matériels.....	125

Introduction Générale

1. Contexte général

Avec l'évolution des processus technologiques permettant la démocratisation et la miniaturisation des systèmes d'acquisition et de traitement de l'information, la vision par ordinateur et l'imagerie numérique ont gagné de plus en plus d'importance dans le secteur industriel (automatisation et contrôle de qualité) comme dans la société civile (vidéo surveillance, reconnaissance de visages, cinéma, photographie numérique, imagerie médicale) et les structures militaires (véhicules autonomes, téléguidage, ...), ainsi que dans la vie de tous les jours. Le nouvel enjeu maintenant est de pouvoir assembler les briques de base développées ces dernières années, pour construire des systèmes de vision complets, performants et robustes en se basant sur ce que les nouvelles technologies nous offrent de point de vue outils matériels ainsi que logiciels. Finalement, on arrive à un point où la théorie et les algorithmes de base sont bien connus et où le matériel est devenu performant et facilement disponible. Dans ce même contexte, la robotique mobile n'a pas cessé d'attirer l'attention, notamment pour remplacer l'homme dans des environnements à risques (zones radioactives, robots militaires) ou des zones qui lui sont inaccessibles (exploration planétaire ou sous-marine), ou à des échelles différentes (robot à l'intérieur d'une canalisation, voire robot chirurgical à l'intérieur du corps humain).

Dans ces travaux de thèse, on s'intéresse plus particulièrement à l'exploration d'environnements sous-marins qui a émergé dans le cadre d'un projet regroupant plusieurs géologues, roboticiens et scientifiques. Elle est issue du besoin de mieux connaître, de caractériser les réseaux sous-marins, appelés des réseaux karstiques. Néanmoins, la présence de courants assez rapides et de configurations topologiques complexes empêche d'envoyer des plongeurs humains tout en restant dans des conditions de sécurité optimales. L'idée est donc d'envisager l'exploration de ces réseaux sous-marins à l'aide d'un robot muni de caméras. Celui-ci devra se déplacer le plus possible au centre du conduit et être capable d'enregistrer des données géologiques sur l'environnement sous-marin.

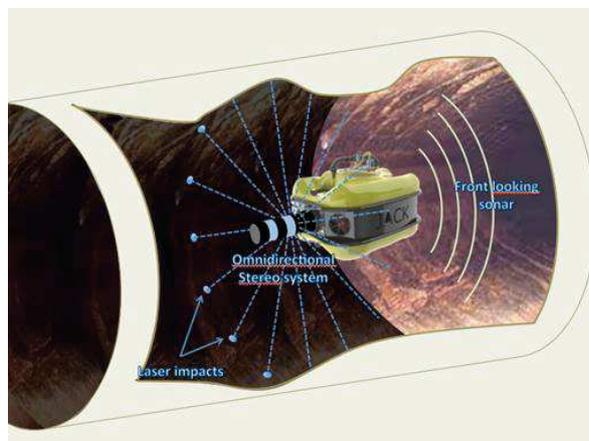


Figure 1 : Système stéréoscopique à ceinture laser omnidirectionnelle dans un environnement confiné

L'objectif de ces travaux de thèse est de mettre au point une application qui sera dédiée à la robotique mobile sous-marine basée sur la notion de la vision et l'approche de la reconstruction 3D permettant de restituer la structure d'un environnement inconnu sous-marin en se basant uniquement sur des données de perception. Afin d'atteindre cet objectif, on a été mené à développer un prototype d'essai nous permettant en premier lieu de proposer tout d'abord un système de vision générant une image compréhensible de l'environnement et de proposer un processus algorithmique permettant de restituer la structure 3D d'un milieu confiné aquatique tout en respectant d'une part les contraintes temps-réel exigées par l'application et d'autres part, les conditions physiques et optiques de l'environnement.

2. Problématique

De point de vue système d'acquisition, il existe de nombreuses technologies permettant de fournir des mesures 3-D. Dans les travaux présentés dans ce mémoire, une étude a été consacrée aux systèmes de vision 3D à large champ. Cette étude permet de choisir judicieusement les éléments du système de vision et sa configuration afin de répondre au mieux au besoin de l'application tout en tenant en considération les contraintes de l'environnement qui sera exploré. La vision omnidirectionnelle est privilégiée car elle augmente le champ de vision des capteurs, notamment l'utilisation des capteurs catadioptriques (combinaison de miroir(s) + caméra(s)) qui représentent une solution simple et rapide pour atteindre une vue large satisfaisante. Une estimation des paramètres intrinsèques et extrinsèques de l'ensemble permet la modélisation géométrique du système de vision représentant au mieux son comportement physique et optique dans l'environnement exploité. Cette modélisation sera par la suite utilisée pour l'estimation de la structure tridimensionnelle de la scène en se basant sur l'approche de la triangulation stéréoscopique. Des données métriques seront aussi estimées et peuvent être utilisées pour le contrôle de la navigation de l'engin et lui assurant une navigation sans danger.

De point de vue algorithmique, plusieurs travaux ont été proposés dans la littérature considérant les contraintes liées aux conditions environnementales d'un milieu sous-marin. En effet, à cause des propriétés physiques de l'eau (les particules d'eau, l'absorption de la lumière et la diffusion) les méthodes standards de traitement d'images ne peuvent plus être utilisées. Les plus grandes longueurs d'onde sont les plus absorbées et par conséquent, les images sont faiblement contrastées et manquent de luminosité et de texture. D'autre part, le phénomène de diffusion de la lumière sous l'eau est différent de celui dans l'air et affecte principalement le comportement physique du capteur d'images. Par conséquent, les paramètres standards (données par le constructeur) caractérisant le capteur ne peuvent plus être utilisés dans ce cas.

3. Contributions

Dans le cadre des travaux de cette thèse, nous proposons de développer un système 3D capable de fournir des informations utiles à analyser et en extraire la structure 3D d'un milieu aquatique de forme

inconnue à l'aide d'une chaîne de traitement d'image. L'objectif est de pouvoir s'approcher le plus possible des conditions d'expérimentations réelles.

Notre contribution concerne différents aspects :

- Développement d'un système 3D permettant de répondre simultanément aux contraintes de l'environnement et au problème de reconstruction 3D de milieux sous-marins.
- Etalonnage du système stéréoscopique omnidirectionnel pour l'estimation de ses paramètres intrinsèques et extrinsèques représentant son comportement dans un environnement aquatique.
- Proposition d'un algorithme permettant d'estimer la structure 3D de l'environnement exploré et ses données métriques : la modélisation géométrique du système de vision permet de remonter à l'espace 3D tout en prenant en compte la nature physique et optique des phénomènes qui interfèrent lors de l'acquisition.
- Proposition d'une architecture matérielle synthétisable de l'algorithme proposé permettant d'assurer les performances nécessaires à notre système en termes de temps d'exécution, de précision et de ressources.

4. Plan du mémoire

Ce rapport de thèse est organisé en quatre chapitres répartis selon deux grandes parties.

- La première partie, "Etude Théorique", est consacrée à l'étude de travaux existants s'intéressant aux applications de vision. Cette partie est subdivisée en deux chapitres (chapitre 1 et 2).

- Le premier chapitre présente un bref historique sur les domaines d'applications basées sur la vision. Nous proposons ensuite une étude sur les systèmes à large champ de vision en s'intéressant particulièrement aux capteurs catadioptriques et leur modélisation géométrique. Quelques exemples de systèmes de vision existant dédiés à la reconstruction 3D d'environnements urbains et sous-marins seront cités.

- Dans le deuxième chapitre, nous présentons une taxonomie des algorithmes liés à l'étalonnage de systèmes omnidirectionnels, les algorithmes de détection et correspondance de points caractéristiques (des lasers en particulier) et les méthodes de reconstruction 3D. Nous nous intéressons aussi à la réfraction des rayons lumineux dans les milieux marins pour essayer de prendre en considération ce phénomène lors du processus de traitement des images.

- La deuxième partie, "Etude Expérimentale", est consacrée à la validation de notre application à travers des simulations logicielles et matérielles et à l'évaluation de ses performances en termes de précision et temps d'exécution. Cette partie est aussi subdivisée en deux chapitres (chapitre 3 et 4).

- Dans le troisième chapitre, nous décrivons le processus de fonctionnement de notre système

en essayant de mettre en ordre les différentes étapes clés liées à notre application. Nous présentons, ensuite, le système stéréoscopique ainsi que l'approche d'étalonnage choisie. Etant donné que les expérimentations sous l'eau sont parfois contraignantes, en particulier pour la phase de l'étalonnage, nous proposons une approche qui montre la possibilité d'exploiter les paramètres d'étalonnage effectué dans l'air pour des applications sous l'eau. Des résultats de simulation sont donnés à la fin de cette section. Nous présentons après des descriptions détaillées des algorithmes utilisés pour la détection des points laser, leurs correspondances et la transformation 2D/3D dans l'eau. Des résultats de simulations, discutant la précision et les performances temporelles de l'algorithme, sont donnés à la fin de ce chapitre.

- Le quatrième chapitre sera consacré à la présentation de l'implémentation matérielle de toute l'architecture de reconstruction 3D conçue sous System Generator. Le flot de conception de l'outil de développement et son intérêt pour le prototypage rapide seront présentés. Toute l'architecture matérielle sera étudiée en détails en décrivant chaque bloc. Les résultats de co_simulation seront présentés pour la validation de l'architecture conçue. Des résultats de synthèse de toute l'architecture sont donnés pour montrer les performances de notre système lors de son implémentation sur FPGA.

5. Contributions scientifiques

Les travaux présentés dans ce mémoire ont donné lieu à une publication dans le journal Applied Science citée ci-dessous et a ouvert l'enquête sur d'autres idées et approches qui seront valorisées dans des publications futurs.

□ Rihab Hmida, Abdesslem Ben Abdelali, Frédéric Comby, Lionel Lapierre, Abdellatif Mtibaa, René Zapata, "Hardware Implementation and Validation of 3D Underwater Shape Reconstruction Algorithm Using a Stereo-Catadioptric System", Applied Sciences Journal, 6(9):247, August 2016.

Partie I.

Etude Théorique : Etat de l'Art

Chapitre I.

Etat de l'Art : Vision, Reconstruction 3D et Robotique Mobile

1. Vision : Généralités et Systèmes de vision

1.1 Généralités

1.1.1. Historique

Au moyen-âge, le mathématicien perse **Ibn al-Haytham** (dit Alhazen) a expliqué le phénomène de la vision. Ce sont, en effet, des rayons lumineux émis par chaque point d'un objet, qui sont en mouvement en ligne droite et formant une image sur le cristallin de l'œil.

En 1637, **René Descartes** a publié son livre "La Dioptrique" dans lequel il détaille quelques phénomènes en optique géométrique, notamment la réfraction. Il décrit aussi le fonctionnement de l'œil où l'image se forme sur la rétine. En 1704, **Isaac Newton**, s'étant beaucoup intéressé aux couleurs, a publié son traité "Opticks" où il expose son explication sur la lumière mécaniquement. Il décrit les différentes couleurs comme un flot de particules ou corpuscules dont la vitesse n'est pas la même. Il explique ainsi le phénomène de réfraction.



Figure I.1. Photos de Alhazen, René Descartes, Isaac Newton, Thomas Young et Samir Zeki

La physiologie de la perception visuelle et des couleurs en particulier a réellement progressé avec **Thomas Young**. Son apport au domaine de l'optique est sans doute son plus grand motif de célébrité, en particulier sa célèbre expérience de la double fente en 1801. Young, considéré comme le fondateur de l'optique physiologique, s'intéressait aussi aux récepteurs sensoriels de l'œil et découvrit les couleurs primaires : rouge, vert et bleu (RVB). Il proposa donc l'hypothèse, confirmée plus tard, que la vision humaine n'utilise que les trois capteurs RVB pour réaliser la synthèse de toutes les autres couleurs [1].

Au cours des années 1970, **Samir Zeki**, un physiologiste turc, a prouvé que le traitement de l'information visuelle dans le cerveau s'organise en différentes voies séparées et spécialisées:

- reconnaissance du mouvement, de la forme, de la séparation figure/fond, organisation de la scène dans l'espace,
- reconnaissance des objets, des visages, et perception des couleurs.

Au cours des années 2000, la puissance des ordinateurs devient considérable, il devient possible à de nombreux ordinateurs de traiter le son et la vidéo, l'apparition de la réalité augmentée, la synthèse d'images de scènes complexes, l'apparition des systèmes d'indexation et de recherche d'images, le développement des approches statistiques et les méthodes temps réel.

1.1.2. Paradigme de Marr

Il y a vingt ans, David Marr proposa une approche computationnelle de la représentation et de la manipulation des informations visuelles chez l'homme et pour la machine. Ses travaux présentent le plus bel exemple d'élaboration d'une "Théorie de la vision" en proposant un modèle pour le traitement et la représentation de l'information visuelle. C'est en effet la première méthodologie complète pour la vision par ordinateur, plus connue sous le nom de "Paradigme de Marr". Ses recherches se sont principalement focalisées sur la reconstruction de scènes basées souvent sur les modèles mathématiques reposant sur des fonctions continues, la géométrie euclidienne, l'analyse différentielle [2].

Trois niveaux de représentations avaient été introduits par Marr :

- ✓ Traitement bas-niveau : la représentation des propriétés de l'image 2-D en termes de primitives. Ce premier niveau de représentation est appelé "première ébauche",
- ✓ Traitement moyen-niveau : la représentation tridimensionnelle des surfaces visibles afin d'extraire une première représentation spatiale de la scène sans la reconstruire en 3D car les informations retrouvées restent insuffisantes pour décrire la localisation des objets dans l'espace tridimensionnel. Cette représentation est appelée "ébauche 2.5D",
- ✓ Traitement haut-niveau : la représentation centrée objet de la structure tridimensionnelle en se basant sur les informations 3D de la scène issues de la caméra et la carte de profondeurs obtenue précédemment.

1.1.3. Vision par ordinateur

Bajcsy a abordé sous un angle différent le problème de la vision, en ne considérant pas la vision comme un problème mal posé, mais plutôt comme un problème de contrôle. Alors que Bajcsy vise le contrôle entre les différents modules d'un système de vision [3], Aloimonos s'intéresse plus aux capteurs [4] en mettant en avant le but pour lequel un système de vision doit être

construit. Cette approche est connue sous le nom de "vision intentionnelle". Une autre approche inspirée de la vision par ordinateur est la "vision animée", présentée par Ballard [5]. Cette approche s'oppose à l'approche reconstructionniste de Marr [6] et s'approche des approches comportementales comme celle présentées par Brooks [7].

Selon Trivedi et Rosenfeld [8], la vision artificielle est abordée par trois écoles scientifiques différentes: la neurophysiologie, psychologie perceptive et la vision par ordinateur. Le premier aspect tente de comprendre le fonctionnement des senseurs et des mécanismes neuronaux des systèmes biologiques. Le deuxième tente de comprendre les aspects psychologiques qui induisent la tâche de perception. La vision par ordinateur, quant à elle, étudie les aspects computationnels et algorithmiques liés à l'acquisition, au traitement et l'interprétation des images. Elle permet en outre d'automatiser le système de perception en question afin de l'adapter au maximum des situations et conditions rencontrées: reconnaissance de forme, reconnaissance de visage, restitution 3D, navigation de robot...etc.

La vision par ordinateur, concerne la compréhension d'une scène ou d'un phénomène à partir des informations contenues dans l'image (Fig. 2). Celle-ci est divisée, en abscisse et ordonnée, en un certain nombre de points appelés pixels. L'extraction des données décrivant la scène 3D est réalisé par analyse des pixels formant son image, appelé traitement d'image. Le processus de traitement d'image est généralement divisé en trois parties. La première est le prétraitement, permettant de corriger les bruits de capture. Le second est le traitement de base qui résume au maximum l'information utile délivrée par l'image d'entrée. Le dernier est appelé post-traitement car il vient corriger subjectivement les défauts de traitement de l'étape précédente et rapprocher les données résultantes aux objectifs tracés au préalable.

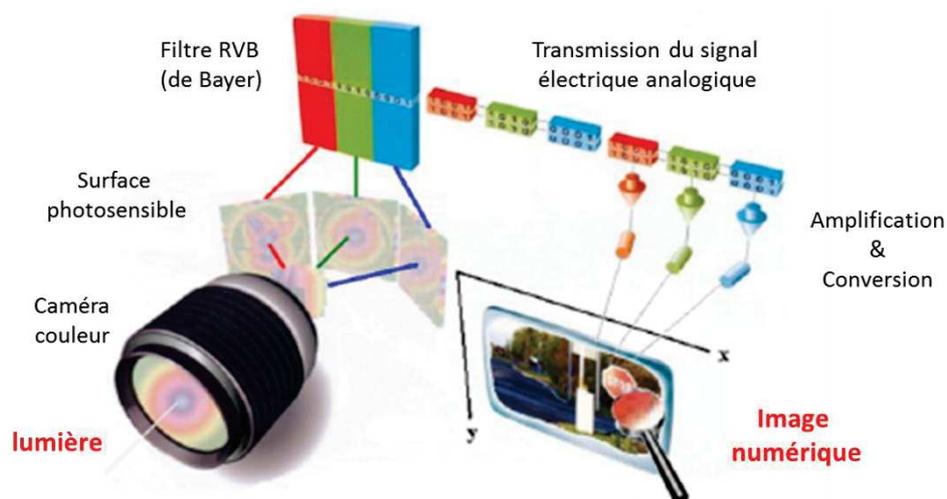


Figure I. 2: Chaîne d'acquisition d'une scène réelle à l'image numérique

L'acquisition d'images est réalisée par l'intermédiaire d'une caméra vidéo. Chaque caméra vidéo comprend un système optique qui forme une projection bidimensionnelle de l'image reçue par sa rétine électronique. Ces capteurs numériques sont omniprésents dans notre vie quotidienne sous différentes formes pour multiples applications: appareil photographique, vidéo-conférence, webcam, et bien sûr dans nos téléphones portables. Les domaines d'application de la vision sont variés, citons par exemple l'utilisation croissante de la vision dans la médecine (IRM, scanner,...), la défense (véhicules autonomes, sécurité,...), l'astronomie, la vidéosurveillance (détection, reconnaissance,...), la robotique (contrôle de navigation, détection d'obstacles,...) ...etc.

1.2 Systèmes de vision

L'objectif visé des premiers travaux de recherche en vision artificielle était de munir les robots d'un système artificiel ayant un fonctionnement qui soit le plus similaire possible au système humain. Plusieurs travaux de recherche sont, jusqu'à ce jour, focalisés sur l'analyse et l'interprétation des phénomènes liés à la vision humaine (œil + cerveau) dans le but de trouver le dispositif matériel qui représente le mieux la rétine humaine et traduire les propriétés du processus visuel humain sous la forme d'un algorithme "intelligent".

Parce que le cerveau est une des plus complexes inventions de la nature, il reste et restera pour longtemps le bastion encore inconnu que les sciences humaines se proposent de conquérir. Du côté du système d'acquisition et avec le progrès technologique, plusieurs systèmes de vision ont été lancés tout en s'inspirant du système de vision animale. Le marché des capteurs explose, offrant à l'utilisateur de plus en plus de produits technologiques pour bien choisir "son capteur" qui répond le plus aux contraintes de son application, notamment la résolution spatiale, la sensibilité, le bruit, le taux de transfert... On cite par exemple, les capteurs d'images monoculaires (Fig. 3. a), les systèmes stéréoscopiques (Fig. 3. b) et les systèmes "tri-focal" (Fig. 3. c).



Figure I. 3: Exemples de systèmes de perception

Les systèmes de perception 2D sont utilisés dans plusieurs applications telles que la vidéo surveillance, la reconnaissance de forme, le contrôle de qualité, néanmoins l'absence de l'information de relief et l'orientation spatiale est une limitation reconnue. En médecine, par

exemple, certaines anomalies s'avèrent être difficiles à visualiser en échographie 2D. Elles sont plus accessibles avec une échographie 3D qui fournit plus de détails sur l'élément concerné et permet de pouvoir orienter le processus de traitement médical (changer ou stopper un traitement). Dans le domaine de la robotique, il est important d'avoir le maximum d'informations de l'environnement visité pour pouvoir autonomiser la navigation du robot, assurer sa localisation précise et détecter et analyser le mouvement d'objets présents dans son entourage. Dans ce contexte, l'idée d'augmenter le champ de vision s'est avérée importante dans l'objectif d'obtenir une image contenant un maximum d'informations de la scène 3D.

1.2.1 Systèmes passifs / actifs

Les systèmes passifs n'utilisent que des caméras pour les acquisitions. Ils peuvent utiliser une ou plusieurs caméras, mais doivent prendre au moins deux images de l'objet de différents angles de vue pour réaliser une reconstruction 3D. Les informations 3D sont obtenues par triangulation (Fig. 4) en faisant correspondre les pixels entre chaque couple d'images et en utilisant la différence entre les coordonnées des correspondants. Ces méthodes utilisent les points caractéristiques de la surface des objets, ce qui rend difficile l'analyse d'objets faiblement texturés.

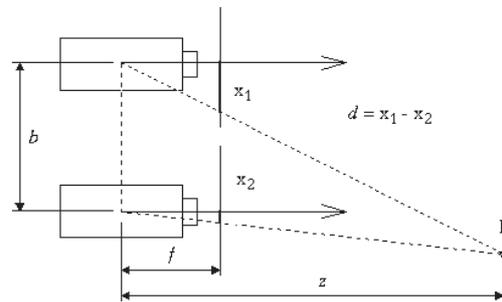


Figure I. 4: Principe de stéréovision passive

Dans [9], un dispositif de vision 3D est mis en œuvre destiné à la robotique médicale pour des manipulations chirurgicales endoscopiques précises. Plusieurs algorithmes de stéréovision passive ont été utilisés pour fournir des images de disparité et permettre d'automatiser et améliorer la précision des interventions chirurgicales. Le banc d'essai est un montage mécanique simple de deux vidéo-endoscopes avec une base entre les centres optiques offrant un champ de vue commun de 60° (Fig. 5). La paire de caméras étant calibrées, un processus de traitement des images stéréoscopiques génère une carte de profondeur en utilisant des algorithmes de détection de points, de correspondance et de reconstruction 3D. Pour que le système développé répond aux contraintes temps-réelles compatibles avec les temps d'exécution des actes chirurgicaux, une architecture décrite en langage VHDL synthétisable a été proposée et implémentée sur une plateforme matérielle.

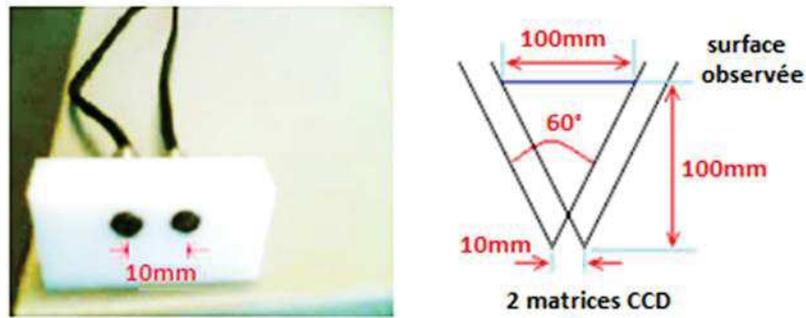


Figure I. 5: Banc stéréoscopique pour endoscopie médicale

Quant aux systèmes actifs, leur principe est de remplacer une des caméras par un appareil qui projette une lumière structurée sur l'objet. Par conséquent, l'extraction d'indices visuels dans la scène est beaucoup plus facile et basée sur la recherche des motifs projetés par la source lumineuse. Cette lumière peut être un rayon laser (point) utilisé avec une caméra classique (Fig. 6. a), un plan laser (ligne) utilisé avec une caméra classique (Fig. 6. b) ou un projecteur de lumière sur un motif structuré (bande) avec la caméra (Fig. 6. c). La lumière peut être aussi générée par un projecteur LCD comme le cas de la "KINECT" qui permet à l'aide de son capteur émetteur infrarouge de projeter des points équidistants sur l'objet 3D (Fig. 6. d). Si le système est calibré géométriquement, il est facile de calculer les informations 3D à très haute qualité des points lumineux à partir de la carte de profondeur générée par la caméra. Même si la caméra KINECT a été initialement conçue pour les jeux vidéo, elle a atteint une grande popularité dans la communauté scientifique où les chercheurs ont développé une énorme quantité d'applications innovantes qui sont liées à différents domaines tels que la reconstruction 3D en ligne [10], [11], [12], les applications médicales et de santé [13], [14], la réalité augmentée [15], etc.

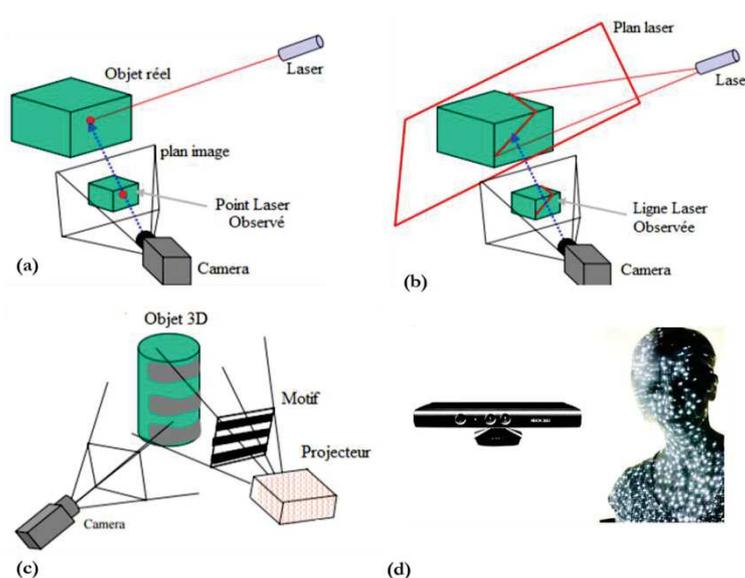


Figure I. 6: Exemples de systèmes actifs à différents motifs de projection laser

Des informaticiens de l'université technique de Braunschweig ont développé une nouvelle technologie et ont reçu le prestigieux prix de l'association allemande pour la reconnaissance des formes. Le premier logiciel "DAVID 4 PRO" a proposé de nombreuses fonctionnalités de numérisation à 360° entièrement automatisée [16]. Il permet de reconstruire toutes formes 3D en utilisant seulement une simple caméra et un laser qui projette une ligne dans la scène photographiée (Fig. 7) Ce logiciel a été amélioré pour donner naissance à une nouvelle solution "SDK DAVID" qui permet de donner des yeux aux robots [17].



Figure I. 7: Exemples de système actif avec projecteur de lumière

1.2.2 Systèmes à large champ

La vision omnidirectionnelle est le procédé de vision qui fournit une sphère de vue du monde observé à partir de son centre. L'image omnidirectionnelle est le résultat d'une vision avec un champ couvrant les 360° par rapport à l'axe vertical. La vision panoramique n'est qu'une réduction de la vision omnidirectionnelle et elle fournit qu'une bande de vue (Fig. 8).

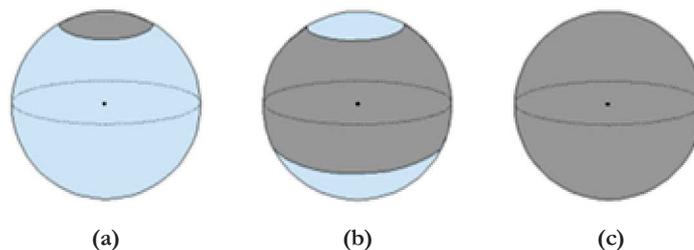


Figure I. 8: Champ de vue des systèmes de vision : (a) Vision classique, (b) Vision panoramique, (c) Vision omnidirectionnelle

Les applications telles que la surveillance, la réalité virtuelle, la vidéo-conférence et bien sûr la navigation des robots autonomes peuvent être grandement améliorées par la disponibilité d'un champ de vue très étendu. Pour obtenir des images omnidirectionnelles, les approches existantes sont initialement inspirées du système de vision animale (Fig. 9). Elles peuvent être distinguées selon les trois catégories suivantes:

- utilisation d'une caméra en rotation autour de son axe,

- utilisation d'images multiples (deux ou plus),
- utilisation d'objectifs spéciaux (fish-eye),
- utilisation de miroirs convexes (caméras catadioptriques).



Figure I. 9: Systèmes à large champ de vue

Ce qui caractérise ces différentes approches est le compromis entre la résolution de l'image obtenue et la vitesse d'acquisition. Un objectif «fish-eye" offre le champ de vue le plus large mais sa modélisation complexe et les distorsions géométriques au niveau des images rend son utilisation assez restrictive. L'emploi des images multiples a l'avantage de fournir une grande résolution, mais au prix temps d'acquisition assez long, les problèmes de synchronisation et la fusion d'images. Avec l'utilisation de miroirs convexes ou d'objectifs spéciaux, l'acquisition des images à la fréquence d'acquisition de la caméra utilisée est possible, au détriment d'une résolution plus faible. Toutefois, pour des applications de navigation, une acquisition rapide est indispensable pendant les déplacements, et une bonne résolution des images n'est pas toujours nécessaire. Dans ce contexte, ce sont les caméras panoramiques catadioptriques qui sont les plus utilisées surtout qu'ils satisfont la théorie du point de vue unique (SVP : "Single View Point") qui sera définie dans les sections suivantes selon le modèle du capteur catadioptrique, en particulier le miroir qui lui est associé.

1.3 Les systèmes catadioptriques

Un capteur catadioptrique repose sur l'association d'une caméra perspective et d'un miroir de révolution, qui en réfléchissant les rayons lumineux provenant de toutes les directions, forme une image omnidirectionnelle une fois projetée sur le capteur. Ces dispositifs se nomment des capteurs catadioptriques : dioptrés pour la réfraction (lentilles) et catoptrique pour la réflexion (miroirs). La réflexion des informations visuelles sur le miroir et puis sa projection sur la caméra

offre une quantité de données plus importante qu'une simple caméra perspective. L'utilisation de caméras catadioptriques est assez populaire dans les applications de robotique mobile. Elle est en général plus légère, plus compacte, plus rapide, plus efficace, moins chère, et de plus, ne demande pas un traitement trop complexe. Comparé aux autres systèmes de vision à large champ, de plus en plus de chercheurs sont intéressés à l'étude des systèmes catadioptriques.

Selon Baker et Nayar, les systèmes catadioptriques peuvent être classifiés en deux catégories: les systèmes catadioptriques centraux et les autres. Les systèmes catadioptriques centraux sont ceux qui vérifient la contrainte du point de vue unique. Celle-ci est définie comme la condition que pour chaque point 3D dans l'espace, une direction bien définie lui est associée, intersectant le miroir et convergeant vers son point focal noté "effective view point" et se projetant sur le plan image en un seul point pixel correspondant. Avec un système vérifiant la contrainte d'unicité, il est donc possible de reconstruire n'importe quelle image perspective à partir de l'image omnidirectionnelle (Fig. 10).

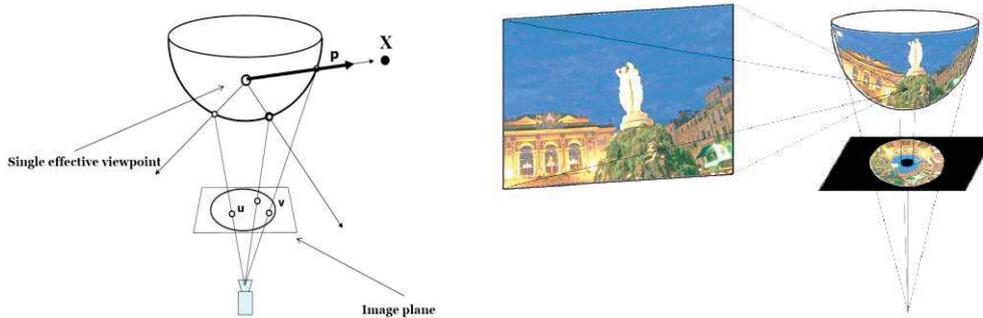


Figure I. 10: Projection d'une image perspective sur le plan image d'un capteur catadioptrique

Les dispositifs catadioptriques centraux regroupent les miroirs paraboliques, hyperboliques, elliptiques et plans. Certains capteurs utilisant des miroirs coniques ou sphériques peuvent également, par approximation, être considérés comme ayant un point de vue unique. L'équation généralisée de la contrainte du point de vue unique (Eq.1) a été proposée par Baker et Nayar et détaillée dans [18]. Cette expression est définie pour tout point $p_s(x, y, z)$, de rayon $r = \sqrt{x^2 + y^2}$ par rapport au centre du miroir positionné à une distance "d", par rapport au centre optique de la caméra, appelée excentricité. La variable "k" est une constante paramétrique positive qui selon sa valeur définit la forme du miroir. Elle est supérieure à 2, dans le cas d'un miroir de forme hyperbolique.

$$\left(z - \frac{d}{2}\right)^2 + r^2 \left(1 - \frac{k}{2}\right) = \frac{d^2}{4} \left(\frac{k-2}{k}\right) \quad (1)$$

La résolution de cette équation aboutit à différents types de miroirs, réalisables ou non, pour la construction de capteurs catadioptriques vérifiant la contrainte du point de vue unique. En d'autre terme, dans l'assemblage caméra-miroir, certaines conditions de construction sont exigées pour que le système vérifie la contrainte de point de vue unique:

- Un système à miroir plan possède un point de vue unique, mais un petit champ de vision.
- Un système à miroir ellipsoïdal possède un centre de projection unique, situé sur le premier foyer du miroir, mais ayant un champ de vision très occulté par la caméra.
- Un système à miroir sphérique ou un miroir conique, doit avoir une caméra sur la surface de la sphère ou sur le sommet du cône ce qui rend inutilisable ces capteurs.
- Un système à miroir parabolique (Fig. 11. a) doit être muni d'une optique télécentrique réalisant une projection orthographique. L'inconvénient réside dans la taille de l'objectif télécentrique.
- Un système à miroir hyperboloïdal (Fig. 11. b) doit coïncider parfaitement le centre optique de la caméra avec le second foyer de l'hyperboloïde ($F=C$).

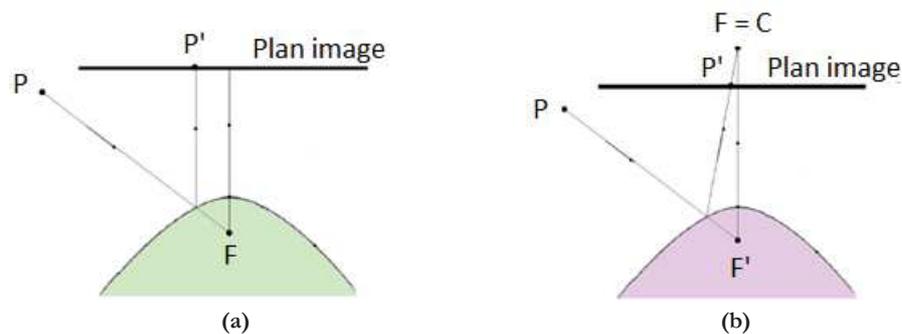
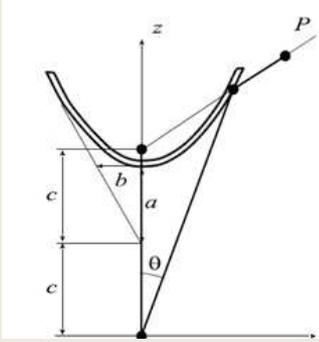
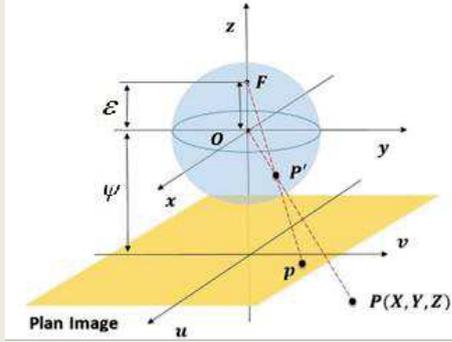


Figure I. 11: Exemple de projection sur un miroir : (a) parabolique et (b) hyperbolique

1.3.1 Modélisation géométrique

Dans [18], les auteurs ont présenté la résolution du modèle géométrique correspondant au système catadioptrique et permettant de définir, pour chaque forme de la quadrique, l'équation de projection d'un point monde sur le miroir. Dans le cas d'un système catadioptrique à point de vue unique, l'équation de projection sur un miroir hyperbolique est issue de l'équation généralisée (Eq.1). Elle peut s'écrire selon le modèle hyperbolique utilisant les paramètres (a, b) du miroir comme donné dans l'expression (Eq.2) ou le modèle générique utilisant les paramètres (ζ, ψ) de la sphère d'équivalence comme donné par les expressions (Eq.3). La variable " p " est un paramètre du miroir appelé "latus rectum" et proportionnel à son diamètre " D " et " d " est l'excentricité permettant d'assurer la contrainte de point de vue unique. Dans [19], Mouaddib a présenté les expressions des différents modèles de projection d'un point 3D en un point image selon le modèle hyperbolique (Eq.4) et le modèle générique (Eq.5). L'inconvénient majeur de ce type de

capteur est la relation non-linéaire qui existe entre le point 3D noté $P(X, Y, Z)$ est sa projection sur l'image noté $p(x, y)$. Cette contrainte rend la projection inverse d'un point 3D à partir de ses projections 2D impossible, à moins d'avoir une relation entre ses coordonnées 3D ou plus d'information sur sa position dans le monde réel (une des composantes X ou Y ou Z est connue).

Modèle hyperbolique	Modèle sphérique
	
$\frac{\left(z - \frac{d}{2}\right)^2}{a^2} - \frac{x^2 + y^2}{b^2} = 1 \quad (2)$ $a = \frac{d}{2} \sqrt{\frac{k-2}{k}} ; \quad b = \frac{d}{2} \sqrt{\frac{2}{k}}$ $\begin{cases} x = \frac{2dp.X}{Z(\sqrt{d^2 + 4p^2} + 2p) \pm d\sqrt{X^2 + Y^2 + Z^2}} \\ y = \frac{2dp.Y}{Z(\sqrt{d^2 + 4p^2} + 2p) \pm d\sqrt{X^2 + Y^2 + Z^2}} \end{cases} \quad (4)$	$\varepsilon = \frac{d}{\sqrt{d^2 + 4.p^2}} ; \quad \psi = \frac{d + 2.p}{\sqrt{d^2 + 4.p^2}} \quad (3)$ $d = 2.\sqrt{a^2 + b^2} ; \quad p = \frac{D}{4}$ $\begin{cases} x = \frac{(\xi + \psi) X}{\xi \sqrt{X^2 + Y^2 + Z^2} - Z} \\ y = \frac{(\xi + \psi) Y}{\xi \sqrt{X^2 + Y^2 + Z^2} - Z} \end{cases} \quad (5)$

1.3.2 Modélisation algébrique

La modélisation algébrique présente une description analogue à la modélisation géométrique, en utilisant les paramètres physiques de la caméra. Pour exprimer les relations de projection, on définit quatre repères de coordonnées: Monde, caméra, image et pixel. La projection 3D-2D s'effectue en trois étapes respectives (Fig.12):

1- Projection dans le repère caméra: Le repère Monde est repère statique attaché à la scène 3D. Il peut être choisi de manière arbitraire pour pouvoir représenter la position des points 3-D et des caméras par rapport à ce repère. La position de la caméra est décrite par un vecteur 3D de translation "T" traduisant la translation de l'origine du repère caméra vers le repère Monde et une rotation "R". Soient X, Y, Z les coordonnées du point 3-D, exprimées dans le repère monde. L'expression de projection du repère monde vers le repère caméra est donnée par l'expression suivante (Eq.6).

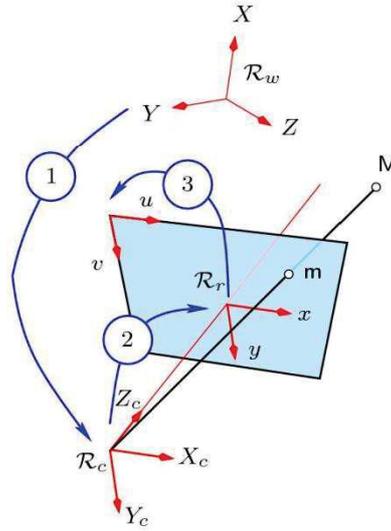


Figure I. 12: Illustration des repères utilisés

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ \mathbf{0} & 1 \end{bmatrix} \times \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (6)$$

2- Projection sur le repère image: Le repère caméra est un repère 3-D orthonormé. Le plan XY est parallèle au plan image. Le repère image est un plan 2-D dont l'origine est le point principal. La distance entre les deux repères correspond à la distance focale notée "f". L'utilisation des coordonnées homogènes permet d'écrire la projection perspective sous forme linéaire (Eq.7).

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (7)$$

3- Projection dans le repère pixels: Pour des traitements d'image, les points sont identifiés par des coordonnées en pixels alors que le repère image est un repère métrique (mesuré en mm). Aussi, certaines caméras génèrent des pixels qui ne sont pas carrés mais ils ont un côté plus grand que l'autre. Le changement de repère nécessite tout d'abord une translation et puis un ajustement selon les directions x et y. La projection des coordonnées métriques (x, y) en des points pixels (u, v) est exprimée par une matrice de paramètres (Eq.8). Les coordonnées (u_0, v_0) présentent les coordonnées du point central de l'image. Les paramètres (k_u, k_v) présentent les densités de pixels en direction respectivement de l'axe des u et l'axe des v (exprimées en nombre de pixels par mm).

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (8)$$

Le système complet est représenté par la transformation donnée comme suit (Eq.9)

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} fk_u & 0 & u_0 & 0 \\ 0 & fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0 & 1 \end{bmatrix} \times \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (9)$$

2. Reconstruction 3D : Systèmes et Méthodes

La première étape dans la reconstruction 3D de structure est l'acquisition des formes qui peut être traitée de différentes manières, comme le montre la taxonomie suivante (Fig. 13).

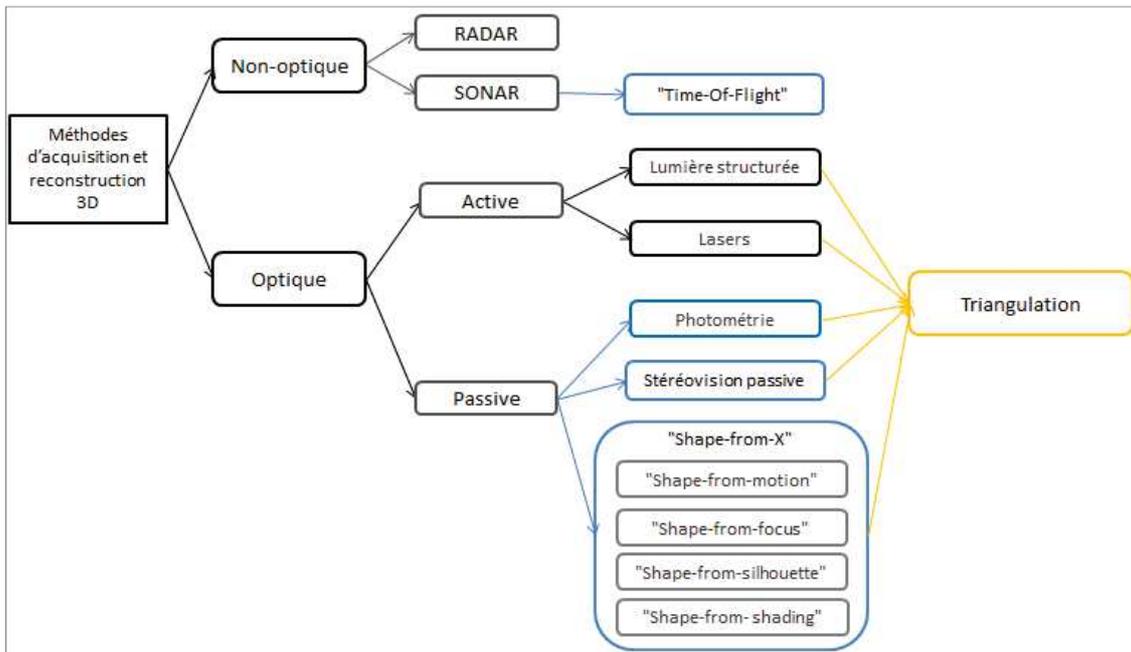


Figure I. 13: Taxonomie des méthodes d'acquisitions de formes

2.1. Les systèmes non-optiques

La robustesse des capteurs non-optiques (Radar, Sonar,...) et plus particulièrement leur portée et leur résistance aux variations environnementales (climat, brouillard, luminosité, etc.) les rend bien adaptés à fonctionner dans des conditions atmosphériques difficiles comme dans le cas des applications de robotiques dans des environnements extérieurs terrestres. Par exemple, les auteurs dans [20] ont introduit une méthode géométrique pour la reconstruction 3D de l'environnement extérieur à l'aide d'une caméra et d'un RADAR panoramique à micro-ondes. Ils se sont appuyés sur la complémentarité de ces deux capteurs, compte tenu de la robustesse du

radar et sa capacité de générer une carte de profondeur et la haute résolution spatiale de la caméra. Dans ce même contexte, plusieurs travaux existent dans la littérature exploitant l'utilisation des SONARs en combinaison, en particulier, avec les méthodes de "Time-of-Flight" surtout pour les longues distances. Ces méthodes sont basées sur l'analyse du temps de propagation du signal qui, en connaissant sa vitesse dans le milieu où il se déplace, nous permet d'estimer la distance de l'objet ayant généré un écho. Ces capteurs sont communément appelés des sonars de profilage, car ils sont principalement utilisés pour recueillir des données bathymétriques [21], [22] et [23]. Néanmoins, l'utilisation de ces dispositifs dans les environnements marins est moins avantageuse à cause des changements de température de l'eau, de la salinité et de la pression [24].

2.2. Les systèmes optiques actifs

L'extraction de primitives à partir d'images de lumière structurée a été utilisée dans plusieurs travaux [25], [26] et [27]. Ces systèmes se composent principalement d'un vidéo projecteur qui projette une séquence de motifs spécifiques de lumière (rayures en noir et blanc ou en niveaux de gris, motifs de lignes colorées, des cibles spécifiques, etc.), tandis qu'une ou plusieurs caméras numériques enregistrent la déformation des motifs projetés sur les objets. Différentes techniques sont proposées pour la reconstruction 3D par triangulation. Dans [28], la combinaison des techniques basées sur l'image et des informations issues du système à laser permet d'offrir une reconstruction très fiable. En face du laser, un élément optique diffractif est inséré permettant de modifier la forme du faisceau en un ensemble de lignes parallèles (Fig. 14). Ces lignes sont projetées sur la scène et récupérées par la caméra. Une fois que la relation entre un pixel et le plan laser est connue, les informations 3D peuvent être calculées par triangulation. L'avantage majeur de ces dispositifs est la grande résolution et la précision obtenue conduisant à la reconstruction de surface 3D de haute qualité. Néanmoins, le principal inconvénient est qu'ils sont limités à des objets statiques et qu'un grand nombre d'images doit être acquis.

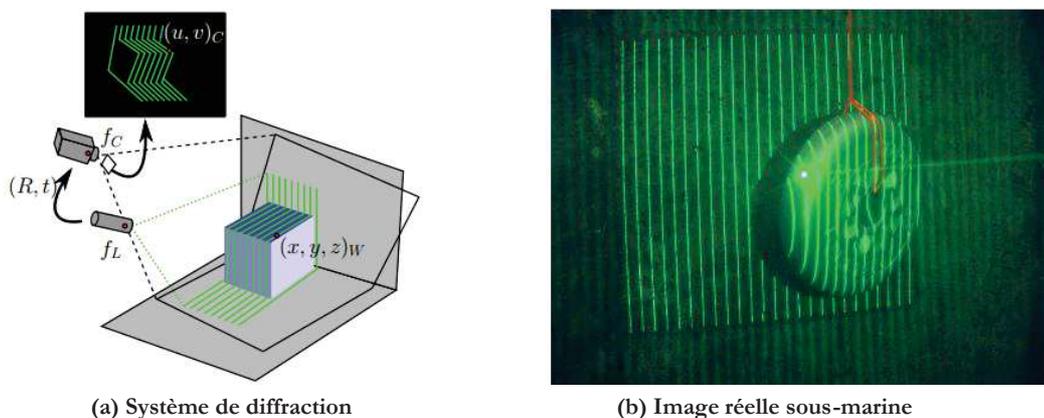


Figure I. 14: Reconstruction 3D avec un système à motifs lasers structurés [28]

2.3. Les systèmes optiques passifs

La reconstruction de la forme 3D d'objets à partir de plusieurs images est l'une des problématiques les plus anciennes du domaine de la vision par ordinateur. La littérature propose peu de méthodes temps réel, qui estiment précisément la forme 3D d'objets filmés par une ou plusieurs caméras (Fig.15). Les approches de reconstruction qui utilisent une seule caméra, s'appuient sur l'extraction de certaines caractéristiques de l'image afin de déterminer l'information de profondeur. Dans le domaine de l'archéologie, l'utilisation de lumière structurée ou les méthodes de scan 3D ont été recommandées pour une meilleure précision de reconstruction à condition de pouvoir transporter l'objet et l'orienter de manière à pouvoir capturer toutes ses faces. Dans le cas contraire, ce sont les méthodes basées sur la stéréovision [29] qui sont les plus utilisées. La stéréoscopie est basée sur l'extraction de points caractéristiques d'une même scène à partir des images acquises simultanément par deux caméras. Ces points sont appariés pour pouvoir calculer leur correspondant 3D. En effet, en connaissant la position relative de la caméra et l'emplacement d'un même point caractéristique dans les deux images, les coordonnées 3D de ce point dans le repère de la scène peuvent être calculées par triangulation. Ces méthodes ont cependant l'inconvénient que la précision est souvent limitée car de nombreux sites ou objets archéologiques sont plutôt sans texture et le problème de correspondance devient difficile.

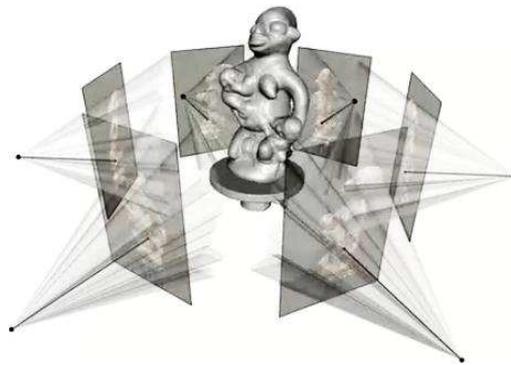


Figure I. 15: Reconstruction à partir de plusieurs images

Afin de pallier au problème de correspondance, les auteurs dans [30] ont proposé de combiner les méthodes de stéréovision avec les techniques de "Shape-from-X" et de photométrie. Le mot "Shape-from-X" est un nom générique pour les techniques d'extraction de formes à partir d'images seulement. La technique "Shape-from-Shading" (SfS) [31], est une méthode de reconstruction qui permet d'améliorer la forme d'un objet à partir des informations d'ombrage issu d'une seule image. Le nombre d'application réelle de cette méthode est assez limité et est principalement dû au fait que les directions des sources lumineuses et les propriétés de réflexion doivent être parfaitement connues pour aboutir à de bons résultats. Dans ce

contexte, des algorithmes de raffinement ont été proposés dans [31], [32] et [33] pour améliorer la qualité de reconstruction. La technique "Shape-from-Motion" (SfM) est une méthode de triangulation qui consiste à prendre plusieurs images d'un objet ou d'une scène en utilisant une seule caméra. À partir des orientations de la caméra, les caractéristiques d'image sont détectées et comparées entre les prises consécutives pour estimer le mouvement relatif de l'appareil et, par conséquent, déterminer sa trajectoire en 3D. Cette approche est la moins chère en termes de matériel et la plus facile à réaliser. La seule restriction est d'avoir suffisamment de mémoire sur l'appareil photo ou l'enregistreur vidéo pour sauvegarder une vue complète de la scène.

La photométrie est une technique de reconstruction qui nécessite un nombre réduit d'images capturées sous différentes conditions d'éclairage. Une sphère est placée dans la scène pour l'étalonnage de la source lumineuse lors de la capture des données (Fig.16, images au-dessous). En modifiant l'emplacement de la source lumineuse, les informations 3D sont obtenues tout en conservant la caméra et l'objet dans une position fixe (Fig.16, images au-dessus). Vu que la profondeur d'une scène dans une image est variable, le volume des particules éclairées est lui aussi variable. Cela implique que, sous une lumière diffusée, la rétrodiffusion est directement liée à la profondeur de la scène ce qui nous permet d'estimer une carte de profondeur [34].

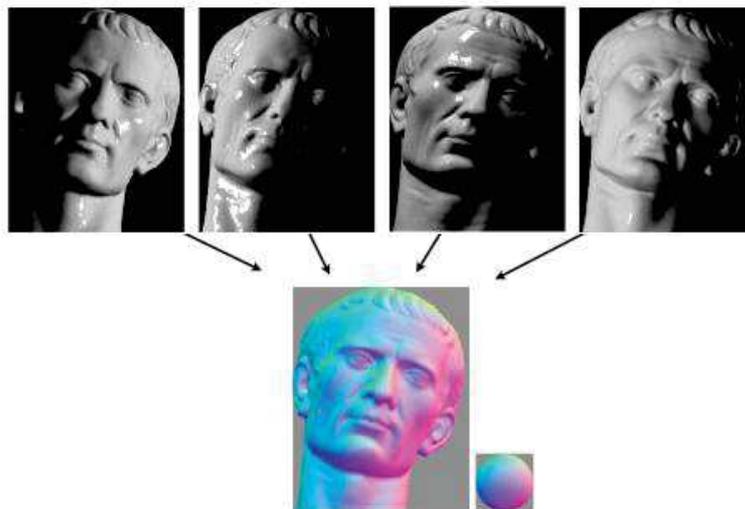


Figure I. 16: Principe de la photométrie [35]

3. Robotique mobile : Généralités et Exemples d'applications

3.1. Généralités

La robotique mobile est une discipline de recherche très variée qui s'intéresse à la conception de robots mobiles pour de nombreuses applications. La première des fonctionnalités que le robot doit réaliser est de percevoir, modéliser et interpréter son environnement, afin d'y adapter ses différentes actions. Le robot est un système mécanique bien équipé lui permettant de

se déplacer dans son environnement en toute sécurité. De même que l'homme se fie à ses cinq sens, un robot autonome exploite pour cela des données fournies par ses capteurs lui permettant de se localiser et de percevoir son environnement et d'actionneurs assurant son mouvement. Pour exécuter une tâche de navigation autonome, le robot mobile est muni d'un système d'algorithmes intelligents qui fusionnent les données recueillies par les capteurs et envoient des signaux de commande aux actionneurs pour réaliser les tâches demandées. Parmi les fonctionnalités attribuées, on peut citer :

- La modélisation de l'environnement: C'est par la construction de modèles de l'environnement que le robot structure et rend exploitable l'ensemble des données qu'il perçoit. Ce problème a été énormément étudié dans la communauté des roboticiens et de nombreuses méthodes de construction de cartes de l'environnement ont été proposées [36], [37], [38] .

- La localisation: C'est l'estimation de la position du robot par rapport à un repère fixe de l'environnement. Selon que le robot se base sur des cartes métriques ou topologiques, le problème de la localisation est abordé différemment [39], [40]. On peut classer les moyens de se localiser en trois grands types de méthodes :

- les techniques d'estimation de mouvement, telles que l'odométrie et la navigation inertielle.
- les techniques de recalage, qui construisent un modèle de l'environnement à partir de données extéroceptives acquises par le robot, et qui utilisent ce modèle pour corriger la position estimée par les techniques d'estimation du mouvement.
- les techniques de localisation absolue, qui mettent en correspondance des données perçues avec un modèle initial, et produisent des estimées de position dont l'erreur est bornée.

- Le SLAM: La réalisation simultanée de la cartographie et de la localisation est un problème fondamental en robotique mobile connu sous le nom de SLAM. Lorsque le robot explore une zone inconnue il doit aussi être capable de se localiser sur la carte qu'il construit. Cette carte servira à la fois au robot à se localiser pour évoluer de manière autonome et à l'utilisateur humain pour récupérer des informations sur les lieux explorés. Plusieurs travaux se sont intéressés à ce sujet comme dans [41], [42], [43] et [44].

3.2. Exemples d'applications

Dans la majorité des applications de robotique mobile, la vision est choisie comme capteur privilégié du fait de la richesse des informations contenues dans une image. L'interprétation de ces informations permet de détecter et extraire des amers, de la surface ou bords de chemin ou d'obstacles afin d'orienter l'avancement du robot sans danger. Les domaines d'applications les plus traités sont la navigation autonome du robot mobile [45], [46], [47], la

reconstruction de scène en 3D [48], [49], [50], la surveillance « intelligente » et le suivi [51], [52] et [53]. Dans la section suivante, on cite quelques exemples d'utilisation de systèmes de vision en robotique.

3.2.1 Exemples de robots terrestres

Dans [54], un système stéréoscopique à base de capteurs catadioptriques à miroirs hyperboliques a été conçu pour la reconstruction dense d'une scène (Fig.17. a). La configuration verticale des deux systèmes permet d'avoir un champ de vue commun entre les deux images ce qui permet d'extraire un nombre important de points caractéristiques. Dans cet article, les auteurs ont proposé un algorithme qui facilite la correspondance entre les points détectés tout en considérant des contraintes globales telles que la contrainte d'unicité, contrainte d'ordre et la continuité de la contrainte de disparité. La carte de disparité est obtenue en utilisant la technique de corrélation croisée normalisée (NCC) et les composantes de l'espace couleur H_1, H_2, H_3 , un nouvel espace défini dans [54].

Un autre exemple de système omnidirectionnel est celui proposé par Christopher Mei dans [55]. Ce système consiste à un assemblage mécanique d'un capteur catadioptrique et un système laser (Fig.17.b) destiné à effectuer de la localisation et de la cartographie simultanée dans des environnements complexes et de grandes tailles. Des algorithmes ont été proposés pour effectuer du SLAM planaire et aussi estimer la position 3D du robot ainsi que la structure de l'environnement. Il a aussi développé une technique d'étalonnage simple des systèmes catadioptriques centraux basée sur le modèle sphérique [56] ouvrant des perspectives dans le domaine de la reconstruction 3D temps-réel soit avec la vision seule soit en combinant vision et laser.

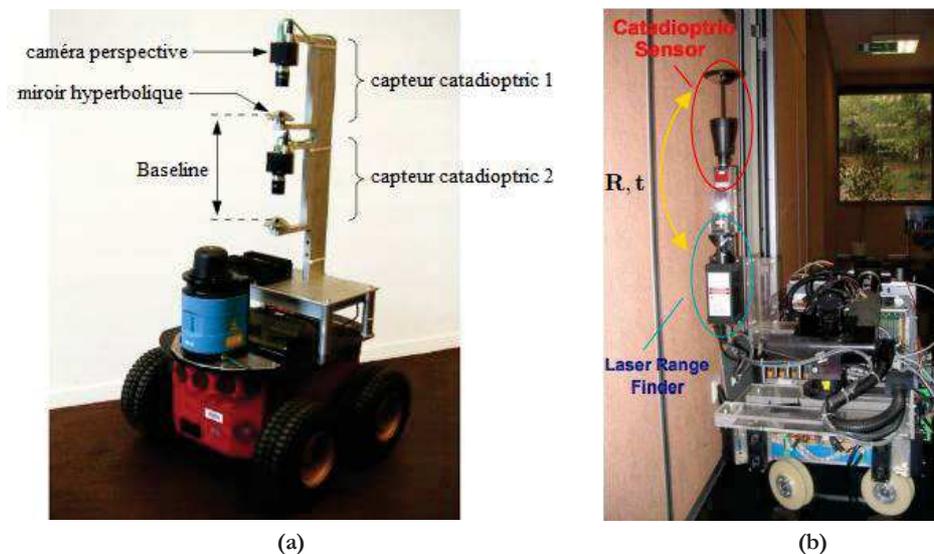


Figure I. 17: Système stéréoscopique de capteurs catadioptriques: (a) Système passif, (b) Système actif

Différentes combinaisons de capteurs catadioptriques à différents type de miroirs et différentes configurations ont été proposées et étudiées dans la littérature. Citons par exemple les travaux présentés dans [57], [41], [36]...etc. Un cas particulier des systèmes catadioptriques est celui qui n'utilise qu'une seule caméra, combinée avec deux miroirs. Dans [58] un système composé d'une caméra unique orientée vers deux miroirs hyperboliques disposés verticalement est conçu pour la détection d'obstacles (Fig. 18. a). Un trou dans le miroir du bas permet d'imager celui au-dessus. Ainsi chaque image acquise, contient deux images catadioptriques concentriques, représentant la réflexion de la scène par chacun des miroirs (Fig. 18. b). La triangulation dans ce cas (Fig. 18. c) est similaire à celle d'un système stéréo binoculaire, la seule différence réside dans l'unicité du plan image.

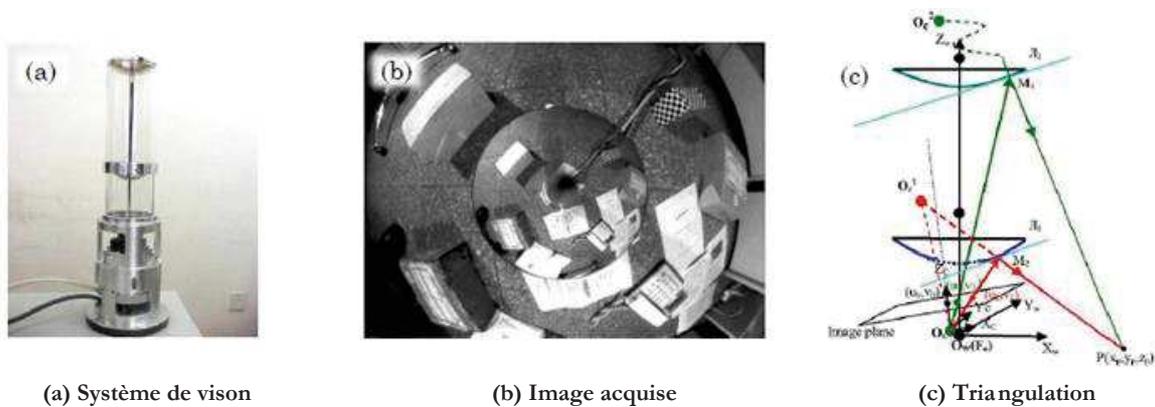


Figure I. 18: Système omni-stéréo binoculaire-V à une seule caméra

Dans [59], un dispositif stéréo destiné à être intégré dans un système d'aide à la conduite permettant d'envisager une perception 3D par stéréovision de la scène routière a été développé. La même combinaison a été mise en œuvre mais avec deux surfaces différentes de miroir: un miroir panoramique extérieur permettant d'avoir une vue sur l'environnement du véhicule et un miroir plan intérieur offrant une image à résolution plus élevée et permettant l'analyse et le suivi des yeux du conducteur (Fig. 19. a). Les points caractéristiques détectés sur le visage du conducteur sont extraits à partir des deux vues (Fig. 19. b), panoramique et dédiée et ils sont utilisés dans le processus de triangulation.

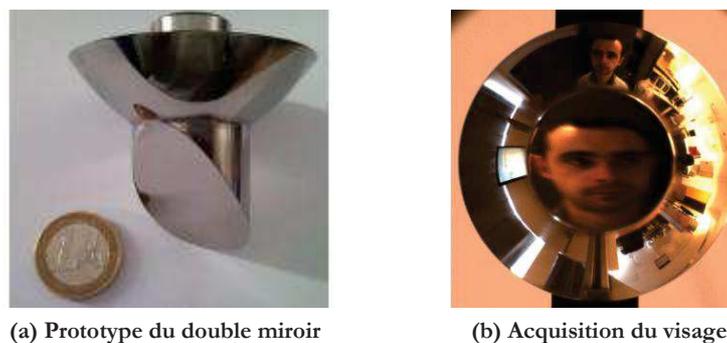


Figure I. 19: Système catadioptrique à double miroir pour observation du conducteur

Un autre système particulier est introduit dans [60] par Orghidan, Mouaddib et Salvi (Fig. 20). Un système générant des faisceaux laser produisant une lumière rouge est assemblé à un système catadioptrique à miroir conique. Ce capteur joue le rôle de la lentille télécentrique au système parabolique d'une part et d'autre part il transforme les rayons laser réfléchis sur sa surface à des points lumineux dans la scène. Le dispositif proposé a été conçu pour donner une information sur la profondeur de la scène observée tout en évitant le problème de la mise en correspondance. La troisième dimension a été calculée en faisant l'intersection entre les rayons laser et les rayons captés par la caméra, autrement dit, par triangulation. Le principal avantage de ce capteur est que la profondeur est calculée simultanément autour d'un champ de vision de 360° ce qui le rend approprié pour des applications en temps réel. Aussi, l'utilisation de la lumière structurée permet de faciliter la détection de points d'intérêt et d'éviter le problème de correspondance.

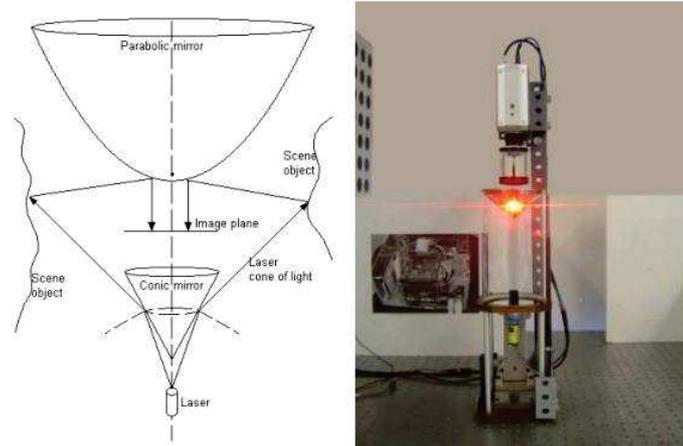


Figure I. 20: Système catadioptrique parabolique et conique avec système laser

3.2.2 Exemples de robots sous-marins

Aujourd'hui, il y a plus que 1000 véhicules sous-marins commerciaux en exploitation effectuant des tâches aussi diverses que les inspections océanographique. Comme dans d'autres domaines, les robots autonomes peuvent entreprendre des missions qui sont trop dangereuses pour l'homme, pour atteindre de plus grandes profondeurs, faire des tâches sous-marines pour des durées plus longues, et d'aller dans des endroits où les humains ne peuvent pas atteindre. Ces véhicules utilisent principalement des inclinomètres ("*Pitch and Roll*") qui indique l'attitude en degrés par rapport à l'horizontale et des caméras panoramiques et d'inclinaison ("*Pan and Tilt*") permettant d'afficher l'orientation de la caméra afin d'aider le pilote téléspectateur de connaître la direction visualisée par la caméra.

Dans les applications de cartographie 3D, les engins sous-marins sont munis de sonars mécaniques numérisés donnant à l'opérateur la possibilité de voir au-delà du système et des

sonars à balayage utilisés souvent en conjonction avec des sonars de profilage. Ces derniers sont idéals pour la mesure et la création de sections de structures sous-marines, néanmoins ils offrent un débit d'informations trop faible pour être utilisé directement dans la boucle de commande du robot. La technologie Vidéo est souvent considérée comme alternative. Cependant, de nombreux facteurs influent sur la qualité de l'image et la marge de visibilité tels que : le type de la caméra, le modèle d'éclairage (type et quantité de lumières), les modes de transmission de données et l'attitude du pilote lors de la navigation du robot sous différentes conditions de l'environnement sous-marin.

Les équipes travaillant dans "*Hibbard Inshore*" ont réalisé une grande variété de projets à l'aide de véhicules télécommandés [61]. Les projets couvrent un large éventail de travaux et d'inspection sous-marine, des applications dans les industries telles que les centrales électriques, les barrages hydroélectriques, les tunnels d'adduction d'eau, les émissaires marins, les ponts,...etc. Ils exploitent de nombreuses tailles et types de systèmes ROV conçus différemment pour pouvoir installer les capteurs et outils existants avec différentes configurations à savoir les caméras à haute définition, les sonars de profilage (Fig.21), les géo-radars (GPR : "*Ground Penetrating Radar*"), des lasers à balayage utilisés pour fournir des images tridimensionnelles d'objets physiques, des lasers de mise échelle utilisés pour estimer la taille d'objet sous-marin à proximité,...etc.

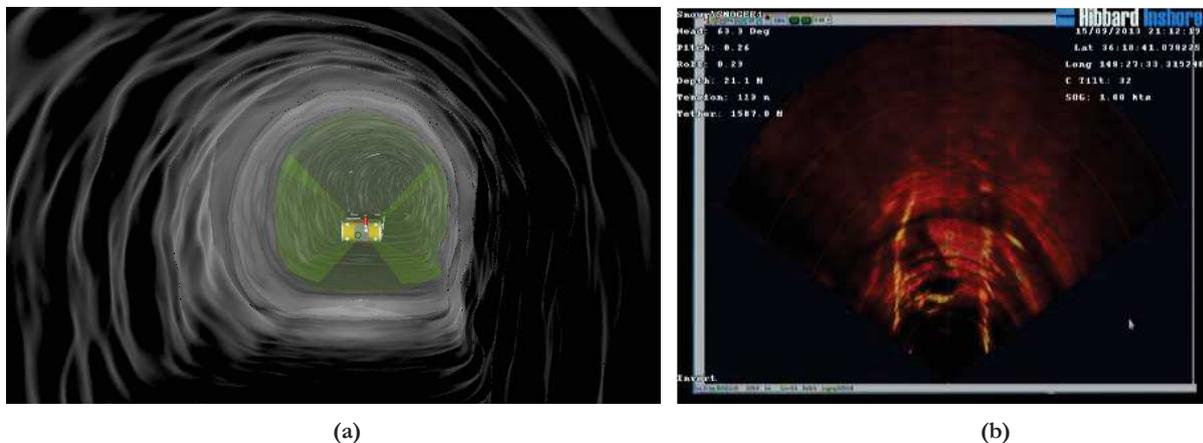


Figure I. 21 : (a) Système d'inspection "Sabertooth" (b) Exemple d'image Sonar

Dans [62], un système à caméras perspectives a été conçu pour la modélisation 3D d'objets archéologiques sous-marins. Ces travaux de recherche ont donné un aperçu sur certains phénomènes de mathématique et de photogrammétrie, en particulier, les influences de l'interface air-verre-eau sur les paramètres intrinsèques de la caméra (Fig.22). Vu que les images sous-marines sont sombres de mauvaise qualité, des algorithmes d'amélioration de la perception sous-marine sont proposés et évalués (Fig.23).



Figure I. 22 : Système stéréoscopique pour application sous-marine



Figure I. 23 : Résultats de l'algorithme de correction sur un exemple d'image sous-marine

Dans [63], un robot autonome sous-marin nommé DEPTHX est conçu pour explorer les cavernes et les tunnels inondés. Pour accomplir cette tâche, le véhicule est équipé d'un enregistreur de vitesse Doppler (DVL), un système laser en anneau à base de gyroscope à navigation inertielle (INS), un capteur de profondeur, et un réseau de 54 transducteurs étroits de sonar à faisceau (Fig.24.a). Afin de produire une carte de son environnement (ici une carrière de calcaire), le véhicule doit effectuer plusieurs missions d'enquête qui se chevauchent, visant à fournir une couverture sonar de toutes les régions visitées. Lors de la construction des cartes 3D, les données sonar recueillies pendant les excursions sont concaténées et utilisées en se basant sur la méthode du plus proche voisin pour construire une carte régulièrement échantillonnée. Figure 24.b représente une carte préliminaire de profondeur représentée par la couleur : de la surface en rouge jusqu'à environ 12 mètres de profondeur en bleu. Ces données ont été collées sur le dessus d'une capture d'écran de 'Google Earth' pour donner un peu de contexte.

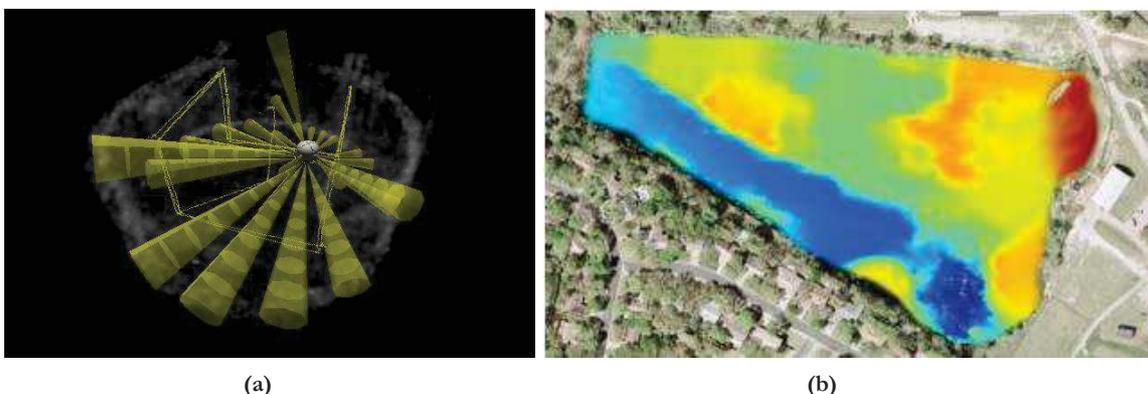


Figure I. 24 : (a) Système de navigation DEPTHX à faisceaux sonar, (b) Carte de profondeur résultante

La plus grande étude qui a utilisé la photogrammétrie sous-marine est l'exploration du Titanic [64]. En 2010, l'expédition "Woods Hole Oceanographic Institution" a réussi l'acquisition de plusieurs images optiques et sonar détaillées du Titanic (Fig.25.a) en utilisant un véhicule télécommandé et deux véhicules sous-marins autonomes. Le résultat de l'expédition était un orthophotomosaïque de haute définition et un modèle 3D du bateau de naufrage basé sur des milliers d'images individuelles et une documentation complexe énorme fournie par les chercheurs (Fig.25.b).

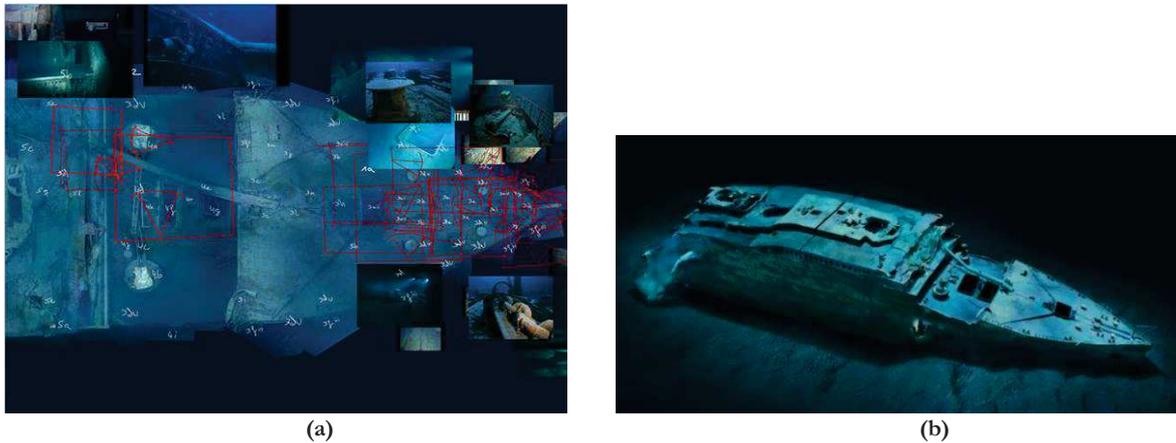


Figure I. 25: Système stéréoscopique pour application sous-marine

Dans [65], un prototype de recherche a été conçu pour une utilisation dans un environnement aquatique (Fig.24.b). Le système est constitué de deux capteurs catadioptriques à miroirs hyperboliques disposés selon une configuration stéréoscopique verticale (Fig.24.a) offrant un important champ de vue commun. Pour la correspondance des points caractéristiques extraits de la paire d'image, une approche basée sur la géométrie épipolaire a été adaptée (Fig.24.c). Une approche de reconstruction 3D, basée sur la méthode du tracé de rayons, a été présentée en considérant les effets de réfraction de lumière et de distorsion causés par les changements des milieux. Cette approche permet, en effet, de corriger les distorsions dans l'image et de fournir une reconstruction plus précise de l'objet immergé dans l'eau [66], [67] et [68].

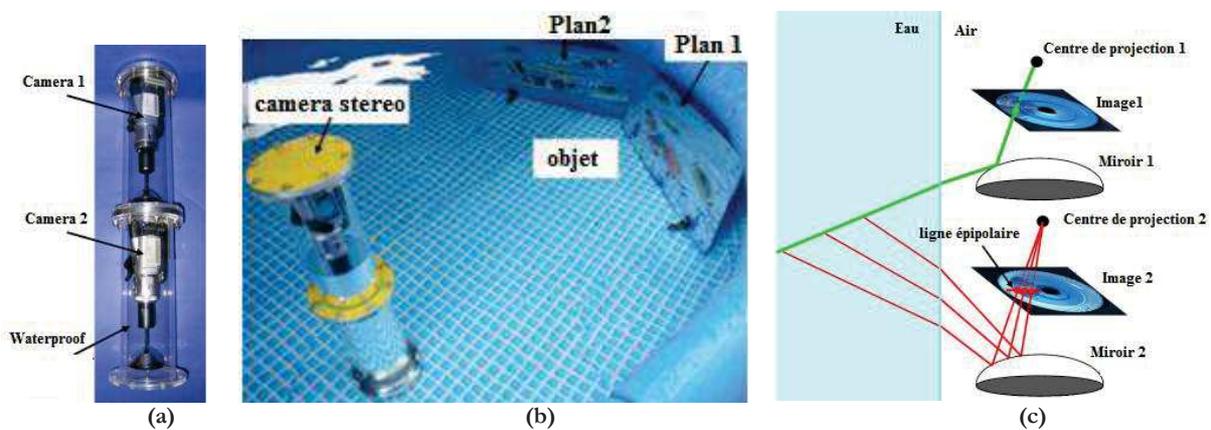


Figure I. 26: Système stéréoscopique pour application sous-marine

4. Conclusion

Dans ce chapitre, nous avons définis les principaux thèmes qui sont liés à nos travaux de recherche à savoir : la vision, la reconstruction 3D et la robotique mobile. Quelques exemples de systèmes de vision ont été présentés tout en mettant en œuvre les avantages d'utilisation des systèmes catadioptriques. Ces dispositifs, caractérisés par leur modélisation moins complexe et offrant un large champ de vision, sont utilisés dans plusieurs domaines d'applications terrestres et sous-marines, en particulier pour la navigation autonome. D'autre part, les méthodes de reconstruction 3D, basées essentiellement sur la stéréovision passive ou active, sont exploitées dans diverses fonctionnalités des systèmes de navigation. Le choix d'une méthode ou d'une autre dépend de la configuration matérielle du système de vision et les conditions de l'environnement de navigation. Dans ce contexte, plusieurs projets ont été mis en œuvre dans le domaine de recherche scientifique aussi bien que industrielle.

Chapitre II.

Taxonomies des Algorithmes Existants

1. Introduction

Quatre étapes sont nécessaires pour retrouver l'information de profondeur:

- Etalonnage/Calibration : Cette étape consiste à déterminer les paramètres internes et externes du modèle géométrique du capteur stéréoscopique.
- Acquisition : Deux images de la même scène sont acquises au même moment par deux caméras différentes ou bien similaires.
- Détection et mise en correspondance stéréo : Il s'agit de retrouver sur les deux images les paires de pixels qui correspondent à la projection d'un même élément de la scène.
- Reconstruction 3D : Elle consiste à calculer pour chaque paire de points 2D la position 3D du point monde dans l'espace.

Ce processus doit se produire idéalement en temps réel, pendant que le robot se déplace à travers le tunnel sous-marin. Une taxonomie des algorithmes de traitement existant en littérature sera présentée dans ce qui suit. Le but principal de cette étude est de définir un algorithme qui soit performant d'une part et qui respecte les contraintes temporelles exigées par l'application, d'autre part.

2. Etalonnage des systèmes omnidirectionnels

Dans cette section, nous présentons les méthodes existantes destinées à calibrer tout système omnidirectionnel central en citant les avantages et les inconvénients de chacun. Tout d'abord, nous présentons une classification de ces méthodes présentant les caractéristiques les plus pertinentes de chaque méthode particulière. Ensuite, nous passons à l'évaluation de la précision d'étalonnage considérant particulièrement la reconstruction métrique 3D.

Le nombre de méthodes d'étalonnage de système catadioptrique central a augmenté ces dernières années. Ces méthodes diffèrent selon le modèle de la caméra : si le système sépare l'étalonnage de la caméra et l'estimation des paramètres du miroir [69], [70] ou s'il est considéré comme un système compact. La plupart des méthodes se réfèrent sur le modèle compact et utilisent soit des vues d'un modèle de test 2D ([71], [72], [73], [74], [75] et [76]), d'un modèle de test 3D ([77], [78]), d'un modèle cylindrique ([79]) et d'autres utilisent des images contenant des lignes caractéristiques ([80], [81], [82], [83], [84] et [85]). D'autres approches sont basées sur l'auto-calibration [86], [87], [88] et [89].

Outre la performance, l'ergonomie, la simplicité d'utilisation, l'algorithme et le type du modèle de test sont à considérer, car ils sont des éléments importants qui peuvent aider l'utilisateur à choisir la meilleure approche. Parmi toutes les approches mentionnées, il existe

plusieurs méthodes d'étalonnage disponibles en ligne sous forme de "*Toolbox Open Source*". Ces méthodes permettent d'économiser en temps et en effort, en particulier lorsque l'utilisateur est plus intéressé par l'obtention des résultats de structure 3D que de traiter des modèles de projection complexes. Ces méthodes d'étalonnage sont classées selon 4 catégories :

- "*Sphere-2D Pattern*"¹ [73], qui utilise le modèle sphérique de la caméra et nécessite plusieurs images d'un modèle de test 2D.
- "*DLT-like*" [77], qui utilise également le modèle sphérique de la caméra. Elle présente une solution analytique basée sur un ensemble de correspondances 3D-2D.
- "*Sphere-Lines*" [81], qui utilise le modèle sphérique de la caméra et nécessite une image omnidirectionnelle unique contenant au minimum trois lignes.
- "*Distortion-2D Pattern*" [71], qui considère les images omnidirectionnelles en tant que images déformées avec des paramètres de distorsion qui doivent être estimés.

2.1. Classification des méthodes d'étalonnage

L'étalonnage des caméras omnidirectionnelles n'est pas simplement différent des caméras classiques, elle est aussi plus difficile. Les caméras ordinaires présentent un problème linéaire car elles projettent une ligne en une ligne alors que les caméras omnidirectionnelles présentent un problème non-linéaire car elles transforment une ligne réelle en un cône sur l'image. Les travaux de recherche qui ont été réalisés sur l'étalonnage de telles caméras essaient donc de trouver une solution approchée linéaire pour résoudre le problème de cette non-linéarité. Plusieurs techniques ont été proposées dans la littérature, adaptées à différents modèles de formation d'images (Tableau 1).

Dans cette partie, nous présentons une taxonomie de classification des méthodes d'étalonnage existantes. On propose cinq catégories en fonction du mode d'étalonnage :

i. Etalonnage basé sur des lignes caractéristiques ("*Line-based calibration*") : L'avantage de cette approche est qu'on a plus besoin d'un modèle de test et principalement que les lignes sont toujours présents dans de nombreuses scènes. Ces approches calculent les paramètres intrinsèques du système catadioptrique à partir de l'image de la conique absolue. Geyer et Daniilidis [80] utilisent seulement trois lignes de chaque image pour étalonner leurs caméras paracatadioptriques. Ying et Hu [90] utilisent la projection des lignes ainsi que la projection de la sphère pour analyser la relation entre les paramètres intrinsèques de la caméra et les contours de la sphère projetée sur l'image. Vasseur et Mouaddib [91] propose un algorithme, valable pour tout système catadioptrique central, qui détecte des lignes omnidirectionnelles à partir d'une scène 3D qui sont ensuite utilisés pour calculer les paramètres intrinsèques.

ii. Etalonnage avec modèle 2D "*2D pattern calibration*": Ces méthodes utilisent un modèle d'étalonnage 2D avec des points caractéristiques (coins, points,...) faciles à extraire des images. Scaramuzza et al. [71] proposent une technique d'étalonnage pour les caméras omnidirectionnelles à point de vue unique. Ils supposent que la fonction de projection sur l'image peut être décrite par une écriture en série de Taylor, dont les coefficients sont estimés par la résolution d'un problème de minimisation. Mei et Rives [73] proposent, comme Scaramuzza, une approche souple pour l'étalonnage des caméras à point de vue unique, mais en se basant sur une fonction de projection théorique équivalente au modèle de la sphère. Deng et al. [74] utilisent le contour de l'ellipse de l'image catadioptrique et le champ de vision ("*FOV*") pour obtenir les paramètres intrinsèques. Ensuite, ils utilisent la relation entre le catadioptrique central et le modèle sténopé de projection pour calculer les paramètres extrinsèques.

iii. Etalonnage basé sur des points 3D ("*3D Point-based calibration*"): Ces méthodes utilisent les positions 3D de points réels observés dans une seule image. Aliaga [92] propose une approche qui ne fonctionne que pour les systèmes paracatadioptriques. Elle permet l'estimation des paramètres intrinsèques et extrinsèques la caméra, sachant que la projection du centre du miroir doit être déterminée manuellement. Puig et al. [93] présentent une approche basée sur la transformation directe linéaire ("*DLT*") valable pour toute caméra catadioptrique centrale à travers une matrice de projection générique. A partir de cette matrice, les paramètres intrinsèques et extrinsèques sont extraits par optimisation non linéaire. Cette approche nécessite une image omnidirectionnelle unique contenant des points répartis dans au moins trois plans différents.

Tableau III. 1 : Classification des méthodes d'étalonnage de système omnidirectionnel

Ref	Modèle de test	Nbre de vues	Type du miroir	Modèle caméra
[69]	—	unique	générique	Modèle hyperbolique central
[70]	3 images polarized	multiple	générique	Modélisation générique
[71]	plusieurs vues de modèle 2D	multiple	générique	Modèle de distorsion
[73]	plusieurs vues de modèle 2D	multiple	générique	Modèle sphérique
[74]	plusieurs vues de modèle 2D	multiple	générique	Modèle sphérique
[80]	3 lignes	unique	parabolique	Modèle sphérique
[81]	3 lignes	unique	générique	Modèle sphérique
[84]	1 ligne + contour miroir	unique	générique	Modèle catadioptrique non-central
[87]	9 correspondances	multiple	générique	Modèle de distorsion
[88]	2 ensemble de translations et rotations	multiple	générique	Modélisation générique
[89]	2 ensemble de rotations	multiple	générique	Modélisation générique
[92]	3D points	multiple	parabolique	Modèle catadioptrique non-central
[93]	Modèle 3D (20 correspondances 3D-2D)	unique	générique	Modèle sphérique
[94]	3 lignes	unique	parabolique	Modèle sphérique

iv. Auto-calibration ("*Self-calibration*") : Ce type de techniques utilise uniquement des points correspondants extraits de plusieurs points de vue, sans avoir besoin de connaître ni leurs

localisations 3D ni leurs emplacements par rapport à la caméra. Micusik et Pajdla [87] proposent une méthode valable pour les objectifs fish-eye et les systèmes catadioptriques. Ils étudient la géométrie de la caméra, ses propriétés, la formation de l'image et aussi la possibilité de réaliser un auto-étalonnage à partir de correspondances de point. Ramalingam et al. [88] utilisent des translations, des rotations et des correspondances images pour calibrer les caméras centrales à partir des contraintes géométriques de projection.

v. Image de polarisation ("*Polarisation Imaging*") : Cette méthode est proposée par Morel et Fofi [70] qui utilisent l'imagerie de polarisation pour une reconstruction précise de la forme de la surface du miroir en se basant sur l'algorithme de Frankot-Chellappa [95].

Dans [96], Puig a réalisé une évaluation des différentes méthodes sur des systèmes omnidirectionnels. L'étude a montré que toutes les approches ont donné de bons résultats, à l'exception de l'approche "*Sphere-Lines*" appliquée sur un système "*fish-eye*". En particulier, pour ces lentilles, le meilleur étalonnage a été réalisé avec l'approche "*DLT-like*". Dans le cas du capteur catadioptrique de paramètres inconnus, l'approche "*Distortion-2DPattern*" a fourni le meilleur résultat. Cette étude a également donné de l'importance à la surface occupée par les modèles d'étalonnage (points, lignes) dans l'image.

2.2. Approches d'étalonnage utilisant un modèle de test 2D

Cette sous-section présente les approches qui nécessitent une mire 2D pour effectuer l'étalonnage à savoir le modèle sphérique et le modèle de distorsion. Ces approches appliquent, en général, une extraction automatique des points caractéristiques (coins de la mire). Mais, dans la plupart du temps, celle-ci ne permet pas d'extraire tous les coins et les points non-détectés doivent être manuellement donnés. Cette phase est la plus fastidieuse dans l'étalonnage puisque nous devons utiliser au minimum une dizaine d'images et que chaque image contient un nombre important de coins. Dans le modèle sphérique, l'utilisateur doit indiquer en supplément le centre de l'image et un minimum de trois points non-radiaux pour estimer la distance focale.

a. Approche du modèle sphérique

Cette approche, proposée par Mei et Rives [73], utilise le modèle de la sphère comme mentionné dans la dernière section. Cette méthode utilise comme entrée plusieurs images d'une même mire contenant autant de points caractéristiques que possible. Le point principal est calculé à partir du centre et la bordure intérieure du miroir. La distance focale est calculée à partir d'au moins trois points colinéaires et non radiaux. Une technique de raffinement subpixellique est appliquée pour extraire le reste des points présents dans la mire. Si la distance focale n'est pas bien estimée alors les points extraits doivent être donnés manuellement. Une fois que tous les paramètres

intrinsèques et extrinsèques sont initialisés, un processus de minimisation non linéaire est effectué.

Cette approche nécessite l'estimation de 17 paramètres à savoir :

- Sept paramètres extrinsèques de (R, T) représentant la relation entre le repère du système de la caméra et le repère monde,
- Un paramètre du miroir ξ (ψ est proportionnel à ξ),
- Quatre paramètres de distorsions : deux tangentielles et deux autres radiales [97],
- Cinq paramètres représentant le modèle générique de projection contenant la distance focale généralisée (δ_x, δ_y), le paramètre de désalignement s et les coordonnées du point central (u_0, v_0).

La matrice intrinsèque K est définie par Barreto and Araujo [81] sous la forme suivante (Eq.1):

$$K = \begin{pmatrix} (\psi - \xi) f_c & s & u_0 \\ 0 & (\psi - \xi) f_c & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{tels que : } \delta = (\psi - \xi) f_c; \quad \psi = \xi + \sqrt{1 - \xi^2} \quad (1)$$

Les paramètres intrinsèques du miroir sont obtenus en appliquant l'approche proposée par Geyer et Daniilidis dans [98]. La distance focale f_c est déterminée par application de la décomposition en valeurs singulières (SVD) [73] sur quatre coins non-radial choisis appartenant à une même ligne de la mire.

b. Approche du modèle de distorsion

Dans cette approche, proposée par Scaramuzza et al. [71], le modèle de projection 2D/3D est différent de celui présenté précédemment. En effet, la fonction de projection est décrite par un polynôme, en se basant sur le développement en série de Taylor. Les coefficients du polynôme sont estimés par la résolution d'un problème de minimisation linéaire de moindres carrés. Cette approche, comme celle présentée précédemment, nécessite plusieurs images acquises pour différentes orientations d'une mire 2D à dimension connue. La précision de la projection dépend du nombre d'images utilisées, le degré du polynôme et la précision de détection des coins de la mire.

Scaramuzza propose une relation entre un pixel $p(u, v)$ et un rayon vecteur 3D correspondant noté $r(X, Y, f(X, Y))$ issu du point de vue situé au centre du miroir jusqu'au point 3D de la scène réelle. Le polynôme d'approximation est donné comme suit (Eq.2) :

$$p \begin{bmatrix} u \\ v \end{bmatrix} \Leftrightarrow \vec{r} \begin{bmatrix} X = u - u_0 \\ Y = v - v_0 \\ Z = f(\sqrt{X^2 + Y^2}) \end{bmatrix} \quad f(X, Y) = a_0 + a_1 \rho + a_2 \rho^2 + \dots + a_n \rho^n \quad (2)$$

ρ représente la distance euclidienne entre les coordonnées du pixel et ceux du centre de l'image. L'étalonnage du capteur catadioptrique est réalisé en estimant les coefficients $\{a_0, a_1, a_2, \dots, a_n\}$ correspondants à la fonction d'approximation $f(X, Y)$. Un algorithme d'optimisation prend comme entrées des correspondances 2D-3D des coins de la mire acquises pour différentes positions. La résolution du problème d'étalonnage est obtenue par minimisation de l'erreur de re-projection en utilisant l'algorithme de Levenberg-Marquardt. L'erreur quadratique est calculée entre les coordonnées des points réels $p(x_p, y_p)$ et les points reconstruits $p'(x_p', y_p')$ définie comme suit (Eq.3) :

$$erreur = \sqrt{(x_p - x_p')^2 + (y_p - y_p')^2} \quad (3)$$

Les paramètres extrinsèques $[R_w, T_w]$ sont obtenus en appliquant la décomposition en valeurs singulières (SVD) [99] sur la matrice correspondante à la relation entre les coordonnées du rayon vecteur $r(X, Y, f(\rho))$ et les coordonnées de son correspondant image $p(u, v)$ comme suit (Eq.4) :

$$\begin{cases} u.(r_{31}X+r_{32}Y+t_3) = f(\rho).(r_{11}X+r_{12}Y+t_1) \\ v.(r_{31}X+r_{32}Y+t_3) = f(\rho).(r_{21}X+r_{22}Y+t_2) \\ u.(r_{21}X+r_{22}Y+t_2) = v.(r_{11}X+r_{12}Y+t_1) \end{cases} \quad R_w = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad \text{et} \quad T_w = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad (4)$$

2.3. Etalonnage sous-l'eau

Toutes les méthodes, mentionnées ci-dessus, ont été développées pour des applications terrestres. Quand il s'agit d'une application sous-marine, plusieurs contraintes liées à la nature de l'environnement s'ajoutent. En effet l'eau de mer introduit des effets variables sur la propagation des signaux dégradant la qualité de l'image. Les principales causes de cette dégradation sont :

- L'intensité de la lumière qui pénètre dans les océans diminue rapidement avec la profondeur. L'utilisation de sources de lumière artificielle devient donc indispensable pour observer la scène avec des caméras CCD.
- La présence de particules en suspension provoque des phénomènes de réfraction et de réflexion des rayons lumineux. La réfraction représente la déviation que subit le rayon lumineux en passant d'un milieu à un autre avec des indices de réfraction différents.
- Les hublots des enceintes utilisées pour protéger les caméras de l'eau et de la pression provoquent généralement des phénomènes de distorsion optique.
- L'indice optique de l'eau est un paramètre variable en fonction des caractéristiques du milieu qui sont la température, la salinité et la pression de l'eau.

La première difficulté apparaît lorsqu'on choisit d'estimer les paramètres de notre système de vision avec une mire de calibration sachant qu'on doit effectuer, à chaque fois, plusieurs acquisitions d'images stéréoscopiques jusqu'à atteindre la bonne estimation. Des travaux antérieurs traitent spécifiquement la calibration sous-marine de caméra [100–102]. Compte tenu des difficultés rencontrées lors d'un étalonnage d'une caméra sous l'eau, les techniques mentionnées ci-dessus traitent le problème d'étalonnage sous-marin de caméra sans l'utilisation d'une mire d'étalonnage. Citons l'exemple de l'auto-étalonnage qui a été adopté dans [102]. Dans ce cas, une séquence d'images acquise par une caméra sous-marine est utilisée pour l'étalonnage tout en exploitant des contraintes spécifiques à la scène. Par conséquent, la mire d'étalonnage n'est plus nécessaire. Dans [100,101], les auteurs ont pensé à trouver une relation analytique entre l'étalonnage terrestre et sous-marin de sorte que le premier soit utilisé pour estimer les paramètres de la caméra lorsque le système est plongé dans un environnement sous-marin. Dans ces deux cas, la précision d'estimation est inférieure à celle obtenue avec une calibration sous-marine réelle où une mire de calibration est immergée et imagée sous l'eau mais elle reste une alternative aux problèmes liés à la marinisation du système. La bonne adéquation entre les paramètres estimés dans l'air et ceux estimés pour des applications sous-l'eau peut être optimisée itérativement en vue d'obtenir un meilleur résultat de reconstruction.

3. Détection et correspondance de points caractéristiques

3.1. Algorithmes de détection de spot(s) laser

L'utilisation de pointeurs laser a l'avantage de faciliter leur détection dans une image, en se basant sur des caractéristiques qui sont en effet définies à partir de la projection du laser utilisé sur l'image. Il est utilisé comme information d'entrée dans plusieurs systèmes à base de caméra et de laser, tels que pour les interfaces interactives [103], [104], [105], les robots guidés par laser [106], [107], [108], les applications technologiques d'assistance [109] et les mesures de distance [110].

Différentes méthodes ont été proposées, dans la littérature, pour l'identification, la détection, le suivi de points caractéristiques issus de la projection d'un pointeur laser sur une surface opaque. On peut classer ces méthodes selon deux catégories : les méthodes qui considèrent que les points d'intérêt sont les points les plus marquants dans toute l'image et les méthodes qui considèrent que les points d'intérêt sont ceux qui ont des caractéristiques prédéfinies auparavant par l'utilisateur. Dans le premier cas, l'utilisation des descripteurs standards tels que SIFT [111], Harris [112] ou SURF [113] est la plus recommandée étant donné que ces descripteurs présentent un bon résultat de détection aussi bien que les méthodes d'apprentissage telles que celles présentées dans [103] et [114].

Dans [115], une description détaillée de ces descripteurs ainsi qu'une étude comparative ont été présentées par Blanco et al., montrant d'une part que le choix d'un détecteur dépend des performances exigées par l'application en termes de rapidité et précision et d'autres part qu'une bonne association de descripteurs peut augmenter le taux de détection. Dans un autre contexte, une méthode systématique qui utilise la différence entre des trames successives d'images et un algorithme d'apprentissage auto-associatif à base de perceptron multicouches (AAMLN) est proposée pour détecter des spots lasers [103]. Afin d'améliorer les performances de détection, l'algorithme AAMLN est entraîné par un nouvel algorithme d'apprentissage. L'augmentation de la sensibilité de détection lors la mise en correspondance entre entrée-sortie de l'algorithme AAMLN permet d'indiquer avec succès un spot laser même à une petite variation de la caractéristique d'entrée. Néanmoins, ces méthodes restent les plus coûteuses en temps d'exécution et ne sont pas adaptées aux images à mauvaises qualités telles que les images sous-marines. Le taux de détection des approches d'apprentissage dépend en effet de la base d'apprentissage utilisée.

Dans [116], la technique proposée est une association de plusieurs spécifications du laser telles que son intensité par rapport aux autres objets de la scène, sa taille, la variation progressive de sa couleur à partir de son centre jusqu'au contour ...etc. Dans [117], la détection est effectuée en associant trois algorithmes : le premier est pour la détection de mouvement afin d'identifier dans l'image la région qui a subi un changement, le deuxième est la reconnaissance de forme afin de détecter la position d'un laser dans l'image et le dernier est une comparaison entre l'histogramme du laser détecté et des histogrammes prédéfinis afin d'extraire celui qui lui est le plus similaire. Dans ce même but, le processus de segmentation apparaît dans la plupart des ouvrages tels que [118–122]. En effet, la segmentation permet de mettre en évidence la projection d'un spot laser par rapport aux autres éléments de l'environnement en se basant sur sa valeur d'intensité. En comparant tous les pixels par rapport à un seuil donné, l'image couleur est binarisée : les pixels blancs représentent la projection du laser et les pixels noirs l'arrière-plan. L'image couleur peut être acquise selon différents espaces couleurs tels que l'espace RVB (rouge/vert/bleu)[123], [124], HSI (teinte/saturation/intensité) [125], HSV (teinte/saturation/valeur) [126] ou autres (Lab, XYZ...etc.). Dans [124], l'algorithme de détection comprend une conversion de l'image RVB en une image représentant les intensités des pixels, sur laquelle un filtre médian est appliqué pour éliminer les bruits et les objets non-intéressants. Ayant appliqué une détection de contour, les spots lasers sont identifiés en se référant sur la surface des cercles extraits de l'image.

Dans [106], l'utilisation du pointeur laser permet d'attirer l'attention du robot vers un objet donné qui sera enregistré dans la mémoire du robot pour un apprentissage ultérieur (Fig.1). Une

méthode de segmentation basée sur la couleur verte du laser puis l'identification du spot le plus brillant est appliquée. Tandis que dans [107], l'application a pour but de commander le mouvement du robot en suivant les objets pointés par le laser. L'algorithme détecte les traces du faisceau laser, calcule le flux optique correspondant, reconnaît les objets et envoie la commande au robot.



Figure II. 1: Apprentissage et commande du robot [37]

A cause des variations de luminosité, d'éclairage ou la mauvaise qualité de l'image, le choix d'un seuil de segmentation peut être contraignant vu qu'il peut varier d'une image à une autre. Le choix d'une seule valeur de seuillage pour tout type d'image peut engendrer de fausses détections. Dans [126], l'approche proposée permet de différencier entre les bonnes et les mauvaises détections en analysant la proportionnalité entre le diamètre de la surface de projection et le nombre de pixels qu'elle contient. Dans [127], un système à base de règles floues ajusté par un algorithme génétique est proposé afin d'améliorer le taux de détection de l'algorithme en définissant le spot laser par un ensemble de caractéristiques telles que sa couleur, son énergie, la forme de sa surface de projection ...etc.

3.2. Algorithmes de correspondances

Le problème de correspondance stéréo présente un grand intérêt dans le domaine de la vision, en particulier pour les applications de reconstruction 3D. Il concerne la mise en correspondance de points, ou toute autre primitive, entre une paire d'images de la même scène. Il est basé, en général, sur l'hypothèse que deux correspondants aient le même voisinage. Les méthodes de correspondances peuvent être classées selon deux catégories : les méthodes globales et les méthodes locales. Les premières sont basées sur la définition d'une fonction objective de correspondance (par exemple, distance de Hamming) qui permet de minimiser l'erreur globale correspondante aux pixels stéréo extraits tout en prenant en compte la totalité de l'image.

L'inconvénient principal de ces méthodes est qu'elle exige plusieurs itérations de raffinement et par conséquent un long temps d'exécution pour obtenir un bon résultat. D'où l'intérêt des méthodes locales, qui permettent de mesurer la similarité entre les voisinages respectifs des pixels stéréo. Plusieurs algorithmes, existant dans la littérature, définissent une fonction qui permet de calculer la similarité telles que la différence absolue d'intensité (AD :

Absolute Intensity Difference) [128], la différence quadratique d'intensité (SD : Squared Intensity Difference) [129] et la corrélation croisée normalisée (NCC : Normalized Cross-Correlation) [130]. Dans [131], Nadir et al., présentent une comparaison entre les différents algorithmes pour une application de détection et de reconnaissance de visage humain. Ils présentent aussi une nouvelle approche, qui est une technique optimisée du calcul de la somme des différences absolues (OSAD : Optimization Of Sum Absolute Difference), permettant de s'adapter aux variations de l'image (rotation, illumination, ...etc.) et d'améliorer le taux de correspondance par rapport aux autres approches.

Dans [132], Daniel et al., ont présenté une taxonomie des algorithmes existants en comparant les performances en termes de temps de calcul et en mettant en évidence les très bons résultats obtenus par les méthodes fondées sur les régions. Ces dernières permettent, en effet, de diviser les images stéréo en plusieurs régions selon des critères bien définis et de faciliter donc la correspondance ce qui permet de réduire le taux des fausses correspondances.

En général, il n'existe pas de méthode générique qui résout tous les problèmes de mise en correspondance mais des propositions de solutions adaptées peuvent être proposées. Celles-ci peuvent accompagner la technique de correspondance en lui apportant plus d'information sur les points stéréo détectés. On peut citer, par exemple, le critère de proximité qui permet de poser l'hypothèse que le point le plus proche possède plus de chance pour être choisi comme éventuel correspondant. Le déplacement d'un point entre les deux images est la distance euclidienne calculée entre les deux pixels stéréo. Dans [133], une contrainte qui nécessite la connaissance des paramètres du capteur stéréoscopique est proposée. Elle permet, en effet, de réduire l'espace de recherche des correspondants d'une image à une droite appelée "droite épipolaire". Cette contrainte est liée à la géométrie du système mais elle présente de bons résultats de correspondances en dépit du temps de calcul mis pour le calcul de la droite épipolaire. Cette contrainte peut être aussi associée à la contrainte d'ordre ou d'unicité. La première est utilisée lorsqu'on cherche plusieurs correspondances sur une même droite épipolaire vu que l'ordre des pixels doit être le même pour l'image droite et gauche. La deuxième est utilisée pour insister sur le fait qu'un seul pixel ne peut être apparié qu'un seul correspondant de l'autre image.

4. Reconstruction 3D

Bien qu'il existe un grand nombre de travaux, dans la littérature, sur d'étalonnage de caméra omnidirectionnelle [71,83,134,135], la localisation [40,51], l'estimation de structure à partir du mouvement [87,136] ou les applications SLAM [41,42,137], moins nombreux sont les recherches qui ont été menées à la reconstruction 3D avec des caméras catadioptriques.

Dans les applications sous-marines, la nature de l'environnement aquatique présente des effets de réfraction de la lumière qui le traverse. Plusieurs travaux ont été menés dans ce même contexte citons l'exemple de [68,138–140], montrant la relation entre la bonne estimation des conditions de navigation et la précision de modélisation 3D de la structure de l'environnement. En outre, les algorithmes de restitution de scène 3D doivent prendre en considération les effets de réfraction soit en simulant une estimation des paramètres du système de vision à partir d'images sous-marine [141], soit en insérant ces effets lors du suivi des rayons traversant les différents milieux à partir du point 3D jusqu'à sa projection 2D sur le plan image [138,139]. L'obtention de l'information de profondeur à partir des données 2D peut s'effectuer par diverses méthodes. Dans les sous-sections suivantes, on va citer quelques exemples existants dans la littérature.

4.1. Géométrie épipolaire

La géométrie épipolaire décrit la relation entre les points 3D et les positions des points correspondants dans une paire d'images. Elle peut être établie à partir de quelques correspondances d'images et est utilisée pour: simplifier la recherche de plus de correspondances, calculer le déplacement entre les caméras, et reconstruire la scène.

Dans [69], Svoboda et Pajdla ont étudié la géométrie épipolaire des systèmes catadioptriques centraux. Ils montrent que les lignes épipolaires correspondent à des coniques dans l'image omnidirectionnelle qui se réduisent à des lignes radiales dans le cas des caméras catadioptriques alignés verticalement. Plusieurs méthodes prennent l'avantage de cette configuration soit par une association de deux caméras catadioptriques l'une au-dessus de l'autre (Fig.2) ou en utilisant un double miroir avec seulement une seule caméra (Fig.3) permettant un appariement stéréo sur une même image.

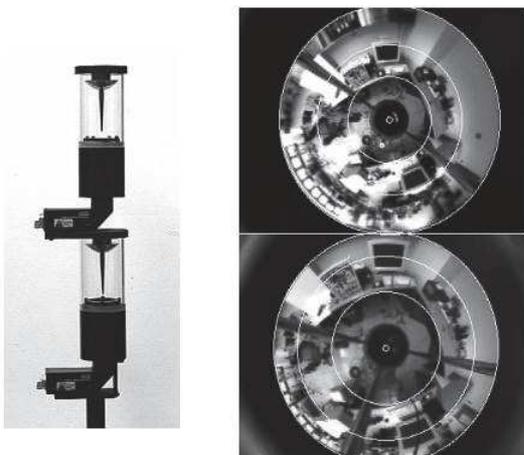


Figure II. 2: Système stéréo à deux caméras [142]

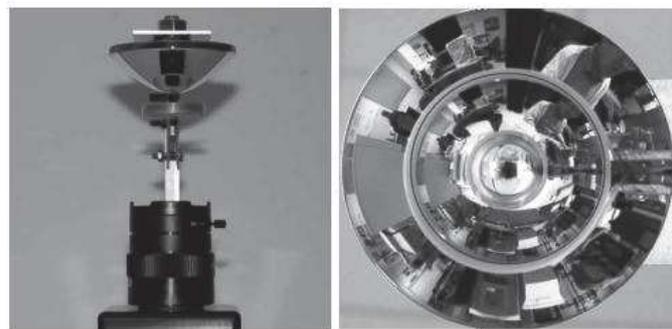
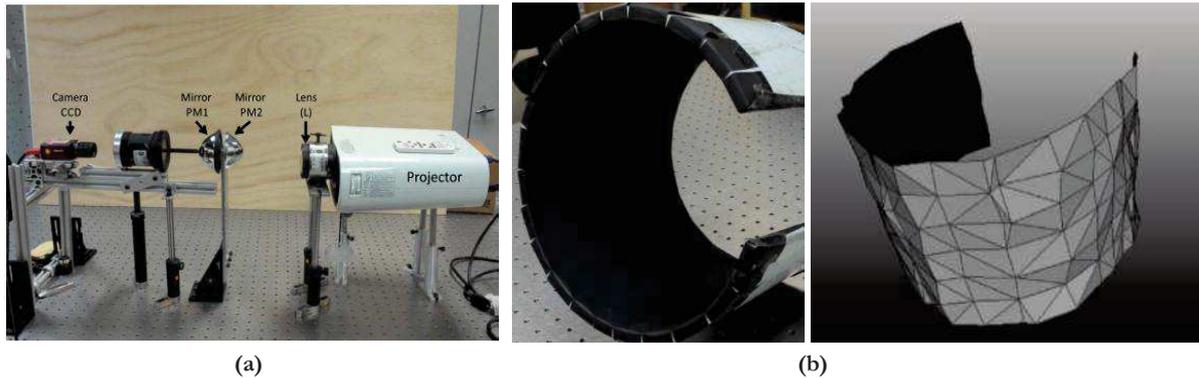


Figure II. 3: Système stéréo à une seule caméra [143]

Le système panoramique, utilisée dans [144], est composé d'une caméra catadioptrique et un projecteur de lumière (Fig.4.a). La méthode de reconstruction 3D utilisée est basée sur la géométrie épipolaire permettant de définir des relations entre les points correspondants dans les images stéréo d'une même scène. Les coordonnées 3D d'un point du monde réel sont calculées en utilisant la distance entre deux points correspondants qui dépend, en effet des paramètres intrinsèques et extrinsèques du système de vision. Une représentation 3D est réalisée en utilisant le logiciel Metrolog XG® (Fig.4.b, à droite).



(a) (b)
Figure II. 4: Reconstruction de formes 3D par géométrie épipolaire

4.2. Triangulation

La position 3D d'un point monde peut être calculée en utilisant la triangulation classique comme on fait traditionnellement dans un système de stéréovision (par exemple : [145], [146], [37] et [147]). La triangulation est une technique qui permet de déterminer la profondeur d'un point de l'espace, à partir de ses projections pixels dans les deux images stéréoscopiques. Dans sa forme la plus simple, on associe un point 3D à deux uniques pixels appariés, chacun est situé sur une image de la paire des images stéréo.

Dans [147], une triangulation est appliquée sur les paires de points extraits des images stéréo (Fig.5.a) afin d'obtenir une estimation de la structure 3D, en utilisant la géométrie du système stéréo et les paramètres internes de la caméra. Afin d'améliorer la structure 3D des objets, une rectification des images stéréo est réalisée et un algorithme d'estimation dense des profondeurs est appliqué (Fig.5.b). La carte de profondeur est transformée en un maillage surfacique à l'aide de l'algorithme de triangulation de Delaunay (Fig.5.c). La représentation 3D est alors visualisée dans un aspect plus réaliste en fournissant un modèle 3D avec mappage de texture (Fig.5.d). Le processus de reconstruction est réalisé hors ligne à partir d'un ensemble d'images stéréo sous-marine vu qu'elle nécessite un long temps de calcul.

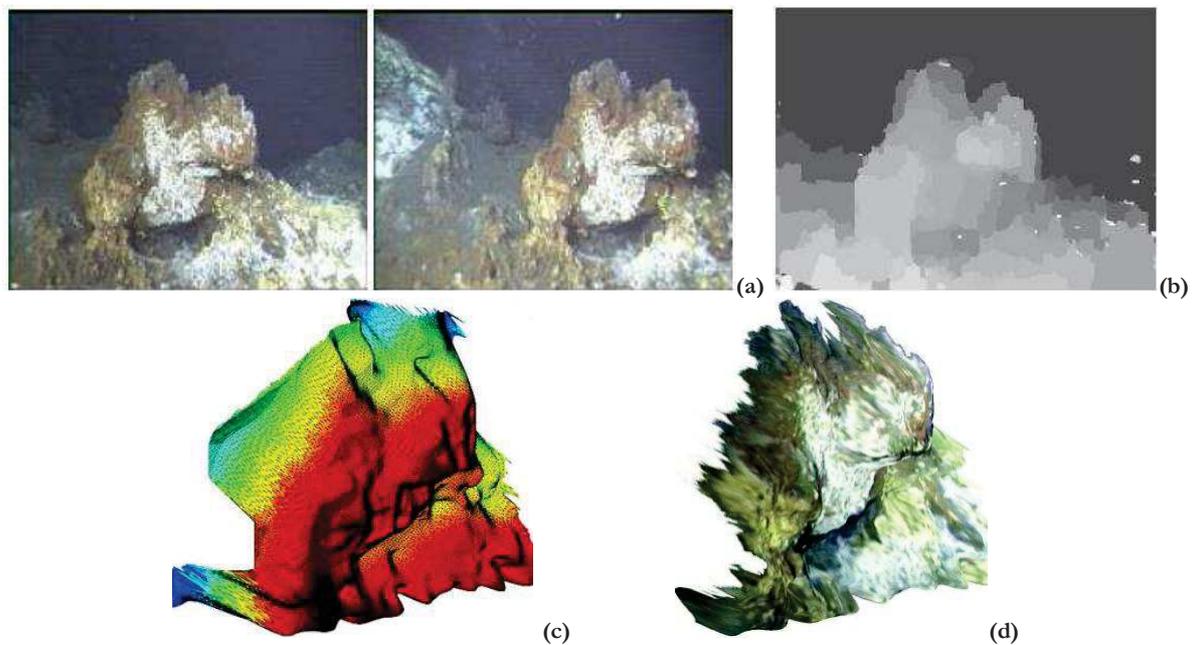


Figure II. 5: Exemple de reconstruction 3D sous-marine [85]

Dans [37], un système de navigation est développé pour créer des modèles 3D d'objets réels photographiés dans un environnement sous-marin. L'algorithme est une succession de plusieurs processus de traitement. Il utilise les informations stéréoscopiques (Fig.6.a) pour créer une carte de disparité et générer les positions 3D des points caractéristiques 2D (Fig.6.b). Entre deux acquisitions successives, un algorithme basé sur l'approche RANSAC [38] est appliqué pour estimer le mouvement de la caméra. La rotation et translation du système sont optimisées en utilisant l'algorithme de minimisation de Levenberg-Marquardt [148]. Le modèle de nuages de points résultant est ensuite converti en une représentation volumétrique et un maillage polygonal texturé est finalement extrait (Fig.6.c).

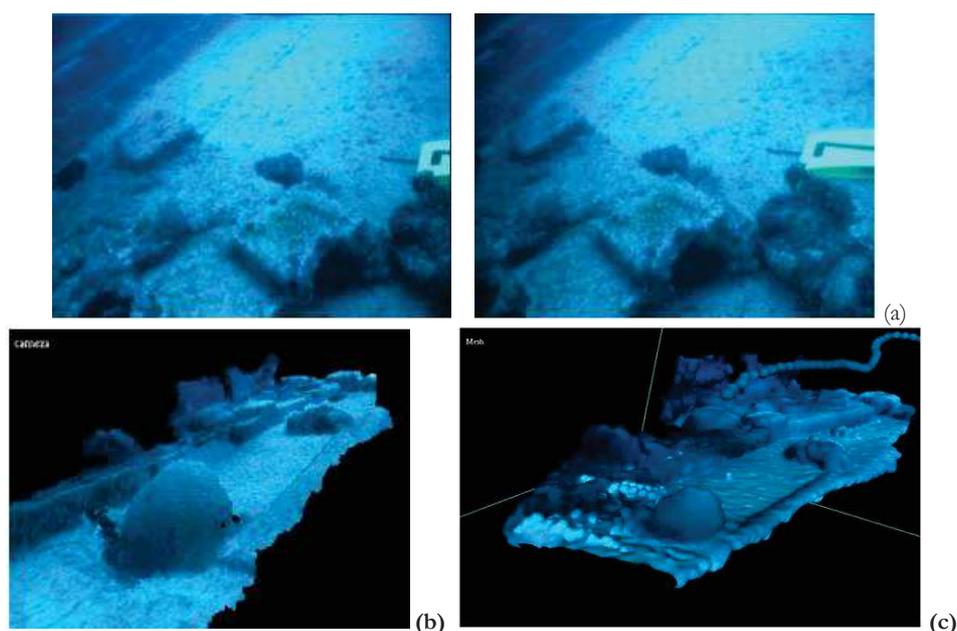


Figure II. 6: Reconstruction 3D sous-marine par triangulation [80]

4.3. Géométrie du tracé de rayon

Plusieurs études sur la formation d'image sous-marine sont développés dans [149–151] afin d'intégrer les propriétés de la réfraction de la lumière dans l'eau. La reconstruction tridimensionnelle de forme/structure peut être réalisée en utilisant des modèles de formation d'images et des techniques numériques spécifiques. Le milieu sous-marin, n'étant pas un milieu homogène, fait qu'un rayon lumineux ne suit pas une trajectoire linéaire lors de son passage d'un milieu à un autre (Fig.7.a). En effet, l'indice de réfraction dépend de la célérité du rayon lumineux qui varie d'un milieu à un autre. Ce dernier subit, à chaque surface séparatrice, une réfraction qui dépend de plusieurs facteurs tels que la température, la salinité et la pression de l'eau. Outre ces caractéristiques, l'indice de réfraction varie aussi dans de sensibles proportions avec la couleur: longueur d'onde du rayonnement considéré (Fig.7.b).

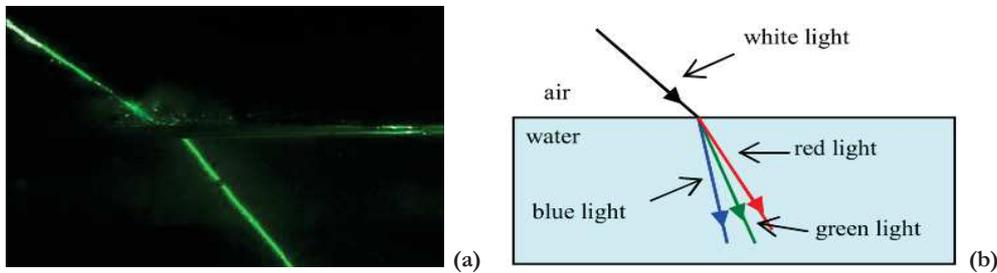


Figure II. 7: Principe de réfraction de rayon laser sous l'eau

Lorsque l'étalonnage du système de vision est réalisé dans l'air, les effets de réfraction sont intégrés dans l'algorithme de reconstruction 3D. Dans ce contexte, l'algorithme du tracé de rayon est le plus recommandés en littérature. En effet, ce processus permet la reconstruction précise de la scène par la modélisation et le suivi des réfractés de la lumière. Il est bien adapté aux réfractés complexes à travers plusieurs interfaces et il exige une connaissance complète du processus de projection 3D. Dans [138], un système de caméras catadioptriques plongé dans un environnement aquatique a été utilisé pour l'acquisition d'images omnidirectionnelles. Les coordonnées 3D du rayon vecteur traversant l'air et réfracté à la surface du miroir est déterminé à partir des paramètres caractéristiques de la caméra (calibrée dans l'air), des paramètres intrinsèques du miroir et les coordonnées images (u, v) en pixels. Les équations donnant la relation entre les rayons vecteurs réfractés sur les différentes surfaces (air/verre et verre/eau) peuvent être déterminées par la méthode du tracé de rayons ("Ray-tracing").

Les algorithmes de restitution 3D, présentés dans [152–155], sont inspirés des relations de géométrie du tracé de rayon. Ils consistent, en général, à utiliser la géométrie du système de vision pour passer des informations 2D à celles en 3D. Ceci est réalisé en suivant la trajectoire des vecteurs porteurs de faisceaux lumineux et utilisant des lois de Snell pour déterminer les angles de

réfraction correspondants à chaque changement de milieu. Dans [156], un nouvel algorithme d'étalonnage dans l'eau ainsi que sous l'eau a été présenté. L'étude des effets de distorsions est basée sur l'approche du tracé de rayon ainsi que les expressions de réfraction du rayon lumineux en traversant les différents milieux (Fig.8). Les angles de réfraction sont déterminés en appliquant la loi de Snell (5) :

$$n_1 \sin(\theta_1) = n_2 \cdot \sin(\theta_2) \quad (5)$$

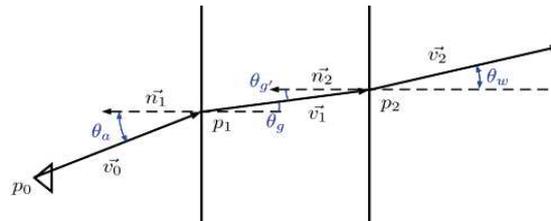


Figure II. 8: Schéma de réfraction d'un rayon lumineux passant par différents milieux

5. Conclusion

Dans ce chapitre, nous avons présenté quelques techniques exposées en littérature qui sont en relation avec les processus d'étalonnage des systèmes catadioptriques en citant les avantages de l'approche qui sera utilisée dans notre algorithme. Aussi, nous avons présenté les méthodes les plus utilisées pour la détection ainsi que la correspondance de points d'intérêts. En ce qui concerne les méthodes de reconstruction 3D dans des environnements plus ou moins complexes, nous nous sommes intéressés aux travaux existants prenant en considération les phénomènes de réfraction de rayons lumineux en traversant des milieux différents.

Partie II.

Etude Expérimentale : Etude, Validation et Implémentation du système proposé

Chapitre III.

Etude et Validation du Système Proposé

1. Introduction

La principale fonctionnalité attribuée à notre robot est la génération d'un modèle métrique 3D approximatif de l'environnement sous-marin. Dans ce travail de thèse, on est intéressé par le centrage du robot dans le conduit évitant ainsi sa collision avec ses parois en se basant seulement sur des données de perception. En effet, tout au long de sa navigation sous l'eau, l'engin sera entraîné de manière automatique dans le centre du conduit en se basant sur les données métriques générées par notre système. Lorsque le robot est positionné à un instant "t", les deux caméras omnidirectionnelles capturent simultanément un flux vidéo de l'environnement visité. Ces données seront envoyées par l'intermédiaire du PC au système de traitement. Ce dernier échantillonne les deux vidéos et construit pour chaque paire d'images la forme métrique du conduit à cette position. Ces coupes vont être assemblées au cours du temps pour former un modèle complet de l'environnement.

Dans ce chapitre, on présente une description détaillée des différentes approches utilisées pour le prétraitement des images (conversion d'espace couleur, segmentation et filtrage), l'analyse des informations en vue d'extraire les positions des projections laser et la construction de la structure 3D de l'environnement de test. Une approche d'étalonnage de notre système stéréoscopique dans un environnement aérien et aquatique est proposée avec des expérimentations réelles. Pour la validation du système complet, plusieurs tests ont été menés pour étudier les performances des algorithmes en termes de précision en comparant entre les mesures réelles et celles estimées.

2. Description des étapes de traitement

Une décomposition des différents traitements numériques aboutissant à la restitution de la forme 3D de l'environnement sous-marin à partir des images stéréoscopiques est illustrée dans la figure 1. Ces étapes sont exécutées dans l'ordre suivant:

- *Estimation de la Géométrie du système:* Le modèle du système de vision est une représentation géométrique des caractéristiques physiques caractérisant le comportement du système de vision. Ce modèle permet d'exprimer les points 3D à travers les données 2D extraites du plan image de la caméra. Ils sont exprimés par un ensemble de paramètres estimés par étalonnage effectué hors ligne dans l'environnement de test.
- *Acquisition des images:* Une séquence de paires d'images stéréoscopiques simultanées est utilisée comme données d'entrée à une fréquence intrinsèque aux caméras utilisées.
- *Extraction de primitives:* Dans cette étape, les éléments les plus marquants de l'image sont extraits. Dans le cas de notre application, il s'agit de l'extraction des projections des pointeurs lasers. Ce

processus est effectué suite à un pré-traitement de l'image RVB permettant de mettre en œuvre les projections lasers. Les positions extraites seront utilisés dans l'étape d'appariement.

- *Mise en correspondance*: Le problème de mise en correspondance consiste à trouver la projection d'un point monde dans les deux images stéréo. Si la géométrie de la caméra est connue, le problème se réduit à un problème plus simple, où les courbes épipolaires deviennent des lignes et qu'un correspondant à un point de l'image droite se trouve le long de la même ligne épipolaire de balayage dans l'autre image.

- *Transformation 2D/3D*: Compte tenu des relations géométriques et des paramètres d'étalonnage caractérisant le système de vision, si deux projections concrètes reflétées dans chaque caméra sont liés au même point 3D, les coordonnées de cet objet dans l'espace peuvent être calculées.

- *Calcul de profondeur*: La profondeur représente la distance entre l'axe optique du système de vision et la surface pointée par le laser. Les valeurs de distances déterminées donnent la position du robot par rapport aux parois du conduit.

- *Mise à jour de la forme 3D*: A chaque acquisition d'image stéréo, le processus de traitement génère une forme 3D correspondante à la structure de l'environnement. Cette forme sera par la suite fusionner avec le modèle 3D restitué en utilisant les acquisitions précédentes.

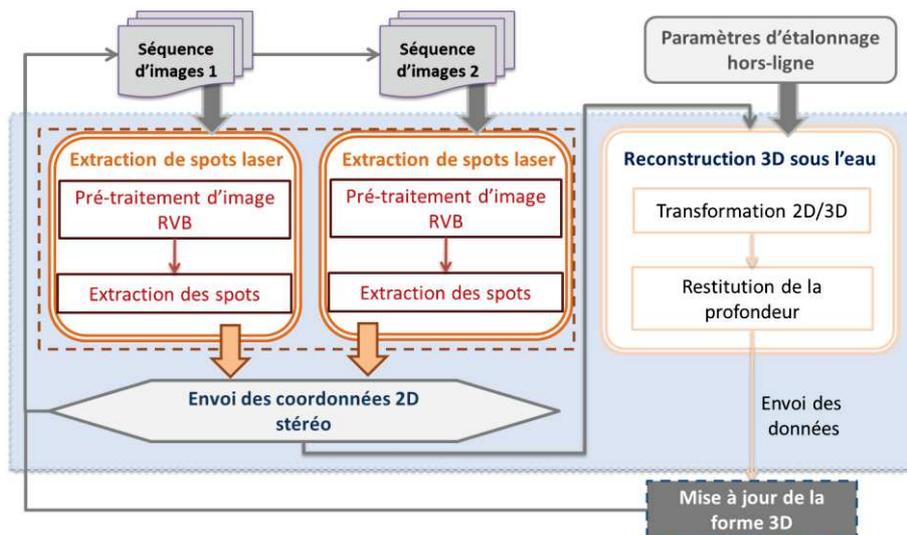


Figure III. 1: Etapes de traitements d'images

3. Etalonnage des systèmes stéréoscopiques catadioptriques

La reconstruction 3D sous-marine est une discipline nettement différente de la celle dans l'air, bien qu'elle se pratique généralement avec le même matériel, sous réserve d'une adaptation. En effet, travailler dans l'eau, milieu dont toutes les caractéristiques physiques sont différentes de celles de l'air nécessite la connaissance des phénomènes de réfraction en matière d'optique, puisque cette loi concerne le trajet entre le sujet et les capteurs en vidéo des rayons lumineux issus du sujet. Dans cette partie, une description des constituants du système omnidirectionnel

sera présentée. Afin d'assurer une bonne reconstruction 3D, ce système sera étalonner en vue d'extraire les paramètres physiques le caractérisant et pouvoir créer un modèle géométrique qui lui est correspondant.

3.1. Description du système stéréoscopique

Grâce à leurs propriétés optiques, les dispositifs catadioptriques avec PVU ont été choisis pour construire notre système omnidirectionnel. La contrainte de la projection centrale est respectée par les miroirs hyperboliques conjointement utilisés avec des caméras perspectives. La solution matérielle proposée pour répondre à notre application est un banc stéréoscopique catadioptrique constitué de deux caméras perspectives pointant chacune sur un miroir hyperbolique. On a opté pour une configuration verticale des deux systèmes catadioptriques vu qu'ils offrent un champ de vue commun important dans les deux images. Le choix d'un capteur catadioptrique à miroir de révolution hyperboloïde a été retenu parce qu'il s'agit d'un capteur à projection centrale (plus simple à modéliser) et que ce capteur est moins coûteux qu'un objectif fish-eye et moins encombrant qu'un capteur à miroir parabolicoïde utilisant une lentille télécentrique.

Un autre point important, qui distingue la vision sous-marine, est l'absorption des couleurs. Le rouge et le jaune sont absorbés très rapidement et il est nécessaire de prévoir un système d'éclairage adéquat si l'on souhaite percevoir les véritables couleurs. Aussi, les images catadioptriques sont caractérisées par une distorsion géométrique et une basse résolution ce qui peut rendre la mise en correspondance encore plus difficile, en particulier pour les images sous-marines. Pour pallier donc à ces problèmes, nous avons proposé de munir le système par un ensemble de pointeurs laser vert-bleus, moins absorbés sous l'eau et permettant d'éclairer des surfaces du conduit. Pour s'assurer du bon choix des éléments du système 3D, nous avons simulé les deux capteurs catadioptriques ainsi que la ceinture laser avec un logiciel de modélisation 3D appelé "POV-Ray". Ce dernier permet de visualiser les images résultantes en dépend du modèle géométrique choisi. Le miroir est défini sous une forme quadrique de paramètres proportionnels aux paramètres intrinsèques du miroir (Fig.2). Afin de garantir la contrainte de point de vue unique, la caméra est placée à une distance égale à l'excentricité notée 'd' dont la valeur est proportionnelle aux paramètres du miroir ("a" et "b"). Les trois systèmes sont montés de telle sorte que leurs axes optiques soient confondus (Fig.3.a). Le système est émergé à l'intérieur dans un environnement confiné à texture presque homogène. Les pointeurs laser doivent être placés devant le système stéréo afin de pouvoir visualiser leur projection sur les deux miroirs sur les rétines des caméras. Les images stéréo permettent de visualiser la projection des lasers qui pointent sur la surface du conduit (Fig.3.b). La configuration verticale des deux capteurs permet de faciliter l'identification de points caractéristiques et leurs mises en correspondance.

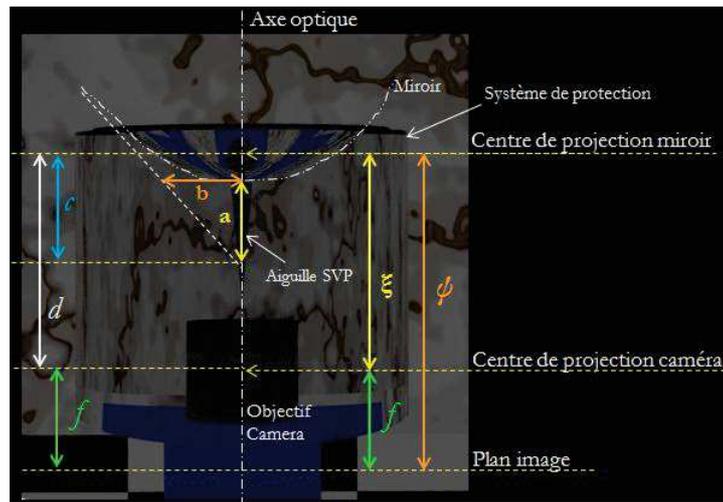


Figure III. 2: Modélisation paramétrique du système catadioptrique

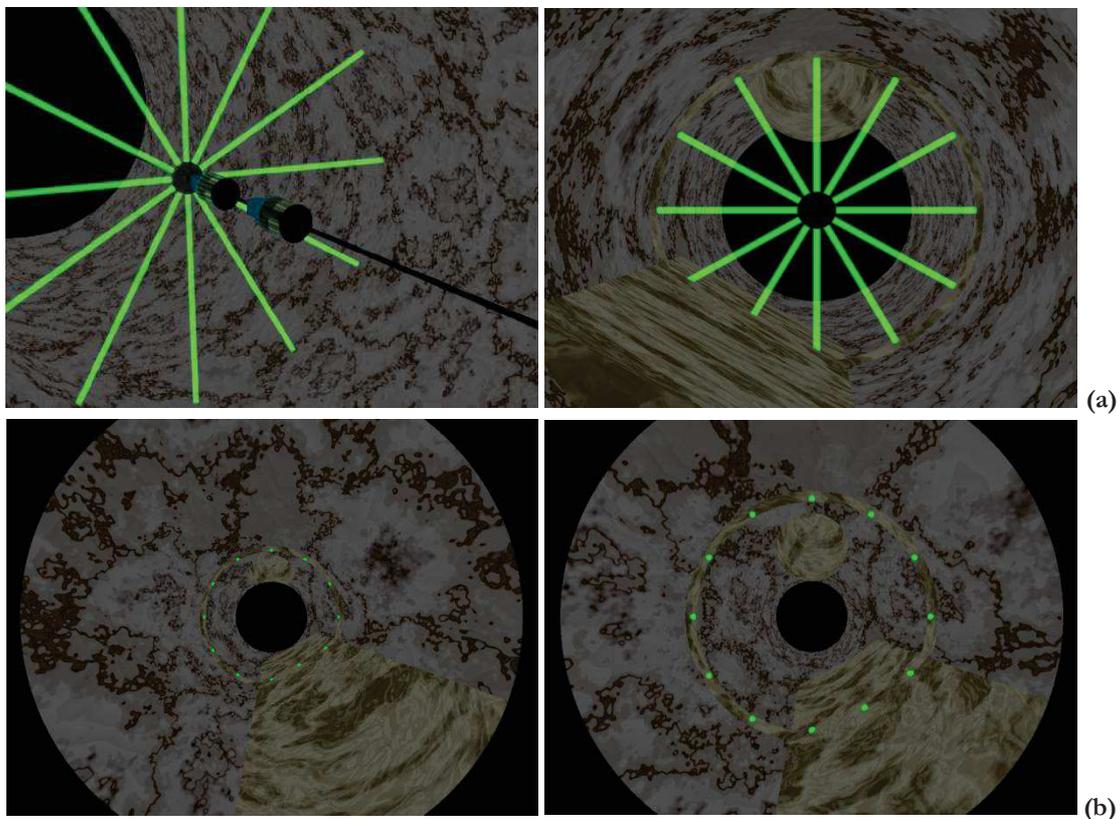


Figure III. 3: Modélisation du système stéréo actif (a) Système 3D (b) Image stéréo acquises

Ayant été satisfait par les résultats de simulation obtenus, on a donc conçu notre système stéréoscopique. Les deux capteurs sont fixés verticalement sur une potence en aluminium avec une distance inter-caméras appelée "*Baseline*". Dans la configuration proposée, les axes optiques des capteurs, également les axes de révolution des miroirs, sont supposés parfaitement confondus (configuration coaxiale verticale). Cette configuration fait que la même géométrie utilisée pour projeter un point 3D de l'espace dans le plan image de la caméra1 est aussi exploitable pour la caméra2. Chaque capteur est muni d'une aiguille PVU qui permet d'éviter que les réflexions des rayons lumineux sur le système de protection en verre ne reviennent perturber l'image (Fig.4.a).

Le capteur CCD est de dimension 1/3" pouce (largeur : 6.4mm, hauteur : 4.8mm). La caméra est de type Flea2 de la famille Point Grey. Elle comporte un connecteur 9-broches de la norme IEEE-1394b qui est utilisée pour la transmission de données, le contrôle de la caméra et la mise sous tension de l'appareil (Fig.4.b). Une carte "PCI" est fournie pour assurer la connectivité caméra – PC (Fig.4.c).

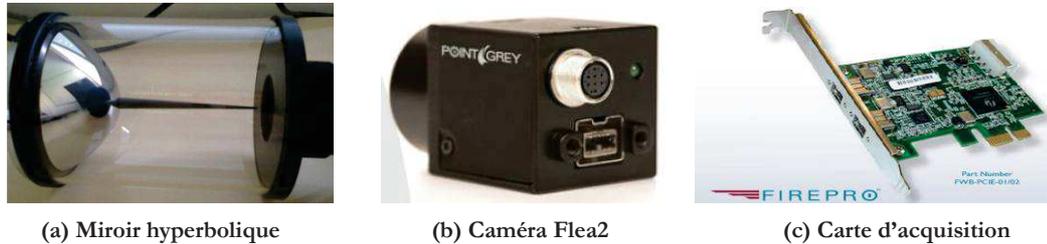


Figure III. 4: Différents périphériques du capteur catadioptrique

Les caméras sont fixées sur la potence et les miroirs sont montés sur des socles fixes suivant les trois axes à l'aide de vis micrométriques. Cette mécanique permet de positionner le miroir à une distance prédéterminée assurant que le système (c'est-à-dire l'ensemble miroir-caméra) respecte la contrainte de point de vue unique (PVU). Lorsque ce positionnement est respecté, que la focale de l'objectif est choisie en adéquation avec le système et que les axes optiques de la caméra et de révolution du miroir sont confondus, la projection du miroir dans l'image omnidirectionnelle est alors un disque qui s'inscrit dans la totalité de la hauteur du CCD et dont le centre du disque est le point principal de l'image (Fig. 5). Le système de vision est connecté à un PC à travers la liaison FireWire IEEE1394 pour lui envoyer les données issues des caméras. La configuration des caméras permet d'offrir un champ de vue commun important entre les deux images permettant une meilleure compréhension et à la visualisation d'une plus large partie de l'environnement observé.

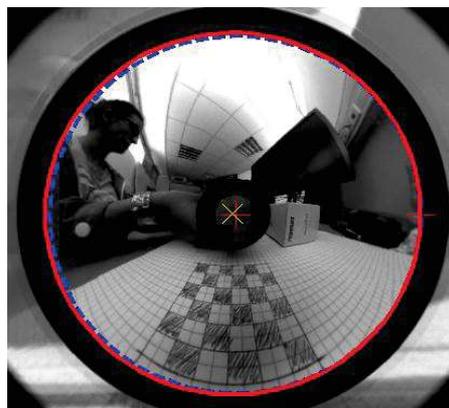


Figure III. 5: Vérification de la satisfaction de la contrainte PVU

Les pointeurs lasers ont été montés sur un tore 3D, positionnés selon une forme cylindrique et distants d'un angle d'inclinaison égal à 30° l'un de l'autre (Fig.6). Les lasers sont des pointeurs standards de longueurs d'ondes égale à 532nm est une puissance de 10mW. Afin d'assurer

l'étanchéité des caméras et des lasers lors des expérimentations sous l'eau, nous avons placé le tout dans un tube en plexi transparent. Le choix de la forme cylindrique a été sélectionné car une forme plate présente une réfraction de lumière plus importante au niveau de la surface séparant les deux milieux [1].

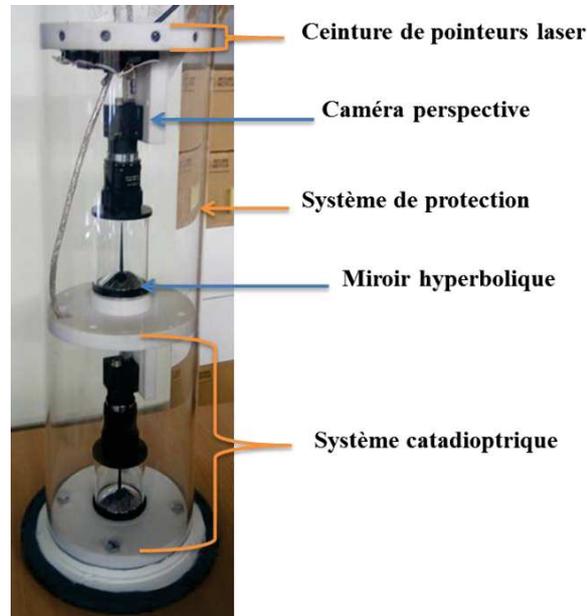


Figure III. 6: Système stéréoscopique à pointeurs laser

3.2. Démarche d'étalonnage du système

La reconstruction 3D à partir de plusieurs images est un thème de recherche de plus en plus étudié dans les dernières années. Avec la connaissance des paramètres de la caméra, on peut déduire les informations 3D à partir d'images acquises sous différents angles de vue. La géométrie d'imagerie d'un système stéréoscopique est principalement définie par la position relative de la lentille et le capteur d'image dans chaque caméra, les propriétés de distorsion optique de chaque lentille, et la position des caméras l'une par rapport à l'autre. En utilisant l'imagerie de la géométrie du système, la position d'un point 3D peut être déterminée par triangulation après la localisation de la projection de ce point dans chacune des images stéréo capturées par deux caméras. Par conséquent, l'analyse et la compréhension de la géométrie d'imagerie d'un système de vision stéréoscopique est un prérequis pour la reconstruction stéréoscopique, qui peut dériver du processus de formation d'image dans chaque caméra, appelée étalonnage ou "*Calibration*".

L'étalonnage permet de trouver la relation entre les dimensions sur l'image et les dimensions réelles. Il peut être réalisé en ayant des informations soit sur la géométrie de la scène (lignes parallèles, forme et dimension réelle de quelques objets connus, ...etc.), soit sur le mouvement de la caméra ou tout simplement en utilisant des mires d'étalonnage. Plusieurs "Toolboxes" (logiciels) ont été proposés en ligne et peuvent donc être testés à savoir : HySCaS :

logiciel de calibration stéréo hybrides [157], Toolbox de Bouguet : calibration de caméra mono ou stéréo perspective et fisheye [158], Toolbox de Mei : calibration de caméra mono fisheye et omnidirectionnelle [159] et le Toolbox de Scaramuzza: calibration de caméra mono omnidirectionnelle centrale [160]. Parmi les approches d'étalonnage qu'on a testé, l'algorithme proposé par Scaramuzza nous est le plus adéquat vu qu'il offre une représentation moins complexe du problème de calibration contrairement à d'autres Toolboxes à savoir :

- il offre une relation polynomiale plus simple permettant d'estimer les coordonnées 3D d'un point monde à partir de ses projections 2D,
- c'est le seul qui offre une extraction automatique des coins de la mire,
- l'algorithme d'estimation ne nécessite pas une connaissance à priori sur la forme du miroir. Il ne nécessite pas l'étalonnage du système catadioptrique séparément: le miroir et la caméra sont considérés comme un système unique et compact qui englobe à la fois les paramètres intrinsèques de la caméra ainsi que ceux du miroir.
- la détection du centre de l'image est effectuée automatiquement. Il ne nécessite pas la visibilité de la limite extérieure circulaire du miroir.

En pratique, la procédure d'étalonnage d'un système stéréo est identique à la procédure de calibrage pour une caméra unique; une mire de test est déplacée dans des directions différentes dans le champ commun de vision des deux caméras afin de couvrir l'intégralité du miroir associé à chaque caméra. Une série d'images de la mire de test, de dimension connue, sont capturées pour chaque caméra donnant d'une part les coordonnées 3D d'un ensemble de points lui appartenant et d'autre part leurs projections 2D correspondantes. La correspondance des points entre l'espace objet et l'espace image constitue l'information d'entrée pour l'algorithme d'estimation des paramètres de calibration standard. Les paramètres de la caméra qui définissent le modèle de la caméra peuvent être classés en deux catégories; les paramètres intrinsèques et les paramètres extrinsèques. Les paramètres intrinsèques décrivent la géométrie interne du dispositif de formation d'image et les propriétés de distorsion optique de la lentille. Ils sont représentés par le polynôme d'approximation dont les coefficients seront estimés. Par contre, les paramètres extrinsèques définissent l'emplacement et l'orientation de la caméra dans l'espace monde. Par conséquent, la géométrie d'imagerie stéréoscopique est déterminée à partir des paramètres intrinsèque et extrinsèques de chacune des caméras stéréo-appariée. La distance entre les deux capteurs (Baseline) et les positions des différents pointeurs laser sont déterminés en dernier lieu. La figure 7 présente les différentes étapes successivement simulées.

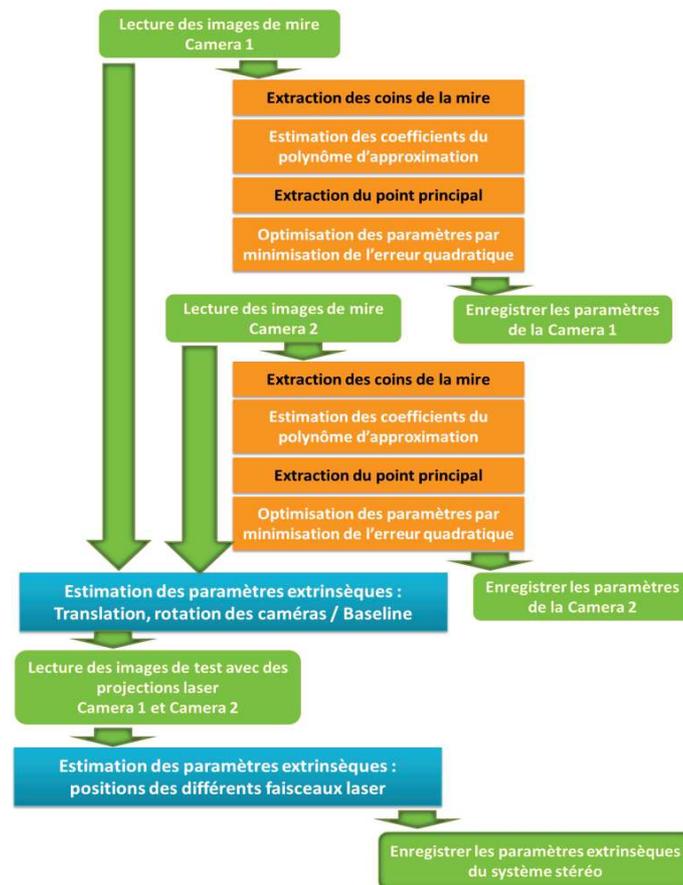


Figure III. 7: Diagramme d'étalonnage stéréo du système catadioptrique et à lasers

Afin de simplifier le test itératif pour l'étalonnage de notre système stéréoscopique à système de lasers structurés, nous avons créé une interface graphique (Fig.8) qui fait appel aux fonctions proposées par Scaramuzza pour chaque capteur catadioptrique séparément puis les fonctions qu'on a proposées pour l'estimation de la Baseline et la position des faisceaux laser par rapport au repère de référence.

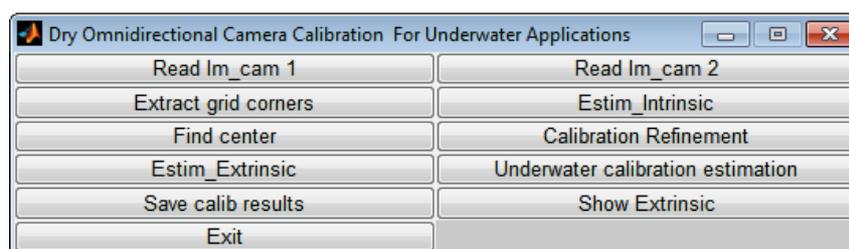


Figure III. 8: Interface graphique de simulation

3.3. Validation de l'approche d'étalonnage

Afin de valider le processus complet d'étalonnage, on a testé l'algorithme d'estimation sur un système stéréoscopique à orientation connue (R, T) à miroirs hyperboliques de paramètres connus et avec des pointeurs lasers à positions connues (Z_{laser}) . Le système de vision est créé sous le logiciel Pov-Ray qui nous permet de fixer le modèle géométrique des deux capteurs ainsi que la

position des pointeurs laser (Fig.9.a). Afin de simplifier la projection, on a supposé que le repère monde (X, Y, Z) est confondu avec celui de la caméra du dessus (caméra1) et que les deux systèmes catadioptriques sont parfaitement colinéaires. Pour un point monde P, on a deux projections p1 et p2 sur les plans images des deux caméras de centres optiques (C1, C2). La distance F1F2 entre les deux foyers des miroirs représente la baseline "Be" du système stéréoscopique. Chaque position laser peut être obtenue soit en utilisant la caméra1 ou la caméra2 comme illustré dans la figure 9.b ou la moyenne des deux. Ceci nous permet de définir et valider l'expression géométrique qui permet de déterminer la position de chaque faisceau laser par rapport au repère caméra (Eq.8).

$$PosLaser = \frac{((Yp.tan(\theta_1)) + ((Yp.tan(\theta_2)) + Baseline))}{2} \quad (8)$$

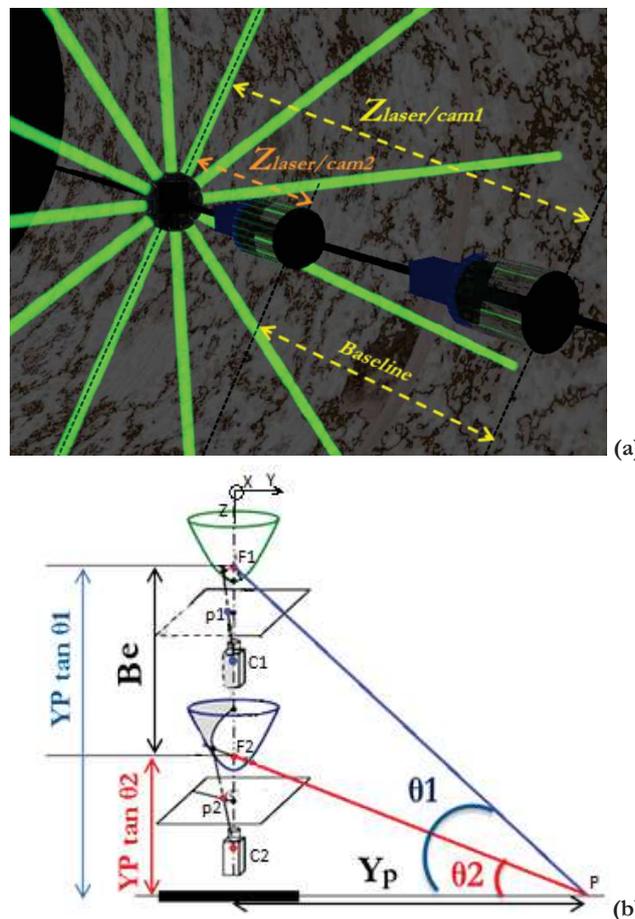
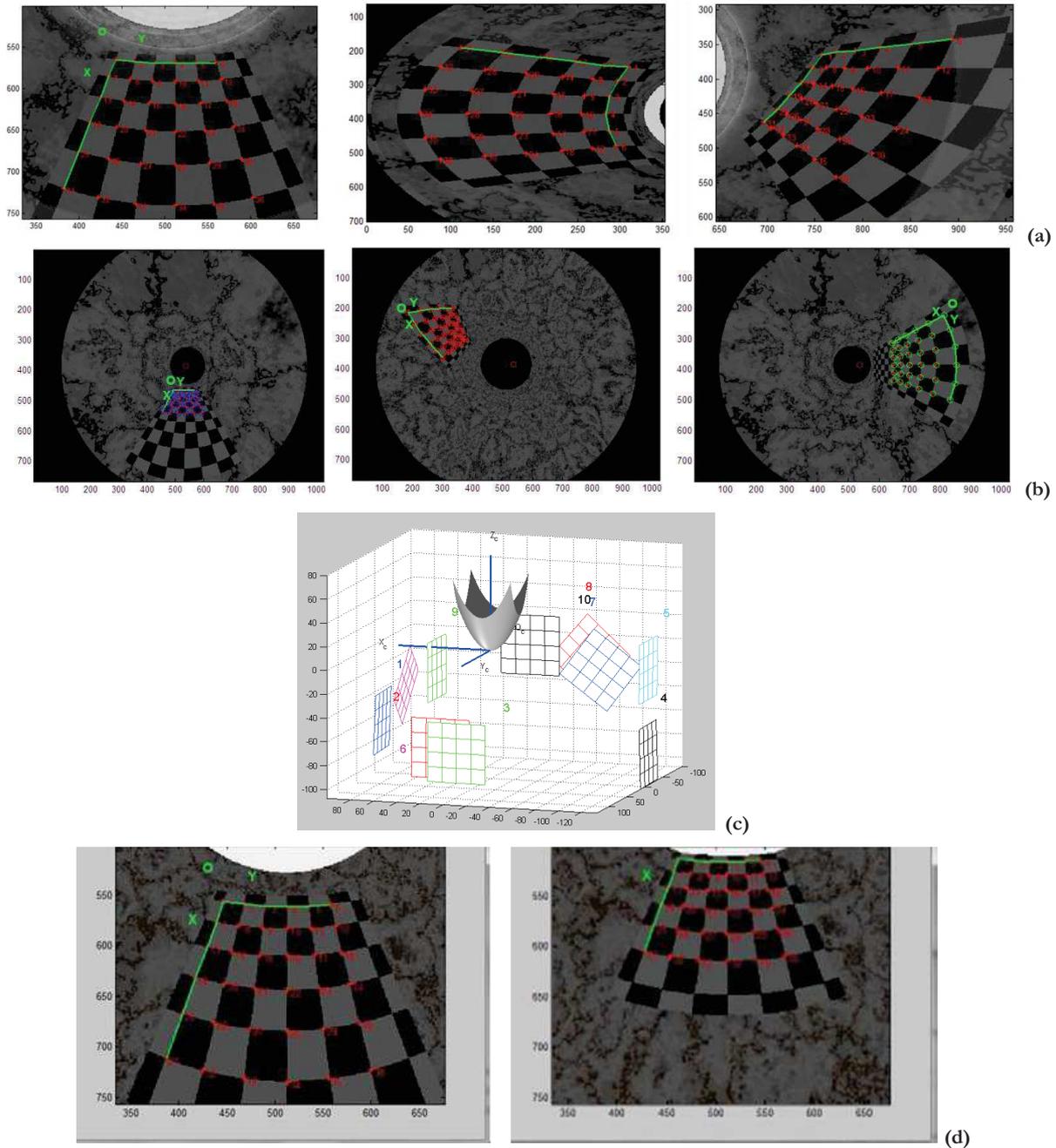


Figure III. 9: Illustration de la relation entre Baseline et les positions des lasers

Plusieurs orientations de la mire ont été capturées par les deux caméras et dont les coins ont été extraits automatiquement (Fig.10.a). Après estimation automatique du centre du miroir indiquant le point principal, un algorithme d'optimisation est appliqué pour minimiser l'erreur entre les points 2D détectés et les points 2D estimés par projection des points 3D (Fig.10.b). Les paramètres extrinsèques permettent de reconstruire et visualiser les mires dans l'espace 3D

(Fig.10.c). La distance entre les deux capteurs " B_e " est déterminée par l'application de la distance euclidienne sur les mesures de translation (Eq.9) d'une même mire (Fig.10.d).

$$B_e = \sqrt{(t_{1x} - t_{2x})^2 + (t_{1y} - t_{2y})^2 + (t_{1z} - t_{2z})^2} \quad (9)$$



En utilisant les paramètres estimés du polynôme d'approximation, on peut créer un ensemble de points 3D et leurs correspondants 2D pour pouvoir déterminer les paramètres du miroir correspondants (ξ_e, ψ_e, d_e) qui sont par la suite comparés par rapport à ceux définis lors de la modélisation des miroirs sur Pov-Ray (ξ_0, ψ_0, d_0) comme présenté dans le Tableau 1.

Tableau III. 1 : Paramètres intrinsèques de la caméra 1

Coefficients du polynôme					Point principal	
$a_0 = -174,4452$	$a_1 = 0$	$a_2 = 5,4172 \text{ e-}04$	$a_3 = 2,5099 \text{ e-}06$	$a_4 = -3,6370 \text{ e-}08$	512.4990	384.3685
Paramètres du miroir correspondant à la caméra 1						
Paramètres définis dans PovRay			Paramètres estimés par étalonnage			
$\xi_0 = 0.9386$	$\psi_0 = 1.2836$		$\xi_e = 0.9457$	$\psi_e = 1.2708$		
$d_0 = 0.0612$	$B_0 = 0.3$		$d_e = 0.0654$	$B_e = 0.2873$		

Dans la figure 11, on présente une vue 3D des pointeurs laser pointant sur la surface d'un objet de forme cylindrique dont les projections 2D sont données sur les deux images vues par les deux caméras (points 2D en vert). En appliquant l'approche d'approximation de Scaramuzza, l'estimation de la "Baseline" et des positions des lasers sur l'ensemble des points 2D (points en vert), on a obtenu un ensemble de points 3D reconstruits (points 3D en rouge) dont leurs projections 2D sont aussi données sur les images acquises (points 2D en rouge). Ces points représentent l'intersection de chaque faisceau laser avec la surface de l'objet. La comparaison entre les points réels et les points estimés nous permet de mettre en évidence la bonne estimation des paramètres du modèle géométrique du système stéréoscopique. L'erreur de re-projection est due à l'erreur d'estimation. Dans le cas d'un système réel, une différence légère peut aussi être rajoutée et qui peut être expliquée par les distorsions dans le modèle qui néglige le faible désalignement entre le miroir et la caméra et qui suppose que les deux capteurs sont parfaitement colinéaires. Des expérimentations utilisant le système réel seront présentées dans les sections suivantes pour évaluer cette erreur.

L'utilisation des paramètres intrinsèques du miroir comme définis et utilisé dans le modèle de projection proposé par Mei ou les paramètres définis et utilisé dans le modèle de projection proposé par Scaramuzza aboutit presque à un même résultat. En effet, on a expérimenté les deux approches de reconstruction sur un même objet 3D (Fig.12.a). Les images stéréoscopiques (Fig.12.b) présentent les projections image des pointeurs laser qui sont détectés et correspondus pour générer la forme intérieure de l'objet et donner une estimation des coordonnées 3D de chaque projection comme indiquée sur la figure 12.c. La comparaison entre le modèle 3D reconstruit en utilisant l'approche de Christopher Mei (en bleu) et le modèle 3D reconstruit en utilisant l'approche de Scaramuzza (en rouge) montre que les deux méthodes génèrent presque la même forme de l'objet confiné. Les mesures en distance données en centimètres présentent une erreur maximale d'environ 1.27 cm pour l'objet mis en test.

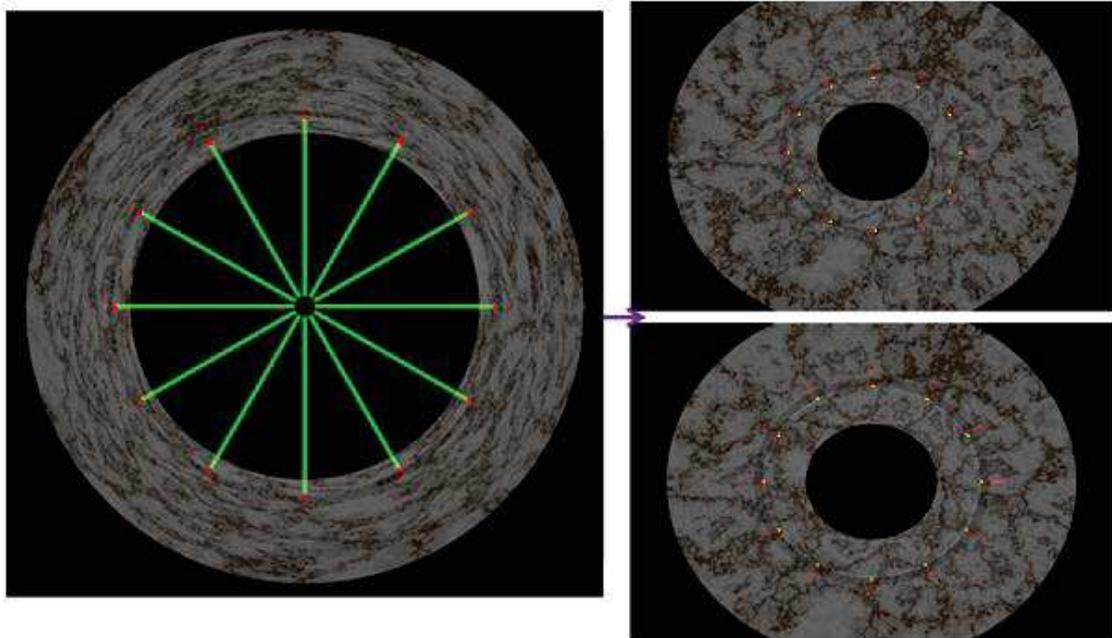
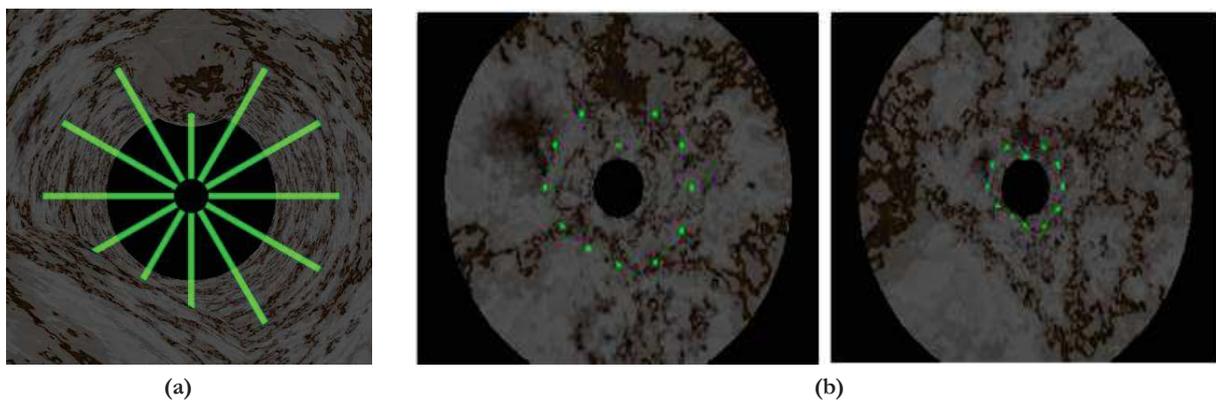
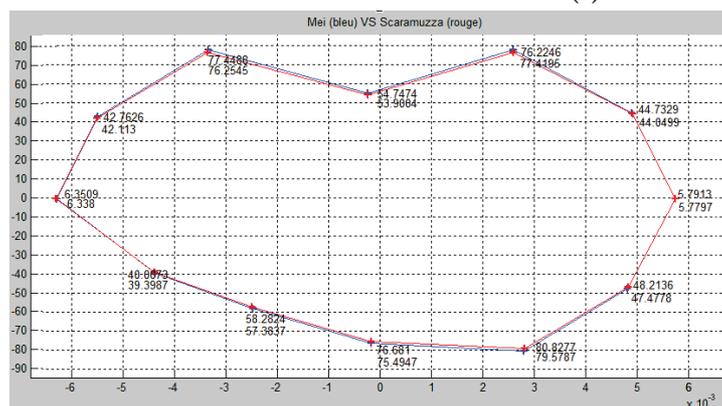


Figure III. 11: Comparaison entre points 3D réels (en vert) et points 3D reconstruits (en rouge)



(a)

(b)



(c)

Figure III. 12: Résultat de comparaison : (a) Modèle 3D réel (b) Modèle reconstruit selon Scaramuzza (en rouge) et selon Mei (en bleu)

3.4. Etalonnage du système stéréoscopique réel

Pour effectuer l'étalonnage du système stéréo catadioptrique, on utilise l'approche d'étalonnage pour modéliser géométriquement notre système complet. En se basant sur les

expérimentations réalisées, on a constaté que plus la mire est grande plus ces carreaux sont visibles par l'algorithme de détection automatique et plus les coordonnées des points 2D correspondants aux coins des mires sont plus précis ce qui permet une bonne estimation des paramètres. On positionne la mire à plusieurs orientations et positions de manière à couvrir toute la surface du miroir et pouvoir estimer la fonction de passage 2D/3D aussi précis pour les points les plus proches que les plus lointains. Ensuite, on applique l'algorithme de détection des points particuliers tout en s'assurant visuellement de la bonne extraction des coins de la mire (Fig.13.a). Le point principal est considéré au début comme étant le centre de l'image pour pouvoir générer une première estimation des paramètres. Puis, en minimisant l'erreur de re-projection par un algorithme itératif de RANSAC [38], on détermine des coordonnées plus précises du point principal (Fig.13.b, image au centre) afin de mettre en évidence le désalignement entre l'axe optique de la caméra et l'axe de révolution du miroir, ce qui permet d'affiner les résultats d'estimation.

La re-projection permet d'identifier les images qui présentent le plus d'imprécision afin de, soit les éliminer ou remplacer soit mettre à jour les coordonnées des coins manuellement. Le processus d'étalonnage est répété jusqu'à atteindre une erreur de re-projection la plus minimale et la plus proche de zéro (Fig.13.b, image à droite).

Les résultats d'estimation des paramètres intrinsèques et extrinsèques des deux capteurs, estimés à partir des paires d'images, sont représentés dans le tableau 2.

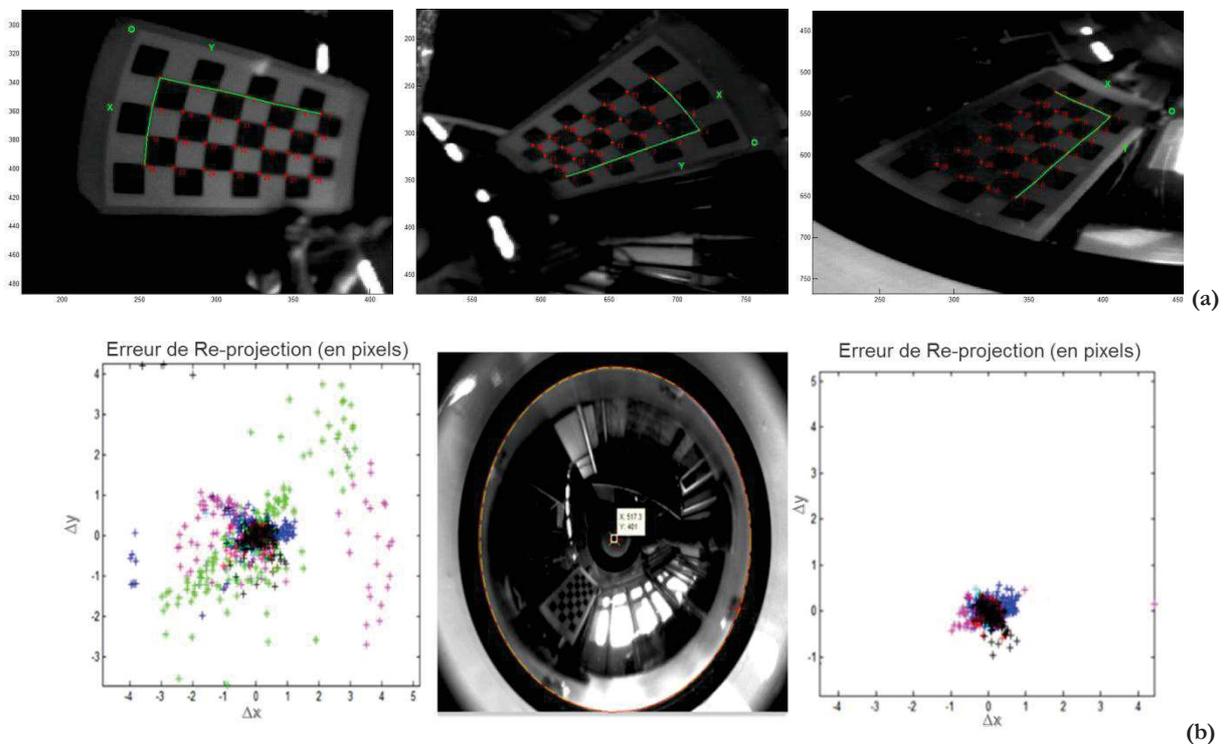


Figure III. 13: Résultat de simulation du processus d'étalonnage

Tableau III. 2 : Paramètres intrinsèques et extrinsèques dans un environnement aérien

Paramètres	Camera 1	Camera 2
u_0 (pix)	401.06	412.26
v_0 (pix)	517.05	489.05
a_0	-133.3703334	-125.45862517
a_1	0	0
a_2	0.000025719	0.0000107122
a_3	-0.051691 e-06	0.4108 e-06
a_4	9.0255 e-09	7.4801 e-09
Rotation matrix	$R_{10} = \begin{bmatrix} 0.9998 & -0.0110 & -0.0089 \\ 0.0109 & 0.9999 & -0.0040 \\ 0.0090 & 0.0039 & 0.9999 \end{bmatrix}$	$R_{21} = \begin{bmatrix} 0,9891 & -0,1253 & -0,0776 \\ 0,1325 & 0,9866 & 0,0953 \\ -0,0646 & 0,1045 & -0,9924 \end{bmatrix}$
Translation vector	$T_{10} = \begin{bmatrix} 1.00 \\ 0.013 \\ 0.006 \end{bmatrix}$	$T_{21} = \begin{bmatrix} 1.00 \\ 0.011 \\ 34.2476 \end{bmatrix}$
PosLaser (m)	0.49	0.28
Erreur de reprojection (pix)	0.342	0.468

3.4.1. Etalonnage du système stéréoscopique réel dans l'eau

La variation de l'indice de réfraction dégrade les images à cause des particules d'eau en mouvement sous l'eau. Il introduit aussi des réflexions et des réfractions des faisceaux lumineux, ce qui modifie le modèle des paramètres des caméras estimés dans l'air et par conséquent toute la reconstruction 3D peut être affecté. Il est donc obligatoire de passer par un processus d'étalonnage avec mire de calibration dans un milieu aquatique qui est, en effet, similaire à celui réalisé dans l'air libre, sauf que les images des mires doivent être acquises dans l'eau (Fig.14). Les paramètres d'étalonnage sont représentés dans le tableau 3.

Tableau III. 3 : Paramètres intrinsèques et extrinsèques dans un environnement aquatique

Paramètres	Camera 1	Camera 2
u_0 (pix)	386.264	387.48
v_0 (pix)	513.18	515.05
a_0	-75.43529834	-52.08678
a_1	0	0
a_2	-0.000405517	0.00010486516
a_3	0.071614 e-06	0.701670 e-06
a_4	17.3038 e-09	11.575 e-09
Rotation matrix	$Re_{10} = \begin{bmatrix} 0.9219 & -0.0354 & -0.1289 \\ 0.0587 & 0.9789 & -0.0406 \\ 0.0466 & 0.0368 & 0.9569 \end{bmatrix}$	$Re_{21} = \begin{bmatrix} 0.9891 & 0.1379 & -0.0524 \\ -0.1058 & 0.2153 & -0.3035 \\ 0.1028 & -0.1992 & 0.9253 \end{bmatrix}$
Translation vector	$T_{10} = \begin{bmatrix} 1.00 \\ 0.087 \\ 0.013 \end{bmatrix}$	$T_{21} = \begin{bmatrix} 1.00 \\ 0.059 \\ 41.07227 \end{bmatrix}$
PosLaser (m)	0.6391	0.4257
Erreur de reprojection (pix)	0.218	0.245

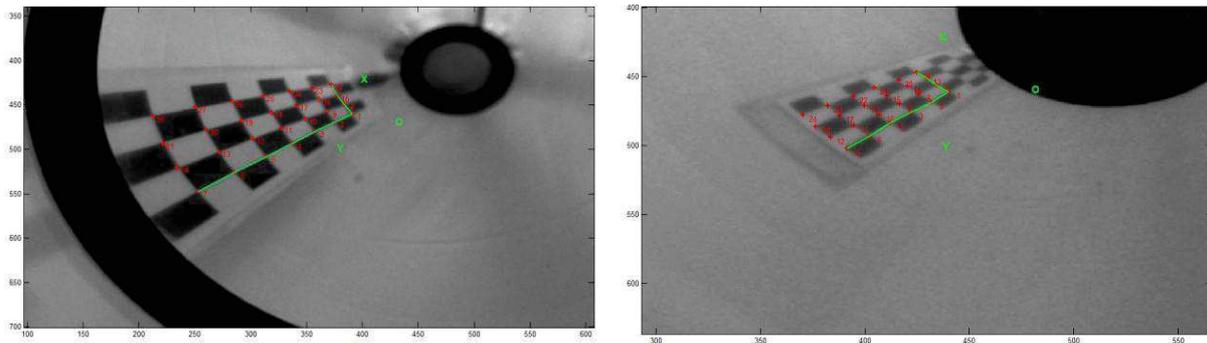


Figure III. 14: Image stéréo d'une même mire vue sous l'eau

3.4.2. Nouvelle approche d'étalonnage dans l'air pour une utilisation sous l'eau

Afin de donner plus de précision aux résultats de reconstruction, les paramètres du système doivent être estimés le plus soigneusement possible et ceux en effectuant plusieurs étalonnages et en analysant l'erreur de re-projection. L'étalonnage du système, sous l'eau, a demandé plusieurs mesures de protection à savoir ; assurer l'étanchéité des caméras, des pointeurs lasers et de la mire de calibration. En plus, à cause de la dégradation des images, la détection automatique des coins de la mire n'a pas pu être effectuée automatiquement, contrairement au processus réalisé dans l'air. Ceci a rendu cette tâche encore plus difficile et contraignante sachant que le processus de détection est semi-automatisé et que pour chaque image, on est amené à détecter une trentaine de coins pour une dizaine d'image pour chaque caméra, c'est-à-dire pointer environ 600 coins à chaque itération. Pour toutes ces raisons, on a proposé une nouvelle approche qui adapte les paramètres d'étalonnage dans l'air pour une application sous l'eau basée essentiellement sur des relations d'optiques. Cette approche exige une maîtrise totale du processus d'étalonnage, les variables utilisées ainsi que la géométrie de notre système.

L'utilisation des paramètres d'étalonnage dans l'air pour une application sous l'eau peut être effectuée selon deux méthodes (Fig.15) :

- La première consiste à suivre la trajectoire du rayon 3D, traversant les différents milieux, étapes par étapes (Fig.15, chemin 1): sa projection dans l'air, son intersection avec la première surface séparatrice air/plexi et déterminer le point d'intersection, l'angle d'inclinaison et le rayon 3D traversant le plexi, son intersection avec la deuxième surface séparatrice plexi/eau et déterminer le point d'intersection, l'angle d'inclinaison et le rayon 3D traversant l'eau et ensuite son intersection avec le faisceau laser correspondant. Cet algorithme exige plusieurs opérations arithmétiques et trigonométriques ce qui peut alourdir le processus en temps d'exécution.

- La deuxième méthode consiste à intersecter le rayon 3D traversant l'air avec une position décalée sur l'axe des Z, par un offset noté "ep" (Fig.15, chemin 2). Cette valeur tient en considération les effets de réfraction et elle dépend des angles de réfractions qui sont déterminés, simplement, par la loi de Snell.

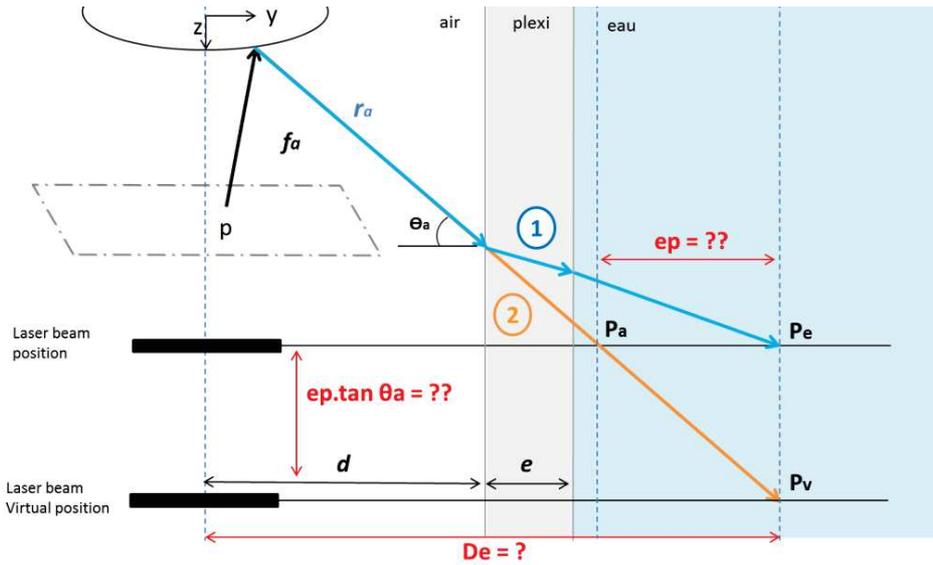


Figure III. 15: Illustration géométrique du passage air/plexi/eau

Les coordonnées du point apparent " P_a " sont représentées géométriquement par rapport au rayon 3D correspondant " r_a " (Eq.9). Les coordonnées du point plongé dans l'eau " P_e ", en considérant les effets de réfraction, sont représentées par l'expression (Eq.10). La détermination de l'offset " e_p " est, en effet, une reformulation plus simple de la première méthode car ce décalage représente la différence entre le point 3D reconstruit avec et sans considération des effets de réfraction (Eq.11). La deuxième expression de " P_e " sera utilisée pour reconstruire les points 3D correspondants à partir des points 2D. Ces correspondances 2D-3D représentent les données d'entrée de l'algorithme d'estimation des paramètres du système de vision pour une application sous l'eau.

$$r_a = \begin{bmatrix} X_r = x - x_c \\ Y_r = y - y_c \\ Z_r = f(X_r, Y_r) \end{bmatrix}; \quad \theta_a = \arctan\left(\frac{Z_r}{\sqrt{X_r^2 + Y_r^2}}\right); \quad P_a = \begin{pmatrix} X_a = \sqrt{\left(\frac{Z_{laser}}{\tan \theta_a}\right)^2 - Y_a^2} \\ Y_a = Z_{laser} / \tan \theta_a \\ Z_a = Z_{laser} \end{pmatrix} \quad (9)$$

$$P_e = \begin{bmatrix} X_e = \sqrt{\left(\frac{Z_{laser}}{\tan \theta_e}\right)^2 - Y_e^2} \\ Y_e = d + e + \left(\frac{Z_{laser} - d \cdot \tan \theta_a - e \cdot \tan \theta_p}{\tan \theta_e}\right) \\ Z_e = Z_{laser} \end{bmatrix} = \begin{bmatrix} \sqrt{\left(\frac{Z_{laser}}{\tan \theta_e}\right)^2 - Y_e^2} \\ e_p + Z_{laser} / \tan \theta_a \\ Z_{laser} \end{bmatrix} \quad (10)$$

$$e_p = Y_{P_e} - Y_{P_a} = Z \cdot \left[\frac{1}{\tan(\theta_e)} - \frac{1}{\tan(\theta_a)} \right] + d \cdot \left[1 - \frac{\tan(\theta_a)}{\tan(\theta_e)} \right] + e \cdot \left[1 - \frac{\tan(\theta_p)}{\tan(\theta_e)} \right] \quad (11)$$

En utilisant cette approche, deux techniques de reconstruction 3D peuvent être réalisées. La première est, bien évidemment, en intersectant le rayon vecteur correspondant " r_a " par rapport à la position du faisceau laser décalée par un offset " e_p ". Celle-ci permet de générer les coordonnées 3D d'un point laser projeté sur l'image en des coordonnées 2D. La deuxième méthode est d'utiliser plusieurs correspondances 2D-3D générés par la première méthode et leur appliquer l'algorithme d'étalonnage qui permet d'estimer des paramètres du système de vision. La fonction de passage 2D-3D obtenue permet de générer les coordonnées réels d'un point monde sous l'eau par application directe de la fonction d'approximation " f_e " (Fig.16).

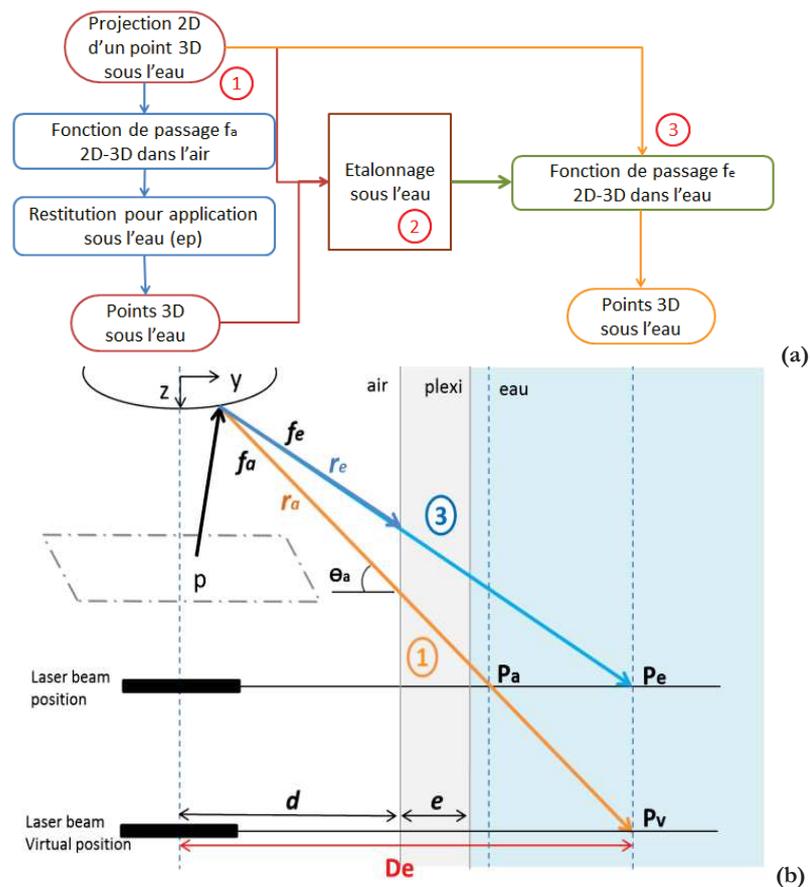


Figure III. 16: Illustration géométrique de la nouvelle approche d'étalonnage

3.5. Estimation de l'indice de réfraction

L'algorithme de reconstruction 3D basé sur l'approche du tracé de rayon nécessite le calcul des angles de réfraction par rapport à chaque surface séparatrice entre deux milieux différents. Ces angles sont déterminés en connaissant les indices de réfraction des milieux traversés. Pour la détermination de l'angle de réfraction dans l'eau " θ_e ", on a besoin de déterminer l'indice de réfraction de l'eau notée " n_{eau} " à partir des indices de réfraction de l'air " n_{air} ", du verre " n_{verre} " et du plexi " n_{plexi} ". Néanmoins, lorsqu'on parle de l'indice de réfraction des milieux transparents, il faut entendre "moyenne des indices de réfraction pour chaque longueur d'onde du spectre

visible". En effet, le système de vision comporte, outre les caméras et les miroirs, un système en verre entourant chaque miroir et assurant sa propreté et un système de protection global en plexi englobant les deux capteurs et la ceinture des lasers assurant leurs étanchéités (Fig.17, image à droite). L'indice de réfraction de chaque élément (air ou verre ou plexi) est connu, sauf que celui du système complet qui doit être estimé afin de pouvoir déterminer l'angle d'inclinaison " θ_e " du rayon vecteur traversant l'eau à sa sortie du système de protection en plexi.

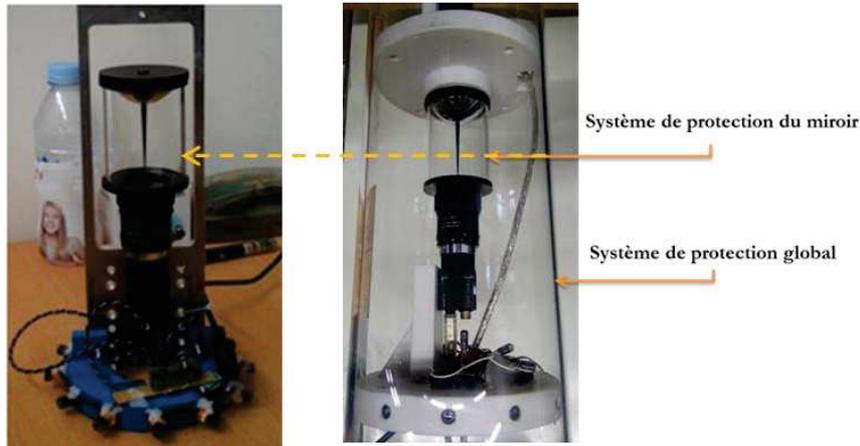


Figure III. 17: Systèmes de vision avec un système d'étanchéité

Afin de donner plus de précision aux résultats de reconstruction 3D, on a proposé une technique qui permet d'estimer les indices de réfraction nécessaires tout en prenant en considération toutes les spécifications liées à notre matériel (type du plexi, du verre, salinité de l'eau...) et conditions de test (température, pression...). Cette estimation est réalisée en utilisant les résultats d'étalonnage du système dans différentes structures : 1/ dans l'air (Fig.17, image à gauche), 2/ dans l'air avec le système en plexi (Fig.17, image à droite), et 3/ dans l'eau avec le système en plexi (Fig.17, image à droite) appliqués sur un même ensemble de point 2D et en utilisant la loi de Snell-Descartes.

Pour la détermination de l'indice de réfraction, on a appliqué les fonctions de re-projection estimées dans l'air (sachant que n_1 : indice de réfraction de l'air) et dans l'air avec le système en plexi et en verre (avec n_{a+v+p} : indice de réfraction de l'air avec le plexi et le verre) sur le même ensemble de points 3D de coordonnées (X, Y, Z). L'angle de réfraction " θ_a " est déterminé en utilisant les coordonnées du rayon vecteur r_a (X_{ra} , Y_{ra} , Z_{ra}). Ayant l'indice de réfraction dans l'air $n_1=1$, on détermine l'indice de réfraction du système complet " $n_{air+verre+plexi}$ " en utilisant la relation de Snell (Eq.12). Les expérimentations ont donné une valeur d'indice de réfraction de 1.26, en moyenne.

$$n_{a+v+p} = \frac{n_1 \times \sin(\theta_a)}{\sin(\theta_{a+v+p})} ; \theta_a = \frac{Z_{ra}}{\sqrt{X_{ra}^2 + Y_{ra}^2}} \text{ et } \theta_{a+v+p} = \frac{Z}{\sqrt{X^2 + Y^2}} \quad (12)$$

Ayant déterminé " n_{a+v+p} ", l'indice de réfraction de l'ensemble air+plexi, on applique le même processus pour déterminer l'indice de réfraction dans l'eau " n_{eau} ". En effet, on utilise les paramètres des caméras estimées lors d'un étalonnage avec le système en plexi pour déterminer les rayons vecteurs $r_{a+v+p}(X_{ra+v+p}, Y_{ra+v+p}, Z_{ra+v+p})$ correspondants à un ensemble de point 3D de coordonnées connues (X, Y, Z) . Ayant comme indice de réfraction du milieu à l'intérieur du système global de protection " $n_{air+verre+plexi} = 1.26$ ", on détermine par application de la loi de Snell (Eq.13) l'indice de réfraction de l'eau qui est égale, après expérimentations, à 1.3392 en moyenne.

$$n_{eau} = \frac{n_{a+v+p} \times \sin(\theta_{a+v+p})}{\sin(\theta_{eau})} ; \theta_{a+v+p} = \frac{Z_{ra+v+p}}{\sqrt{X_{ra+v+p}^2 + Y_{ra+v+p}^2}} \text{ et } \theta_{eau} = \frac{Z}{\sqrt{X^2 + Y^2}} \quad (13)$$

4. Détection et correspondance de points caractéristiques

4.1 Détection des points laser

Dans le chapitre précédent, on a présenté plusieurs techniques de détection de points caractéristiques résultants de la projection de pointeurs laser sur une surface d'un objet de la scène en se basant sur sa couleur, forme, texture, surface occupée, ...etc. Afin d'améliorer le taux de détection, dans notre approche, nous allons combiner deux informations : la couleur et la région (droite) de projection. L'information couleur permet de définir la marge de couleur qui représente une projection du laser. Cette marge est définie en analysant plusieurs images dans différentes conditions de luminosité car l'information couleur varie selon l'éclairage. L'analyse de l'information couleur du pointeur vert nous a permis aussi de constater que la projection du laser est, en effet, un nuage de points blancs entouré par un voisinage vert, ce qui nous permet de différencier entre un laser et tout autre objet de la scène (Fig.18.a et Fig.18.b).

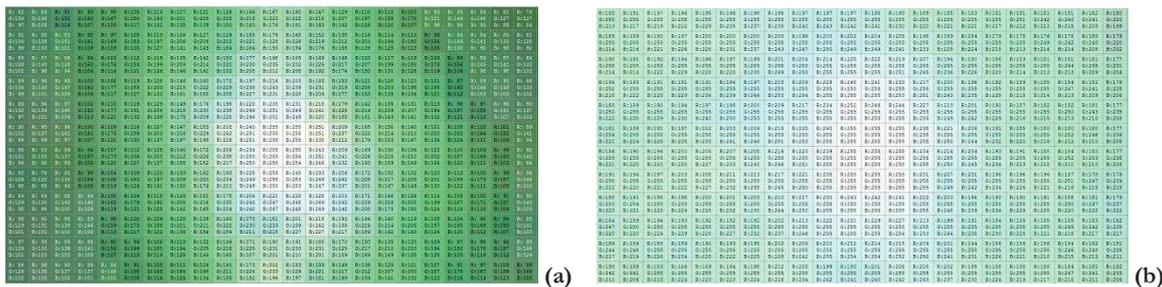


Figure III. 18: Projection du laser (a) environnement éclairé (b) environnement sombre

Afin de mettre en évidence les projections laser dans l'image, on a procédé par une binarisation de l'image en fonction de la composante V de l'espace RVB et des composantes de l'espace HSV. Le principe de l'espace HSV est de caractériser les couleurs de façon plus intuitive: la teinte H représente la couleur et qui dépend de la longueur d'onde, la saturation S représente le taux de pureté de la couleur qui varie entre la couleur éclatante et le niveau de gris, et la valeur V

qui représente l'intensité lumineuse de la couleur et qui varie entre le noir et le blanc. L'expression de conversion de l'espace RVB à l'espace HSV est donnée par (Eq.14).

$$\begin{aligned}
 Hue &= \begin{cases} 0 + \frac{43 \times |V - B|}{\text{Max}(R, V, B) - \text{Min}(R, V, B)}, & \text{if } \text{Max}(R, V, B) = R \\ 85 + \frac{43 \times |B - R|}{\text{Max}(R, V, B) - \text{Min}(R, V, B)}, & \text{if } \text{Max}(R, V, B) = V \\ 171 + \frac{43 \times |R - V|}{\text{Max}(R, V, B) - \text{Min}(R, V, B)}, & \text{if } \text{Max}(R, V, B) = B \end{cases} \\
 Sat &= \left[\frac{\text{Max}(R, V, B) - \text{Min}(R, V, B)}{\text{Max}(R, V, B)} \right]
 \end{aligned} \tag{14}$$

$$Val = \text{Max}(R, V, B)$$

En analysant les valeurs des composantes par rapport à une scène sombre ou éclairée, on a constaté que: les pixels verts des lasers dans l'image couleur ont la même valeur de teinte "Hue", que les composantes "Val" et "Sat" subissent une variation négligeable en variant l'éclairage et que les pixels qui entourent le centre du nuage de points ont une composante V supérieure à 200. L'expression de binarisation (Eq.15) a été sélectionnée suite à plusieurs expérimentations sur des images à différents éclairages, afin de fixer les seuils de segmentation pour chaque composante.

$$\begin{aligned}
 & \text{if } \{v(i, j) > 220 \text{ et } Val(i, j) > 0.7 \text{ et } Hue(i, j) > 0.4 \text{ et } Sat(i, j) < 0.4\} \\
 & \quad pixel = 1; \\
 & \text{else} \\
 & \quad pixel = 0; \\
 & \text{end}
 \end{aligned} \tag{15}$$

Les images résultantes présentent, outre les nuages de points correspondants aux projections des lasers, des réfractions des faisceaux lumineux laser, traversant l'eau, qui sont considérées comme fausses détections. Afin d'éliminer ces projections, on s'est basé sur les hypothèses suivantes :

- 1- seuls les pixels blancs sont analysés,
- 2- seules les réfractions qui sont projetées dans la région entourée par le contour du miroir sont considérées,
- 3- les réfractions qui n'appartiennent pas aux régions de projection estimées sont éliminées. Ces régions sont définies par rapport à l'emplacement des pointeurs laser vis-à-vis du repère caméra.
- 4- dans chaque région, une seule projection est considérée. Dans le cas contraire, le nuage de points le plus loin du centre de l'image est considéré car il représente la projection de l'extrémité du faisceau, alors que les autres nuages points sont des réflexions du faisceau lors de sa projection.

La recherche des projections laser passe, par conséquent, d'une recherche dans la totalité de l'image à une recherche plus spécifique : recherche par région. Ces régions sont définies par rapport à l'axe du faisceau lumineux correspondant. On fait correspondre à chaque faisceau, un angle $\theta \in \{0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330\}$ qui représente son inclinaison par rapport à l'horizontale. En pratique, l'assemblage des pointeurs laser peut manquer de précision ce qui nous mène à définir 12 sections délimitées par $\theta \pm \Delta\theta$ au lieu de 12 directions (Fig.19) pour garantir la détection de leurs projections dans l'image. Les résultats de segmentation ont permis de mettre en évidence les projections laser en utilisant la même expression de binarisation que ce soit pour une scène sombre (Fig.20.a) ou éclairée (Fig.20.b).

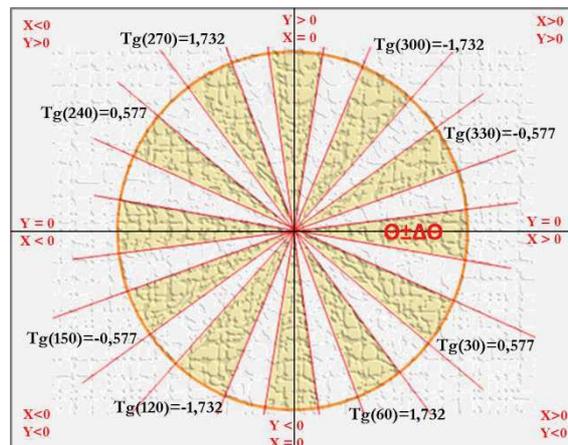


Figure III. 19: Illustration des régions de recherche

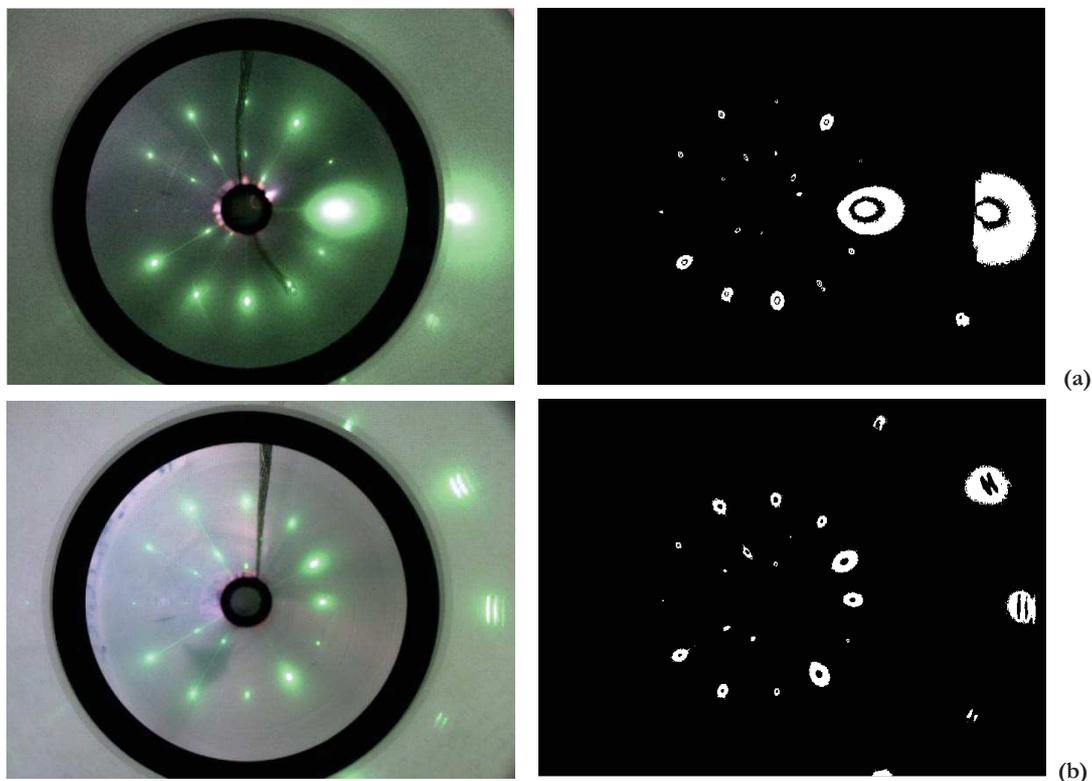


Figure III. 20: Résultat de segmentation de l'image couleur pour une scène (a) sombre et (b) éclairée

Ayant éliminé les fausses détections, on passe à la détermination du centre de gravité de chaque nuage de points. Pour cette fin, on va procéder d'abord par une classification des points appartenant au même nuage de points et à la même région. En parcourant les pixels blancs, on détermine l'angle d'inclinaison " θ " pour chaque pixel (x, y) . Selon la valeur de " θ ", les coordonnées (x, y) seront sauvegardées dans une mémoire correspondante à l'ensemble de points dans le nuage appartenant à une région donnée. Ayant parcouru toute l'image, on détermine la moyenne des valeurs de tous les pixels sauvegardés dans chaque mémoire. Cette valeur sera considérée comme le centre de gravité du nuage de points dans cette région (Fig.21.b). A la fin de la lecture de toute l'image, les coordonnées en pixels de toutes les projections sont générées simultanément.

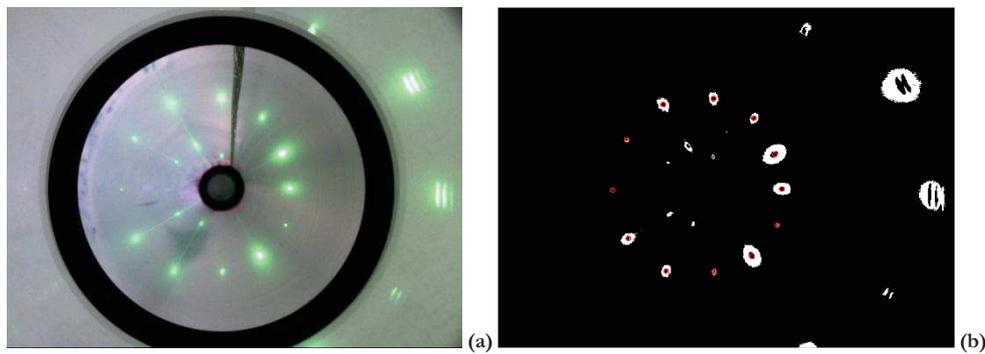


Figure III. 21: Résultats de simulation : (a) image couleur, (b) détection des points laser et (c) extraction des coordonnées 2D

4.2 Correspondance des points laser stéréo

Plusieurs techniques de correspondance existent dans la littérature, néanmoins celles-ci dépendent des caractéristiques d'un voisinage donné. Puisque la texture du conduit est presque homogène, une méthode de correspondance basée sur la région de projection sur l'image est adoptée. La configuration verticale de notre système d'imagerie stéréo simplifie cette tâche vu que les points caractéristiques correspondants existent dans la même région/direction indépendamment de la distance entre l'objet et la caméra (Fig.22).



Figure III. 22: Résultat de correspondance

5. Reconstruction 3D sous-marine

La reconstruction 3D sous-marine est une discipline nettement différente de celle dans l'air, bien qu'elle se pratique généralement avec le même matériel, sous réserve d'une adaptation. La connaissance des phénomènes de réfraction est fondamentale en matière d'optique, puisque cette loi concerne le trajet entre le sujet et les capteurs en vidéo des rayons lumineux issus du sujet. Dans cette section, on va prendre en considération l'effet de réfraction sous l'eau à travers les résultats d'étalonnage du système stéréoscopique sous l'eau et on va présenter et discuter les différentes approches de reconstruction 3D. Tout d'abord, on a commencé par l'étude des différents scénarios d'acquisition stéréoscopique. En effet, la restauration de la position 3D du point laser peut se faire suivant trois situations (Fig.23):

1- la projection du faisceau laser existe sur les deux images. Dans ce cas, le point monde est le résultat d'une reconstruction stéréo active. D'abord, chaque capteur est traité séparément : la détection des points laser dans l'image, la détermination des rayons vecteurs traversant l'air, le plexi et l'eau et la détermination des angles de réfraction correspondants. Chaque point laser 3D est représenté par deux vecteurs, issus des deux caméras, dont la source est le point central du miroir et la destination est le point monde. L'intersection de ces deux vecteurs donne les coordonnées 3D du point laser issus d'une reconstruction stéréo passive. Ensuite, chaque rayon vecteur issu de chaque caméra est intersecté avec la direction du faisceau laser pour donner les coordonnées 3D du point laser issus d'une reconstruction stéréo active. La moyenne de toutes ces reconstructions est considérée comme étant le résultat final.

2- la projection laser existe sur une seule image. Dans ce cas, nous utilisons l'algorithme de reconstruction mono-vision active. Les vecteurs rayons 3D correspondants sont obtenus en appliquant la fonction d'approximation estimée sur chaque point pixel extrait de l'image. Ces vecteurs sont intersectés avec les directions correspondantes des faisceaux laser pour donner les coordonnées 3D du point laser.

3- aucune projection n'existe sur les deux images. Nous supposons, dans ce cas, que dans cette position du robot et dans la direction du faisceau laser non visionné, il existe une fissure ou un trou profond ou une nouvelle voie menant à un autre réseau d'eau.

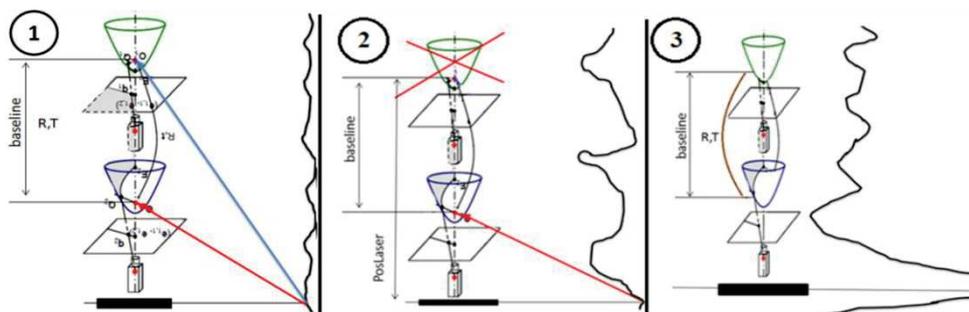


Figure III. 23: Illustration des scénarios possibles

5.1. Equations de reconstruction 3D dans l'eau avec étalonnage dans l'air

Dans ce cas, les effets de réfraction sont pris en considération en suivant la trajectoire du rayon lumineux et sa réfraction sur la surface séparatrice en plexi entre l'air et l'eau (Fig.24). Pour chaque couple $p(u, v)$ extrait de l'image, le rayon vecteur 3D est déterminé par application de la fonction d'approximation estimée dans l'air (Eq.16) et à chaque vecteur, correspond un angle d'inclinaison " θ_1 ". L'angle de réfraction " θ_2 " est déterminé en utilisant la loi de Snell et les indices de réfraction des milieux traversés (Eq.17). Les coordonnées 3D du point monde P plongé dans l'eau sont déterminées en intersectant la direction décalée du faisceau laser " Z_v " déterminé selon l'expression (Eq.18) et le rayon vecteur correspondant à sa projection. Les coordonnées du point monde données dans Eq.19 sont fonction de la distance " D_e " qui représente la distance entre l'axe optique du système de vision et la surface de l'objet pointé par le laser.

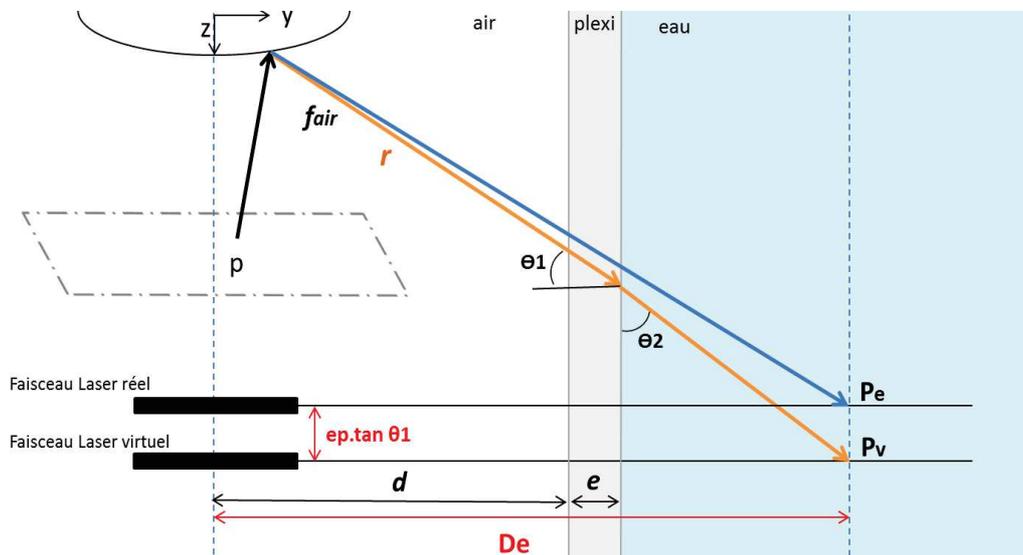


Figure III. 24: Géométrie de reconstruction 3D avec étalonnage dans l'air

$$p \begin{bmatrix} u \\ v \end{bmatrix} \leftrightarrow \vec{r} \begin{bmatrix} x_r = u - u_0 \\ y_r = v - v_0 \\ z_r = f(\sqrt{x_r^2 + y_r^2}) \end{bmatrix}; \quad \theta_1 = \arctan \left(\frac{z_r}{\sqrt{x_r^2 + y_r^2}} \right) \quad (16)$$

$$\theta_2 = \arcsin \left(\frac{n_1}{n_2} \times \sin \theta_1 \right); \quad \text{sachant que : } n_1 = n_{\text{air}} \text{ et } n_2 = n_{\text{air+verre+plexi}} \quad (17)$$

$$Z_v = Z_{\text{laser}} + e_p \times \tan(\theta_1); \quad \text{sachant que : } e_p = Z_{\text{laser}} \times \left[\frac{1}{\tan(\theta_2)} - \frac{1}{\tan(\theta_1)} \right] \quad (18)$$

$$P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \left(\frac{Z_v}{Z_r} \right) \times X_r \\ \text{sign}(y_r) \cdot \sqrt{D_e^2 - X^2} \\ Z_{laser} + e_p \times \tan(\theta_1) \end{bmatrix} ; \text{ sachant que : } D_e = \frac{Z_v}{\tan(\theta_1)} = \sqrt{X^2 + Y^2} \quad (19)$$

5.2. Equations de reconstruction 3D avec étalonnage dans l'eau

En utilisant les paramètres d'estimation du modèle géométrique du système de vision, la reconstruction des coordonnées 3D du point monde est effectuée en profitant des résultats stéréoscopiques 3D et des données relatives à notre système à savoir la position relative des pointeurs lasers. En effet, le résultat final de la reconstruction est une moyenne pondérée du résultat issu de la reconstruction stéréo passive par re-projection 2D-3D et celui de la reconstruction active en bénéficiant des données de positions des lasers.

Dans les deux cas, la transformation 2D-3D est la même. En effet, pour chaque point pixel $p(u, v)$, le rayon vecteur 3D est déterminé par application de la fonction d'approximation estimée lors d'un étalonnage sous l'eau en se basant sur la relation définie par l'expression présentée par (Eq.20).

$$p_1 \begin{bmatrix} u_1 \\ v_1 \end{bmatrix} \leftrightarrow \vec{r}_1 \begin{bmatrix} x_1 = u_1 - u_{01} \\ y_1 = v_1 - v_{01} \\ z_1 = f(\sqrt{x_1^2 + y_1^2}) \end{bmatrix} ; \quad p_2 \begin{bmatrix} u_2 \\ v_2 \end{bmatrix} \leftrightarrow \vec{r}_2 \begin{bmatrix} x_2 = u_2 - u_{02} \\ y_2 = v_2 - v_{02} \\ z_2 = f(\sqrt{x_2^2 + y_2^2}) \end{bmatrix} \quad (20)$$

➤ Equations de reconstruction stéréo passive

Dans ce cas, les pointeurs lasers ne sont utilisés que pour l'identification de points caractéristiques. Les coordonnées du point laser 3D sont obtenues par l'intersection des vecteurs rayons correspondants aux projections d'un même ensemble de points 3D sur les images stéréo. Dans l'espace 3D (Fig.25), ce point "P" est le point le plus proche du vecteur rayon issu de la caméra1 de centre optique C1 et simultanément celui issu de la caméra2 de centre optique C2.

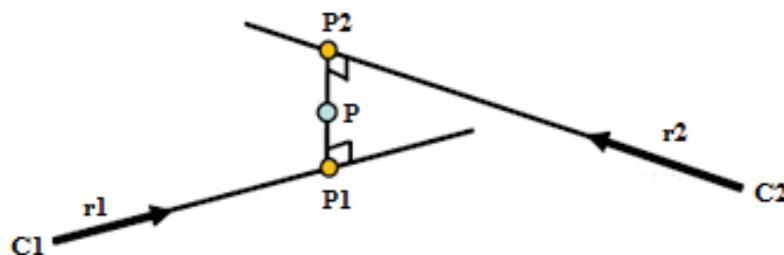


Figure III. 25: Schéma d'approximation de l'intersection de deux vecteurs 3D dans l'espace

Les expressions permettant de déterminer ce point utilisent les coordonnées 3D des rayons vecteurs $r_1(x_1, y_1, z_1)$ et $r_2(x_2, y_2, z_2)$ et les coordonnées 3D des extrémités des vecteurs qui sont, dans ce cas, les points centraux des miroirs $C_1(xc_1, yc_1, zc_1)$ et $C_2(xc_2, yc_2, zc_2)$. Les variables "La" et "Ca" représentent les distances optimales entre les deux vecteurs (r_1, r_2) et le point d'intersection optimal "P". L'algorithme de détermination du point optimum est donné comme suit : on calcule les distances optimales "La" et "Ca" (Eq.21), puis on détermine les points appartenant aux deux vecteurs qui sont les plus proches du point d'intersection "P1" et "P2" dont la moyenne représente notre point 3D reconstruit par stéréovision passive (Eq.22).

$$\begin{aligned} \cos(\theta_p) &= x_1 \times x_2 + y_1 \times y_2 + z_1 \times z_2 \\ L_a &= (x_1 - x_2 \times \cos \theta_p)(xc_2 - xc_1) + (y_1 - y_2 \times \cos \theta_p)(yc_2 - yc_1) + (z_1 - z_2 \times \cos \theta_p)(zc_2 - zc_1) \\ C_a &= (x_2 - x_1 \times \cos \theta_p)(xc_1 - xc_2) + (y_2 - y_1 \times \cos \theta_p)(yc_1 - yc_2) + (z_2 - z_1 \times \cos \theta_p)(zc_1 - zc_2) \\ P_1 &= \frac{L_a}{\cos^2 \theta_p - 1} \times \left(\begin{array}{c} x_1 \\ y_1 \\ z_1 \end{array} \right) + \left(\begin{array}{c} xc_1 \\ yc_1 \\ zc_1 \end{array} \right) \\ P_2 &= \frac{C_a}{\cos^2 \theta_p - 1} \times \left(\begin{array}{c} x_2 \\ y_2 \\ z_2 \end{array} \right) + \left(\begin{array}{c} xc_2 \\ yc_2 \\ zc_2 \end{array} \right) \end{aligned} \quad \Rightarrow \quad P = \frac{1}{2}(P_1 + P_2) \quad (22)$$

► Equations de reconstruction stéréo active

Pour chaque couple correspondant $\{p1(u_1, v_1) ; p2(u_2, v_2)\}$ extrait de l'image1 et l'image2, le même processus de calcul est appliqué en respectant les propriétés d'étalonnage (respectivement, f_{w1} et f_{w2}) dans l'eau de chaque système (Fig.26). Les angles notés " θ_{e1} " et " θ_{e2} ", représentent les angles de réfraction et sont déterminés à partir des coordonnées respectives des rayons $\{r_1, r_2\}$ issus de la projection du point monde sur le plan image de la caméra1 et la caméra2 (Eq.23). A partir de chaque projection image, on détermine les distances " D_{e1} " et " D_{e2} " en utilisant respectivement les distances estimées " Z_{laser1} " et " Z_{laser2} " de chaque faisceau par rapport à chaque repère caméra (Eq.24). Chaque projection image nous permet d'estimer les coordonnées 3D du point monde "P" qui est approximé par la moyenne du point 3D " P_{e1} " estimé à travers la projection sur le plan image de la caméra1 et celui " P_{e2} " estimé à partir de la projection sur le plan image de la caméra2. Les coordonnées 3D du point monde résultent de l'intersection des deux rayons vecteurs par rapport à la direction du faisceau laser traversant l'eau et à la position dans l'axe des Z du repère de référence. Les expressions de passage sont données comme suit (Eq.25).

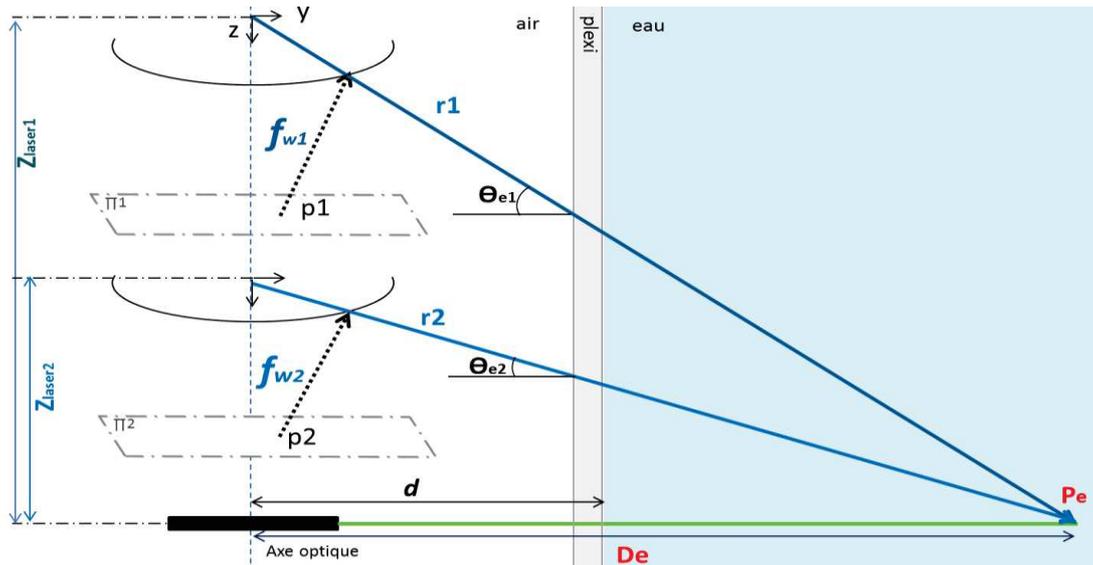


Figure III. 26: Schéma illustrant la reconstruction stéréo active d'un point 3D

$$\theta_{e1} = \arctan\left(\frac{z_{r1}}{\sqrt{x_{r1}^2 + y_{r1}^2}}\right) ; \quad \theta_{e2} = \arctan\left(\frac{z_{r2}}{\sqrt{x_{r2}^2 + y_{r2}^2}}\right) \quad (23)$$

$$D_{e1} = Z_{laser1} / \tan(\theta_{e1}) \quad ; \quad D_{e2} = Z_{laser2} / \tan(\theta_{e2}) \quad (24)$$

$$\left. \begin{aligned} P_{e1} \begin{bmatrix} X_{e1} \\ Y_{e1} \\ Z_{e1} \end{bmatrix} &= \begin{bmatrix} \left(\frac{Z_{laser1}}{Z_{r1}}\right) \times X_{r1} \\ \text{sign}(Y_{r1}) \sqrt{D_{e1}^2 - X_{e1}^2} \\ Z_{laser1} \end{bmatrix} \\ P_{e2} \begin{bmatrix} X_{e2} \\ Y_{e2} \\ Z_{e2} \end{bmatrix} &= \begin{bmatrix} \left(\frac{Z_{laser2}}{Z_{r2}}\right) \times X_{r2} \\ \text{sign}(Y_{r2}) \sqrt{D_{e2}^2 - X_{e2}^2} \\ Z_{laser2} \end{bmatrix} \end{aligned} \right\} \Rightarrow P = \frac{1}{2}(P_{e1} + P_{e2}) \quad (25)$$

6. Expérimentations et interprétation des résultats

L'organigramme, présenté dans la figure 27, donne le déroulement des opérations de traitement des images depuis l'acquisition jusqu'à l'affichage de la forme 3D accompagnée par les métriques mesurées.

A chaque position du robot, une paire d'images est acquise et une forme 3D métrique est générée. La pose successive des formes 3D permet de donner une estimation sur la structure 3D de l'environnement. Le flot commence par un étalonnage hors-ligne permettant d'estimer le modèle géométrique qui représente avec précision le comportement physique du système de

vision dans l'environnement de test. Une succession de méthodes d'analyse d'image est appliquée afin de mettre en œuvre les projections laser, éliminer les fausses détections et extraire les coordonnées des projections lasers. Les résultats obtenus, en pixels, représentent les données d'entrée de l'algorithme de reconstruction. Celui-ci consiste à une transformation 2D-3D des données puis une normalisation des vecteurs reconstruits et enfin une estimation des coordonnées 3D des points monde et leurs distances métriques par rapport à l'origine.

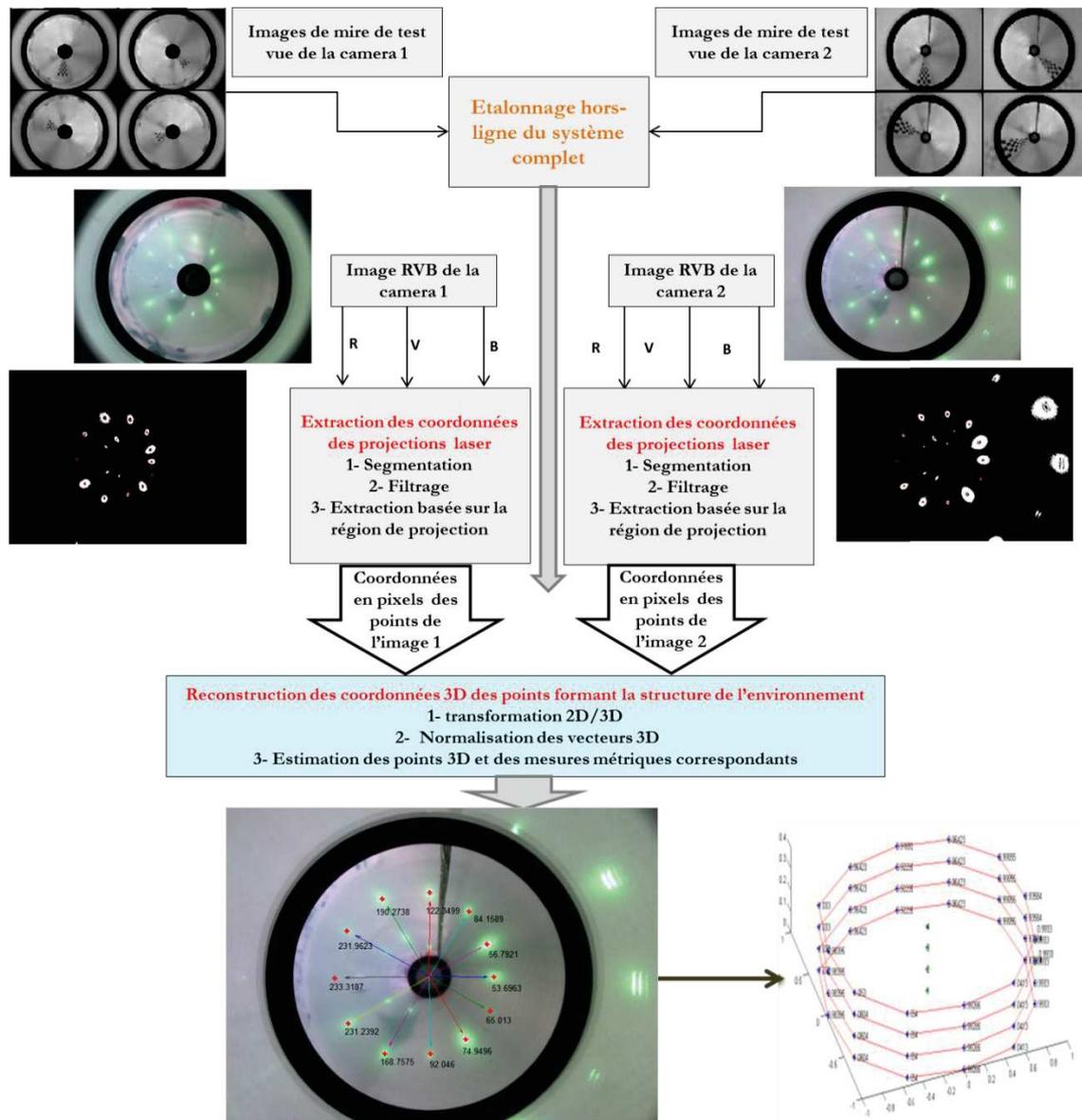


Figure III. 27: Organigramme du processus de traitement

6.1 Conditions expérimentales

La fonctionnalité principale attribuée au système développé (prototype et algorithmes) est de pouvoir effectuer des reconstructions 3D instantanées de l'environnement dynamique et confiné sous l'eau et aussi de pouvoir donner ensuite des mesures métriques les plus précises possibles, en se basant seulement sur les images. C'est la raison pour laquelle on a utilisé en premier temps

des exemples de formes géométriques simples et à dimensions connues, pour pouvoir évaluer et valider la méthode de reconstruction proposée.

Les expérimentations initiales ont été effectuées dans une piscine remplie d'eau construite au sein du laboratoire LIRMM, de forme hexagonale mesurant comme indiqué dans la figure 28. Tenant en considération les propriétés de l'eau dans cet environnement (température, pression, salinité...etc.) et après expérimentations d'étalonnages, l'indice de réfraction de l'eau obtenu est égal à 1.3392 en moyenne. Cette valeur va être utilisée dans l'algorithme de reconstruction.

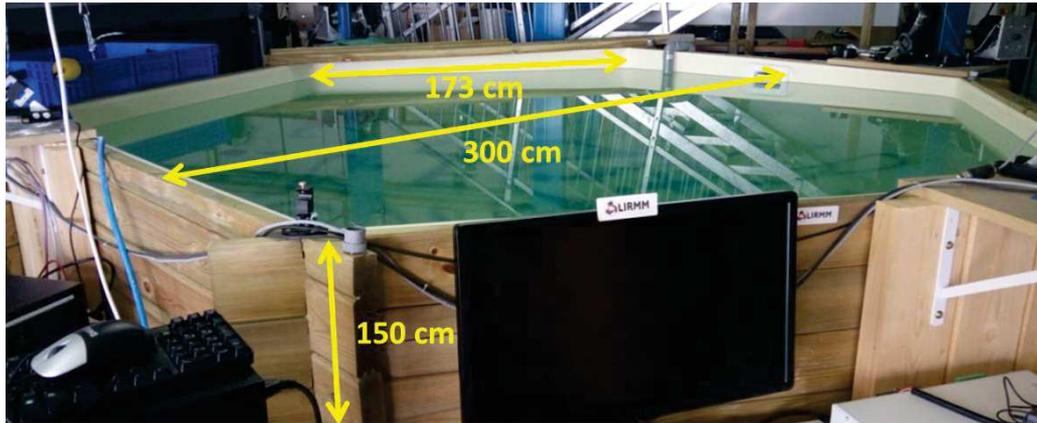


Figure III. 28: Environnement de test

6.2 Evaluation de la précision des mesures sur la structure 3D

Une mire étanche a servi d'abord pour l'étalonnage permettant la détermination des paramètres du modèle géométrique du système stéréoscopique dans l'environnement de test (Fig.29). Pour valider les paramètres estimés, on a procédé par des expérimentations de mesures pour pouvoir comparer entre les valeurs de distances réelles et estimées. Un exemple de mesures est donné, pour deux positions différentes du système de vision dans l'environnement de test. Les mesures réelles sont données par les figures 30.a et les mesures estimées sont indiquées sur la reconstruction de la forme 3D de la figure 30.b. Pour cet environnement hexagonal, l'erreur quadratique moyenne est d'environ 5.451mm^2 , ce qui pourrait être considéré comme satisfaisant vu la mauvaise résolution des images sous-marines et les effets de réfraction de la lumière.

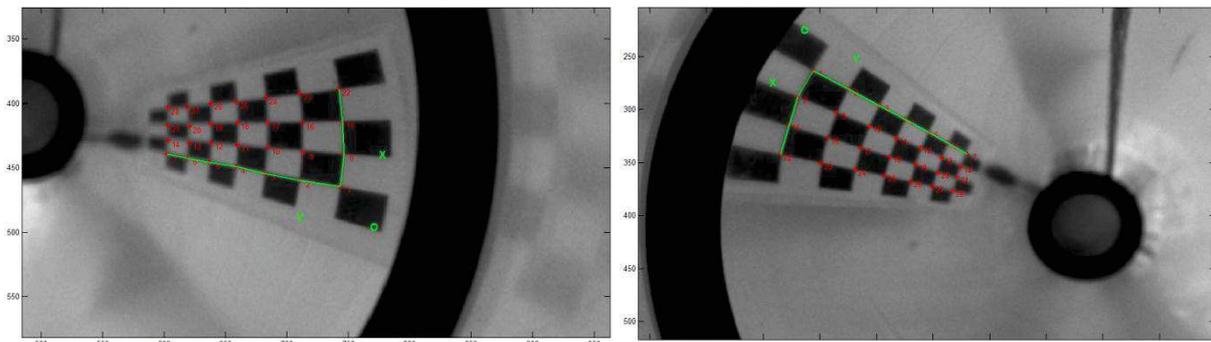


Figure III. 29: Exemple d'images prises sous l'eau avec différentes orientations de la mire d'étalonnage

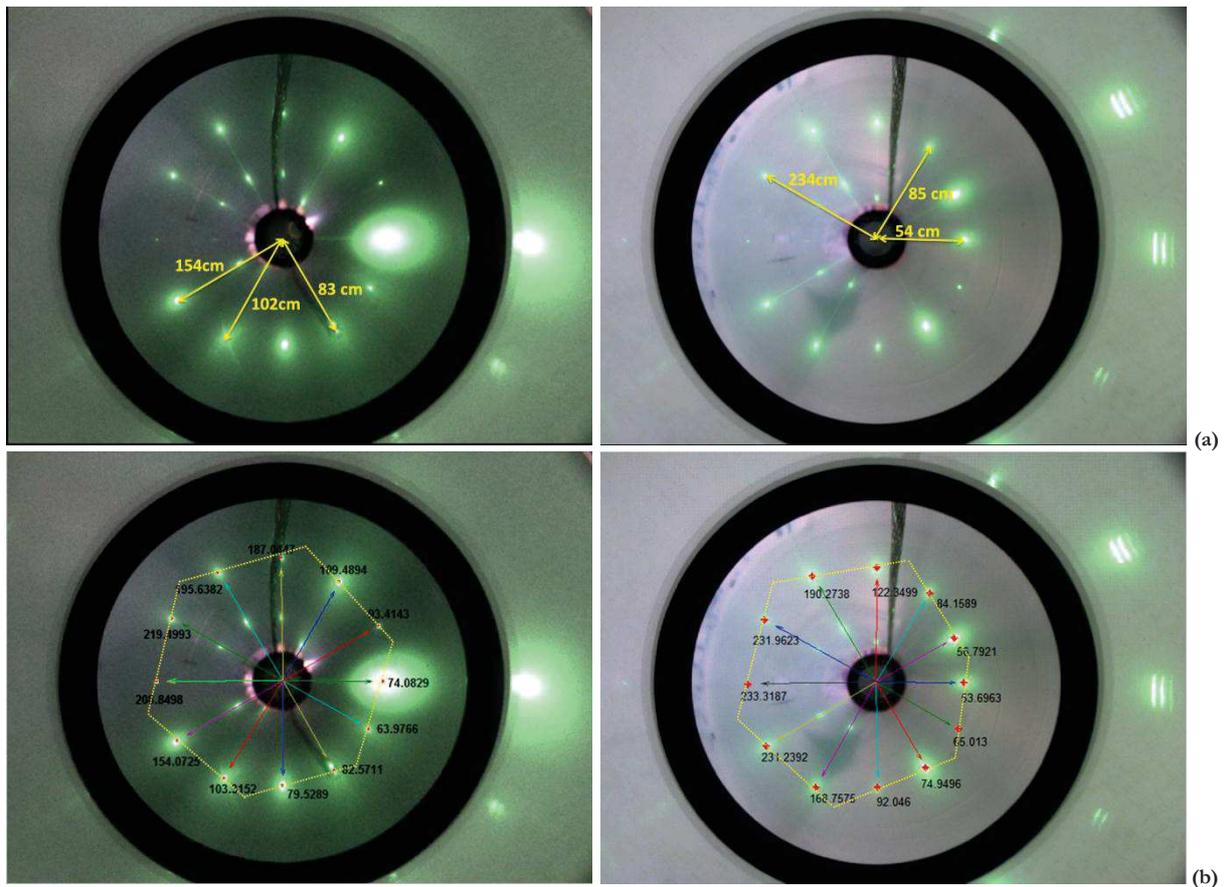


Figure III. 30: Mesures métriques sur deux différentes structures 3D sous l'eau

L'erreur obtenue entre les distances estimées et des mesures réelles peut être soit suite à l'erreur de re-projection qui a été déjà minimisée lors de l'étalonnage, ou par l'erreur de détection du centre de gravité d'un spot laser. Une analyse sur l'effet de cette dernière sur les résultats de mesures de distances a été effectuée. Etant donnée la symétrie radiale des mesures métriques, une évaluation a été effectuée sur une direction d'un même faisceau laser. Pour chaque point $p(x, y)$ d'un ensemble de points laser donnés, on calcule son rayon $r = \sqrt{x^2 + y^2}$ puis la distance lui correspondant. La figure 31 présente des mesures de distance (jusqu'à 480 cm) par rapport aux coordonnées polaires de centre de gravité dans l'image pour différentes valeurs de rayons (de 68 à 320 pixels). On présente trois exemples de mesures : pour chaque mesure, on a mesuré la distance correspondante à un centre de gravité de coordonnées (x, y) et la distance correspondante à un centre de gravité de coordonnées $(x+10, y)$. Cette courbe nous permet aussi d'évaluer l'effet d'une erreur sur la détection de la position d'un spot laser sur la distance résultante calculée. L'erreur obtenue entre les deux distances nous permet d'évaluer l'effet d'une erreur de 10 pixels sur l'une des composantes du centre de gravité. Le tableau 4 donne un exemple d'évaluation d'erreur pour une variation de $\Delta x = 1$ pixel et $\Delta x = 10$ pixels. Pour des distances allant jusqu'à 480 cm, l'erreur maximale de la distance relative n'excède pas 1,3%.

Tableau III. 4 : Exemple d'évaluation d'erreur

		Point 1 $D=40\text{cm}$	Point 2 $D=115\text{cm}$	Point 3 $D=415\text{cm}$
$\Delta x = 1 \text{ pixel}$	Δr	0,88	0,93	0,97
	ΔD	0,81	0,96	2
	%	1,02%	1,01%	1%
$\Delta x = 10 \text{ pixels}$	Δr	8,83	9,28	10,73
	ΔD	10,38	12	27
	%	1,26%	1,10%	1,07%

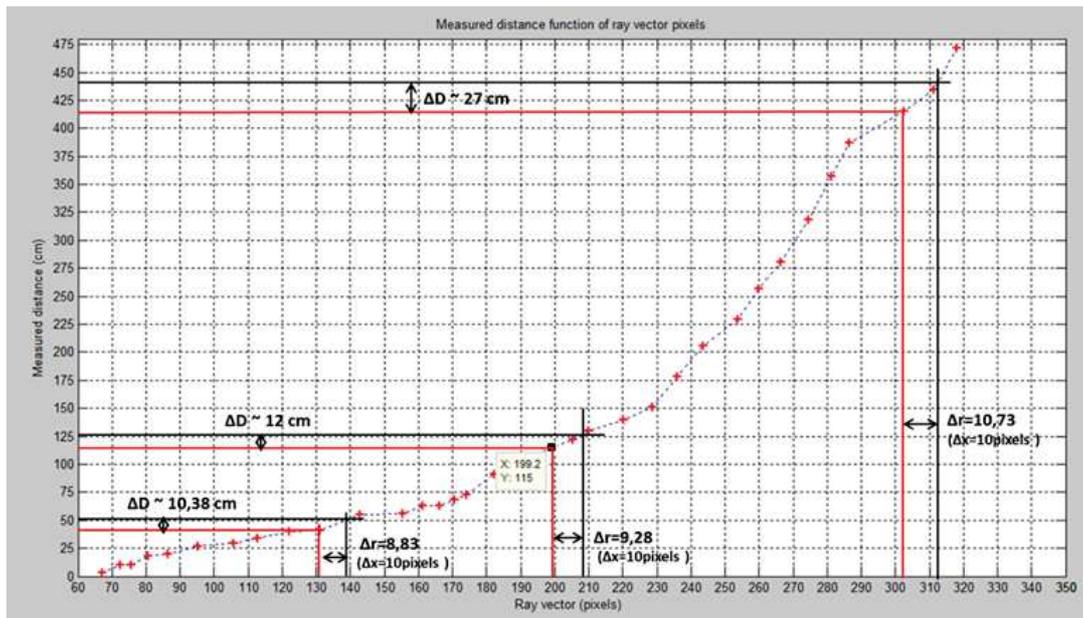


Figure III. 31: Mesures métriques de distances par rapport à la position du spot laser dans l'image

6.3 Projection des expérimentations pour application dans un environnement sous-marin réel

Le système proposé a été évalué en se basant sur des expérimentations sous l'eau dans un environnement donné avec des dimensions et des formes connues pour être en mesure de comparer les résultats obtenus par rapport aux mesures réelles et parvenir ensuite à évaluer la précision du système développé. La question qui se pose est : qu'elle est la différence entre les conditions d'essai connues et celles de l'environnement karstique réel où le système sera exploité et quel est l'effet de cette différence sur la précision du système ? L'évaluation sera considérée comme irréaliste si une différence significative existe et qu'elle affecte la précision évaluée.

Pour cette fin, les trois principales différenciations doivent être discutées: la forme, l'indice de réfraction et les propriétés de réflexion.

1- Forme : L'algorithme de reconstruction 3D de points particuliers extraits de la scène à l'aide des projections des faisceaux laser fournit des mesures de distance indépendamment de la forme

géométrique de l'environnement. L'évaluation de la précision de l'algorithme développé a été réalisée principalement sur des mesures de distances, indépendamment de la forme de l'environnement exploré.

2- Indice de réfraction de l'eau : La différence entre l'environnement de test et un environnement sous-marin réel (grottes karstiques) peut également être liée aux caractéristiques de l'eau, en particulier l'indice de réfraction. Chaque environnement sous-marin est caractérisé par une valeur spécifique d'indice de réfraction de l'eau, même à des niveaux différents de profondeurs du même milieu sous-marin. En effet, l'indice de réfraction dépend de la température, la pression, la salinité de l'eau ... etc. La meilleure solution est de pouvoir réaliser à chaque fois un étalonnage répétitif du système de vision pour raffiner les paramètres estimés. Néanmoins, cette solution est coûteuse en temps d'exécution et exige que le système de navigation soit muni, pendant tout le long de sa navigation, d'une mire d'étalonnage qui soit visible par le système stéréoscopique.

Pour ces raisons, l'effet de l'indice de réfraction sur la valeur de la précision des mesures de distance a été analysé afin d'estimer l'erreur due à des variations de l'indice de réfraction de l'eau. La courbe de la figure ci-dessous (Fig.32) présente quelques mesures de distance pour différents indices de réfraction de l'eau ($n = 1.32$, $n = 1.33$, $n = 1.34$ et $n = 1.35$). Les marqueurs noirs représentent les distances estimées à l'aide de l'algorithme de reconstruction 3D basé sur des paramètres d'étalonnage sous l'eau de l'environnement de test. Pour ces distances, une erreur relative maximale est d'environ 1.05cm pour un indice de réfraction de 1.35 (cette valeur d'indice de réfraction de l'eau ne peut pas être dépassée pour toute longueur d'onde lumineuse). En se basant sur les valeurs de distances évaluées, on peut prévoir une erreur relative maximale d'environ 1% entre la piscine et les environnements réalistes.

3- Effet de réflexion du faisceau Laser : Un faisceau laser obéit aux mêmes lois de réfraction et réflexion comme tout autre type de lumière. Cela signifie que, lorsqu'un faisceau laser se réfléchit sur une surface réfléchissante, le faisceau lumineux se réfléchit en faisant le même angle de réflexion par rapport à la surface de projection que celui incident. Cependant, le rayon réfléchi perd de la vitesse et de la puissance, même dans un environnement aérien (Fig.33.a). En effet, les surfaces réceptrices de faisceaux lumineux peuvent réfléchir, absorber et/ou transmettre des rayons lumineux selon leurs matériaux.

Dans le cas de notre application, les différentes possibilités de réflexion sont présentées dans la figure 33.b. La réflexion dépend de l'angle d'incidence du faisceau laser et de la surface réfléchissante:

- A la surface du système étanche à l'eau (en plexi), le faisceau laser est toujours perpendiculaire à la surface de séparation ce qui correspond à une réflexion totale du rayon réfléchi. Ce dernier

revient au pointeur laser dans la même direction (l'angle de réflexion est égal à l'angle incident). Dans les images stéréoscopiques acquises, la réflexion du faisceau laser à la surface du système étanche est visible parce que le système en plexi est très proche des pointeurs laser (distance maximale de 2cm). Ces réflexions sont considérées et traitées dans l'algorithme de détection des spots laser dans les sections précédentes.

- A la surface de la grotte karstique ou la piscine de test (surfaces opaques), la lumière laser est partiellement absorbée. Aussi, les photons du faisceau laser réfléchi sont dispersés à cause des propriétés physiques de l'eau (comme la viscosité). En effet, les radiations réfléchies ne sont pas projetées sur les miroirs du système catadioptrique et par conséquent elles ne sont pas visibles sur les images stéréoscopiques acquises.

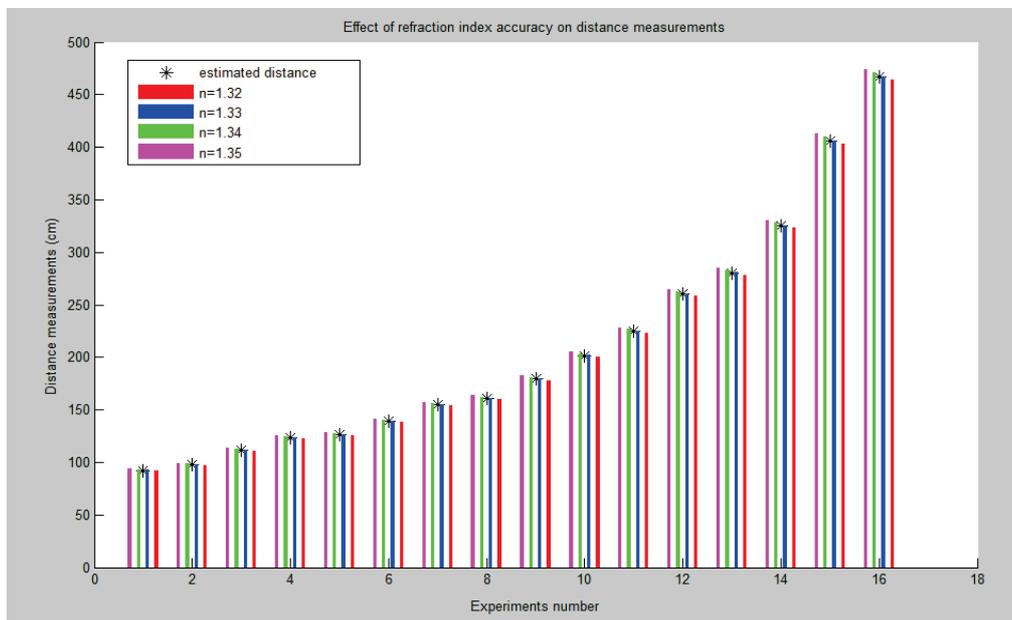


Figure III. 32: Effet de la variation d'indice de réfraction sur des mesures métriques

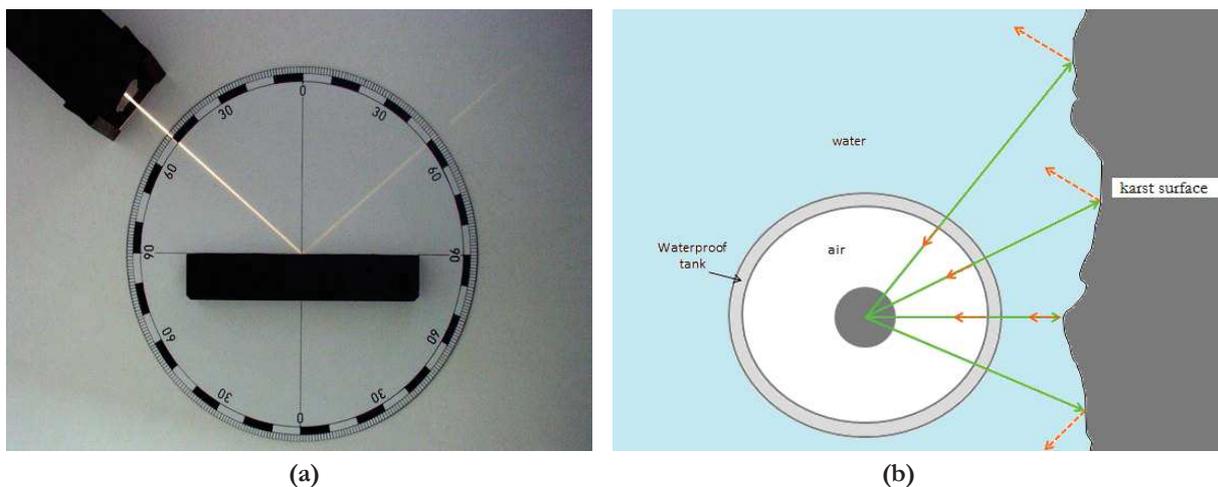


Figure III. 33: Illustration du phénomène de réflexion dans (a) l'air et (b) l'eau

7. Conclusion

L'objectif principal de notre application est de développer un système stéréoscopique offrant une vision la plus large de l'environnement aquatique et un processus permettant la reconstruction 3D d'un milieu de forme inconnue qui soit le plus robuste possible et aussi adaptables (système et méthode) pour pouvoir être utilisés en conditions sous-marines réelles. Plusieurs expérimentations ont donc eu lieu pour pouvoir tester le passage de la théorie avec des simulations à la pratique en conditions réelles. Des simulations dans des environnements aquatiques de dimensions connues ont permis en premier lieu de valider l'approche d'exploitation d'un étalonnage à sec pour application sous-l'eau et ensuite de mettre en place les expérimentations de test de l'algorithme de restitution de structure 3D de milieu aquatique. Les essais dans l'air, que ce soit dans la phase d'étalonnage ou de restitution 3D, ont été les plus nombreux car ceux dans l'eau nous ont nécessité beaucoup de logistiques qui sont liés en particulier avec la mise en place d'un système de protection pour assurer l'étanchéité du système de vision.

Chapitre IV.

Architectures et Implémentation matérielle

1. Introduction

Pour l'intégration des systèmes complexes dédiés au traitement d'images, multiples approches d'implantation sont proposées dans la littérature. Pour les applications fortement contraignantes, le choix d'un circuit dédié tel que un ASIC (Application-Specific Integrated Circuits) est recommandé étant donné que cette solution conduit aux meilleures performances en termes de consommation, de vitesse et de surface de silicium. Néanmoins, le coût de conception d'un ASIC est très élevé et les étapes d'optimisation sont longues à tous les niveaux de la conception. Contrairement à un circuit ASIC figé et qui n'est pas réutilisable, la flexibilité et la puissance de calcul des processeurs tels que les GPP (General Purpose Processor) représentent le grand intérêt de ces circuits. Néanmoins, la nature séquentielle de calcul de ces circuits et leur forte consommation fait que les processeurs ne sont pas adaptés aux applications embarquées temps réel de traitement d'image complexe. Une autre approche alternative réside dans l'utilisation des circuits reconfigurables FPGA (Field-Programmable Gate Arrays) qui permettent de répondre le mieux à la demande en capacité de calcul tout en assurant un haut niveau de flexibilité. En outre, la tendance actuelle est de remplacer les circuits intégrés à application spécifique ASIC et les processeurs DSP par des dispositifs FPGA de plus grande flexibilité même dans les applications complexes. En effet, les circuits FPGA étant reconfigurables, offrent une solution intermédiaire entre un circuit ASIC à performances élevées mais à fonctionnalité fixe, et les processeurs qui sont hautement flexibles mais pas très efficace (Fig.1).

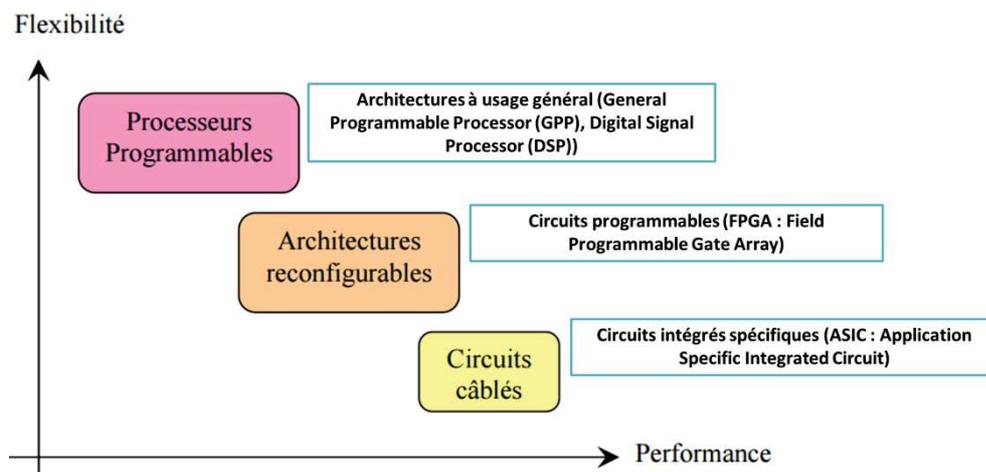


Figure IV. 1 : Comparaison entre les technologies d'implémentation

Dans notre travail, nous nous intéressons plus particulièrement aux traitements numériques des images en temps réel sur cible FPGA. Une architecture matérielle évolutive sera présentée faisant face à la complexité de calcul dans l'algorithme de traitement d'images que nous avons proposé. L'architecture sera synthétisée et convertie en un fichier physique prêt à être testé

sur un circuit FPGA. Les résultats obtenus seront comparés à l'implémentation logicielle en termes de précision et temps d'exécution. Notre objectif est l'amélioration du temps nécessaire obtenu par l'implémentation logicielle de référence, tout en gardant un taux de précision satisfaisant. On vise que l'architecture proposée soit considérée comme le point de départ pour le développement d'un système complet encore plus performant et qui peut atteindre les performances des systèmes de balayage à grande vitesse et temps-réel.

2. Environnement de développement et Plateforme FPGA

2.1. Environnement de développement

Pour un prototypage rapide et efficace, un environnement de développement dans lequel les modules logiciels et matériels peuvent être à la fois conçus, testés et validés est favorisé. A cette fin, l'outil Xilinx System Generator (XSG) pour le développement DSP (Digital Signal Processor) a été utilisé. C'est un outil de développement de haut niveau offert par Xilinx. Il est entièrement intégré dans l'environnement MATLAB/Simulink pour la conception à haute performance d'architecture synthétisable dédiée pour l'implémentation sur des circuits FPGA. Le concepteur peut profiter de la bibliothèque "Xilinx Blocksets" qui contient une large gamme de fonctions primitives et de fonctions avancées de traitement du signal développé par Xilinx. Ces blocksets lui permettent de développer ses propres algorithmes et bénéficier au même temps de l'environnement MATLAB lors de la phase de test pour la visualisation et l'analyse des résultats obtenus. System Generator permet, à travers l'environnement de synthèse ISE (Integrated Synthesis Environment), de convertir automatiquement l'architecture logique conçue en un "bitstream" qui sera chargé sur le dispositif FPGA à travers une connexion JTAG (Joint Test Action Group).

L'avantage principal de l'utilisation de l'outil XSG est qu'il intègre à la fois un simulateur logiciel et un simulateur matériel et permet ce qu'on appelle "la co_simulation". Durant la phase de co_simulation, le processus de traitement des images est implémenté sur le circuit alors que l'acquisition et l'affichage des données se fait à travers l'environnement Simulink (Fig.2). Aussi, la simulation sous System Generator permet de tester l'architecture développée avec les mêmes conditions que si elle fonctionne sur un circuit FPGA physique en utilisant la même largeur de bit (bit width) des variables et des opérateurs. La simulation est réalisée "bit à bit" et à cycle d'horloge précis, ce qui signifie que chaque bit est traité avec la même précision et chaque opération prend exactement le même nombre de cycles à traiter dans les deux cas. Ainsi, il n'y aura presque pas de surprises lorsque le système sera exécuté directement sur un circuit FPGA.

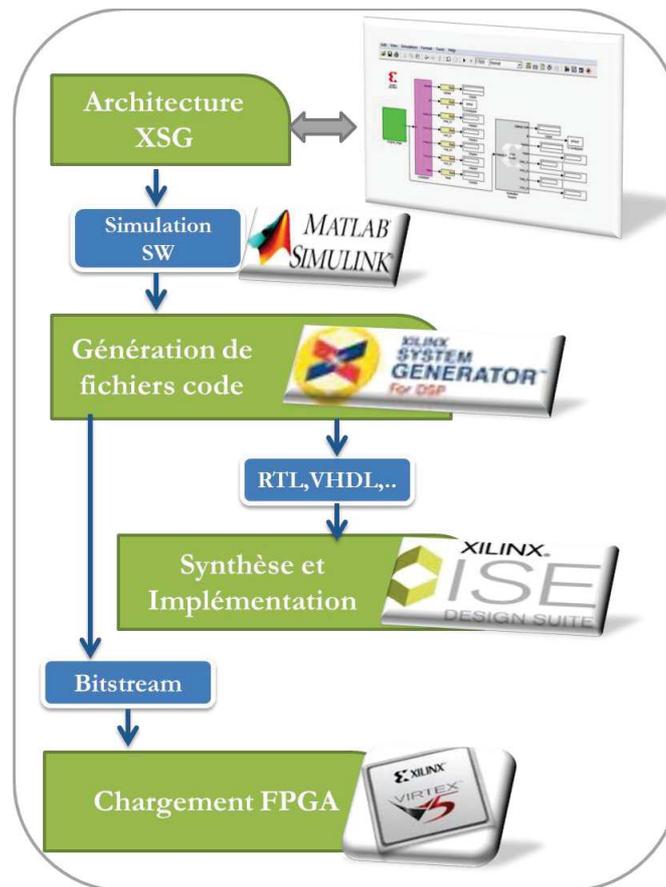


Figure IV. 2 : Principe de conception sous XSG

2.1.1. Flot de conception

Dans cette thèse, la démarche de conception suivie, présentée par la figure 3, commence par l'implémentation d'un algorithme en Matlab exécuté sur un processeur en vue d'obtenir des résultats logiciels à haute précision. Ensuite, nous développons une architecture matérielle traduisant le fonctionnement de notre système sous forme de blocs XSG codés en langage matériel (VHDL). L'objectif est de pouvoir profiter de la flexibilité et des ressources matérielles disponibles dans le FPGA afin d'optimiser les performances en termes de temps d'exécution jusqu'à atteindre celles d'un système embarqué temps réel. Le passage par un code Matlab, avant la conception de l'architecture, permet essentiellement de gagner en temps de prototypage. Il permet, en effet, de valider le fonctionnement de l'algorithme proposé en essayant d'atteindre un bon niveau de précision en optimisant à chaque fois les paramètres d'entrée. La validation de l'optimisation est effectuée en comparant entre les données métriques réelles, considérées comme métriques de référence, et les métriques estimées par l'algorithme exécuté sur un processeur software. De même pour l'architecture matérielle proposée, elle sera validée pour une implémentation sur FPGA dès l'obtention d'une marge d'erreur relativement acceptable. L'optimisation au niveau de l'architecture sera détaillée dans les sections qui suivent.

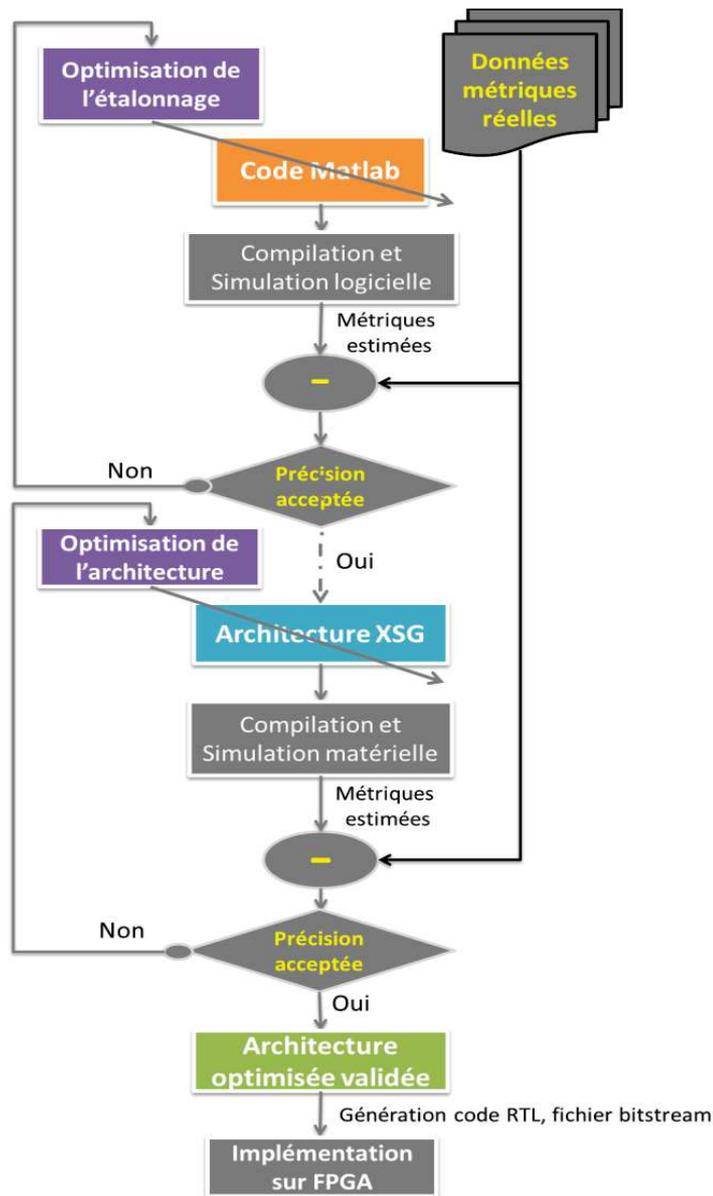


Figure IV. 3 : Approche de conception suivie sous Matlab/XSG

2.1.2. Exploration architecturale et optimisation

a. Extraction de parallélisme

L'extraction du parallélisme nous conduit à un nombre minimal de cycles d'exécution. Si nous analysons la dépendance de données présentes dans la suite d'instruction, nous pouvons voir qu'un certain nombre d'opérations peuvent être exécutées de façon parallèle. En analysant la séquence d'instructions et en respectant les contraintes de dépendance de données et d'accès à l'espace mémoire, nous obtenons une suite de traitements parallèles comme représentée pour l'exemple de graphe de la figure 4.

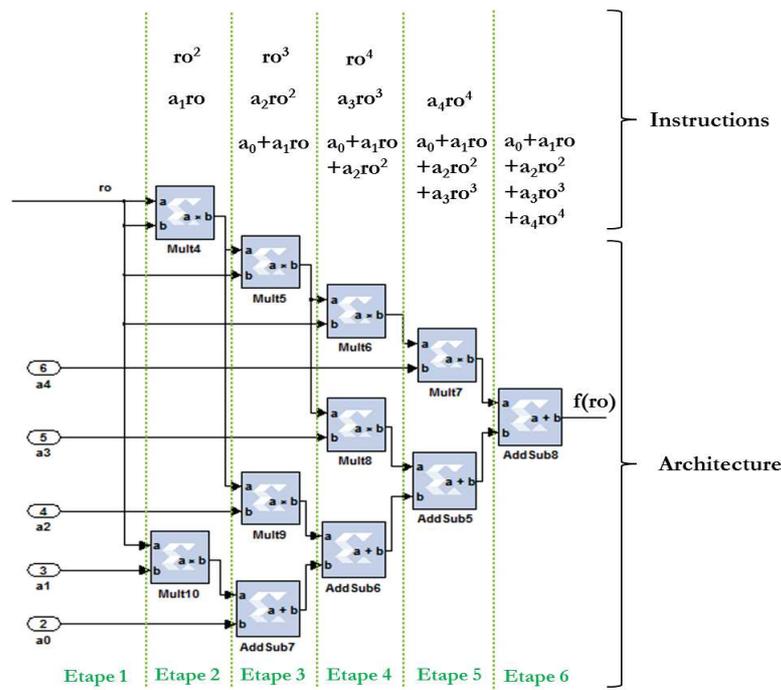


Figure IV. 4 : Exemple d'extraction de parallélisme au niveau des instructions

b. Optimisation algorithmique

L'objectif principal d'une conception est de pouvoir créer un meilleur compromis entre précision, temps d'exécution et ressources utilisés. Il est évident qu'il est impérativement impossible d'optimiser ces trois aspects en même temps car l'optimisation de l'un se fait toujours au détriment des autres. Dans notre cas, la précision nécessaire pour valider l'architecture est définie et fixée lors de la simulation logicielle. Elle est, en effet, vérifiée à chaque modification de l'architecture. Aussi, le temps d'exécution correspondant au temps de réponse du système de traitements est limité par un seuil maximal prédéfini en se basant sur les contraintes temporelles requises par l'application et la carte FPGA: la fréquence d'acquisition des images, la fréquence maximale d'horloge, etc.... Ayant ces deux métriques prédéfinies (précision et temps d'exécution), l'objectif de cette optimisation est de pouvoir profiter des primitives Xilinx et de proposer une alternative de quelques fonctions (instructions algorithmiques) qui seront moins coûteuse en ressources matérielles tout en essayant de garder la même précision.

En simulation, Matlab/Simulink utilise le format "double" pour représenter les signaux et les variables alors que System Generator ne manipule que des nombres à virgules fixes. La conversion du virgule flottant au virgule fixe est effectuée au niveau des blocs d'entrées spécifiques "Gateway-In". Le concepteur est amené à choisir la bonne configuration de ces blocs à l'entrée comme à l'intérieur de l'architecture de manière à garantir une bonne utilisation des ressources du FPGA à travers une meilleure représentation des données assurant une précision optimale. L'optimisation de la longueur du mot à l'entrée de chaque bloc matériel présente une étape clé qui permet d'utiliser efficacement les ressources du circuit FPGA. Néanmoins, pour les

algorithmes fonctionnant avec des nombres réels, il est impossible de représenter exactement les chiffres et des erreurs peuvent exister et peuvent être autorisés par le concepteur. L'optimisation en nombre de bits s'effectue pour chaque sous-système suivant deux étapes :

1- Première simulation en utilisant la configuration "full_output" : cette étape permet au concepteur d'acquérir des informations sur la progression des données en nombre de bits et d'avoir une estimation sur le résultat de sortie à précision maximale. Cette configuration ne peut pas être toujours adoptée car le design est limité par le nombre de ressources disponibles dans le circuit FPGA.

2- Plusieurs simulations sont effectuées en minimisant à chaque fois le nombre de bits utilisés et analysant l'effet de réduction sur la précision des résultats. La configuration optimale considérée sera celle qui offre la plus précise que celle obtenue avec une configuration "full_output".

Dans la figure 5, on présente un exemple de simulation d'un même module avant et après optimisation du nombre de bits utilisé. Les deux blocs présentent des valeurs proches de sortie mais à types de données différents. Le premier, étant configuré pour fonctionner à précision maximale, présente une sortie à 145 bits alors que celui optimisé présente une sortie à 32 bits. Dans le tableau 1, on donne les ressources utilisées et on remarque une optimisation importante en nombre de slices, de Luts et en Multiplicateurs/DSP48.

Tableau IV. 1 : Comparaison des ressources utilisées avant et après optimisation du bloc hardware

	Module avant optimisation	Module après optimisation
Slices	2391	226
FFs	0	11
LUTs	8316	771
Mults/DSP48	214	2

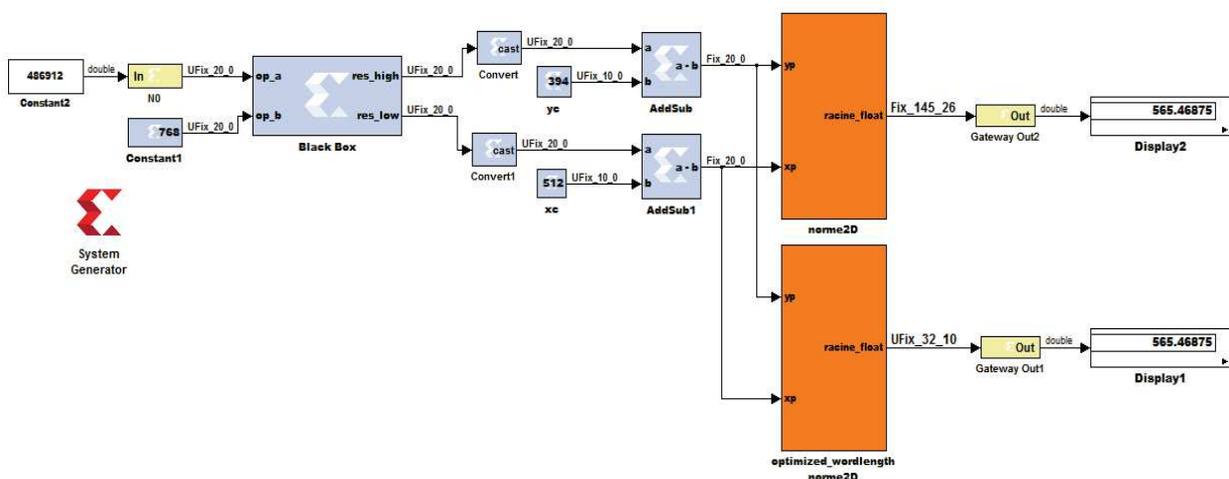


Figure IV. 5 : Comparaison entre le même module avant et après optimisation

2.2. Plateforme FPGA

L'évolution rapide des techniques de fabrication et des technologies ont permis de voir la naissance de FPGA de plus en plus performants et de plus en plus complexes. La plateforme de prototypage qui a été utilisée en conjonction avec MATLAB et Xilinx System Generator est une carte ML605 de la firme Xilinx comportant un circuit FPGA Virtex-6 XC6VLX 240T (Fig. 6).

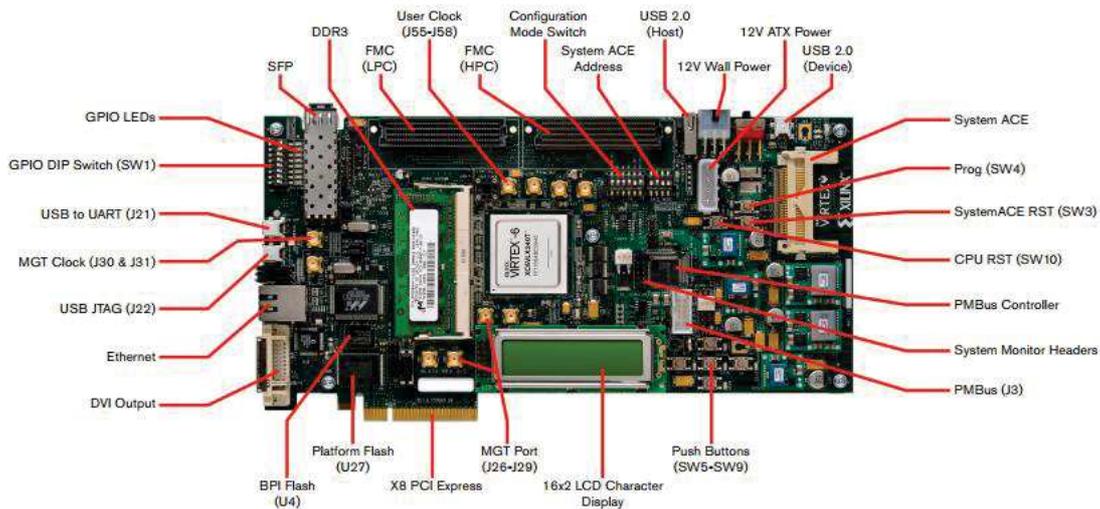


Figure IV. 6 : La carte de test à base de FPGA : plateforme ML605

La carte ML605 contient un oscillateur interne qui permet de générer une horloge de fréquence 100 MHz. Nous avons testé notre architecture XSG en utilisant de la co_simulation matérielle, c'est à dire en utilisant la communication JTAG pour effectuer les I/O et en testant avec le matériel en boucle.

Ce circuit FPGA XC6VLX 240T de la carte ML605 intègre plusieurs ressources dédiées (Fig.7.a):

- Des **CLB** (Configurable Logic Bloc) : Le circuit FPGA utilise le concept de blocs logiques configurables (CLB). Ces CLB permettent un routage et configuration optimaux de chaque bloc logique dans le FPGA. Dans la famille Virtex, un CLB est formé de 2 slices, comportant chacun un ensemble de 4 LUT à 6 entrées et 2 sorties, 2 bascules D et de la logique de contrôle (Fig.7.b). Les LUTs sont programmés pour exécuter la fonction logique requise. Xilinx indique souvent la capacité de ses circuits en Slices. Dans le cas du FPGA Virtex 6-XC6VLX 240T, celui-ci dispose de 37680 Slices. Les CLB sont disposés dans une matrice bidimensionnelle et des interconnexions programmables sont organisées horizontalement et verticalement entre eux. Sur les côtés de la matrice existent des blocs d'I/O qui sont configurés en tant que les entrées ou les sorties du circuit.
- Des **IOB** (Input Output Bloc) : ces blocs d'entrées et de sorties sont principalement composés de buffers qui permettent d'interfacer le FPGA avec l'environnement extérieur.

Certaines entrées sont plus particulièrement dédiées aux signaux d'horloge. Le Virtex utilisé sur notre carte dispose de 720 entrées-sorties.

- Des blocs de calcul adaptés au traitement du signal (**DSP48**): En plus des blocs CLB, les circuits Virtex-6 dispose de slices DSP. Ces blocs permettent d'effectuer différentes opérations arithmétiques (Fig.7.c) Il effectue la majorité des traitements puissants sur FPGA tels que : multiplication 25x18 bits, multiplication avec accumulation (MAC), additionneur à 3 bloc d'entrées 48 bits, registre à décalage 48 bits,...etc. Le circuit FPGA utilisée (Virtex 6-XC6VLX 240T) dispose de 768 blocs DSP48. Chaque DSP est placé à côté d'un bloc mémoire RAM (BRAM).

- Les ressources d'interconnexion des cellules : Ce sont des bus qui permettent de réaliser des fonctions complexes à partir des fonctions de base que représentent les CLBs.

- Des blocs mémoires (**BlocRAM** : block random access memory) : Pour pouvoir stocker une somme plus importante de variables sans avoir à accéder à des mémoires externes, certains fabricants de FPGAs ont introduit des blocs mémoires à l'intérieur des FPGAs. Aujourd'hui, tous les circuits Virtex comprennent des blocs de mémoires intégrées (BRAMs) qui représentent un module de mémoire destiné aux grands éléments de stockage. Elles peuvent être utilisées en simple ou double accès. Le bloc mémoire du Virtex 6-XC6VLX 240T dispose de 416 BlocRAM à 32Kbits chacun, ce qui donne une capacité totale de stockage interne de 14976 Kbits.

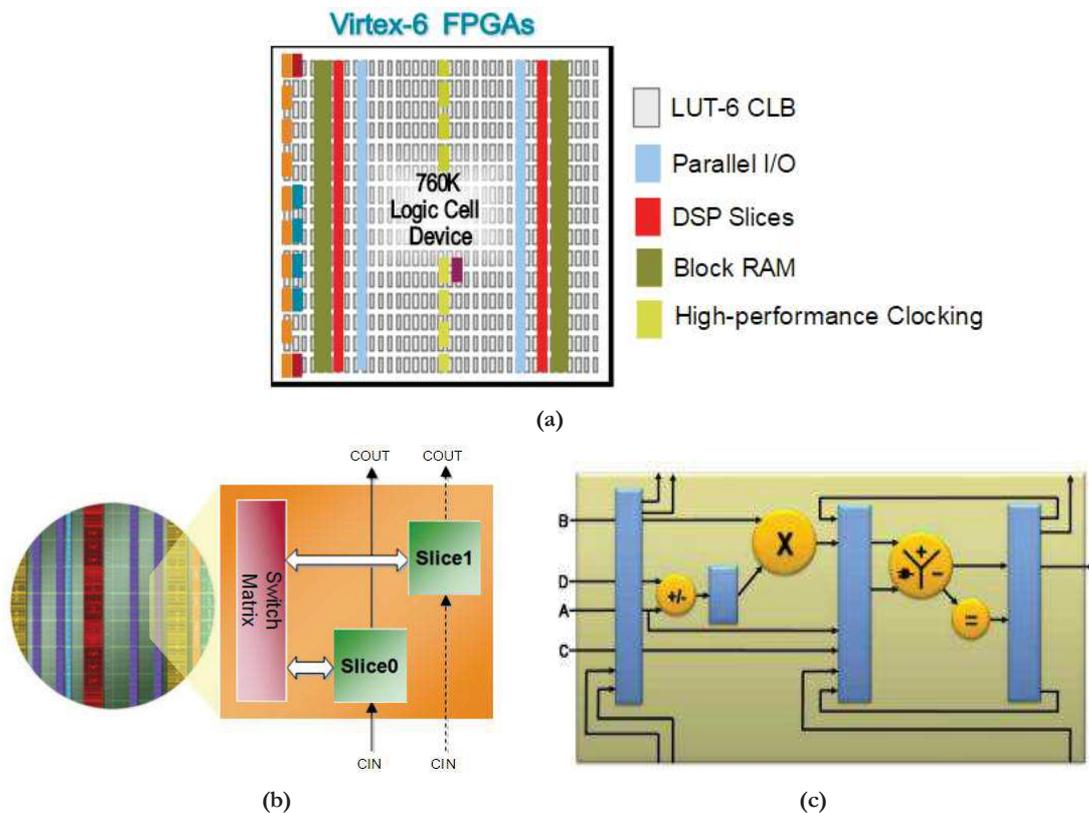


Figure IV. 7 : Structure interne d'un FPGA

3. Architecture matérielle proposée pour la reconstruction 3D

3.1. Bibliothèque Xilinx

L'architecture XSG est conçue en utilisant des éléments de bases, offerts par Xilinx et rassemblés dans la bibliothèque de Simulink. La section "Memory" contient des éléments de mémoire standards comme les mémoires RAMs, les ROMs, les FIFOs et les registres. Les bibliothèques de traitement des signaux et de communication contiennent des modules pour l'implémentation des algorithmes sophistiqués comme la transformée de Fourier, les filtres "FIR", etc.... La bibliothèque "Control Logic" contient des modules spécifiques plus complexes en particulier, le "Black Box", le "EDK processor", le bloc "MCode" et le "microcontrôleur MicroBlaze" permettant d'importer ou exporter des projets, des codes Matlab embarqué, des fichiers VHDL, ou des architectures pré-développées sous forme d'IP (Intellectual Property). Aussi, des primitives d'entrées/sorties sont utilisées pour passer du domaine continu en virgule flottante au domaine en temps discret à point fixe et qui peuvent être reconnectées aux blocs standards de Simulink. Les blocs d'entrées appelés "Gateway-In" permettent de fixer le type de donnée et sa précision ainsi que la fréquence à laquelle le signal sera échantillonné. En effet, vu que XSG fonctionne uniquement en temps discret, celui-ci introduit le principe de la période d'échantillonnage ("sampling period") : si par exemple la période d'échantillonnage d'un élément est fixée à 'm', la sortie au niveau de cet élément n'est calculée qu'après 'm' cycles d'horloge. Une autre propriété est aussi importante qui est la latence qui peut être spécifiée pour certains blocs. Si dans un circuit la valeur par défaut de la latence est gardée à zéro, le circuit est considéré totalement combinatoire ce qui signifie que le module implémenté sur FPGA génère une sortie en un seul cycle d'horloge.

3.2. Bibliothèque d'opérateurs personnalisés et optimisation

Comme indiqué dans le chapitre précédent, l'algorithme de reconstruction 3D utilise plusieurs opérations de division, racine carrée et des fonctions trigonométriques. Ces fonctions sont appliquées sur des nombres réels qui seront traités, par System Generator, comme des nombres à virgule fixe spécifiés et configurés par le concepteur. Xilinx offre, dans sa bibliothèque Simulink, des noyaux IP appelés des CORDICs (COordinate Rotation DIgital Computer) qui permettent d'approximer les fonctions de division, logarithme, sinus, cosinus, racine carré et l'inverse de la tangent à partir d'opérations élémentaires (additions, soustractions et multiplications). Ce module repose sur le calcul répétitif des coordonnées de vecteurs auxquels on applique une rotation bien choisie. L'avantage du CORDIC est qu'il permet d'obtenir une

précision déterminée à l'avance en effectuant un nombre d'itération donné. Afin d'obtenir une précision élevée, le concepteur peut modifier les configurations de ce module (la latence de chaque élément, le nombre de bits de l'entrée...) ou augmenter le nombre d'itération. L'obtention d'une bonne approximation peut demander parfois un long temps d'exécution et plusieurs ressources matérielles.

3.2.1. Fonctions personnalisées

Dans le cas de notre architecture, plusieurs expérimentations ont été effectuées pour juger la fonctionnalité du CORDIC à différentes configurations. Les résultats d'approximation de la fonction de racine carrée sur un nombre à 32 bits ont été plus précis que ceux sur un nombre à 8 bits. Les résultats obtenus pour la division étaient plus satisfaisants mais à une vitesse d'exécution importante à cause du nombre de cycles d'horloge demandé. Les deux fonctions demandent de nombreux ressources pour atteindre une bonne précision ce qui rend son utilisation moins avantageuse en particulier pour une application complexe qui nécessite déjà beaucoup de ressources FPGA. Par conséquent, une solution alternative utilisant des algorithmes d'approximation est définie pour chacune des fonctions citées ci-dessous:

a. Fonction de division

L'architecture qui permet d'approximer la fonction de la division est basée sur un code VHDL. Ce bloc, codé pour effectuer une division entre deux entiers positifs (à 20bits ou à 32bits), a été adapté pour pouvoir effectuer une division entre deux nombres réels quelconques. Ceci a été réalisé en utilisant des blocs élémentaires, essentiellement la multiplication. Dans la figure 8, on donne les étapes successivement citées avec un exemple concret. D'abord, le produit des signes des entrées réelles est sauvegardé avant l'opération de division. On multiplie ensuite les deux valeurs d'entrées par un facteur de 10^N qui permet d'intégrer N chiffres de la partie décimale à la partie entière. Le choix du nombre de bit des entrées est fixé selon la marge des données qui seront traitées. Le facteur de multiplication N est choisi aussi selon les données à l'entrée afin de garantir que le numérateur soit plus grand que le dénominateur et par conséquent assurer un résultat précis à la sortie du code VHDL. Le résultat obtenu est multiplié par le produit des signes sauvegardé et par le facteur de différence (fixée dans ce cas à 10^{-4}). Afin de valider cette architecture et le choix des facteurs utilisés, différents nombres réels et entiers à longueurs de bits différents ont été simulés et une sortie à erreur près (dans ce cas, elle est égale à 10^{-4}) est obtenue par rapport à une division complète.

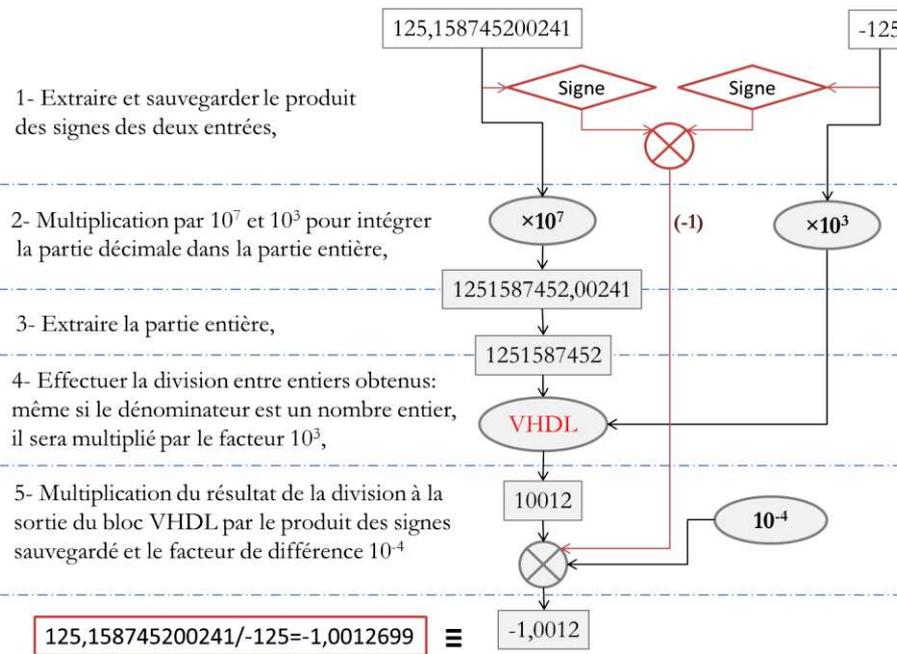


Figure IV. 8 : Organigramme de l'architecture proposée pour la fonction de division

b. Fonction Racine Carrée

L'algorithme qui permet d'estimer une valeur approchée de la racine carrée est donné par la figure 9. L'architecture combine entre l'utilisation d'un code VHDL et quelques blocs opératoires tels que la multiplication, l'addition, la soustraction et la comparaison. Le code VHDL permet de calculer la racine carrée d'un entier à 20, 32 ou 64 bits. Il permet, en effet, de donner une première approximation de la racine carrée d'un nombre réel. Un module logique est conçu pour rapprocher progressivement le résultat estimé à chaque itération de la valeur exacte de la racine carrée à erreur près.

L'idée dans le processus d'approximation est de limiter, d'abord, la racine réelle (\sqrt{A}) entre la racine de sa partie entière (B) et la racine de la partie entière incrémentée par 1 ($C=B+1$). A chaque itération, on calcule les variables $D1$ et $D2$ qui représentent la différence entre les racines afin de déterminer à quelle valeur la racine réelle est la plus proche. Si l'erreur d'estimation E , qui représente le minimum des deux différences, est inférieure à une valeur de seuil fixe (0.01, dans ce cas), le résultat final d'approximation sera celui correspondant à la différence minimale (B si $D1 < D2$ et C sinon). Si l'erreur d'estimation n'est pas acceptée, une autre itération est effectuée. Dans chaque nouvelle itération, la nouvelle valeur de C est la valeur moyenne entre les deux anciennes entrées (B et C) et la nouvelle valeur de B est celle de la plus proche estimation (B ou C). La sortie finale de la boucle peut avoir lieu en fixant soit le nombre d'itération ou le seuil d'erreur de manière à avoir une approximation suffisamment précise.

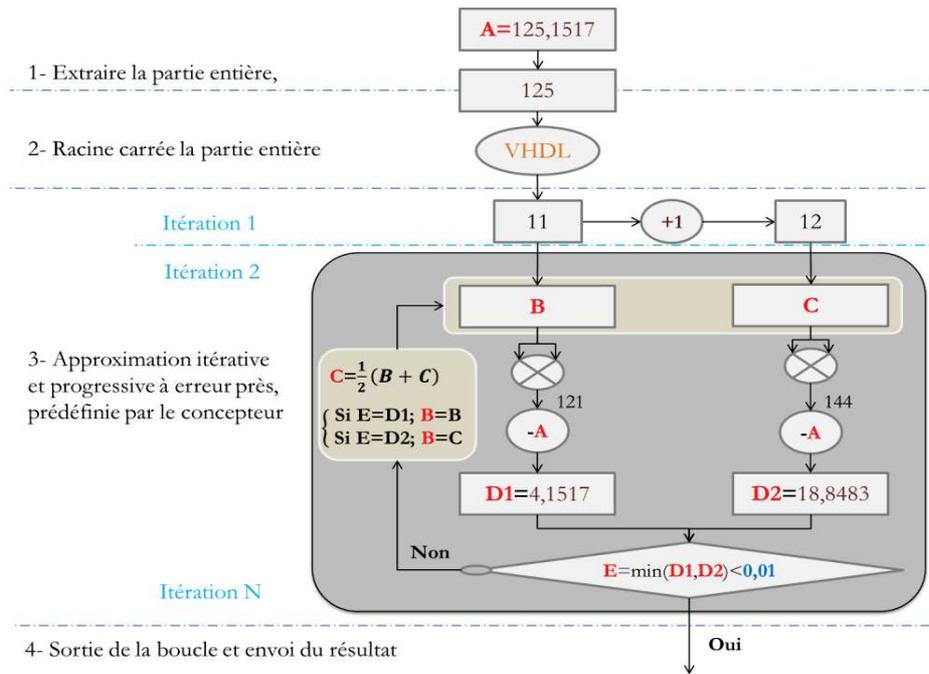


Figure IV. 9 : Organigramme de l'architecture proposée pour la fonction de la racine carrée

c. Fonctions trigonométriques

En ce qui concerne l'approximation des fonctions trigonométriques, le choix a été opté pour l'utilisation du développement limité de Taylor étant donné que les angles de réfractions sont toutes dans un voisinage de 0 ce qui coïncide avec les conditions d'exploitation de cette méthode. L'ordre du polynôme d'approximation est choisi selon les exigences en taux d'erreur. L'architecture (Fig.10) conçue utilise des primitives de calcul simples telles que l'addition, la soustraction et la multiplication. Les coefficients des polynômes sont fixés d'avance et exploités en tant que constantes pour éviter l'utilisation de la fonction de division.

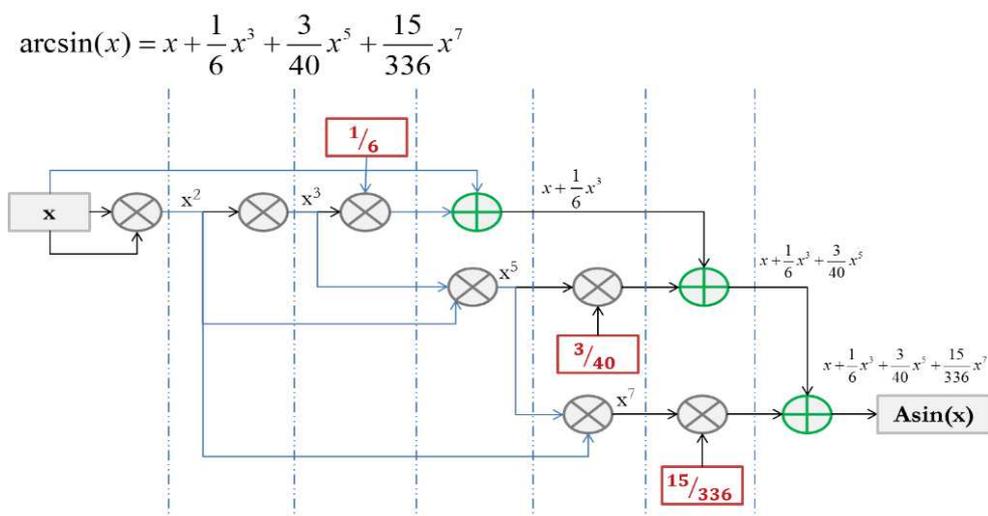


Figure IV. 10 : Organigramme de l'architecture proposée pour l'approximation de la fonction ArcSinus

3.2.2. Comparaison entre le CORDIC et les algorithmes proposés après optimisation

Une évaluation a été effectuée en termes de ressources matérielles utilisées et fréquence maximale des fonctions citées tout en gardant une bonne précision par rapport à un résultat logiciel. Le tableau suivant donne une comparaison de performances entre les opérateurs CORDIC et NON CORDIC proposés pour un même résultat de sortie (Tableau 2).

Tableau IV. 2 : Comparaison de performances entre CORDIC et le bloc XSG proposé

	Sqrt 20bit CORDIC	Sqrt 20bit Non CORDIC	Division 32bit CORDIC	Division 32bit Non CORDIC	Atan 32bit CORDIC	Atan 32bit Non CORDIC
Slices	494	226	867	779	118	100
FFs	1547	11	3111	0	392	0
LUTs	1675	771	2955	2489	431	389
Mults/DSP48	4	2	6	0	0	15
Période (ns)	4.365	1.408	2.693	0.843	1.592	1.080
Fréquence (MHz)	229.095	710.227	371.315	1186.24	628.220	925.314

3.3. Architecture proposée en XSG

Le processus de traitement illustré par la figure 11 présente la chaîne complète de traitement : depuis la lecture simultanée des images couleurs stéréoscopiques jusqu'à la génération des coordonnées 3D donnant la forme du conduit à la position d'acquisition.

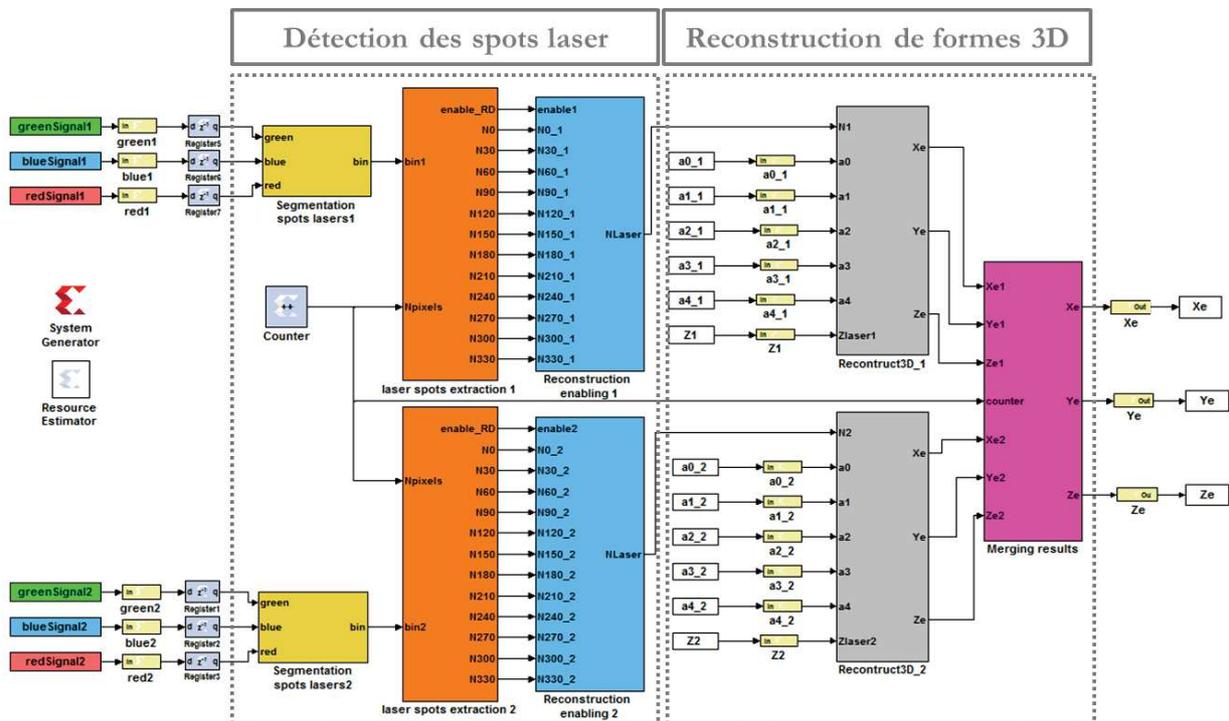


Figure IV. 11 : Processus de détection de points laser et de reconstruction 3D

Ce processus est subdivisé en deux principaux étages : extraction des points laser et reconstruction 3D. Le premier module contient la segmentation en images binaires des images omnidirectionnelles RVB lues à partir du système stéréoscopique pour mettre en valeur les points d'intérêt qui sont essentiellement des projections des faisceaux laser, dont les coordonnées 2D

sont extraites en se basant sur des contraintes géométriques. Le second module est consacré au passage 2D/3D, par application de relations géométriques spécifiques et en utilisant les paramètres d'étalonnage de notre système de vision.

3.4. Module de détection et extraction des spots laser

L'algorithme de détection des spots laser consiste à mettre en évidence les projections des faisceaux laser sur les images stéréo RVB pour faciliter l'extraction des coordonnées 2D correspondants. La première étape est la binarisation des images stéréo en se basant sur l'information couleur des pixels. La segmentation est effectuée sur des images converties en premier lieu à l'espace HSV. Les relations de segmentation, présentées dans le chapitre précédent, permettent de mettre en valeur les projections laser sur les images. Les pixels blancs représentent les projections des lasers et ceux en noir, l'arrière-plan. L'architecture matérielle traduisant ces relations est présentée par la figure 12. La recherche des valeurs minimale et maximale des composantes RVB a été effectué au niveau d'un bloc "MCode" intégrant un code Matlab de comparaison. Pour effectuer les opérations de division, on a profité du code VHDL, appelé par le "Black-Box", programmé pour réaliser des divisions entières. Des éléments de comparaison, mathématique et logique ont été utilisés pour traduire les expressions de conversion et de segmentation sous forme de blocs matériels. Cette architecture génère en sortie une image binaire mettant en valeur les projections des faisceaux lasers.

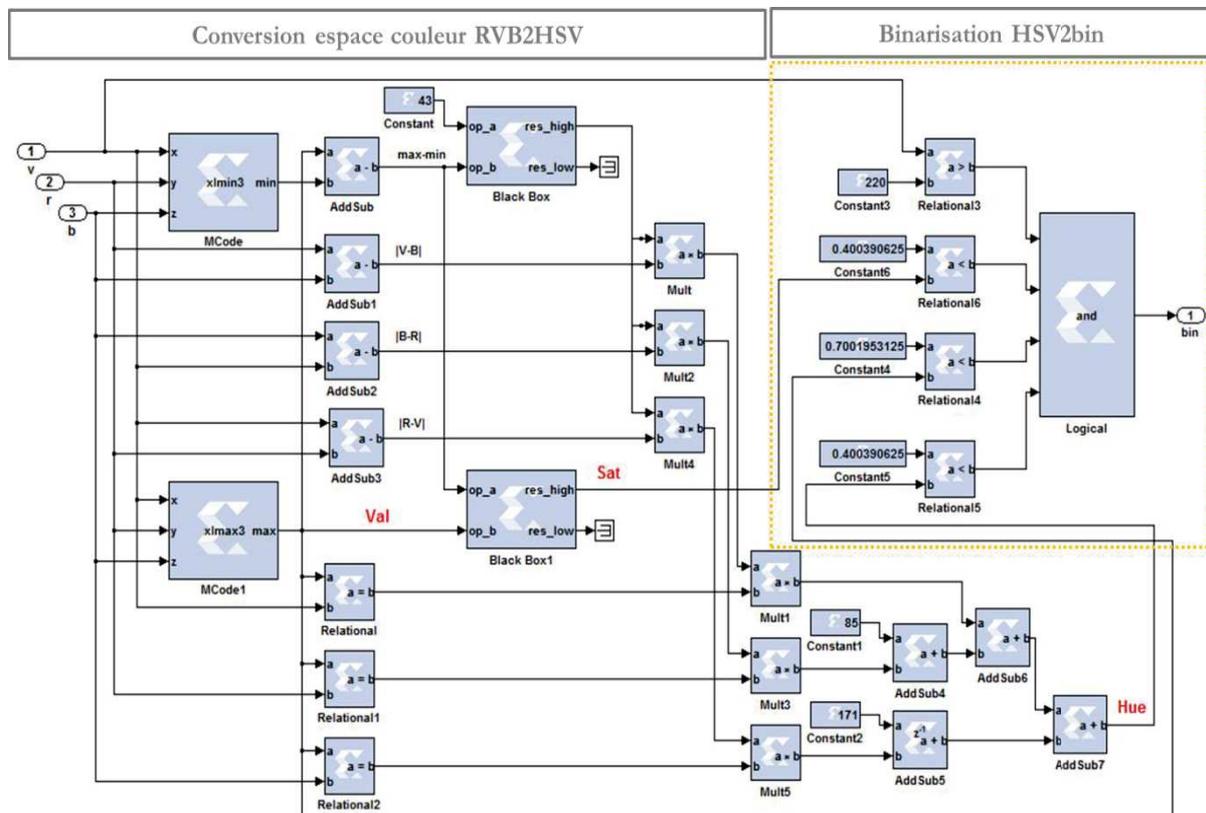


Figure IV. 12 : Module de segmentation des images RVB converties en HSV

Le résultat de segmentation présente des points d'intérêt utiles et quelques fausses détections dues essentiellement aux phénomènes de réflexion lors de la projection des faisceaux laser sur des surfaces réfléchissantes. Ces fausses détections doivent être filtrées pour ne pas erroner la restitution de la forme reconstruite du conduit. Une élimination de ces projections est réalisée en se basant sur les contraintes_régions, comme définies dans le chapitre 3. La figure 13.a présente les trois modules matériels réalisant les étapes de filtrage: changement de repère, application de la contrainte région et extraction des points saillants utiles selon leurs régions de projections prédéfinies. L'architecture en blocs Xilinx détaillée dans la figure 13.b, présente la détermination des coordonnées cartésiennes centralisées $p(x, y)$ par rapport au point principal estimé lors de l'étalonnage ($x_c=518, y_c=394$) et la détermination de la composante polaire notée "rayon". L'architecture XSG exploite le code VHDL permettant de réaliser la division entière pour retrouver la position en coordonnées cartésiennes des spots laser et un autre bloc programmé pour déterminer le rayon polaire qui correspond à la distance du spot par rapport au point principal.

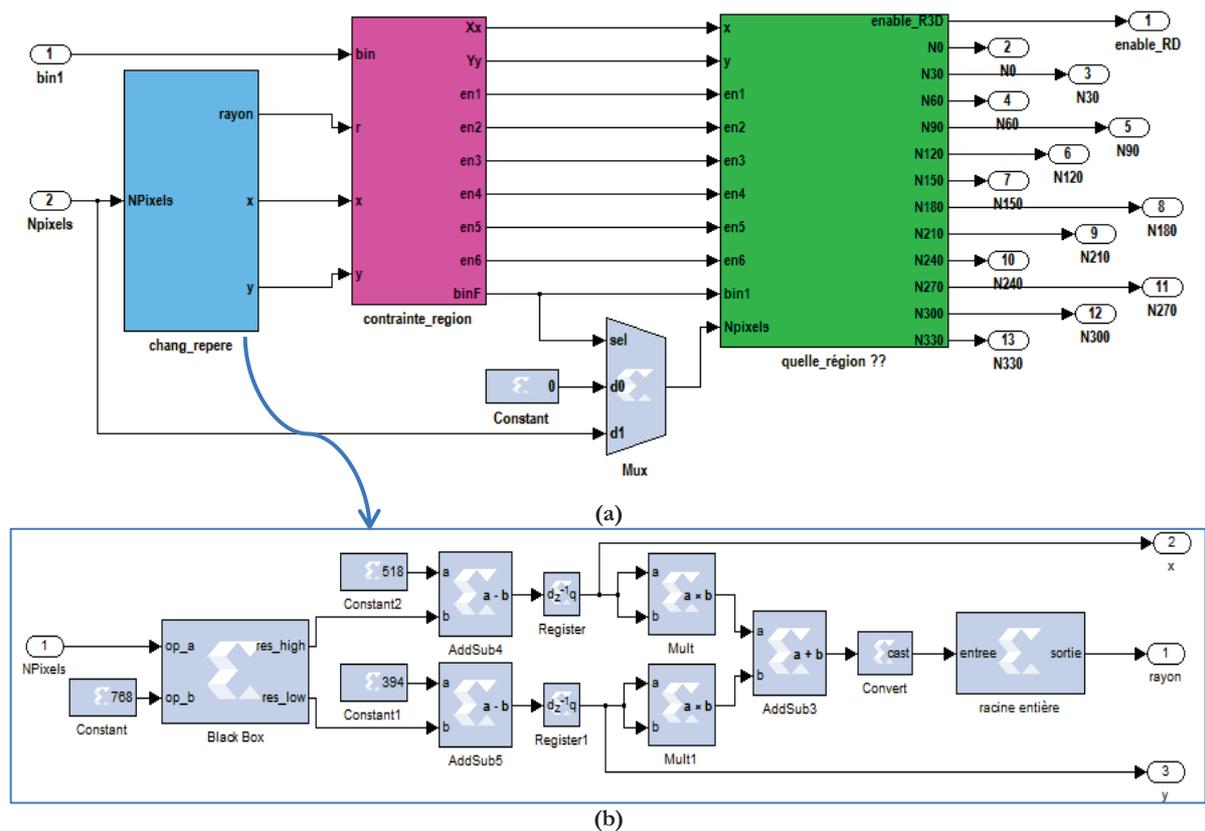


Figure IV. 13 : Module d'extraction des coordonnées des points lasers

Dans la figure 14, on présente l'architecture réalisée pour l'élimination des projections dues aux réflexions imprévues. Un algorithme de sélection, basé sur les valeurs des coordonnées polaires (r, θ) , est appliqué pour extraire les projections utiles et éliminer celles erronées. L'angle de projection " θ " est calculé à travers les coordonnées en pixels (x, y) . La fonction

d'appartenance, prédéfinie dans le chapitre III et illustrée par la figure 19, permet de sélectionner les nuage de points laser ayant comme angle d'inclinaison appartenant à l'ensemble $\{0, 30, 60, \dots, 330\}$ correspondant aux tangentes plus ou moins proche de $\{\pm 0.577, \pm 1.732\}$, dont la valeur du pixel du point sélectionné est à "1" logique ("bin"="1") et appartenant à la zone de projection du miroir définie par $\{r < r_{\text{miroir}}\}$. Ce point de coordonnées (x, y) est considéré comme point saillant utile. Le résultat de ce traitement est une image binaire composée de nuages de points présentant la projection directe des faisceaux laser.

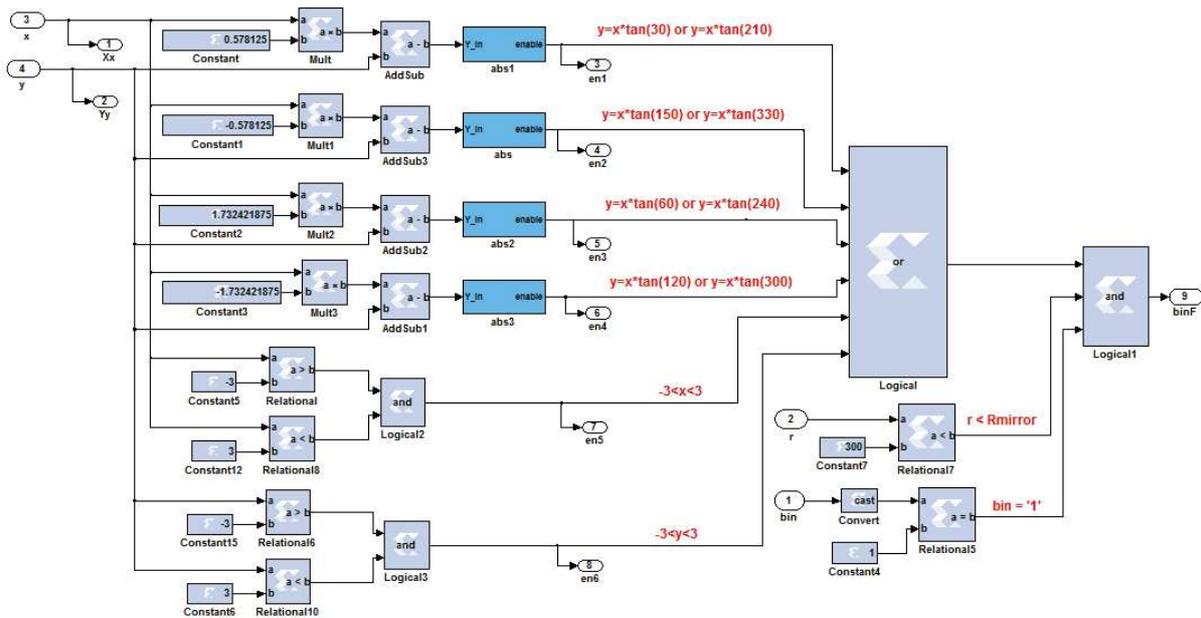


Figure IV. 14 : Sous-système 2: Elimination des fausses projections

Les coordonnées des pixels retenus sont ensuite réparties selon la région de projection, c'est-à-dire l'angle d'inclinaison correspondant, pour pouvoir classifier les projections laser et après extraire le centre de gravité correspondant. Douze blocs matériels sont conçus pour s'exécuter en parallèle définissant chacun les caractéristiques de chaque région. Chaque sous-bloc permet d'accumuler le nuage de points appartenant à une même région pour déterminer le centre de gravité lui représentant (Fig.15.a). Le centre de gravité trouvé pour chaque région sera considéré comme point de départ pour la reconstruction des coordonnées 3D du laser dans le repère monde. Les douze sorties sont simultanément envoyées au bloc suivant. Dans la figure 15.b, on présente un exemple de sous-système d'extraction du centre de gravité du nuage de points saillants qui se trouvent dans le voisinage droite des axes des abscisses (" $x > 0$ ") et qui présente la projection du faisceau laser se trouvant sur l'horizontale et ayant un angle d'inclinaison " $\theta = 0$ ". Un bloc matériel appelé "MCode" est utilisé pour faire le tri des valeurs de positions sélectionnées et vérifiant (" $en6 = '1'$ ") afin de déterminer les valeurs minimales et maximales. Celles-ci sont sauvegardées à chaque mise à jour dans des "registers". La sortie "N0" représente la position du centre de gravité du spot projeté à " $\theta = 0$ ".

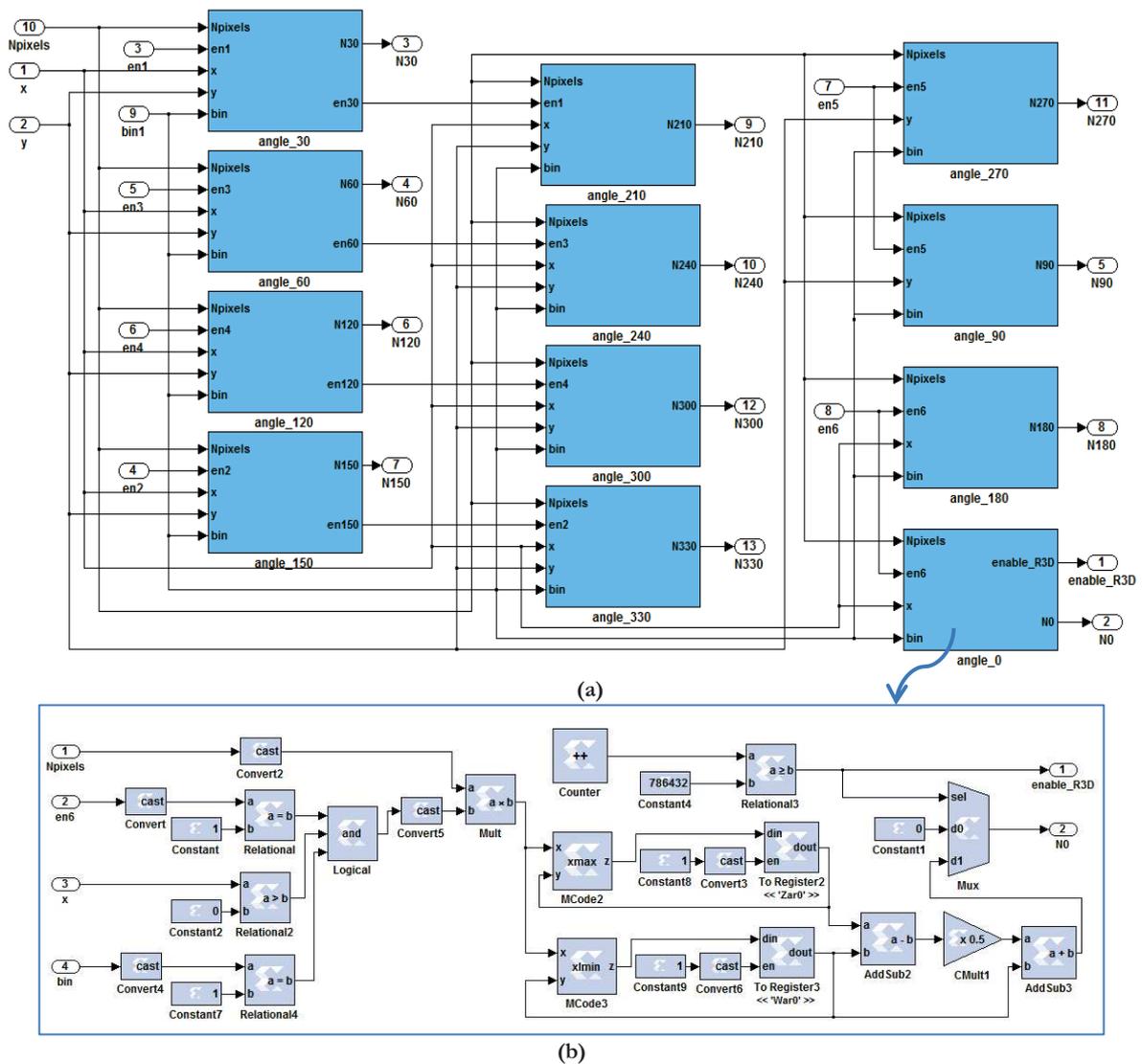


Figure IV. 15 : Sous-système 3: Extraction parallèle des coordonnées des spots laser

3.5. Module de reconstruction 3D depuis les projections images 2D

L'architecture donnée dans la figure 11 présente un module (module bleu) permettant de suivre la phase de détection des spots laser et d'enclencher la reconstruction des points 3D à partir des coordonnées 2D. Cette fonction commence dès que tous les spots sont détectés, c'est-à-dire dès que tous les pixels de la zone prédéfinie de projection sont parcourus. En outre, vu que les deux phases de détection et de reconstruction sont complètement indépendantes l'une de l'autre, on peut paralléliser les deux étages : lorsque le deuxième commence à traiter les données entrantes, le premier traite la nouvelle paire d'images. Etant donné que le nombre de cycles d'horloge demandé par la phase d'extraction de spots est plus important que celui de reconstruction, ce dernier est conçu de manière à traiter les données entrantes (les coordonnées des projections laser) en boucle en utilisant l'architecture donnée dans la figure 16.

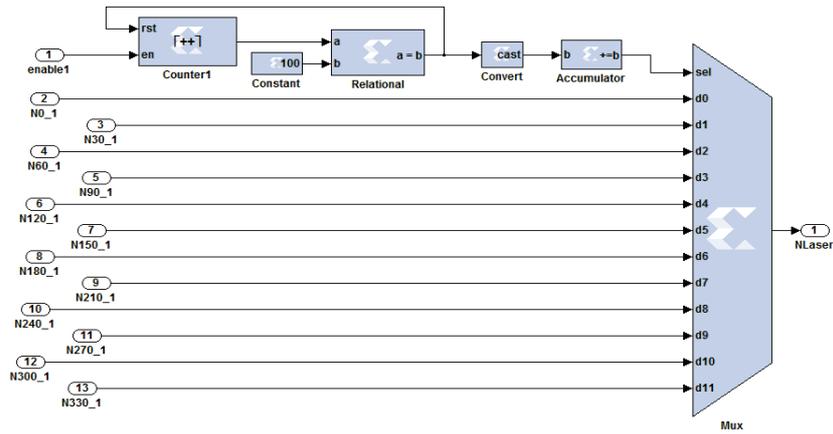


Figure IV. 16 : Module de passage entre les deux étages : Détection et Reconstruction

Le second étage du processus complet de traitement est celui de la reconstruction 3D. Le même bloc hardware est parcourue douze fois (nombre de lasers utilisés) pour la génération successive des coordonnées 3D des points lasers à partir de leurs projections 2D extraits des images stéréoscopique. Dans ce qui suit, on présente en détails la composition de l'architecture matérielle qui traduit l'algorithme de passage 2D/3D. La figure 17 présente le module XSG de la reconstruction 3D : il prend comme entrée la projection du faisceau laser sur l'image ainsi que les paramètres de modélisation du système de vision estimées lors de la phase d'étalonnage pour générer la mesure de distance ainsi que les coordonnées 3D de chaque point correspondant.

Le processus est subdivisé en quatre sous-blocs permettant chacun de réaliser une fonction donnée qui sont successivement :

- Module de détermination de la fonction d'approximation "Scaramuzza",
- Module de normalisation des rayons vecteurs correspondants "Normalisation",
- Module de détermination des angles de réfractions par la méthode de Snell "refraction_angle",
- Module de détermination des coordonnées 3D du point réel "Distance measurement".

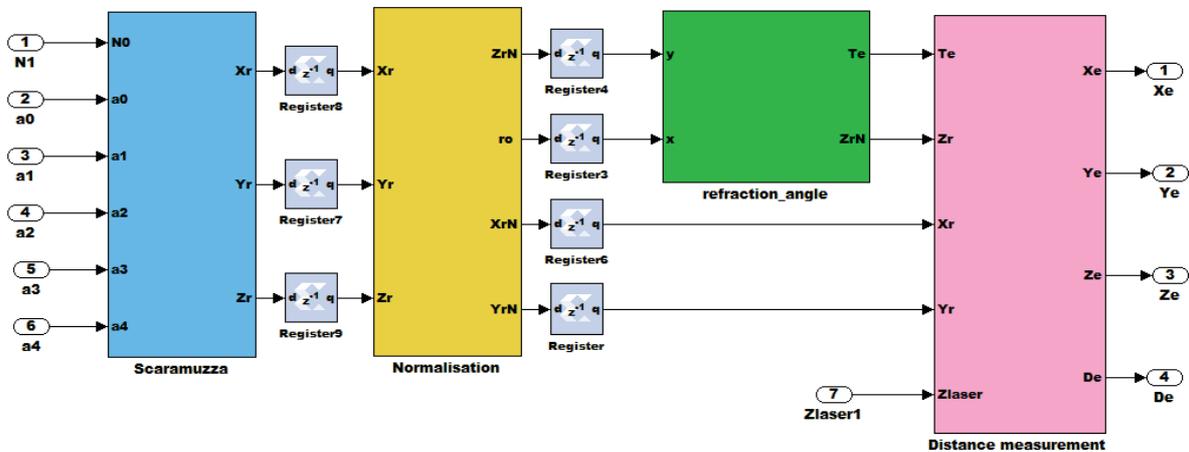


Figure IV. 17 : Module de reconstruction 3D

3.5.1. Module de détermination de la fonction d'approximation "Scaramuzza" :

Dans la phase d'étalonnage, l'estimation des paramètres est réalisée en utilisant l'approche proposée par Scaramuzza. La fonction d'approximation est appliquée sur la composante " ρ ", déterminée à partir des coordonnées pixelliques centralisées (x_p, y_p) par rapport au point principal (u_0, v_0), comme donnée par les expressions suivantes (Exp.1).

$$\vec{r} = \begin{bmatrix} X_r = x_p = u - u_0 \\ Y_r = y_p = v - v_0 \\ Z_r = f(\rho) \end{bmatrix}; \rho = \sqrt{(x_p^2 + y_p^2)}; f(\rho) = a_0 + a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 \quad (1)$$

La figure 18 présente les blocs matériels traduisant ces expressions. Les coordonnées en pixels sont obtenues en divisant la position "N0" dans l'image par le nombre de ses lignes (L=768, dans notre cas). Pour effectuer la division entière, on utilise le même block matériel "Black Box" défini précédemment permettant de diviser entre deux nombres entiers de 20 bits. Il génère en sortie un quotient noté "y" représentant l'ordonnée du pixel dans l'image de résolution 1024×768 et le reste "x" représentant son abscisse. La fonction d'approximation s'applique sur la composante " ρ " qui est égale à la norme 2D des coordonnées centralisés tels que $x_p = u - u_0$ et $y_p = v - v_0$ sachant que les valeurs de ($u_0=518, v_0=394$) ont été obtenues lors de l'étalonnage.

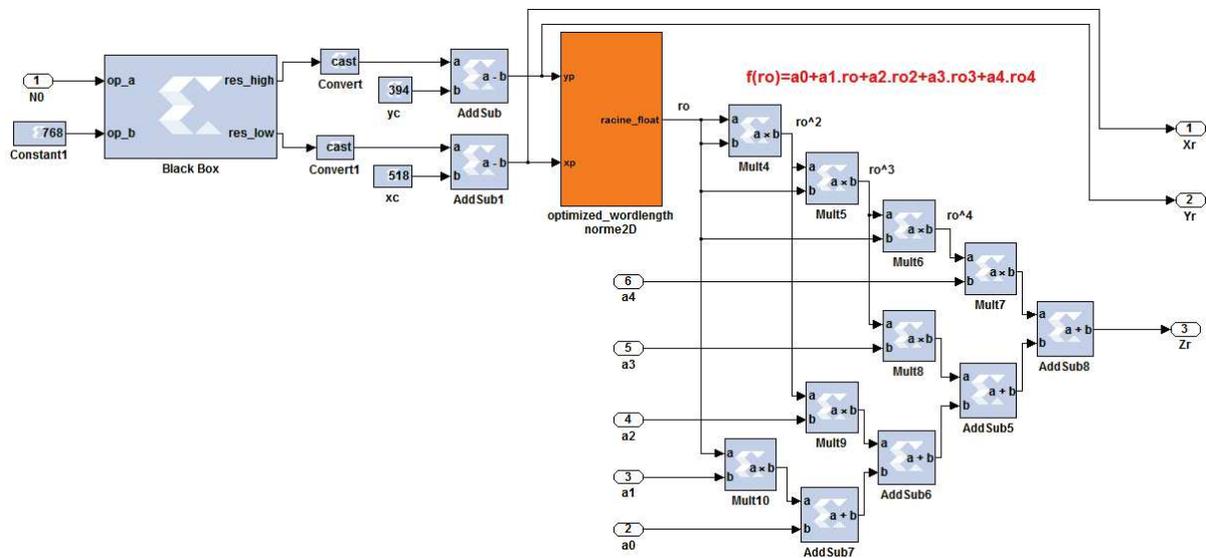
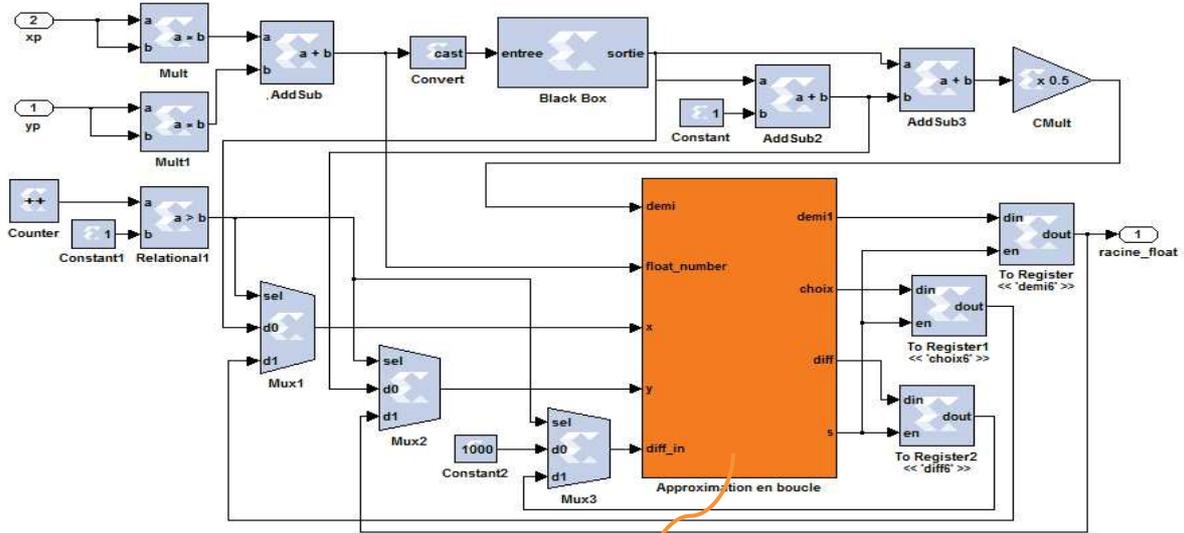


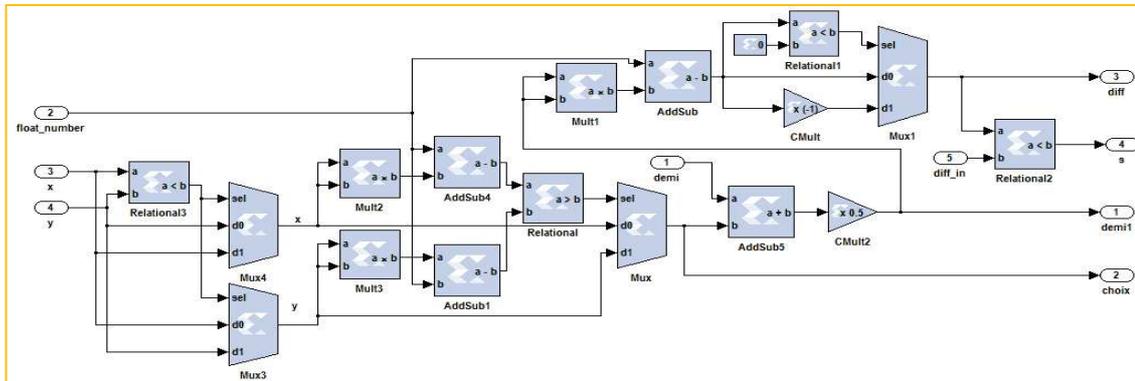
Figure IV. 18 : Module de Scaramuzza

Le principe de détermination de racine carrée d'un nombre réel est donné dans la figure 8. La figure 19.a présente l'architecture matérielle du module XSG permettant l'approximation de la racine carrée d'un nombre réel. Un blockset dénoté "convert" permet d'extraire la partie entière du nombre réel dont la racine carrée est déterminée en utilisant un code VHDL, testé et synthétisé sous ISE, appelé via le "Black box". Des itérations successives sont effectuées au sein

d'une architecture conditionnée à répéter le même processus quand la précision adéquate n'est pas atteinte. A chaque itération, la différence notée par "diff" entre la valeur réelle et la valeur estimée est comparée par rapport à sa précédente. Tant que l'erreur est en cours de diminution, une nouvelle itération est lancée et les entrées sont mises à jour. Des registres sont utilisés comme mémoire pour sauvegarder les dernières données. La figure 19.b présente le module XSG du processus d'approximation itérative.



(a)



(b)

Figure IV. 19 : Module d'approximation de la racine carrée d'un nombre réel

3.5.2. Module de normalisation des rayons vecteurs correspondants "Normalisation" :

Les coordonnées 3D obtenues du vecteur rayon sont ensuite normalisées en les divisant par leur norme comme donné dans l'expression (Eq.2).

$$\begin{bmatrix} X_r^N \\ Y_r^N \\ Z_r^N \end{bmatrix} = \frac{1}{\sqrt{X_r^2 + Y_r^2 + Z_r^2}} \begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix} \quad (2)$$

De même que les coordonnées 2D en pixels, la norme des coordonnées 3D obtenues est approximée en utilisant une architecture d'approximation proposée. Un sous-système est conçu pour l'approximation de la racine carrée d'un nombre réel dont la taille de la partie entière est de 20 bits ou 32 bits, selon le nombre de bits des entrées. Ce sous-système est utilisé d'abord pour donner la valeur de la composante " ρ " et ensuite pour déduire la norme des coordonnées 3D du rayon vecteur $r(X_r, Y_r, Z_r)$ correspondant aux coordonnées 2D. Le module de normalisation est présenté dans la figure 20. Des coefficients de multiplication sont utilisés pour accompagner le bloc de division. La racine carrée de $XN^2 + YN^2$ est enfin approximée avec le même processus détaillé précédemment pour être exploitée dans le bloc qui suit.

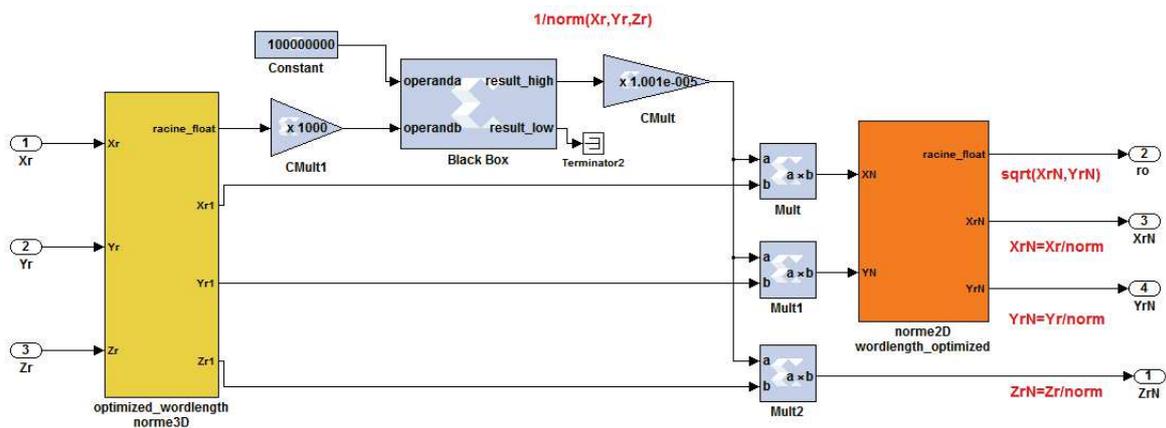


Figure IV. 20 : Module de normalisation

3.5.3. Module de détermination des angles de réfractions "refraction_angle"

L'angle de réfraction " θ_1 " est d'abord estimé à travers les coordonnées du vecteur rayon vu qu'il présente l'inclinaison de la projection du faisceau laser par rapport à la normale à la surface du plexi. L'architecture matérielle permettant de déterminer la valeur de l'arc-tangente dénoté " T_e " est donnée dans les figures suivantes :

➤ Dans le cas d'utilisation des paramètres d'étalonnage dans l'eau, cette variable sera utilisée par la suite pour estimer la mesure métrique de distance et les coordonnées 3D du point Monde illuminé par le laser (Fig. 21).

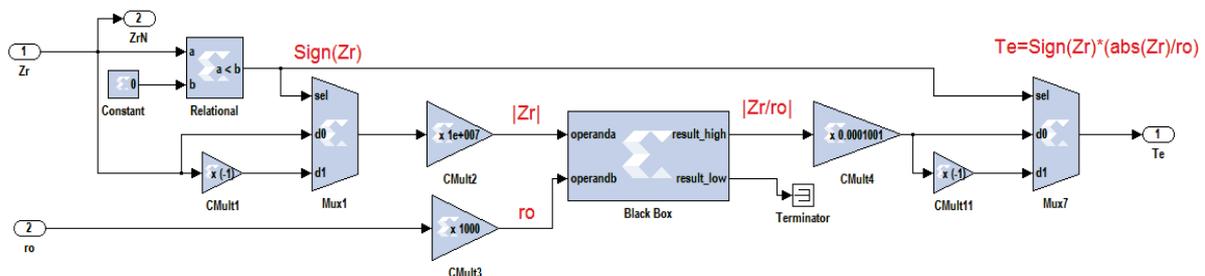


Figure IV. 21 : Module de détermination de l'angle de réfraction entre la surface du plexi et l'eau

➤ Dans le cas d'utilisation des paramètres d'étalonnage dans l'air, on a besoin de déterminer les angles de réfraction à chaque surface séparatrice entre les trois milieux (air/plexi/eau) en utilisant, comme défini dans le chapitre 3, les lois optiques de Snell (Fig.22).

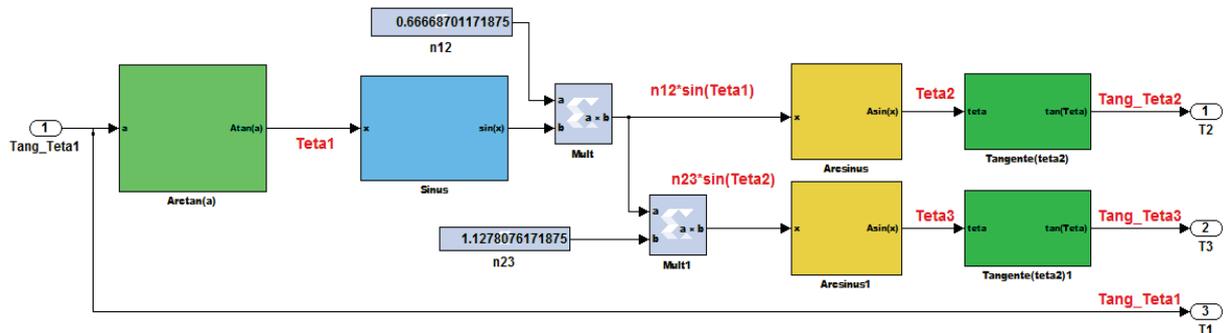


Figure IV. 22 : Module de détermination des angles de réfraction à chaque surface séparatrice

3.5.4. Module de détermination des coordonnées 3D du point réel "Distance measurement" :

Après avoir déterminé toutes les composantes (rayon vecteur traversant l'air et les angles de réfraction sur les surfaces intérieurs et extérieurs du plexi), on présente l'architecture conçue avec les blocs Xilinx pour la détermination des coordonnées 3D du point Monde.

➤ Dans le cas où on utilise les paramètres d'étalonnage dans l'eau, le passage est plus simple. La figure 23 donne l'architecture matérielle des équations données dans le chapitre 3. Elle exploite deux modules de division entre nombre réels et un module d'estimation de racine carré réelle. Des blocs primitifs sont aussi utilisés pour la détermination de signe et des multiplications.

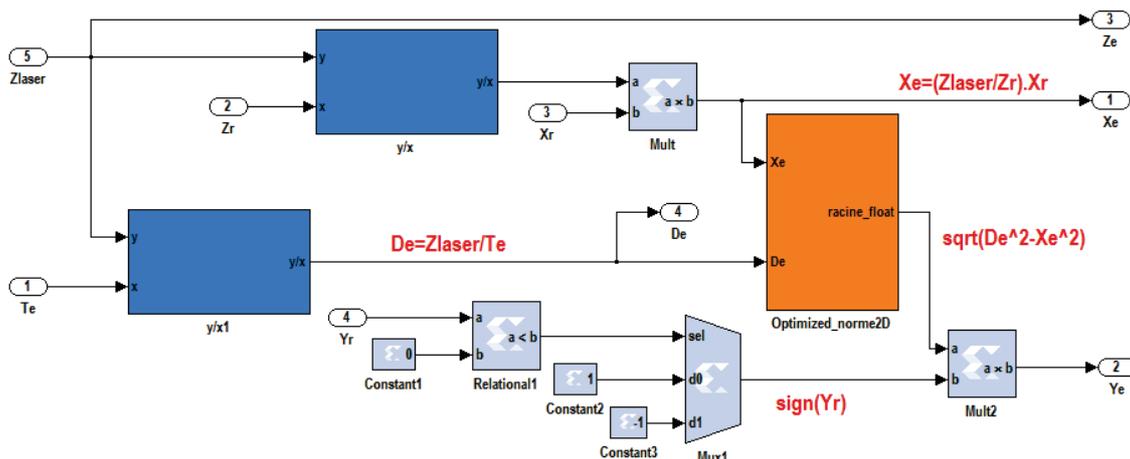


Figure IV. 23 : Module de détermination des coordonnées 3D avec étalonnage dans l'eau

➤ Dans le cas où on utilise les paramètres d'étalonnage dans l'air, le passage aux coordonnées 3D est plus complexe. La figure 24 donne l'architecture matérielle des équations (18-19) données dans le chapitre III. L'architecture comprend deux sous-modules : le premier permet la détermination du décalage "ep" et de la position virtuelle du faisceau laser dénoté "Zv" (Fig.24.a)

et le deuxième permet la restitution des coordonnées 3D du point Monde à travers la position virtuelle estimée pour chaque faisceau laser (Fig.24.b).

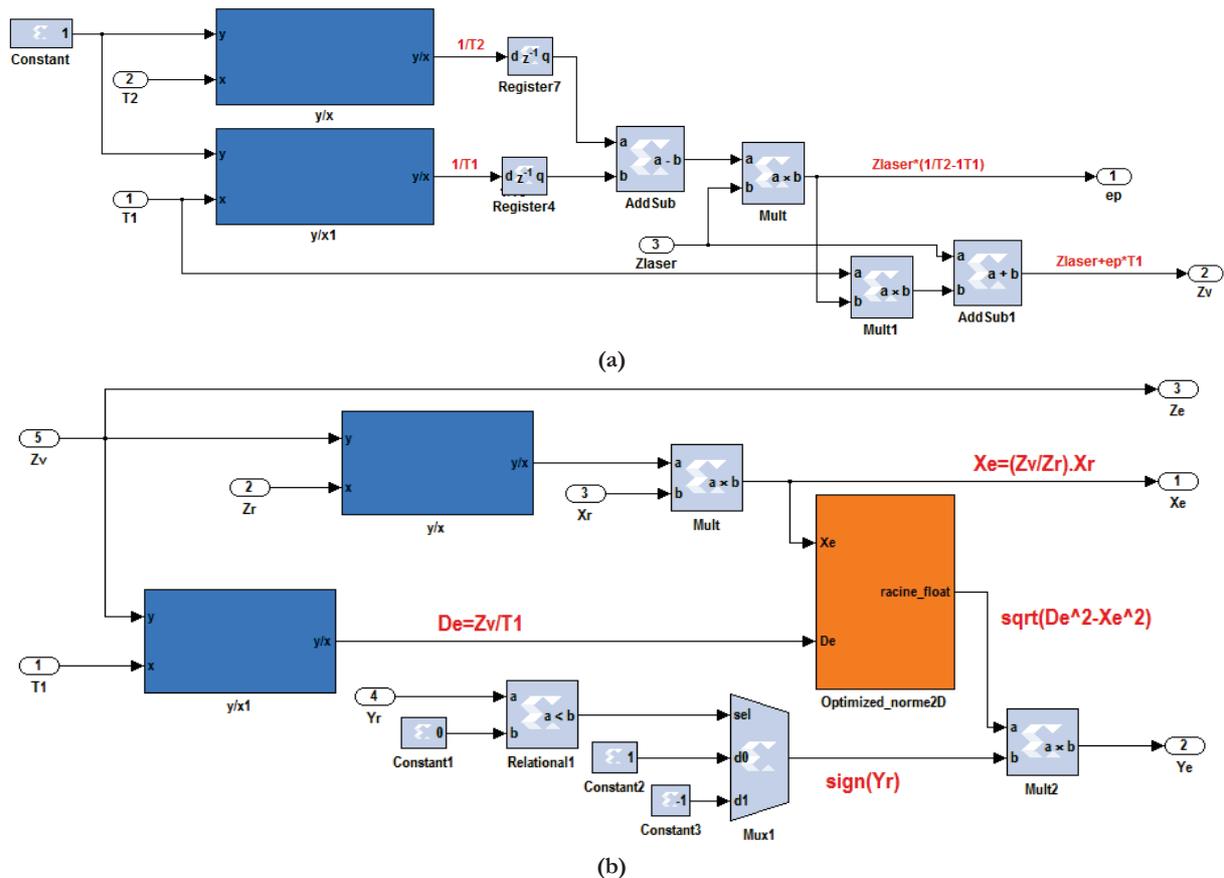


Figure IV. 24 : Module de détermination des coordonnées 3D avec étalonnage dans l'air

Deux modules de reconstruction 3D sont conçus identiquement pour traiter les données issues de l'image1 ainsi que l'image2. La seule différence réside dans les paramètres d'entrées qui représentent les paramètres d'étalonnage intrinsèques à chaque système catadioptrique. Le résultat final est la moyenne obtenue entre les deux estimations.

4. Expérimentation et analyse de performances

Une simulation d'un système XSG conçu peut être effectuée sans considérations temporelles, mais le prototype ne peut être validé que lorsqu'une réponse en latence et précision adéquates sont obtenues.

4.1. Résultats de co_simulation

Afin de vérifier si le comportement du modèle XSG correspond à celui du code Matlab initialement préparé pour la vérification du bon fonctionnement de l'algorithme théorique, les résultats des deux modèles sont comparés en notant qu'il pourrait y avoir des erreurs relatives en fonction du nombre de bits utilisés. Pour accéder à la validation et la mise en œuvre de la conception sur le circuit FPGA, l'outil ISE et System Generator permettent de générer

automatiquement le fichier binaire d'implémentation. Aussi, un bloc hardware est créé pour assurer le transfert du fichier de configuration (.bit) sur le circuit FPGA via un câble JTAG. Le nouveau bloc de co_simulation pourrait remplacer le modèle XSG et être prêt à l'exécution sur le dispositif FPGA. La figure 25 donne le bloc de co_simulation matériel généré par les outils XSG et ISE pour la plate-forme Xilinx ML605. Le bloc présente les mêmes entrées et sorties que l'architecture XSG : en entrée, les paramètres de modélisation des deux systèmes ainsi que les composantes couleurs des pixels des images stéréo et en sortie, les coordonnées 3D des points lasers envoyés successivement pour former la structure 3D du conduit à l'instant de capture et la fusionner avec la structure 3D construite antérieurement.

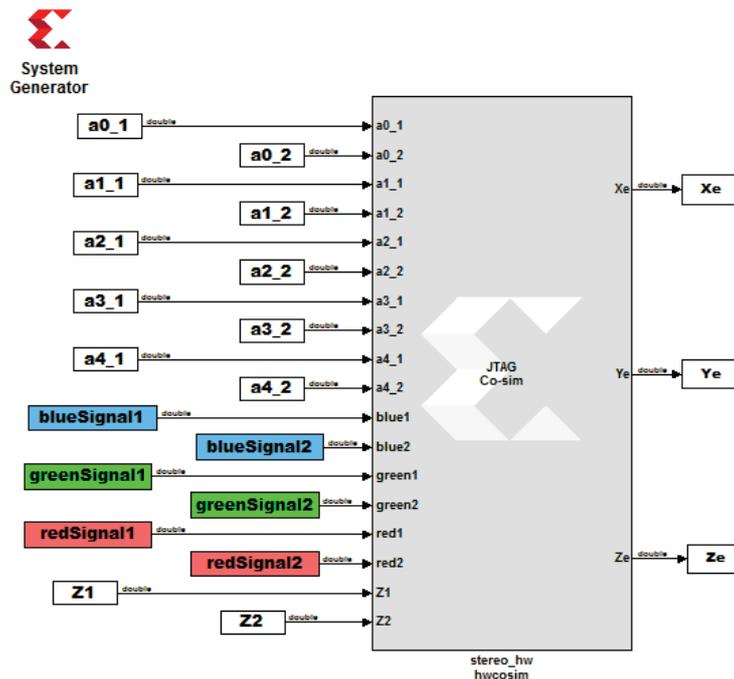


Figure IV. 25 : Module de Co_simulation

La co_simulation sous System Generator permet de tester l'application avec les mêmes conditions que si elle est réellement implémentée sur un FPGA physique, avec la même largeur de bit des variables et des opérateurs. Le résultat de la co_simulation est effectué "bit-to-bit" à "cycle d'horloge" précis, ce qui signifie que chaque bit est traité avec précision et chaque opération prend exactement le même nombre de cycles. Ainsi, il n'y aura presque pas de surprises lorsque l'application est finalement en cours d'exécution sur le FPGA.

Les premières étapes de traitement: segmentation, identification des spots laser et l'extraction sont présentés dans la figure 26. Celle-ci présente respectivement les résultats du code Matlab (Fig.26.a) et les résultats correspondants à l'architecture XSG (Fig.26.b). Une petite différence existe entre les deux processus en raison de la conversion du type de données de double précision en virgule flottante à virgule fixe spécifié par l'utilisateur. L'identification des nuages de points utiles qui entourent la projection sur laser est effectuée par co_simulation sur la

cible matérielle sélectionnée et est présentée dans la figure 26.c. L'extraction des coordonnées en pixels des spots laser est présentée dans la figure 26.d, les marques en vert sont les centres de gravité calculés par Matlab et ceux en rouges sont ceux calculés par implémentation. Une erreur moyenne d'environ de ± 2 pixels entre le logiciel et le matériel est estimé.

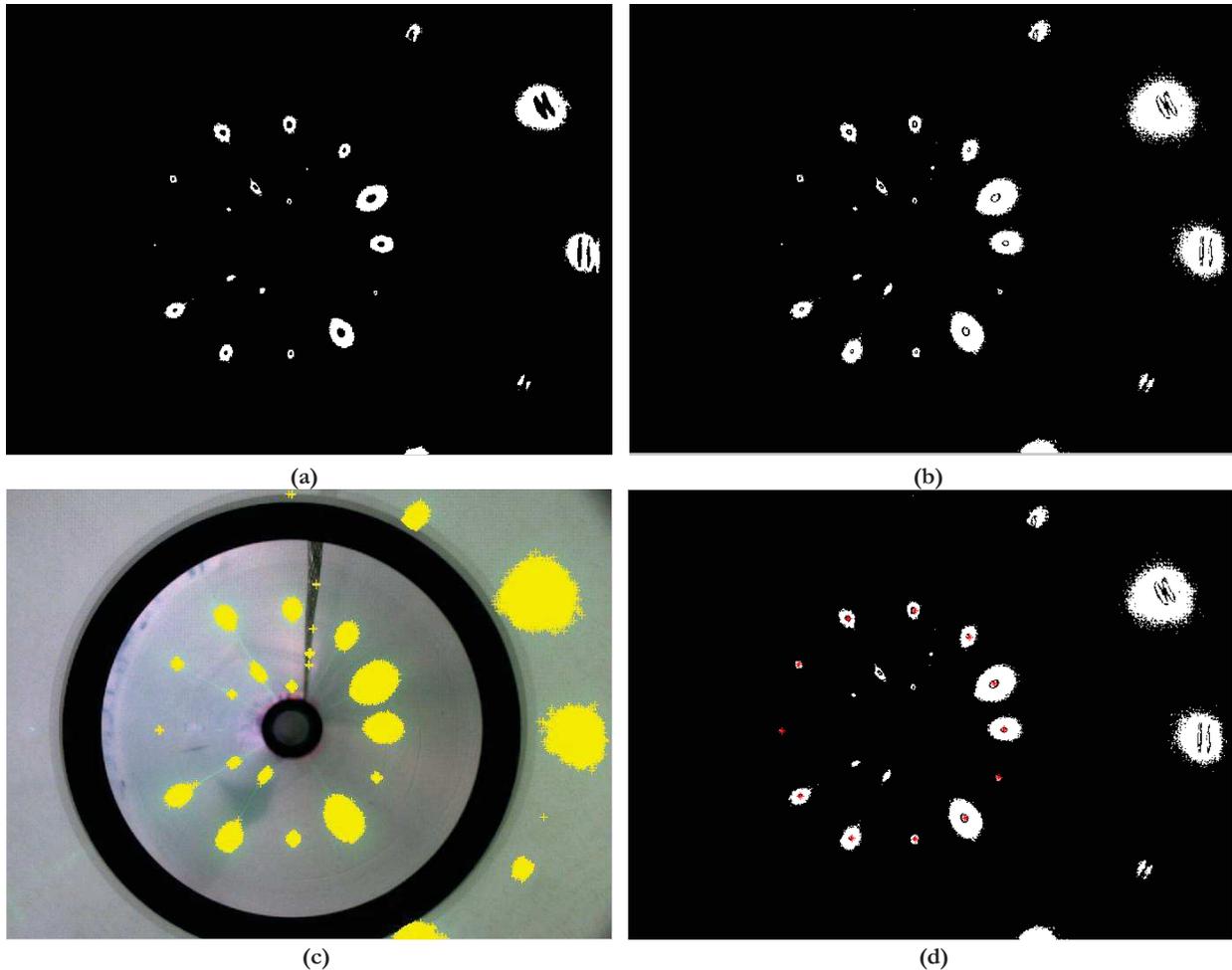


Figure IV. 26 : Résultats de pré-traitement d'image (a) Segmentation par code Matlab, (b) Segmentation par co_simulation, (c) Identification des spots Laser par Hw_cosimulation, et (d) Extraction des centres des spots par code Matlab et blocs XSG

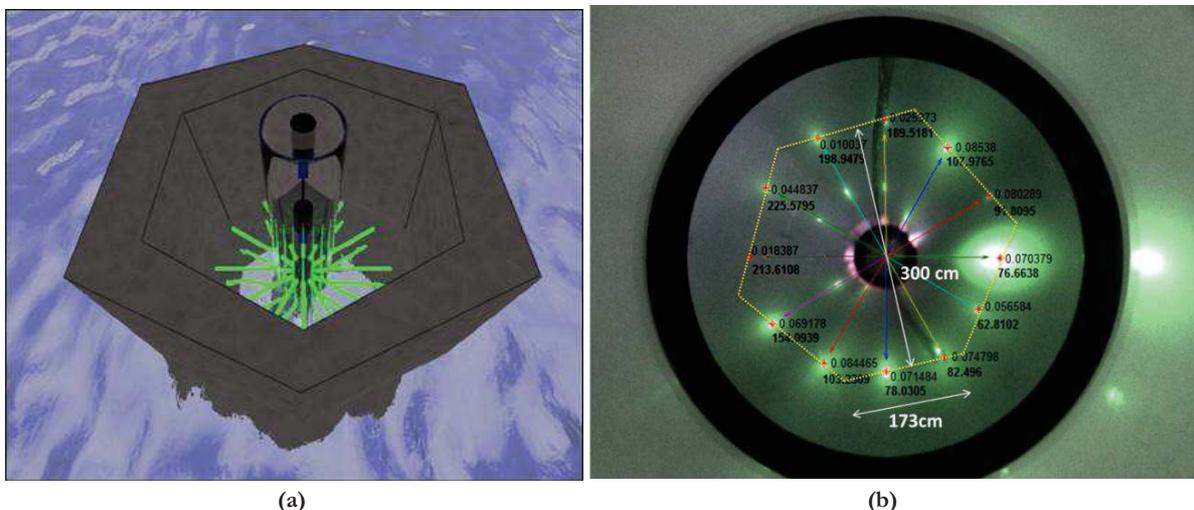


Figure IV. 27 : Résultats de co_simulation hardware sur FPGA

Les résultats de co_simulation hardware de l'algorithme de reconstruction de forme 3D à partir de points 3D particuliers sont donnés dans la figure suivante. L'environnement de test, de forme hexagonale régulière (à six côtés égaux), est une piscine remplie d'eau dont les dimensions sont indiquées sur la figure. Le système catadioptrique à lasers est plongé sous l'eau et positionné à des emplacements prédéfinies (Fig.27.a). La figure 27.b donne les mesures de distances obtenues par simulation sur FPGA et l'erreur relative par rapport aux dimensions réelles déterminées à partir des données mesurées. Pour l'objet hexagonal, l'erreur quadratique moyenne est d'environ $5,45 \text{ mm}^2$ qui est tout à fait satisfaisant si on considère la mauvaise résolution des images sous-marines et les effets de réfraction de la lumière. Une erreur moyenne, pour les différents objets utilisés dans les expérimentations, est quantifiée à 2,5%.

4.2. Résultats de synthèse

La synthèse architecturale de la conception XSG est réalisée sous l'outil de développement ISE générant de manière automatique une description structurelle au niveau transfert de registre (RTL) à partir d'une description comportementale. Ce sera au minimum des portes NAND, NOR, NOT, des bascule D, des buffers d'entrées-sorties, mais cela peut être aussi des compteurs, des multiplexeurs, des mémoires RAM, des registres, des additionneurs, des multiplieurs ou tout autres blocs particuliers. Elle décrit le chemin de données représentant la partie opérative du système et constituée d'unités fonctionnelles (additionneurs, multiplieurs, ALUs et unités logiques), d'unités de mémorisation (RAMs, ROMs, Registres ou Flip-flop, etc.) et d'unités de communication interconnectées (bus et multiplexeurs, etc.). La translation transforme la Netlist issue de la phase de synthèse en des primitives FPGA. Elle est généralement optimisée de façon à tirer profit au mieux des ressources FPGA ciblé.

Outre la description VHDL, le concepteur a la possibilité de fournir au synthétiseur des contraintes temporelles réalistes en termes d'ordre de coût, de temps d'exécution ou de performance, de délai de propagation, de limitation de consommation de puissance ou de ressources, de mémoire, etc. Le synthétiseur, par un certain nombre d'itérations, essaiera d'optimiser le produit surface-vitesse. Il peut aussi choisir de réaliser les mémoires et les multiplieurs par des LUTs ou bien par des BRAMs et des DSP48E. En outre, l'optimisation dépend du type de l'application à synthétiser, d'une fonction optimisation et d'une bibliothèque de modules permettant de mettre en œuvre les opérations de la description selon une solution répondant aux objectifs respectant les contraintes spécifiées.

Pour vérifier le bon fonctionnement de l'architecture, le design a été déployé sur la plateforme de prototypage ML605 (XC6VLX240T). La figure 28 présente un exemple d'architecture RTL correspondante au module de reconstruction 3D donné dans la figure 23. Les modules de division conçus sont indiqués en rouge et le bloc d'approximation de la norme 2D est

indiqué en jaune. Dans certaines architectures, entre les différents modules principaux, nous avons ajouté des registres pour assurer des architectures en pipeline, ce qui explique le nombre de registres utilisés (14% de la plateforme).

Le tableau 3 résume les résultats de synthèse de notre architecture en utilisant la plateforme ML506. Ce module utilise 28% des LUTs, 14% des slices-registres de la ML605. En plus des LUTs et des slices-registres, le module proposé utilise 11 BRAM à 36 Kbits chacune et 318 DSP48E ce qui représente (3%, 41%). L'implémentation fonctionne à une fréquence maximale de 275.179 MHz.

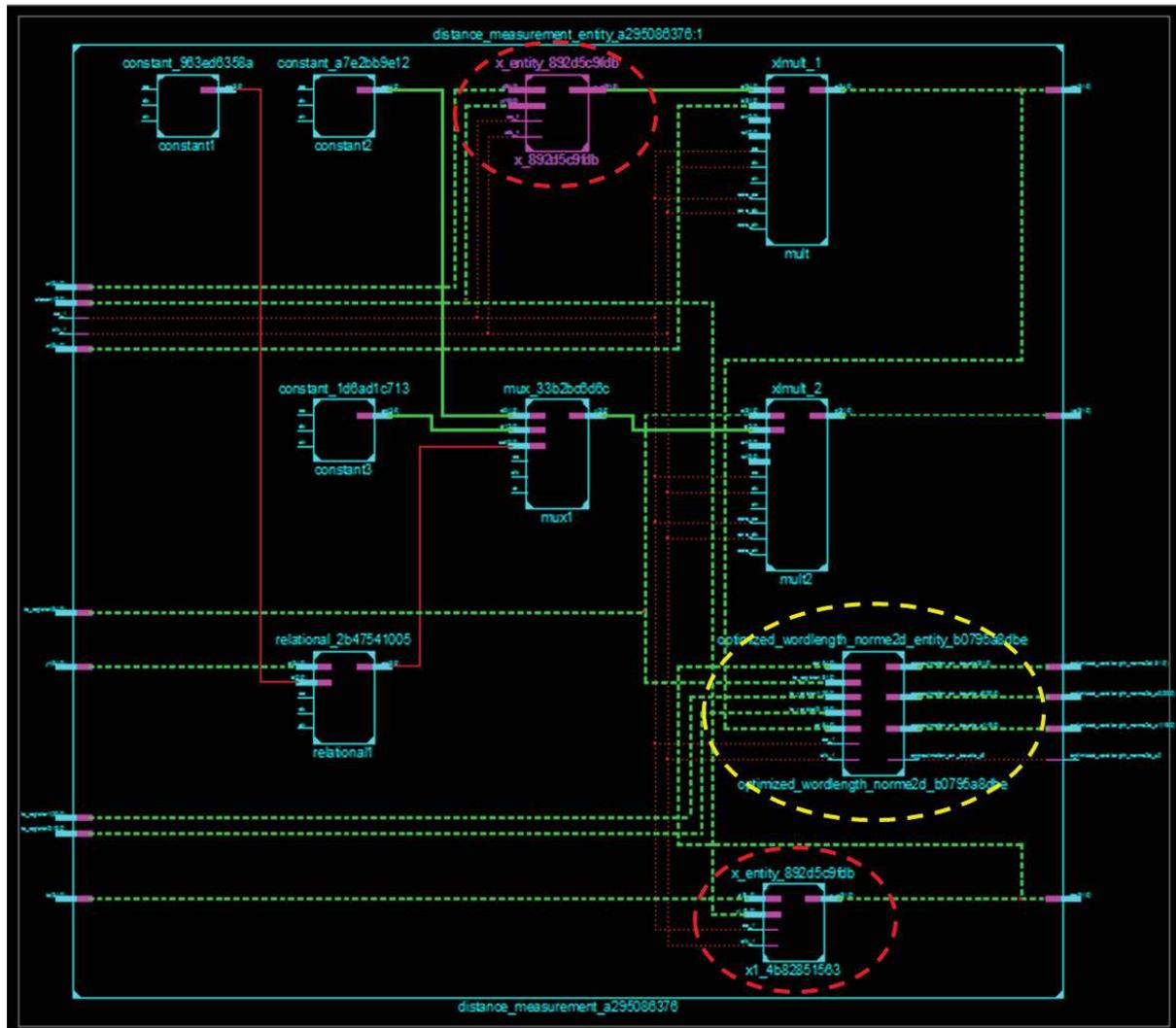


Figure IV. 28 : Schéma RTL du module de reconstruction 3D

Tableau IV. 3 : Reliquat des ressources FPGA utilisées pour toute l'architecture

	Slice-registres	LUTs	fully-used Flip Flops	BRAM	DSP48E	Fréquence max (MHz)
Utilisé	5381	43704	42630	11	318	275.179
Disponible	37680	150720	301440	416	768	
Taux d'utilisation	14%	29%	14%	3%	41%	

4.3. Analyse temporelle de la conception matérielle

L'engin de navigation traversant le conduit sous-marin est muni d'un système de vision stéréoscopique actif (Fig. 29). A chaque position du robot (" $t-T$ ", " t ", " $t+T$ ",...), les deux caméras omnidirectionnelles capturent simultanément deux images stéréo décrivant l'environnement aquatique visité. Ces données seront envoyées par l'intermédiaire du PC au système de traitement. Ce dernier échantillonne les deux images et construit une forme métrique du conduit à cette position. Cette coupe est assemblée au cours du temps avec le modèle précédent pour former un modèle complet de l'environnement avec des mesures métriques précisant la position du robot par rapport aux parois de l'environnement confiné. La période de traitement dépend d'une part de la distance " D " entre l'extrémité du robot et la ceinture laser et de la vitesse de navigation du robot " v ". Elle dépend aussi de la fréquence d'acquisition des caméras car pendant cette période " T ", le système de traitement doit s'exécuter pour envoyer les données métriques au système de contrôle de navigation pour pouvoir prendre la décision de déplacement nécessaire afin de centrer le robot avant l'acquisition des images suivantes.

Pour que notre système satisfasse les performances temporelles exigées par l'application, le temps d'exécution du processus de traitement " T_p " ne doit pas dépasser la période entre deux acquisitions successives " T_F " dépendante du système d'acquisition et la période " T " qui peut être variable en fonction de la vitesse du robot. Une analyse temporelle de l'architecture matérielle est présentée dans la suite de cette section pour pouvoir étudier son efficacité.

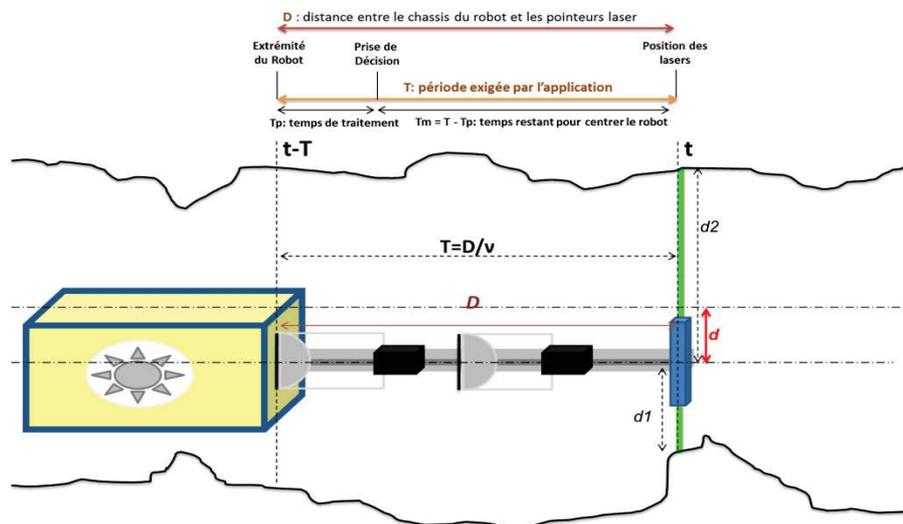


Figure IV. 29 : Illustration du processus de contrôle du robot

4.3.1. Fréquence maximale

Etant donné que l'analyse temporelle de l'architecture complète est longue surtout lorsque le système est complexe, nous procédons à une vérification de chaque sous-système séparément : analyser, ajuster et vérifier. Avant optimisation, un rapport donnant une analyse temporelle de la

conception XSG a été généré à l'aide de l'outil du développement ISE. La fréquence d'horloge maximale supportée par la conception était autour de 80,912 MHz. Cette valeur ne peut pas supporter les exigences temporelles de certaines applications. D'où la nécessité d'une phase d'optimisation permettant d'augmenter la fréquence d'horloge de la plateforme matérielle. System Generator propose un outil appelé "Timing analysis tool" qui aide l'utilisateur à travers des fichiers graphiques et textuels générés après la phase de synthèse et lui permettent d'analyser et d'extraire les chemins critiques. Le chemin critique a été identifié dans le rapport généré donnant les valeurs de latences de certains blocs qui ont été ensuite étudiés et partiellement mis en parallèle pour diminuer la latence. Des techniques d'optimisation ont été appliquées dans les bons endroits pour minimiser la latence des chemins critiques, telles que l'ajout manuel de "delays" et de "registers", l'ajustement de la latence de certains opérateurs comme ceux de multiplication,...etc. Après vérification de chaque sous-système, une analyse de l'architecture complète a été réalisée. Après optimisation, la nouvelle fréquence de fonctionnement est à 275,179 MHz qui est beaucoup plus compatible avec les systèmes à temps-réel. Dans le tableau 4, nous présentons la fréquence maximale nécessaire pour accomplir chaque bloc après optimisation et celle obtenue pour l'architecture matérielle complète.

Tableau IV. 4 : Temps d'exécution des blocs matériels

Modules XSG	Période minimale (ns)	Fréquence maximale (MHz)
Pré-traitement des images Extraction des spots laser	1.388	720,636
Transformation 2D-3D Reconstruction 3D	3.594	278.276
Architecture complète	3.634	275.179

4.3.2. Analyse de performances

Pour chaque acquisition stéréo, l'algorithme de restitution génère une forme 3D correspondante à la structure du karst à la position actuelle du robot. L'algorithme temps-réel proposé est mis en œuvre pour fournir, en plus de la forme 3D, des mesures de distances entre l'axe optique du système de vision et chaque projection du faisceau laser sur la paroi du karst. Cette information pourra être utilisée pour localiser le robot par rapport à la surface du conduit et alerter le dispositif de commande de navigation en cas de présence d'obstacles pour pouvoir varier la trajectoire de l'engin au bon moment. Afin de satisfaire les contraintes temporelles, l'algorithme de traitement doit être effectué au moins à la fréquence d'acquisition de trame qui dépend, elle-même, de la fréquence de capture du capteur utilisé.

Dans nos travaux, nous avons utilisé deux caméras *FireWire* fabriquées par Point Grey (Point Grey Research Inc., Richmond, BC, Canada), compatible avec la norme IEEE-1394b

(FireWire b). Pour la famille *Flea2* de type FL2-08S2C, la fréquence d'horloge maximale d'envoi de pixel pourrait atteindre $66,67\text{ MHz}$, ce qui donne un débit d'un pixel chaque 15 ns . La période minimale générée par l'algorithme de traitement est d'environ $3,634\text{ ns}$ qui est inférieur au débit pixel. Pour cette famille de caméra, le nombre maximal d'image par seconde (FPS : frame per second) est égal à 30 FPS pour une résolution maximale de 1024×768 pixels. Ceci exige que le processus doit s'exécuter pendant une période de temps qui doit être inférieure à la période de $T = 33\text{ ms}$ entre deux trames successives.

L'analyse temporelle de l'architecture complète a été effectuée en vue d'obtenir la valeur de la latence correspondante à toute la chaîne de traitement, qui est par définition le délai entre l'introduction des données ou de commandes dans un système et le résultat souhaité en sortie. Elle est le résultat de multiplication du nombre de cycles d'une architecture (algorithme) et la période d'horloge (période de chaque cycle). Dans le cas de notre application, l'architecture XSG proposée pour le pré-traitement des images et l'extraction des coordonnées des spots laser demande 1024×768 cycles d'horloge qui est en effet égal à la résolution de l'image. Pour ce qui est de la reconstruction 3D, le nombre de cycles demandé pour générer les coordonnées 3D des 12 pointeurs laser est de 1200 cycles d'horloge. Ce module à plus courte latence va s'exécuter en même temps que le module d'extraction de spots laser. Par conséquent, la latence de toute l'architecture sera égale à celle du module d'extraction des spots laser. En profitant de la possibilité de la mise en parallèle, la fréquence maximale de l'architecture complète atteint $275,18\text{ MHz}$. Pour cette fréquence, le système proposé peut atteindre les performances suivantes :

- Débit maximal = **275 Mégapixels par seconde**
- Temps d'exécution du module de pré-traitement et extraction de spots laser = **2.86 ms**
- Temps d'exécution du module de reconstruction 3D et génération de résultat = **0.00436 ms**
- Temps d'exécution total considérant le parallélisme = **2.86 ms**
- Fiabilité de reconstruction, pour une vitesse de navigation de 2 ms^{-1} , égale à **60mm**.

Avec l'analyse temporelle présentée ci-dessus, on présente dans cette section les possibilités d'amélioration des performances de système. Étant donné que notre système respecte largement les contraintes temporelles déterminées en se basant sur les caractéristiques des caméras à 30 FPS actuellement utilisées. Avec un temps d'exécution de 2.86 ms pour une résolution d'image de 1024×768 pixels et pour une période de 33 ms , le nombre d'images par seconde peut atteindre 300 FPS. Ceci nous permet d'utiliser des caméras à haute vitesse et à résolution maximale afin d'aboutir à une reconstruction plus précise et plus fiable. Pour un système de caméras à 300 FPS, le processus de traitement pourra générer une forme 3D correspondante à chaque acquisition toute les 6 mm , pour une vitesse de navigation de 2 ms^{-1} . Aussi, étant donné que la durée de la phase d'extraction des spots lasers égale à 2.86 ms ne dépend pas du nombre de lasers utilisés

(architecture parallélisée), on peut envisager l'utilisation de plus de 12 pointeurs lasers dans le but d'obtenir une forme 3D plus dense. Par exemple, l'utilisation d'une ligne laser omnidirectionnelle nous permet d'extraire jusqu'à 3600 pointeurs lasers à 0.1° d'angle d'écartement (au lieu de 30°) et même plus.

5. Conclusion

La vision stéréo est une solution considérée comme difficile vu qu'elle exige parfois un taux de calcul énorme dans de nombreuses applications utiles. Plusieurs travaux ont montré que un circuit FPGA est la plupart du temps le plus adapté pour l'implémentation d'une classe spécifique d'algorithmes de stéréovision et plus efficace que la plupart des autres plateformes de calcul, tels que des processeurs spécifiques (DSP) ou des processeurs programmables à usage général (GPP). Dans ce chapitre, nous avons présenté une architecture synthétisable prête à être implémentée sur FPGA et permettant de restituer des informations 3D à partir de données 2D en temps-réel. Nous avons montré que cette conception matérielle exploitant le parallélisme au plus possible permet d'atteindre des performances en termes de précision et temps d'exécution satisfaisant les contraintes exigées par l'application et l'environnement de test. En effet, pour une cadence de 30FPS, relative aux caméras utilisées, notre processus est capable de s'exécuter en un temps (2.86ms) largement inférieur à la période exigée de 33ms entre deux trames successives. Ce qui nous permet d'étudier la possibilité d'améliorer davantage la fiabilité de reconstruction du système, à savoir ; passer à une reconstruction temps-réel plus dense.

Conclusion Générale

Les travaux de cette thèse s'intègrent dans le cadre de l'exploration sous-marine ayant comme objectif principal la caractérisation topographique et morphologique de réseaux karstiques sous-marins en se basant sur des données de perception. Ceci touche divers domaines, notamment la vision omnidirectionnelle, les systèmes catadioptriques, l'imagerie numérique et l'implémentation d'architecture matérielle sur circuit reconfigurable pour la mise en œuvre d'un système temps-réel. Étant donné la diversité des aspects liés à l'objectif principal, des contributions ont été apportées dans cette thèse suivant les différents aspects : mise en place d'un système de vision adéquat, développement d'une approche d'étalonnage du système de vision, mise en œuvre d'un algorithme de reconstruction 3D et l'implémentation sur FPGA de l'architecture matérielle dédiée.

Le premier aspect étudié dans cette thèse vise la mise en place d'un système de vision qui répond aux contraintes de l'environnement sous-marin et permet l'acquisition d'informations utiles pour une restitution la plus fiable de la structure métrique de l'environnement. Pour ce faire, nous avons effectué une étude portant sur les systèmes de vision en s'intéressant plus particulièrement aux systèmes à large champ de vision: les systèmes catadioptriques et leur modélisation géométrique. Pour notre application, nous avons envisagé l'utilisation d'une paire de capteurs catadioptriques constitué chacun d'une caméra standard pointant sur un miroir hyperbolique permettant ainsi de récupérer une vue panoramique de l'environnement confiné. On a opté pour une configuration verticale des deux systèmes offrant ainsi un champ de vue commun plus important. À cause du manque de luminosité et de texture dans les fonds marins, on a muni notre système stéréoscopique d'un système de pointeurs laser placé de manière à avoir une projection omnidirectionnelle des lasers dans l'espace. Ceci nous permet d'avoir un minimum d'amers visuels dans la scène pour pouvoir les extraire et reconstruire la forme 3D du milieu visité.

La deuxième contribution apportée est de développer une approche d'étalonnage du système de vision proposé en tenant en considération d'une part les dispositifs mis en place permettant l'étanchéité du matériel électronique (caméras + objectifs + lasers) et d'autre part les phénomènes optiques tels que les lois de réfraction de la lumière. En étudiant les différentes méthodes d'étalonnage proposées dans la littérature, nous nous sommes inspirés de l'approche proposée par Scaramuzza qu'on a adapté à notre système étant donné qu'elle présente une méthode intelligente ; moins complexe et plus générique. Multiples acquisitions et expérimentations ont été effectuées, minutieusement, sur le système complet pour la validation de l'approche mise en œuvre et dans l'objectif d'obtenir la modélisation géométrique la plus précise, à une erreur de re-projection la plus faible.

Le troisième aspect étudié consiste à la mise en œuvre d'un algorithme de reconstruction 3D à partir de points saillants extraits des images stéréoscopiques en utilisant les paramètres

d'étalonnage du système proposé. Pour ce faire, nous avons défini une chaîne de traitements numériques correspondants aux différentes étapes de la démarche classique d'une reconstruction 3D. En étudiant les caractéristiques des différentes techniques proposées dans la littérature, une architecture séquentielle adaptée à notre système a été proposée. Cette architecture se décompose en deux parties dédiées respectivement à la détection et l'extraction de points saillants et à la reconstruction 3D de la scène à partir des primitives appariées. Le premier élément consiste à la détection et extraction de primitives. Etant donné que les points laser représentent les points saillants, les résultats de simulations des algorithmes standards d'extraction de ces primitives (ex. SIFT, SURF, Harris, etc.) n'étaient pas satisfaisants. Ainsi, nous avons proposé, suite à plusieurs recherches, une approche qui tient compte des caractéristiques intrinsèques de notre système de vision et qui permet effectivement d'extraire les points utiles à la restitution de la structure 3D de l'environnement confiné. L'approche proposée exploite la redondance des régions de projection des faisceaux laser au sein d'une séquence d'images omnidirectionnelles, qui permet en outre d'identifier les projections laser pertinentes et supprimer ceux provoqués par des réflexions non-désirables. Quant à la reconstruction des points 3D, second élément de l'architecture, nous nous sommes inspirés des approches de triangulation stéréo à partir de données appariées 2D. Celles-ci ont été adaptées à notre géométrie en essayant de prendre en considération les phénomènes de réfraction de la lumière en passant d'un milieu à un autre différent. Dans ce même cadre, nous avons proposé différentes démarches pour la détermination des coordonnées 3D d'un point plongé dans l'eau. Chaque démarche dépend des conditions d'étalonnage (dans l'air / l'eau), de l'approche de reconstruction utilisée (stéréovision active / passive). Aussi, nous avons présenté une approche d'estimation d'indice de réfraction d'un milieu d'indice de réfraction inconnu à partir de multiples expérimentations d'étalonnage. Les performances de l'approche d'extraction de points laser ainsi que celle des approches de reconstruction 3D ont été évaluées en termes de taux d'erreur en comparant entre les résultats estimés et les données réelles. Les performances du système proposé dans ces travaux de thèse sont satisfaisantes malgré la mauvaise résolution des images sous-marines et les effets de réfraction de la lumière.

La dernière contribution apportée à cette thèse consiste à concevoir une architecture matérielle du système de reconstruction 3D. L'environnement de développement exploité nous a permis un prototypage rapide du système proposé. Toute la conception a été développée de manière à profiter du parallélisme et des ressources disponibles dans un circuit FPGA. Néanmoins, les performances de certaines fonctions, de la bibliothèque Xilinx, n'ont pas été satisfaisantes. Dans ce manuscrit, nous avons proposé une solution plus optimisée de ces fonctions. Les résultats de co_simulation ont fourni des résultats prometteurs qui répondent aux objectifs du travail, notamment au niveau des points saillants extraits des images et aussi au niveau des coordonnées 3D reconstruits. La synthèse de toute l'architecture nous a permis

d'obtenir de bonnes performances en termes de précision et de temps d'exécution. Ces résultats montrent que le système implémenté satisfait largement les contraintes temporelles exigées par l'application ce qui nous permet d'envisager d'autres améliorations.

Cette thèse représente une "brique" d'un projet de grande ampleur, qui intègre plusieurs aspects à savoir la commande, la téléopération et la vision. Les premières performances obtenues nous permettent de "voir" au-delà des objectifs prédéfinis. Parmi les perspectives actuelles de ce travail, on peut citer :

- Augmentation de la fréquence d'acquisition par des optimisations logicielles et l'utilisation de meilleures technologies matérielles. Le succès des technologies FPGAs et l'apparition de plateformes matérielles qui supportent la reconfiguration partielle et dynamique fait que la majorité des travaux de recherche en prototypage rapide sont orientés vers la conception de tels systèmes.

- Amélioration des performances temporelles de la phase d'extraction des points saillants (phase la plus coûteuse en temps d'exécution). Etant donné que les points saillants sont des projections issues des pointeurs laser, en particulier. Il est envisageable d'introduire certaines contraintes telles que les régions possibles de projections afin de converger plus rapidement vers des solutions plus précises. Cette solution consiste à l'acquisition simultanée de seulement des régions d'intérêt de résolution pixellique plus faible et de lancer la recherche en parallèle des points saillants dans chaque région.

- Amélioration de la robustesse du système développé en effectuant de divers tests et expérimentations sur des séquences d'images omnidirectionnelles impliquant des conditions différentes, notamment en milieu sous-marin. Ces tests permettront d'adapter notre approche aux différents milieux aquatiques et de généraliser son application.

- Les données métriques estimées accompagnant la structure 3D de l'environnement peuvent être exploitées dans un outil informatique permettant de regrouper toutes les données historiques accumulées lors de l'évolution de l'engin sous-marin dans son environnement. Ceci permet de fournir un moyen de sauvegarde de données et l'observation spatio-temporelle intéressante pour diverses applications marines, comme par exemple l'observation de l'évolution du milieu, élaboration de cartes 3D, ...etc.

- Les données métriques peuvent être utilisées comme source d'information pour la commande intelligente du robot au sein de l'environnement, en essayant à travers ces mesures de détecter de possibles obstacles, de suivre l'état du conduit et d'assurer une navigation sans risque de l'engin en le positionnant le plus possible au centre du milieu confiné.

Bibliographies

- [1] Lumières et couleurs. Web Page: <http://www.profil-couleur.com/lc/004-couleur-young.php>.
- [2] Jolion J-M. Sur la méthodologie de conception des systèmes de vision. Chapitre du livre. Les systèmes de vision. Hermès (ICC), pp. 97-131, Paris, **2000**.
- [3] Ballard DH. Active Perception. Encyclopedia of Neuroscience, pp. 31–37, **2010**. doi:10.1016/B978-008045046-9.01436-4.
- [4] Aloimonos J. Purposive and qualitative active vision. Proceedings of 10th International Conference on Pattern Recognition, IEEE Computer Society Press, pp. 346–360, **1990**. doi:10.1109/ICPR.1990.118128.
- [5] Ballard DH. Reference Frames for Animate Vision. 11th International Joint Conference on Artificial Intelligence, vol. 2, pp. 1635–1641, **1989**.
- [6] Marr D. and Poggio T. A computational theory of human stereo vision. Proceedings of the Royal Society of London. Series B, Biological Sciences, vol. 204 , pp. 301–328, **1979**.
- [7] Brooks R-A. Achieving Artificial Intelligence through Building Robots. Technical Report, AI Memo n°899. MIT Artificial Intelligence Laboratory, Cambridge MA, **1986**.
- [8] Trivedi M.M. and Rosenfeld A. On Making Computers See. IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 19(6), pp. 1333–1335, **1989**.
- [9] NAOULO A. Architectures pour la stéréovision passive dense temps réel: Application à la stéréo-endoscopie. Thèse soutenue le 5 septembre **2006**.
- [10] Keller M., Lefloch D., Lambers M., Izadi S., Weyrich T. and Kolb A. Real-time 3D Reconstruction in Dynamic Scenes using Point-based Fusion. 2013 International Conference on 3D TV, pp. 1–8, **2013**. doi:10.1109/3DV.2013.9.
- [11] Newcombe R-A. et al. KinectFusion: Real-Time Dense Surface Mapping and Tracking. IEEE International Symposium on Mixed and Augmented Reality, pp. 127–136, **2011**. doi:10.1109/ISMAR.2011.6092378.
- [12] Niessner M., Zollhöfer M., Izadi S. and Stamminger M. Real-time 3D Reconstruction at Scale Using Voxel Hashing. ACM Transactions on Graphics, vol. 32(6), **2013**. doi:10.1145/2508363.2508374.
- [13] Gallo L., Placitelli A-P. and Ciampi M. Controller-free exploration of medical image data: Experiencing the Kinect. Proceedings on IEEE Symposium on Computer-Based Medical Systems, pp. 1-6, **2011**. doi:10.1109/CBMS.2011.5999138.
- [14] Bauer S., Seitel A., Hofmann H-G. and Blum T. Real-Time Range Imaging in Health Care: A Survey. Chapter in: Time-of-Flight and Depth Imaging Sensors, Algorithms, and Applications, vol 8200, pp. 228–254, Springer, Heidelberg **2013**. doi:10.1007/978-3-642-44964-2_11.

- [15] Vera L., Gimeno J., Coma I. and Fernández M. Augmented mirror: Interactive augmented reality system based on kinect. In *Human-Computer Interaction - Interact 2011. Lecture Notes in Computer Science (LNCS)*, vol. 6949, pp. 483–486, Springer, Berlin, **2011**. doi:10.1007/978-3-642-23768-3_63.
- [16] DAVID. DAVID 3D SCANNING, Web site: <http://www.david-3d.com/fr>.
- [17] David SDK-Tool. Web site: <http://www.david-3d.com/fr/products/david-sdk>.
- [18] Baker S. and Nayar S. A Theory of Single-Viewpoint Catadioptric Image Formation. *International Journal of Computer Vision*, vol. 35(2), pp. 175-196, **1999**. doi:10.1023/A:1008128724364
- [19] Mouaddib E-M. Introduction à la vision panoramique catadioptrique. *Traitement de Signal*, vol. 22(5), pp. 409-417, **2005**.
- [20] El Natour G., Ait-Aider O., Rouveure R., Berry F. and Faure P. Toward 3D reconstruction of outdoor scenes using an MMW radar and a monocular vision sensor. *Sensors*, vol. 15, pp. 25937–25967, Switzerland **2015**. doi:10.3390/s151025937.
- [21] Pathak K., Birk A. and Vaskevicius N. Plane-based registration of sonar data for underwater 3D mapping. *IEEE International Conference on Intelligent Robots and Systems*, pp. 4880–4885, **2010**. doi:10.1109/IROS.2010.5650953.
- [22] Hurtos N., Cufi X. and Salvi J. Calibration of optical camera coupled to acoustic multibeam for underwater 3D scene reconstruction. In: *OCEANS'10 IEEE SYDNEY*, pp. 1–7, **2010**. doi:10.1109/OCEANSSYD.2010.5603907.
- [23] Leiva J-M., Martínez P., Pérez E-J., Urdiales C. and Sandoval F. 3D reconstruction of static indoor environment by fusion of sonar and video data. *Proceedings on International Symposium on Intelligent Robotic Systems (SIRS), Toulouse (France)*, **2001**.
- [24] Massot-Campos M. and Oliver-Codina G. Optical sensors and methods for underwater 3D reconstruction. *Sensors*, vol. 15, pp. 31525–31557, Switzerland **2015**. doi:10.3390/s151229864.
- [25] Daval V. et al. Extraction de primitives à partir d'images de lumière structurée. *Orasis, Congrès Des Jeunes Chercheurs En Vision Par Ordinateur, Cluny, France*, **2013**.
- [26] Kalisperakis I., Grammatikopoulos L., Petsa E. and Karras G. A structured-light approach for the reconstruction of complex objects. *Geoinformatics*, vol. 6, pp. 259–266, **2011**. doi:10.14311/gi.6.32.
- [27] Dias P., Sequeira V., Vaz F. and Gonçalves J-G-M. Combining intensity and range images for 3D modelling. *Proceeding of the International Symposium on Virtual and Augmented Architecture*, pp. 139-145, Dublin **2001**. doi: 10.1007/978-1-4471-0337-0_16.
- [28] Massot M. and Oliver G. One-Shot Underwater 3D Reconstruction. In *proceeding of the IEEE Emerging Technology and Factory Automation Conference*, pp.1-4, **2014**.

doi:10.1109/ETFA.2014.7005282

- [29] Gruen A., Remondino F. and Zhang L. Image-Based Automated Reconstruction of the Great Buddha of Bamiyan, Afghanistan. *Computer Vision and Pattern Recognition Workshop on Application of Computer Vision in Archaeology*, Madison, USA, **2003**.
- [30] Fassold H., Danzl R., Schindler K. and Bischof H. Reconstruction of archaeological finds using shape from stereo and shape from shading. *Proceeding of the 9th Computer Vision Winter Workshop*, pp. 21–30, Piran, Slovenia, **2004**.
- [31] Zhang R., Tsai P., Cryer J-E and Shah M. Shape from shading: A survey introduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21(8), pp. 690-706, Washington, DC, USA, **1999**.
- [32] Samaras D. and Metaxas D. Incorporating illumination constraints in deformable models for shape from shading and light direction estimation. *Proceeding of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 247–264, **2003**. doi:10.1109/TPAMI.2003.1177155.
- [33] Zheng Q. and Chellappa R. Estimation of illuminant direction, albedo, and shape from shading. *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '91)*, pp. 540 - 545, **1991**. doi: 10.1109/CVPR.1991.139750.
- [34] Tsiotsios C., Angelopoulou M-E., Kim T. and Davison A-J. Backscatter Compensated Photometric Stereo with 3 Sources. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2259–2266, **2014**. doi:10.1109/CVPR.2014.289.
- [35] Lun W., Arvind G., Boxin S., Yasuyuki M., Yongtian W., and Yi M. Robust Photometric Stereo via Low-Rank Matrix Completion and Recovery. In *Proceedings of Asian Conference on Computer Vision*, vol. 6494, pp. 703-717, Berlin **2010**. doi:10.1007/978-3-642-19318-7_55
- [36] Miura J., Negishi Y. and Shirai Y. Mobile robot map generation by integrating omnidirectional stereo and laser range finder. *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and System*, vol. 1, pp. 250–255, **2002**. doi:10.1109/IRDS.2002.1041397.
- [37] Hogue A., German A. and Jenkin M. Underwater environment reconstruction using stereo and inertial data. In *IEEE International Conference on Systems, Man and Cybernetics*, pp. 2372–2377, October **2007**. doi:10.1109/ICSMC.2007.4413666.
- [38] Fischler M-A., Bolles R-C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, vol. 24, pp. 381-395, **1981**. doi:10.1145/358669.358692.
- [39] Carreras M., Ridao P., Garcia R. and Nicosevici T. Vision-based localization of an underwater robot in a structured environment. In *proceeding of the IEEE International*

- Conference on Robotics and Automation, vol. 1, pp. 971–976, Taipei, Taiwan **2003**. doi:10.1109/ROBOT.2003.1241718.
- [40] Burguera A., Bonin-Font F. and Oliver G. Trajectory-based visual localization in underwater surveying missions. *Sensors*, vol. 15, pp. 1708–1735, Switzerland **2015**. doi:10.3390/s150101708.
- [41] Jac-Hean K. and Myung J-C. Slam with omni-directional stereo vision sensor. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 1, pp. 442–447, **2003**. doi:10.1109/IROS.2003.1250669.
- [42] Kim A. and Eustice R-M. Real-time visual SLAM for autonomous underwater hull inspection using visual saliency. *IEEE Transactions on Robotics*, vol. 29, pp. 719-733, **2013**. doi:10.1109/TRO.2012.2235699.
- [43] Aulinas J., Petillot Y-R., Llandó X., Salvi J. and Garcia R. Vision-based underwater SLAM for the SPARUS AUV. *Proceeding of the International Conference on Computer and IT Applications in the Maritime Industries (COMPIT)*, pp. 171–180, **2011**.
- [44] Palomeras N., Nagappa S., Ribas D., Gracias N. and Carreras M. Vision-based localization and mapping system for AUV intervention. *MTS/IEEE OCEANS*, pp. 1–7, Bergen **2013**. doi:10.1109/OCEANS-Bergen.2013.6608058.
- [45] Tommaselli A., Marcato J-J., Olivete A. and Moraes M. A mobile platform with a catadioptric sensor. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS)*, vol. 1, pp. 407-413, **2014**. doi:10.5194/isprsarchives-XL-1-407-2014.
- [46] Aliakbarpour H., Tahri O. and Araujo H. Visual servoing of mobile robots using non-central catadioptric cameras. *Robotics and Autonomous Systems*, vol. 62, pp. 1613–1622, **2014**. doi:10.1016/j.robot.2014.03.007.
- [47] Adorni G., Mordonini M., Cagnoni S. and Sgorbissa A. Omnidirectional stereo systems for robot navigation. In *Proceeding of IEEE Workshop on Omnidirectional Vision and Camera Networks (Omnivis-03)*, pp. 79–79, **2003**. doi:10.1109/CVPRW.2003.10076.
- [48] Peng Q., Tu L., Zhang K. and Zhong S. Automated 3D scenes reconstruction using multiple stereo pairs from portable four-camera photographic measurement system. *International Journal of Optics*, vol. 2015, pp. 1-9, **2015**. doi:10.1155/2015/471681.
- [49] Ragot N., Rossi R., Savatier X., Ertaud J-Y. and Mazari B. 3D volumetric reconstruction with a catadioptric stereovision sensor. *Proceeding of the IEEE International Symposium on Industrial Electronics (ISIE)*, pp. 1306–1311, **2008**. doi:10.1109/ISIE.2008.4677128.
- [50] Rossi R., Savatier X., Ertaud J-Y. and Mazari B. Real-time 3D reconstruction for mobile robot using catadioptric cameras. *Proceeding of the IEEE International Workshop on Robotic and Sensors Environments (ROSE)*, pp. 104–109, **2009**.

doi:10.1109/ROSE.2009.5355981.

- [51] Sogo T., Ishiguro H. and Trivedi M-M. Real-time target localization and tracking by Nonocular stereo. Proceedings of the IEEE Workshop on Omnidirectional Vision, IEEE Computer Society, pp. 153–160, **2000**. doi:10.1109/OMNVIS.2000.853823.
- [52] Jong W-L., Suya Y. and Neumann U. Tracking with omni-directional vision for outdoor AR systems. Proceedings of the International Symposium on Mixed and Augmented Reality, IEEE Computer Society, pp. 47–56, **2002**. doi:10.1109/ISMAR.2002.1115069.
- [53] Tang Y., Luo J., Li Y-F. and Zhou X. Rotational Kinematics Model Based Adaptive Particle Filter for Robust Human Tracking in Thermal Omnidirectional Vision. Mathematical Problems in Engineering, vol. 2015, **2015**. doi:10.1155/2015/347497.
- [54] Boutteau R., Savatier X. and Ertaud J-Y. A Dynamic Programming Algorithm Applied to Omnidirectional Vision for Dense 3D Reconstruction. Proceeding of the 2nd International IEEE Asian Conference on Pattern Recognition (IAPR), pp. 927–931, **2013**. doi:10.1109/ACPR.2013.170.
- [55] Mei C. Couplage Vision Omnidirectionnelle et Télémétrie Laser pour la Navigation en Robotique. Thèse soutenue le 9 Février **2007**.
- [56] Mei C. et Rives P. Calibrage non biaisé d'un capteur central catadioptrique. 14ème Congrès Francophone AFRIF-AFIA Reconnaissance des Formes & Intelligence Artificielle, **2004**.
- [57] Spacek L. and Burbridge C. Instantaneous robot self-localization and motion estimation with omnidirectional vision. Robotics and Autonomous Systems, vol. 55, pp. 667-674, **2007**. doi:10.1016/j.robot.2007.05.009.
- [58] Su L., Luo C. and Zhu F. Obtaining Obstacle Information by an Omnidirectional Stereo Vision System. Proceeding of the IEEE International Conference on Information Acquisition (ICIA), pp. 48–52, **2006**. doi:10.1109/ICIA.2006.305769.
- [59] Layerle J., Savatier X., Ertaud J. et Mouaddib E-M. Conception d'un capteur catadioptrique pour une observation optimale du conducteur et de la scène de conduite. Congrès Des Jeunes Chercheurs En Vision Par Ordinateur, ORASIS'09, Trégastel, France, **2009**.
- [60] Orghidan R., Mouaddib E-M. and Salvi J. A computer vision sensor for panoramic depth perception. In 2nd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA), vol. 3522, pp. 153-160, Estoril, Portugal, Juin **2005**. doi: 10.1007/11492429_19.
- [61] Hibbard Inshore, Web Page: <http://www.hibbardinshore.com/rov-sonar-inspection.html>.
- [62] Gawlik N. 3D modelling of underwater archaeological artefacts. Thèse soutenue en Juin 2014. Université du Norvège, **2014**.
- [63] Kantor G., Fairfield N., Jonak D. and Wettergreen D. Experiments in navigation and

- mapping with a hovering AUV. 6th International Conference on Field and Service Robotics (FSR), vol. 42, pp. 115-124, Chamonix, France, **2008**. doi:10.1007/978-3-540-75404-6_11.
- [64] Titanic on Behance. National Geographic Magazine **2014**. Produced by Woods Hole Oceanographic Institution. Web Page : <https://www.behance.net/gallery/ixtract-Titanic/3422941>.
- [65] Yamashita A., Kawanishi R., Koketsu T., Kaneko T. and Asama H. Underwater sensing with omni-directional stereo camera. Proceeding of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops), pp. 304–311, **2011**. doi:10.1109/ICCVW.2011.6130257.
- [66] Atsushi Y., Akira F. and Toru K. Three dimensional measurement of objects in liquid and estimation of refractive index of liquid by using images of water surface with a stereo vision system. Proceeding of the IEEE International Conference on Robotics and Automation (ICRA 2008), pp. 974–979, **2008**. doi:10.1109/ROBOT.2008.4543331.
- [67] Yamashita A., Ikeda S. and Kaneko T. 3-D Measurement of Objects in Unknown Aquatic Environments with a Laser Range Finder. Proceedings of the IEEE International Conference on Robotics and Automation, pp. 3912–3917, **2005**. doi:10.1109/ROBOT.2005.1570718.
- [68] Kawai R., Yamashita A. and Kaneko T. Three-dimensional measurement of objects in water by using space encoding method. Proceedings of the IEEE International Conference on Robotics and Automation, pp. 2830–2835, **2009**. doi:10.1109/ROBOT.2009.5152430.
- [69] Svoboda T. and Pajdla T. Epipolar geometry for central catadioptric cameras. International Journal of Computer Vision, vol. 49, pp. 23-37, **2002**. doi:10.1023/A:1019869530073.
- [70] Morel O. and Fofi D. Calibration of catadioptric sensors by polarization imaging. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'07), pp. 3939–3944, **2007**. doi:10.1109/ROBOT.2007.364083.
- [71] Scaramuzza D., Martinelli A. and Siegwart R. A flexible technique for accurate omnidirectional camera calibration and structure from motion. Proceedings of the 4th IEEE International Conference on Computer Vision Systems, ICVS'06, pp. 2006-2045, **2006**. doi:10.1109/ICVS.2006.3.
- [72] Kannala J. and Brandt S-S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, pp. 1335-1340, **2006**. doi:10.1109/TPAMI.2006.153.
- [73] Mei C. and Rives P. Single View Point Omnidirectional Camera Calibration from Planar

- Grids. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 3945–3950, **2007**. doi:10.1109/ROBOT.2007.364084.
- [74] Deng X-M., Wu F-C. and Wu Y-H. An Easy Calibration Method for Central Catadioptric Cameras. *Acta Automatica Sinica*, vol. 33, pp. 801-808, **2007**. doi:10.1360/aas-007-0801.
- [75] Frank O., Katz R., Tisse C-L. and Durrant-Whyte H. Camera calibration for miniature, low-cost, wide-angle imaging systems. Proceedings of the British Machine Vision Conference (BMVC), vol. 31, pp. 1-31, **2007**. doi:10.5244/C.21.31.
- [76] Gasparini S. and al. Plane-Based Calibration of Central Catadioptric Cameras. Proceeding of the IEEE 12th International Conference on Computer Vision (ICCV), pp. 2146-2153, **2009**. doi: 10.1109/iccv.2009.5459336.
- [77] Puig L. and Guerrero J-J. Scale space for central catadioptric systems: Towards a generic camera feature extractor. Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1599–1606, **2011**. doi:10.1109/ICCV.2011.6126420.
- [78] Bastanlar Y., Puig L., Sturm P. and Guerrero J. DLT-Like Calibration of Central Catadioptric Cameras. OMNIVIS 8th Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras, pp. 14, Marseille, France, **2008**.
- [79] Toepfer C. and Ehlgen T. A Unifying Omnidirectional Camera Model and its Applications. Proceeding of the IEEE 11th International Conference on Computer Vision, pp. 1–5, **2007**. doi: 10.1109/ICCV.2007.4409207.
- [80] Geyer C. and Daniilidis K. Paracatadioptric camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(5), pp. 687–695, May **2002**. doi:10.1109/34.1000241.
- [81] Barreto J-P and Araujo H. Geometric properties of central catadioptric line images and their application in calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1327-1333, **2005**. doi:10.1109/TPAMI.2005.163.
- [82] Xianghua Y. and Hongbin Z. Simultaneously calibrating catadioptric camera and detecting line features using Hough transform. Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 412–417, **2005**. doi:10.1109/IROS.2005.1545166.
- [83] Wu Y., Li Y. and Hu Z. Easy Calibration for Para-catadioptric-like Camera. Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems, p. 5719–5724, **2006**. doi:10.1109/IROS.2006.282377.
- [84] Caglioti V., Taddei P., Boracchi G. and Gasparini S. Single-Image Calibration of Off-Axis Catadioptric Cameras Using Lines. Proceeding of the 7th Workshop on Omnidirectional Vision, pp. 1–6, Rio de Janeiro, Brazil, October **2007**. doi:10.1109/ICCV.2007.4409192.
- [85] Wu F., Duan F., Hu Z. and Wu Y. A new linear algorithm for calibrating central

- catadioptric cameras. *Pattern Recognition*, vol. 41(10), pp. 3166–3172, October **2008**. doi:10.1016/j.patcog.2008.03.010.
- [86] Kang S. Catadioptric self-calibration. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 201 - 207, **2000**. doi: 10.1109/CVPR.2000.855820
- [87] Micusik B. and Pajdla T. Structure from Motion with wide circular field of view cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1135–1149, **2006**. doi: 10.1109/TPAMI.2006.151.
- [88] Ramalingam S., Sturm P. and Lodha S-K. Towards Generic Self-Calibration of Central Cameras. In *proceeding of the 6th Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras*, pp. 20-27, Beijing, China, **2005**.
- [89] Espuny F. and Burgos G-J-I. Generic Self-calibration of Central Cameras from Two Rotational Flows. *International Journal of Computer Vision*, vol. 91, pp. 131-145, **2011**. doi:10.1007/s11263-010-0335-9.
- [90] Ying X. and Hu Z. Catadioptric camera calibration using geometric invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1260-1271, **2004**. doi:10.1109/TPAMI.2004.79.
- [91] Vasseur P. and Mouaddib E-M. Central catadioptric line detection. *British Machine Vision Conference (BMVC)*, pp. 57–66, Kingston, **2004**. doi:10.5244/C.18.8.
- [92] Aliaga D-G. Accurate catadioptric calibration for real-time pose estimation in room-size environments. *Proceedings of the 8th IEEE International Conference on Computer Vision*, vol. 1, pp. 127–134, **2001**. doi:10.1109/ICCV.2001.937508.
- [93] Puig L., Bastanlar Y., Sturm P., Guerrero J-J. and Barreto J. Calibration of central catadioptric cameras using a DLT-like approach. *International Journal of Computer Vision*, vol. 93, pp. 101-114, **2011**. doi:10.1007/s11263-010-0411-1.
- [94] Vandeportaele B., Cattoen M., Marthon P. and Gurdjos P. A New Linear Calibration Method for Paracatadioptric Cameras. *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, pp. 647–451, **2006**. doi:10.1109/ICPR.2006.121.
- [95] Frankot R-T. and Chellappa R. Method for Enforcing Integrability in Shape From Shading Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 439–451, **1988**. doi:10.1109/34.3909.
- [96] Morales L-P. *Omnidirectional Multiview Systems: calibration , features and 3D information*. Thèse soutenue en décembre **2011**.
- [97] Heikkilä J. and Silvén O. A Four-step Camera Calibration Procedure with Implicit Image Correction. *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1106-1112, Washington **1997**. doi:10.1109/CVPR.1997.609468.

- [98] Geyer C. and Daniilidis K. A unifying theory for central panoramic systems and practical implications. *Computer Vision Chapter*, vol. 1843, pp. 445–461, **2000**. doi:10.1007/3-540-45053-X_29.
- [99] Lourakis M. and Deriche R. Camera self-calibration using the singular value decomposition of the fundamental matrix: From point correspondences to 3D measurements. *Rapport de recherche INRIA n°3748*, **1999**.
- [100] Lavest J-M., Rives G. and Lapresté J-T. Dry camera calibration for underwater applications. *Machine Vision and Applications*, vol. 13, pp. 245-253, **2003**. doi:10.1007/s00138-002-0112.
- [101] Lavest J., Rives G. and Lapresté J. Underwater camera calibration. *Computer Vision Chapter*, vol. 1843, pp. 654-668, **2003**.
- [102] Gracias N. Underwater Video Mosaics as Visual Navigation Maps. *Computer Vision and Image Understanding*, vol. 79(1), pp. 66-91, July **2000**. doi:10.1006/cviu.2000.0848
- [103] Jung C. Laser spot detection-based computer interface system using autoassociative multilayer perceptron with input-to-output mapping-sensitive error back propagation learning algorithm. *Optical Engineering*, vol. 50(8), pp. 1-11, **2011**. doi:10.1117/1.3610985.
- [104] Berman S. and Stern H. Sensors for Gesture Recognition Systems. *IEEE Transactions on Systems, Man, and Cybernetics, (Applications and Reviews)*, vol. 42, pp. 277– 290, **2012**. doi:10.1109/TSMCC.2011.2161077.
- [105] Soetedjo A., Mahmudi A., Ashari M-I. and Nakhoda Y-I. Detecting Laser Spot in Shooting Simulator Using. *Journal on Smart Sensing and Intelligent*, vol. 7(1), pp. 423-441, **2014**.
- [106] Rouanet P., Oudeyer P-Y., Danieau F. and Filliat D. The Impact of Human-Robot Interfaces on the Learning of Visual Objects. *IEEE Transactions on Robotics*, vol. 29, pp. 525-541, **2013**. doi:10.1109/TRO.2012.2228134.
- [107] Tsujimura T., Minato Y. and Izumi K. Shape recognition of laser beam trace for human–robot interface. *Pattern Recognition Letters*, vol. 34, pp. 1928- 1935, **2013**. doi:10.1016/j.patrec.2013.03.023.
- [108] Shibata S., Yamamoto T. and Jindai M. Human-robot interface with instruction of neck movement using laser pointer. *Proceeding of the IEEE/SICE International Symposium on System Integration (SII)*, p. 1226–1231, **2011**. doi:10.1109/SII.2011.6147624.
- [109] Chávez F. et al. Hybrid laser pointer detection algorithm based on template matching and fuzzy rule-based systems for domotic control in real home environments. *Applied Intelligence*, vol 36, pp. 407–423, **2012**. doi:10.1007/s10489-010-0268-6.
- [110] Portugal-Zambrano C-E. and Mena-Chalco J-P. Robust range finder through a laser pointer and a webcam. *Electronic Notes in Theoretical Computer Science*, vol. 281, pp.

- 143-157, **2011**. doi:10.1016/j.entcs.2011.11.031.
- [111] Zerr B. and Stage B. Three-dimensional reconstruction of underwater objects from a sequence of sonar images. Proceedings of the 3rd IEEE International Conference on Image Processing, vol. 3, pp. 927–930, **1996**. doi:10.1109/ICIP.1996.560946.
- [112] Johnson-Roberson M., Pizarro O., Williams S-B. and Mahon I. Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys. Journal of Field Robotics, vol. 27, pp. 21-51, **2010**. doi:10.1002/rob.20324.
- [113] Gokhale R. and Kumar R. Analysis of three dimensional object reconstruction method for underwater images. International Journal of Scientific and Technology Research, vol. 2 (6), pp. 85-88, **2013**.
- [114] Chavez F., Fernandez F., Alcalá R., Alcalá-Fdez J. and Herrera F. Evolutionary learning of a laser pointer detection fuzzy system for an environment control system. Proceeding of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 256–263, **2011**. doi:10.1109/FUZZY.2011.6007528.
- [115] Blanco J-L., Gonzalez J. and Fern J-A. An Experimental Comparison of Image Feature Detectors and Descriptors applied to Grid Map Matching, Technical Report, pp. 1-13, University of Malaga, **2010**.
- [116] Ladha S-N., Smith-Miles K. and Chandran S. Multi-user natural interaction with sensor on activity. Proceeding of the 1st IEEE Workshop on User-Centered Computer Vision (UCCV) pp. 25–30, **2013**. doi:10.1109/UCCV.2013.6530804.
- [117] Kirstein C., Muller H. and Heinrich M. Interaction with a projection screen using a camera-tracked laser pointer. Multimedia Modeling Proceedings (MMM '98), p. 191–192, **1998**. doi:10.1109/mulmm.1998.723001.
- [118] Ahlborn B., Thompson D., Kreylos O., Hamann B. and Staadt O-G. A Practical System for Laser Pointer Interaction on Large Displays. Proceedings of the International Symposium on Virtual Reality Software and Technology (VRST'05), pp. 106–109, N-Y, USA, **2005**. doi:10.1145/1101616.1101637.
- [119] Brown M-S and Wong W-K-H. Laser pointer interaction for camera-registered multi projector displays. Proceeding of the IEEE International Conference on Image Processing, vol. 1, p. 913–916, **2003**. doi: 10.1109/ICIP.2003.1247112.
- [120] Davis J. and Chen X. Lumipoint: multi-user laser-based interaction on large tiled displays. Displays, vol. 23, pp. 205-211, **2002**. doi:10.1016/S0141-9382(02)00039-2.
- [121] Latoschik M-E and Bomberg E. Augmenting a Laser Pointer with a Diffraction Grating for Monoscopic 6DOF Detection. Journal of Virtual Reality and Broadcasting, vol. 4(14), pp. 1860-2037, **2007**. doi: 10.20385/1860-2037/4.2007.14.
- [122] Triebel R. and Burgard W. Improving Simultaneous Mapping and Localization in 3D

- Using Global Constraints. Proceeding of the National Conference on Artificial Intelligence (AAAI), vol. 20(3), pp. 1330-1335, Cambridge, London **2005**.
- [123] Liang H. and Kong B. A Shooting Training and Instructing System Based on Image Analysis. Proceeding of the IEEE International Conference on Information Acquisition, pp. 961–966, August **2006**. doi:10.1109/ICIA.2006.305866.
- [124] Soetedjo A. and Nurcahyo E. Developing of Low Cost Vision-Based Shooting Range Simulator. Proceeding of the International Journal of Computer Science and Network Security (IJCSNS), vol. 11, pp. 109-1, **2011**.
- [125] Kim N-W., Lee S-J., Lee B-G and Lee J-J. Vision Based Laser Pointer Interaction for Flexible Screens. Proceedings of the 12th International Conference on Human-Computer Interaction: Interaction Platforms and Techniques (HCII'07), vol. 4551, pp. 845–853, **2007**. doi:10.1007/978-3-540-73107-8_93.
- [126] Meško M. and Toth Š. Laser spot detection. Journal of Information, Control and Management Systems, vol. 11(1), pp. 35-42, **2013**.
- [127] Chavez F. et al. Genetic tuning of a laser pointer environment control device system for handicapped people with fuzzy systems. Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ), pp. 1–8, **2010**. doi: 10.1109/FUZZY.2010.5584876.
- [128] Atallah M-J. Faster image template matching in the sum of the absolute value of differences measure. IEEE Transactions on Image Processing, vol. 10, pp. 659–663, **2001**. doi:10.1109/83.913600.
- [129] Marghany M., Razlan M., Tahar B. and Hashim M. 3D stereo reconstruction using sum square of difference matching algorithm. Scientific Research and Essays, vol. 6, pp. 6404–411, **2011**. doi:10.5897/SRE11.1661.
- [130] Hii A-J-H., Hann C-E., Chase J-G. and Van-Houten E-E-W. Fast normalized cross correlation for motion tracking using basis functions. Computer Methods and Programs in Biomedicine, vol. 82, pp. 144–156, **2006**. doi:10.1016/j.cmpb.2006.02.007.
- [131] Nourain-Dawoud N., Belhaouari B. and Janier J. Fast Template Matching Method based Optimized Sum of Absolute Difference Algorithm for Face Localization. International Journal of Computer Applications, vol. 18, pp. 30-34, **2011**. doi:10.5120/2302-2912.
- [132] Scharstein D., Szeliski R. and Zabih R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Proceedings of the IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV), pp. 131–140, **2001**. doi:10.1109/SMBV.2001.988771.
- [133] Ma Y. et al. Epipolar Geometry and the Fundamental Matrix. Multiple View Geometry in Computer Vision, Chapter section, pp. 239-261, **2004**.
- [134] Aliakbarpour H. et al. An Efficient Algorithm For Extrinsic Calibration Between a 3D

- Laser Range Finder And A Stereo Camera For Surveillance. Proceeding of the 14th International Conference on Advanced Robotics, pp. 1-6, Munich, Germany, **2009**. doi:10.1016/S0263-2241(96)00065-6.
- [135] Scaramuzza D., Martinelli A. and Siegwart R. A Toolbox for Easily Calibrating Omnidirectional Cameras. Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5695–5701, **2006**. doi:10.1109/IROS.2006.282372.
- [136] Pagani A. and Stricker D. Structure from Motion using full spherical panoramic cameras. Proceeding of the International Conference on Computer Vision Workshops (ICCV), pp. 375–382, **2011**. doi:10.1109/ICCVW.2011.6130266.
- [137] Rituerto A., Puig L. and Guerrero J-J. Visual SLAM With An Omnidirectional Camera. Proceeding of the 20th International Conference on Pattern Recognition, pp. 348–351, **2010**. doi:10.1109/ICPR.2010.94.
- [138] Yamashita A., Hayashimoto E., Kaneko T. and Kawata Y. 3-D Measurement of Objects in a Cylindrical Glass Water Tank with a Laser Range Finder. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 1578-1583, **2003**. doi: 10.1109/IROS.2003.1248869.
- [139] Yamashita A., Fujii A. and Kaneko T. Stereo Measurement of Objects in Liquid and Estimation of Refractive Index of Liquid by Using Images of Water Surface. Proceeding of the IEEE International Conference on Robotics and Automation (ICRA 2008), pp. 974 - 979, **2008**. doi: 10.1109/ROBOT.2008.4543331.
- [140] Tatsuya-Naruse T-K. 3-D Measurement Of Objects In Water Using Fish-Eye Stereo Camera. Proceeding of the 19th IEEE International Conference on Image Processing, pp. 2773–2776, **2012**. doi:10.1109/ICIP.2012.6467474.
- [141] Bruno F., Bianco G., Muzzupappa M., Barone S. and Razionale A-V. Experimentation of structured light and stereo vision for underwater 3D. International Journal of Photogrammetry and Remote Sensing (ISPRS), vol. 66, pp. 508-518, **2011**. doi:10.1016/j.isprsjprs.2011.02.009.
- [142] Gluckman J-M., Nayar S-K. and Thoresz K-J. Real-Time Omnidirectional and Panoramic Stereo. Computer Vision and Image Understanding, vol. 1, pp. 299–303, **1998**.
- [143] Yi S. and Ahuja N. An omnidirectional stereo vision system using a single camera. Proceeding of the 18th International Conference on Pattern Recognition (ICPR), vol.4, pp. 861-865, **2006**. doi:10.1109/ICPR.2006.263.
- [144] Cordova-Esparza D-M. et al. A panoramic 3D reconstruction system based on the projection of based on the projection of patterns. International Journal of Advanced Robotic Systems, vol. 11(1), **2014**. doi:10.5772/58227.
- [145] Trucco E., Verri A. Introductory Techniques for 3-D Compute Vision. Chapter section

- (8.1), pp. 178-194, University of New-Jersey **1998**.
- [146] Heyden A and Pollefeys M. Multiple View Geometry. Chapter section (3), pp. 45-107, In Emerging Topics in Computer Vision, **2004**.
- [147] Brandou V. et al. 3D Reconstruction of Natural Underwater Scenes Using the Stereovision System IRIS. Proceeding of the IEEE Europe Conference, OCEANS 2007, pp. 1–6, **2007**. doi:10.1109/OCEANSE.2007.4302315.
- [148] Press W. et al. Numerical Recipes in C: The Art of Scientific Computing. Chapter, vol. 29, University of Cambridge **1987**. doi:10.2307/1269484.
- [149] Mc-Glamery B. Computer analysis and simulation of underwater camera system performance. Chapter, vol. 75 (2), University of California **1975**.
- [150] Jaffe J-S. Computer modeling and the design of optimal underwater imaging systems. IEEE Journal of Oceanic Engineering, vol. 15, pp. 101-111, **1990**. doi:10.1109/48.50695.
- [151] Schechner Y-Y and Karpel N. Clear underwater vision. Proceedings of the IEEE Computer Society Conference in Computer Vision and Pattern Recognition, vol. 1, pp. 536-543, **2004**. doi: 10.1109/CVPR.2004.1315078.
- [152] Sorensen S., Kolagunda A., Saponaro P. and Kambhamettu C. Refractive stereo ray tracing for reconstructing underwater structures. Proceeding in IEEE International Conference on Image Processing (ICIP), pp. 1712–1716, **2015**.
- [153] Bräuer-Burchardt C., Heinze M., Schmidt I., Kühmstedt P. and Notni G. Underwater 3D Surface Measurement Using Fringe Projection Based Scanning Devices. Sensors, vol. 16(1), pp.13, **2015**. doi:10.3390/s16010013.
- [154] Moore K-D., Jaffe J-S. and Ochoa B-L. Development of a new underwater bathymetric laser imaging system: L-Bath. Journal of Atmospheric and Oceanic Technology, vol. 17, pp. 1106-1117, **2000**. doi:10.1175/1520-0426.
- [155] Massot-Campos M., Oliver-Codina G., Kemal H., Petillot Y., Bonin-Font F. Structured light and stereo vision for underwater 3D reconstruction. Proceeding of the IEEE Conference, OCEANS-Genova, pp. 1–6, **2015**. doi: 10.1109/OCEANS-Genova.2015.7271433.
- [156] Bosch J., Gracias N., Ridao P. and Ribas D. Omnidirectional Underwater Camera Design and Calibration. Sensors, vol. 15, pp. 6033–6065, Switzerland **2015**. doi:10.3390/s150306033.
- [157] Caron G. and Eynard D. HySCaS: Hybrid Stereoscopic Calibration Software. SPIE newsroom in Electronic Imaging and Signal Processing, **2011**. doi:10.1117/2.1201106.003738.
- [158] Camera Calibration Toolbox for Matlab. Web Page: https://www.vision.caltech.edu/bouguetj/calib_doc/

- [159] Omnidirectional Calibration Toolbox. Web Page:
<http://www.robots.ox.ac.uk/~cmei/Toolbox.html>
- [160] OCamCalib: Omnidirectional Camera Calibration Toolbox for Matlab. Web Page:
<https://sites.google.com/site/scarabotix/ocamcalib-toolbox>.

Résumé

Les systèmes de navigation destinés plus particulièrement à l'exploration sous-marine suscitent de plus en plus l'intérêt de plusieurs géologues, roboticiens et scientifiques en vue de mieux connaître et caractériser les réseaux sous-marins. Pour des raisons de sécurité, de nouvelles technologies (Radar, Sonar, système de caméras,..) ont été développées pour remplacer les plongeurs humains. C'est dans ce cadre que s'intègre les travaux de cette thèse ayant comme objectifs la mise en œuvre d'un système de vision stéréoscopique pour l'acquisition d'informations utiles et le développement d'un algorithme pour la restitution de la structure 3D d'environnement confiné aquatique. Notre système est composé d'une paire de capteurs catadioptriques et d'une ceinture de pointeurs lasers permettant d'identifier des amers visuels de la scène et d'une plateforme pour l'exécution du processus de traitement des images acquises. La chaîne de traitement est précédée par une phase d'étalonnage effectuée hors-ligne pour la modélisation géométrique du système complet. L'algorithme de traitement consiste à analyser les images stéréoscopiques pour l'extraction des projections lasers 2D et à reconstruire leurs correspondants 3D en se basant sur les paramètres d'étalonnage.

La mise en œuvre du système complet sur une plateforme logicielle demande un temps d'exécution supérieur à celui exigé par l'application. La dernière partie des travaux présentés dans ce mémoire s'adressent à cette problématique, et proposent une solution permettant de simplifier le développement et l'implémentation d'applications temps-réel sur des plateformes basées sur un dispositif FPGA. La mise en œuvre de notre application a été effectuée et une étude des performances a été présentée tout en considérant les besoins de l'application et ses exigences de point de vue précision et rapidité.

Mots-clés :

- Robotique mobile
- Etalonnage
- FPGA
- Système stéréoscopique
- Reconstruction 3D
- Performance.
- Pointeurs lasers
- Implémentation