



HAL
open science

Algorithmes, architecture et éléments optiques pour l'acquisition embarquées d'images totalement focalisées et annotées en distance

Simon Emberger

► **To cite this version:**

Simon Emberger. Algorithmes, architecture et éléments optiques pour l'acquisition embarquées d'images totalement focalisées et annotées en distance. Micro et nanotechnologies/Microélectronique. Ecole nationale supérieure Mines-Télécom Atlantique, 2017. Français. NNT : 2017IMTA0055 . tel-01800048

HAL Id: tel-01800048

<https://theses.hal.science/tel-01800048>

Submitted on 25 May 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.



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

**UNIVERSITÉ
BRETAGNE
LOIRE**

THÈSE / IMT Atlantique

sous le sceau de l'Université Bretagne Loire

pour obtenir le grade de

DOCTEUR D'IMT Atlantique

Spécialité : Ondes et matériaux

École Doctorale 3M

Présentée par

Simon Emberger

Préparée au département Optique
Laboratoire L3I - CEA/LETI

Algorithmes, architecture et
éléments optiques pour
l'acquisition embarquées
d'images totalement focalisées
et annotées en distance

Thèse soutenue le 13 décembre 2017

devant le jury composé de :

Pierre Pellat-Finet

Professeur, Université de Bretagne-Sud / président

Michel Paindavoine

Professeur, Université de Dijon / rapporteur

Jérôme Primot

Directeur de recherche, Onera - Palaiseau / rapporteur

Laurent Alacoque

Ingénieur/Docteur, CEA/LETI - Grenoble / examinateur

Antoine Dupret

Ingénieur / Docteur, CEA Tech - Gif-sur-Yvette / examinateur

Isabelle Hardy

Ingénieur d'études, IMT Atlantique / examinatrice

Jean Louis de Bougrenet de la Tocnaye

Professeur, IMT Atlantique / directeur de thèse

Nicolas Fraval

Ingénieur/Docteur, Evosens - Plouzané / invité

Remerciements

Je remercie Pierre PELLAT-FINET, Professeur à l'Université de Bretagne Sud de me faire l'honneur de présider ce jury.

Je remercie Michel PAINDAVOINE, Professeur à l'université de Dijon, et Jérôme PRIMOT, directeur de recherche à l'ONERA d'avoir accepté la charge de rapporteur.

Je remercie à mon directeur de thèse, Jean Louis de BOUGRENET de la TOCNAYE, Professeur et chef du département d'optique à l'IMT-Atlantique, de m'avoir accueilli au sein de son département, pour ses conseils avisés durant ces travaux et pour relecture de mes essais dont ce manuscrit.

Je remercie également Isabelle HARDY, ingénieur de recherche au sein du département optique de l'IMT-Atlantique d'avoir accepté la charge d'examinatrice.

Un très grand merci à mon encadrant au CEA Grenoble, Laurent ALACOQUE, Docteur, pour ses conseils, sa bonne humeur et sa disponibilité, en particulier pour les heures passées à la relecture de mes essais dont ce manuscrit.

Un grand merci à mon co-encadrant, Antoine DUPRET, Docteur et chef du laboratoire LFIC au CEA de Saclay, pour ses conseils et le temps qui m'a été accordé.

Egalement un grand merci à Nicolas FRAVAL, Docteur de l'entreprise Evosens et spécialiste des lentilles à cristaux liquides, pour m'avoir formé à la réalisation de lentilles et pour toute l'aide qu'il m'a apportée au cours de ces travaux.

Je remercie chaleureusement Florian SCHOLKOPF, Samir ABBAS et Stephen LE GALL du département optique de l'IMT-ATANTIQUE pour leur aide en salle blanche lors de la fabrication de lentilles à cristaux liquide et leur bonne humeur au travail.

Un grand merci aux collègues et doctorants du département optique de l'IMT-Atlantique pour leur accueil.

Un très grand merci à Fabrice GUELLEC, chef du laboratoire Laboratoire Imageur Intégré Intelligent (L3I), ainsi qu'aux collègues et doctorants, pour leur convivialité au travail et leur soutien. Merci également aux collègues et doctorants du Département Architecture Conception Logiciel Embarqué (DACLE) du CEA.

Enfin je remercie tous mes proches, famille et amis, pour tout ce que nous partageons depuis si longtemps.

Résumé

L'acquisition de la profondeur d'une scène en plus de son image est une caractéristique souhaitable pour de nombreuses applications qui dépendent de l'environnement proche. L'état de l'art dans le domaine de l'extraction de profondeur propose de nombreuses méthodes, mais très peu sont réellement adaptées aux systèmes embarqués miniaturisés. Certaines parce qu'elles sont trop encombrantes en raison de leur système optique, d'autres parce qu'elles nécessitent une calibration délicate, ou des méthodes de reconstructions difficilement implantables dans un système embarqué. Dans cette thèse nous nous concentrons sur des méthodes à faible complexité matérielle afin de proposer une solution algorithmique et optique pour réaliser un capteur permettant à la fois d'extraire la profondeur de la scène, de fournir une évaluation de pertinence de cette mesure et de proposer des images focalisées en tout point. Dans ce sens, nous montrons que les algorithmes du type Depth from Focus (DfF) sont les plus adaptés à ces contraintes. Ce procédé consiste à acquérir un cube d'images multi-foyers d'une même scène pour différentes distances de focalisation. Les images sont analysées afin d'annoter chacune des zones de la scène d'un indice relatif à sa profondeur estimée. Cet indice est utilisé pour reconstruire une image nette en tout point.

Nous avons travaillé sur la notion de netteté afin de proposer des solutions peu complexes, uniquement basées sur des additions et comparaisons, et de fait, facilement adaptables pour un portage sur une architecture matérielle. La solution proposée effectue une analyse bidirectionnelle de contraste local puis combine les meilleures estimations de profondeur en fin de traitement. Elle se décline en trois approches avec une restriction de la complexité de plus en plus forte et ainsi une aptitude de plus en plus marquée pour l'embarqué. Pour chaque méthode, des cartes de profondeurs et de confiances sont établies, ainsi qu'une image totalement focalisée constituée d'éléments issus de l'ensemble du cube multi-foyers. Ces approches sont comparées en qualité et en complexité à d'autres méthodes de l'état de l'art de complexité similaire. Une architecture est proposée pour une implantation matérielle de la solution la plus prometteuse.

La conception de ces algorithmes soulève le problème de la qualité d'image. Il est en effet primordial d'avoir une évolution remarquable du contraste ainsi qu'une invariance de la scène lors de la capture du cube multi-foyers. Un effet très souvent négligé dans ce type d'approche est le zoom parasite provoqué par la lentille responsable de la variation de la mise au point. Ce zoom de focalisation fragilise l'aspect invariance de la scène et provoque l'apparition d'artefacts sur les trois informations Profondeur, Image et Confiance. La recherche d'optiques adaptées au DfF constitue donc un second axe de ces travaux. Nous avons évalué des lentilles

liquides industrielles et des lentilles modales expérimentales à cristaux liquides nématiques conçues durant cette thèse. Ces technologies ont été comparées en termes de rapidité, de qualité d'image, d'intensité de zoom de focalisation engendré, de tension d'alimentation et enfin de qualité des cartes de profondeur extraites et des images totalement focalisées reconstruites.

La lentille et l'algorithme répondant le mieux à cette problématique DfF embarqué ont ensuite été évalués via le portage sur une plateforme de développement CPU-GPU permettant l'acquisition d'images et de cartes de profondeurs et de confiances en temps réel.

Table des matières

Introduction.....	3
Chapitre 1 Etat de l'art.....	5
1.1 Méthodes d'extraction de profondeur d'une scène	5
1.1.1 Introduction.....	5
1.1.2 Technologies actives.....	6
1.1.3 Technologies passives	9
1.1.4 Combinaison de technologies	20
1.1.5 Synthèses des méthodes d'extraction de profondeur	20
1.1.6 Conclusion sur les méthodes d'extraction de profondeur	20
1.2 Technologies Optiques	22
1.2.1 Introduction.....	22
1.2.2 Lentilles liquides	22
1.2.3 Technologie à cristaux liquides	24
1.2.4 Conclusion sur les technologies optiques	29
Chapitre 2 Développements algorithmiques et architecture pour l'extraction de la profondeur	31
2.1 Introduction.....	31
2.2 Avant-propos sur l'évaluation de la netteté.....	32
2.2.1 Analyse de l'évolution du flou	33
2.2.2 Choix d'un critère de netteté	34
2.2.3 Influence de la mise au point sur l'invariance de l'image ou « zoom de focalisation »	38
2.3 Présentation du banc de test	40
2.4 Algorithme d'évaluation de distance par mesure bidirectionnelle de contraste local	41
2.4.1 Approche intra-image haute résolution.....	43
2.4.2 Approche intra-image basse résolution	50
2.4.3 Approche intra-image basse résolution par fenêtres de traitement non recouvrantes	53
2.5 Comparaison des algorithmes de Depth from Focus	56
2.5.1 Comparaison des cartes de profondeur	57
2.5.2 Comparaison des images All in Focus.....	60
2.6 Amélioration du rendu	63

2.7	Proposition d'architectures	65
2.7.1	Architecture de construction de pixel RGB-Z-C par blocs distincts.....	67
2.7.2	Evaluation de la complexité de l'architecture.....	72
2.8	Conclusion	75
Chapitre 3 Développements optiques et réalisation de lentilles modales à cristaux liquides nématiques 77		
3.1	Introduction.....	77
3.2	Éléments dimensionnants pour l'optique	78
3.2.1	Importance de la profondeur de champ	78
3.2.2	Invariance de la scène	81
3.2.3	Qualité optique.....	81
3.3	Lentilles modales expérimentales à cristaux liquides nématiques	82
3.3.1	Potentiel de variation	83
3.3.2	Paramètres qui délimitent les choix à faire.....	84
3.3.3	Calcul de l'épaisseur de la couche de cristal liquide	85
3.3.4	Etapes de réalisation	86
3.4	Caractérisation de lentille modulable	89
3.4.1	Présentation du banc de test	90
3.4.2	Caractérisation de la loi de commande.....	92
3.4.3	Mesure des vitesses de stabilisation du foyer.....	94
3.4.1	Analyse de la diffusion.....	99
3.4.2	Mesure de zoom de focalisation	103
3.4.3	Evaluation des lentilles avec un algorithme DfF.....	104
3.4.4	Observations.....	106
3.5	Synthèse des résultats de mesure.....	107
3.6	Conclusion	108
Chapitre 4 Prototype expérimental d'acquisition d'images 3D..... 111		
4.1	Introduction.....	111
4.2	Description de la plateforme	112
4.2.1	Point de vue matériel	112
4.2.2	Point de vue logiciel	114
4.3	Portage de l'algorithme	114
4.4	Résultats obtenus	117
4.4.1	Influence de la taille des fenêtres de traitement.....	117

4.4.2	Mesure des temps de traitement.....	119
4.4.3	Evolution temporelle des images totalement focalisées et leurs cartes de profondeur et de confiance associées	120
4.4.4	Influence de l'exposition	121
4.4.5	Sensibilité au mouvement.....	122
4.5	Conclusions sur le portage matériel	124
	Conclusion et perspective	125
	Références.....	129
	Annexe 1 : Etude du cercle de confusion	139
	Annexe 2 : Evaluation de distance par inflexion de la luminosité	145
	Annexe 3 : Etude de la profondeur de champ	149
	Annexe 4 : Calcul de l'épaisseur d'une lentille à cristaux liquides.....	155
	Annexe 5 : Etapes de fabrication d'une lentille à cristaux liquides	159
	Annexe 6 : Réalisation d'un driver de lentille à cristaux liquides	165

Contributions

Poster aux Journées de l'Imagerie Optique Non Conventionnelle (JIONC), Paris 2016

Présentation au Workshop on Architecture of Smart Cameras (WASC), Dijon, 2016

Article de revue à Electronic Imaging, San Francisco, 2017 [1]

Article de revue à International Conference on Distributed Smart Cameras (ICDSC), Stanford, 2017 [2]

Article de revue à Electronic Imaging, San Francisco, 2018 [3]

Article de Journal dans Applied Optics [4]

Brevet en cours de dépôt

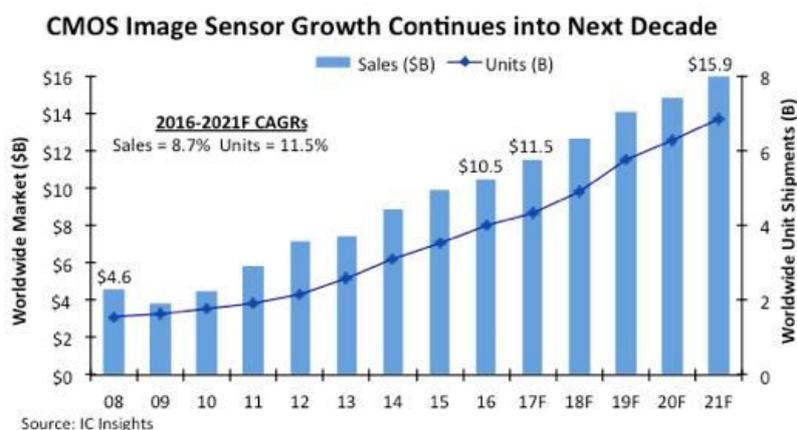
Algorithmes,
architecture et éléments
optiques pour
l'acquisition embarquée
d'images totalement
focalisées et annotées
en distance

Algorithms, architecture and optics components for embedded all in focus and distance-annotated image acquisition system

Introduction

Le marché de l'électronique embarquée ne cesse d'augmenter depuis quelques années. Avec le développement de l'internet des objets (IoT) et des véhicules autonomes, cet engouement va continuer à prendre de l'ampleur, et de nouvelles applications plus spécifiques, moins coûteuses et miniaturisées vont apparaître. Le monde des capteurs étant le premier maillon de la chaîne d'interaction entre le réel et le virtuel, il se retrouve particulièrement sollicité pour le développement de toutes ces nouvelles applications.

Les capteurs d'images sont très certainement les capteurs les plus polyvalents qui soient, ils permettent par exemple de reconnaître des gestes, des mouvements, ou encore les formes et même les émotions, et tout cela à la fois, la seule limite étant la capacité d'interprétation du système hôte. On a d'ailleurs vu une augmentation rapide de l'utilisation de ces capteurs autant par l'industrie que par le grand public, notamment avec la démocratisation des smartphones. Et selon une étude d'IC insight (graphe ci-dessous), le marché des imageurs va encore grandement s'intensifier.



Les marchés porteurs ne sont plus seulement les smartphones mais bien l'automobile, la sécurité, la vision artificielle, le médical, la réalité virtuelle, la robotique, soit des applications capables de traiter elles-mêmes des images pour fournir de nouvelles informations et en faire usage intelligemment.

Dans le domaine des caméras, la grande tendance est de fournir plus d'informations qu'une simple image. Cela permet à la fois de limiter la transmission de données à seulement ce qui est important, de baisser la consommation des systèmes lorsque les traitements sont effectués par des architectures dédiées et optimisées, ce qui diminue également les latences de calcul, et permet aux systèmes contraints en capacité de calcul d'avoir accès à des informations qu'ils n'auraient jamais pu extraire d'eux-mêmes.

Avec le développement des objets autonomes, la connaissance de l'environnement proche devient capitale. L'acquisition de la profondeur d'une scène en plus de son image est alors une caractéristique plus que souhaitable pour de tels systèmes. La miniaturisation des objets fait apparaître des contraintes d'encombrement physique des composants et amène à la miniaturisation des systèmes d'acquisition et donc à une simplification des algorithmes d'extraction de la profondeur et un rapprochement des organes de calcul au plus près de la source d'acquisition.

Dans cette thèse, nous nous sommes intéressés à l'acquisition de la profondeur d'une scène, en ayant à l'esprit l'aspect miniaturisation afin de proposer une solution adaptée pour l'automobile, les smartphones et la robotique. De plus, avec la forte tendance à la fusion de capteurs, la redondance d'information est devenue primordiale et l'autoévaluation de la qualité de ses propres données un avantage certain pour une reconstruction fidèle de l'environnement. C'est pourquoi nous avons également travaillé sur la pertinence des données de profondeur et proposons de calculer un coefficient de confiance pour pondérer les évaluations de distance. Enfin, l'optimisation d'imageur nécessite également une parfaite adaptation des composants optiques qui le composent. Nous avons donc mené une étude optique puis réalisé et utilisé des lentilles en adéquation avec la solution d'acquisition de profondeur proposée.

L'objectif de cette thèse est de proposer une solution algorithmique et optique pour réaliser un prototype de capteur de très faible volume permettant à la fois d'extraire la profondeur de la scène, d'autoévaluer la pertinence de cette mesure et de fournir des images sans zones floues.

Le premier chapitre de ce manuscrit présente un état de l'art des méthodes d'acquisition de l'environnement en trois dimensions.

Dans le second, une solution algorithmique, du type Depth from Focus, basée sur une analyse bidirectionnelle de contraste local est proposée. Elle se décline en trois approches, avec une restriction de la complexité de plus en plus forte et ainsi une aptitude de plus en plus marquée pour l'embarqué. Nous les comparons avec des méthodes de l'état de l'art de complexité similaire et proposons une architecture d'implantation matérielle.

Le troisième chapitre présente l'étude optique à travers la réalisation de lentilles modales à cristaux liquides nématiques, puis leur comparaison avec des lentilles liquides industrielles au regard de critères propres à l'imagerie et au calcul de profondeur.

Le quatrième et dernier chapitre réunit algorithmique et optique pour la réalisation d'un prototype expérimental temps réel d'acquisition de profondeur, de calcul de la pertinence et de construction d'images totalement focalisées.

Chapitre 1 Etat de l'art

1.1	Méthodes d'extraction de profondeur d'une scène	5
1.1.1	Introduction.....	5
1.1.2	Technologies actives.....	6
1.1.3	Technologies passives	9
1.1.4	Combinaison de technologies	20
1.1.5	Synthèses des méthodes d'extraction de profondeur	20
1.1.6	Conclusion sur les méthodes d'extraction de profondeur	20
1.2	Technologies Optiques	22
1.2.1	Introduction.....	22
1.2.2	Lentilles liquides	22
1.2.3	Technologie à cristaux liquides	24
1.2.4	Conclusion sur les technologies optiques	29

1.1 Méthodes d'extraction de profondeur d'une scène

1.1.1 Introduction

L'acquisition de la profondeur d'une scène en plus de son image est fondamentale pour de nombreuses applications notamment dans le domaine de la robotique, et plus généralement la navigation autonome, ainsi que le médical, le multimédia ou encore la sécurité. En effet, la perception tridimensionnelle (3D) de l'environnement est primordiale pour le concept d'interaction avec le monde réel dont le sens de la vue est la base pour la grande majorité des êtres vivants.

Les méthodes d'extraction d'informations 3D [5] sont étudiées depuis de nombreuses années, mais ce domaine reste un challenge notamment avec l'arrivée de l'internet des objets et de tous les petits modules autonomes. Ce chapitre présente un panorama représentatif de ces méthodes d'appréhension de l'environnement à travers deux grandes thématiques que sont les technologies actives et passives (Figure 1).

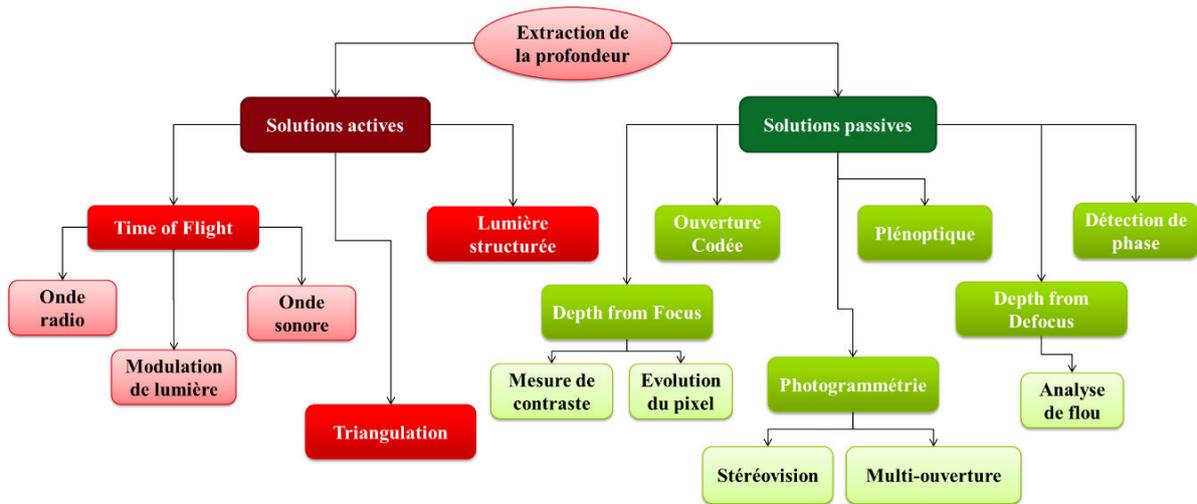


FIGURE 1 : ARBRE ETAT DE L'ART

1.1.2 Technologies actives

Les technologies classées actives [6] analysent la scène par l'intermédiaire d'un signal radiofréquence, lumineux, ou encore sonore qui permet de scanner l'environnement.

Parmi les techniques actives d'extraction de carte de profondeur, on distingue 3 principes récurrents : la triangulation active, la mesure du temps de vol et les technologies à projection de lumière structurée. Quelques autres méthodes existent, notamment l'interférométrie ou encore des méthodes d'analyse d'ombre (Shape From Shading), mais elles sont peu adaptées aux applications embarquées.

1.1.2.1 Méthodes de triangulation

Les méthodes de triangulation actives utilisent un laser asservi en position pour balayer l'ensemble de la scène à modéliser (Figure 2).

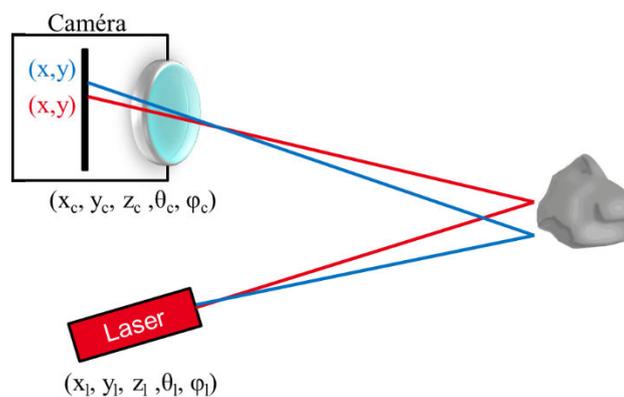


FIGURE 2 : ILLUSTRATION DE LA TRIANGULATION ACTIVE

Pour chaque position du laser (différentes orientations sur la Figure 2), une image est enregistrée par un capteur matricé dont la position relative par rapport laser est parfaitement connue. L'analyse des différentes positions du spot lumineux sur les images permettent de calculer par triangulation la distance aux objets [7] [8].

Ces méthodes nécessitent un balayage précis du laser sur la scène ce qui peut prendre beaucoup de temps lorsqu'une haute résolution spéciale est exigée. La scène doit pouvoir être considérée comme fixe par rapport à la vitesse de déplacement du laser et à la cadence image de la caméra.

1.1.2.2 Le Time of flight, ou mesure du temps de vol.

Le Time of flight (ToF), ou mesure du temps de vol, est une méthode d'extraction de distance par mesure du temps entre l'émission d'un faisceau et son retour sur un capteur, après rebond sur les différents objets de la scène (Figure 3).

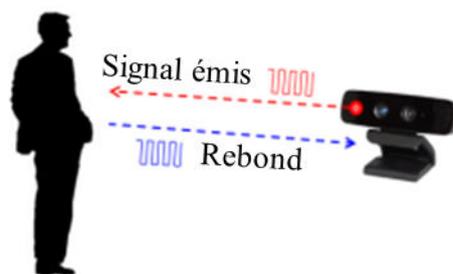


FIGURE 3 : ILLUSTRATION DU TOF

Le signal émis peut être sonore : un transducteur émet une impulsion, généralement en haute fréquence pour augmenter la précision tout en restant imperceptible par l'homme, et un signal de synchronisation qui permet de lancer un compteur jusqu'à réception du signal rebond. La vitesse du son dans le milieu étant connue, la distance peut être aisément calculée. C'est un principe historiquement utilisé pour les sonars de bateaux [9] [10] ainsi que dans le secteur de la médecine pour les échographies [11]. Le principe à base d'ultrasons est également utilisé dans des capteurs simples de proximité, pour de l'évitement d'obstacle en robotique ou encore pour l'autofocus de certains appareils photos numériques. Dans l'air, ce type de technologie est limité à quelques mètres à cause des bruits ambiants et du manque de directivité.

Les RADAR (Radio Detection and Ranging) utilisés depuis la seconde guerre mondiale, mesurent le temps de propagation d'impulsions radio à haute fréquence pour détecter des objets et créer une carte d'obstacles [12]. Ils sont cependant relativement directs et de ce fait doivent soit avoir une antenne mobile pour réussir à quadriller une zone de l'espace, soit être constitués de nombreuses antennes, ce qui les rend assez encombrants.

Avec l'amélioration de la rapidité des circuits électroniques, il est devenu possible de détecter le temps de vol de rayons lumineux. Dans la littérature récente le terme ToF semble désigner uniquement les méthodes de mesure de la profondeur basées sur l'émission de signaux lumineux dans le proche infrarouge (~850nm). Quelques-unes de ces méthodes reprennent l'architecture utilisée pour la triangulation active en remplaçant l'analyse de la position du spot lumineux par une mesure du temps entre l'émission et la réception d'une impulsion laser [7]. L'obtention de la distance est plus directe mais garde l'inconvénient du laser mobile. Cela a donné lieu à l'invention du LIDAR (Light Detection and Ranging) de nos jours très utilisé par les véhicules autonomes [13].

D'autres techniques [14] [15] consistent à éclairer l'ensemble de la scène en même temps et à mesurer la distance pour chaque pixel grâce à l'acquisition de la lumière réfléchi. Deux principes sont couramment utilisés [16]. Un premier consiste à faire une mesure directe du temps de vol grâce à un compteur ultra-rapide et un éclairage par lumière pulsée. Ces compteurs sont cependant difficiles à intégrer sur silicium ; à titre d'exemple, une précision de 1mm nécessite une horloge de l'ordre de 151 GHz. Le second permet une mesure directe de phase d'un éclairage modulée par lecture d'intensité lumineuse. Cette solution est plus simple à mettre en œuvre mais d'une part, elle est sensible à la luminosité ambiante, d'autre part, la portée est limitée à cause de l'ambiguïté de la phase ; c'est particulièrement vrai lors de modulation à hautes fréquences. Comment savoir à quelle période du signal émis le signal rebond correspond ? Ces méthodes fournissent une mesure directe de la profondeur, avec peu de calculs et pour un très bon résultat (Figure 4). En revanche la résolution des capteurs actuels est assez limitée (inférieure au VG, soit 640×480).



FIGURE 4 : EXEMPLE DE RESULTAT DE TOF [16]

1.1.2.3 Analyse de la lumière structurée

L'analyse de la profondeur par lumière structurée consiste à projeter un motif connu sur la scène puis à analyser ses distorsions en fonction des formes et de la disposition des objets présents dans le champ de vision [8] [17] [18]. C'est le principe utilisé par la première version de la Kinect de Microsoft [19], bien connue du monde de la recherche. Pour une bonne évaluation de profondeur, la projection successive de différents motifs est nécessaire (Figure 5), ce qui limite la vitesse d'extraction et impose un sujet relativement fixe. De plus, les interférences d'éclairage sont problématiques et restreignent ce type de méthode aux

environnements intérieurs. En revanche, la profondeur peut être évaluée avec une très grande résolution (\sim mm) [6].

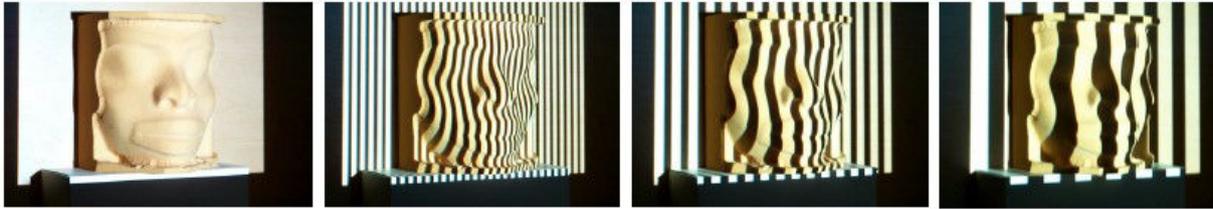


FIGURE 5 : MESURE DE PROFONDEUR PAR PROJECTION DE MOTIFS

1.1.3 Technologies passives

Les méthodes passives d'extraction de profondeur utilisent seulement l'intensité des pixels pour déduire la disposition spatiale des objets dans une scène. Les systèmes de mesure n'émettent donc aucun signal ce qui permet d'être plus discret, de réduire la consommation du système et également d'éviter les potentielles perturbations du signal rebond par les signaux déjà présents dans le milieu extérieur. Cela permet également de limiter le nombre de pannes potentielles du système en supprimant la nécessité d'un module émetteur. Ces méthodes nécessitent cependant qu'un minimum de détails soit discernable. Les zones d'une scène avec peu de contraste sont tout simplement impossibles à estimer.

Nous avons classé les systèmes passifs d'extraction de carte de profondeur suivant 5 grandes catégories : les systèmes multi-ouvertures (photogrammétrie), utilisant plusieurs caméras, ou plusieurs points de vue, pour capturer des images et extraire une carte de profondeur ; les caméras plénoptiques qui permettent un échantillonnage de l'ensemble du champ lumineux ; les systèmes à ouverture codée qui filtrent optiquement le flux de photons ; la détection de phase qui offrent la possibilité d'avoir un système proche de la stéréovision sur un unique capteur, notamment grâce au développement récent des pixels à détection de phase ; les approches Depth from Defocus (DfD) dont l'objet est une analyse locale d'une seule image dans le but de quantifier le degré de flou par zone et ainsi d'évaluer les distances ; et les méthodes Depth from Focus (DfF) qui complètent le DfD par l'analyse de plusieurs images d'une même scène à différents niveaux de focalisation afin d'évaluer le degré de netteté par pixel ou zone de pixels. Ces trois dernières catégories sont d'ailleurs très liées aux problématiques d'autofocus des appareils photo numériques (APN).

1.1.3.1 Photogrammétrie

La photogrammétrie est la science qui consiste à modéliser un environnement à partir de la parallaxe obtenue entre des images capturées avec des points de vue différents. Ces images

peuvent être acquises simultanément par un ensemble de caméras [20], ou à l'issue d'un déplacement d'une unique caméra comme cela est proposé avec les algorithmes SLAM [21].

Les systèmes de mesure photogrammétriques fonctionnent dans l'ensemble selon le principe illustré Figure 6. Plusieurs images sont acquises puis analysées de sorte à faire une triangulation passive des éléments de la scène. Dans tous les cas, la position des caméras et leur orientation doivent être connues ou déductibles. Cela nécessite une première étape de calibration du système [22].

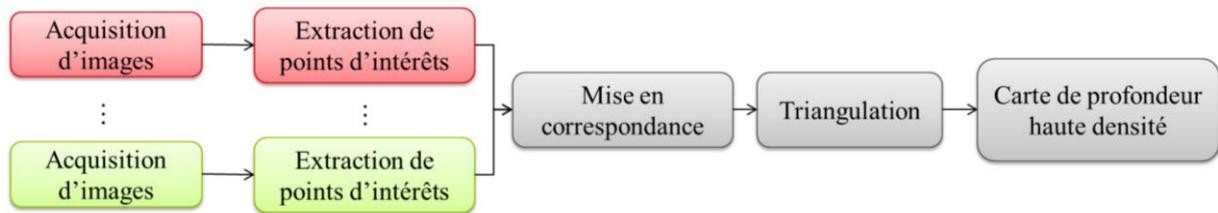


FIGURE 6 : FONCTIONNEMENT DES METHODES MULTI-OUVERTURES

Parmi ces systèmes, les caméras stéréoscopiques sont les plus connues et les plus utilisées. Elles découlent directement du fonctionnement de la vision humaine. Deux images d'une même scène sont capturées par deux caméras avec deux points de vue différents, ce qui permet ensuite d'extraire la profondeur de la scène par triangulation [23]. La Figure 7 illustre en 2 dimensions le principe de la stéréovision.

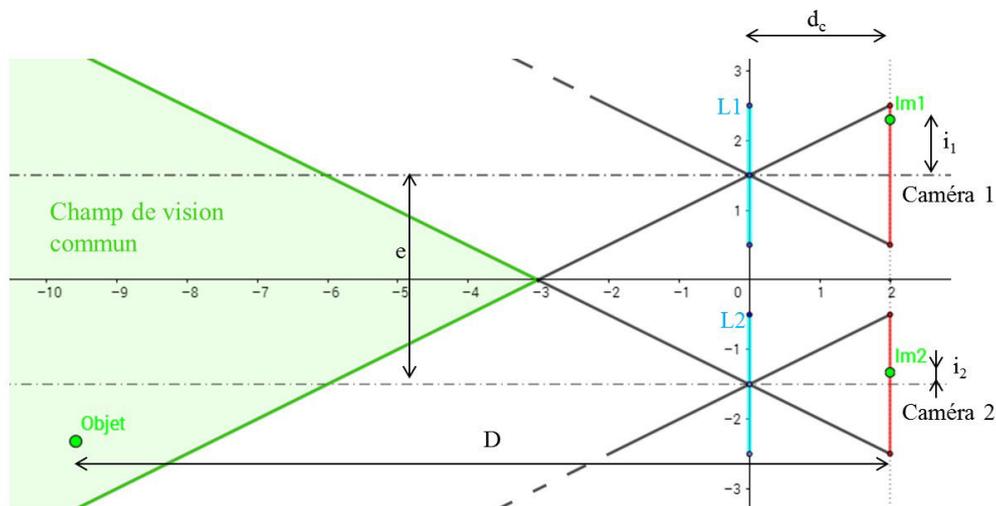


FIGURE 7 : PRINCIPE DE LA STEREOVISION

La considération des positions i_1 et i_2 des images d'un même objet sur chacun des capteurs et l'espacement connu e entre les 2 caméras permettent de calculer la distance via l'équation (1.1)

$$D = \frac{e}{i_1 - i_2} d_c \quad (1.1)$$

La principale difficulté de la stéréovision consiste à trouver et fiabiliser les correspondances entre les deux images [24]. Cela demande l'extraction de points d'intérêts et de descripteurs locaux qui servent à retrouver les coordonnées de mêmes objets dans les deux images par mesure de disparité. Une calibration minutieuse des caméras permet d'utiliser les modèles géométriques épipolaires qui limitent les zones de recherche lors de la mise en correspondance.

De nombreux travaux s'emploient à rendre plus accessibles les approches stéréoscopiques pour les systèmes embarqués, que ce soit par l'optimisation des algorithmes d'appariement [25], par des propositions d'ASIC [26] [27] ou encore d'architectures dédiées [28] [29] principalement basées sur l'utilisation de FPGA. Les méthodes stéréoscopiques ont l'avantage d'acquérir un champ de la scène relativement grand. En revanche, la fiabilisation de la mise en correspondance nécessite beaucoup de ressources calculatoires et une calibration précise des caméras ce qui, en plus de leur taille incompressible due à l'entraxe des caméras, complexifie le portage vers des systèmes miniatures.

1.1.3.2 Les caméras plénoptiques

Une caméra plénoptique (de *plenus*, complet ou plein, et *optic*) est un dispositif inventé dans le but de capturer la totalité des informations contenues dans un flux lumineux. Elle mesure à la fois la position et l'angle d'incidence des rayons lumineux sur le plan du capteur [30] [31]. Cela confère des propriétés très intéressantes, comme la possibilité de faire la mise au point par post traitement, de corriger les aberrations chromatiques ou encore de reconstruire la scène capturée en trois dimensions.

L'objet de la photographie classique est de focaliser sur un élément en particulier. Les faisceaux qui en émanent se retrouvent sur un même pixel d'un capteur d'image (cible 1 sur la Figure 8). Un appareil plénoptique permet d'enregistrer chaque pinceau lumineux rouge de la cible 2 séparément. L'intensité lumineuse d'un pinceau est associée aux coordonnées du point dans le plan du capteur, mais aussi à ses angles d'incidences. En retrouvant les pinceaux émis par une même source ponctuelle, il devient possible de synthétiser une image refocalisée par somme des pixels correspondants (Pixels touchés par le faisceau rouge).

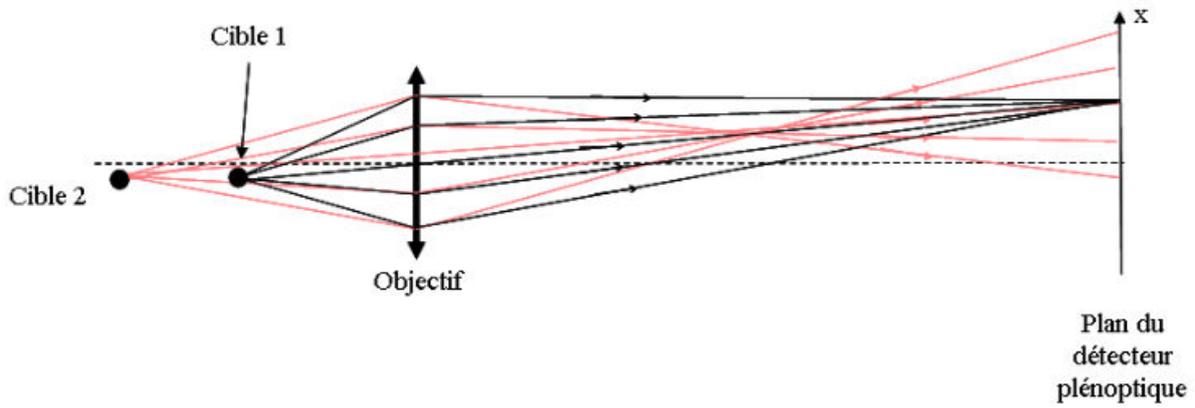


FIGURE 8 : ILLUSTRATION DU CONCEPT PLENOPTIQUE

La difficulté consiste à retrouver les points d'impacts de ces pinceaux sources. Avec un objectif standard, c'est impossible, car des rayons provenant de différentes sources se retrouvent sur un même pixel sans pouvoir être isolés. Pour répondre à ce problème, une matrice de microlentilles est placée entre le capteur et l'objectif (Figure 9).

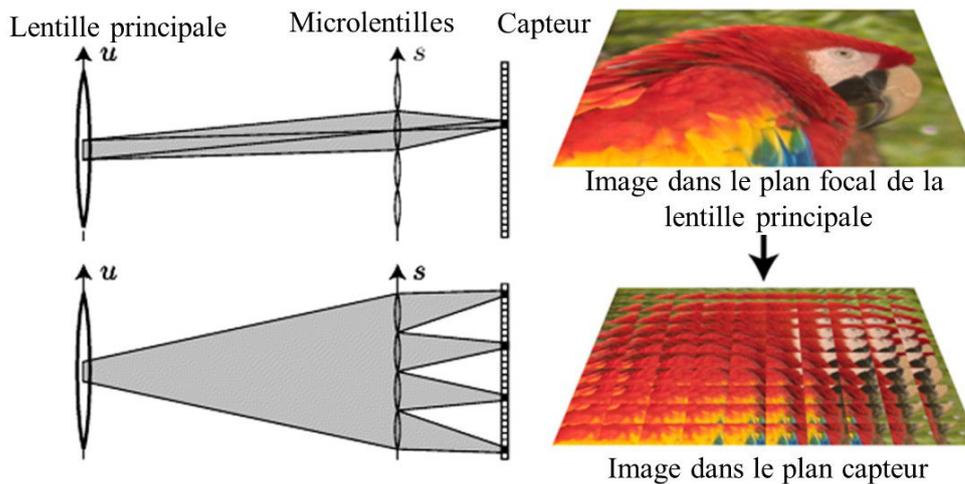


FIGURE 9 : PRINCIPE PLENOPTIQUE

Chaque microlentille correspond à un macro-pixel constitué de N^2 pixels qui permettent de capturer séparément les différents pinceaux et ainsi d'effectuer un échantillonnage angulaire du champ lumineux. Malheureusement, la résolution de l'image reconstruite est souvent limitée à la taille de cette matrice de microlentilles.

Le concept de base est né de Lippmann en 1908 [32]. Ce n'est qu'en 1991, qu'Adelson et Bergen [33] concrétisent mathématiquement la fonction plénoptique (1.2) par une analyse logique des paramètres qui semblaient nécessaires pour une reconstruction complète d'un flux lumineux.

$$P(\theta, \varphi, \lambda, t, V_x, V_y, V_z) \quad (1.2)$$

P représente la densité de lumière provenant d'un faisceau lumineux traversant une lentille observé depuis une position (V_x, V_y, V_z) . (ϑ, φ) sont les angles polaires d'incidences de ce même faisceau au temps t pour une longueur d'onde λ .

Très vite, en 1992, l'intérêt pour extraire la profondeur est montré [34] mais avec des résultats qui restent mitigés et qui nécessitent beaucoup de calculs.

En considérant à la fois l'aspect acquisition d'image et le dispositif nécessaire pour capter les données plénoptiques, Marc Levoy et Pat Hanrahan [35] ont démontré l'intérêt de représenter la fonction plénoptique avec seulement 4 dimensions (1.3). (x, y) est le point d'impact d'un faisceau sur le plan capteur et (u, v) un autre point de ce même faisceau dans le plan de la lentille. La longueur d'onde est connue grâce au filtre de Bayer (matrice colorée disposée au-dessus des pixels) et le temps par l'instant de capture.

$$P(x, y, u, v) \tag{1.3}$$

Partant de cette équation, ce sont les travaux de Ren Ng [36] [37], puis Lytro, qui ont relancé l'engouement pour le plénoptique en 2005. Son approche par analyse spectrale a permis de simplifier grandement la complexité calculatoire pour la génération d'image. D'autres travaux ont suivi, notamment [38] dans lesquelles la matrice de lentilles est remplacée par un réseau optique, moins coûteux. Cela supprime quelques aberrations optiques mais diminue la luminosité et fait apparaître des effets supplémentaires de diffraction. Les travaux les plus aboutis en termes de reconstruction 3D sont ceux menés par la société Raytrix [39] [40]. A l'instar de la stéréoscopie, ces auteurs utilisent une méthode de comparaison pour faire correspondre les informations redondantes obtenues grâce aux microlentilles, puis trianguler la scène et remonter à la profondeur.

Des méthodes de capture multi-ouvertures ont également été développées [41] et intégrées sur un seul capteur [42]. Cela revient en quelque sorte à utiliser une matrice de lentille très proche de la lentille principale et ainsi d'avoir des imageries plénoptiques contenant la quasi-totalité de la scène. La camera PiCam [41] constitue une caméra plénoptique miniature adaptable aux tablettes et smartphones, donc répondant parfaitement aux contraintes d'encombrement de l'embarqué. Puisqu'elle capture 16 fois la même scène avec des points de vue légèrement différents, des considérations de parallaxe permettent de remonter à l'information de profondeur.

Globalement les méthodes plénoptiques donnent des résultats très corrects. Mais elles souffrent d'une forte complexité de calcul, que ce soit pour la reconstruction d'image conventionnelle ou d'image en 3 dimensions. De plus la résolution finale est souvent bien plus faible que la résolution effective de la matrice de pixel ; plus la résolution angulaire augmente, plus la résolution spatiale diminue. A titre d'exemple, pour une résolution initiale de 100MPix, la résolution finale est au mieux de 4MPix avec l'appareil haut de gamme Lytro.

1.1.3.3 Système à ouverture codée

Le développement de systèmes mixtes (optique et électronique) pour la capture d'image nous amène aux technologies exploitant les ouvertures codées. Il s'agit de positionner devant la lentille ou le capteur un masque qui permet d'effectuer des prétraitements optiques.

Des chercheurs du MIT ont inventé une ouverture codée qui leur permet à la fois de capturer une photo totalement focalisée après reconstruction d'image et correction de flou, et de construire une carte de profondeur [43]. Le but de cette étude était d'optimiser un filtre codé pour la restitution de la profondeur : il y a un compromis à faire entre résolution spatiale et résolution de profondeur. Les auteurs ont notamment construit un modèle statistique de génération d'image qui leur a permis de tester un grand nombre de filtres et d'optimiser la forme de l'ouverture pour la mesure de distance. L'obtention d'images totalement focalisées et de la profondeur permet de synthétiser des images focalisées dans un plan particulier à l'instar des systèmes plénoptiques.

Anita Sellent et Paolo Favaro ont proposé une technique de génération de masques d'ouverture codée pour de la stéréovision et le DfD [44]. L'idée est de chercher l'ouverture qui maximise les différences entre des niveaux de flou correspondant à des distances précises. Une fois le critère de différenciation trouvé, un algorithme cherche l'ouverture par simulation de rendu, calcul de PSF (Point Spread Fonction, ou fonction d'étalement d'un point lumineux) et convolution de scènes. La thèse de master [45] donne plus de détails sur l'optimisation de la forme des ouvertures.

Comme pour l'imagerie plénoptique, le fait de capturer une image optiquement filtrée nécessite une capacité de calcul conséquente pour la reconstruction d'image et freine le développement de telle solution pour l'embarqué.

1.1.3.4 Détecteur de phase

Les détecteurs de phase ont été développés pour les dispositifs d'autofocus des APN. La Figure 10 montre le fonctionnement d'un tel système dans un appareil réflexe.

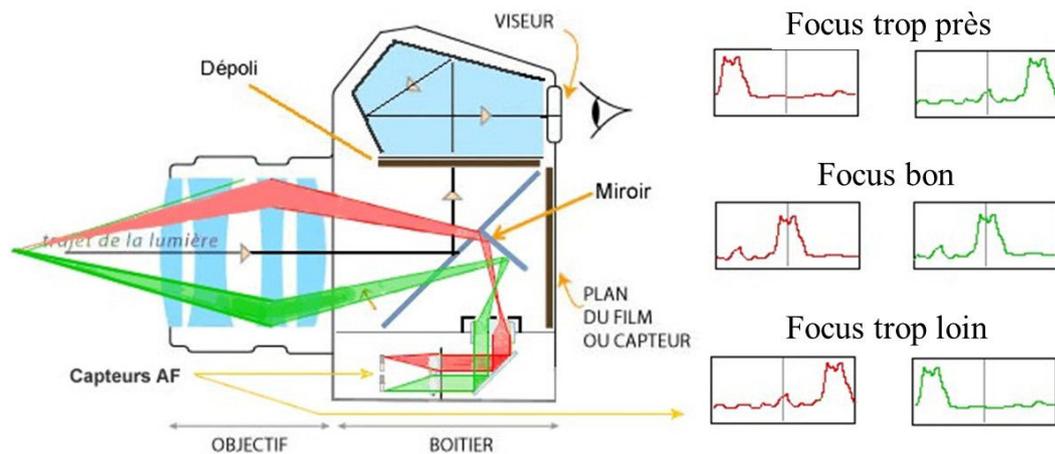


FIGURE 10 : DISPOSITIF D'AUTOFOCUS D'UN APPAREIL REFLEX

L'idée de la détection de phase est d'utiliser les deux extrémités d'un faisceau lumineux pour déduire le degré de focalisation. Ces bords de faisceaux sont isolés et orientés vers deux petits imageurs dédiés à la détection de phase. Lorsque l'APN est focalisé, les imagerie formées par ces morceaux de faisceaux se superposent (Graphes Figure 10). Dans le cas contraire, le déphasage observable permet de déduire le sens d'évolution de la mise au point et la distance par mesure précise du décalage. Ce dispositif est encombrant mais de nouvelles solutions, principalement portées par Canon et Fuji, permettent d'intégrer ce type de traitement directement sur le capteur principal. Les pixels détection de phase sont apparus autour de l'année 2010. La détection de phase devient intégrable directement au sein du capteur ce qui la rend adaptée aux caméras miniatures. Il existe actuellement deux principales astuces exploitées pour séparer le faisceau en plusieurs composantes utiles.

Une première solution consiste à cacher partiellement des pixels de façon à ce qu'ils ne captent que les rayons provenant d'un seul côté de la lentille [46] [47] [48] (cadre de gauche Figure 11). Seulement 40% du capteur [34] est concerné par cette occultation métallique partielle. Le déphasage n'est plus mesurable dès que les angles d'incidence des rayons lumineux sont trop importants. Cela induit également une perte de luminosité.

La seconde technologie consiste à placer deux photodiodes sous une même microlentille [49] [50] (cadre de droite Figure 11). Celles-ci permettent alors de construire deux images à partir desquelles le déphasage est observable. L'image finale est obtenue par somme des intensités recueillies par ces 2 photodiodes. Avec cette technologie, il n'y a d'une part aucune perte de luminosité, d'autre part la zone efficace est plus grande, jusqu'à 64% de la surface du capteur permet de mesurer un déphasage potentiel. En revanche la séparation d'un pixel en deux photodiodes réduit le rapport signal sur bruit.

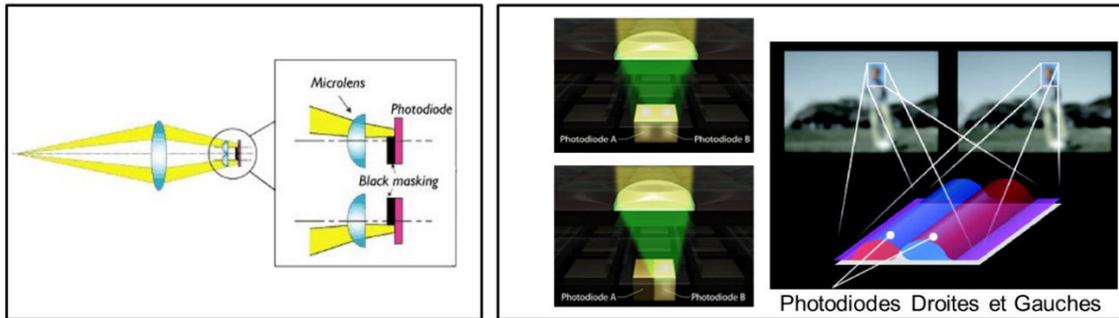


FIGURE 11 : CAPTEUR DETECTION DE PHASE PAR COUVERTURE PARTIELLE (GAUCHE) ET PIXEL DUAL PHOTODIODE (DROITE)

A partir des deux images acquises, des opérations de mise en correspondance, similaires à celles proposées pour la stéréovision, permettent de calculer la profondeur. L'utilisation d'un seul et même capteur pour les 2 images rend le système auto-calibré. L'amplitude du décalage peut alors être évaluée ligne par ligne (pour un déphasage horizontal) par corrélation croisée. Ce type de solution est étudiée dans [51] pour l'extraction de la profondeur et semble donner des résultats relativement intéressants mais avec une capacité de calcul non négligeable.

La détection de phase est sujette à brevet par les grands acteurs des capteurs d'images et donc difficilement exploitable. Aussi, cela conduit aux problématiques de la stéréovision, avec tout de même l'avantage de la suppression du besoin de calibration et le gain de place. En revanche, seulement la partie centrale du capteur est exploitable.

1.1.3.5 Depth from Defocus

Le Depth from Defocus consiste à analyser le niveau de flou d'une unique image afin d'arriver à remonter à l'information de profondeur. L'idée est d'évaluer le diamètre C du cercle de confusion (Figure 12), ou tache de flou, pour en déduire la distance d d'un objet de la scène par des considérations géométriques.

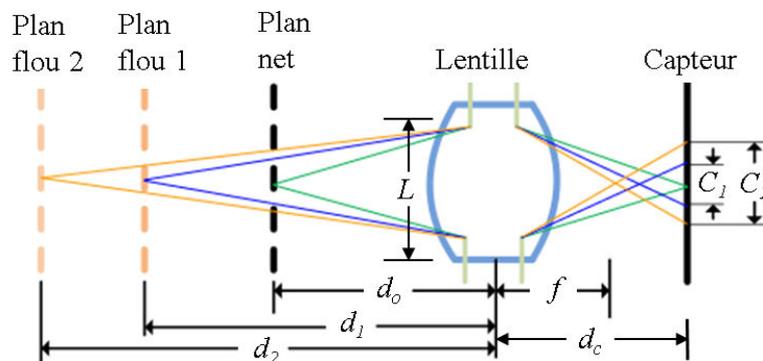


FIGURE 12 : VISUALISATION DU CERCLE DE CONFUSION

L'évaluation de C est particulièrement compliquée dans le cas de scènes réelles. En effet une image correspond à une infinité de taches lumineuses qui sont difficilement discernables les unes des autres. Les contours dans l'image sont en revanche riches en informations ; l'évaluation de leur épaisseur [52] permet de remonter à l'information de distance.

Les auteurs de [53] [54] proposent une première étape de filtrage de contour par filtrage de Canny [55] puis font une analyse fréquentielle autour des contours pour estimer C . Le résultat semble pertinent mais l'extraction nécessite plus de 20 secondes de calcul sur un PC pour une image de 400×300 pixels, ce qui dénote une forte complexité calculatoire.

D'autres travaux [56] [57] estiment les PSF (noyaux de convolution de flou), par l'intermédiaire d'une analyse fréquentielle ou de comparaison avec des PSF précalculées, afin de retrouver les niveaux de flou d'une image. Une fois encore cela semble complexe à mettre en œuvre pour une solution embarquée.

Les travaux de Chao Liu [58] sont plutôt originaux pour du DfD car 2 images sont utilisées. L'une d'elle est focalisée « partout », elle sert de référence, la seconde contient du flou. L'image de référence est floutée par 256 noyaux de convolution gaussiens. Chacun de ces noyaux étant représentatif d'une distance, la profondeur est déduite par minimisation de la différence entre les images totalement focalisées filtrées et l'image source contenant du flou. Plusieurs itérations permettent d'améliorer la carte de profondeur en fonction du contenu fréquentiel de la scène. Si cette technique semble plus prometteuse pour une solution embarquée, encore faut-il trouver une solution optique permettant de prendre une image totalement focalisée et une image avec une faible profondeur de champ, ce qui est souvent incompatible.

Yusuke Moriuchi et al. [59] combinent en quelque sorte la détection de phase et le DfD en séparant le faisceau incident sur la lentille avec deux filtres colorés semi-circulaire. Les filtres colorés utilisés sont les couleurs complémentaires aux traditionnelles rouge, verte et bleue car, d'une part, ils ont une meilleure transmittance, et d'autre part, ils laissent passer le vert qui est majoritairement présent sur les capteurs d'image. Une analyse du niveau de netteté est faite par convolution sur les 3 couleurs primaires, les résultats sont ensuite combinés pour extraire les distances effectives. L'avantage d'avoir un semblant de détection de phase permet de savoir si l'objet est placé avant ou après le plan net. Les méthodes conventionnelles de DfD conventionnelles se placent dans un seul de ces deux cas.

Globalement, ces méthodes DfD sont efficaces mais très complexes à mettre en œuvre directement au sein d'un capteur.

1.1.3.6 Depth from Focus

Le Depth from Focus [60], également appelée Shape from Focus dans la littérature [61], consiste à analyser des images d'une même scène, correspondantes à différents plans de focalisation (Figure 13). Les distances focales pour chaque point de la scène peuvent être extraites, et les distances effectives retrouvées par des considérations d'optique géométrique. C'est une méthode qui existe dans la nature, les caméléons semblent utiliser un tel mécanisme pour appréhender leur environnement [62].

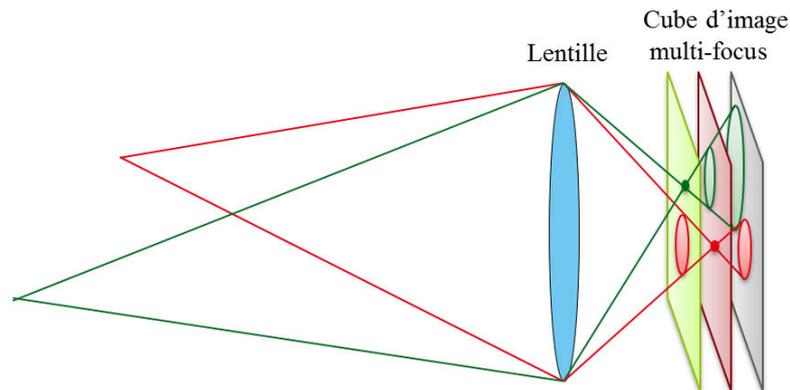


FIGURE 13 : PRINCIPE DU DfF

En supposant que chaque faisceau lumineux traversant la lentille est centré sur un même pixel de l'imageur quel que soit la distance de focalisation, l'analyse de l'évolution des pixels en fonction de cette distance permet d'extraire l'information de profondeur [63]. En revanche, cette méthode est très sensible au bruit de mesure et aux variations de la scène.

Divers travaux se basent sur une approche par analyse de contraste locale [64] [65] [66] [67] [68] dont le maximum à travers un cube d'images multi-foyers indique l'image la plus nette, et permet de construire une carte de profondeur. Celle-ci peut être filtrée par seuillages pour un résultat rapide [64] ou par une méthode itérative [65] [66] afin de supprimer les mauvaises estimations. La densité des cartes ainsi obtenues est relativement faible, cependant des solutions de remplissage, telles que proposées dans [69] et [70], permettent d'obtenir un rendu plus complet et fidèle à la réalité. Mais leur portage matériel est difficilement envisageable.

Yi-Che Chen et al. proposent de soustraire chaque image du cube multi-foyers à elle-même floutée par filtrage gaussien [71]. Une zone floue dans l'image d'origine va très peu ressortir alors que le résultat de la différence pour une zone nette sera important. La carte de profondeur est alors construite par sélection du maximum des différences en fonction de l'indice de mise au point. Cette méthode se distingue des analyses plus classiques du contraste mais les bords des objets dans les cartes de profondeur résultantes sont difficilement

discernables ce qui peut poser des problèmes lors de l'interprétation des données par un système hôte.

Une solution de DfF permettant une auto calibration de la lentille est proposée pour un Samsung Galaxy S3 [67]. Un premier traitement permet de réaligner les images et ainsi de compenser les tremblements de la main. L'approche consiste à extraire une image totalement focalisée du cube d'images multi-foyers puis à flouter cette image nette en tout point par de multiples PSF correspondant à des distances théoriques connues. Ces images floutées permettent ensuite d'estimer les distances focales réelles correspondantes à chacune des images du cube par une analyse différentielle similaire à [57]. La profondeur peut ensuite être déduite. Cette approche est très intéressante pour le développement logiciel de DfF sur des architectures non calibrées. Cependant, leur solution complète testée en simple tâche sur CPU i7 à 3.40GHz pour un cube de 25 images VGA prend 20 minutes pour fournir un résultat, ce qui est particulièrement long.

Dans [65] et [66], les auteurs cherchent à créer un capteur d'image intégrant du DfF. Un calcul proche du Laplacien permet d'extraire un coefficient représentatif de la netteté. Pour chaque pixel, le maximum du coefficient à travers le cube multi-foyers indique la distance correspondante. L'ASIC est prévu pour fonctionner avec une cadence image de 1500Hz afin de pouvoir fournir 30 cartes de profondeur par seconde. L'estimation de distance est assez grossière mais une méthode itérative d'amélioration du résultat par suppression des erreurs puis interpolation des données est proposée. Elle est en revanche difficilement implantable au sein même du capteur.

Une architecture a été proposée par le CEA [72] ; celle-ci comporte des unités de calcul programmable associées à des blocs de pixels qui permettent d'analyser la netteté en temps réel. Chaque macro-pixel se voit *in fine* attribué un indice relatif à la profondeur. L'avantage d'une telle architecture est qu'elle est programmable, ce qui permet d'évaluer diverses stratégies d'extraction de netteté ainsi que d'être robuste face à tout type d'optique et aux petites variations de scène pendant une mesure.

Le plus récent des articles traitant sur le DfF a été présenté en 2017 [64]. Les auteurs montrent des résultats intéressants à partir d'un critère de netteté basé sur le LoG (Laplacian of Gaussian [73]), en particulier en termes de vitesse de traitement (25fps pour 32 niveaux de profondeur). Le traitement est exécuté sur une architecture de capteur SCAMP-5 combinant acquisition et traitement [74]. La carte de profondeur est cependant peu dense et l'image totalement focalisée comporte de nombreux artefacts et des zones floues.

Dans l'ensemble, pour un système calibré, la complexité de calcul des solutions de DfF reste faible. Ces derniers exemples de l'état de l'art montrent qu'elles sont propices aux systèmes embarqués. De plus, l'utilisation d'une unique caméra permet de réduire l'encombrement et

le coût du système. Cependant, l'analyse sur plusieurs images d'un cube multi-foyers rend le DfF particulièrement sensible aux variations parasites dans la scène.

1.1.4 Combinaison de technologies

Avec l'avènement des véhicules autonomes, les recherches sur la fusion de capteurs ont grandement proliféré. Il existe de plus en plus de méthodes de mixage d'information, basées sur diverses technologies d'acquisitions [75] [76] [77]. Ces approches sont plutôt adaptées à une exécution côté système et de ce fait en aval de ces travaux de thèse.

1.1.5 Synthèses des méthodes d'extraction de profondeur

Les deux familles, actives et passives, de technologie d'extraction de profondeur offrent des solutions intéressantes. Un net avantage de la technologie active est son efficacité sur les zones uniformes d'une scène. En revanche, ces systèmes sont généralement encombrants et consomment plus d'énergie qu'un système passif à cause de l'émission de signaux et sont également moins discret. De plus, ils sont sensibles au milieu ambiant et peuvent poser des problèmes d'interférences lorsque plusieurs de ces systèmes sont utilisés simultanément. Les technologies passives proposent des méthodes très diverses mais exigent des ressources de calcul conséquentes pour la majorité d'entre elles. Elles sont cependant totalement démunies face à des éléments très uniformes. Les solutions mono caméra sont très compactes, ce qui confère un vrai plus pour concevoir des systèmes embarqués miniatures.

Le Tableau 1 résume les avantages et inconvénients de chacune des technologies vues précédemment.

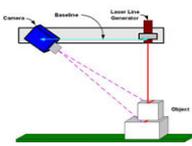
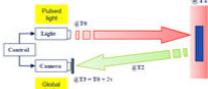
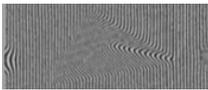
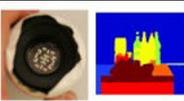
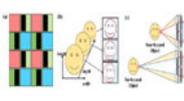
1.1.6 Conclusion sur les méthodes d'extraction de profondeur

Au regard de cet état de l'art, il semble exister peu de solutions pour les petits systèmes embarqués offrant de faibles ressources calculatoires. Les solutions mono-caméras, par leur compacité et leur faible dépendance énergétique, sont particulièrement attractives pour de tels systèmes. De toutes les méthodes présentées, la technique Depth from Focus semble la plus adaptée aux contraintes d'intégration et de limitation des ressources de calcul.

Des différents travaux vus précédemment, aucune référence ne porte d'attention particulière aux variations de l'image provoquées par l'optique. Un changement de foyer optique provoque notamment un grandissement, ou effet de zoom, étudié plus loin dans ce manuscrit. Celui-ci peut avoir un impact important sur des techniques DfF. Une question se pose alors, quel type d'optique doit être utilisé pour une application DfF ? La section suivante présente

une revue de méthodes optiques permettant la réalisation de la fonction de balayage de foyers.

Tableau 1 : Récapitulatif de l'état de l'art de l'extraction de profondeur

		Avantage	Inconvénient	Ref
Active	Passive			
Triangulation active		<ul style="list-style-type: none"> - Résolution configurable - Précision 	<ul style="list-style-type: none"> - Sécurité suivant la classe du laser - Encombrement - Sensibilité à la forme (zone d'ombre) - Temps de mesure - Sensibilité à la luminosité ambiante 	[7] [8]
Temps de vol		<ul style="list-style-type: none"> - Précision - Temps réel 	<ul style="list-style-type: none"> - Faible résolution - Interférences - Sensibilité à la luminosité ambiante 	[13] [14] [15] [16]
Lumière structurée		<ul style="list-style-type: none"> - Précision - Résolution 	<ul style="list-style-type: none"> - Sensibilité à la luminosité ambiante - Complexité de calcul - Sensibilité aux variations de la scène - Coût - Encombrement 	[6] [8] [17] [18] [19] [75]
	Photogrammétrie	<ul style="list-style-type: none"> - Résolution - Précision 	<ul style="list-style-type: none"> - Complexité de calcul - Calibration - Dépendance aux textures - Encombrement 	[2] [24] [25] [26] [28] [29] [24] [75]
	Caméra Plénoptique	<ul style="list-style-type: none"> - Capture du champ lumineux - Précision 	<ul style="list-style-type: none"> - Forte complexité de calcul - Faible résolution - Dépendance aux textures 	[34] [36] [37] [39] [40] [41]
	Ouverture codée	<ul style="list-style-type: none"> - Mono image - Résolution - Précision 	<ul style="list-style-type: none"> - Forte complexité de calcul - Dépendance aux textures 	[43] [44] [45]
	Pixel à détection de phase	<ul style="list-style-type: none"> - Résolution - Système intégré - Mono image 	<ul style="list-style-type: none"> - Complexité de calcul - Dépendance aux textures 	[46] [47] [49] [50] [51] [59]
	Depth from Defocus	<ul style="list-style-type: none"> - Précision - Mono image - Résolution 	<ul style="list-style-type: none"> - Forte complexité de calcul - Dépendance aux textures 	[53] [54] [58] [57] [69]
	Depth from Focus	<ul style="list-style-type: none"> - Résolution - Faible complexité de calcul - Système intégré - Précision - Faible coût 	<ul style="list-style-type: none"> - Sensibilité aux variations de la scène - Dépendance aux textures 	[61] [63] [64] [65] [66] [67] [68] [72]

1.2 Technologies Optiques

1.2.1 Introduction

Pour des systèmes embarqués contraints, il est nécessaire d'avoir une optique proposant une forte intégrabilité, une variabilité de la distance focale de l'ordre du centimètre afin de garantir une résolution de profondeur intéressante, et une grande rapidité des changements de distance focale. La plupart des caméras actuelles avec lesquelles il est possible d'effectuer un balayage de foyers optiques sont basées sur un mouvement mécanique de lentille. Dans le cadre de cette thèse, nous visons une implémentation DfF la plus compacte possible, de ce fait, nous avons choisi de nous concentrer sur des optiques intégrées qui permettent une variation de foyers sans aucun mouvement mécanique. A cette fin, deux principales technologies que sont les lentilles liquides et celles à cristaux liquides sont présentées dans cet état de l'art. Citons les travaux [58] qui utilisent une lentille liquide dans une application DfD, ou encore l'approche DfF [71] ainsi que l'étude [78] pour lesquelles des lentilles à cristaux liquides sont utilisées, qui confirment la pertinence de ces choix.

1.2.2 Lentilles liquides

Les lentilles liquides sont des dispositifs optoélectriques qui jouent sur la déformation d'un liquide pour générer un dioptre permettant une focalisation de la lumière. Il en existe deux catégories technologiques aujourd'hui commercialisées par les sociétés Varioptic et Optotune.

Une première catégorie de lentilles liquides [79], breveté par Varioptic [80], utilise le phénomène d'électromouillage pour générer une forme de lentille à partir d'un liquide transparent (Figure 14).

Une goutte d'huile quasiment sphérique est placée dans une cellule de verre hermétique remplie d'un liquide, d'indice optique différent, non miscible avec l'huile. En appliquant un courant, la forme de la goutte d'huile change, modifiant ainsi l'interface entre les deux liquides, ce qui provoque une variation de la distance focale de la lentille.

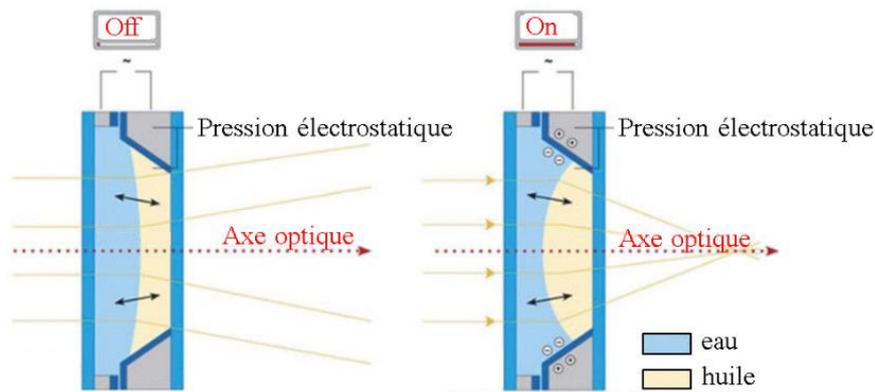


FIGURE 14 : LENTILLE LIQUIDE UTILISANT LE PHENOMENE D'ELECTROMOULLAGE [80]

Ce type de lentille permet d'effectuer une excursion de foyers de -15 à +25 dioptries en quelques millisecondes. De plus, un contrôle intelligent de leur polarisation permet également de distordre légèrement la lentille et ainsi aider à la stabilisation d'image. Un inconvénient d'utiliser l'électromouillage est la limitation de l'ouverture à quelques millimètres à cause, d'une part, de la saturation de la force capillaire, et d'autre part, de la gravité qui peut venir perturber la forme du dioptre [79]. Bien que la tension de commande de ces lentilles soit grande, 10V à 70V, le faible courant nécessaire la rend utilisable dans des systèmes embarqués grâce à une micro-pompe de charge.

La seconde catégorie de lentilles liquides utilise une membrane élastique entre deux chambres contenant chacune des liquides d'indices différents. La différence de pression entre ces deux chambres impose une déformation de la membrane et crée ainsi un dioptre sphérique (Figure 15).

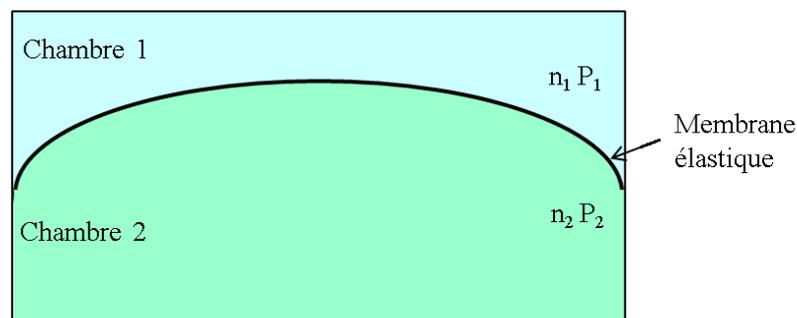


FIGURE 15: LENTILLE LIQUIDE A MEMBRANE ELASTIQUE

Le changement de pression peut être provoqué de diverses façons : mécaniquement par mouvement du support de la membrane, hydrauliquement par régulation de fluide dans l'une des deux chambres, électrostatiquement avec l'utilisation de polymères électroactifs [81]. Les lentilles liquides Optotune [82] sont construites selon ce second modèle. L'utilisation d'une membrane permet de concevoir des lentilles avec des pupilles bien plus grandes qu'avec l'électromouillage, jusqu'à 25mm d'après la gamme Optotune contre maximum 5mm pour Varioptic. En revanche, la puissance optique semble plus limitée. Il reste cependant largement suffisant pour une application DfF.

Lei Li et al. [83] proposent un système à double membrane qui permet d'augmenter la puissance optique d'une lentille liquide ou de diminuer l'épaisseur de la lentille pour une distance focale donnée, ce qui est intéressant dans un contexte de miniaturisation de système. Mais un doute demeure sur l'homogénéité de la déformation.

1.2.3 Technologie à cristaux liquides

1.2.3.1 Les cristaux liquides

Les cristaux liquides sont des matériaux présentant des propriétés physiques intermédiaires entre les phases solides et liquides [84]. Ces composés s'écoulent comme des liquides plus ou moins visqueux, mais leurs molécules sont ordonnées de façon similaire à un solide. Cette fluidité leur confère une anisotropie de leurs propriétés généralement plus importante que celle que l'on peut retrouver dans la plupart des cristaux et notamment en termes de biréfringence. C'est cette propriété de biréfringence, associée à l'état « semi-liquide » des cristaux, qui place ces matériaux sur un piédestal pour de nombreuses applications électro-optiques. Ils présentent notamment des effets électro-optiques parmi les plus importants des matériaux existants.

Il existe plusieurs phases « cristal liquide » définies par l'agencement structurel de leurs molécules, la plus commune étant la phase nématique. Seule la phase nématique uniforme sera utilisée dans le cadre de cette thèse.

Une phase nématique est une phase dans laquelle les molécules ne présentent aucun ordre d'orientation : celles-ci peuvent « couler » librement les unes sur les autres, de ce fait, c'est une phase liquide. Les molécules du cristal liquide nématique ont tendance à s'orienter spontanément de manière parallèle les unes aux autres. Afin de privilégier une direction, une mince couche présentant de minuscules sillons parallèles réalisés par brossage, permet aux molécules de cristal liquide proches des parois de s'ancrer suivant cette direction privilégiée. Ainsi, en l'absence de champ électrique, les molécules s'orientent naturellement suivant cette direction. En présence d'un champ électrique supérieur à une valeur critique, les molécules de cristal liquide éloignées des parois tendent à s'orienter parallèlement au champ. Celles voisines des parois restent solidement ancrées dans les sillons pratiqués sur les couches d'orientation. Lorsque le champ électrique n'est plus présent, les molécules de cristal liquide tendent à redevenir parallèles aux molécules proches des parois. La direction moyenne locale des axes moléculaires est communément appelée le directeur de la phase nématique, il est souvent repéré par un vecteur unitaire \vec{n} .

La phase nématique est un milieu uniaxe ou biréfringent, c'est à dire qu'elle présente deux indices de réfraction différents, n_o et n_e , respectivement les indices ordinaire et extraordinaire.

L'indice de réfraction du milieu sera donc différent selon la direction de propagation de la lumière, de sa polarisation et de l'orientation moléculaire. On considère une biréfringence positive, n_e est donc orienté suivant le vecteur directeur \vec{n} et n_o lui est perpendiculaire.

Le principe des cellules à cristaux liquides est de créer une interaction entre le champ électrique de commande et l'orientation moléculaire \vec{n} . Les molécules nématiques sont très anisotropes et, lors de l'application d'un champ électrique \vec{E} , elles se comportent comme un dipôle induit par ce champ électrique. Un moment dipolaire perpendiculaire à l'axe des molécules et centré sur celles-ci apparaît. Cela génère un couple électrique qui entraîne une rotation des molécules de cristal liquide.

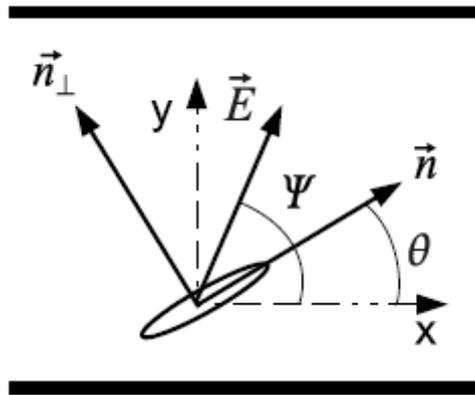


FIGURE 16 : ORIENTATION D'UNE MOLECULE SUIVANT LE CHAMP ELECTRIQUE

Avec les conventions de la Figure 16, le module du couple électrique imposé au cristal lors de l'application d'un champ est donné par l'équation (1.5).

$$\Gamma_E = \frac{\Delta\epsilon \cdot E^2}{2} \sin(2(\psi - \theta)) \quad (1.5)$$

Cette formule indique que l'orientation des cristaux liquides nématiques est dépendante du carré du champ et de son anisotropie diélectrique $\Delta\epsilon$. Il s'avère que plus $\Delta\epsilon$ est grand, moins le champ électrique doit l'être pour induire une rotation des molécules. Cela influence directement les tensions de commande qui seront appliquées pour l'exploitation des cristaux liquides.

Un faisceau lumineux non polarisé incident à une couche de cristaux liquides à biréfringence positive est donc séparé en deux composantes, une composante ordinaire, perpendiculaire à \vec{n} et une composante extraordinaire, parallèle à \vec{n} . Chaque composante se propage dans le milieu biréfringent à deux vitesses différentes suivant l'indice considéré, n_o pour la composante ordinaire et un indice $n(\theta)$ résultant de la projection le long des deux axes propres pour la composante extraordinaire. L'équation (1.6) nous donne la valeur de cet indice en fonction de l'inclinaison de la molécule de cristal liquide par rapport à l'axe optique.

$$n(\theta) = \frac{n_e \cdot n_o}{\sqrt{n_e^2 \cdot \cos^2(\theta) + n_o^2 \cdot \sin^2(\theta)}} \quad (1.6)$$

Cette différence de vitesse de propagation engendre un déphasage qui peut être calculé à partir de l'équation (1.7).

$$\Delta\varphi = \frac{2\cdot\pi}{\lambda} \left(\int_0^e n(\theta) \cdot dy - n_o \cdot e \right) \quad (1.7)$$

Avec e l'épaisseur de cristal liquide à traverser et λ la longueur d'onde du faisceau lumineux.

C'est ce déphasage de la lumière qui est utilisé dans le cadre des lentilles à cristaux liquides. Lors de la rotation moléculaire, l'indice optique du milieu est modifié. Sa valeur suivant l'axe optique est alors contrôlable électriquement.

Une particularité des cristaux liquides est qu'ils ne supportent pas les tensions continues. En effet, celles-ci peuvent provoquer des réactions d'électrolyses sur les cristaux liquides, réduisant ainsi fortement leur durée de vie. Puisque le couple électrique (1.5) dépend du carré du champ, les cristaux liquides sont commandables par un signal alternatif centré. Pour le maintien du couple électrique à une valeur constante, il est préférable d'appliquer un signal créneau de rapport cyclique égal à 50%.

Ce sont là seulement quelques grandes lignes concernant les cristaux liquides. Plus de détails concernant leur mécanique sont précisés dans la thèse [85]. Quant à leur procédé de fabrication, il est présenté dans la thèse [86].

1.2.3.2 Les lentilles à cristaux liquides nématiques

Les lentilles à cristaux liquides nématiques sont des dispositifs optiques électriquement commandables qui permettent de moduler la phase d'un signal lumineux, et ainsi, d'émuler la fonction de lentille à focale variable. Elles sont composées de deux électrodes entre lesquelles se trouve une couche de cristaux liquides d'épaisseur calibrée par l'intermédiaire de cales micrométriques. La constitution et la géométrie de ces électrodes permettent de générer un gradient de champ électrique qui module l'indice de la couche nématique de façon à émuler une lentille selon une direction de polarisation de la lumière correspondant à la direction de rotation des directeurs.

Dans l'état de l'art, on rencontre des lentilles dites non-homogènes dont au moins une des électrodes est incurvée (cadre de droite sur la Figure 17) et des lentilles homogènes, dans lesquelles la couche de cristal liquide est d'épaisseur constante (cadre de gauche sur la Figure 17). Les premières ont l'avantage de fonctionner avec des électrodes simples, le caractère sphérique du substrat provoquant à lui seul une convergence du faisceau lumineux. Les lentilles homogènes requièrent des électrodes particulières pour induire la fonction de lentille. Elles sont plus complexes à concevoir mais les lentilles ainsi réalisées peuvent être plus fines.

De par la nature biréfringente des cristaux liquides, ces différentes structures nécessitent de polariser la lumière incidente parallèlement au directeur du cristal liquide ce qui limite le taux de transmission. Une solution consistant à empiler deux cellules de sorte à ce que leurs axes optiques soient parfaitement perpendiculaires fut proposée par Williams et al. en 1984 [87].

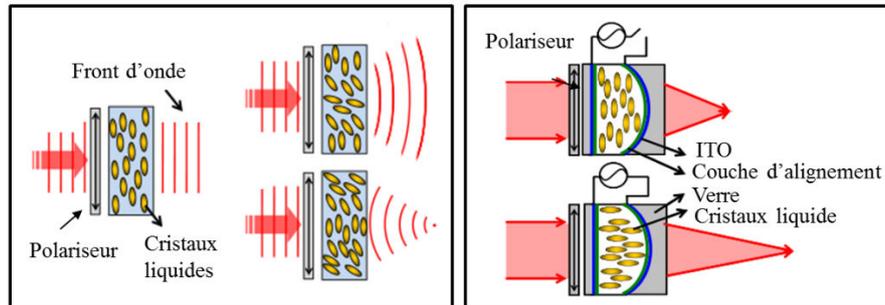


FIGURE 17 : ARCHITECTURE DE LENTILLE A CRISTAL LIQUIDE HOMOGENE (GAUCHE) ET NON HOMOGENE (DROITE)

Les premières lentilles à cristaux liquides non homogènes furent inventées par Sato dans les années 1970 [88]. Celles-ci avaient un diamètre de 20mm et une épaisseur de cristal liquide très importante, de l'ordre de 500 μm , ce qui entraîne de longs temps de réponses liés aux contraintes élastiques, ainsi qu'une diminution de la transmission lumineuse due à la fois à l'absorption et à la diffusion de la couche nématique. De plus, pour de fortes épaisseurs, les molécules centrales sont très désordonnées à cause d'une faible interaction avec les surfaces d'ancrage. Cela dégrade la qualité optique.

Depuis de nombreuses études ont été menées, on peut en apprécier l'étendue dans [89]. Les lentilles homogènes sont apparues avec diverses propositions de forme d'électrodes (Figure 18).

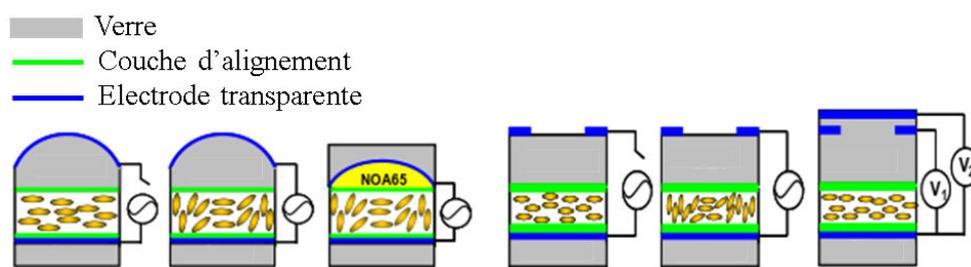


FIGURE 18 : DIFFERENTES ARCHITECTURES DE LENTILLES HOMOGENES

Une méthode consiste à utiliser une électrode courbe (3 images de gauche Figure 18) qui induit un gradient de champ grâce à l'évolution du diélectrique au sein de la lentille. Cependant, le gain en épaisseur que nous pourrions obtenir avec une technologie homogène se retrouve limité. Les premières lentilles avec électrodes percées [90] (partie droite de la Figure 18) permettaient d'avoir une répartition non uniforme du champ électrique. L'inconvénient de ces structures est la nécessité de tensions élevées (plusieurs dizaines de volt) pour que le champ puisse se propager jusqu'au centre la lentille, ce qui limite les

ouvertures maximum à environ 1mm. L'augmentation du diélectrique par l'ajout de verre entre l'électrode percée et la couche de cristaux liquides (3 images de droite Figure 18) permet d'augmenter l'ouverture, tout comme l'utilisation d'une deuxième électrode placée au-dessus de l'électrode percée. En outre, les travaux [91] montrent que la solution double électrode réduit le risque d'apparition de lignes de déclinaison dues à un manque d'uniformité du champ électrique dans la cellule. Cependant ces astuces augmentent l'épaisseur finale de la lentille ce qui a tendance à réduire l'ouverture des systèmes imageurs dans lesquels elles peuvent être utilisées.

Les travaux [92] présentent une architecture de lentille à cristaux liquides constituée de plusieurs électrodes concentriques ce qui permet de réduire les tensions d'alimentation (-10V) et d'augmenter le diamètre de pupille (4mm). Le profil de la lentille ainsi modulée est tout de même perturbé au niveau des interfaces inter-électrodes et au centre de la pupille (Figure 19). La lentille est plus mince mais telles des lentilles de Fresnel, la distance focale minimale ne peut être très courte car cela nécessiterait un grand nombre d'électrodes et complexifierait de ce fait la réalisation et l'adressage des tensions.

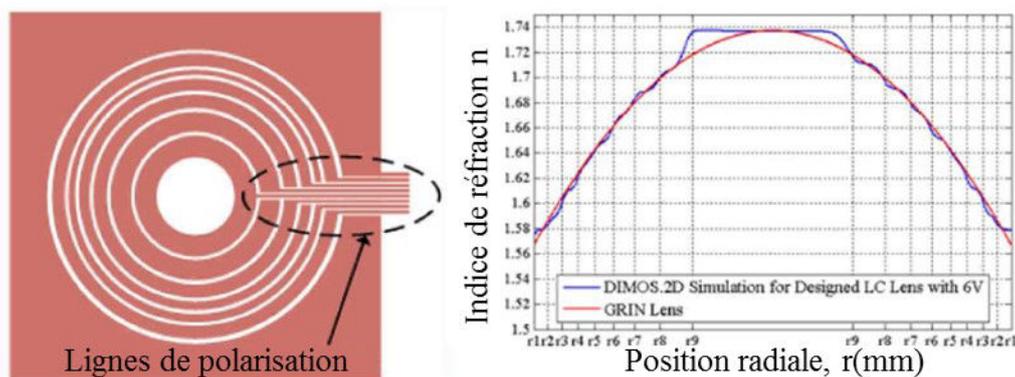


FIGURE 19 : LENTILLE AVEC DE MULTIPLES ELECTRODES CIRCULAIRES [92]

D'autres travaux [93] proposent d'appliquer les tensions appropriées à une série d'électrodes ligne liées par une résistance électrique importante (16M Ω) de sorte à créer un gradient de tension et ainsi provoquer la modulation d'indice de nature à émuler une lentille cylindrique. La proximité des électrodes provoque également de la diffraction qui réduit la qualité de telles lentilles [94]. D'après l'article [95], les lentilles à électrodes pixélisées créent un champ non continu ce qui limiterait les distances focales atteignables à des distances relativement importantes. De plus, les résultats de [92] (Figure 19) montrent l'altération de l'indice à cause de l'espacement inter-électrodes, ce qui empire avec une électrode pixélisée. La solution proposée dans [96] associe deux lentilles cylindriques créées à partir d'électrodes lignes pour réaliser l'équivalent d'une lentille à électrode pixélisée. La variation d'indice dans la couche de cristal liquide semble plus homogène mais l'inconvénient de cette structure est son épaisseur : il faut l'équivalent de 4 lentilles pour focaliser sans polariseur un faisceau lumineux, ce qui limite l'ouverture du système.

L'augmentation du diélectrique est limitée, les solutions multi-électrodes ne semblent pas suffisantes, ce qui amène à la réalisation de lentilles modales utilisant des électrodes résistives [97] [98] [92] [85]. Avec des électrodes résistives, la variation de la distance focale peut être provoquée par un changement de la fréquence de polarisation ou par une modification de la tension. L'idéal étant de jouer avec les deux grandeurs pour limiter les aberrations comme cela est montré dans [92].

De nos jours, de nombreux travaux tendent vers la miniaturisation de cette technologie afin de proposer une solution de lentille de contact pour la correction de la presbytie [99] [100] [101], c'était notamment l'une des applications originelles proposées par Sato [88].

1.2.4 Conclusion sur les technologies optiques

Le Tableau 2 montre un récapitulatif des technologies de lentilles modulables.

Tableau 2 : Récapitulatif de l'état de l'art es lentilles liquides et à cristaux liquides

Type de lentille	Diamètre de pupille	Distance focale	Epaisseur	Référence
Lentille Optotune	2.3 à 25mm	-20cm \rightarrow ∞ \leftarrow +15cm	12 à 24.5mm	[82]
Lentille Varioptic	0.6 à 3.9mm	-20cm \rightarrow ∞ \leftarrow +6cm	1.8 à 5.25mm	[80]
Lentilles modales PEDOT-PSS	1 à 5mm	+1cm \rightarrow ∞	10 à 500microns + épaisseur de verre	[85] [95]
Lentilles CL avec substrat incurvés	20mm	+16cm \rightarrow 80cm	10 à 500microns + épaisseur de verre incurvé	[88]
Lentilles modales (ITO)	2.5mm et 6.5mm	+50cm \rightarrow ∞	10 à 500microns + épaisseur de verre	[97] [98]
Multi-électrode circulaire	4mm	+18cm \rightarrow ∞	10 à 500microns + épaisseur de verre	[92]
Réseau d'électrodes	20mm	+1000cm \rightarrow ∞	10 à 500microns + épaisseur de verre	[96]

En terme d'encombrement, les lentilles Optotune semblent les moins adaptées pour une miniaturisation d'un système DfF. En revanche leur gamme d'ouverture permet d'avoir une profondeur de champ réduite et ainsi une meilleure résolution de profondeur que les lentilles Varioptic. Les lentilles à cristaux liquides modales présentent également des caractéristiques

en adéquation avec la plupart des contraintes citées en introduction de ce chapitre. Elles semblent en revanche relativement lentes comme tous les dispositifs à cristaux liquides nématiques, en particulier dès que l'épaisseur devient importante.

Notre choix c'est donc porté sur l'évaluation comparative des lentilles Optotune et de lentilles à cristaux liquides modales.

Chapitre 2 Développements algorithmiques et architecture pour l'extraction de la profondeur

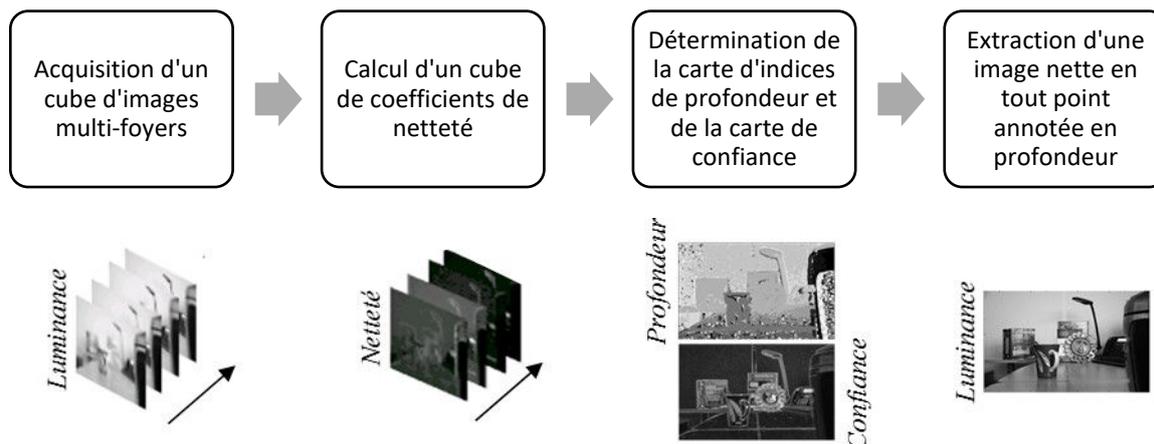
2.1	Introduction	31
2.2	Avant-propos sur l'évaluation de la netteté	32
2.3	Présentation du banc de test	40
2.4	Algorithme d'évaluation de distance par mesure bidirectionnelle de contraste local	41
2.5	Comparaison des algorithmes de Depth from Focus	56
2.6	Amélioration du rendu	63
2.7	Proposition d'architectures	65
2.8	Conclusion	75

2.1 Introduction

L'objet de ce chapitre est de présenter des solutions algorithmiques dans un premier temps puis un portage matériel de la solution la plus prometteuse au regard des compromis qui doivent être faits pour une intégration au plus près du pixel. La solution retenue doit donc être capable de traiter des données brutes (RAW) colorées tout en limitant l'empreinte mémoire, la complexité de calcul et la consommation. Le chapitre précédent a montré que les algorithmes DfF semblaient particulièrement adaptés aux architectures embarquées ; c'est donc dans cet axe que se porteront nos développements.

Le synoptique ci-après présente les étapes typiques d'un algorithme DfF.

Un cube d'images multi-foyers est acquis en faisant varier la distance focale entre chaque prise. Chacune des images du cube est analysée aussi tôt que possible pour fournir une matrice à trois dimensions représentative de la netteté, également appelée critère de netteté dans ces travaux. Celle-ci permet ensuite d'extraire la profondeur de la scène et la carte de confiance. Une image nette en tout point peut alors être construite à partir des indices de profondeur et du cube d'images sources.



Nous présentons dans ce chapitre une première section sur la notion de netteté et ce qui en découle pour l'analyse de contraste intra-image. Nous décrivons ensuite les solutions algorithmiques retenues et les comparons avec d'autres méthodes de l'état de l'art. Enfin, nous proposons une micro-architecture pour l'implémentation matérielle de la solution la plus pertinente.

2.2 Avant-propos sur l'évaluation de la netteté

La netteté est définie dans le dictionnaire Larousse par « l'absence de flou dans les contours ». En langage naturel, la netteté est donc un critère binaire qui est défini comme une absence de flou. Mathématiquement, elle se traduit dans l'image par l'apparition de discontinuités locales entre les éléments de différents tons et couleurs. La netteté peut donc être étudiée en considérant le voisinage de chaque pixel.

En optique géométrique, une zone est dite nette si l'image de chacun de ses points lumineux pris indépendamment est réduite à un point. En pratique, la résolution de tout capteur d'image quel qu'il soit ne peut être infinie. L'œil, les pellicules argentiques ou encore les capteurs des appareils photos numériques les plus performants, ont un grain limite de résolution. La netteté représente alors la capacité qu'a un système optique à rendre du détail dans une image. De ce fait, les données visuelles que nous percevons sont toujours échantillonnées, donc massivement discontinues, ce qui rend impossible l'évaluation de netteté par une simple étude de discontinuité. Une solution pour détecter la netteté d'une zone dans une image passe par l'analyse du flou. Celui-ci représente un critère mesurable complémentaire de la netteté et varie en fonction de la distance de l'objet et de la distance de focalisation. La manière dont le flou évolue en fonction des caractéristiques optiques du capteur n'est cependant pas triviale et nous avons souhaité préalablement étudier la sensibilité du flou par rapport aux variations de la distance focale d'une lentille, ainsi que la notion d'hyperfocale. Cette étude est présentée en Annexe 1 et 3, nous en reprendrons les résultats dans ce qui suit.

2.2.1 Analyse de l'évolution du flou

Un résultat important de cette étude est l'équation qui lie le diamètre C de la tache lumineuse formée sur le capteur (aussi appelée cercle de confusion) par une source ponctuelle en fonction des paramètres optiques et de la distance de la source (Figure 20). Cette équation est reproduite ci-dessous (2.1).

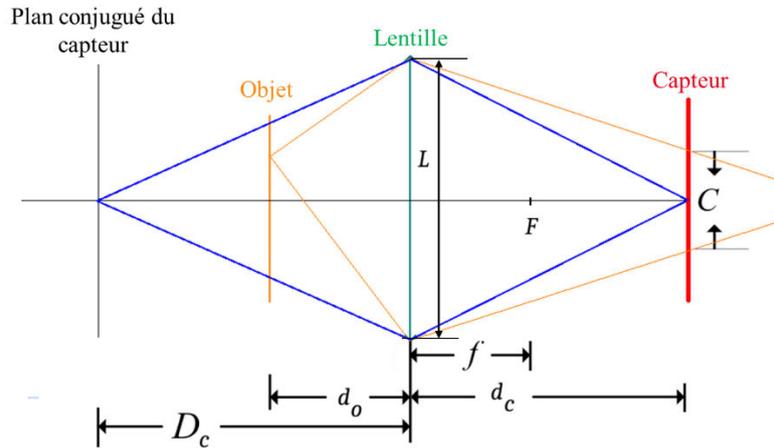


FIGURE 20 : REPRESENTATION D'UN DISPOSITIF IMAGEUR

$$C = L \cdot d_c \cdot \left| \frac{1}{d_o} - \frac{1}{D_c} \right| \quad d_o \in [f + \infty[\quad (2.1)$$

L représente l'ouverture du système, d_c la distance capteur lentille, D_c la distance conjuguée du capteur fonction de la distance focale f de la lentille et d_o la distance lentille-objet.

En supposant que nous arrivions à mesurer C , il serait possible d'accéder très simplement à l'information de distance en remaniant l'équation (2.1). Dans la réalité, une scène est constituée d'une infinité de sources lumineuses qui génèrent chacune sa tache lumineuse. Celles-ci se mélangent sur le capteur ce qui rend presque impossible la mesure directe de leur diamètre. De plus, l'effet de pupille provoque une tache d'Airy (alternance de franges circulaires) qui complique d'autant plus la délimitation de la bordure du cercle de confusion. L'estimation précise du flou est donc relativement ardue. En revanche, sa variabilité en fonction de la distance de mise au point offre la possibilité de construire un critère évolutif permettant la détection du « minimum de flou ». Nous parlerons dans la suite de critère de netteté.

Ce type de critère est généralement conçu pour admettre un maximum lorsqu'une image est nette. Il existe trois évolutions théoriques d'un tel critère pour un balayage de foyers de proche à l'infini. Ceux-ci sont présentés sur la Figure 21.

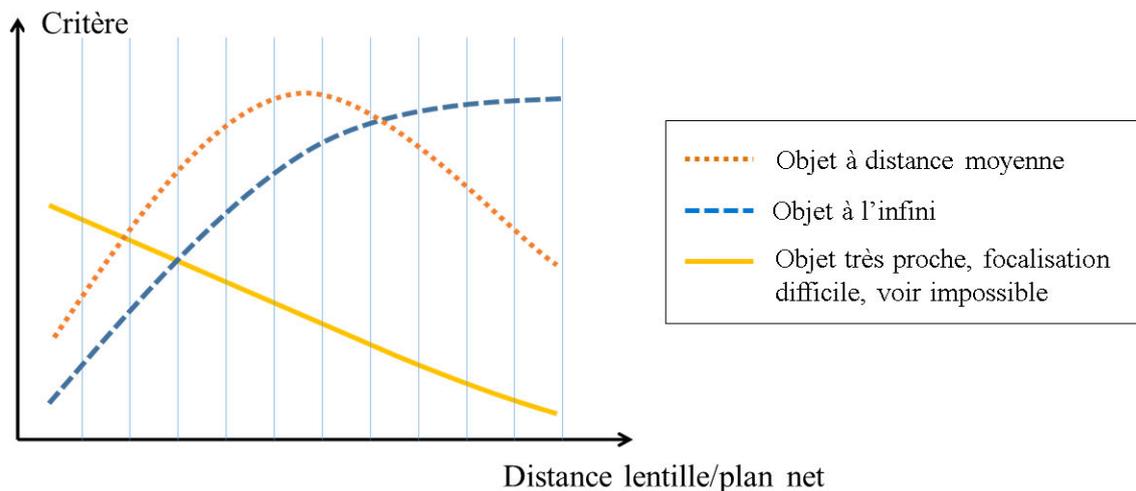


FIGURE 21: EVOLUTION IDEALE D'UN CRITERE DE NETTETE

Un objet placé à une distance moyenne est flou au plus proche foyer, tend vers une image nette puis redevient flou. Un objet situé au-delà de la distance hyper focale ($D_c > L \cdot d_c / P_{pix}$) dit à l'infini est flou au départ puis atteint un horizon de netteté avec l'augmentation de la distance focale. Enfin le flou d'un objet trop proche ne fait qu'empirer avec l'augmentation de la distance focale. Les deux premiers cas sont réellement exploitables par un algorithme DfF. Le dernier quant à lui est une situation désespérée, les seules solutions pour rendre un objet proche net sont de reculer physiquement l'appareil photo ou bien de changer d'objectif.

2.2.2 Choix d'un critère de netteté

Dans l'état de l'art, deux types d'évaluation de la netteté ressortent : les analyses fréquentielles qui permettent d'évaluer la netteté à partir de la transformée de Fourier 2D locale, et les analyses de contrastes locaux. La complexité de calcul liée à l'analyse fréquentielle est cependant un inconvénient vis-à-vis de notre objectif de faible complexité.

Une troisième vision intéressante, bien que peu courante dans l'état de l'art, consiste à observer l'évolution de l'intensité de chaque pixel en fonction de la distance focale, indépendamment de son voisinage. En effet, aux alentours de zones contrastées, un pixel n'acquiert pas la même intensité lumineuse suivant qu'il est focalisé ou pas. La Figure 22 montre l'évolution de 4 pixels en fonction de l'indice de mise au point pour un cube d'images alignées.

Les pixels situés dans un voisinage uniforme (bleu clair et foncé) ont une évolution difficilement mesurable. En revanche, ceux situés dans des zones contrastées (rouge et vert) admettent un extremum qui correspond à la meilleure mise au point. Remarquons que sur le point vert, il semble y avoir un plateau minimum. Celui-ci s'explique tout simplement avec la considération de la profondeur de champ présentée en annexe 3.

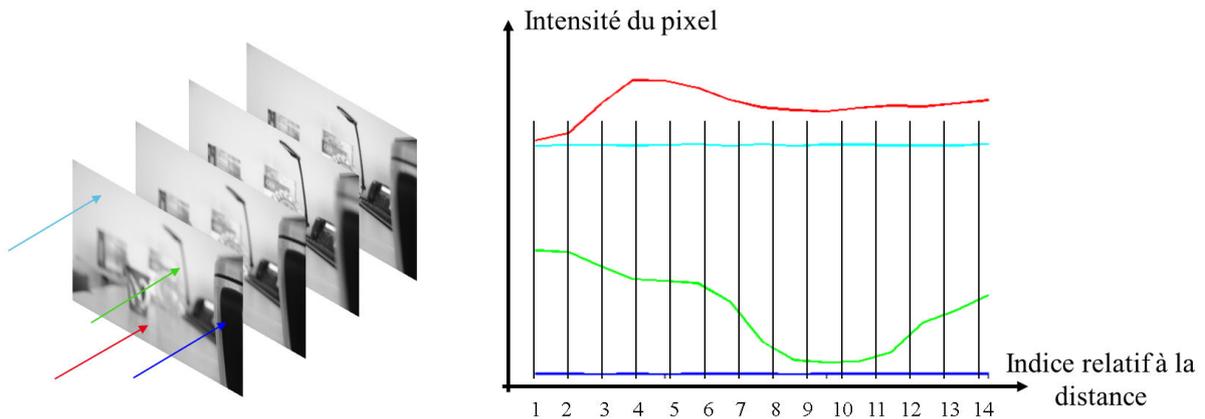


FIGURE 22 : EVOLUTION DE L'INTENSITE DU PIXEL EN FONCTION DE LA DISTANCE DE MISE AU POINT

L'analyse de l'évolution de chaque pixel peut donc aussi être un bon indicateur de netteté. Une méthode d'extraction de profondeur que nous avons développée à partir de cette solution est présentée dans l'annexe 2, mais sa sensibilité vis-à-vis de l'artéfact de « zoom de focalisation » présenté plus bas, ainsi qu'aux variations de luminosité en limite son usage à des cas de figures particuliers, par exemple en vision microscopique [63]. Nous avons donc préféré ne pas poursuivre cette étude dans le cadre de ce travail de thèse.

Nous nous sommes intéressés aux analyses de contraste pour leur faible complexité. Le but d'une analyse locale est de pouvoir associer à chaque pixel, ou groupe de pixels, un scalaire représentatif de sa netteté. L'objet d'une telle analyse est d'étudier le proche voisinage de chaque pixel. Pour tous les cas considérés dans la suite de ce manuscrit, nous avons utilisé des fenêtres de traitement composées de $w_h \times w_w$ pixels. Lorsque les dimensions des images ne sont pas un multiple de $w_h \times w_w$, on ajoute des zéros sur les bords droit et bas de l'image pour compléter les derniers rangs. Ainsi on « dégrade » les informations sur deux bords de l'image plutôt que sur l'ensemble du contour.

Il existe dans l'état de l'art divers traitements pour effectuer une analyse de contraste [64] [102] [66] [67] [68]. L'article [103] présente notamment une étude comparative entre divers critères de netteté. La méthode que ces auteurs préconisent consiste à faire la somme de valeurs absolues des différences (SMD) de pixels deux à deux suivant deux directions (2.2).

$$SMD = \sum_x \sum_y |I(x, y) - I(x, y - 1)| + \sum_x \sum_y |I(x, y) - I(x - 1, y)| \quad (2.2)$$

Des méthodes, également présentées dans [103] pour certaines, sont basées sur un prétraitement de l'image par un filtre passe-haut plus classique, du type détection de contour. Des travaux se basent notamment sur l'extraction du Laplacien pour chaque image du cube multi-foyers [104] [66] [105] puis conservent l'indice de l'image pour lequel celui-ci est maximum. Il est également possible d'utiliser le module du gradient de l'image puis d'extraire la profondeur par analyse du maximum des maximums locaux des gradients à travers le cube pour chaque pixel.

Nous avons remarqué au cours de nos expériences que l'utilisation du seul maximum par fenêtre de traitement avait un inconvénient qui est illustré ci-dessous.



FIGURE 23 : ILLUSTRATION DU CHOIX DE NOTRE CRITERE DE NETTETE

Sur la Figure 23, on peut voir 2 barres lumineuses, une floue à gauche et une nette à droite. Une extraction du maximum dans les fenêtres illustrées par les cadres rouges nous donne deux mêmes valeurs et ne permet donc pas de discriminer la netteté. On a donc cherché à utiliser un autre descripteur de netteté pour pallier ces cas.

En statistique, l'histogramme représente la distribution des pixels par valeur. Pour une image, cette distribution dépend de l'exposition, des couleurs de la scène... qu'en est-il du niveau de flou d'une image ? Nous pouvons supposer qu'une image nette aura un histogramme plus étalé qu'une image floue pour laquelle les faisceaux lumineux ont tendance à se mélanger pour former une zone relativement uniforme. Le support de l'histogramme représente donc un critère intéressant pour juger de la netteté d'une zone d'intérêt. Cette technique est d'ailleurs également référencée dans [103].

Nous avons réalisé une expérience en utilisant un cube de 15 imagerie monochromes correspondant à une même zone fortement contrastée d'une scène pour différentes distances de mise au point. La Figure 24 présente ces imagerie et leurs histogrammes respectifs.

L'évolution du flou sur ces blocs de pixels est visible sans aucune ambiguïté, l'imagerie numéro 7 est la plus nette. Il est bien moins évident d'extraire à partir du support de ces histogrammes lequel des échantillons est le plus net. Les histogrammes 7, 8 et 9 sont difficiles à différencier.

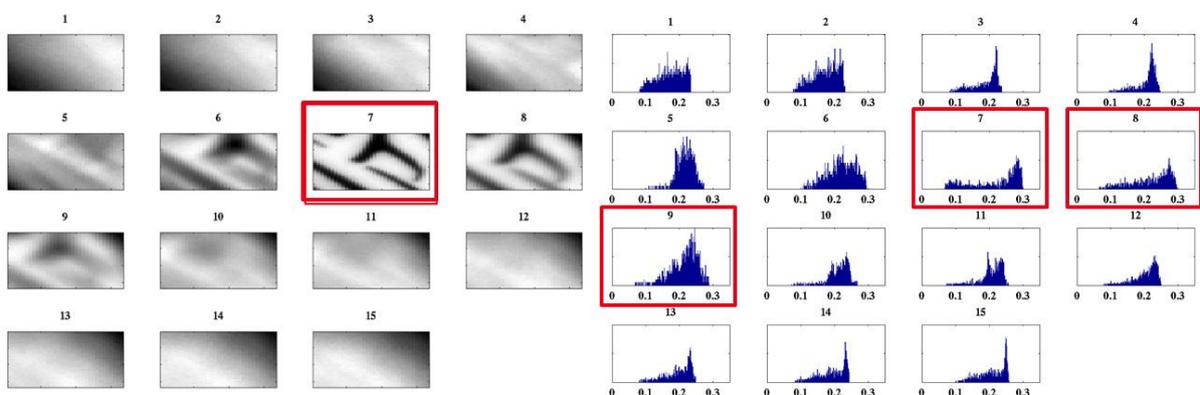


FIGURE 24 : BLOCS 32x32 PIXELS CENTRES SUR LE PIXEL (753;945) POUR CHACUNE DES IMAGES D'UN CUBE MULTI-FOYERS (GAUCHE) ET LEURS HISTOGRAMMES RESPECTIFS (DROITE)

L'observation des hautes fréquences de ces mêmes imagerie (Figure 25) semble plus pertinente. L'histogramme le plus étendu est sans aucun doute celui correspondant à

l'échantillon 7. Il y a donc un grand intérêt à faire une analyse après un filtrage passe haut sur l'image. On pourrait même augmenter la dynamique du critère, et le rendre ainsi encore plus différenciant, en utilisant les données issues de filtrage passe-haut signées (comme la dérivée d'une image suivant une direction), c'est ce qui sera fait avec la méthode bidirectionnelle proposée plus loin.

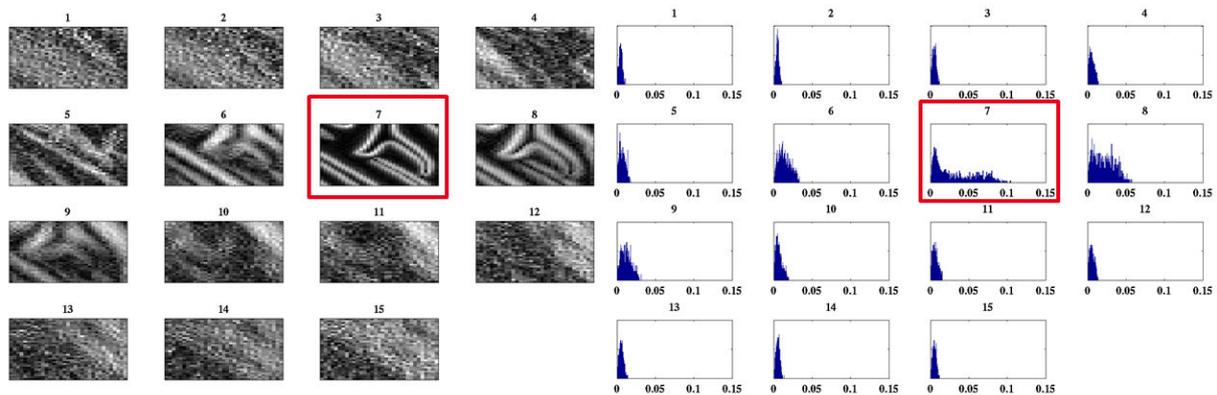


FIGURE 25 : BLOCS 32x32 COEFFICIENTS PASSE-HAUT CENTRES SUR LE PIXEL (753;945) POUR CHACUN DES GRADIENTS D'UN CUBE MULTI-FOYERS (GAUCHE) ET LEUR HISTOGRAMME RESPECTIF (DROITE)

Après analyse des histogrammes du module du gradient pour diverses zones de la scène, il s'avère que celui-ci est généralement le plus étendu dans les zones fortement contrastées (contour, texture marquée...).

Si on reprend la Figure 23, l'étendue de l'histogramme entre les deux cadres est bien différente et permet de discriminer de tels cas de figure.

Dans un système contraint, on ne peut se permettre de calculer l'histogramme pour chaque bloc d'une image. En revanche, l'extraction de son support au sens large, c'est-à-dire du maximum et minimum par bloc de traitement, en est finalement très proche.

Nous proposons donc de travailler avec le critère de netteté décrit par l'équation (2.3).

$$S_{crit}(x, y) = \max_{X,Y}(D(X, Y)) - \min_{X,Y}(D(X, Y)) \quad (2.3)$$

Où (X, Y) est un élément de la fenêtre de dimension $w_h \times w_w$ centrée sur (x, y) et D une image filtrée par un filtre passe haut. Ce critère sera comparé à quelques autres de l'état de l'art dans le cadre d'application DfF plus loin dans ce manuscrit. L'expérience nous montre bien souvent que le calcul de la différence des valeurs extrema est équivalent à la recherche du maximum. Mais il permet d'avoir un meilleur résultat pour les cas restants, les zones relativement uniformes mais légèrement texturées par exemple.

Notons que quelle que soit la méthode, l'analyse de netteté a une principale limitation : elle est quasiment impossible dans les zones uniformes. Quelle que soit la mise au point d'un système optique, l'image d'un corps uniforme reste uniforme quelle que soit la distance focale de l'objectif. Il est donc très difficile de faire une quelconque distinction de mise au point pour

ces objets. L'utilisation d'éléments optiques de très bonne qualité peut aider à faire ressortir les plus petits détails, comme les textures, et ainsi percevoir des détails dont l'évolution de netteté est détectable, mais ce type de matériel peut coûter très cher sans apporter de solution sur l'ensemble de l'image. Il reste donc que les images de profondeurs obtenues par l'emploi de méthodes Depth from Focus contiennent forcément un certain nombre de zones de profondeur inconnue.

2.2.3 Influence de la mise au point sur l'invariance de l'image ou « zoom de focalisation »

Durant cette thèse, nous avons pris l'habitude d'appeler « zoom de focalisation » le grandissement, ou l'effet de zoom, provoqué par un déplacement de la lentille lors de la mise au point. Je n'ai trouvé aucune publication dans l'état de l'art faisant référence à ce phénomène. Il peut pourtant être très impactant pour le traitement d'images capturées à différentes distance de focalisation : la recherche du maximum de netteté dans le cube multi-foyers implique une hypothèse d'invariabilité de la scène qui est malmenée par cette déformation. On se doit donc d'en tenir compte, ou du moins d'en être conscient. Nous en verrons dans ce chapitre les effets directs sur l'extraction de carte de profondeur.

Le zoom de focalisation est très simple à visualiser par une expérience d'optique géométrique. Considérons un « objectif » à lentille simple tel que représenté sur la Figure 26. La mise au point se fait en translatant légèrement la lentille le long de l'axe optique du système ($d_{c1} < d_{c2}$). Ce déplacement est l'essence même du zoom ; il réduit l'angle délimitant le champ de vision du capteur comme on peut le voir sur la Figure 26. Pour une même surface de capture, une plus petite portion de la scène est imagée lorsque la distance capteur lentille est plus grande.

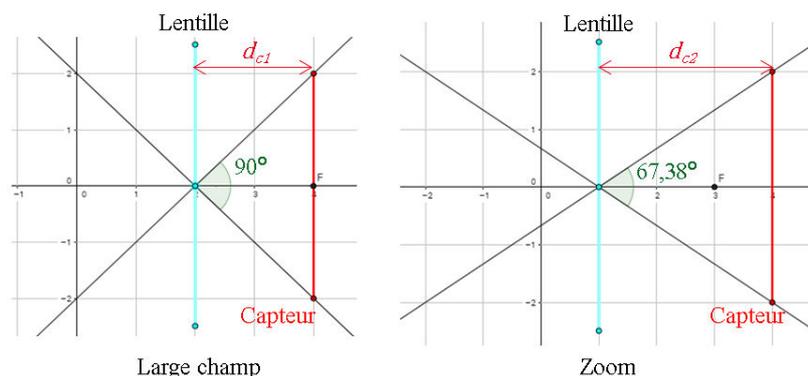


FIGURE 26 : ILLUSTRATION DU ZOOM DE FOCALISATION

Sur la Figure 27, on peut voir la superposition en vert et violet des deux images extrêmes d'un cube d'images multi-foyers. L'effet de zoom est particulièrement marqué au niveau de la lampe de bureau (cadre orange).

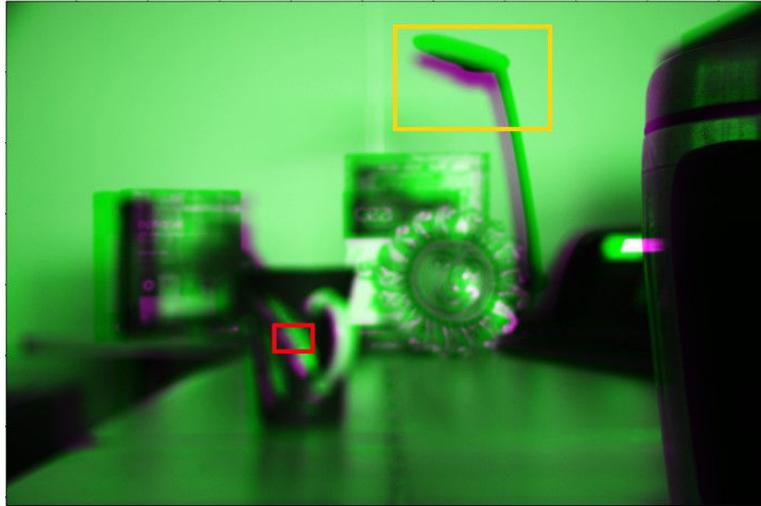


FIGURE 27 : SUPERPOSITION DES DEUX IMAGES EXTREMES D'UN CUBE D'IMAGES MULTI-FOYERS

Une petite zone de 32×32 pixels centrée sur une même coordonnée a été extraite (Figure 28) d'un cube d'images multi-foyers afin de montrer les effets locaux du zoom de focalisation.

Les flèches rouges représentent le « mouvement » global observable par comparaison de ces imagettes.

Les points rouges sont disposés sur un même objet, il est indéniable que celui-ci change de coordonnées dans l'image en fonction du foyer. La netteté devant être analysée localement, on déduit que la dimension de la fenêtre d'analyse peut jouer sur le résultat en fonction de l'amplitude du zoom de focalisation. Si celle-ci est trop petite, on risque de perdre de vue l'objet de la scène entre les différentes fenêtres à travers le cube.

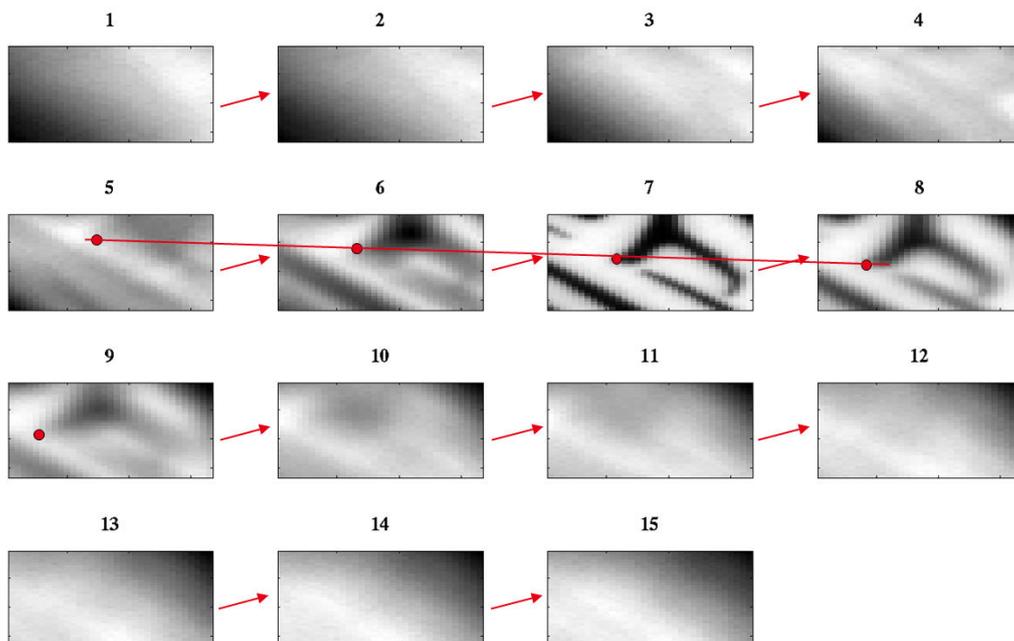


FIGURE 28 : EFFET LOCAL DU ZOOM DE FOCALISATION

Ce zoom de focalisation diffère d'un système optique à un autre. Nous le quantifierons dans le troisième chapitre de ce manuscrit pour quelques optiques et nous chercherons à le minimiser.

2.3 Présentation du banc de test

Afin d'évaluer les performances de nos algorithmes de construction de carte de profondeur nous avons pris 3 séries de 14 photos en format RAW (non dématricées), pour 3 scènes différentes, avec un appareil photo Nikon D300S équipé d'un objectif Nikkor AF-S DX 16-85 mm f/3.5-5.6G ED VR. Afin de garantir l'invariance de la scène, l'appareil a été fixé sur un trépied et l'éclairage néon neutralisé pour éviter les modulations lumineuses à 50Hz. La distance focale de l'appareil a été ajustée manuellement, par pas réguliers, de proche à l'infini. Les distances de mise au point ne sont pas exactement connues mais dans un premier temps, nous souhaitons extraire l'image la plus nette pour chaque zone de la scène, et ainsi construire une carte contenant des indices d'images acquises. Ces indices pourront par la suite être mis en correspondance avec des distances réelles par des méthodes de calibration ou par une étude optique. Pour cette raison, on associera dans la suite de ce manuscrit la carte d'indice des images dans le cube d'images multi-foyers et la vraie carte de distance des éléments de la scène par les termes « carte de profondeur ». C'est légitime car le passage de l'un à l'autre est une relation bijective paramétrée par les propriétés optiques du dispositif de capture.

Nous avons vu qu'un critère de netteté est souvent basé sur une analyse locale du contraste, or la présence de filtres colorés (comme le motif de Bayer) sur le capteur impose une forte modulation d'intensité pour les pixels proches ce qui a un impact important sur l'analyse des contrastes locaux, qui est plutôt adaptée à des images en niveaux de gris. C'est pourquoi nous avons fait le choix de moyenniser les 4 pixels du motif de Bayer pour obtenir une image monochrome. Cela permet également de réduire le bruit thermique du capteur. Les nouveaux cubes d'images sont alors équivalents à des images en niveaux de gris sur 12 bits, 4 fois moins résolues spatialement que les images RAW. Ainsi, on conserve une certaine fidélité des images pour l'analyse de netteté, tout en gardant la possibilité d'utiliser les images couleurs pour la reconstruction d'images totalement focalisées. Quelques échantillons sont présentés sur la Figure 29. Ils seront respectivement dénommés *soleil*, *panda* et *chaise* dans la suite de ce manuscrit.

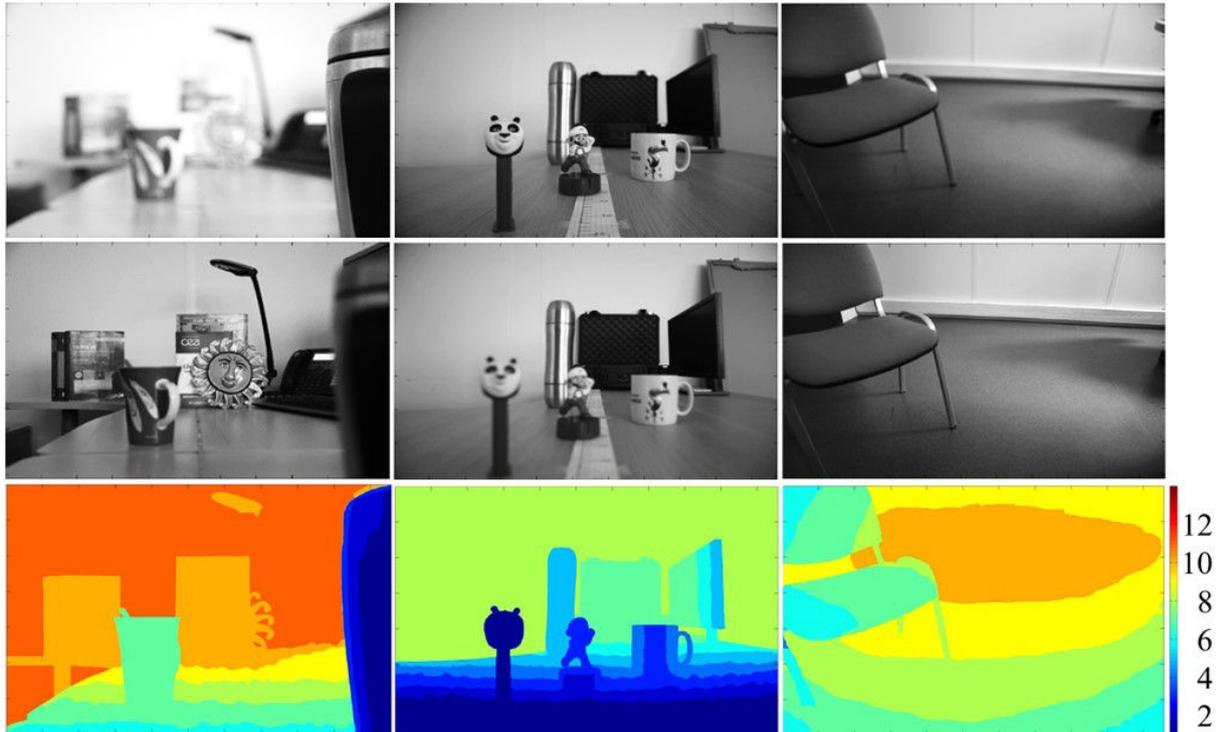


FIGURE 29 : DEUX IMAGES EXTRAITES DES 3 CUBES D'IMAGES SOURCES (DE GAUCHE A DROITE : SOLEIL, PANDA, CHAISE). DISTANCE FOCALE PROCHE EN HAUT ET ELOIGNEE AU MILIEU. VERITES TERRAIN EN BAS.

Pour chacun des cubes d'images tests, nous avons utilisé un algorithme DfF afin d'extraire des cartes de profondeur. Ces cartes ont ensuite été retouchées manuellement, grâce à un éditeur d'image (Gimp) afin de créer des cartes vérités terrain correspondantes à des cartes de profondeur issues d'un algorithme DfF idéale. Celles-ci nous serviront pour analyser nos résultats.

Ces scènes sont composées de divers objets placés à des distances allant de 20 à 300 centimètres. On a choisi volontairement deux types de scènes. L'une est relativement vide, avec seulement une chaise et un arrière-plan légèrement texturé ; elle permettra de valider l'efficacité des algorithmes dans des conditions proches de ce que pourrait voir un véhicule en navigation autonome. Les autres contiennent divers objets permettant de mieux se rendre compte de la profondeur réelle d'une part, mais également d'évaluer l'efficacité d'une potentielle détection d'obstacle.

2.4 Algorithme d'évaluation de distance par mesure bidirectionnelle de contraste local

Comme expliqué dans l'introduction de ce chapitre, la détection de profondeur repose sur la détermination de l'indice du maximum du critère de netteté sélectionné. Nous avons évalué les performances du module du Laplacien d'une part et du SMD d'autre part. Lors de l'évaluation de ces critères, nous avons remarqué que leur nature multidirectionnelle, bien

qu'efficace dans les zones texturées, présentait un handicap dans les zones de transitions horizontales ou verticales marquées : dans ces zones, la composante orthogonale à la transition présente une faible valeur qui atténue la valeur du critère.

Partant de ce constat, nous avons souhaité développer un procédé permettant de tirer le meilleur des deux mondes : textures et transitions directionnelles.

Pour ceci, nous proposons d'extraire la profondeur suivant deux directions, verticale et horizontale, puis de réunifier les données en conservant les estimations les plus pertinentes afin de fournir une carte de profondeur unifiée. La sélection de la pertinence étant basée sur une carte de vraisemblance, auto-évaluation de la performance de la mesure, présentée plus bas. Une représentation schématique de l'algorithme est montrée Figure 30. Nous verrons par la suite que l'utilisation de deux directions est plus performante que les critères multidirectionnels proposés dans l'état de l'art.

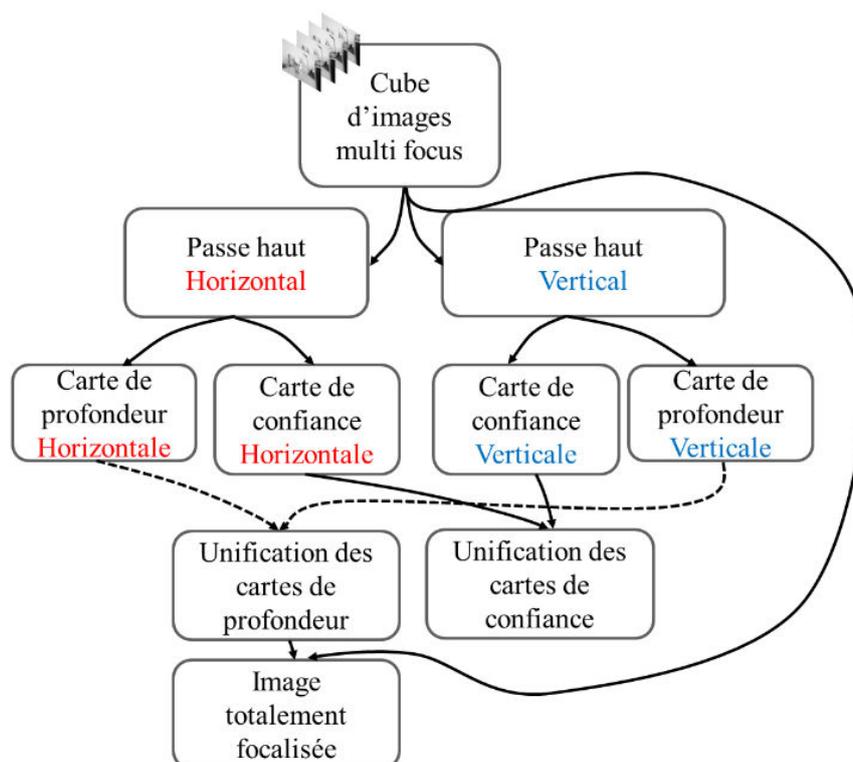


FIGURE 30 : REPRESENTATION SCHEMATIQUE DE NOTRE FAMILLE D'ALGORITHMES

Nous présentons dans cette section trois approches de complexités décroissantes qui reposent sur la mesure du contraste local pour la détermination de cartes d'indice de profondeur et de confiance. Ces cartes permettent ensuite de reconstruire une image totalement focalisée, ou All in Focus (AiF), à partir du cube d'images sources.

Une première approche, dite haute résolution, réalise l'analyse de netteté par fenêtres glissantes. Ses images intermédiaires sont de taille équivalente aux images sources.

La seconde, dite basse résolution, utilise des résultats intermédiaires qui ont des dimensions réduites, permettant ainsi de gagner en temps de calcul et en mémoire utilisée.

La dernière reprend la seconde approche mais cette fois en utilisant des fenêtres de traitement distinctes, ou non recouvrantes, ce qui permet un nouveau gain important en quantité de calculs et en mémoire.

Prises dans l'ordre énoncé précédemment, ces approches sont de plus en plus compatibles avec une implémentation matérielle embarquée ; on tend vers une réduction de la complexité algorithmique et un allègement du besoin mémoire. Les sous paragraphes suivants présentent ces trois approches et la section qui suit propose une comparaison qualitative et quantitative entre elles et avec l'état de l'art.

2.4.1 Approche intra-image haute résolution

2.4.1.1 Extraction de la profondeur suivant deux directions.

Le critère de netteté que nous avons développé a été présenté à la section 2.2.2, équation (2.3), il repose sur un filtre passe haut D . Afin d'extraire des cartes de profondeur suivant deux directions, nous avons fait le choix d'utiliser les dérivées premières de l'image (2.4).

$$D_H = \frac{\partial Im(x,y,z)}{\partial x} \quad D_V = \frac{\partial Im(x,y,z)}{\partial y} \quad (2.4)$$

Où Im représente un cube d'images en niveau de gris. Les valeurs sont conservées signées, afin d'augmenter la dynamique du support de l'histogramme qui est utilisé pour le calcul de ce critère. Les deux directions produisent chacune un cube de coefficients représentant l'évolution de la netteté locale par zone de la scène en fonction de l'indice relatif à la distance au plan net z .

Nous extrayons alors le support de l'histogramme de ces dérivées pour construire des cubes contenant les critères de netteté horizontaux S_{critH} et verticaux S_{critV} .

Le critère de netteté est construit de sorte que les maxima suivant la troisième dimension z pour chacun des cubes S_{critH} et S_{critV} correspondent à l'indice de l'image la plus nette. La carte de profondeur DM peut alors être extraite avec l'équation (2.5).

$$DM_{H/V}(x, y) = \underset{z}{\operatorname{argmax}}(S_{crit\ H/V}(x, y, z)) \quad (2.5)$$

Il peut arriver que plusieurs indices d'images correspondent au même maximum du critère de netteté pour un même pixel. Lorsque cela se produit, nous avons fait le choix de conserver celui qui correspond à la distance la plus proche, soit le plus petite indice lors d'un balayage avec une distance focale qui augmente. Ce choix est motivé par l'utilisation d'un tel système

pour des applications critiques telle que la navigation autonome. Dans ce cas, il est préférable d'éviter un objet qui n'existe pas, ou qui est plus loin que prévu, plutôt que de heurter un objet qui n'a pas été correctement estimé.

La Figure 31 représente les cartes de profondeurs horizontales et verticales extraites avec cette méthode pour des fenêtres glissantes de taille 16×16 pixels.

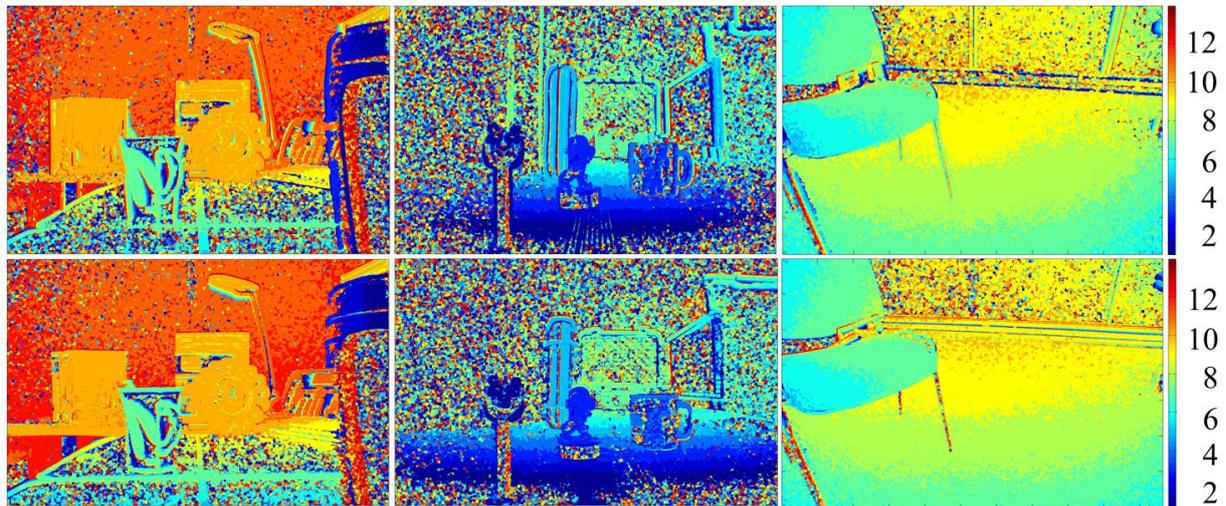


FIGURE 31 : CARTES DE PROFONDEUR HORIZONTALES (HAUT) ET VERTICALES (BAS) POUR LES TROIS SCENES ($w_x=w_y=16$)

Sur les trois scènes, on distingue la direction utilisée pour extraire les cartes de profondeur en observant les contours verticaux et horizontaux des objets. Notons également que les zones très texturées, ou possédant un contraste local important, sont relativement peu bruitées. C'est particulièrement flagrant sur les cartes de profondeurs de la scène « chaise » qui a priori contenait peu d'éléments. Cela s'explique par la qualité des images, le sol et la chaise elle-même sont recouverts de micro-motifs qui ont été captés lors de l'acquisition des images sources.

2.4.1.2 Calcul des cartes de confiance

Les profondeurs suivant chacune des deux directions ont été extraites indépendamment ; il nous faut maintenant un moyen d'unifier les résultats les plus pertinents. Une fusion des cartes de profondeurs par moyennage vient très vite à l'esprit mais le résultat d'une telle opération donnerait des indices pas toujours entiers qui ne correspondraient alors à aucune distance focale utilisée lors du balayage de foyers. Cela compliquerait l'évaluation de la distance réelle. Il s'agit plutôt de trouver une stratégie d'autoévaluation de la qualité de l'estimation de profondeur pour chacune des deux directions afin de conserver seulement la plus pertinente.

D'autre part, il est intéressant de fournir à l'application hôte le résultat de cette autoévaluation pour éclairer le post-traitement, faciliter l'amélioration et l'interprétation des

résultats, ou encore améliorer la fusion de données multi-capteurs qui devient un enjeu important pour les véhicules autonomes. Cela permettrait, par exemple, de marquer les zones uniformes comme peu fiables.

Nous proposons donc de produire une carte de confiance, de même dimension que la carte de profondeur, dans le but de pondérer les différentes estimations de distance. On confondra dans ce manuscrit les termes pertinence, confiance et vraisemblance.

Les images de profondeurs suivant les deux directions nous montrent que l'évaluation est peu pertinente dans les zones uniformes et ce parce qu'elles présentent peu de variations d'intensité et de contraste. La dynamique de variation du critère de netteté en fonction de l'indice relatif à la distance au plan net est faible pour ces zones uniformes. À l'inverse les zones fortement contrastées montrent une variabilité de leur critère de netteté au-dessus de la moyenne. Nous avons donc choisi d'utiliser la dynamique du critère de netteté comme scalaire représentatif de la pertinence de l'estimation. Celle-ci est calculée à partir de l'équation (2.6) pour chacune des directions.

$$C_{h/v}(x, y) = \max_z \left(S_{crit\ h/v}(x, y, z) \right) - \min_z \left(S_{crit\ h/v}(x, y, z) \right) \quad (2.6)$$

Un tel coefficient de confiance est facile à extraire par la mise à jour du minimum et du maximum du critère de netteté au fur et à mesure du traitement.

La Figure 32 présente les cartes de confiance horizontales et verticales pour chacune des scènes.

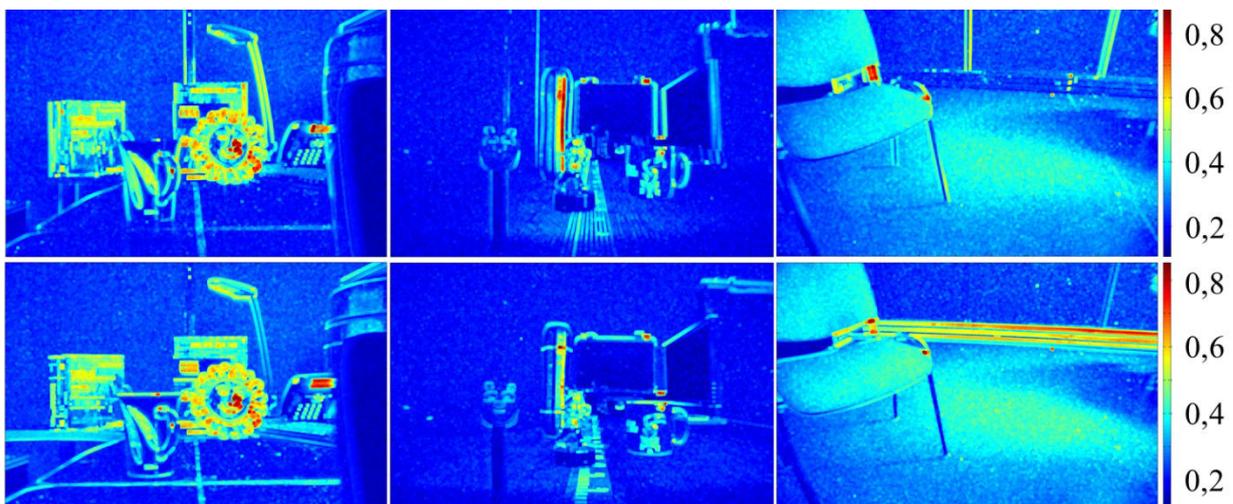


FIGURE 32 : CARTES DE CONFIANCE HORIZONTALES (HAUT) ET VERTICALES (BAS) POUR LES TROIS SCENES
($w_x=w_y=16$)

Une fois encore on distingue très clairement les deux directions de traitement. Les zones qui ressortent le plus correspondent à celles contenant des textures marquées comme supposé précédemment.

Les valeurs de vraisemblance doivent à ce stade être considérées en valeurs relatives. Elles sont très dépendantes de la luminosité de la scène. Dans ces travaux, on se contente de fournir une donnée représentative de l'estimation, laissant leur interprétation et utilisation libre pour de futures études qui pourront indiquer à partir de quelle valeur de vraisemblance l'information peut être considérée comme réaliste. Cependant nous sommes persuadés que l'utilisation de cette information de pertinence dans des modèles probabilistes de fusion de capteur, ou encore avec des approches réseau de neurones peut être un vrai atout pour améliorer la perception de la profondeur.

2.4.1.3 Unification des cartes de profondeur et de confiance

Nous avons créé des cartes de profondeur et de confiance horizontales et verticales qu'il faut désormais fusionner. L'idée est de conserver la meilleure estimation de profondeur en regard des coefficients de confiance horizontaux et verticaux. Ceux-ci sont donc utilisés pour fusionner les cartes de profondeur et également les cartes de confiance par l'intermédiaire des équations (2.7) et (2.8).

$$DM(x, y) = \begin{cases} DM_h(x, y) & \text{si } C_v(x, y) \leq C_h(x, y) \\ DM_v(x, y) & \text{si } C_v(x, y) \geq C_h(x, y) \end{cases} \quad (2.7)$$

$$C(x, y) = \begin{cases} C_h(x, y) & \text{si } C_v(x, y) \leq C_h(x, y) \\ C_v(x, y) & \text{si } C_v(x, y) \geq C_h(x, y) \end{cases} \quad (2.8)$$

DM et C sont respectivement les cartes profondeurs et de confiances unifiées. C_v , DM_v et C_h , DM_h sont respectivement leur composantes horizontales et verticales précédemment calculées.

Les cartes de profondeurs et de confiances unifiées sont présentées Figure 33 avec leurs échelles de couleur respectives.

Dans l'ensemble, le meilleur des deux directions semble effectivement avoir été conservé. On devine la forme des objets présents dans ces scènes avec les cartes de profondeur sans pour autant conserver trop de détails. Ceux-ci sont d'ailleurs voués à disparaître dans une carte de profondeur.

Le résultat est assez proche des cartes de profondeurs idéales même s'ils restent tout de même relativement bruités. La pertinence est généralement faible lorsque les estimations de profondeurs sont mauvaises, cela confirme l'intérêt de générer des cartes de confiance associées aux cartes de distance.

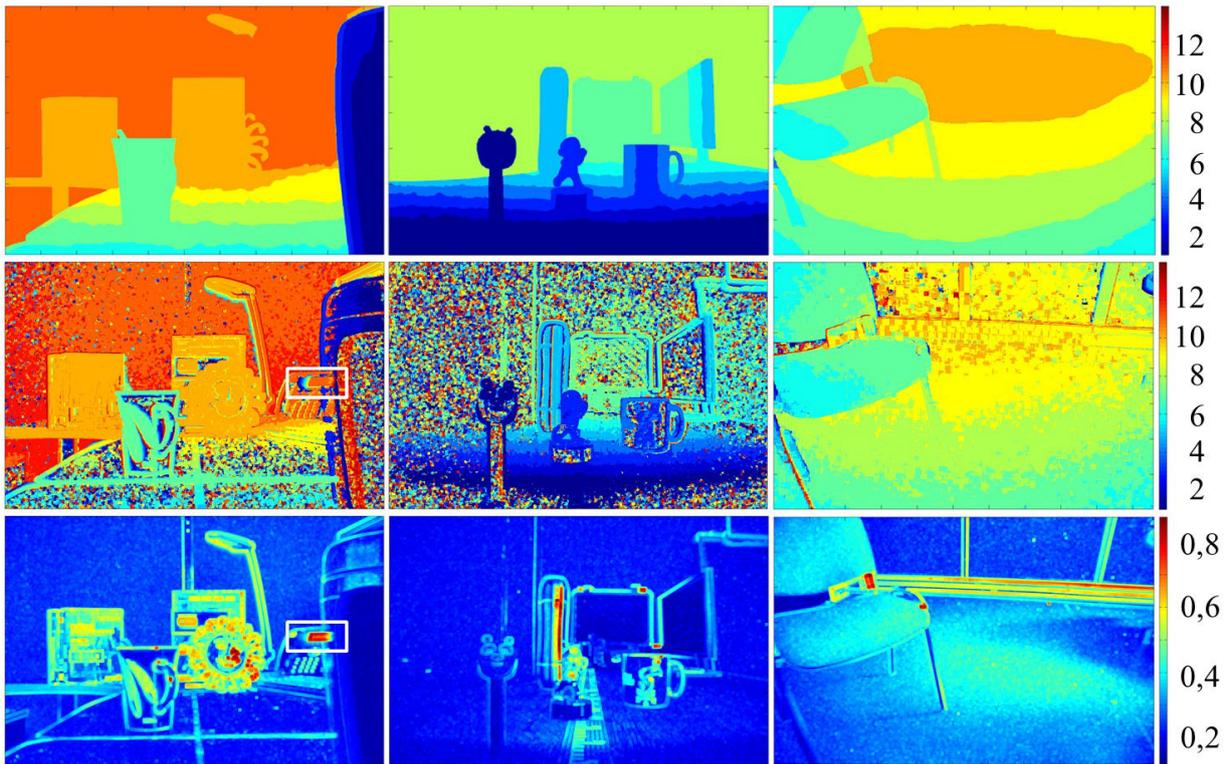


FIGURE 33 : CARTES DE PROFONDEURS IDEALES (HAUT) ET EXTRAITES UNIFIEES (MILIEU) AINSI QUE LEUR CARTES DE CONFIANCES UNIFIEES ASSOCIEES (BAS) POUR LES 3 SCENES ($w_x=w_y=16$)

2.4.1.4 Observations

Des effets de superpositions d'objets sont visibles sur ces images. C'est flagrant sur les cartes Soleil dans lesquelles l'écran du téléphone semble traverser le thermos qui est au premier plan (cadre blanc Figure 33). Ces effets, principalement dus au zoom de focalisation, sont particulièrement impactants pour des objets éloignés les uns des autres dans la scène qui se chevauchent dans l'image.

Le zoom de focalisation est également responsable d'un effet de traîne autour des bords des divers objets de la scène. Pour exemple, une imagerie autour de la lampe de bureau est présentée sur la Figure 34 et montre une distance erronée (halo cyan) autour de la lampe.

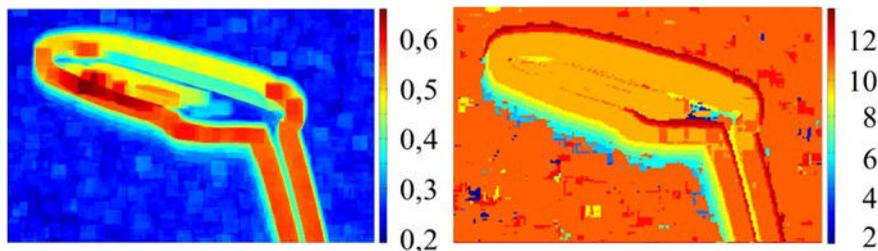


FIGURE 34 : EXEMPLE D'EFFET DU ZOOM DE FOCALISATION, CONFIANCE (DROITE) PROFONDEUR (GAUCHE) ($w_x=w_y=16$)

La vraisemblance est aussi légèrement impactée mais, mis à part le fond uniforme, on voit qu'elle reste la plus forte pour l'évaluation qui est la meilleure.

Si l'application le permet, c'est-à-dire si les capacités de calcul sont suffisantes ou encore si le temps de calcul n'est pas critique, une première étape de réaligement des images pourrait être réalisée avant l'extraction de profondeur. Dans la Figure 35, les images du cube ont précédemment été réalignées via l'outil imregister de Matlab. On peut voir par comparaison des Figure 34 et Figure 35 que cela réduit les effets de traîne autour des objets.

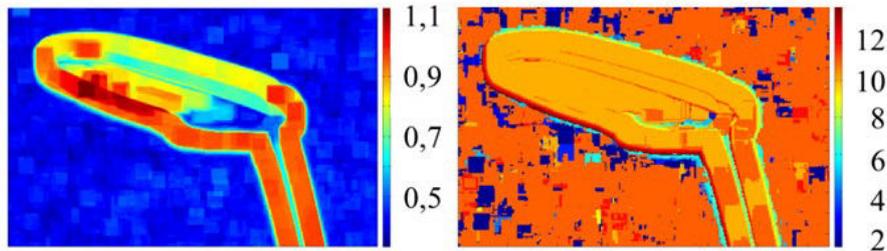


FIGURE 35 : EXEMPLE AVEC LES IMAGES SOURCES REALIGNEES CONFIANCE (DROITE) PROFONDEUR (GAUCHE)
($w_x=w_y=16$)

Notons également que les contours des objets se retrouvent épaissis dans ces cartes de profondeur et de confiance. Cela provient du traitement par fenêtre glissante et du calcul du critère de netteté. En effet, pour une coordonnée particulière de l'image, et une fenêtre de traitement 16×16 , il existe 256 blocs de 16×16 pixels incluant cette coordonnée. Le contenu des images étant relativement homogène, ces 256 blocs de pixels ont de fortes chances de se voir attribuer la même estimation de profondeur. C'est un défaut lié à l'extraction du maximum et du minimum par bloc. Plus la fenêtre de traitement est grande, plus la dilatation du contour des objets dans les cartes de profondeur et de confiance sera importante. Ce défaut devient cependant un avantage pour l'uniformisation des cartes dans les zones texturées puisque qu'il constitue une estimation véridique pour de nombreux pixels.

2.4.1.5 Construction d'images totalement focalisées

Pour une bonne sensibilité de l'algorithme, la profondeur de champ, qui est étudiée en annexe 3, doit être réduite le plus possible. Ceci conduit à des images contenant de grandes zones floues, d'où la nécessité de recombinaison de différents plans de focalisation afin de construire une image nette en tout point, dite All in Focus (AiF).

Avec des configurations optiques particulières, comme l'utilisation de très petites ouvertures, on peut capturer directement de telles images. Cela nécessite que l'ensemble de la scène soit à une distance à laquelle la profondeur de champ est infinie. La plupart des systèmes imageur n'ont pas une distance hyperfocale très proche. Il est donc souvent impossible d'avoir à la fois des objets nets proches et éloignés sur une même image. En revanche, avec une connaissance de la carte de profondeur et un cube d'images multi-foyers, il est possible de générer une

image totalement focalisée, quelle que soit la distribution des objets dans la scène. En effet, la carte de profondeur donne des indices correspondant aux images les plus nettes du cube par zone du champ de vision. Il suffit donc de rassembler au sein d'une unique image tous les éléments nets du cube multi-foyers en utilisant la carte de profondeur. L'équation (2.9) résume ce procédé.

$$AiF(x, y) = Im(x, y, DM(x, y)) \quad (2.9)$$

AiF représente l'image totalement focalisée, Im le cube d'images sources et DM la carte de profondeur unifiée.

Les images AiF (Figure 36 bas) présentent des objets nets pour chacune des scènes en comparaison avec les échantillons des cubes multi-foyers présentés au-dessus.

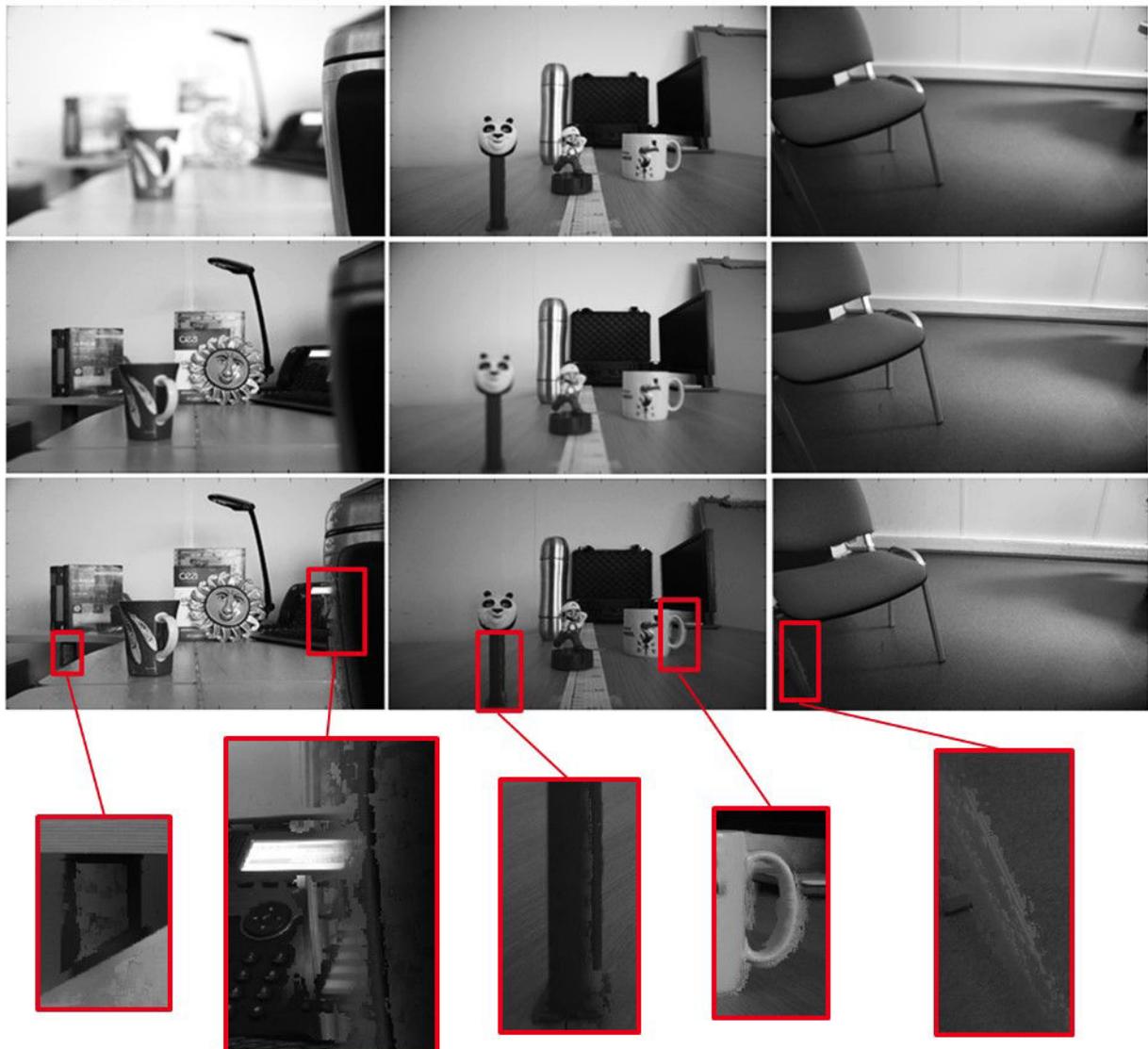


FIGURE 36 : DISTANCE FOCALE PROCHE (HAUT) ET ELOIGNEE (MILIEU) ET IMAGE ALL IN FOCUS CONSTRUITE AVEC NOTRE METHODE INTRA-IMAGE HAUTE RESOLUTION (BAS)

Cependant on observe des artefacts caractérisés par des textures imparfaitement rendues (cadres rouges) et quelques discontinuités sur les contours. Cela s'explique à la fois par le zoom de focalisation et le traitement par bloc qui grossit les objets sur la carte de profondeur et de ce fait, provoque des discontinuités sur l'image AiF en particulier lors de proximité dans l'image d'objets éloignés dans l'espace.

A titre indicatif, la Figure 37 montre une image AiF construite à partir du cube d'images « Soleil » réaligné sous Matlab. Le résultat est dans l'ensemble plus propre, ce qui nous motive à essayer de trouver une solution optique dans le chapitre 3 pour limiter le zoom de focalisation.



FIGURE 37 : IMAGE ALL IN FOCUS CONSTRUITE A PARTIR D'UN CUBE D'IMAGES ALIGNEES

2.4.2 Approche intra-image basse résolution

Nous venons de voir que notre méthode utilise deux gradients unidirectionnels pour extraire des données suivant deux directions avant de les combiner dans le but de produire des pixels RGB-Z-C.

Pour créer une architecture de traitement proche capteur, la méthode proposée est pertinente en termes de complexité de calcul, cependant elle est trop gourmande en espace mémoire. En effet, les deux chaînes d'extraction de données nécessitent chacune autant de mémoire qu'une chaîne classique utilisant le calcul du Laplacien [104] [66] [105] pour laquelle on aurait un résultat au prix d'une chaîne unique, bien que légèrement moins bon comme cela sera montré dans quelques pages. On va donc chercher dans ce paragraphe à minimiser l'empreinte mémoire.

2.4.2.1 Extraction des cartes de profondeur et de confiance

Plutôt que d'exploiter les dérivées suivant les deux axes principaux de l'image, nous proposons d'utiliser les coefficients horizontaux et verticaux au premier ordre d'une transformée en ondelettes de Haar 2D. Ainsi, on divise par 4 la taille des matrices de coefficients à partir desquelles le critère de netteté est calculé. L'extraction des coefficients de Haar est très peu complexe. Chaque coefficient est la moyenne par bloc non recouvrant de taille 2x2 pixels du gradient suivant les deux directions (2.10).

$$\begin{aligned}
 H_h(x, y) &= \frac{1}{4} \left(\begin{array}{l} Im(2x + 1, 2y) - Im(2x, 2y) + \\ Im(2x + 1, 2y + 1) - Im(2x, 2y + 1) \end{array} \right) \\
 H_v(x, y) &= \frac{1}{4} \left(\begin{array}{l} Im(2x, 2y + 1) - Im(2x, 2y) + \\ Im(2x + 1, 2y + 1) - Im(2x + 1, 2y) \end{array} \right) \quad (2.10)
 \end{aligned}$$

Le traitement sur deux chaînes directionnelles nécessite désormais plus de 2 fois moins de mémoire qu'un algorithme plus classique.

On calcule le critère de netteté comme précédemment, suivant les deux directions, avec l'équation 2.11.

$$S_{crit\ h/v}(x, y, z) = \max_{X,Y} \left(H_{h/v}(X, Y, z) \right) - \min_{X,Y} \left(H_{h/v}(X, Y, z) \right) \quad (2.11)$$

Où (X,Y) est un élément de la fenêtre de dimension $w_h \times w_w$ centrée sur (x,y). Les cartes de profondeurs et de vraisemblances sont ensuite unifiées par la méthode du maximum de vraisemblance présentée plus haut.

Les résultats sont présentés sur les Figure 38, pour un traitement par fenêtres glissantes de 16x16 pixels.

En tenant seulement compte des cartes de profondeur de la Figure 38, on constate que cette méthode « basse résolution » donne des résultats plus homogènes que celle « haute résolution » (Figure 33 page 47). Les cartes de distance semblent également plus proches de la réalité. En particulier sur les zones texturées telles que la table sur les images du soleil (droite) et du panda (milieu) ou le mur sur l'image du soleil. Les arrière-plans sont aussi moins bruités et mieux estimés, ce qui confère un net avantage pour la génération d'images AiF.

On peut remarquer aussi que la résolution est plus grossière augmentant ainsi l'épaisseur du contour des objets.

Les cartes de confiance semblent avoir des valeurs légèrement inférieures qu'avec la méthode haute résolution. Cela s'explique par une diminution de la dynamique due au moyennage effectué lors de la transformée de Haar.

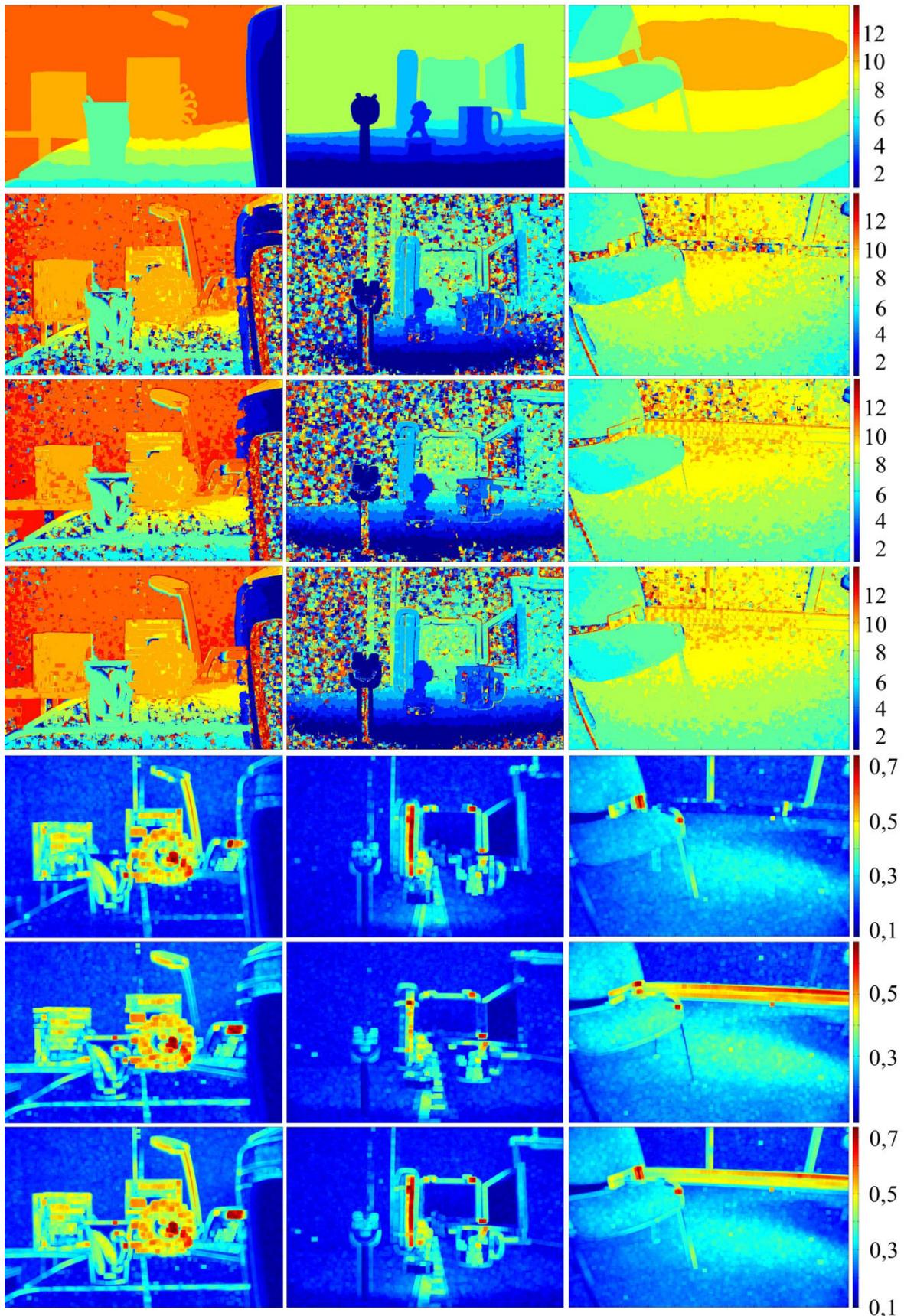


FIGURE 38 : 1ERE LIGNE, CARTES DE PROFONDEUR IDEALES ; LIGNES 2 A 3, CARTES DE PROFONDEUR HORIZONTALES, VERTICALES ET UNIFIEES ; LIGNES 4 A 7 CARTES DE CONFIANCE HORIZONTALES, VERTICALES ET UNIFIEES ($w_x=w_y=16$)

2.4.2.2 Construction d'images totalement focalisées

Une fois encore, on utilise l'équation (2.9) pour reconstruire les images focalisées en tout point de la scène. Pour cela, on effectue un rééchantillonnage de la carte de profondeur afin de retrouver une résolution identique aux images d'entrées. Celui-ci consiste en un produit de Kronecker simple avec une matrice 2x2 composée uniquement de 1 (2.12).

$$DM_{resampled} = DM_{orig} \otimes \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (2.12)$$

Les images All in Focus (Figure 39) générées avec cette méthode sont de bonne qualité. On observe toujours des artefacts de textures (cadres rouges) mais dans l'ensemble moins flagrants qu'avec la méthode haute résolution. Notons d'ailleurs l'amélioration du pied de la chaise et du support du panda. Les discontinuités sur certains contours sont en revanche légèrement plus marquées (flèche orange), ceci étant dû à l'utilisation de fenêtre de traitement de même taille entre les deux méthodes alors que les cubes de critère de netteté sont plus petits suite à l'utilisation des coefficients de Haar.

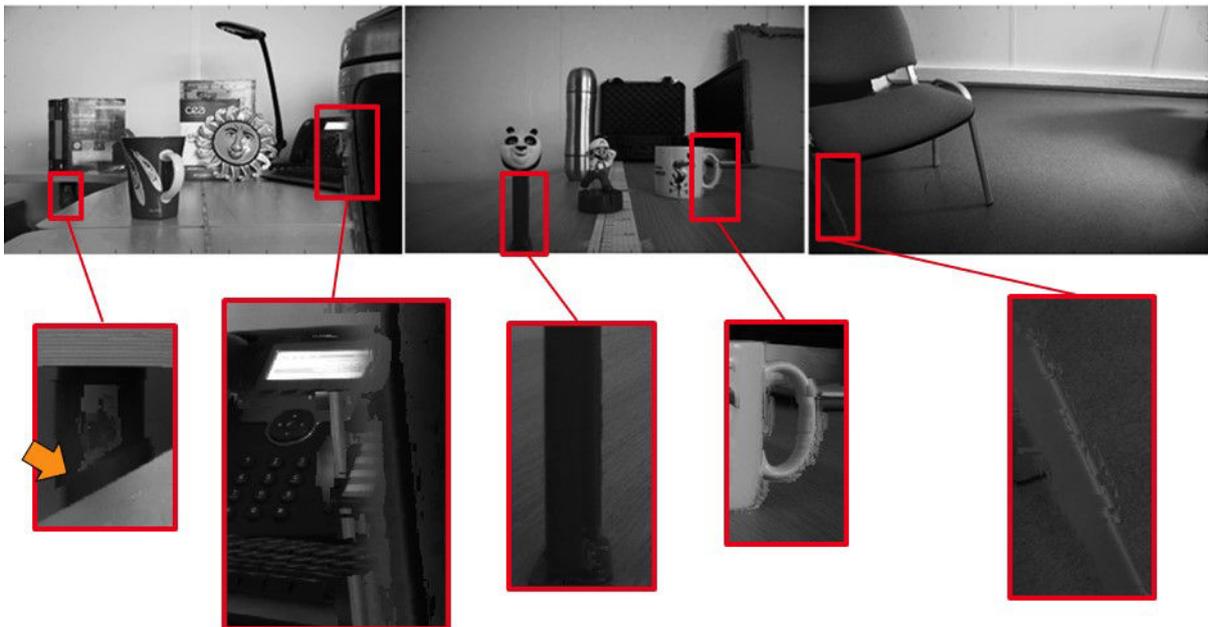


FIGURE 39 : IMAGE ALL IN FOCUS CONSTRUITE AVEC NOTRE METHODE INTRA-IMAGE BASSE RESOLUTION

2.4.3 Approche intra-image basse résolution par fenêtres de traitement non recouvrantes

L'extraction du critère de netteté était jusqu'à présent réalisé dans une fenêtre glissante sur une matrice contenant les hautes fréquences de l'image. Les mémoires nécessaires pour le traitement (carte de profondeur et de confiance) sont donc quasiment de la même taille que la matrice des coefficients de Haar. Pour des raisons économiques, nous avons tout intérêt à réduire le plus possible la mémoire nécessaire. Nous proposons donc d'extraire le critère de

netteté à partir de fenêtres non recouvrantes, c'est-à-dire par blocs distincts de pixels. Le gain en mémoire est énorme, en revanche, il faut vérifier la qualité du résultat de cette simplification.

2.4.3.1 Extraction des cartes de profondeur et de confiance

Nous nous sommes donc placés dans les mêmes conditions que pour l'évaluation des algorithmes, en prenant des fenêtres de traitement de 16×16 pixels et avons extrait les cartes de profondeur pour nos trois familles d'images (Figure 40) en utilisant des fenêtres de traitement non recouvrantes plutôt que des fenêtres glissantes. Les cartes de confiance sont montrées sur la Figure 41.

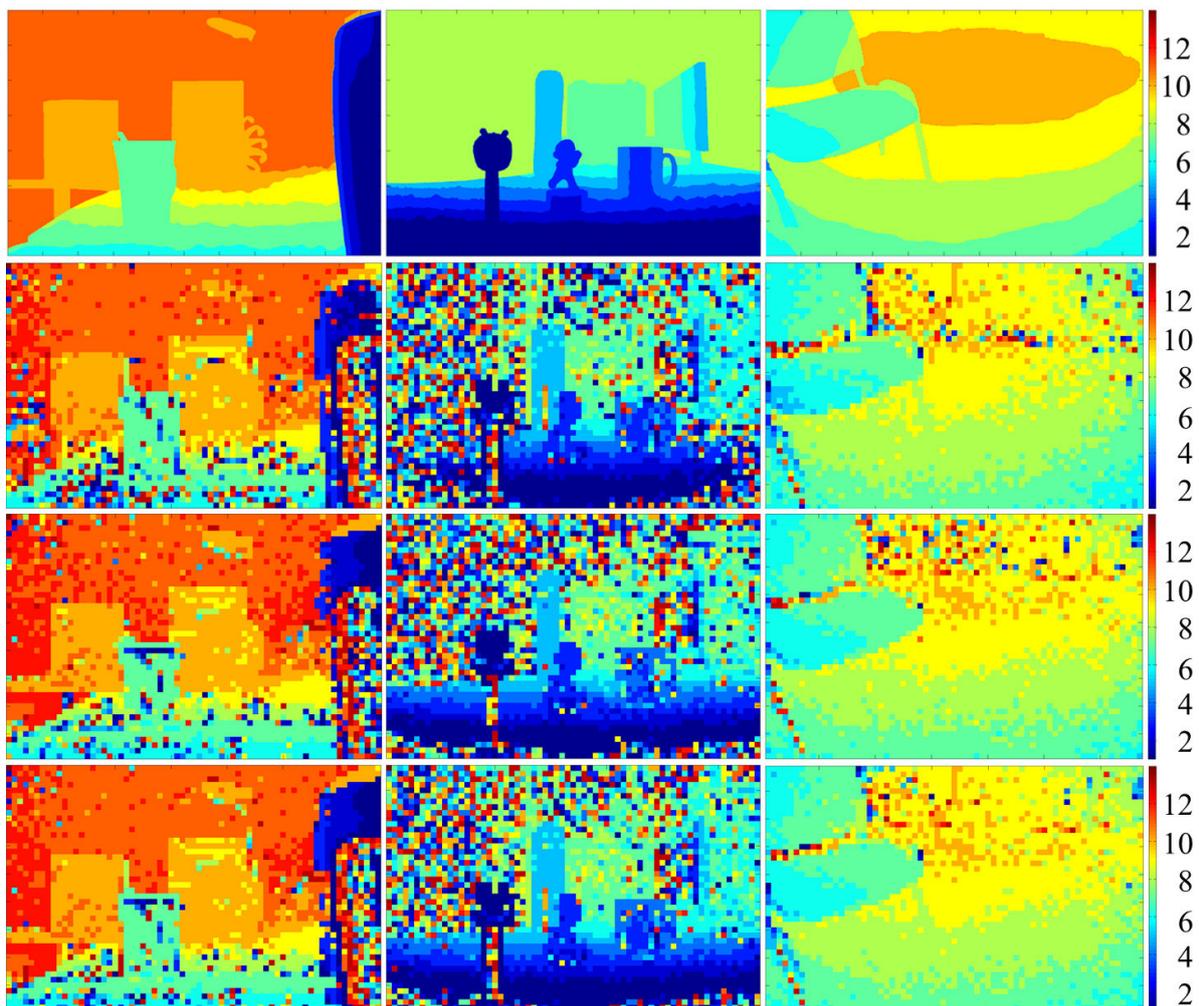


FIGURE 40 : 1ERE LIGNE, CARTES DE PROFONDEUR IDEALES ; LIGNES 2 A 3, CARTES DE PROFONDEUR HORIZONTALES, VERTICALES ET UNIFIEES (FENETRES NON RECOUVRANTES, $w_x=w_y=16$)

On observe une perte importante de résolution cependant une grande partie de l'information est conservée et reste correctement estimée.

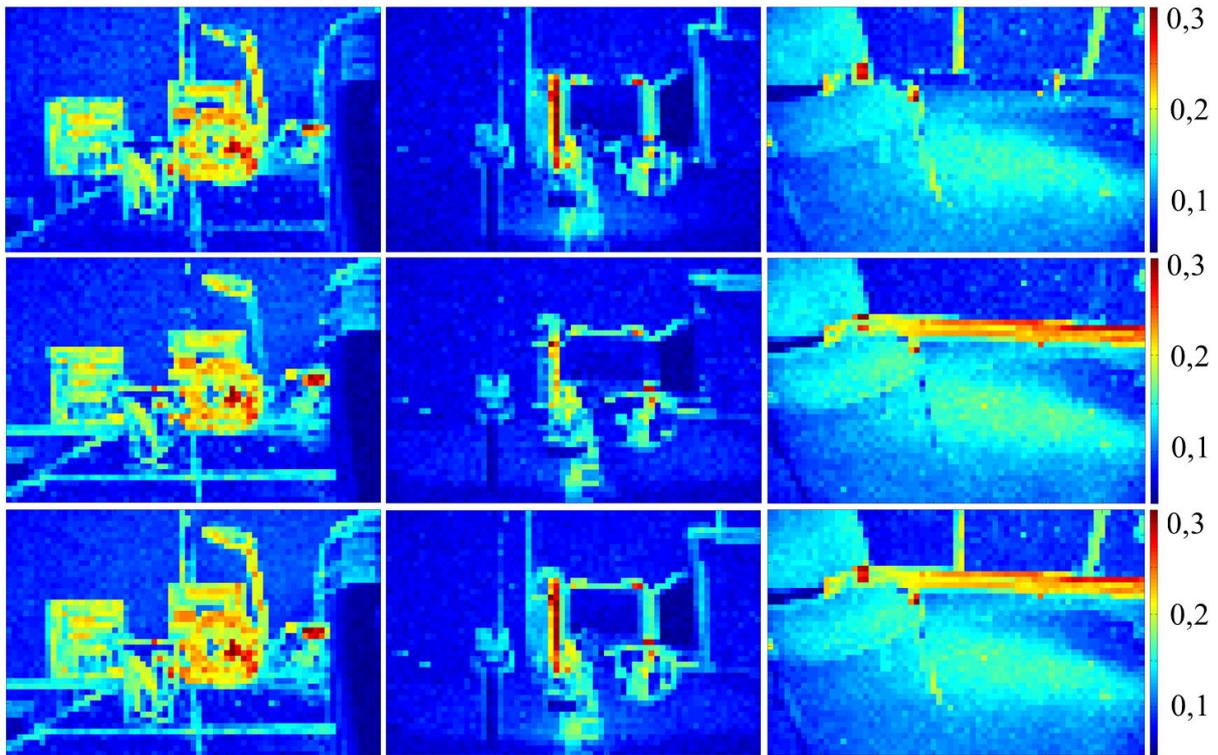


FIGURE 41 : LIGNES 1A 3, CARTES DE CONFIANCE HORIZONTALES, VERTICALES ET UNIFIEES (FENETRES NON RECOUVRANTES, $w_x=w_y=16$)

2.4.3.2 Construction d'images totalement focalisées

L'équation (2.9) permet une fois de plus de reconstruire les images totalement focalisées (Figure 42) après sur-échantillonnage des cartes de profondeur.

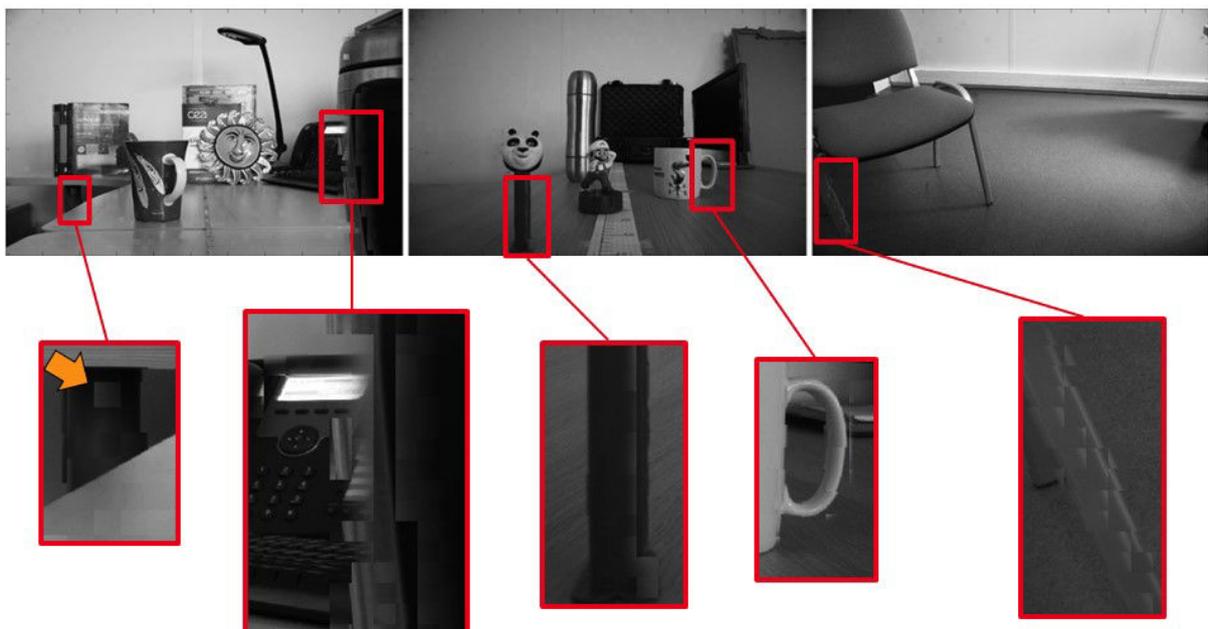


FIGURE 42 : IMAGE ALL IN FOCUS CONSTRUITE AVEC NOTRE METHODE INTRA-IMAGE BASSE RESOLUTION PAR FENETRES DE TRAITEMENT NON RECOUVRANTES

Une fois encore, les images focalisées en tout point reconstruites sont très correctes. Cependant, il semble qu'utiliser des fenêtres non recouvrantes fasse apparaître de plus gros artefacts ; ceux-ci se manifestent par l'apparition de carrés très marqués (flèche orange sur la Figure 42). Si on compare les images totalement focalisées issues de ces deux méthodes intra-image basse résolution (Figure 39 et Figure 42), on note qu'il n'y a pas de très grosses différences, les discontinuités de certains contours sont presque moins marquées. Des objets, par exemple l'anse de la tasse, sont mieux reconstruits. Dans l'ensemble on reconnaît sans problème les éléments de chacune des scènes. Cette proposition d'utiliser des fenêtres de traitement non recouvrantes semble être un bon compromis pour diminuer l'empreinte mémoire sans trop diminuer les performances qualitatives du système.

Dans la suite de ce manuscrit, lorsque cela n'est pas précisé, le traitement est effectué par fenêtre glissante.

2.5 Comparaison des algorithmes de Depth from Focus

Nous avons comparé nos trois méthodes d'extraction de profondeurs entre elles ainsi qu'avec quatre autres de l'état de l'art. La plupart sont connues depuis longtemps mais contrairement à des méthodes plus récentes, elles sont assez similaires aux nôtres en termes de complexité calculatoire.

Nous avons repris la méthode de Yokota et al [104] [66] dans un premier temps sans tenir compte des stratégies proposées pour améliorer le rendu de la carte de profondeur. Ces auteurs se sont aussi intéressés de près à l'implémentation matérielle d'algorithme de DfF. Leur méthode se rapproche de l'extraction d'un Laplacien par pixel avec sélection du maximum à travers le cube d'images pour chacun d'eux. Cela se traduit mathématiquement par l'équation (2.13)

$$DM(x, y) = \underset{z}{\operatorname{argmax}} \left(\begin{array}{l} |I(x+1, y, z) + I(x-1, y, z) - 2I(x, y, z)| + \\ |I(x, y+1, z) + I(x, y-1, z) - 2I(x, y, z)| \end{array} \right) \quad (2.13)$$

L'article [103] montre une étude exhaustive de critères de netteté pour des applications d'autofocus. Les auteurs concluant que le critère SMD (équation (2.2) page 35) donne les meilleures performances, nous allons la comparer aux méthodes proposées dans ces travaux. Nous avons expérimenté des techniques plus classiques basées sur l'extraction du maximum dans une fenêtre glissante de taille 16x16 sur le Laplacien des images extrait par filtrage avec le noyau de convolution (2.14), ou encore sur la norme du gradient (2.15). Soit en tout 4 méthodes complémentaires des 3 présentées dans ce mémoire, 4 si on tient compte de l'étude en annexe 2.

$$\frac{1}{8} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (2.14)$$

$$N_{grad}(x, y) = \sqrt{\left(\frac{\partial Im(x,y,z)}{\partial x}\right)^2 + \left(\frac{\partial Im(x,y,z)}{\partial y}\right)^2} \quad (2.15)$$

Dans ce qui suit, nous comparerons qualitativement et quantitativement les performances des algorithmes à la fois pour la carte de profondeur et les images focalisées en tous points.

2.5.1 Comparaison des cartes de profondeur

2.5.1.1 Comparaison quantitative

Afin d'évaluer les performances quantitativement de tous ces algorithmes, et de pouvoir les comparer, nous avons calculé le taux d'estimations correctes τ_c avec l'équation (2.16). Ce coefficient calcule la similitude entre les cartes de profondeur extraite et les carte idéales.

$$\tau_c(DM, s) = \frac{1}{h \cdot w} \cdot \text{card}(|DM - tDM| \leq s) \quad (2.16)$$

tDM représente une carte de profondeur idéale (vérité de terrain), DM une carte de profondeur obtenue par l'une des méthodes comparées, h et w respectivement la hauteur et la largeur des cartes de profondeur et s un seuil d'erreur acceptable. Pour les cas « basse résolution », les cartes de profondeur ont été sur-échantillonnées de sorte à avoir des dimensions identiques à celles des cartes idéales.

Les cartes de profondeur idéales utilisées dans ces travaux ont été réalisées à partir de cartes de profondeur lissées puis manuellement éditées avec un logiciel de dessin. Nous avons essayé de rester aussi proches que possible de la réalité en nous basant sur des cartes extraites avec certains des algorithmes présentés, et une analyse visuelle du cube d'images. Elles ne sont par conséquent pas à 100% fidèles à la réalité mais cela semble suffisant pour faire des comparaisons visuelles et quantitatives. Elles représentent des profondeurs qui seraient extraites par un algorithme DfF avec une résolution de distance à 14 niveaux de profondeurs, pour un balayage de focus correspondant aux cubes d'images tests.

Le paramètre s , représentant la tolérance à l'erreur, est légitime en raison de la réalisation manuelle de ces cartes idéales. Les frontières entre les niveaux de profondeur sont relativement approximatives, il est donc pertinent d'accepter une marge d'erreur. De plus, en fonction du système optique utilisé, de la profondeur du cube multi-foyers, de la précision du changement de mise au point, certaines évaluations peuvent être légèrement décalées alors qu'avec un système parfaitement calibré ce décalage disparaîtrait. Le paramètre s permet donc aussi de tenir compte de ces biais de mesure.

La mesure τ_c a permis de faire une première évaluation en % pour les 7 méthodes mentionnées (Tableau 3) avec une erreur tolérée de plus ou moins 1 niveau de profondeur, soit pour $s=1$.

Tableau 3 : Taux d'estimations correctes en % pour différentes méthodes d'extraction de ($s=1$)

Image	<i>Soleil</i>	<i>Panda</i>	<i>Chaise</i>
Notre méthode basse résolution blocs distincts	79.60	60.38	92.56
Notre méthode intra-image en basse résolution	80.54	61.96	93.32
Notre méthode intra-image en pleine résolution	72.33	52.83	91.05
Max du module du gradient	65.58	46.44	88.14
Max du Laplacien	65.13	43.34	89.16
SMD	70.14	57.77	91.17
Yokata (première étape)	35.11	25.20	58.45

D'après le Tableau 3, quatre méthodes ressortent, les trois approches intra-image présentées dans ce manuscrit et celle basée sur le SMD. Ces résultats restent à généraliser mais sur les trois scènes testées, ces techniques donnent des estimations toujours meilleures que les 3 autres techniques expérimentées. L'utilisation des coefficients d'ondelette de Haar (méthode basse-résolution) permet d'améliorer le score τ_c , malgré une légère perte de résolution. L'utilisation de blocs de pixels distincts donne des résultats très corrects : on reste très proche du meilleur résultat précédemment obtenu.

2.5.1.2 Comparaison visuelle

Les comparaisons quantitatives présentées plus haut ne tiennent pas compte des cartes de confiance extraites en même temps que la profondeur. Nous avons jugé important de mesurer les performances en fonction de celles-ci aussi. Pour cela, nous avons extrait des cartes de confiance pour chacune des méthodes, excepté Yokota et al. qui dispose de son propre mécanisme de filtrage, en considérant la dynamique de leur critère de netteté. Celles-ci ont permis de construire des filtres binaires appliqués sur les cartes de profondeur correspondantes : seules les estimations de distance dont la confiance est supérieure à la médiane des confiances sont conservées. Les cartes de profondeur résultantes sont présentées Figure 43.

Pour toutes ces images, comme nous nous y attendions, les éléments qui ressortent systématiquement correspondent aux zones les plus contrastées.

Le Tableau 4 montre le taux d'estimations correctes τ_c mesuré en considérant seulement les données filtrées. De ce fait, 50% représente le taux maximum atteignable pour toutes les cartes de profondeur filtrées par la vraisemblance.

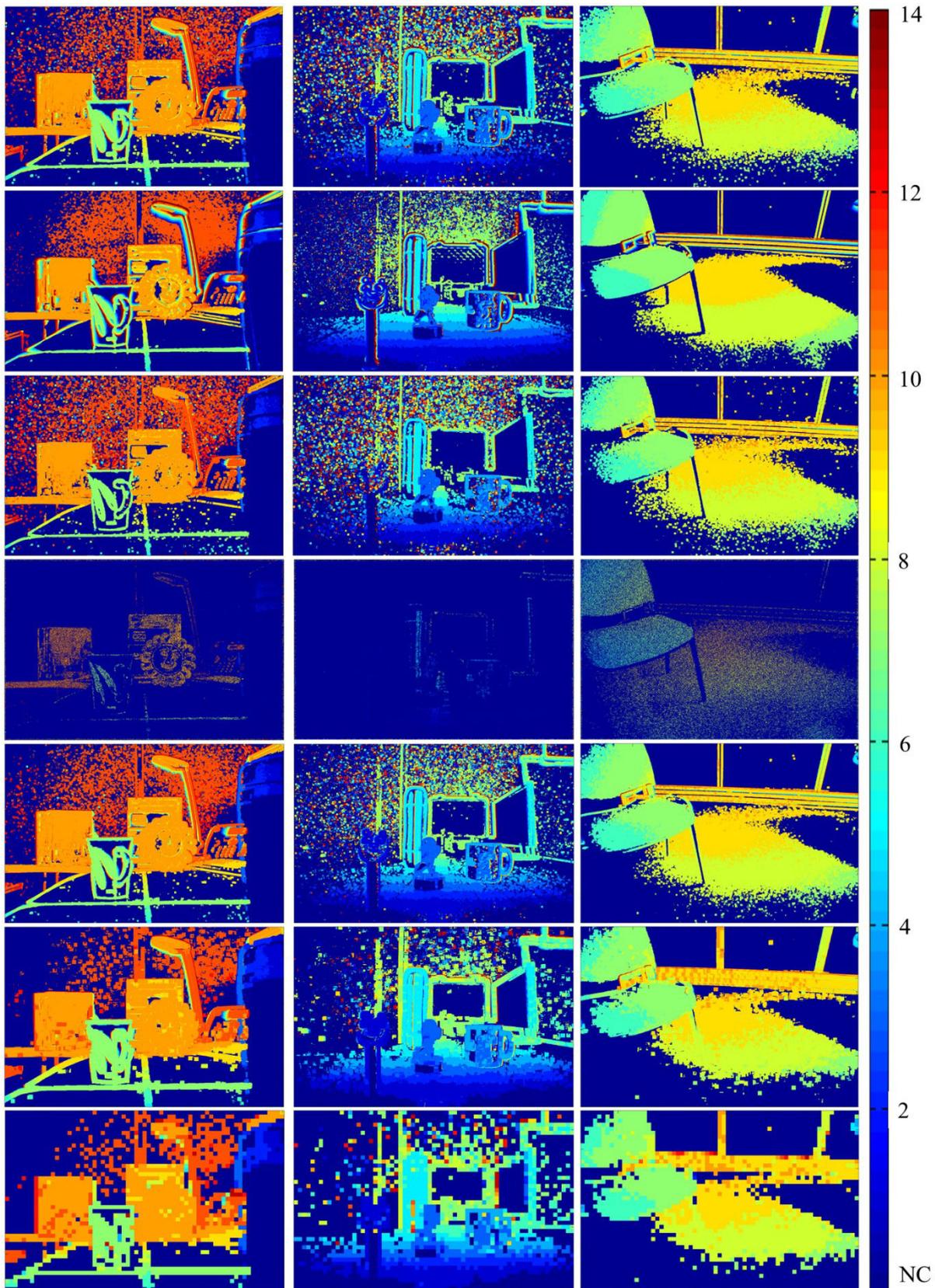


FIGURE 43 : CARTES DE PROFONDEUR FILTRES PAR LE NIVEAU DE CONFIANCE POUR CHACUNE DES SCENES. DE HAUT EN BAS : MAXIMUM DU MODULE DU GRADIENT, MAXIMUM DU SMD, MAXIMUM DU LAPLACIEN, METHODES YOKOTA, METHODE INTRA-IMAGE HAUTE RESOLUTION, METHODE INTRA-IMAGE BASSE RESOLUTION ET METHODE INTRA-IMAGE BASSE RESOLUTION AVEC FENETRES NON RECOUVRANTES. LA COULEUR BLEU MARINE REPRESENTE UNE ABSENCE D'ESTIMATION.

Tableau 4 : Taux d'estimations correctes en % après filtrage ($s=1$)

Image	<i>Soleil</i>	<i>Panda</i>	<i>Chaise</i>
Notre méthode basse résolution blocs distincts	42.47	38.21	45.18
Notre méthode intra-image en basse résolution	43.62	38.50	45.19
Notre méthode intra-image en pleine résolution	40.99	31.48	45.71
Max du module du gradient	39.22	29.40	45.12
Max du Laplacien	41.09	27.17	46.47
SMD	36.20	33.07	44.80
Yokata (première étape)	5.46	1.46	15.10

Le calcul du maximum du module du gradient donne des résultats visuels assez similaires à la méthode bidirectionnelle haute résolution, bien qu'ils soient légèrement plus bruités. Toutes deux sont légèrement affectées par le zoom de focalisation. Le taux d'estimations correctes donne cependant l'avantage à l'algorithme développé au cours de cette thèse.

Malgré un bon score pour la comparaison quantitative de cartes de profondeur non filtrées, le traitement à partir du SMD est très sensible au zoom de focalisation. On voit en effet un artefact de dégradé de profondeurs autour de la plupart des contours (2^{ème} ligne Figure 43). Cela détériore le taux τ_c comparativement aux autres méthodes comme le montre le Tableau 4. Sachant que les données affichées sur la Figure 43 sont celles avec une vraisemblance relativement importante, cela risque d'être problématique dans le cadre d'applications critiques qui auraient pour seule capacité d'utiliser ces données brutes.

Les cartes de profondeur basées sur la recherche du maximum du Laplacien ne semblent quasiment pas affectées par le zoom de focalisation. On remarque qu'elles sont légèrement plus bruitées que les autres dans les zones uniformes. En outre, avec le lot d'images Panda (Figure 43), elle ne parvient pas à faire ressortir le panda au premier plan ce qui peut être très problématique dans le cadre d'évitement d'obstacle.

Avec la méthode Yokota, le résultat semble précis au niveau de l'estimation mais la carte de profondeur est très creuse avec un faible nombre d'estimations comme le prouve la mesure de τ_c . Cela demande un effort de reconstruction conséquent.

L'utilisation des coefficients de Haar avec des blocs toujours de taille 16×16 pour le calcul du critère dilate visiblement les contours. En revanche, on observe moins de bruit qu'avec les autres méthodes. Elles présentent dans l'ensemble les meilleurs taux d'estimations correctes.

2.5.2 Comparaison des images All in Focus

Dans ce paragraphe, les méthodes sont comparées sur leur capacité à fournir une image AiF. Les cartes de profondeurs extraites sont utilisées sans aucun filtre afin de construire les images

totallement focalisées correspondantes aux trois cubes d'images monochromes. La majorité du bruit dans les cartes de profondeur n'a en réalité qu'un faible impact sur l'image AiF. En effet, nous avons vu que les zones bruitées ont une faible vraisemblance ce qui signifie qu'elles sont relativement uniformes à travers le cube multi-foyers. N'importe quelle image du cube peut donc être utilisée pour générer les pixels de l'image totalement focalisée dans ces zones-là. Les images résultantes sont présentées sur la Figure 45.

Au premier regard, les images reconstruites avec les approches intra-image développées dans ces travaux semblent de meilleure qualité que celles générées avec les autres méthodes. Les approches basses résolutions paraissent même moins bruitées, malgré une moins bonne résolution de leur carte de profondeur et la dilatation des contours qu'elles engendrent.

La méthode de Yokota et al. (4^{ème} ligne Figure 45) ne donne pas des images qui puissent être qualifiées de totalement focalisées, le panda par exemple semble complètement flou. Une fois encore, si l'on utilise leur méthodologie complète de filtrage et de remplissage, on aurait probablement une image de bien meilleure qualité mais la complexité de ce type de filtrage serait alors bien trop grande pour permettre qu'il soit embarqué.

L'utilisation du critère SMD (2^{ème} ligne Figure 45) donne un résultat extrêmement impacté par le zoom de focalisation comme le montre les flèches sur les images de la Figure 44. Ce type d'effet est également visible avec l'extraction basée sur le module du gradient (lasso rouge Figure 44), bien que cela soit moins flagrant.

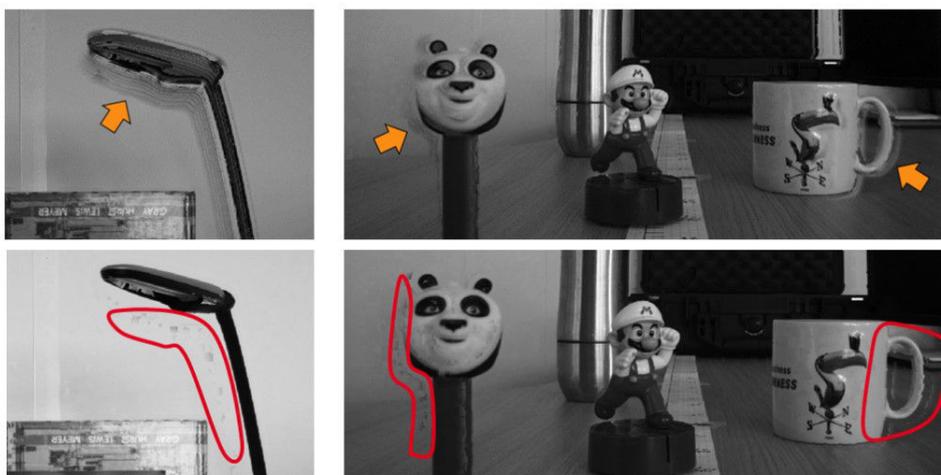


FIGURE 44 : ILLUSTRATION DE L'IMPACT DU ZOOM DE FOCALISATION AVEC LA METHODE SMD (HAUT) ET L'EXTRACTION A PARTIR DU MODULE DU GRADIENT (BAS)

Les quatre autres méthodes permettent une génération d'images totalement focalisées d'assez bonne qualité. En regardant de plus près, l'image issue du calcul du maximum du Laplacien présente des artefacts dans des zones d'intérêt comme sur le visage du panda. Ceux-ci se trouvant dans la carte de profondeur, ils impactent nécessairement l'image AiF et en dégradent la qualité.



FIGURE 45 : IMAGES TOTALEMENT FOCALISEES ISSUES DU MAXIMUM DU MODULE DU GRADIENT (1^{ERE} LIGNE), SMD (2^{EME} LIGNE), DU MAXIMUM DU LAPLACIEN (3^{EME} LIGNE), DE YOKOTA ET AL (4^{EME} LIGNE), DES APPROCHES INTRA-IMAGE HAUTE RESOLUTION (5^{EME} LIGNE), BASSE RESOLUTION (6^{EME} LIGNE) ET FENETRES NON RECOUVRANTES (7^{EME} LIGNE)

On conclut de ces comparaisons que même si on a une perte de résolution apparente, les méthodes intra-image basse résolution présentées dans ce manuscrit donnent des résultats très corrects et amplement suffisants pour être utilisés directement et rapidement par un système autonome.

2.6 Amélioration du rendu

Les cartes de profondeur présentées sont toutes relativement bruitées dans les zones peu texturées et il peut être avantageux d'améliorer leur rendu. L'intérêt d'une amélioration est évident pour une visualisation humaine, que ce soit simplement sur la carte de profondeur ou pour la construction d'image 3D. Pour une machine, cela peut améliorer l'appréhension de la scène mais il faut être certain que la stratégie d'amélioration est efficace pour l'application, et ne vient pas introduire d'erreur critique comme la suppression de petits obstacles.

Des méthodes performantes existent dans la littérature pour la suppression de bruit de mesure et l'amélioration de l'estimation. Ces méthodes tentent de deviner les valeurs manquantes ou aberrantes par diverses stratégies. On part toujours de différents a priori selon lesquels les formes « régulières » (ronds carrés et rectangles) appartiennent au même plan de profondeur, les zones de teintes uniformes font généralement partie d'un même objet... et ainsi on peut faire de la segmentation d'image pour combler les vides [69]. La carte de confiance peut aider à l'amélioration des estimations par exemple pour élaborer une stratégie de propagation itérative de valeurs « sûres » du type [70]. Yokota et al, par exemple, présente dans [66] et [104] une solution en 3 étapes. Les calculs proposés comprennent une recherche itérative de conditions particulières pour chaque pixel dans un voisinage de taille évolutive. Le nombre d'itérations pour un même pixel dépend du contenu de la scène. L'inconvénient des solutions itératives est qu'elles nécessitent une puissance de calcul importante dès lors qu'on souhaite un traitement rapide. Elles sont par conséquent peu adaptées pour faire du traitement matériel à la volée.

Les filtres moyenneur (pondérés ou non) ont également été proscrits car ils conduisent à des indices de profondeur réels qui ne correspondent à aucune distance connue. Nous avons donc opté pour l'utilisation d'un filtre médian qui permet de lisser les fluctuations d'une matrice tout en conservant des valeurs déjà présentes dans celle-ci. Le noyau du filtre médian ne doit pas être trop grand pour ne pas déformer outre mesure les éléments de la scène d'une part, encore que cela puisse être intéressant dans le cas de navigation autonome pour générer rapidement une marge de sécurité par exemple, mais surtout pour éviter de faire disparaître de petits objets proches. La Figure 46 présente deux exemples de lissage avec deux tailles de noyau de filtrage, 21×21 et 41×41 pixels pour la méthode intra image basse résolution avec fenêtre glissante.

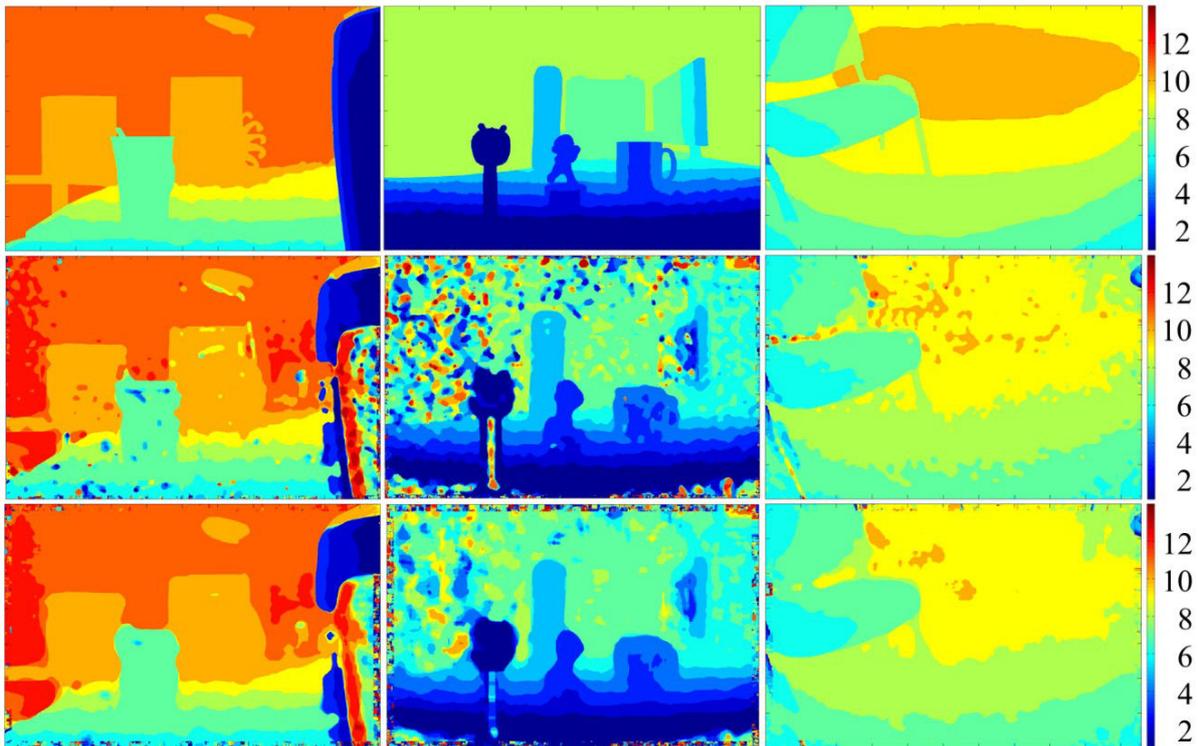


FIGURE 46 : CARTES DE PROFONDEUR IDEALES (1^{ERE} LIGNE) ET EXTRAITES PAR LA METHODE INTRA-IMAGE BASSE RESOLUTION FILTRES PAR UN FILTRE MEDIAN DE NOYAU 21×21 (2^{EME} LIGNE) ET 41×41 (3^{EME} LIGNE)

Nous avons réutilisé notre critère de comparaison quantitatif $\tau_{c,vu}$ dans la section 2.5.1.1 pour évaluer l'amélioration apportée par le filtrage médian sur les estimations de profondeur. Le Tableau 5 reporte l'ensemble des mesures pour toutes les méthodes présentées. On voit qu'avec le filtre médian on se rapproche encore plus de la vraie profondeur de la scène.

Tableau 5 : Taux d'estimations correctes en % (s=1)

Image	<i>Soleil</i>	<i>Panda</i>	<i>Chaise</i>
Notre méthode en basse résolution avec le filtre médian 41x41	86.64	84.00	95.96
Notre méthode en basse résolution avec le filtre médian 21x21	85.31	75.32	95.91
Notre méthode intra-image en basse résolution	80.54	61.96	93.32
Notre méthode intra-image en pleine résolution	72.33	52.83	91.05
Max du module du gradient	65.58	46.44	88.14
Max du Laplacien	65.13	43.34	89.16
SMD	70.14	57.77	91.17
Yokata (première étape)	35.11	25.20	58.45

A l'aide de carte de profondeur « améliorées » nous pouvons reconstruire des modèles 3D focalisés en tous points, ils sont présentés sur la Figure 47.

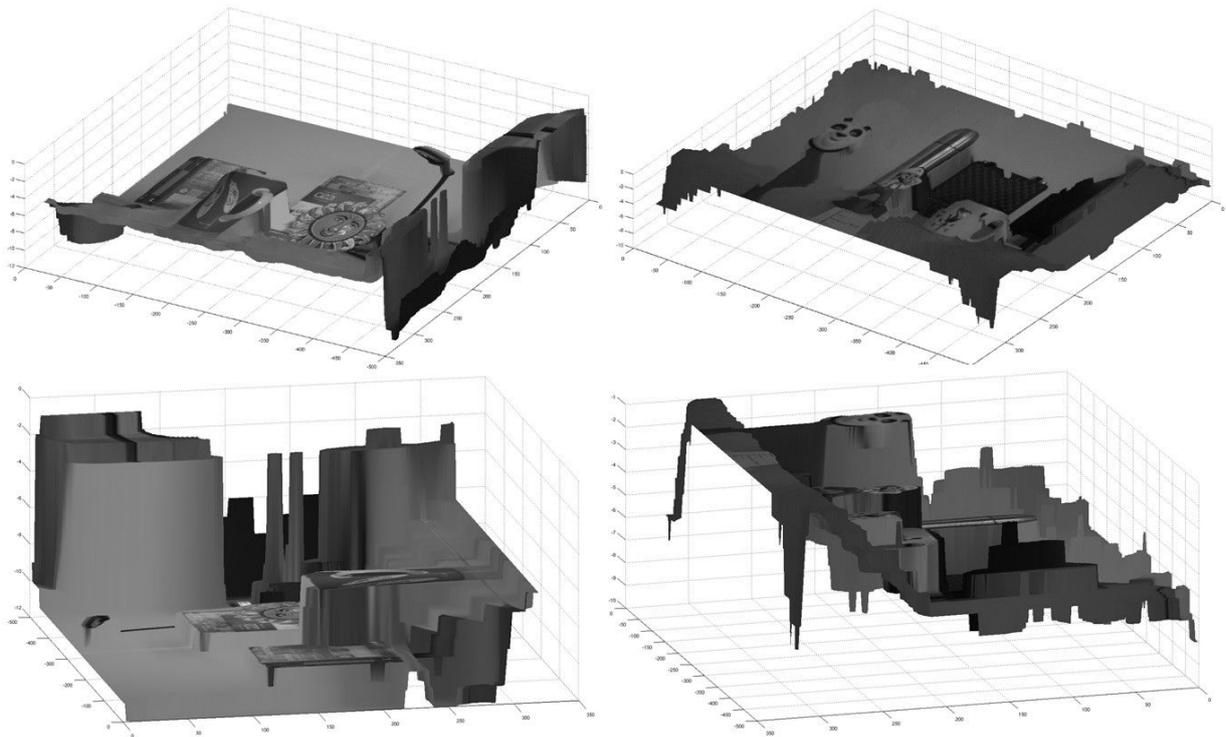


FIGURE 47 : PROJECTION DE SCENES EN 3 DIMENSIONS

Dans le cas de la Figure 47 des cartes de profondeurs filtrées avec un noyau de taille 61×61 ont été utilisées. Elles font disparaître une grande quantité de bruit dans les zones uniformes ce qui permet d'avoir un rendu plus proche de la réalité. Quelques erreurs perdurent mais dans l'ensemble, la représentation 3D des scènes est très correcte.

2.7 Proposition d'architectures

Des paragraphes précédents, il ressort que les solutions intra-images basses résolutions combinent l'intérêt d'une faible complexité et d'une qualité suffisante pour un grand nombre d'applications. Elles sont donc relativement bien adaptées pour être portées vers une architecture matérielle embarquée à faibles ressources.

Il est important de considérer la dépendance de données (Figure 48) pour commencer à construire une architecture. Dans notre cas, la dépendance de donnée la plus pénalisante provient du fait que nous devons obtenir deux cartes de profondeur et de confiance complètes pour chacun des deux axes horizontal et vertical avant de pouvoir produire des cartes unifiées.

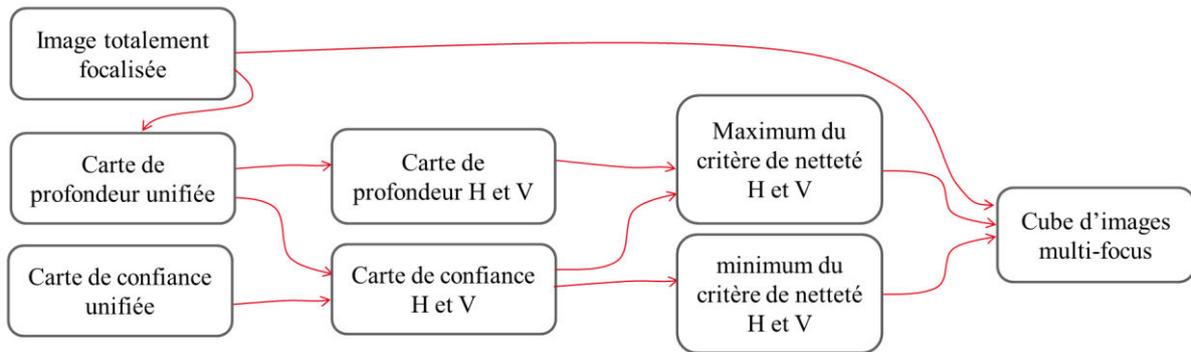


FIGURE 48 : GRAPHE REPRESENTANT LA DEPENDANCE DES DONNEES

Le DfF nécessite de traiter un cube d'images multi-foyers. Il est cependant inconcevable de stocker à l'intérieur d'un ASIC l'ensemble de ces images. On s'impose donc de proposer une architecture qui traite les données à la volée, et éventuellement capable d'adapter son pas de focalisation et sa résolution de profondeur en fonction de l'application. Le cube d'images acquises n'étant plus conservé et le choix de la meilleure direction ne pouvant se faire qu'une fois le traitement de toutes les images terminé, il devient alors nécessaire de sauvegarder une image AiF par direction. C'est une différence notable entre l'implémentation matérielle et le traitement purement logiciel vu jusqu'à présent.

Au regard de la Figure 48 et du paragraphe précédent, on déduit qu'il va falloir mémoriser deux cartes de profondeur (horizontale et verticale), quatre matrices contenant les extrema des critères de netteté par bloc et deux images totalement focalisées. Les matrices d'extrema serviront notamment à calculer les cartes de confiance en fin de procédé. En effet, la confiance étant basée sur la dynamique du critère de netteté, elle ne peut être extraite qu'une fois le cube multi-foyers entièrement traité. Pour des raisons d'économie de mémoire, il est préférable de ne sauvegarder que les extrema à travers le cube du critère de netteté plutôt que tous les critères de netteté pour chaque image. Chacune de ces données sera mise à jour si nécessaire au fur et à mesure du balayage de la scène, avant d'être réunies et extraites en fin d'acquisition d'un cube multi-foyers. Ces différentes mémoires sont bien plus petites avec l'utilisation de fenêtres de traitement non recouvrantes en raison du facteur de réduction $W_h X W_v$.

Nous présentons dans ce paragraphe une architecture qui limite le plus possible l'empreinte mémoire puis faisons une comparaison de ressources avec une architecture basée sur des fenêtres glissantes et l'architecture récemment proposée par Julien N.P. Martel et al [105], à travers l'estimation du nombre d'opérations nécessaires par pixel de profondeur et l'empreinte mémoire.

2.7.1 Architecture de construction de pixel RGB-Z-C par blocs distincts

Considérons une matrice de pixels fournissant des images couleurs au format RAW. Notons I_w et I_h les dimensions horizontale et verticale de ces images. w_w et w_h sont respectivement la largeur et la hauteur d'une fenêtre de traitement. nb_x et nb_y , définis par les équations (2.17), représentent les dimensions horizontale et verticale des cartes de profondeur et de confiance. Ce sont également le nombre de fenêtres de traitement suivant les directions horizontale et verticale. Les facteurs 2 présents au dénominateur résultent de l'opération de dématricage par binning et du sous-échantillonnage induit par le calcul des coefficients de Haar 2D qui divisent chacun par deux la résolution.

$$nb_x = \frac{I_w}{2 \cdot 2 \cdot w_w} \qquad nb_y = \frac{I_h}{2 \cdot 2 \cdot w_h} \qquad (2.17)$$

L'architecture de traitement que nous proposons est présentée sur la Figure 49.

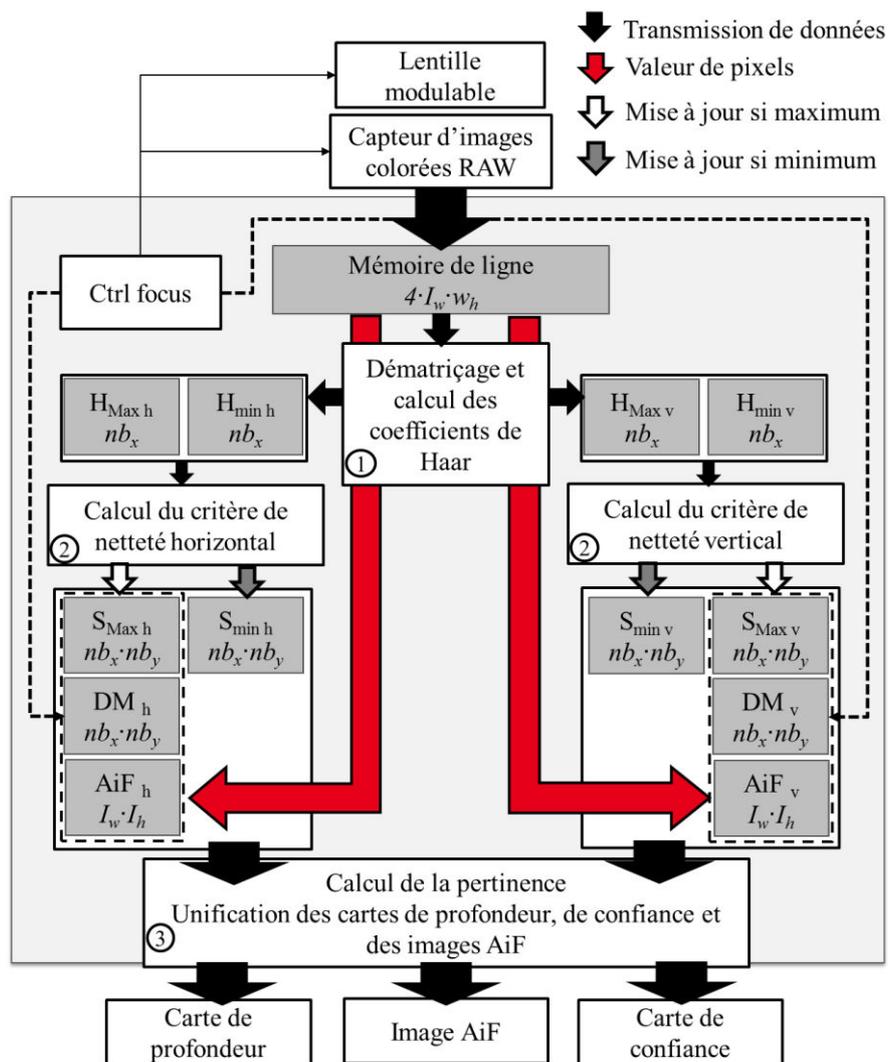


FIGURE 49 : ARCHITECTURE DE GENERATION DE CARTES DE PROFONDEUR ET PERTINENCE ET D'IMAGE TOTALEMMENT FOCALISEE. LES RECTANGLES GRIS REPRESENTENT LES BLOCS MEMOIRE. LEURS TAILLES SONT DONNEES EN ITALIQUE.

Le contrôleur de focus permet d'incrémenter la distance focale de la lentille entre chaque prise d'image et d'incrémenter également le compteur d'indice qui permet de mettre à jour la carte de profondeur. Il comprend plusieurs registres qui permettent notamment de régler la cadence pixel, la résolution de profondeur et les distances focales à appliquer.

Pour la partie traitement à proprement parler, cette architecture est divisée en deux chaînes d'extraction de données, une par direction, elles-mêmes composées de trois étages de traitements. Ces blocs de calcul, annotés par un numéro encerclé sur la Figure 49, fonctionnent de manière pipelinée.

L'étape ① permet de dématricer les données capteurs (moyenne de blocs de 2x2 pixels) et de calculer les coefficients de Haar suivant les deux directions H et V. L'exécution de ce bloc s'effectue aussitôt que possible (toutes les deux lignes d'image RAW pour le calcul du dématricage, toutes les quatre lignes pour les coefficients de Haar). Ses sous-produits (Extrema des coefficients de Haar par fenêtre de traitement non recouvrantes) sont mis à jour toutes les quatre lignes RAW et considérés définitifs une fois leur fenêtre respective entièrement traitée, soit toutes les $4 \times w_h$ lignes RAW.

L'étape ② correspond à l'extraction du critère de netteté suivant les deux directions qui repose sur les extrema calculés au ①. Cette opération ne s'exécute que toute les $4 \times w_h$ lignes. Si les extrema calculés au ① sont plus importants que ceux stockés pour la zone en cours, ils sont mis à jour. Les extrema du critère de netteté, cartes de profondeur et les images totalement focalisées sont également mis à jour, mais seulement lorsque S est supérieur à sa valeur précédente stockée dans S_{max} .

Enfin, l'étape ③ permet d'unifier les valeurs de confiance, de profondeur et les images AiF par sélection des données correspondant au maximum de confiance entre les deux directions. Cette étape se déroule à la même cadence que la précédente uniquement lors de l'acquisition de la dernière image du cube multi-foyers.

Ces étapes sont reprises dans le chronogramme fonctionnel représentant l'enchaînement des tâches Figure 50 et le détail de leur fonctionnement est expliqué ci-dessous.

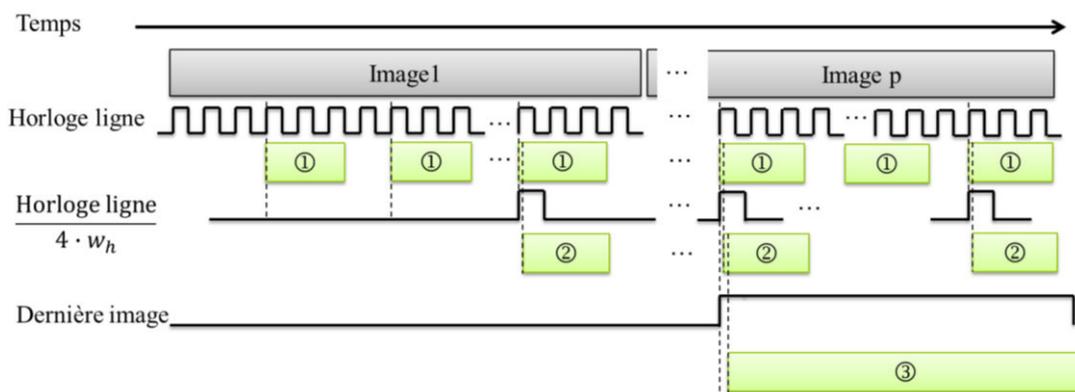


FIGURE 50 : CHRONOGRAMME FONCTIONNEL DE L'ENCHAÎNEMENT DES OPERATIONS DU PIPELINE

2.7.1.1 Etape de préparation du calcul du critère de netteté (étape ①)

Pendant l'étape ①, le système procède à un dématricage du motif de Bayer en sommant les pixels 2×2 afin d'obtenir, à faible coût matériel, un pixel monochrome correspondant à l'énergie des 4 pixels colorés. La résolution pour le traitement est alors réduite d'un facteur 4 ce qui va permettre de gagner en encombrement mémoire. Dans le même temps, on calcule les coefficients de Haar à l'aide de 3 additionneurs et un décalage de 2 bits vers la droite (division par 4) pour chaque lot distinct de 2×2 pixels monochromes (soit 4×4 pixels RAW). 4 lignes d'image source doivent donc être acquises pour résoudre la dépendance de données de ①, même si les sous-produits du calcul sont effectués à chaque nouvelle ligne horizontalement et à chaque paire de ligne verticalement.

Pour effectuer ce prétraitement, la taille minimum nécessaire de la mémoire de ligne est donc de $4 \times l_w$ mots. Cependant, afin de pouvoir *in fine* construire une image AiF, la taille de cette mémoire de ligne doit être d'au moins $4 \times l_w \times w_h$ mots. En effet, les données sources doivent être conservées jusqu'à ce qu'une fenêtre de calcul de $w_w \times w_h$ coefficients de Haar ait été entièrement analysée, puis traitée par l'étape ②.

Dans le cas où des images couleur ne seraient pas nécessaires, un imageur monochrome peut être utilisé. De ce fait, une des opérations de prétraitement serait supprimée et l'étape ① pourrait s'exécuter toutes les 2 lignes au lieu de 4. La mémoire de ligne serait alors de $2 \times l_w \times w_h$.

Chaque coefficient de Haar $H_{h/v}$ subit le traitement présenté Figure 51 qui permet la mise à jour du maximum $H_{Max\ h/v}$ et du minimum $H_{min\ h/v}$ des coefficients de Haar par ligne de blocs.

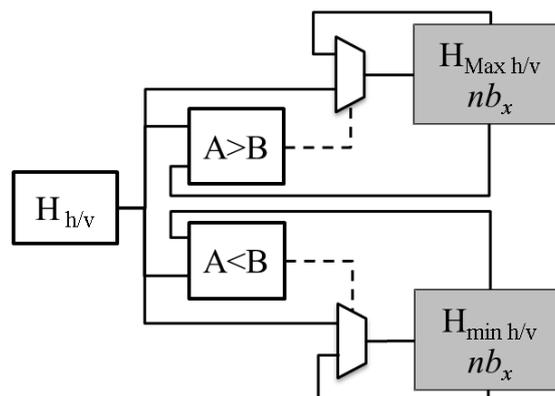


FIGURE 51 : ARCHITECTURE DE PRETRAITEMENT DU CRITERE DE NETTETE : MEMORISATION DES EXTREMA DES COEFFICIENTS DE HAAR. LES RECTANGLES GRIS REPRESENTENT LES BLOCS MEMOIRE. LEUR TAILLE EST DONNEE EN ITALIQUE.

Grâce à des compteurs de synchronisation (non représentés), le coefficient sortant est comparé avec les extrema précédemment calculés appartenant à une même fenêtre de $w_w \times w_h$ pixels. Ce coefficient est ainsi mis à jour au fil de l'acquisition. Les extrema des

coefficients par bloc seront ensuite utilisés dans l'étape ② une fois la fenêtre d'analyse entièrement parcourue. La sauvegarde des maximums et minimums par fenêtres non-recouvrantes pour les deux directions requiert donc 4 blocs mémoires de nb_x mots chacun.

2.7.1.2 Calcul du critère de netteté puis extraction et sauvegarde des cartes de profondeur, des extrema du critère et des images AiF temporaires (étape ②)

L'architecture de traitement de l'étape ② de la Figure 49 est présentée Figure 52.

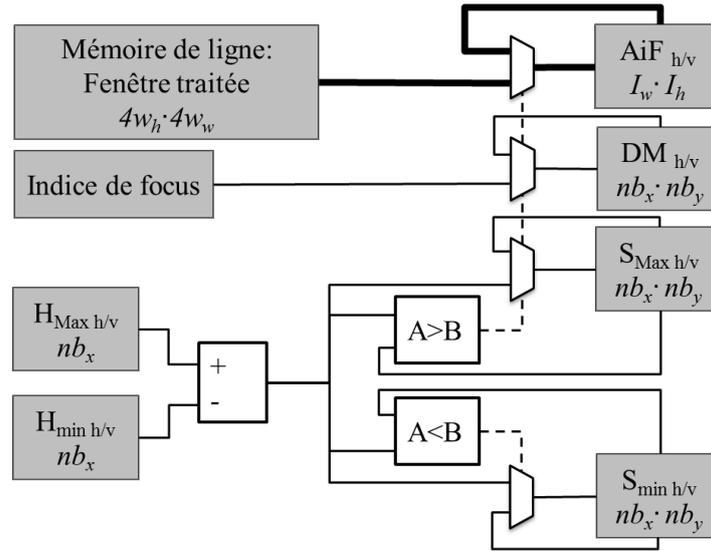


FIGURE 52 : ARCHITECTURE D'EXTRACTION DU CRITERE DE NETTETE ET DE MISE A JOUR DES MEMOIRES AIF ET DM. LES RECTANGLES GRIS REPRESENTENT LES BLOCS MEMOIRE. LEUR TAILLE EST DONNEE EN ITALIQUE.

Elle correspond aux équations (2.11), (2.5) et (2.9) vues précédemment avec quelques modifications apportées pour en optimiser le portage matériel.

$$S_{crit\ h/v}(x, y, z) = \max_{X,Y} (H_{h/v}(X, Y, z)) - \min_{X,Y} (H_{h/v}(X, Y, z)) \quad (2.11)$$

$$DM_{h/v}(x, y) = \underset{z}{\operatorname{argmax}} (S_{crit\ h/v}(x, y, z)) \quad \text{si } S_{Max\ h/v} < S_{crit\ H/V} \quad (2.5)$$

$$AiF(x, y) = Im(x, y, S_{crit\ h/v}(x, y, z)) \quad \text{si } S_{Max\ h/v} < S_{crit\ H/V} \quad (2.9)$$

Le calcul du critère de netteté s'effectue à chaque fois qu'une fenêtre de traitement est entièrement analysée. Le résultat est comparé avec les extrema, S_{Max} et S_{min} . S_{Max} représente la plus grande valeur du critère de netteté de l'image rencontrée jusqu'à présent dans cette zone et pour ce balayage de foyers, tandis que S_{min} est la plus petite valeur rencontrée. Par exemple, pour une fenêtre de coordonnées (n_x, n_y) de l'image d'indice z , $S_{crit}(n_x, n_y, z)$ est comparé avec les extrema, $S_{Max}(n_x, n_y)$ et $S_{min}(n_x, n_y)$.

Etant donné que la carte de confiance est évaluée par la dynamique du critère de netteté, les deux extrema du critère à travers le cube d'images multi-foyers, S_{Max} et S_{min} , doivent être sauvegardés jusqu'à la fin du balayage de foyers pour les $nb_x \times nb_y$ fenêtres distinctes, ceci pour chacune des deux directions de traitement. Quatre blocs mémoire de $nb_x \times nb_y$ mots chacun sont donc nécessaires à cette dépendance de données.

Si un des S_{Max} est mis à jour, l'indice de profondeur de l'image courante est simultanément copié dans le bloc mémoire correspondant de la carte de profondeur $DM_{h/v}$. Deux blocs mémoire de $nb_x \times nb_y$ mots sont également nécessaires pour sauvegarder ces cartes de profondeur horizontale et verticale.

Afin d'économiser de la mémoire, les différentes images du cube multi-foyers ne sont jamais sauvegardées. C'est pourquoi deux blocs mémoires image haute résolution ($l_w \times l_h$) sont nécessaires pour conserver les images AiF. Ils permettent de construire au fil de l'eau une image entièrement focalisée pour chaque direction. Ces images sont mises à jour par bloc de $4w_w \times 4w_h$ pixels dès lors qu'un critère maximum correspondant à leur direction est mis à jour.

Dès qu'un bloc de l'image est traité, ses mémoires H_{Max} et H_{min} sont remises à zéro et à infini respectivement afin d'être utilisées sans interférence par un nouveau bloc de pixels. La mémoire de lignes est également libérée pour les nouvelles lignes entrantes.

2.7.1.3 Unification et production des pixels RGB-ZC (étape ③)

Lors de la dernière image du balayage de foyers, les valeurs produites pour les sous-produits $AiF_{h/v}$, $DM_{h/v}$, $S_{Max\ h/v}$ et $S_{min\ h/v}$ deviennent définitives. Il est alors possible de calculer les cartes de confiances $C_{h/v}$ qui sont la différence entre $S_{Max\ h/v}$ et $S_{min\ h/v}$.

L'étape d'unification repose sur la comparaison des cartes C_h et C_v pour produire trois images unifiées AiF , DM et C sélectionnées à partir de la direction dont la confiance $C_{h/v}$ est la plus grande.

L'architecture proposée est représentée sans ses blocs de synchronisation sur la Figure 53. Elle correspond aux équations (2.6), (2.7), (2.8) et (2.18). X et Y représentent les coordonnées des pixels au sein d'une fenêtre de $4w_w \times 4w_h$ pixels.

$$C_{h/v}(x, y) = \max_z \left(S_{crit\ h/v}(x, y, z) \right) - \min_z \left(S_{crit\ h/v}(x, y, z) \right) \quad (2.6)$$

$$DM(x, y) = \begin{cases} DM_h(x, y) & \text{si } C_v(x, y) \leq C_h(x, y) \\ DM_v(x, y) & \text{si } C_v(x, y) \geq C_h(x, y) \end{cases} \quad (2.7)$$

$$C(x, y) = \begin{cases} C_h(x, y) & \text{si } C_v(x, y) \leq C_h(x, y) \\ C_v(x, y) & \text{si } C_v(x, y) \geq C_h(x, y) \end{cases} \quad (2.8)$$

$$AiF(X, Y) = \begin{cases} AiF_h(X, Y, DM(x, y)) & \text{si } C_v(x, y) \leq C_h(x, y) \\ AiF_v(X, Y, DM(x, y)) & \text{si } C_v(x, y) \geq C_h(x, y) \end{cases} \quad (2.18)$$

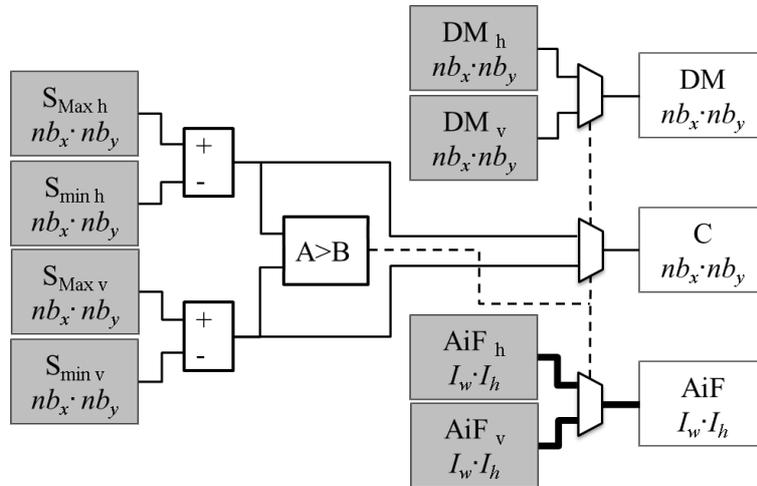


FIGURE 53 : ARCHITECTURE D'UNIFICATION DES VOIES HORIZONTALE ET VERTICALE. LES RECTANGLES GRIS REPRESENTENT LES BLOCS MEMOIRE. LEUR TAILLE EN NOMBRE DE MOTS EST INDIQUEE EN ITALIQUE.

2.7.2 Evaluation de la complexité de l'architecture

Nous avons élaboré cette architecture avec l'a priori que l'utilisation de fenêtres non recouvrantes réduisait l'empreinte mémoire. Nous allons dans ce paragraphe évaluer la mémoire effective qu'elle nécessite et faire quelques comparaisons en fonction de choix de paramètres. Nous estimerons également la complexité de calcul.

2.7.2.1 Empreinte mémoire

Les blocs gris de la Figure 49 permettent de comptabiliser le nombre de mots mémoires nécessaires à la résolution de la dépendance de données pour l'architecture proposée par fenêtres distinctes. L'équation (2.19) reprend ce calcul en fonction des choix architecturaux principaux (taille de l'image et des fenêtres de traitement).

$$M = 4 \cdot I_w \cdot w_h + \frac{I_w}{w_w} + \frac{3 \cdot I_h \cdot I_w}{8 \cdot w_w \cdot w_h} + 2 \cdot I_w \cdot I_h \quad (2.19)$$

Il est important de remarquer que l'encombrement mémoire est totalement indépendant du nombre de niveaux de profondeur souhaité. Cela permet d'adapter la résolution en profondeur en fonction de l'application, si besoin en temps réel.

On peut également noter que si l'application n'a aucunement besoin de l'image AiF, le dernier terme de cette équation $2 \times I_w \times I_h$ correspondant aux mémoires trames nécessaires à la

construction d'images totalement focalisées peut être supprimé, réduisant ainsi drastiquement l'empreinte mémoire aux deux premiers termes de l'équation.

Par analogie, l'équation (2.20) présente les besoins mémoires nécessaires à la même architecture opérant par fenêtres glissantes.

$$M = 4 \cdot I_w \cdot w_h + 2I_w + 6 \left(\frac{I_w}{2} - w_w \right) \left(\frac{I_h}{2} - w_h \right) + 2 \cdot I_w \cdot I_h \quad (2.20)$$

Pour illustrer ces résultats, considérons un imageur capturant des images 1080×1920 pixels et utilisant des fenêtres de traitement 16×16. Le Tableau 6 suivant montre la mémoire nécessaire, exprimée en nombre de trames images équivalentes, pour 6 architectures : celle présentée précédemment pour 4 cas différents : fenêtre glissante ou non, avec ou sans génération d'image AiF et celle proposée dans [105] avec et sans génération d'image AiF.

Tableau 6 : Comparaison des architecture en termes d'empreinte mémoire

	Besoins mémoire (en nombre de trames images)	
	Productions d'images AiF	Pas d'image AiF
Notre architecture Fenêtres glissantes	3,45	1,45
Notre architecture Fenêtres distinctes	2,06	0,06
Méthode [105]	5,00	4,00

L'architecture avec traitement par fenêtre non recouvrantes permet un gain de mémoire allant jusqu'à 96% comparé à la méthode avec fenêtres glissantes. En comparaison avec le système [105], le gain atteint 98.5% lorsque les images AiF ne sont pas nécessaires.

En termes d'optimisation mémoire, l'architecture proposée ici semble donc être un compromis intéressant pour une implémentation matérielle embarquée.

2.7.2.2 Complexité calculatoire

Il est assez difficile d'évaluer la complexité calculatoire sans effectuer le design complet de l'architecture, ce qui aurait dépassé le cadre de cette thèse, d'autant plus lorsqu'il s'agit de faire des comparaisons avec d'autres travaux, car le nombre d'opérations est finalement très dépendant de la méthode d'implémentation. Cette architecture n'ayant pas encore été transcrite en format HDL, nous ne pouvons indiquer l'évaluation précise des ressources matérielles nécessaires. Cependant, nous avons tout de même souhaité fournir un ordre de grandeur du nombre d'opérations nécessaires à l'extraction des informations de profondeur, de confiance et d'image AiF.

Pour cela, nous avons compté le nombre d'additionneurs et de comparateurs nécessaires pour évaluer un pixel de profondeur. Les Tableau 7 et Tableau 8 donnent le nombre d'opérations théoriques pour les cas d'utilisation de fenêtres non recouvrantes et glissantes avec l'architecture proposée.

Tableau 7 : Nombre d'opérations nécessaires par pixel RGB-Z-C pour une architecture avec fenêtres non recouvrantes

	Additionneurs	Comparateurs
Étape ①	$\frac{9 \cdot p}{8}$	$\frac{p}{4}$
Étape ②	$\frac{p}{8 \cdot w_h \cdot w_w}$	$\frac{p}{4 \cdot w_h \cdot w_w}$
Étape ③	$\frac{1}{8 \cdot w_h \cdot w_w}$	$\frac{1}{16 \cdot w_h \cdot w_w}$
Total	$\frac{9p}{8} + \frac{p+1}{8 \cdot w_h \cdot w_w}$	$\frac{p}{4} + \frac{4 \cdot p + 1}{16 \cdot w_h \cdot w_w}$

Tableau 8 : Nombre d'opération nécessaire par pixel RGB-Z-C pour une architecture avec fenêtre glissante

	Additionneurs	Comparateurs
Étape ①	$\frac{9 \cdot p}{8}$	$\frac{p \cdot w_h \cdot w_w}{4}$
Étape ②	$\frac{p}{8}$	$\frac{p}{4}$
Étape ③	$\frac{1}{8}$	$\frac{1}{16}$
Total	$\frac{5 \cdot p}{4} + \frac{1}{8}$	$\frac{p}{4} + \frac{p \cdot w_h \cdot w_w}{4} + \frac{1}{16}$

p représente la résolution de profondeur, soit le nombre total d'images dans le cube multi-foyers.

On peut remarquer qu'une comparaison est souvent réalisée par la prise en compte du signe d'une soustraction. Dans ces conditions, les opérations de comparaison et d'addition ont une complexité équivalente et on peut donc s'intéresser au nombre d'opérations total comme la somme des deux colonnes de chaque tableau (Tableau 9).

Tableau 9 : Nombre total d'opération nécessaires par pixel RGB-Z-C

Fenêtres glissantes	$\frac{p}{4}(w_h \cdot w_w + 6) + \frac{3}{16}$
Fenêtres non recouvrantes	$\frac{11p}{8} + \frac{6p+3}{16 \cdot w_h \cdot w_w}$

Si l'on considère le produit $w_h \times w_w$ grand devant p et $w_h = w_w = w$, l'ordre de grandeur de la complexité est donné par le Tableau 10.

Tableau 10 : Ordre de grandeur du nombre total d'opération nécessaires par pixel RGB-Z-C

Fenêtres glissantes	$\frac{p}{4}(w^2 + 6)$
Fenêtres non recouvrantes	$\frac{11p}{8}$

La complexité calculatoire de l'algorithme par fenêtres glissantes évolue d'une part linéairement avec la résolution de profondeur et d'autre-part en carré de la taille de la fenêtre de traitement. En ce qui concerne les fenêtres non-recouvrantes, on constate qu'à partir d'une taille de fenêtre donnée, la complexité calculatoire est uniquement linéairement dépendante de la résolution de profondeur.

Pour illustrer ces résultats, considérons une application avec 14 niveaux de profondeur p , utilisant des fenêtres de traitement 16×16 . L'architecture avec fenêtre glissante atteint un total de 917,19 opérations par pixel RGB-Z-C alors qu'il n'en faut seulement 19,27 pour la solution basée sur des blocs de traitement non recouvrant, soit presque 50 fois moins.

De plus, aucune opération complexe du type division, multiplication ou encore extraction de FFT ne sont nécessaires ce qui rend son implémentation matérielle plus simple et dans le même temps minimise la surface de silicium requise et donc le coût.

2.8 Conclusion

Nous avons vu dans ce chapitre une analyse de la notion de netteté qui nous a amenés à construire une nouvelle solution pour extraire les éléments nets d'un cube d'images multi-foyers. Un algorithme DfF bidirectionnel a été développé afin d'améliorer la réponse au voisinage des transitions verticales et horizontales. Cette approche repose sur l'évaluation séparée des deux directions, la création d'un critère de confiance basé sur l'autoévaluation de la pertinence des informations captées et une unification des deux canaux basée sur ce critère de confiance.

Trois versions à complexités décroissantes de cet algorithme ont été décrites et comparées quantitativement et qualitativement entre elles ainsi qu'à l'état de l'art. Il a été montré que l'approche intra-image basse résolution est le meilleur compromis pour une intégration matérielle en termes de complexité, de résultat et d'encombrement mémoire. En particulier, l'utilisation de fenêtres non recouvrantes pour l'analyse permet de réduire drastiquement l'empreinte mémoire, le nombre d'opérations par pixel, et ainsi, le coût et la consommation du circuit. Cette solution a été déclinée en une micro-architecture qui a été étudiée et nous a permis d'évaluer finement les besoins en calcul et en mémoire.

L'architecture étudiée permet la création de cartes de profondeur et de confiance au rythme de l'acquisition des images avec des ressources aussi limitées que 0,06 image mémoire et une vingtaine d'opérations par pixel. La production d'images totalement focalisées AiF peut également être réalisée au prix de deux mémoires image supplémentaires.

Nous avons également pris conscience du problème lié au zoom de focalisation qui est très peu abordé dans l'état de l'art et qui peut s'avérer très impactant sur l'extraction de netteté, et donc sur la construction de carte de profondeur, de confiance et d'image totalement focalisées.

Nous allons travailler dans le chapitre suivant sur la recherche d'optiques adaptés au DfF en essayant notamment de limiter le zoom de focalisation. Les lentilles modulables nous ont particulièrement intéressés car elles permettent d'obtenir une variation de la distance de focalisation sans aucun mouvement, ce qui est un atout vis-à-vis de la réduction de l'encombrement du capteur.

Chapitre 3 Développements optiques et réalisation de lentilles modales à cristaux liquides nématiques

3.1	Introduction	77
3.2	Éléments dimensionnants pour l'optique	78
3.3	Lentilles modales expérimentales à cristaux liquides nématiques	82
3.4	Caractérisation de lentille modulable	89
3.5	Synthèse des résultats de mesure	107
3.6	Conclusion	108

3.1 Introduction

Les chapitres précédents montrent que l'utilisation d'algorithme de type Depth From Focus est intéressante pour extraire la profondeur d'une scène dans le cadre d'applications embarquées pour des environnements contraints. Le principe des méthodes de DfF est d'évaluer à partir d'un cube d'images multi-foyers la netteté par zone d'une scène. Les optiques utilisées doivent par conséquent être capables de rendre une qualité d'image suffisante, avec peu de diffusion lumineuse, le moins d'aberrations optiques possible et un zoom de focalisation limité de sorte à faciliter le travail de l'algorithme d'extraction de coefficients de netteté.

Afin de limiter l'encombrement du capteur final, nous visons l'utilisation de lentilles avec des ouvertures relativement petites, de l'ordre de 4-5mm et une distance focale minimum avoisinant les 10cm.

Enfin la troisième caractéristique qui nous intéresse est la vitesse de variation de la position du foyer. La lentille doit être la plus rapide possible pour qu'on puisse, dans l'idéal, fournir une image RGB-Z-C à une cadence vidéo.

Pour répondre à ces trois points, nous avons comparé deux types de lentilles modulables, des lentilles liquides industrielles et des lentilles à cristaux liquides expérimentales. Les lentilles liquides Optotune utilisées pour ces travaux sont calibrées, documentées et parfaitement fonctionnelles ce qui en fait de bons candidats pour le développement rapide d'une

application DfF. Elles sont cependant relativement volumineuses. De plus, nous n'avons pas les moyens techniques à disposition pour faire de l'ingénierie sur la technologie des lentilles liquides ; en revanche, nous disposons de moyens pour réaliser et expérimenter des cellules modulables à cristaux liquides. Nous nous sommes donc intéressés à cette technologie, notamment aux travaux de Nicolas Fraval [106], pour réaliser des lentilles modales à cristaux liquides nématiques et évaluer leur potentiel pour des applications DfF.

3.2 Eléments dimensionnants pour l'optique

Nous avons vu dans le chapitre précédent qu'une image nette se forme sur le capteur seulement si la distance focale du système optique est inférieure à la distance capteur lentille. Dans le cadre de petits systèmes, cela exige une extrême précision du placement de la lentille pour modifier la distance focale. Une méthode courante pour faire évoluer la distance au plan net en imagerie est d'utiliser une lentille fixe focalisée à l'infini devant le capteur et une seconde lentille, qui elle est modulable, placée devant la première lentille. Cela permet de relâcher la contrainte sur la commande focale : on passe d'une précision nécessaire de l'ordre du micron à une précision nécessaire suffisante d'environ un centimètre. En effet, lorsque la lentille fixe est focalisée à l'infini, le plan net sur le capteur n'est autre que le plan focal de la lentille modulable. Une lentille modulable capable d'effectuer une excursion de foyers d'une dizaine de centimètres à l'infini est donc amplement suffisante.

Pour la réalisation d'un système miniature, des considérations géométriques doivent être étudiées ; d'une part, cela permet notamment d'évaluer la zone de fonctionnement du système, d'autre part le facteur matériaux doit également être pris en compte, en particulier avec l'utilisation des cristaux liquides dont l'anisotropie peut varier drastiquement d'une solution à l'autre.

Pour une solution embarquée, il faut également prendre garde aux tensions d'alimentations. Les cristaux liquides pouvant fonctionner avec des tensions de plusieurs centaines de volts, il est primordial de trouver une solution de cristal liquide adaptée à des tensions inférieures à 20V.

3.2.1 Importance de la profondeur de champ

En optique géométrique, il existe théoriquement un seul plan de l'espace qui est net pour une distance focale donnée. Les photo-sites des capteurs d'image ayant une taille minimum définie, il existe en réalité une tranche de l'espace, appelée profondeur de champ, pour laquelle la tache lumineuse est comprise dans le pixel, et donc pour laquelle la netteté est maximale. La profondeur de champ est problématique pour le DfF car elle est une source d'imprécision sur la mesure de la netteté et donc de la distance. La Figure 54 permet

d'appréhender la profondeur de champ à partir de deux sources lumineuses ponctuelles dans les conditions de Gauss. Plus de détails sont disponibles en Annexe 3.

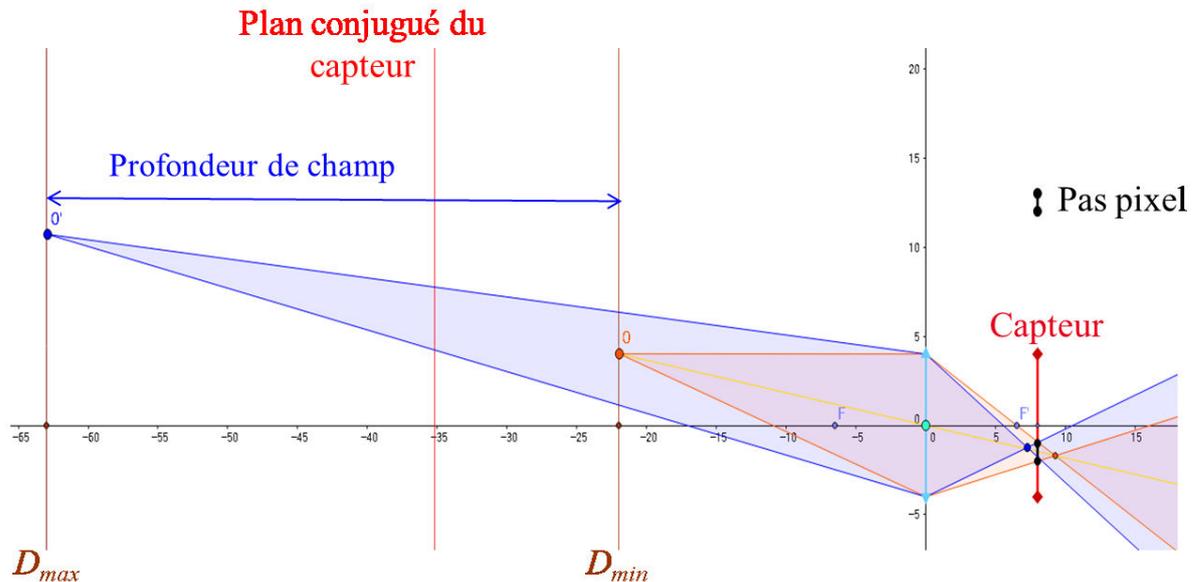


FIGURE 54 ILLUSTRATION DE LA PROFONDEUR DE CHAMP

Le pas pixel (segment noir sur le capteur) est volontairement gros afin de faire ressortir le phénomène observé. Les sections des cônes lumineux des deux sources lumineuses avec le capteur correspondent exactement au pas pixel. Bien qu'une certaine distance les sépare, les images de chacune de ces sources sont nettes dans le cadre d'un raisonnement géométrique.

Dans la réalité, tous les faisceaux lumineux n'arrivent pas forcément sur des pixels indépendants. Même focalisés, certains faisceaux peuvent rencontrer entre 1 et 4 pixels, comme cela est représenté sur la Figure 55. Cela nous amène à considérer un pas pixel de la taille de 2 pixels pour réaliser des simulations plus proches d'une situation réelle.

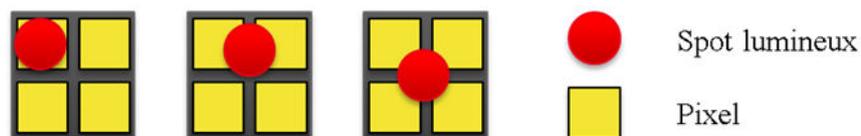


FIGURE 55 : POSITION POSSIBLE POUR LE SPOT LUMINEUX

A cela s'ajoute les aberrations optiques et l'effet de pupille, non considérées ici, qui viennent dégrader davantage la forme et la répartition de la tache lumineuse. Ce modèle va cependant nous permettre d'avoir une première approximation de la profondeur de champ.

Pour conserver une résolution de profondeur intéressante, la profondeur de champ ne doit pas être petite afin de garantir une différence de netteté entre deux images d'un cube multi-foyers.

Définissons K par l'équation (3.1). Ce coefficient représente les paramètres intrinsèques d'un système imageur à longueur focale variable. L est l'ouverture du système, d_c la distance

capteur lentille et P_{pix} le pas pixel. Le coefficient 2 vient de la considération précédente d'une tache lumineuse au minimum égale à 2 fois le pas pixel.

$$K = \frac{L \cdot d_c}{2 \cdot P_{pix}} \quad (3.1)$$

Les considérations de doublet de lentilles et de résolution vues précédemment permettent d'estimer l'ensemble $\{D_{min}, D_{max}\}$ délimitant la profondeur de champ par les équations (3.2).

$$D_{min} = \frac{KD_c}{K+D_c} \quad D_{max} = \begin{cases} \frac{KD_c}{K-D_c} & \text{pour } f < D_c < K \\ +\infty & D_c \geq K \end{cases} \quad (3.2)$$

La profondeur de champ (3.3) est alors déduite par différence de ces deux distances.

$$D_{prof} = \begin{cases} 2K \left(\frac{1}{1 - \left(\frac{D_c}{K}\right)^2} - 1 \right) & \text{pour } D_c \in [0, K[\\ \infty & \text{sinon} \end{cases} \quad (3.3)$$

Lorsque D_c tend vers K , D_{prof} tend vers l'infini. K est appelé distance hyperfocale, c'est la distance à partir de laquelle la profondeur de champ est infinie.

La Figure 56 montre l'évolution des distances D_{min} , et D_{max} en fonction de la distance focale de la lentille modulable pour un pas pixel 3.75microns, une distance capteur lentille de 8mm et une ouverture de 5mm. Cette configuration correspond à celle du module caméra Leopard Imaging utilisé plus loin dans ce manuscrit.

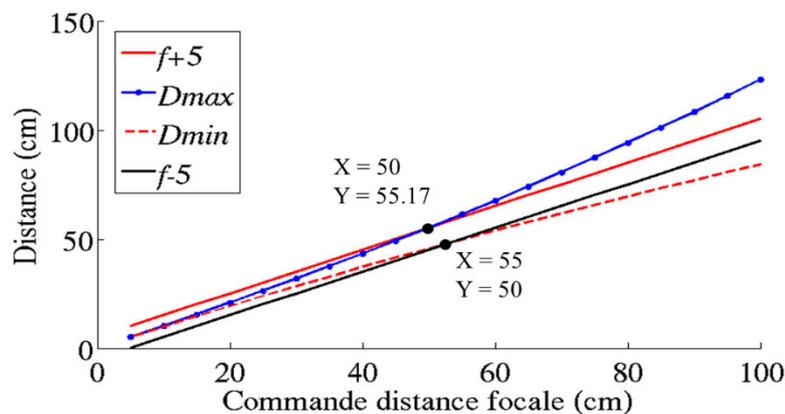


FIGURE 56 : TRACE DES DISTANCES REPRESENTATIVES DE LA PROFONDEUR DE CHAMP ET DE LA LOI DE COMMANDE FUTURE

Les droites rouge et noire, référencée $f+5$ et $f-5$ dans la légende, représentent respectivement la prochaine distance focale à appliquer et la précédente lors d'un balayage de foyers avec un pas de 5cm. Ce graphique montre que pour une construction d'un cube d'images multi-foyers avec un pas de profondeur de 5cm, les objets éloignés de moins de 50 cm de la lentille sont nets sur une unique image du cube et donc aisément estimables.

Une application DfF avec une configuration telle que celle utilisée pour cette simulation sera donc efficace jusqu'à une soixantaine de centimètres.

Une lentille adaptée doit donc être capable de focaliser proche, à quelques centimètres. Cependant, à moins de 10 centimètres, la luminosité des objets est bien souvent limitée, en particulier lorsqu'aucune solution d'illumination n'est déployée. Une distance focale minimum de 10 à 20 centimètres paraît être un bon compromis.

3.2.2 Invariance de la scène

Pour garantir l'invariance de la scène dans un cube multi-foyers complet, la lentille doit être rapide. La vitesse exacte est très dépendante de l'application ; en effet tout dépend de la vitesse relative des objets de la scène par rapport à la caméra. L'idéal étant d'avoir une absence de mouvement. On va donc plutôt se placer dans un cas d'imagerie plus classique et chercher à avoir une cadence de sortie d'image RGB-Z-C de 30 images par seconde avec une dizaine de niveaux de profondeur. Il faut pour cela un minimum de 300 changements de foyers par seconde. Ce qui donne un temps de capture et traitement de 3ms par image. Le temps de réponse de la lentille doit donc être de l'ordre de la milliseconde.

Le « zoom de focalisation » est l'effet de grossissement observable entre deux photographies d'une même scène focalisées à des distances différentes. C'est donc un facteur limitant de l'invariance de la scène. Corriger un zoom, même modéré, requiert des transformations homothétiques complexes pour remettre en correspondance les images. Ceci a un fort impact sur la complexité et l'intégration du capteur. Les lentilles de focalisation introduisent un zoom de puissance variable en fonction de leur technologie. Ce zoom peut provenir du mouvement de la lentille de focalisation mais également de sa déformation dans le cas de lentille modulable liquides par exemple. On peut espérer réduire ce zoom de focalisation avec la technologie à cristaux liquides car d'une part, aucun mouvement n'est nécessaire pour effectuer une variation de distance focale, d'autre part, l'épaisseur de la couche de cristal liquide est fixe et très petite ce qui nous rapproche des conditions de Gauss pour lesquelles aucun zoom n'apparaît sans mouvement de lentille.

3.2.3 Qualité optique

L'évaluation de la qualité de mise au point repose sur une mesure du contraste local dans l'image. La transmittance de la lentille a un impact sur la quantité de lumière reçue par le capteur. Ceci influence le facteur d'amplification (ISO) et donc directement le bruit temporel des images acquises. En fonction de la sensibilité de la mesure de distance, ce bruit peut conduire à des erreurs sur la position des maxima de netteté et à une imprécision sur la mesure de distance.

Il est donc important d'avoir une limitation des aberrations optiques, de la diffusion ou encore de l'absorption lumineuse qui peuvent être provoquées par le matériau constituant les lentilles. Tout phénomène qui tend à réduire la fonction de transfert de modulation (FTM) a un impact sur la sensibilité de la mesure. Les cristaux liquides étant des matériaux diffusants, il faut limiter le plus possible l'épaisseur de leur couche pour garantir une FTM correcte.

3.3 Lentilles modales expérimentales à cristaux liquides nématiques

Les lentilles modales à cristaux liquides nématiques réalisées au cours de ces travaux sont du type de celles précédemment développées par Nicolas Fraval dans sa thèse [106]. Nous avons poursuivi ses travaux pour réaliser de nouvelles lentilles et évaluer leur performance dans le cadre d'une application DfF embarquée.

La Figure 57 est une représentation schématique des lentilles réalisées. Elles sont composées de deux substrats de verre formant deux électrodes résistives entre lesquelles résident quelques microns de cristal liquide nématique.

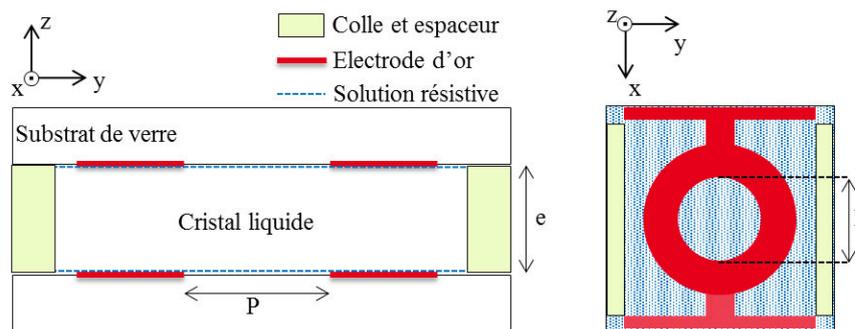


FIGURE 57 : SCHEMA DES LENTILLES REALISEES : VUE EN COUPE (GAUCHE) ET VU DE DESSUS (DROITE)

Grâce à la solution résistive déposée sur les substrats de verre, un gradient de champ électrique apparaît entre les électrodes. Celui-ci se transforme en gradient d'indice de réfraction au sein de la couche de cristal liquide nématique, ce qui permet, au même titre que l'évolution de l'épaisseur de verre dans une lentille, de déphaser la lumière progressivement et ainsi de modifier le front d'onde et de le rendre sphérique.

La modulation du gradient d'indice, et donc de la puissance optique de la lentille, dépend d'une part, du carré de l'amplitude du champ électrique, qui influence le couple électrique (1.5) et de ce fait l'inclinaison des molécules nématiques et donc l'indice de réfraction (1.6), et d'autre part de la fréquence du signal de polarisation.

$$\Gamma_E = \frac{\Delta\epsilon \cdot E^2}{2} \sin(2(\psi - \theta)) \quad (1.5)$$

$$n(\theta) = \frac{n_e \cdot n_o}{\sqrt{n_e^2 \cdot \cos^2(\theta) + n_o^2 \cdot \sin^2(\theta)}} \quad (1.6)$$

L'aspect fréquentiel provient de l'utilisation d'une couche résistive qui permet de ralentir la propagation du potentiel électrique dans chacune des électrodes.

Notons χ^2 le rapport entre la résistance surfacique R_s de l'électrode et l'impédance de la couche de cristal liquide défini par l'équation (3.4), g et c sont respectivement la conductance surfacique et la capacité surfacique de la couche de cristal liquide.

$$\chi^2 \approx \frac{R_s}{(g - i\omega c)} \quad (3.4)$$

Nicolas a étudié en détail dans sa thèse [106] la répartition du potentiel électrique dans les électrodes : pour avoir une distribution de tension linéaire, il faut que $|\chi| \ll 1$. Dans ce cas, la répartition du potentiel électrique dans la cellule peut être approximée par l'équation (3.5) fonction de χ et donc de la fréquence.

$$U(r) \approx U_0 \cdot \left(1 - \frac{\chi^2}{4} \left(\frac{p^2}{4} - r^2 \right) \right) \quad (3.5)$$

avec $r = \sqrt{x^2 + y^2}$, x et y référencés par rapport au centre de la lentille.

C'est ainsi qu'en faisant varier la fréquence du signal de polarisation, on modifie la répartition du potentiel dans l'électrode et ainsi le gradient d'indice. Plus la résistance est faible, plus le potentiel aura tendance à s'installer rapidement dans la pupille et donc plus la fréquence devra être rapide pour conserver ce gradient d'indice de réfraction. A l'inverse, si la résistance est grande, le potentiel risque de rester confiné dans l'électrode circulaire, et aucun champ n'apparaîtra dans la pupille ($|\chi| \gg 1$). Il est également souhaitable de conserver R_s suffisamment important afin de limiter la puissance dissipée.

3.3.1 Potentiel de variation

Nous avons choisi les lentilles à cristaux liquides car nous avons l'avantage de pouvoir faire de l'ingénierie avec cette technologie. En effet, divers paramètres sont exploitables et peuvent changer drastiquement les caractéristiques d'une lentille.

Voici quelques paramètres influençant la lentille et leur impact sur la réalisation :

- La permittivité diélectrique ϵ du cristal liquide permet d'estimer la capacité C d'une cellule, une grande permittivité augmente la capacité de la cellule ce qui influence sa fréquence de coupure et ainsi sa fréquence de commande.
- L'anisotropie de la permittivité diélectrique $\Delta\epsilon$ du cristal liquide détient un rôle important sur le couple électrique permettant la rotation des molécules de cristal

liquide (1.5). Plus $\Delta\epsilon$ est grand, plus la tension de commande de la cellule peut être basse, ce qui est un avantage dans le cadre d'une solution embarquée.

- L'anisotropie optique Δn influence la puissance optique de la lentille à cristaux liquides. Plus elle est forte, plus le déphasage pour une couche d'épaisseur donnée de cristal est important. Et donc plus la lentille peut être mince pour une distance focale donnée. L'anisotropie des cristaux liquides varie entre 0.1 et 0.35. En revanche plus elle est grande, plus la diffusion augmente dans un modèle standard de Mie.
- La viscosité du cristal liquide joue sur sa vitesse de rotation et donc sur le temps de réponse de la lentille.
- La tension de polarisation admet un minimum pour vaincre le couple naturel qui retient les molécules de cristal dans leur position de repos. Plus la polarisation dépasse ce minimum, plus le couple électrique est fort et plus la vitesse de rotation des molécules est grande. Il y existe une limite qui est la tension de claquage mais celle-ci est relativement élevée pour les cristaux liquides, souvent supérieure à 10V/micron.
- L'épaisseur de la lentille dépend du Δn et également de la distance focale minimum désirée. Plus celle-ci est petite, moins la lentille est mince pour un cristal donné. En revanche, les cristaux liquides sont relativement diffusants. Plus la couche est épaisse, plus la lumière diffuse ce qui pose problème pour des applications d'imagerie.
- Le diamètre de la pupille influence la répartition du potentiel électrique dans l'électrode. Il va donc intervenir dans le calcul de la résistance nécessaire pour la couche résistive recouvrant les substrats de verre.
- La résistance de la solution résistive est primordiale. Elle doit être optimisée en fonction du diamètre de la pupille et de la permittivité diélectrique du cristal liquide. Le potentiel ne doit pas se propager trop rapidement, autrement l'électrode devient un équipotentiel. Le gradient d'indice de réfraction disparaît alors et la fonction de lentille avec.

3.3.2 Paramètres qui délimitent les choix à faire

Au regard du potentiel de variation des lentilles à cristaux liquides, nous devons trouver des paramètres pour orienter nos choix expérimentaux. Quelques caractéristiques ont déjà été mises en avant dans ce manuscrit vis-à-vis d'une application DfF embarquée : ouverture de l'ordre de 5mm pour conserver le caractère miniature, une diffusion minimale pour garantir une évaluation correcte de la netteté, une distance focale minimale de l'ordre de 10 à 20 centimètres et une polarisation par des faibles tensions, typiquement $\pm 20V$ grand maximum. Nous allons donc voir dans ce paragraphe en quoi ces 4 caractéristiques vont nous permettre de limiter le champ expérimental. La vitesse est également importante à prendre en compte pour du DfF. Cependant d'une part, les cristaux liquides nématiques ne sont pas très rapides, et d'autre part une vitesse de commutation « importante » nécessite une couche très fine ce qui est difficilement compatible avec une distance focale courte. Nous plaçons donc la caractéristique « vitesse » en bas de l'échelle des priorités.

Partant du constat que les cristaux liquides nématiques sont des matériaux diffusants, nous voulons limiter le plus possible l'épaisseur de la couche de cristal, donc pour atteindre une faible distance focale avec une fine couche de cristaux, il est nécessaire de présélectionner des cristaux liquides avec un Δn important (>0.20) en gardant à l'esprit qu'un Δn trop grand augmentera la diffusion et détériora la qualité d'image. On va alors pouvoir calculer en fonction du cristal liquide et de la distance focale minimale souhaitée, l'épaisseur des lentilles.

Parmi les cristaux présélectionnés, on pourra choisir ceux avec une forte permittivité diélectrique $\Delta\epsilon$ afin d'avoir des tensions de polarisation relativement basses.

En imposant un diamètre de pupille, la résistance de la couche résistive nécessaire au bon établissement du champ électrique peut être calculée. Celle-ci influence directement la commande et les performances attendues de la lentille. La formulation mathématique de la résistance en fonction de la pupille et du cristal liquide nématique a été brevetée par la société Evosens et Telecom Bretagne. Nicolas Fraval nous a fourni les mélanges adéquats.

Nous avons à notre disposition deux lentilles (L1 et L2 du Tableau 11) et en avons réalisé quelques autres dans le but d'appréhender la technologie d'une part, mais aussi pour pouvoir construire un comparatif qui nous permettra de choisir les caractéristiques géométriques et matérielles les plus adaptées à notre application.

Tableau 11 : Caractéristiques souhaitées des lentilles modales à cristaux liquides nématiques

	Diamètre de Pupille (mm)	Distance focale minimum (cm)	Δn	$\Delta\epsilon$	Viscosité (mPa.s)
L1	2	10	0.26	12	203
L2	2.5	5	0.26	12	203
L3	4	15	0.25	47	464
L4	4	10	0.25	47	464
L5	4	15	0.35	4	47
L6	4	10	0.35	4	47
L7	5	15	0.25	47	464

3.3.3 Calcul de l'épaisseur de la couche de cristal liquide

Pour concevoir une lentille à cristal liquide, il faut connaître les caractéristiques des cristaux, la distance focale minimum f_{min} utile pour l'application visée, et le diamètre de la pupille souhaitée P . Ces paramètres vont permettre de calculer l'épaisseur de la couche de cristal liquide nématique.

Ces lentilles sont modélisables par une lame à retard de phase à distribution sphérique (Figure 58).

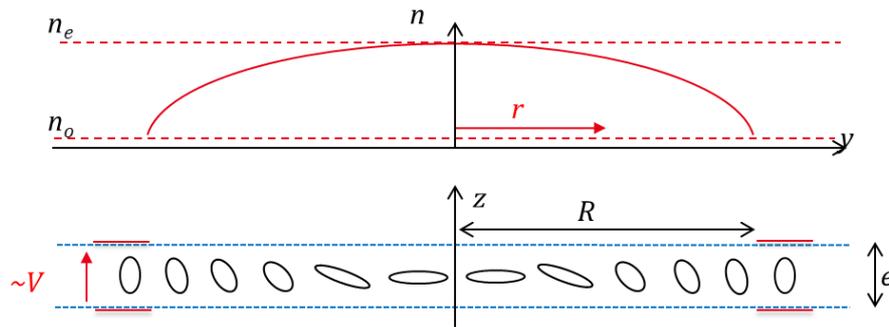


FIGURE 58 : REPARTITION DE L'INDICE DE REFRACTION SELON UN DISTRIBUTION SPHERIQUE DANS UNE LENTILLE A CRISTAUX LIQUIDES.

Cette modélisation est légitime car le champ électrique appliqué aux bornes des électrodes provoque une rotation des cristaux donc une variation locale d'indice dans la couche de cristaux liquides. Il n'y a aucune déformation de dioptre contrairement aux lentilles liquides.

Par analogie entre cette modélisation d'une lentille à cristaux liquides et une modélisation similaire d'une lentille mince vue comme une lame à retard de phase constituée d'air et de verre (Annexe 4), on obtient l'équation (3.6) qui donne l'épaisseur e de la couche de cristal liquide en fonction de l'anisotropie optique Δn des cristaux, du diamètre de la pupille P et de la distance focale courte f_{min} .

$$e = \frac{P^2}{8 \cdot \Delta n \cdot f_{min}} \quad (3.6)$$

Le Tableau 12 reprend le Tableau 11 avec cette fois les distances focales minimales théoriques obtenues avec les tailles des espaceurs disponibles dans le laboratoire.

Tableau 12 : Caractéristiques des lentilles modales à cristaux liquides nématiques évaluées.

	Diamètre de Pupille (mm)	Distance focale minimum (cm)	Epaisseur (μm)	Δn	$\Delta \epsilon$	Viscosité (mPa.s)
L1	2	10.6	18	0.26	12	203
L2	2.5	2.4	125	0.26	12	203
L3	4	16	50	0.25	47	464
L4	4	11,4	70	0.25	47	464
L5	4	11.4	50	0.35	4	47
L6	4	8.1	70	0.35	4	47
L7	5	17.8	70	0.25	47	464

3.3.4 Etapes de réalisation

Ces lentilles sont réalisées en salle blanche. La propreté des substrats est primordiale à tout moment de la réalisation. Ils sont nettoyés et conservés dans des boîtes fermées entre chacune des cinq étapes de la fabrication.

Une première phase consiste à préparer des substrats de verre recouverts d'un conducteur (ITO ou Chrome) pour la gravure des électrodes. Ceux-ci sont recouverts de 50 nanomètres d'or afin d'augmenter la conductivité. Les électrodes sont ensuite gravées à l'aide d'un masque tel que celui présenté sur la Figure 59.



FIGURE 59 : EXEMPLE DE MASQUE POUR UNE LENTILLE DE 5MM DE PUPILLE

Une fois les électrodes gravées, une fine couche de solution résistive composée de (3,4-ethylenedioxythiophene)-poly(styrene-sulfonate) (PEDOT-PSS) et d'alcool polyvinyl est déposée sur les électrodes à l'aide d'une centrifugeuse.

Lors de notre succincte description des cristaux liquides (section 1.2.3.1), nous avons parlé d'ancrage des molécules. Pour s'ancrer, celles-ci nécessitent une surface d'accroche particulière, par exemple des sillons monodirectionnels. Nous avons réalisé de tels sillons par brossage de la couche résistive pour forcer l'orientation des molécules de cristal liquide nématique dans la cavité de la lentille. Les substrats sont ensuite assemblés 2 à 2 en opposition de brossage via des joints de colle. L'épaisseur des cellules est contrôlée par addition de microbilles calibrées dans la colle.

Le cristal liquide (MLC-2140 [$\Delta n=0.25$], MDA-98-1602 [$\Delta n=0.26$] and QYDLC-DH002 [$\Delta n=0.35$]) est ensuite introduit dans la cavité par effet de capillarité. Enfin, la cellule est scellée par un nouveau joint de colle afin de limiter les apparitions de bulles et la fuite du cristal liquide. Plus de détails sont disponibles en Annexe 5 et dans la thèse [106].

Au cours de ces travaux, nous avons utilisé des épaisseurs de couche de cristaux liquides supérieures à 100 microns. Il est assez difficile et très onéreux de se procurer des espaceurs. Pour pallier à la fois au coût et au temps de livraison, nous avons mis au point un procédé de fabrication à base de film résine photosensible pour créer nos propres espaceurs. Il s'agit des mêmes films résines que ceux destinés au vernissage des cartes électroniques.

Sur l'une des électrodes préalablement gravées, des couches de résine sont déposées pour obtenir l'épaisseur souhaitée, puis gravées à l'aide d'un masque à bande. Ainsi, seulement deux bandes de résine sont conservées sur les bords d'une des deux électrodes. La Figure 60 montre une électrode avec des bandes de résine gravées (deux couches de 40 microns) et le profil de ces bandes mesuré avec un profileur.

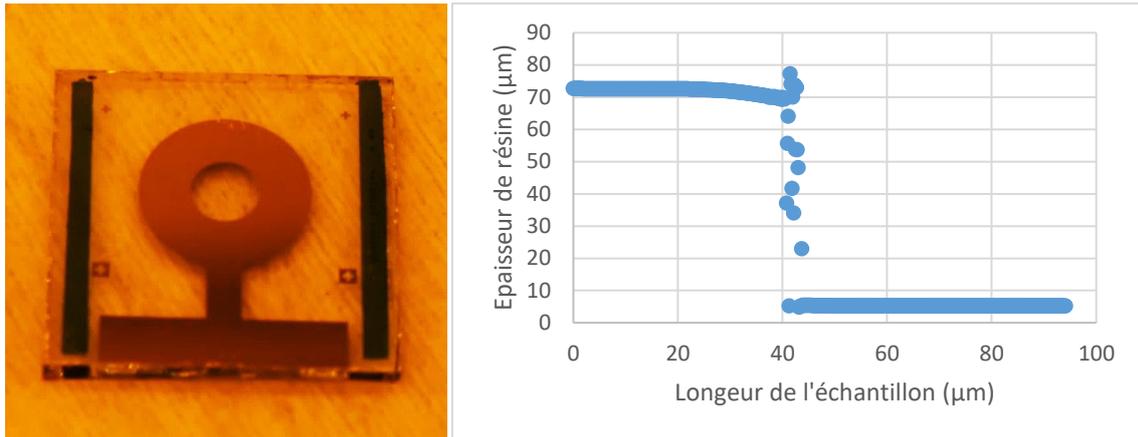


FIGURE 60 : ELECTRODE AVEC BANDE DE RESINE (GAUCHE) ET PROFILE D'ÉPAISSEUR DE LA COUCHE DE RESINE PHOTOSENSIBLE DÉPOSÉE (DROITE).

Le profil est plutôt régulier avec une épaisseur finale de résine de 67,5μm. Malgré une épaisseur plus faible qu'annoncée, le profil est suffisamment plat pour avoir un bon maintien de l'électrode supérieure et une épaisseur de cristal liquide relativement constante.

Une fois la résine gravée, la solution résistive peut être déposée. Les bandes de résine perturbent légèrement l'établissement de la couche de PEDOT-PSS à la centrifugeuse, elles empêchent notamment la matière en surplus d'être éjectée par force centrifuge. Cependant, lorsque les bandes sont suffisamment éloignées du centre de la lentille, le dépôt résistif au sein de la pupille est homogène, ce qui garantit un bon fonctionnement des lentilles ainsi réalisées.

La Figure 61 montre une gamme de lentilles réalisées avec les deux types d'espaceurs.

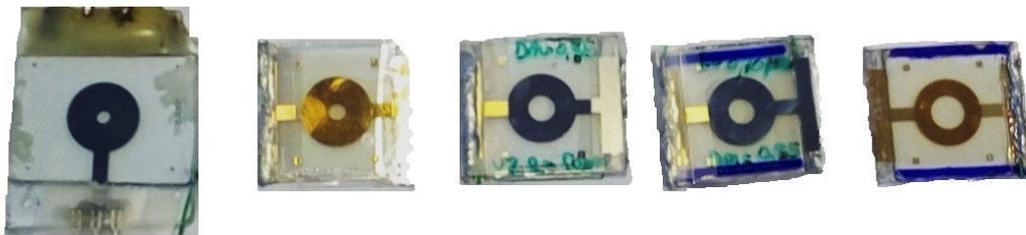


FIGURE 61 : EXEMPLE DE LENTILLE A CRISTAUX LIQUIDES, DE GAUCHE A DROITE, PUPILLE DE 2MM, 2.5MM, 4MM ET 5 MM

Pour contrôler le fonctionnement et la qualité de la lentille après la fermeture d'une cellule, celle-ci est observée sous loupe binoculaire entre polariseurs croisés. Le déphasage induit par la couche de cristaux liquides crée l'effet d'une lame à retard qui entre polariseurs croisés module l'intensité lumineuse. Nos électrodes étant circulaires les lignes équidéphasages sont circulaires et centrées sur l'axe optique. La Figure 62 montre les franges d'interférences aperçues pour une lentille de 4mm de pupille polarisée avec une tension de 5V d'amplitude.

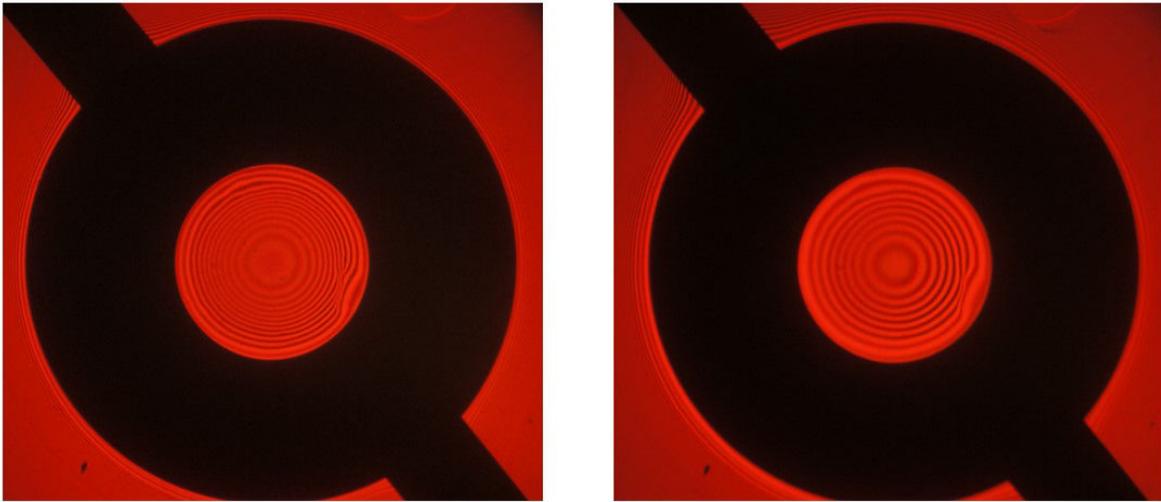


FIGURE 62 : LENTILLE L3, 4MM DE PUPILLE AVEC UN CRISTAL LIQUIDE $\Delta n=0.25$, UN EPAISSEUR DE 50 MICRONS, COMMANDEE PAR UNE TENSION DE $\pm 5V$ A 450kHz ET 330kHz

La qualité de cette lentille est relativement bonne malgré un petit défaut visible sur ces photos au niveau du bord droit de la pupille. Son origine est difficile à analyser, il se peut qu'il y ait un défaut local dans la couche résistive provoqué par la présence d'un grain de poussière.

De façon générale, l'interprétation des dysfonctionnements est complexe. Et ce, principalement parce que nous ne pouvons faire pratiquement aucun test avant l'assemblage final. Certaines lentilles présentaient un axe optique non centré sur la pupille. Cela peut provenir d'une irrégularité dans la solution résistive, d'un mauvais alignement des deux substrats de verre ou encore d'une variation d'épaisseur de la couche de cristaux liquides à cause d'un manque de fluidité de la colle ou d'une concentration en espaceurs trop grande. D'autres avaient visiblement subi un frottement trop fort provoquant un arrachement de la couche résistive. Nous avons également eu quelques difficultés à cause de la solution de PEDOT-PSS dont la résistance ne semblait pas toujours adaptée. Il semble qu'en fonction des fournisseurs, les propriétés des solutions de PEDOT-PSS ne soient pas rigoureusement les mêmes. Certains défauts observés au cours de ces travaux sont montrés en fin d'annexe 5.

3.4 Caractérisation de lentille modulable

L'objectif de cette caractérisation est d'évaluer les lentilles à cristaux liquides en termes de vitesse de fonctionnement, de diffusion lumineuse, de zoom de focalisation puis dans une application DfF. Nous aurions également pu nous intéresser à l'absorption lumineuse mais Shin-Tson Wu et Khoon-Cheng Lim [107] ont montré que la diffusion lumineuse était deux fois plus importante que l'absorption avec les cristaux liquides nématiques, c'est pourquoi nous ne portons que peu d'attention à cette dernière qui reste négligeable pour les épaisseurs considérées.

Pour effectuer cette caractérisation, il a fallu dans un premier temps mesurer précisément la caractéristique commande versus distance focale. Les évaluations de vitesse, de zoom de focalisation et de qualité optique ont ensuite pu être appréhendées.

Nous avons comparé au total 7 lentilles à cristaux liquides et 2 lentilles liquides de la société Optotune. Les lentilles liquides EL-10-30 et EL-16-40 ont des pupilles respectives de 10 et 16 mm, elles sont référencées dans la suite de ce manuscrit par Opto10 et Opto16. Leur épaisseur et vergence respectives sont de 24.6mm et 11.9mm, et [-1.5D à +5D] et [-2D à +3D].

3.4.1 Présentation du banc de test

Nous avons fait le choix d'utiliser un module caméra RAW Leopard Imaging M021C qui constitue une caméra peu encombrante fournissant des images couleurs en format RAW avec une résolution de 1280×960 pixels à une cadence de 60 images par seconde. Une lentille à focale fixe avec une ouverture de 5mm permet de focaliser la lumière sur le capteur, les lentilles modulables ont été accolées au plus près de celle-ci afin de limiter la réduction d'ouverture. La Figure 63 montre un schéma du dispositif d'évaluation qui est utilisé pour toute la caractérisation.

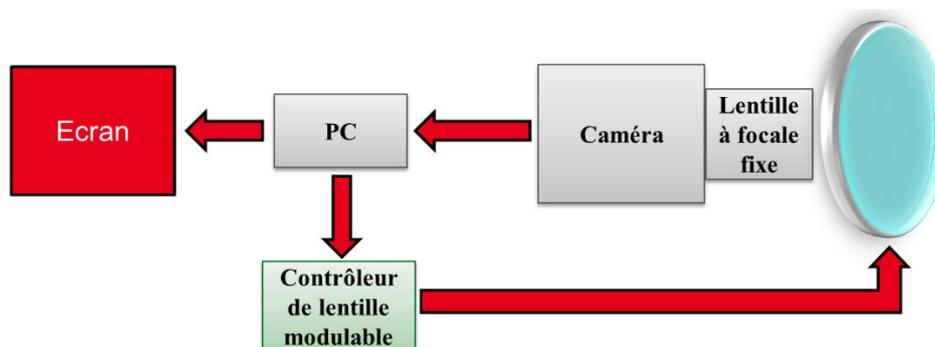


FIGURE 63 : PRESENTATION DU BANC DE TESTS

Les lentilles modulables nécessitent un contrôleur afin de faire le lien avec l'ordinateur qui exécute les programmes de caractérisation. Optotune en propose un très simple d'utilisation pour leurs lentilles. Il crée une liaison série via un port USB qui permet de commander directement la vergence des lentilles, aucune autre source d'alimentation que l'USB n'est nécessaire ce qui est très avantageux.

Pour les lentilles à cristaux liquides, nous avons créé un contrôleur communiquant également par liaison série avec l'ordinateur. Il comporte un module FTDI pour créer la liaison USB/série, un STM32 pour générer un signal PWM de fréquence ajustable, et un amplificateur à pont en H avec des transistors Gan permettant de fournir un signal carré alternatif symétrique d'amplitude ajustable entre 0 et 100V pour des fréquences jusqu'à 2MHz. Les détails concernant la mise en place de ce driver sont donnés en annexe 6.

Les cristaux liquides étant sensibles à la polarisation de la lumière, deux possibilités représentées sur la Figure 64 permettent de fournir une image correcte.

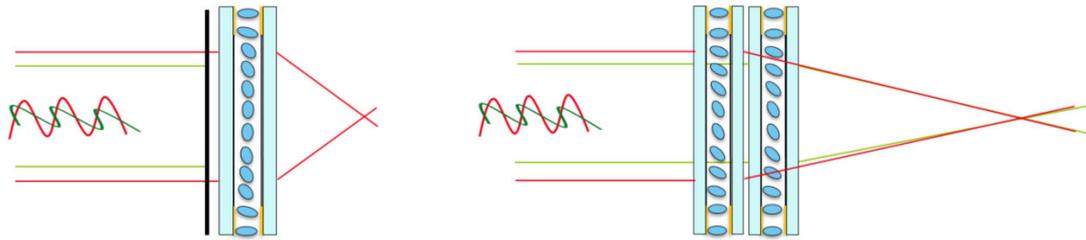


FIGURE 64 : SCHEMA DES SOLUTIONS AVEC POLARISEUR (GAUCHE) OU DOUBLET DE LENTILLE (DROITE) EN LUMIERE NON COHERENTE

Un polariseur peut être placé devant une unique lentille à cristaux liquides (schéma de gauche), permettant ainsi d'avoir un faisceau lumineux orienté seulement suivant la direction extraordinaire des cristaux. L'inconvénient étant une perte de luminosité importante, de l'ordre de 60%, à cause du polariseur.

Une deuxième possibilité consiste à utiliser deux lentilles à cristaux liquides accolées l'une à l'autre (schéma de droite), avec des directions extraordinaires orientées perpendiculairement. Il est important pour cette solution que les lentilles soient les plus fines possible afin de limiter la réduction du champ de vision. En effet, plus la pupille extérieure est éloignée du capteur, plus l'ouverture du système optique diminue. Cela impacte directement la qualité de l'image par une réduction de la luminosité et une augmentation de la profondeur de champ. En outre, l'écartement entre ces deux lentilles n'est pas négligeable. Les focales équivalentes des deux doublets lentille à cristal liquide et lentille à focale fixe sont totalement différentes, ce qui nécessite la génération d'une double commande et complique la caractérisation.

La solution avec polariseur a donc été privilégiée, la Figure 65 présente l'ensemble optique avec leur support de maintien imprimé pour les deux technologies, Optotune et cristal liquide.



FIGURE 65 : MODULE IMAGEUR DU BANC DE TEST OPTOTUNE+CAMERA (GAUCHE), POLARISEUR+LENTILLE A CRISTAUX LIQUIDES+CAMERA (DROITE)

Pour l'analyse de netteté, nous avons utilisé des cibles fortement contrastées, une mire USAF1951 et des scènes avec de multiples objets. Une lampe à LED sur batterie a permis d'avoir une source de lumière fiable sans clignotement.

3.4.2 Caractérisation de la loi de commande

Quelle que soit l'application, il est nécessaire de connaître la loi de commande des lentilles. Dans notre cas, il s'agit d'obtenir les fréquences correspondantes aux distances focales lors d'une polarisation des lentilles à cristaux liquides avec une amplitude constante. Le DfF n'ayant pas de boucle de retour, cette mesure doit être précise et fiable afin de garantir une la répétabilité de la mesure de distances.

Pour cela, nous avons utilisé une méthode itérative de recherche d'optimum de netteté. Pour des distances allant de 20 à 60 cm, notre programme réalise automatiquement une série de mesures pour une excursion de fréquence. Ces mesures permettent ensuite d'identifier la commande la plus adaptée à chacune des distances. La Figure 66 montre un organigramme du programme d'identification de commande pour une distance.

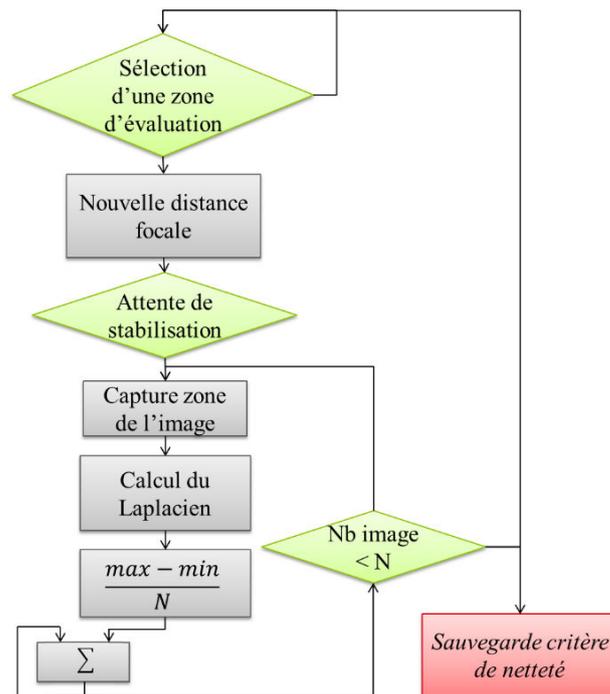


FIGURE 66 : ORGANIGRAMME DE L'ALGORITHME DE CARACTERISATION DE COMMANDE

Une cible à fort contraste est placée à une distance de caractérisation. L'utilisateur sélectionne une zone dans l'image qui est alors analysée avec un critère de netteté pour chaque fréquence du signal de commande. Il s'agit ici d'extraire le support de l'histogramme du Laplacien. Les critères de N images (30 dans notre cas) sont moyennés afin de garantir une certaine robustesse face au bruit. Une fois ces données recueillies pour un balayage de fréquence

conséquent, la recherche du maximum du critère permet d'obtenir la fréquence de commande optimale pour chaque distance.

Cet algorithme a dans un premier temps été validé avec les lentilles Optotune qui sont parfaitement caractérisées.

Les lentilles liquides ont été accolées le mieux possible à la lentille fixe de la caméra. On s'attend alors à obtenir des résultats qui concordent avec l'équation (3.7) estimant théoriquement la distance de l'objet d_o en fonction des distances image d_i et focale de la lentille fixe f_{fixe} , de la vergence V de la lentille modulable et de l'écartement entre les deux lentilles e .

$$d_o = \frac{f_{fixe}}{V \cdot (f_{fixe} - e) + 1 - \frac{f_{fixe}}{d_c}} \quad (3.7)$$

La Figure 67 montre les résultats expérimentaux pour les lentilles liquides et la courbe théorique associée ($f_{fixe} = 8mm$). Les mesures correspondent relativement bien à la théorie pour $e=0.8mm$ et $d_c = 8.0045mm$. Ce n'est pas aberrant étant donné que ces réglages sont effectués manuellement.

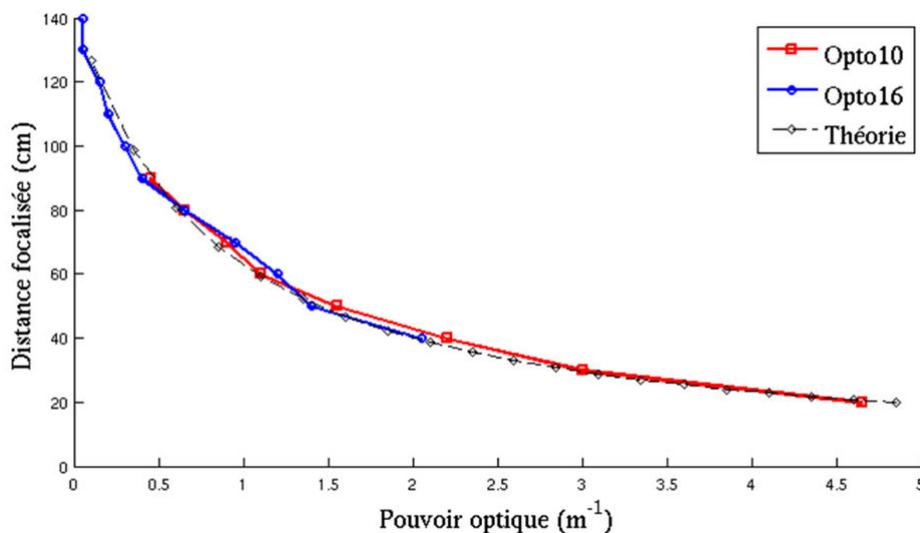


FIGURE 67 : DISTANCE FOCAL DETERMINEE PAR L'ALGORITHME POUR LES DEUX LENTILLES OPTOTUNE

Cette méthode est validée et peut donc être utilisée pour caractériser les distances nettes de la scène D en fonction des fréquences de fonctionnement des lentilles à cristaux liquides. Le Tableau 13 ci-dessous résume les commandes de distance focale pour l'ensemble des lentilles pour des distances de 20 à 60cm par pas de 5cm.

Ces fréquences sont très différentes d'une lentille à l'autre. Les lentilles L3 à L7 ont des couches résistives réalisées à partir du même mélange de PEDOT-PSS.

Tableau 13 : Loi de commande des lentilles modulables.

Lentille	Tension	D (cm)	20	25	30	35	40	45	50	55	60
L1	10V	Freq (kHz)	8.166	7.344	6.630	6.042	5.538	5.082	4.692	4.362	4.068
L2	10V		22.55	20.40	18.45	16.75	15.20	13.75	12.40	11.10	9.800
L3	5V		70	64	59	56.02	54	52.32	50.88	49.35	47.68
L4	5V		108.2	103	98.35	94.1	89.65	84.6	80	77	74.07
L5	20V		700	628	592.2	567.4	545.5	528.1	509.0	487.6	465.9
L6	20V		954.5	902	858.4	822.8	792.5	754.8	730.4	712.4	696.8
L7	5V		336	313.9	291.7	272	256.1	248	230.7	221.3	213.3
Opto	5V	1/f	4,7	3,845	3,095	2,505	2,11	1,8	1,52	1,255	1,075

Des couples (L3, L4) et (L5, L6), pour lesquelles seule l'épaisseur de la lentille diffère, on peut conclure que plus la couche de cristal liquide est épaisse, plus les fréquences nécessaires pour une même amplitude de polarisation sont élevées. Aussi, plus la fréquence est haute, plus la lentille focalise près.

Les tensions de polarisation ont été choisies empiriquement de sorte à être relativement faibles tout en garantissant une réponse correcte. Remarquons que plus le cristal liquide à un $\Delta\epsilon$ élevé, plus la tension est faible.

Ces lois de commande vont désormais permettre d'évaluer les lentilles en ce qui concerne leur vitesse de focalisation, leur zoom de focalisation et leur qualité pour une application d'imagerie.

3.4.3 Mesure des vitesses de stabilisation du foyer

L'intérêt de mesurer la vitesse de stabilisation du foyer pour une lentille donnée est de pouvoir évaluer la cadence d'image que pourrait avoir une application DfF.

Pour mesurer la vitesse de stabilisation de la focalisation des lentilles, nous proposons d'utiliser un critère de stabilité construit sur la différence de deux images successives. Quand la focalisation est stable, les différences observables sont seulement dues aux variations de la scène et au bruit pixélique. Les variations de la scène sont annihilées par l'utilisation d'une scène statique et la suppression de toute source lumineuse clignotante donc seul le bruit de mesure persiste. Dès que l'ordre de changement de foyer est donné, le flou évolue : cela impacte la différence inter-image. L'objet de la mesure de vitesse est donc d'arriver à capturer cette variation et à la mesurer dans le temps. L'équation (3.8) décrit le calcul de ce critère, où t représente le temps, Δt le temps entre deux acquisitions (16ms à 60fps), I_w et I_h les dimensions des images.

$$C_{stab} = \sum_{I_w, I_h} |Im(t + \Delta t) - Im(t)| \quad (3.8)$$

La Figure 68 présente les organigrammes de la méthode de mesure proposée.

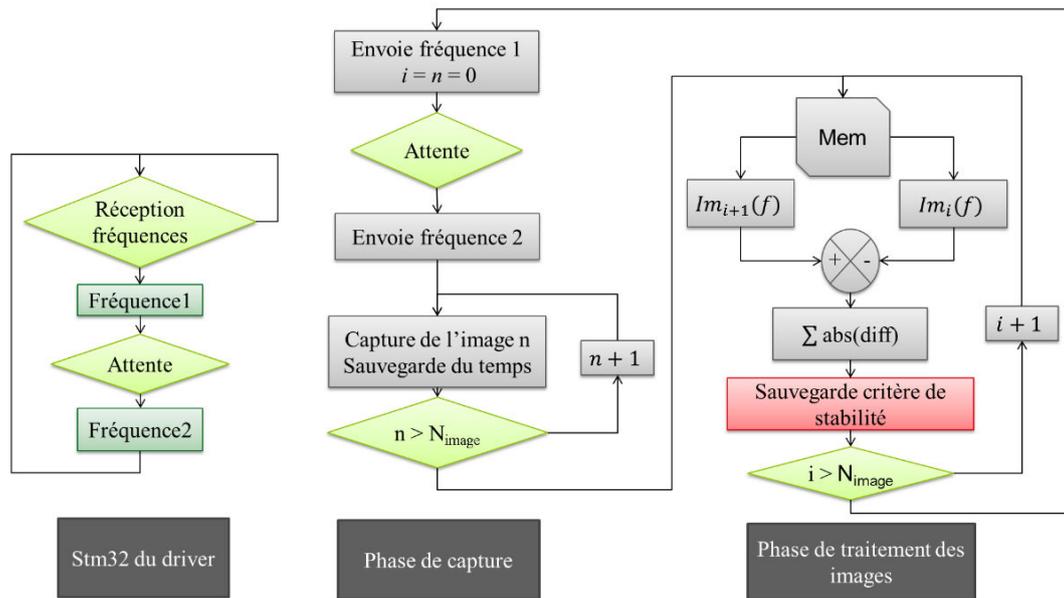


FIGURE 68 : ALGORITHME DE MESURE DU TEMPS DE STABILISATION DU Foyer. PROGRAMME DU CONTROLEUR DE LENTILLES A CRYSTALS LIQUIDES (DROITE), PHASE DE CAPTURE D'IMAGE (MILIEU) ET PHASE D'ANALYSE D'IMAGE ET D'EXTRACTION DU CRITERE DE STABILISATION (GAUCHE)

Le premier organigramme (gauche sur la Figure 68) correspond au programme exécuté par le contrôleur de lentille à cristaux liquides.

Le programme principal est divisé en 2 phases qui s'exécutent l'une après l'autre, une phase de capture et une phase de traitement. La première envoie deux commandes de foyer au contrôleur de lentille, puis commence à capturer et sauvegarder des images. Le temps d'attente entre ces deux commandes permet de laisser le temps à la lentille de se stabiliser sur la première distance focale.

Le driver applique la première fréquence immédiatement à sa réception, puis la seconde après une temporisation d'environ 1 seconde. Ainsi, le programme principal couvre la totalité d'un changement de foyer en capturant N_{image} images.

Le programme de caractérisation est exécuté sur une distribution linux non temps réel. C'est pour cette raison qu'il envoie les deux commandes dès le début, puis passe dans un mode de capture pendant lequel il ne fait rien de plus que récupérer les images du capteur. Cela permet d'augmenter la robustesse de la mesure en limitant les tâches de l'OS et ainsi d'éviter de perturber l'acquisition.

La phase de traitement commence une fois l'acquisition terminée. N_{image} ont été conservées (600 dans notre cas) pour calculer la somme des valeurs absolues de la différence pixel à pixel entre deux images. On trace alors ce critère de stabilisation en fonction du temps. La Figure

69 montre l'évolution du critère au cours d'une variation de foyer de la lentille L2 pour deux sauts de celui-ci.

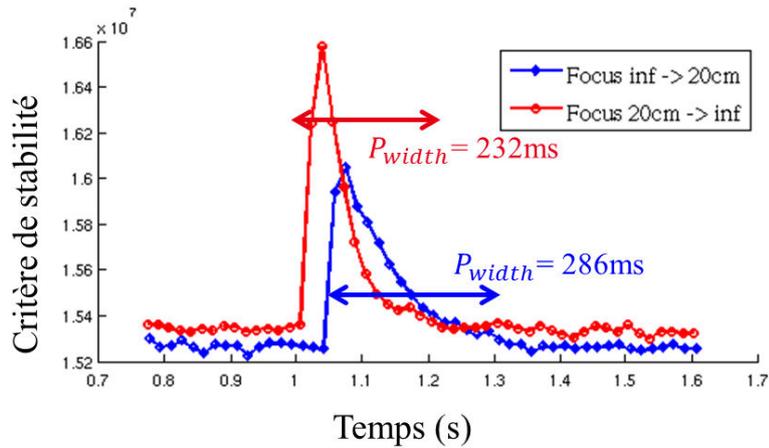


FIGURE 69 : MESURE DU TEMPS DE FOCALISATION POUR L2

La variation du critère de stabilité est parfaitement visible pour une variation de foyer de l'infini à 20 cm et inversement. La largeur des pics du critère P_{width} va permettre d'évaluer le temps nécessaire à la lentille pour changer de distance focale.

Avant de donner les temps de focalisation pour chacune des lentilles, analysons le résultat ci-dessus. Cette méthode permet effectivement de lire un temps qui correspond à la largeur des pics, mais un biais de mesure est introduit par le simple fait d'échantillonner. En effet lorsque le temps de focalisation est proche de la cadence image, ou inférieur, l'imprécision devient très grande comparée au temps prédit. La Figure 70 illustre un tel cas.

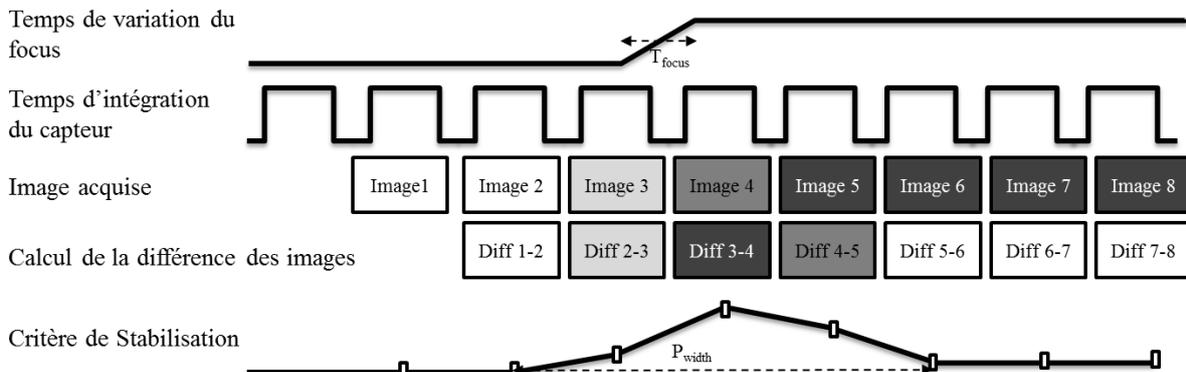


FIGURE 70 : DETAIL DE LA MESURE DE VITESSE DE STABILISATION POUR DES TEMPS DE VARIATION ET DE PRISE D'IMAGE SIMILAIRES.

Pour une focalisation rapide se produisant entre deux images on aurait pu s'attendre à un pic très étroit. Mais parce qu'il est très compliqué de synchroniser le changement de foyer et la capture d'image dans notre situation expérimentale, la mesure est bien plus longue que le temps effectif de la variation de foyer. Une variation de foyer sur deux images provoque une estimation à 3 temps images au lieu de 1. De ces considérations, on déduit que le temps de focalisation T_{focus} doit être approximé par l'inéquation (3.9).

$$P_{width} - 4 \cdot T_{image} < T_{focus} < P_{width} - 2 \cdot T_{image} \quad (3.9)$$

Où T_{image} représente le temps d'acquisition d'une image, 16ms dans notre cas (60fps).

Fort de ceci, des temps de focalisation pour les 9 lentilles sont donnés dans le Tableau 14.

Tableau 14 : Temps de stabilisation du foyer des lentilles modulables en milliseconde et leur viscosité respective.

Distance de focalisation (cm)	20↔25		20↔inf		Viscosity (mPa.s)
	20←25	20→25	20←inf	20→inf	
L1	N.C.	N.C.	110	76	203
L2	N.C.	N.C.	1900	200	203
L3	6393	4898	6225	466	464
L4	6957	5363	10259	731	464
L5	8667	7935	3969	184	47
L6	9929	5546	6077	500	47
L7	7123	6310	5695	1510	464
Opto10	N.C.	N.C.	10*	6*	
Opto16	N.C.	N.C.	15*	15*	

Ces deux sauts de foyer particuliers ont été choisis car l'un représente l'excursion maximale de foyer en commun entre toutes ces lentilles, le second correspond au plus gros saut de fréquence d'après les lois de commande vues précédemment. Nous avons alors supposé qu'il s'agissait de la mise au point la plus lente pour un balayage de foyers entre 20 et 60 cm par pas de 5cm. Pour chacun d'eux, le premier chiffre représente un rapprochement de la distance focale et le second la variation opposée.

Les données pour les lentilles Optotune proviennent de leur documentation, elles sont en effet trop rapides pour pouvoir être évaluées avec cette caméra. Ceci étant sans considérer le léger délai introduit par le contrôleur mais celui-ci est difficile à appréhender à cause de l'aspect non temps réel de l'OS. Les lentilles liquides sont cependant, de loin, les plus rapides des lentilles expérimentées.

Les temps de focalisation courts pour les lentilles L1 et L2 n'ont pas pu être caractérisés à cause du bruit de mesure. En effet, pour de si petites ouvertures la profondeur de champs est telle (~5cm à 25cm pour L2) que la variation des images se retrouve totalement masquée par

le bruit. Leur temps de focalisation à grande excursion est plutôt court comparé aux autres lentilles à cristaux liquides.

Il est étrange de constater que le temps de relaxation des lentilles à cristaux liquides est plus court que le temps de focalisation. En observant la Figure 71 qui représente l'évolution du critère de stabilité de la lentille L3 pour une grande excursion, on voit un pic puis une lente stabilisation vers la distance focale.

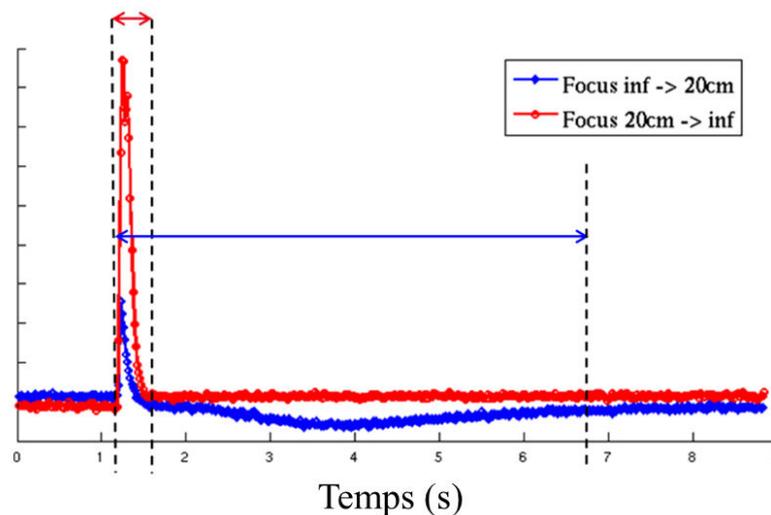


FIGURE 71 : EVOLUTION DU CRITERE DE STABILISATION POUR L3

Nous expliquons ce phénomène par la considération du temps d'activation des molécules de cristal liquide. Lors de leur première polarisation, les cristaux liquides nécessitent normalement un temps d'adaptation au champ électrique qui peut prendre quelques secondes.

Il ressort de ce tableau que pour une même taille de pupille et un même cristal liquide, (L3 et L4) et (L5 et L6) les lentilles les plus « rapides » sont les plus fines.

Pour un même cristal liquide et une même épaisseur (L4 et L7), la lentille avec la plus petite ouverture semble plus rapide.

En théorie les solutions de cristal liquide les plus visqueuses devraient être plus lentes mais cela ne semble pas être le cas sur ce Tableau 14. Cependant les valeurs des tensions appliquées sont différentes à cause du $\Delta\epsilon$, ce qui peut influencer la vitesse de rotation des molécules. Les aléas de réalisation peuvent également être responsables de ces aberrations.

Il ressort de cette étude que les lentilles à cristaux liquides nématiques de ce type sont lentes. Cependant elles restent compatibles avec des applications pour lesquelles on aurait des scènes statiques, ce qui peut être le cas de robot d'exploration comme le Mars Rover qui est extrêmement lent (environ 90m/h). L'utilisation de tensions de polarisation plus élevées permettrait d'accélérer quelque peu la commutation des cristaux mais au détriment de

l'intérêt des basses tensions pour les systèmes embarqués. Les lentilles Optotune en revanche montrent un potentiel intéressant pour le DfF.

La suite de ce manuscrit présente d'autres critères dont il faut tenir compte pour le DfF et qui peuvent être plus discriminants que la vitesse de stabilisation du foyer quant au choix de lentille.

3.4.1 Analyse de la diffusion

La diffusion est un paramètre critique pour une application DfF car elle augmente l'aspect de flou dans l'image, diminuant ainsi le contraste. Les photographies d'une scène (Figure 72) avec les lentilles Opto10 et L2 montrent que le flou avec les lentilles à cristaux liquides est plus prononcé. Les lentilles sont ici focalisées à 50cm.

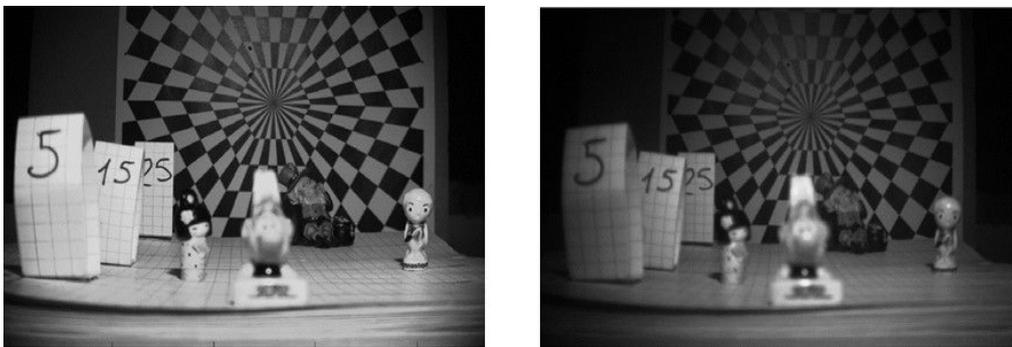


FIGURE 72 : PHOTOS D'UNE SCENE PRISES AVEC LA LENTILLE OPTO10 (GAUCHE) ET LA LENTILLE L2 (DROITE)

Nous verrons que cela peut être mieux avec d'autres caractéristiques que celle de L2 ; cependant ces images permettent de bien faire ressortir la différence. L'anisotropie du cristal liquide provoquant une diffusion de la lumière, ce qui n'est pas le cas de la technologie des lentilles liquides, nous allons donc concentrer cette étude de la diffusion sur les lentilles à cristaux liquides.

La diffusion en tant que telle est un phénomène très complexe. N'ayant pas de moyens techniques de la mesurer précisément, nous nous contenterons de la mettre en évidence et de la quantifier par la mesure de la résolution d'une image. Nous proposons également de regarder les fonctions de transfert de modulation des lentilles comme indicateur de qualité optique. Celles-ci tiennent notamment compte de l'absorption, de l'effet de pupille et des aberrations, en plus de la diffusion.

3.4.1.1 Mesure de la perte de résolution.

Une première expérience pour évaluer la dégradation de l'image à cause des cristaux liquides consiste à mesurer la différence de résolution avec et sans lentille. Pour cela, une mire USAF1951 a été placée à 40 cm de la caméra devant une source lumineuse uniforme. Une

première image témoin a été capturée avec la caméra « seule » et le polariseur (pour conserver les mêmes conditions de luminosité qu'avec les lentilles à cristaux liquide), l'objectif étant manuellement réglé pour focaliser à 40cm. Les lentilles à cristaux liquides ont ensuite été accolées à cet objectif (refocalisé à l'infini) puis commandées de sorte à focaliser sur la cible. La Figure 73 montre les images obtenues sans lentille modulable puis avec L3 et L2.

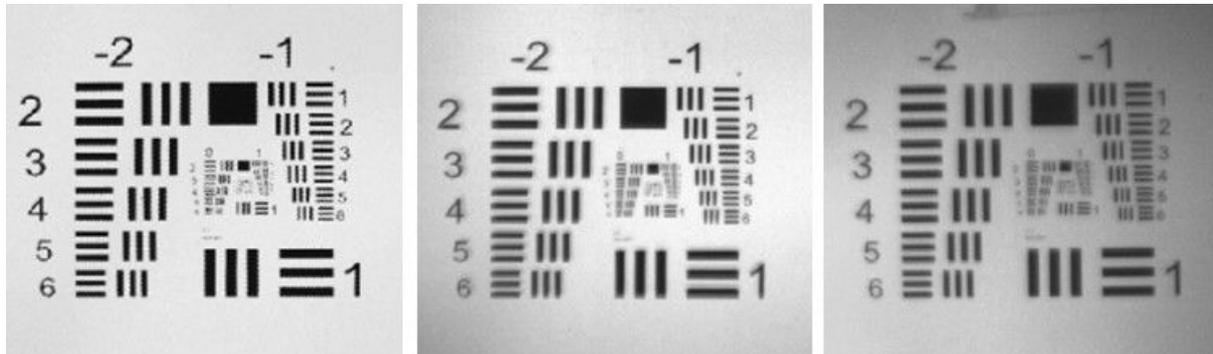


FIGURE 73 : IMAGE D'UNE MIRE USAF1951 PLACÉE A UNE DISTANCE DE 40CM DE LA CAMERA, CAMERA PLUS POLARISEUR (GAUCHE), L3 ET POLARISER (MILIEU) ET L2 ET POLARISEUR (DROITE)

La résolution dans l'image prise caméra seule (gauche sur la Figure 73) est proche de 2.83 paires de ligne par mm (pl.mm^{-1}). Avec les lentilles L3 et L2, elle est respectivement de 2.24 pl.mm^{-1} et de 1.54 pl.mm^{-1} . La luminosité décroît d'une image à l'autre principalement à cause du vignettage introduit par la réduction d'ouverture. Même si cette dernière joue sur la résolution, la différence de netteté due aux cristaux liquides reste remarquable sur ces images par l'observation du seul groupe -2.

Une seconde expérience intéressante consiste à voir l'influence d'une source ponctuelle lumineuse hors axe sur la diffusion lumineuse due à la couche de cristal liquide. La Figure 74 montre le banc expérimental d'une telle expérience et la Figure 75 les images obtenues pour les lentilles L3 à L6 (même pupille).



FIGURE 74 : PHOTO DU BANC DE TEST AVEC SOURCE LUMINEUSE HORS AXE

Une LED blanche est utilisée comme source lumineuse parasite. La cible USAF1951 est placée à 30cm de la caméra, la cible contrastée à 70 cm et la palette de couleurs à 140 cm.

Un voile blanc est présent sur les images prises lorsque la source lumineuse hors axe est allumée (colonne de droite).

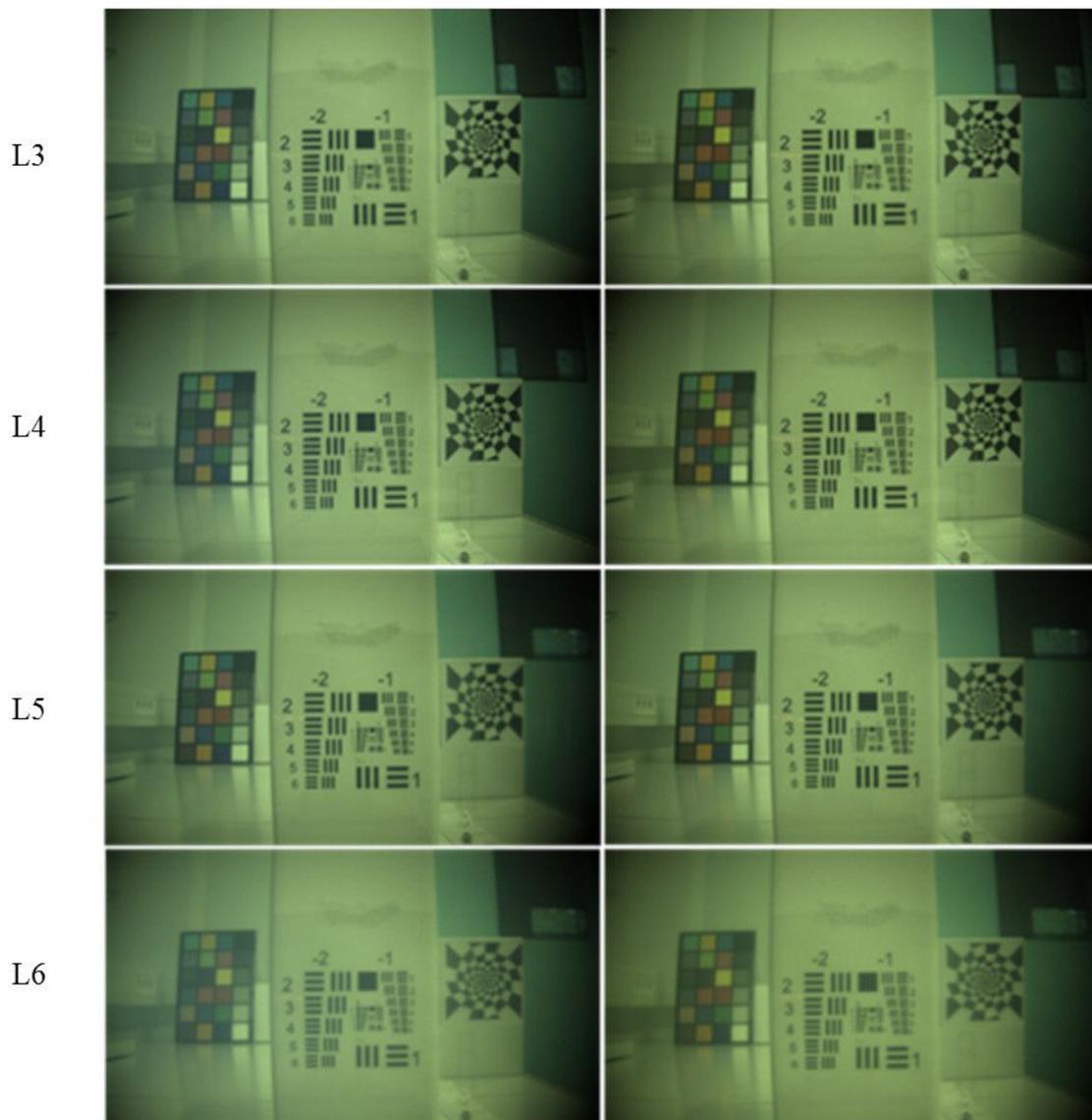


FIGURE 75 : COMPARAISON DES IMAGES AVEC (DROITE) ET SANS (GAUCHE) SOURCE LUMINEUSE HORS AXE POUR LES 4 LENTILLES DE PUPILLE 4MM. LES IMAGES SONT FOCALISEES A 30CM SUR LA MIRE USAF, LA MIRE DE DROITE EST A 70CM ET LA PALETTE COLOREE A 140CM

Cela confirme le caractère diffusant des cristaux liquides. Sur la figure L3, la résolution est de 0.793 pl.mm^{-1} avec la source lumineuse hors axe alors qu'elle est de 1 pl.mm^{-1} sans cette source de lumière. Par comparaison de L3 avec L4 et L5 avec L6, il ressort que l'augmentation de l'épaisseur de la couche de cristaux amplifie le voile lumineux. Les images avec L3 et L5 ainsi que celle avec L4 et L6, montrent que la diffusion augmente également avec la biréfringence du cristal.

Nous venons de voir que l'introduction de cristaux liquides provoque une diminution de la résolution. Plus l'épaisseur de la couche est grande, plus la lumière diffuse, ce qui peut devenir problématique pour des applications d'imageries. Au regard de ces seuls résultats, la lentille

L3 semble la plus adaptée des lentilles à cristaux liquides réalisées au cours de cette thèse pour une application DfF.

3.4.1.2 Fonction de transfert de modulation des lentilles à cristaux liquides

Pour mesurer l'impact direct des différents paramètres des lentilles (largeur pupille, épaisseur et nature de la couche de cristaux liquides), et faire une comparaison en termes de qualité d'image, nous avons mesuré les MTFs des 7 lentilles à cristaux liquides. A l'aide de la mire USAF1951 placée à 40cm, nous avons mesuré le contraste pour chacune des résolutions des groupes -2, -1 et 0. La MTF de la caméra avec le polariseur a été mesurée pour servir de témoin, puis celles des systèmes complets (caméra, lentille à cristaux liquides et polariseur). La Figure 76 montre l'ensemble de ces fonctions de transfert.

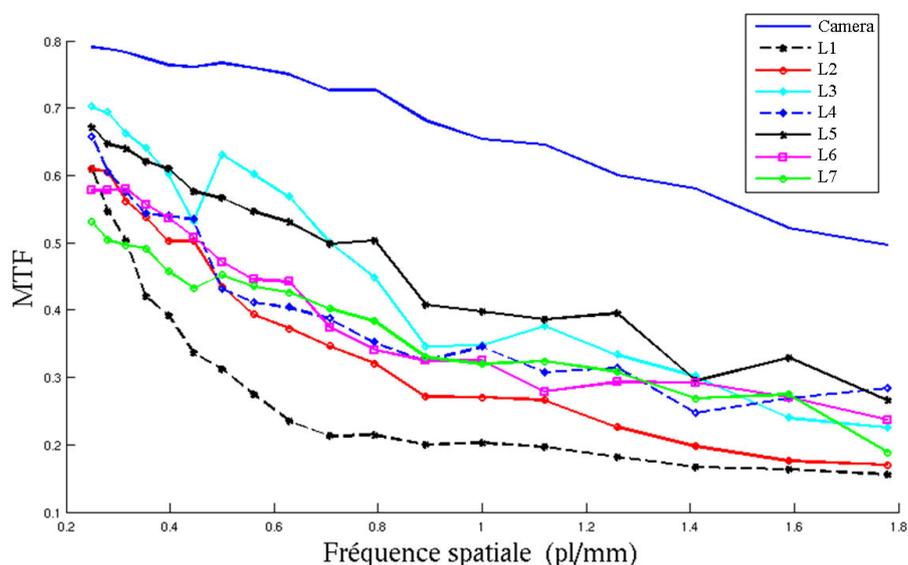


FIGURE 76 : MTF OPTIQUE DES SYSTEMES IMAGEURS, CAMERA + POLARISEUR ET CAMERA + LENTILLE + POLARISEUR, MESURER A PARTIR D'UNE MIRE USAF1951 PLACEE A 40CM

Les MTF confirment que les lentilles à cristaux liquides ont un fort impact sur la qualité des images. L1 et L2 sont les lentilles les moins performantes. Leur faible ouverture est très probablement en cause ici, en plus de la diffusion.

Si on considère les lentilles avec des ouvertures plus grandes, L3 et L5 semblent les plus performantes, ce qui montre une fois encore l'impact de l'épaisseur de la couche de cristaux liquides sur la qualité d'image.

L3 présente une MTF avec une chute de contraste pour les fréquences inférieures 0.4 pl.mm^{-1} , celle-ci est due à la présence d'aberrations locales et au vignettage qui ont influencé l'image. On voit ici l'inconvénient de mesurer la MTF avec une mire USAF : chacune des fréquences n'est pas strictement analysée au même endroit dans l'image et par conséquent est sujette

aux défauts locaux. De ce fait, la comparaison de L3 et L5, pour lesquelles seule la nature du cristal est différente, est moins évidente.

La qualité de L7 est discutable, principalement à cause de défaut de fabrication. (Voir image 5 entre polariseurs croisés en Annexe 5).

3.4.2 Mesure de zoom de focalisation

Afin d'avoir une mesure quantitative du zoom de focalisation, notre programme compare des images successives deux à deux d'un cube d'image multi-foyers, calcule le facteur d'échelle puis trace son évolution cumulée à travers le cube d'images. Nous avons utilisé l'outil *imregister* de matlab et extrait le facteur d'échelle via la structure *tForm*. La scène présentée sur la Figure 72 a été utilisée pour construire un cube d'images à 9 niveaux de profondeur, 20 à 60 cm par pas de 5 cm. La Figure 77 montre l'évolution du zoom de focalisation en fonction de la distance focale pour les 9 lentilles évaluées. Le calcul de facteur de zoom correspondant à 2 images successives, soit deux profondeurs successives, est attribué à la médiane de ces 2 distances.

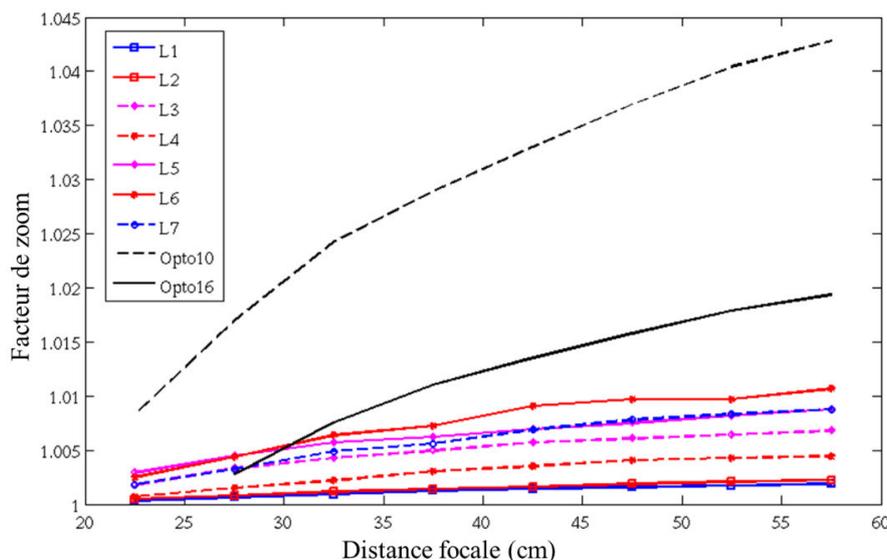


FIGURE 77 : FACTEUR DE ZOOM CUMULE EN FONCTION DE LA DISTANCE FOCAL.

Dans l'ensemble, les lentilles à cristaux liquides présentent un zoom de focalisation réduit comparé aux lentilles liquides. Cela s'explique par le fait que les lentilles liquides ont un dioptre déformable donc une épaisseur non constante à l'inverse des lentilles à cristaux liquides dont l'épaisseur est fixe. L'effet de zoom reste inférieur à 1% sur cette excursion de mise au point entre 20 et 60 cm pour les lentilles à cristaux liquides alors qu'il s'élève au-delà de 4% pour la lentille Opto10. C'est un avantage significatif des lentilles à cristaux liquides.

A titre d'information, le zoom cumulé pour la scène soleil utilisée dans le chapitre 2, capturée avec un appareil réflex avec un objectif à lentille mobile, dépasse les 8%.

Globalement, les lentilles réalisées avec une même solution de cristal liquide ont des évolutions de zoom de focalisation relativement proches. Pour des pupilles similaires, il semblerait que le cristal liquide le plus biréfringent soit sujet à un zoom de focalisation légèrement supérieur. Pour un même cristal et une épaisseur similaire (L4 et L7), une augmentation de l'ouverture semble également amplifier le zoom de focalisation.

On remarque que le facteur d'échelle pour le passage de 20 à 25cm est relativement important pour les lentilles L3, L5, L6 et L7 comparé à leur évolution pour les distances suivantes. C'est un phénomène difficile à expliquer, il se pourrait que cela vienne d'aberrations du champ, et donc de phase, dues à des couples (fréquence, tension) non parfaitement adaptés aux lentilles.

3.4.3 Evaluation des lentilles avec un algorithme DfF

Nous avons utilisé l'algorithme intra-image haute résolution présenté dans le chapitre précédent, avec des fenêtres glissantes 16×16, afin d'extraire la profondeur, la confiance et les images AiF à partir des lentilles évaluées dans ce chapitre. Par considération des résultats précédents, la lentille L1 a été écartée. La scène utilisée est montrée Figure 78. Chacune des étiquettes chiffrées est placée à 10cm l'une de l'autre. La cible du fond est située à 60cm de la caméra.

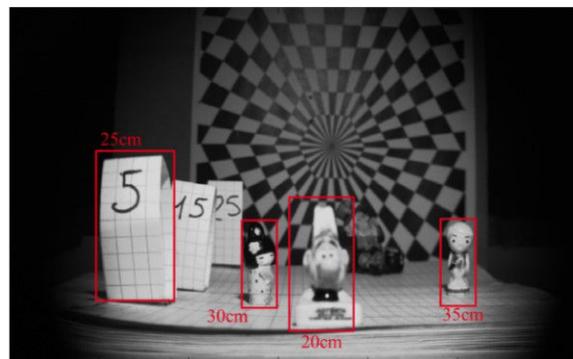


FIGURE 78 : SCENE D'ÉVALUATION DE PERFORMANCE POUR L'EXTRACTION DE PROFONDEUR

Les cartes de profondeur extraites avec chacune des lentilles ont été filtrées de sorte à ce qu'elles soient plus facilement comparables. Seuls les éléments dont la confiance est supérieure à la médiane de leur carte de confiance respective sont conservés (Figure 79).

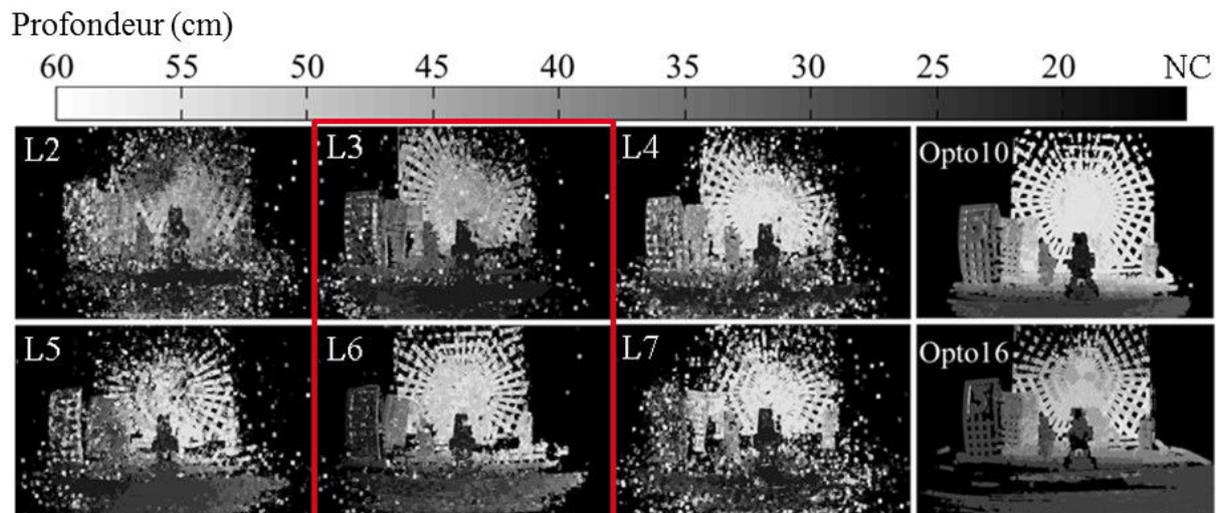


FIGURE 79 : CARTES DE PROFONDEUR FILTRES EXTRAITES AVEC DES LENTILLES MODULABLES

Les lentilles Optotune fournissent des cartes de profondeur particulièrement pertinentes, en particulier Opto10. C'est principalement grâce à l'absence de diffusion lumineuse au sein des matériaux utilisés. On observe un léger décalage entre la réalité et Opto16, qui s'explique par le fait que cette lentille ne permet pas de faire une mesure à moins de 25cm (30cm d'après la documentation). D'autre part, le positionnement étant manuel et la forme des boîtiers différente, il peut y avoir un léger écart des modèles théoriques de chacune des lentilles qui peut impliquer des commandes de foyers légèrement différentes.

Malgré des pupilles plus grandes pour les lentilles liquides, l'ouverture totale du système est très proche de celle obtenue avec les lentilles à cristal liquide pour des pupilles de 4 et 5mm, du moins pour Opto10. Preuve en est le vignettage apparent sur l'image Figure 78 qui provient de l'utilisation d'Opto10.

En ce qui concerne les lentilles à cristaux liquides, L3 et L6 semblent donner les meilleurs résultats (notons pour L3 une petite erreur dans l'arrière-plan qui est tout simplement due à un décalage intempestif de la cible pendant la mesure). Il semblerait que L5 soit plus sujette aux aberrations au regard de la périphérie de l'image. Celle-ci est moins bien évaluée comparé à l'utilisation d'autres lentilles (L3 et L4). Cela peut provenir de défauts de fabrication d'une part mais potentiellement à cause d'une tension de polarisation qui pourrait ne pas être bien adaptée. Rappelons que la même amplitude de 20V est imposée à L5 qui, contre toute attente, donne un meilleur rendu. Une tension non adaptée pourrait causer un aplatissement du gradient d'indice, supposé sphérique, et ainsi une réduction de la pupille effective de la lentille. On observe une amélioration notable de la qualité entre la lentille L2 et les autres, grâce à une réduction de la profondeur de champ par l'augmentation de la pupille et une réduction d'épaisseur de la couche de cristal liquide.

Les images AiF présentées Figure 80 sont obtenues à partir des cubes d'images multi-foyers et des cartes de profondeur lissées par un filtre médian 21×21 .

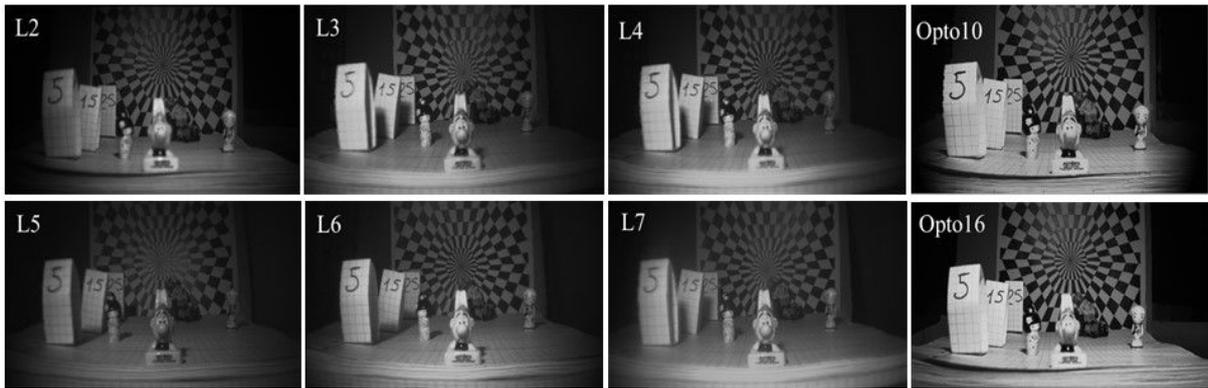


FIGURE 80 : IMAGE AiF GÉNÉRÉE À PARTIR DES LENTILLES L2 À L7 ET OPTOTUNE.

Les images AiF générées à partir des lentilles L6, Opto10 et Opto16 présentent les meilleures qualités. Opto16 a l'avantage de contenir très peu d'artefact de zoom de focalisation comparé à Opto10 comme le montre l'exemple indiqué par la flèche orange sur la Figure 81 ; en revanche le premier plan (carré rouge) de la scène n'est pas net car il est trop proche pour la vergence max de la lentille.

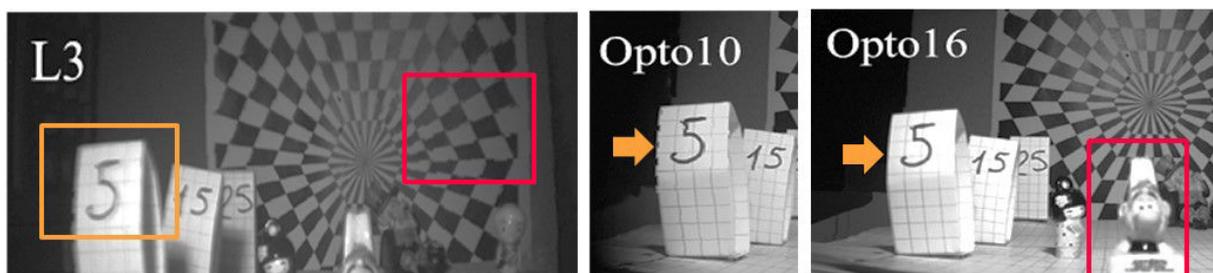


FIGURE 81 : DÉFAUTS DE RECONSTRUCTION D'IMAGE

La qualité de L3 est altérée par la présence d'artefacts de mouvement sur la cible d'arrière-plan qui s'est décalée pendant la capture du cube multi-foyers (Figure 81). Aussi, malgré une estimation correcte de l'étiquette 5 dans la carte de profondeur, celle-ci n'est pas particulièrement nette sur l'image AiF ce qui dénote la présence de défaut sur ce bord de l'image (surexposition locale ou défaut de lentille). L6 est ternie par le voile blanc caractéristique de la diffusion ; cependant l'ensemble de la scène est plutôt bien reconstruit.

3.4.4 Observations

Au cours de nos expériences, nous avons remarqué que lors d'utilisation de haute tension (>20V) et de hautes fréquences, la commande en fréquence des lentilles a tendance à dériver. A titre d'exemple, après une heure d'utilisation à une amplitude de polarisation de 40V, la lentille L6 nécessitait une fréquence de 1.58MHz au lieu de 1.19MHz pour avoir l'image nette à 27cm. Cela peut s'expliquer par un échauffement des cristaux liquides qui modifierait les propriétés physiques, notamment la viscosité ou encore par la présence de charge piégée dans la couche de cristaux liquides. Après une période de repos la lentille a de nouveau fonctionné à 1.19MHz.

Nous avons également noté que les lentilles avaient tendance à se dégrader avec le temps. Pour une même amplitude de fonctionnement, nous avons observé un décalage de fréquence vers le bas, et une perte de netteté. Et ce, même entre deux phases de test espacées de quelques semaines, donc sans stress dû à la polarisation. C'est particulièrement vrai pour les lentilles L5 et L6. Cela pourrait s'expliquer par une évolution chimique irréversible dans la cellule : il se peut que des microbulles d'air présentes dans le cristal liquide viennent réagir avec la couche résistive, changeant ainsi ses caractéristiques. Ce phénomène étant plus avancé avec les lentilles L5 et L6, la polarisation (20V et hautes fréquences) a pu créer un échauffement de la solution résistive qui aurait joué le rôle de catalyseur. Cela pourrait également venir d'une réaction entre les solutions de cristal liquide elles-mêmes et la couche résistive. De futures études permettraient de mettre au clair le vieillissement des lentilles à cristaux liquides de ce type réalisées à partir d'une couche de PEDOT-PSS. De futures études permettraient de mettre au point un modèle tenant compte de ces observations

Il est également arrivé que des bulles visibles apparaissent dans certaines lentilles probablement à cause d'un manque de colle lors de la fermeture de la cellule.

Enfin ne disposant pas de matériel adéquat, nous avons été sujets à un manque de précision sur le positionnement des lentilles devant la caméra. Cela provoque des divergences d'intensité lumineuse sur le capteur. Et cela peut également dégrader légèrement les résultats. Ceci explique notamment les différences d'éclaircissement visibles sur les images AiF du paragraphe précédent.

3.5 Synthèse des résultats de mesure

Le Tableau 15 ci-après fait une synthèse de la comparaison des lentilles en reprenant les résultats de mesure ainsi que les considérations importantes précédemment explicitées pour une application DfF embarquée.

En termes de résultat, en considérant seulement la contrainte « qualité » des images AiF et de la carte de profondeur, les lentilles L3, L6, Opto10 et Opto16 sont prometteuses. En ajoutant la contrainte d'embarquée, donc faible tension, on élimine L6.

Le choix final dépend de la contrainte considérée comme majeure pour l'application. S'il s'agit du zoom de focalisation toléré, L3 paraît alors le meilleur choix. Si c'est de la distance focale minimum atteignable, dans ce cas Opto16 est éliminée. Enfin si on privilégie la qualité de l'image, Opto16 semble le meilleur compromis.

Tableau 15 : Comparaison des lentilles sur l'ensemble des critères.

	L1	L2	L3	L4	L5	L6	L7	Opto10	Opto16
Tension de commande	+	+	++	++	-	-	++	++	++
Vitesse	-	-	---	---	---	---	---	++	++
Qualité d'image	---	--	+	-	+	-	-	+++	+++
Distance focale courte	++	++	++	++	++	++	++	++	-
Zoom	+++	+++	++	++	+	+	+	---	-
Restitution Profondeur	---	--	+	-	-	++	-	++	++
Image AiF	---	---	+	+	-	++	-	++	++

3.6 Conclusion

Nous avons vu dans ce chapitre les différents paramètres dont il faut tenir compte pour choisir un module optique pour une application DfF (ouverture, excursion de foyers, profondeur de champ, magnitude du zoom de focalisation...). Ces paramètres nous ont conduits à l'utilisation de lentilles modulables construites à partir de 2 technologies : des lentilles liquides d'une part avec l'expérimentation de 2 modèles de la société Optotune, et des lentilles modales à cristaux liquides nématiques sur lesquelles nous avons la capacité de faire de l'ingénierie. Des lentilles à cristaux liquides ont été réalisées puis évaluées en considérant les contraintes du DfF embarqué.

Il s'avère qu'une lentille modale de 4mm de pupille, avec le cristal liquide MLC-2140 [$\Delta n=0.25$] constitue un bon compromis pour une extraction de profondeur dans un proche environnement invariant par tranche d'espace de 5cm. Sa vitesse est certes relativement lente, mais peut être qu'en retravaillant la commande il deviendrait possible de capturer des images au cours d'une grande excursion de mise au point et ainsi d'extraire le cube d'images multi-foyers pendant seulement un temps de focalisation. Aussi avec une optimisation des étapes de fabrication, ces lentilles présenteront un potentiel intéressant pour des applications d'imagerie. C'est une voie d'exploration pour de futurs travaux.

Il serait également intéressant d'ajouter à cette comparaison des lentilles de 4 et 5mm de pupille avec le cristal liquide MDA-98-1602 [$\Delta n=0.26$] qui est moins visqueux et donc potentiellement plus rapide. Si on tolère une polarisation d'amplitude 10V, celui-ci devrait donner une lentille plus fine que L3, donc probablement moins diffusante et un zoom de focalisation certainement très similaire. L'utilisation de cristaux liquides plus rapides, par exemple comportant des composés fluorés synthétisés spécifiquement pour l'application, constitue aussi une solution à considérer pour améliorer la vitesse de variation de focalisation

de ces lentilles et ainsi alléger la contrainte sur l'invariance de la scène. Le MCL-2172 permettrait, selon de très récentes expérimentations de multiplier par 10 la vitesse.

Pour l'évaluation en temps réel de l'algorithme DfF intra-image basse résolution décrite dans le chapitre suivant, nous avons fait le choix d'utiliser une lentille liquide. En effet, malgré des résultats intéressants de la lentille L3, sa dégradation avec le temps, d'une part, ainsi que la vitesse de commutation des cristaux liquides, d'autre part, rendent difficile l'évaluation temps réel de notre algorithme avec cette lentille. La cadence de sortie serait alors limitée par la lentille et l'impact de la carte de traitement ne pourrait être mesuré.

Les deux lentilles liquides présentent une qualité d'image et des vitesses de focalisation similaires. En revanche, bien que plus avantageuse en termes de zoom de focalisation, la puissance optique de la lentille Opto16 est plus réduite que celle d'Opto10. Nous avons donc fait le choix d'utiliser la lentille Opto10 pour le portage temps réel sur plateforme embarquée.

Chapitre 4 Prototype expérimental d'acquisition d'images 3D

4.1	Introduction	111
4.2	Description de la plateforme	112
4.3	Portage de l'algorithme	114
4.4	Résultats obtenus	117
4.5	Conclusions sur le portage matériel	124

4.1 Introduction

Dans ce chapitre, nous combinons les résultats des parties algorithmique et optique de cette thèse afin de construire un prototype de capteur permettant l'acquisition d'images annotées en profondeur et en confiance. Ce prototype est réalisé à l'aide d'une carte électronique développée spécialement pour des applications d'imagerie embarquée. Cela permet d'une part d'évaluer la cadence image minimale que pourrait avoir un tel système mais aussi d'observer la robustesse de la méthode proposée sur des scènes variables ou encore d'analyser rapidement l'influence d'autres paramètres, telle que l'exposition de la caméra sur le critère de netteté.

Suite aux analyses présentées, nous avons choisi l'algorithme de DfF intra-image basse résolution avec une analyse par fenêtre glissante et la lentille Opto10 afin de proposer une implémentation temps réel d'extraction de carte de profondeur.

Nous présentons le développement du prototype puis montrons quelques images totalement focalisées ainsi que leurs profondeurs et confiances respectives. Les temps de calcul et la cadence de sortie des données ont également été mesurés pour cette architecture.

4.2 Description de la plateforme

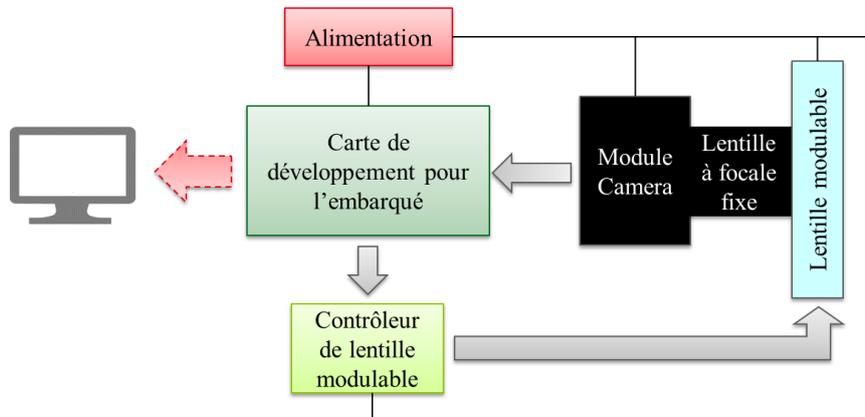


FIGURE 82 : ARCHITECTURE DU SYSTEME

La Figure 82 résume l'ensemble des éléments qui constituent le prototype, à savoir une caméra, une lentille modulable et son driver associé, une carte de calcul en adéquation avec les contraintes de l'embarqué sur laquelle est porté l'algorithme et une source d'alimentation.

4.2.1 Point de vue matériel

En ce qui concerne l'acquisition d'images, nous réutilisons le module caméra Leopard Imaging M021C présenté dans le chapitre précédent. Il est peu encombrant et fournit des images couleurs en format RAW avec une résolution de 1280×960 pixels, à une cadence de 60 images par seconde. Il opère via une liaison USB3, ce qui lui fournit une alimentation fiable. En revanche cela contraint légèrement le choix de la carte de traitement qui doit nécessairement proposer une connexion USB.

Suite à la comparaison des performances optiques du chapitre précédent, nous avons fait le choix d'utiliser la lentille Opto10 pour remplir la fonction de balayage de foyers. Elle permet de focaliser plus près que la lentille Opto16 tout en ayant une excellente qualité d'image malgré un zoom de focalisation relativement prononcé. Son contrôleur dispose également d'une liaison USB qui sert à la fois pour l'alimentation de la lentille et la communication des vergences de commande.

Le choix de la carte de traitement dépend entièrement de la façon dont le portage de l'algorithme est réalisé. Pour permettre un prototypage rapide dans le cadre du traitement d'image, le meilleur moyen est de s'appuyer sur des bibliothèques de code existantes. Parmi les bibliothèques graphiques existantes, OpenCV est très certainement la plus connue en traitement d'image. Elle est très bien documentée, également suivie par une active communauté internationale et permet l'accélération matérielle par GPU. En revanche elle est

faite pour fonctionner sur systèmes d'exploitation (OS), ce qui la rend sujette aux aléas d'ordonnancement, notamment aux aspects non temps réel. Cela donne cependant l'avantage d'une mise en place rapide et une adaptabilité à de nombreux systèmes mais ajoute une contrainte supplémentaire pour le choix de la plateforme embarquée. Celle-ci doit supporter un OS compatible avec OpenCV.

La bibliothèque OpenCV intègre tout un panel de fonctions adaptées pour les architectures de traitement massivement parallèles CUDA. Ce type d'architecture est conçu par NVIDIA pour les processeurs graphiques (GPU). Cela permet notamment de faire de nombreux traitement par blocs très rapidement, ce qui semble être une bonne option au regard de ce qui a été dit précédemment concernant l'analyse de contraste.

Nous avons opté pour la plateforme de développement Nvidia Jetson Tegra K1 [108] (référéncée sous le terme Jetson dans la suite de ce manuscrit). Elle comprend des outils permettant d'exploiter la puissance des GPUs pour des applications embarquées. Elle est notamment basée sur le système sur puce (SOC) NVIDIA Tegra K1 SoC, et intègre les mêmes cœurs de calcul NVIDIA Kepler que ceux qui équipent de puissants supercalculateurs mais avec un format relativement adapté aux systèmes embarqués (Figure 83).

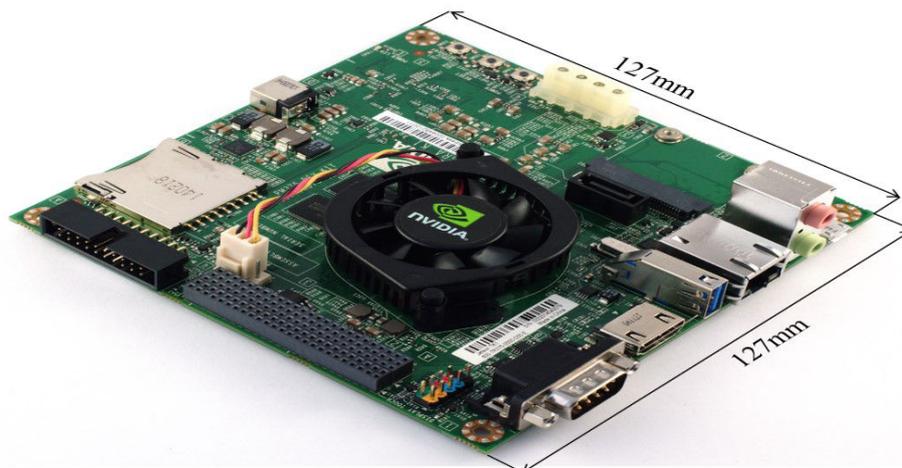


FIGURE 83 : NVIDIA JETSON TK1

Ses composants sont finalement proches de ce qu'on peut trouver sur une carte graphique d'un PC moderne tout en ayant une consommation relativement réduite. NVIDIA propose une distribution Linux spécialement adaptée pour cette carte. Bien que non temps réel, cette distribution permet de simplifier la mise en place du prototype.

En ce qui concerne l'alimentation, la source de la Jetson est capable de fournir une cinquantaine de watts ce qui est amplement suffisant pour faire fonctionner l'ensemble du système.

Un moniteur est également utilisé pour la visualisation des résultats d'une part, et également pour avoir accès à l'interface graphique de la distribution linux installée sur la Jetson.

4.2.2 Point de vue logiciel

Il existe plusieurs distributions linux compatibles avec la Jetson TK1. Pour notre part, nous avons opté pour la distribution Linux4Tegra (version « officielle » fournie par Nvidia) qui est une version d'Ubuntu 14.4 avec des drivers préconfigurés spécialement pour le matériel présent sur cette carte de développement.

Pour le traitement d'image, il existe une version précompilée d'OpenCV2, OpenCV4tegra 2.4, qui a été modifiée par NVIDIA pour s'adapter tout particulièrement à la Jetson. Cependant la configuration de transmission de pixel de la caméra Leopard Imaging n'est pas standard dans le monde des caméras USBs. Celle-ci fournit des images RAW 1280×720 codées sur 12bits via une liaison USB3.0. Ces données sont rangées sur 2 mots de 8bits, les 4 MSBs sont forcés à 0, les 12 autres correspondent à l'intensité des pixels, puis encapsulées au format YUY2. Cela implique un problème de récupération des données lors de l'utilisation de cette version précompilée d'openCV2. En effet, les données reçues sont soit tronquées, soit mal interprétées par un organe logiciel sur lequel nous n'avons pas la main. Pour pallier ce problème, nous avons pris l'initiative de modifier les sources d'openCV, afin de rendre compatible le logiciel de récupération d'image avec notre caméra. L'avantage d'une telle manœuvre est qu'elle permet d'utiliser OpenCV3 qui est une version plus récente de la bibliothèque de traitement d'image que celle précompilée par Nvidia. En revanche, il est possible qu'il y ait une légère diminution d'efficacité à cause de la perte de l'optimisation du code « spéciale Jetson » pour la couche GPU.

Nous avons codé les algorithmes en C++, afin d'avoir une exécution directe et d'éviter toute interprétation de code par une quelconque machine virtuelle qui pourrait ralentir le traitement.

4.3 Portage de l'algorithme

L'algorithme intra-image faible résolution vu précédemment permet d'extraire des cartes de profondeurs et de coefficients de confiance ainsi qu'une image totalement focalisée à partir des coefficients d'ondelette de Haar du premier ordre. L'extraction est faite par évaluation du support de l'histogramme par fenêtre glissante de $w_x \times w_y$ coefficients, puis par recherche du maximum de ce support en fonction de la distance de focalisation. Le traitement par bloc est hautement parallélisable et est donc particulièrement adapté au traitement de données sur GPU.

Ubuntu est un système d'exploitation non temps réel et de ce fait, des temps de latence sont susceptibles d'apparaître et de perturber le bon déroulement de l'algorithme. L'affichage

temps réel de la vidéo peut ainsi provoquer des perturbations, en particulier avec la Jetson qui semble avoir des difficultés à maintenir à jour plusieurs flux d’affichage. Cependant, le principal problème qui se pose consiste à synchroniser correctement la prise d’image avec la commande de la lentille. En effet, la commande de la lentille repose sur l’écriture d’un foyer dans son driver et il s’écoule un temps inconnu avant que l’ordre soit exécuté. D’autre part, au niveau de l’acquisition d’image, le driver Video for Linux (V4L) est employé et là encore, il existe une incertitude temporelle sur le flux d’images envoyé par la caméra. La cohérence des images n’est pas claire non plus et il subsiste une incertitude sur le fait que les images acquises proviennent d’une image tampon ou non.

De ce fait, nous proposons une architecture logicielle qui permet de contrer ces aléas et d’obtenir un résultat logique sans avoir à se plonger dans l’ensemble du code source de l’OS (Figure 84).

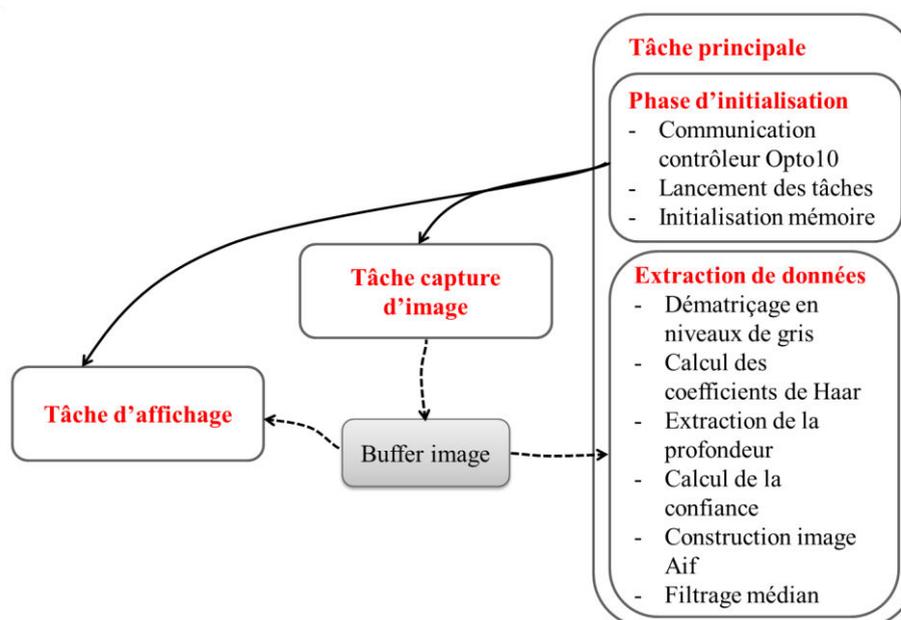


FIGURE 84 : ARCHITECTURE DE L’ALGORITHMIQUE EN TROIS THREADS (1 PRODUCTEUR 2 CONSOMMATEURS)

Une première phase d’initialisation ouvre la communication avec le contrôleur de lentille, prépare les différentes instances mémoires nécessaires au traitement puis lance deux tâches supplémentaires, exécutées en parallèle du traitement principal en utilisant le multithreading permis par le processeur multi cœurs. Celles-ci permettent la capture d’images en flux tendu et l’affichage à l’écran des images capturées en temps réel.

La tâche capture d’image permet de récupérer chaque image en provenance de la caméra afin d’éviter qu’elle ne soit stockée dans une mémoire tampon du driver qui pourrait restituer une image en retard vis-à-vis de la variation de focalisation. Cela provoquerait inévitablement des erreurs dans la carte de profondeur. Pour cela, un mécanisme à base de mutex et de variables de synchronisations permet de garantir la transmission d’une unique image par niveau de

focalisation au programme principal. Cette image est également transmise à la tâche d'affichage qui montre en temps réel l'évolution de focalisation.

La tâche d'extraction de données récupère l'image afin de l'analyser. Celle-ci est convertie en format GPU afin de subir l'ensemble du traitement sur le processeur graphique de la Jetson. Le transfert entre CPU et GPU constitue le plus souvent le goulot d'étranglement dans le calcul. Il est important de limiter le plus possible ces transferts afin d'éviter le gaspillage de temps de calcul. De ce fait, chaque étape de calcul doit être réfléchie afin de trouver le meilleur compromis calculatoire. Dans le cas présenté ici, nous sommes parvenus à ce que toutes les étapes soient réalisées sur GPU et donc les seuls transferts mémoires nécessaires sont en début et fin de processus d'extraction de profondeur.

L'algorithme est construit sur le modèle proposé dans la section architecture (section 2.7.1), si ce n'est que le dématricage est effectué non pas par binning mais par la fonction *bayer2gray*. Le traitement est effectué à la volée, avec une chaîne de calcul pour chacune des directions analysées. Pour chaque indice de focalisation, les mémoires de profondeur, d'images totalement focalisées et d'extrema des critères de netteté sont mises à jour. La réunification puis l'affichage des cartes de profondeur et de confiance sont réalisés en fin de processus. Avec OpenCV3, l'utilisation du GPU est fortement simplifiée et intuitive ce qui permet de retranscrire aisément l'architecture matérielle présentée dans le chapitre précédent. En revanche les traitements des deux directions s'effectuent séquentiellement et non en parallèle. Une considération également importante à prendre en compte est le fait que certaines fonctions CUDA, comme par exemple les filtrages maximum et minimum par bloc de pixels, supportent seulement des données 8bits. La caméra fournissant des pixels 12bits, l'algorithme conserve seulement les 8 bits de poids fort des coefficients de Haar pour effectuer ces opérations. Cela introduit un léger biais de mesure et ralentit quelque peu le calcul par l'ajout d'une fonction réalisant un décalage de 4 bits vers la droite, mais le gain en vitesse reste très important comparé à un passage des données au format CPU puis calcul sur CPU. Les fonctions CUDA ne permettent pas de faire simplement du ré-échantillonnage, de ce fait, les variations de dimensions des mémoires, par exemple le grossissement du masque utilisé pour la réunification des images totalement focalisées, est exécuté par la fonction *resize* avec une interpolation linéaire. Il est possible que ces adaptations nécessaires au portage affectent légèrement le résultat.

En fonction de la valeur d'une variable globale, le GPU peut également exécuter un filtrage médian sur la carte de profondeur. En revanche, étant donné que les deux images totalement focalisées sont mises à jour au fur et à mesure du traitement, et donc que les images du cube de multi-foyers ne sont pas conservées, la carte de profondeur filtrée ne peut être utilisée pour améliorer le rendu de l'image AiF comme proposé dans le chapitre 2 (section 2.6). Le filtrage médian permet donc seulement d'améliorer le rendu de profondeur. La dimension du noyau médian est ajustable avant compilation du code via la définition d'une macro.

4.4 Résultats obtenus

Pour expérimenter ce prototype nous avons utilisé la scène présentée Figure 85. Elle contient des mires planes contrastées placées à 25, 35 et 60cm de la lentille Opto10, trois figurines et trois étiquettes numérotées, placées respectivement à 20 cm, 30cm, 40 cm, 25 cm, 35 cm et 45 cm de cette même lentille. Le programme exécuté par la Jetson fait un balayage de foyers à 9 niveaux, de 20 à 60 centimètres par pas de 5 centimètres.

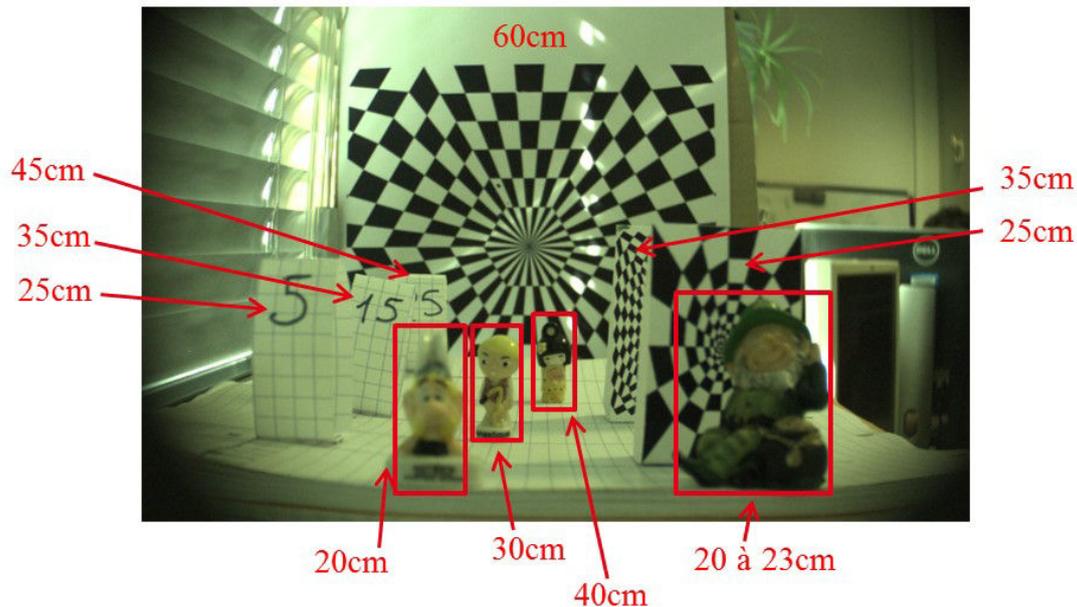


FIGURE 85 : SCENE POUR LES EVALUATIONS EXPERIMENTALES

4.4.1 Influence de la taille des fenêtres de traitement

Jusqu'à présent, nous avons empiriquement choisi d'utiliser des fenêtres de traitement comprenant 16×16 pixels. Le capteur d'image étant plus petit et moins résolu que celui d'un appareil reflex, il se peut qu'en fonction de l'application une autre taille de fenêtre soit plus adaptée. Ce prototype permet très simplement d'expérimenter de nouvelles valeurs par simple définition d'une constante macro. Ainsi, la Figure 86 montre des cartes de profondeur et de confiance pour 3 tailles de fenêtres de traitement, 5×5, 9×9 et 13×13.

Les cartes avec un traitement par fenêtre 5×5 présentent beaucoup de détails, presque trop pour une carte de profondeur qui, de par sa nature, est censée être relativement homogène. Comme cela a été montré dans le chapitre 2, ces noyaux ont tendance à dilater les transitions dans l'image. Cela peut être un désavantage pour des objets de dimensions proches des noyaux mais cela permet d'avoir des cartes de profondeur plus uniformes et plus représentatives de la réalité. En outre, l'expérience montre que les images AiF reconstruites à partir de fenêtres de traitement contenant 100 à 400 pixels contiennent moins d'artefacts.

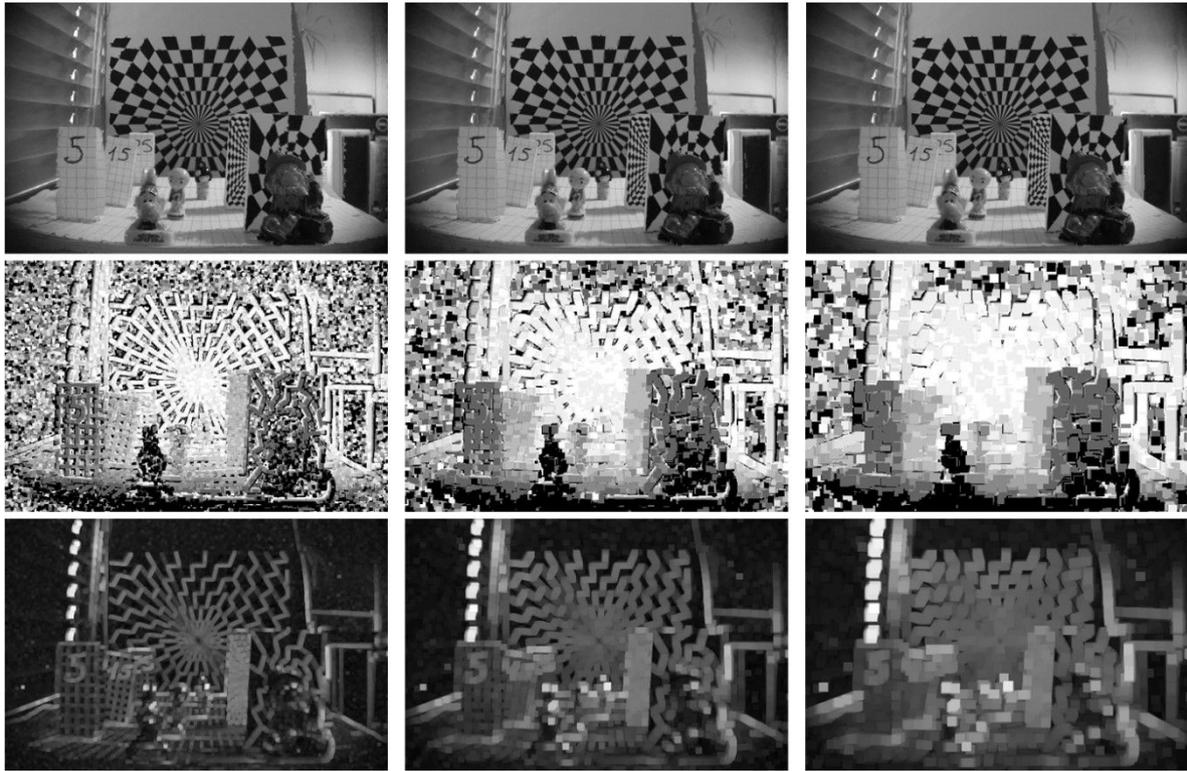


FIGURE 86 : IMAGES TOTALEMENT FOCALISEES ET SES CARTES DE PROFONDEURS NON FILTRES ET DE CONFIANCES ACQUISES SUR LE PROTOTYPE. DANS LE SENS DE LECTURE, LES FENETRES DE TRAITEMENT UTILISEES SONT DE TAILLE 5×5 9×9 ET 13×13

On note l'apparition de dégradé d'indice sur les contours de zones à fort contraste dans les cartes de profondeur, cela dénote une sensibilité au zoom de focalisation. Un filtrage médian, avec un noyau arbitrairement choisi carré de taille 11, permet d'atténuer cette sensibilité au zoom tout en améliorant le rendu des cartes de profondeur (Figure 87).

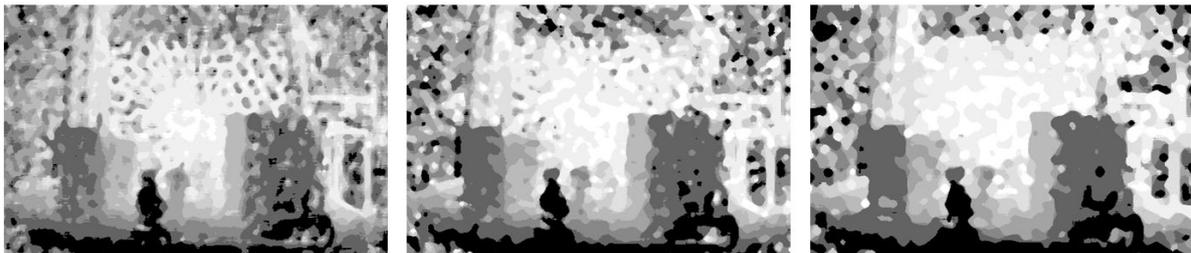


FIGURE 87 : CARTES DE PROFONDEURS LISSEES PAR UN FILTRAGE MEDIAN 11×11. DANS LE SENS DE LECTURE, LES FENETRES DE TRAITEMENT UTILISEES SONT DE TAILLE 5×5 9×9 ET 13×13

L'utilisation d'un tel noyau médian avec des fenêtres de taille 13×13 semble faire disparaître des niveaux de profondeur au niveau de la table. On privilégiera donc dans la suite des traitements par bloc 9×9.

Sans application visée, il est assez difficile de choisir la taille des fenêtres de traitement. Une étude itérative, basée sur la minimisation du PSNR, ou encore la maximisation du taux d'évaluation correcte proposé au chapitre 2 (section 2.5.1.1), avec une base de données

conséquente pourrait permettre de tendre vers des dimensions de fenêtre et de filtre médian optimums.

4.4.2 Mesure des temps de traitement

Ce prototype temps-réel permet d'évaluer le temps d'extraction des cartes de profondeur et de confiance avec une plateforme existante, et destinée à des applications embarquées. Plusieurs sondes ont été ajoutées au programme afin de récupérer l'heure courante à chaque itération de calcul puis de mesurer, en fin d'acquisition d'un cube multi-foyers, le temps d'exécution de certaines étapes. A cause de l'OS, ces temps sont relativement instables. De ce fait, les chiffres présentés, à titre indicatif, dans le Tableau 16 sont des moyennes de temps d'exécution. Pour chaque nouvelle acquisition d'image, le programme incrémente les temps d'exécution de chaque étape sondée et calcule leur moyenne globale à chaque itération de balayage de foyers. Nous avons choisi de montrer des temps de calcul pour des tailles très différentes de fenêtres glissantes et de noyaux médian.

Tableau 16 : Mesure des temps moyens de différentes étapes de calcul pour une extraction sur 9 niveaux de profondeur, la première ligne avec des fenêtres de traitement glissante 21×21 et un filtrage médian 31×31, la seconde avec des fenêtres de traitement glissantes 5×5 et un filtrage médian 11×11

Opération	Extraction AiF DM non filtrée et C	Traitement d'une image du cube multi foyers	Conversion GPU→CPU et dématricage en niveaux de gris	Traitement d'une direction	Réunion des données	Filtrage médian
Temps moyen 21-31	1.09s	84.1ms	3.7ms	42.7ms	3.7ms	131.5ms
Temps moyen 5-11	0.97s	62.25ms	4.8ms	29.5 ms	7.36ms	252.5ms

Globalement, lorsque la taille de la fenêtre de traitement augmente, le temps de traitement croît légèrement. Ce n'est pas nécessairement le cas avec l'augmentation de la taille du noyau médian, probablement pour des raisons d'optimisations algorithmiques lors de la compilation. Pour 9 niveaux de profondeur, ce prototype permet d'extraire des cartes de profondeurs et de confiances et leur image AiF associée à une cadence moyenne de 1 lot par seconde. L'évolution en fonction du nombre de niveaux est plutôt linéaire. En revanche, les modifications des paramètres d'extraction que sont la tailles des fenêtres et noyaux de calcul ne change pas drastiquement le temps total, celui-ci reste autour de la seconde. Cela confère

une certaine liberté d'adaptation des paramètres de calcul en fonction de la situation et de l'application.

Si on supprime la construction d'image AiF, on gagne 5ms par chaîne de traitement, soit 10ms en raison de leur portage en séquentiel, ce qui est peu comparé au temps de traitement total. Le temps de transfert de données CPU vers GPU prend en moyenne autant de temps que l'opération de dématricage, soit entre 1.5 et 2ms.

En raison de la cadence maximum de la caméra utilisée, on pouvait espérer sortir, au mieux, 6.7 couples de cartes de profondeur et de confiance par seconde pour 9 niveaux de profondeur, soit un couple toutes les 15ms. En supposant que nous puissions traiter les deux directions en parallèle, ce dispositif serait toujours trop lent pour suivre la cadence maximale de la caméra. Quand à extraire, à l'aide d'une caméra plus rapide, les données de profondeur et de confiance à une cadence de 30 couples par seconde, c'est tout simplement impossible avec un tel dispositif au regard du seul temps pris par le transfert de données CPU vers GPU et l'opération de dématricage.

Ces chiffres justifient l'intérêt d'une implémentation matérielle future. En effet, malgré une architecture particulièrement adaptée au traitement d'image embarqué, la Jetson ne permet pas de suivre la cadence caméra.

4.4.3 Evolution temporelle des images totalement focalisées et leurs cartes de profondeur et de confiance associées

Un aspect que nous n'avons pas encore abordé mais qui est néanmoins très important pour une application temps réel est l'évolution temporelle des données extraites pour une même scène. En outre, nous avons vu que le bruit des cartes de profondeur pouvait être traité par diverses méthodes à partir d'une seule image et d'a priori sur la scène (section 2.6 du chapitre 2) mais également par l'analyse de l'évolution des cartes au cours du temps. La solution inter-image évoquée dans l'annexe 2 a en partie été écartée à cause d'un manque de fiabilité temporelle. Nous avons donc observé l'évolution des cartes de profondeur et de confiance au cours du temps pour une scène statique. La Figure 88 montre dans le sens de la lecture 8 lots de données extraits les uns après les autres en une dizaine de secondes grâce à ce prototype expérimental.

Dans l'ensemble les données sont homogènes et fiables. Quelques valeurs semblent varier entre deux indices de profondeur mais pour les zones dont le taux de confiance est élevé, il s'agit d'indices précédent ou suivant dans le cube multi-foyers, ce qui représente une erreur minime qui peut être traitée par filtrage temporel. Ces erreurs sont dues à la fois à la profondeur de champ qui est légèrement plus grande que ce que nous avons calculé,

notamment à cause d'une réduction d'ouverture, des aberrations optiques, ainsi qu'aux variations causées par le bruit thermique intrinsèque du capteur.



FIGURE 88: EVOLUTION TEMPORELLE, DANS LE SENS DE LECTURE, DES IMAGES AIF, DES CARTES DE PROFONDEURS ET DE CONFIANCES POUR DES FENETRE DE TRAITEMENT 9×9 ET UNE FILTRE MEDIAN DE NOYAU 11×11

4.4.4 Influence de l'exposition

La méthode proposée dans cette thèse étant basée sur une analyse de contraste, l'exposition de la scène peut être un problème pour la mesure de profondeur. La Figure 89 présente une même scène et ses cartes de profondeur et de vraisemblance respectives pour trois expositions de la caméra. La première est très surexposée, la seconde est sous-exposée, et la dernière est correctement exposée.

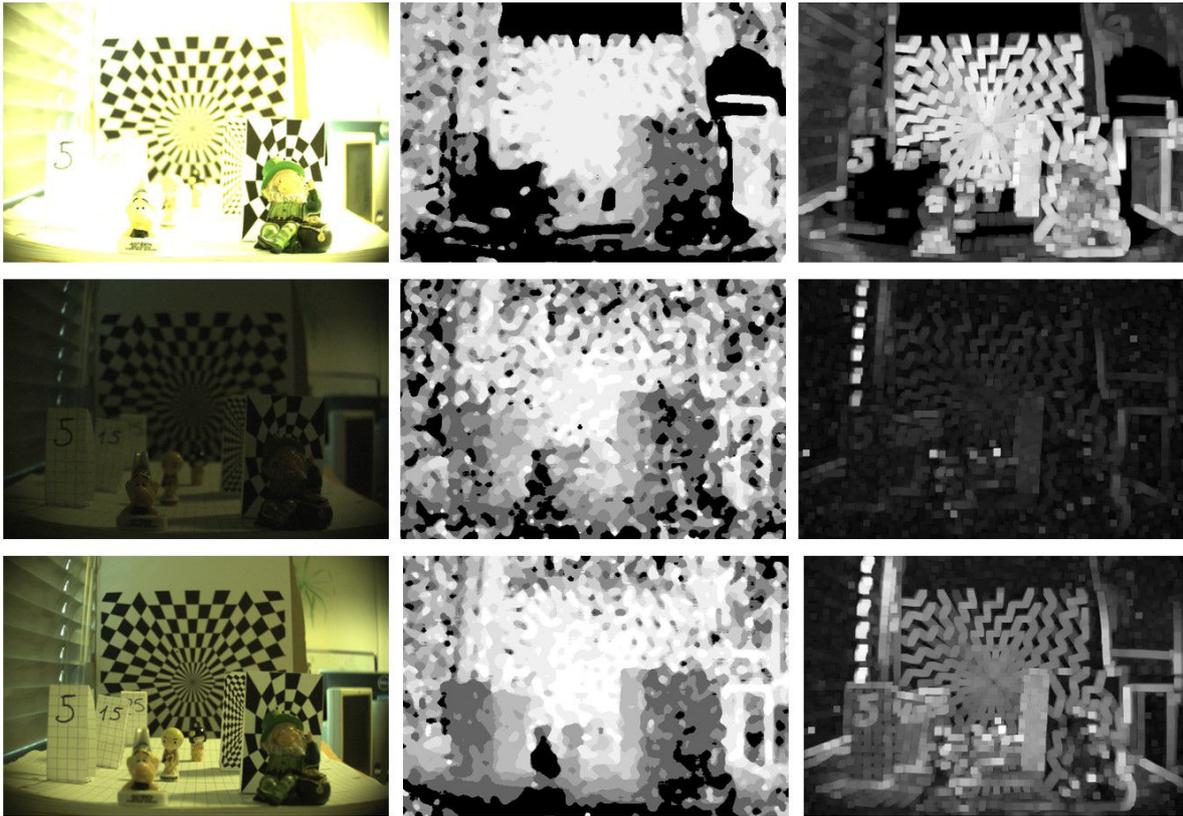


FIGURE 89 : SCENE AVEC TROIS EXPOSITIONS DIFFERENTES, EN HAUT, SUREXPOSITION AVEC UN REGLAGE A 1000, AU MILIEU, SOUS-EXPOSITION AVEC UN REGLAGE A 20, EN BAS EXPOSITION CORRECTE REGLEE A 60. DANS L'ORDRE DE LECTURE IMAGE COLOREE, CARTE DE PROFONDEUR ET CARTE DE CONFIANCE

Sur ces images, l'impact de l'exposition est très marqué. Une grande partie de la scène semble être à une distance de 20cm sur la carte de profondeur du haut alors que la carte du bas montre une évaluation plus fidèle à la réalité. A l'inverse, la sous-exposition rend très confuse la carte de profondeur. Notons que la carte de vraisemblance semble conforme à ce qu'on attend : les mauvaises estimations pour les cubes d'images incorrectement exposées sont dans l'ensemble peu pertinentes.

En raison de cette sensibilité, une méthode d'auto exposition telle que [109] devra à termes être combinée avec l'algorithme DfF, comme tâche supplémentaire par exemple, afin de corriger ce type d'incident et ainsi de rendre plus robuste ce système. C'est également important à prendre en compte pour une implémentation matérielle.

4.4.5 Sensibilité au mouvement

Le système fournissant un couple d'estimations par seconde, tout mouvement dans la scène devient impactant sur la mesure de profondeur et le rendu AiF. Nous avons enregistré un mouvement de bras de quelques secondes, celui-ci est annoté d'une flèche pointillée sur les cartes de profondeur et de confiance de la Figure 90. Ce mouvement se passe dans le plan

perpendiculaire à l'axe optique à environ 55cm de la lentille modulable, juste devant la cible contrastée la plus éloignée.

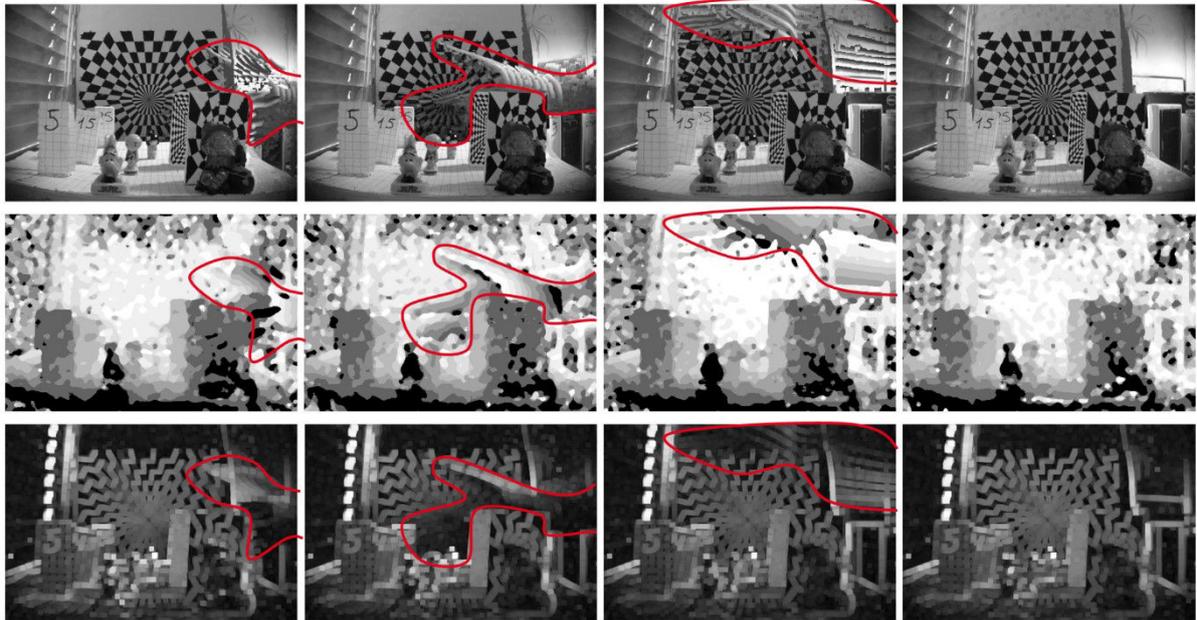


FIGURE 90 : IMPACTE DU MOUVEMENT D'UNE MAIN DANS LA SCENE SUR LA CARTE DE PROFONDEUR

Ce mouvement provoque l'apparition de dégradé d'indices dans les zones de la scène concernées, cela est particulièrement visible dans les cadres de la Figure 91. Sur les images AiF, on observe une redondance de l'objet en mouvement.

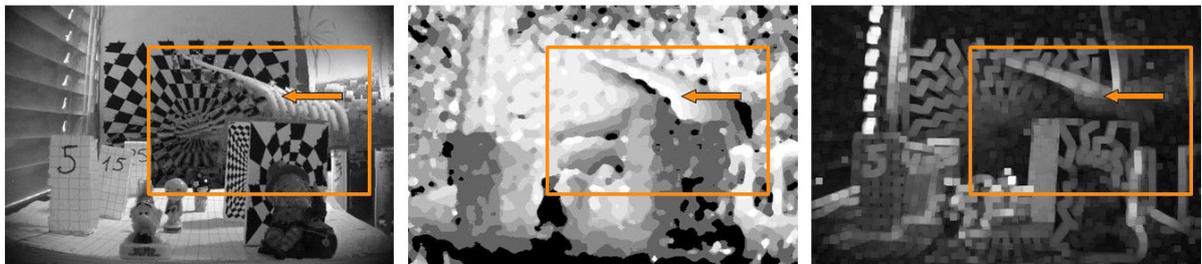


FIGURE 91 : MISE EN EVIDENCE DES EFFETS DU MOUVEMENT

La confiance est également affectée, les estimations correspondant au mouvement ressortent comme pertinentes ce qui s'avère être particulièrement dérangeant pour l'utilisation des données. Notons tout de même que le sens des dégradés de profondeur peut être exploité pour déduire la direction du mouvement comme l'indique les flèches sur ces images.

La seule solution pour augmenter la robustesse du système face aux variations de la scène est d'accélérer le traitement et pour cela tendre vers un portage matériel.

Cette sensibilité au mouvement dénote également l'importance de réussir à réaliser des lentilles rapides dans le cas d'utilisation de lentilles à cristaux liquides. En effet, celles-ci

deviendraient alors le facteur limitant de la chaîne d'acquisition et pourraient, si elles sont trop lentes, provoquer le même type de défaut sur les données capturées.

4.5 Conclusions sur le portage matériel

Le prototype expérimental présenté dans cette section nous a permis d'évaluer l'aspect temps réel d'un portage d'une solution DfF basée sur une analyse bidirectionnelle de contraste local.

Les images focalisées en tout point ainsi que leurs cartes de profondeur et de confiance sont extraites à une cadence de 1 lot par seconde, ce qui est intéressant pour des applications dans un environnement évoluant à faible vitesse. Cependant la forte sensibilité au mouvement confirme le besoin d'un portage vers une solution matérielle, en particulier pour une augmentation de la cadence de traitement et ainsi une plus grande robustesse face aux mouvements dans la scène.

Nous avons vu que l'algorithme proposé est fiable dans le temps ; il nécessite cependant l'ajout d'un module pour la gestion de l'exposition de la scène.

Malgré ses grandes capacités de calculs permises par la GPU, la Jetson ne parvient pas à réaliser l'acquisition de profondeur à la cadence vidéo. Cependant, en s'appuyant sur le temps d'exécution de certaines opérations standard comme le dématricage, on constate que ce type de plateforme reste bien moins efficace qu'un processeur d'image et nous conforte dans l'idée que la réalisation d'un ASIC, intégrant l'architecture que nous proposons, permettrait d'atteindre des vitesses d'acquisition plus concordantes avec des situations temps réel tout en réduisant drastiquement l'encombrement du capteur.

Conclusion et perspective

Au cours de cette thèse, nous avons exploré l'état de l'art des méthodes d'acquisition de la profondeur d'une scène et conclu qu'il existait peu de solutions pour les systèmes embarqués disposant de faibles ressources de calcul. Nous nous sommes alors orientés dans cette voie pour trouver une méthode d'estimation peu complexe adaptée à de nombreuses applications. Par leur compacité et leur faible complexité de calcul, les techniques dites Depth from Focus sont les plus à même de répondre aux contraintes de l'embarqué.

Nous avons proposé un nouvel algorithme DfF bidirectionnel qui permet d'extraire des cartes de profondeur suivant deux directions, horizontale et verticale, puis de les combiner grâce à un critère de pertinence fournissant une auto-évaluation de la mesure. L'analyse de la notion de netteté nous a permis de mettre au point un critère représentatif de celle-ci. Ce critère correspond au support de l'histogramme d'un cube d'images multi-foyers convoluées avec un filtre passe haut. Son maximum en fonction de la distance focale permet de déterminer les indices relatifs à la profondeur. L'autoévaluation des estimations proposée dans ces travaux est extraite à partir de la dynamique du critère de netteté correspondant à chaque pixel de profondeur. Cela permet de différencier simplement les zones uniformes des zones contrastées. Chaque estimation de profondeur est ainsi pondérée de sorte à fournir à l'application hôte un outil pour faciliter l'interprétation des données. En outre, les données recueillies permettent de construire des images focalisées en tous points de la scène, et ainsi de compenser l'utilisation d'optiques à faible profondeur de champ nécessaires pour garantir une discrimination correcte des niveaux de profondeur.

Trois versions de cet algorithme de complexités décroissantes ont été décrites et comparées quantitativement et qualitativement entre elles ainsi qu'à l'état de l'art. Il a été montré que l'approche intra-image basse résolution, basée sur les coefficients horizontaux et verticaux des ondelettes de Haar, est le meilleur compromis pour une intégration matérielle en termes de complexité, de qualité du résultat et d'encombrement mémoire. En particulier, on recommande l'utilisation de fenêtres de traitement non recouvrantes, qui permettent une réduction importante de l'empreinte mémoire, du nombre d'opérations par pixel, et ainsi, du coût et de la consommation du circuit. Elle peut finalement s'accommoder d'architectures de traitement très modestes. Cette solution a été déclinée en une architecture qui permettrait d'obtenir des résultats très corrects avec seulement l'équivalent de 20 additions par pixel sortant et une mémoire équivalente à 2.06 images RAW lorsque les images totalement focalisées sont reconstruites.

Pour le DfF, les optiques doivent être capables de rendre une qualité d'image suffisante, pour garantir une évolution remarquable du contraste local dans les images et ainsi faciliter

l'extraction de netteté ; être peu encombrantes, de façon à pouvoir s'adapter à des systèmes de faibles dimensions ; et être rapide afin de préserver l'invariance de la scène. Un dernier effet très souvent négligé dans ce type d'approche est le zoom parasite provoqué par la lentille responsable de la variation de la mise au point. Celui-ci a été étudié au cours de ces travaux. Il induit notamment une variation de la scène qui peut perturber la mesure de profondeur. Pour répondre à ces quatre points, nous avons comparé des lentilles liquides industrielles et des lentilles modales expérimentales à cristaux liquides nématiques réalisées au cours de cette thèse. Nous avons présenté les paramètres dimensionnant de ces dernières et fait des choix, d'épaisseur et de nature des couches de cristaux liquides, et de diamètres de pupille, qui ont mené à la réalisation de 7 lentilles différentes. Nous avons montré que les lentilles modales à cristaux liquides nématiques sont sujettes à des effets de diffusion lumineuse qui impactent la qualité des images et de ce fait l'analyse de contraste. En revanche, elles engendrent très peu de zoom lors d'un balayage de foyers ce qui s'avère être un net avantage pour améliorer l'invariance de la scène. D'un point de vue optique, les résultats que nous avons eus avec les lentilles à cristaux liquides de pupille 4-5 mm sont encourageants et montrent que celles-ci peuvent être intéressantes pour des applications d'imagerie. De plus, des expérimentations récentes avec de nouveaux cristaux liquides ont montré qu'on pouvait potentiellement accélérer les vitesses de focalisation d'un facteur 10. L'utilisation de tels cristaux liquides combinée à une amélioration des étapes de réalisation des lentilles et une meilleure compréhension des phénomènes physiques au sein de la couche de cristal liquide, à travers une modélisation électrique complète des lentilles, permettrait d'améliorer grandement l'efficacité de ces lentilles modales pour une application DfF. A cause du grand nombre de paramètres matériaux et de fabrication à prendre en compte, un tel modèle est complexe à élaborer et constitue une étude à mener à l'issue de ces travaux. Egalement, la dégradation temporelle observée au cours de cette thèse pour des lentilles à cristaux liquides expérimentales mérite d'être étudiée en profondeur à l'avenir ; cela permettrait notamment de garantir la répétabilité des résultats, ce qui est primordiale pour des applications effectuant une correction temps réel de la focalisation de la lentille sans boucle de retour. Une lentille de 4mm de diamètre, avec le cristal liquide MLC-2140 constitue cependant un bon compromis pour une extraction de profondeur dans un proche environnement statique.

En raison d'une qualité d'image supérieure et d'un temps de variation de focalisation plus rapide, une lentille liquide a été préférée pour réaliser un prototype DfF embarqué à partir d'une carte de développement NVIDIA équipée d'un GPU. Ce prototype expérimental nous a permis d'évaluer l'aspect temps réel du portage de l'algorithme intra-image développé pendant cette thèse. Une image totalement focalisée ainsi que ses cartes de profondeur et de confiance ont pu être extraites à une cadence d'un lot par seconde. Nous avons montré suite à de multiples estimations d'une même scène que cette méthode DfF était stable dans le temps, mais reste sensible aux variations de l'environnement, mouvement et exposition, pendant l'acquisition d'un cube d'images multi-foyers. Nous avons constaté que ce prototype

est bien moins efficace qu'un processeur d'images ; l'idée de la réalisation d'un ASIC intégrant l'architecture que nous proposons est donc cohérente. Cela permettrait notamment d'accélérer le traitement, de réduire drastiquement l'encombrement du capteur et d'améliorer la robustesse face au mouvement de la scène.

De futurs travaux pourraient permettre d'effectuer un portage de l'algorithme sur un ASIC intégrant un capteur d'image rapide, les éléments de synchronisation de la mise au point et de l'acquisition d'image, ainsi qu'une gestion de l'exposition de la scène. De plus, l'utilisation de lentilles à cristaux liquides très rapide permettrait de proposer *in fine* une solution d'extraction de profondeur, de pertinence et d'images totalement focalisées très compacte.

Références

- [1] Emberger Simon, Alacoque Laurent, Dupret Antoine, Lecat-Mathieu de Boissac Capucine, de Bougrenet de la Tocnaye Jean Louis, Fraval Nicolas, "Liquid Crystal lens characterization for integrated depth sensing and all in focus imaging application," in *Electronic Imaging, Mobile Devices and Multimedia: Enabling Technologies, Algorithms, and Applications*, San Francisco, 2017, pp. 34-39
- [2] Simon Emberger, Laurent Alacoque, Antoine Dupret, Jean Louis de Bougrenet de la Tocnaye, "Low Complexity Depth Map Extraction and All-in-Focus Rendering for Close-to-the-Pixel Embedded Platforms," in *International Conference on Distributed Smart Cameras*, Stanford, 2017, pp. 29-34
- [3] S. Emberger, L. Alacoque, A. Dupret, G. Sicard, J.-L. de Bougrenet de La Tocnaye, "A near pixel Depth from Focus architecture for video rate depth estimation," in *Electronic Imaging, Image sensor*
- [4] S. Emberger, L. Alacoque, A. Dupret, N. Fraval, J.-L. de Bougrenet de La Tocnaye, "Evaluation of the key design parameters of Liquid crystal tunable lenses for Depth-from-Focus algorithm," *Applied Optics*, vol. 57, no. 1, pp. 85-91, 2018.
- [5] Manap N.A., Darsono A.M., Yusof Z.M. Aboali M., "Review on three dimensional (3-D) acquisition and range imaging techniques," *International Journal of Applied Engineering Research*, vol. 12, no. 10, pp. 2409-2421, 2017.
- [6] Francois Blais, Luc Cournoyer, Guy Godin and Marc Rioux J-Angelo Beraldin, "Active 3D sensing," , 2000
- [7] Thierry M. Bosch, Marc Lescure, Risto A. Myllylae, Marc Rioux Markus-Christian Amann, "Laser ranging: a critical review of unusual techniques for distance measurement," *Optical Engineering*, vol. 40, no. 1, 2001.
- [8] Stjepan Jecić Nenad Drvar, "The assessment of structured light and laser scanning methods in 3D shape measurements," in *4 th International Congress of Croatian Society of Mechanics*, 2003
- [9] A. Trucco and S. Curletto, "Extraction of 3D information from sonar image sequences," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 33, no. 4, pp. 687-699, 2003.

- [10] B. Kamgar-Parsi, E. Belcher and O. Engelsen L. Rosenblum, "Acoustic imaging: the reconstruction of underwater objects," in *Visualization*, San Diego, 1991
- [11] Jackie L. Whittaker and Deydre S. Teyhen and James M. Elliott and Katy Cook and Helene M. Langevin and Haldis H. Dahl and Maria Stoke, "Rehabilitative Ultrasound Imaging: Understanding the Technology and Its Applications," *Journal of Orthopaedic & Sports Physical Therapy*, vol. 37, no. 8, pp. 434-449, 2007.
- [12] Vikas Yadav, Shikha Sharma Rajni Rawal, "RADAR- A BRIEF STUDY," *International Journal of Innovative Research and Technology*, vol. 1, no. 12, pp. 1017 - 1020, 2015.
- [13] A. N. Catapang and M. Ramos, "Obstacle detection using a 2D LIDAR system for an Autonomous Vehicle," in *6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Batu Ferringhi, 2016
- [14] G. Alenya and C. Torras S. Foix, "Lock-in Time-of-Flight (ToF) Cameras: A Survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917-1926, 2011.
- [15] H. Yalcin and C. Bamji, S. B. Gokturk, "A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions," in *Conference on Computer Vision and Pattern Recognition Workshop*, 2004
- [16] Larry Li, Texas Instruments, "Time-of-Flight Camera – An Introduction," 2014.
- [17] M. Watanabe and M. Noguchi S. K. Nayar, "Real-time focus range sensor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 12, pp. 1186-1198, 1996.
- [18] Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light.," *Computer Visio and Pattern Recognition (CVPR)*, 2003.
- [19] N. H. B. Ismail and S. N. B. Basah, "The applications of Microsoft Kinect for human motion capture and analysis : A review," in *2nd International Conference on Biomedical Engineering*, Penang, 2015
- [20] R. Tron and R. Vidal, "Distributed 3-D Localization of Camera Sensor Networks From 2-D Image Measurements," *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3325-3340, 2014.
- [21] S. Bourgeois, F. Dekeyser, E. Royer and M. Dhome P. Lothe, "Towards geographical referencing of monocular SLAM reconstruction using 3D city models: Application to

- real-time accurate vision-based localization," in *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 2009
- [22] M. R. Petraglia and A. Petraglia L. C. Silva, "A robust method for camera calibration and 3-D reconstruction for stereo vision systems," in *12th European Signal Processing Conference*, Vienna, 2004
- [23] U. R. Dhond and J. K. Aggarwal, "Structure from stereo-a review," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1489-1510, 1989.
- [24] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision (IJCV)*, 2002.
- [25] M. B. Nikiforov, A. B. Rusakov and V. S. Gurov Y. R. Muratov, "Estimation of distance to objects by stereovision," in *2015 4th Mediterranean Conference on Embedded Computing (MECO)*, Budva, 2015, pp. 155-158
- [26] Z. Li et al., "3.7 A 1920×1080 30fps 2.3TOPS/W stereo-depth processor for robust autonomous navigation," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, San Francisco, CA, 2017, pp. 62-63
- [27] M. Hariyama and M. Kameyama, "VLSI processor for reliable stereo matching based on window-parallel logic-in-memory architecture," in *2004 Symposium on VLSI Circuits. Digest of Technical Papers*, 2004
- [28] G. Camellini et al, "3DV — An embedded, dense stereovision-based depth mapping system," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, 2014
- [29] J. Y. Mori, C. H. Llanos and E. Fortaleza C. Sánchez-Ferreira, "Development of a stereo vision measurement architecture for an underwater robot," in *2013 IEEE 4th Latin American Symposium on Circuits and Systems (LASCAS)*, Cusco, 2013
- [30] Wilfried Glastre, "Imagerie plénoptique à travers des milieux complexes par synthèse d'ouverture optique," Grenoble, Thèse 2013.
- [31] J. Restrepo and L. Mignard-Debise I. Ihrke, "Principles of Light Field Imaging: Briefly revisiting 25 years of research," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 59-69, 2016.
- [32] G. Lippmann, *Epreuves reversibles. photographies integrales*, 1908.

- [33] Edward H. Adelson and James R. Bergen, "The Plenoptic Function and the Elements of Early Vision," in *Computational Models of Visual Processing*.: MIT Press, 1991.
- [34] Edward H. Adelson and John Y.A. Wang, "Single lens stereo with a plenoptic camera.," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 1992.
- [35] Marc Levoy and Pat Hanrahan, "Light Field Rendering," *ACM SIGGRAPH '96*, 1996.
- [36] Ren Ng, "Digital Light Field Photography," Stanford university, Palo Alto, Thèse 2006.
- [37] Ren Ng, "Fourier Slice Photography," *ACM Transactions on Graphics*, 2005.
- [38] Ramesh Raskar, Amit Agrawal, Ankit Mohan, Jack Tumblin Ashok Veeraraghavan, "Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing," *ACM SIGGRAPH 2007*, 2007.
- [39] Christian Perway and Lennart Wietzke, "Single lens 3d-camera with extended depth of field," 2012.
- [40] Raytrix, "Raytrix light field camera," 2012.
- [41] Jacques Duparre Andrew McMahon Gabriel Molina Priyam Chatterjee Robert Mullis Kartik Venkataraman, Dan Lelescu. Pelican Imaging Corporation, Picam : An ultra-thin high performance monolithic camera array., 2013.
- [42] Todor Georgeiv and Chintan Intwala, "Light Field Camera Design for Integral View Photography," San Jose, 2003.
- [43] Rob Fergus, Frédo Durand, William T. Freeman Anat Levin, "Image and depth from a conventional camera with a coded aperture," *ACM SIGGRAPH*, vol. 26, no. 3, 2007.
- [44] Paolo Favaro Anita Sellent, "Optimized aperture shapes for depth estimation," *In Pattern Recognition Letters*, vol. 40, pp. 96-103, 2014.
- [45] Chun Wang, "Design and analysis of coded aperture for 3D scene sensing," Tampere University of Technology, Thèse de master 2015.
- [46] Fuji, "Evolution into the x-trans Cmos ii sensor with phase detection pixels af speed achieving the level of single lens reflex cameras".
- [47] Isao Hirota, Isao Ohdaira, Masao Matsumura, Atsushi Morimitsu, Hiroaki Takahashi,Toshio Yamazaki, Hideki Oyaizu, Yalcin Incesu, Muhammad Atif, Yoshikazu

- Nitta Sozo Yokogawa, "A 4M pixel full-PDAF CMOS image sensor with 1.58 μ m 2 \times 1 On-Chip Micro-Split-Lens technology," Sony Corporation,.
- [48] Paweł Wachel Przemysław Śliwiński, "A Simple Model for On-Sensor Phase-Detection Autofocusing Algorithm," *Journal of Computer and Communications*, vol. 1, pp. 11-17, 2013.
- [49] M.Johnson Y.Wada H.Tsuboi T.Ono H.Takada-K.Togo T.Kishi H.Takashi T.Ichikawa S.Inoue M.Kobayashi, "A low noise and high sensitivity image sensor with imaging and phase-difference detection af in all pixels," 2015.
- [50] Canon, Dual pixel cmos autofocus.
- [51] Sangwoo Park, Jieun Jo, and Joonki Paik Jinbeum Jang, "Depth map generation using a single image sensor with phase masks," *Opt. Express*, vol. 24, no. 12, 2016.
- [52] W. Zhang and W. K. Cham, "Single-Image Refocusing and Defocusing," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 873-882, 2012.
- [53] Chunping Hou Chang Tang and Zhanjie Song, "Depth recovery and refinement from a single image using defocus cues," *Journal of Modern Optics*, 2014.
- [54] Chunping Hou, and Zhanjie Song Chang Tang, "Defocus map estimation from a single image via spectrum contrast," *Optics Letters*, vol. 38, no. 10, pp. 1706-1708, 2013.
- [55] J. Canny, "A Computational Approach to Edge Detection," *EEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-698, 1986.
- [56] S. Cohen, S. Schiller and P. Milanfar X. Zhu, "Estimating Spatially Varying Defocus Blur From A Single Image," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4879-4891, 2013.
- [57] F. Champagnat, G. Le Besnerais and J. Idier P. Trouvé, "Single image local blur identification," in *18th IEEE International Conference on Image Processing*, Brussels, 2011
- [58] Chao Liu and L. Christopher, "Three dimensional moving pictures with a single imager and microfluidic lens".
- [59] Yusuke and Sasaki, Takayuki and Mishima, Nao and Mita, Takeshi Moriuchi, "Depth from Asymmetric Defocus using Color-Filtered Aperture," *SID Symposium Digest of Technical Papers*, vol. 48, pp. 325-328, 2017.

- [60] J. B. Kulkarni and C. M. Sheelarani, "Generation of Depth Map Based on Depth from Focus: A Survey," in *2015 International Conference on Computing Communication Control and Automation*, Pune, 2015
- [61] Shree K. Nayar, "Shape from focus," 1989.
- [62] L. Harkness, "Chameleons use accommodation cues to judge distance," *Nature*, vol. 267, pp. 346–349, 1977.
- [63] N. Mitsumoto and N. Kawahara T. Kaneko, "A new smart vision system using a quick-response dynamic focusing lens," in *IEEE Thirteenth Annual International Conference on Micro Electro Mechanical Systems*, Miyazaki, 2000
- [64] Lorenz K. Müller, Stephen J. Carey and Piotr Dudek Julien N. P. Martel, "High-speed Depth from Focus on a Programmable," in *IEEE International symposium on circuits & systems (ISCASS)*, Baltimore, 2017
- [65] T. Yoshida, H. Kashiyama, and T. Hamamoto A. Yokota, "High speed depth estimation by smart image sensor system," 2004.
- [66] T. Yoshida, H. Kashiyama, and T. Hamamoto A. Yokota, "High-speed sensing system for depth estimation based on depth-from-focus by using smart imager," 2005.
- [67] Supasorn and Hernandez, Carlos and Seitz, Steven M. Suwajanakorn, "Depth From Focus With Your Mobile Phone," 2015.
- [68] H. Fujiwara, "Passive range sensor using multi-focus images," in *intelligent Robots and Systems '96, IROS 96*, Osaka, 1996
- [69] D. Lischinski, and Y. Weiss A. Levin, "A closed-form solution to natural image matting".
- [70] M. Benetti, A. Peizerat, A. Dupret and P. Vandergheynst W. Guicquero, "On compensating unknown pixel behaviors for image sensors with embedded processing," in *48th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, 2014
- [71] Yi-Chin Wu, Chih-Hung Liu, Wei-Chih Sun and Y. C. Chen Y. C. Chen, "Depth map generation based on depth from focus," in *International Conference on Electronic Devices, Systems and Applications*, Kuala Lumpur, 2010
- [72] CEA, "Integrated three-dimensional vision sensor".

- [73] L. K. Müller, S. J. Carey and P. Dudek J. N. P. Martel, "Parallel HDR tone mapping and auto-focus on a cellular processor array vision chip," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, 2016
- [74] Alexey Lopich, David R.W. Barr, Bin Wang, Piotr Dudek Stephen J. Carey, "A 100,000 fps Vision Sensor with Embedded 535GOPS/W 256x256 SIMD Processor Array," , 2013
- [75] A. Alfarouq, S. Sukaridhoto and D. Pramadihanto S. A. Waskitho, "FloW vision: Depth image enhancement by combining stereo RGB-depth sensor," in *International Conference on Knowledge Creation and Intelligent Computing (KCIC)*, 2016
- [76] Zhaoqing Liu, Minshan Jiang, Min Chang Xuedian Zhang, "Fast and accurate auto-focusing algorithm based," *Opt Express*, vol. 22, no. 15, pp. 31237-47, 2014.
- [77] Ardiansyah Alfarouq, Suryo Aji Waskitho, Sritrusta Sukaridhoto Dadet Pramadihanto, "Merging of Depth Image Between Stereo Camera and Structure Sensor on Robot "FloW" Vision ," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 7, no. 3, pp. 1014-1025, 2017.
- [78] Myungjin Cho, Yi-Pai Huang, and Bahram Javidi Chih-Wei Chen, "Three-dimensional imaging with axially distributed sensing using electronically controlled liquid crystal lens," *Opt. Letters*, vol. 37, no. 19, pp. 4125-4127, 2012.
- [79] Inès Sahnoun, Anaïs Sezgin, Altan Soran William Ren, "Les lentilles liquides,".
- [80] Bruno Berge, Franck Fillit, Hilario Gaton, Martin Guillet, Olivier Jacques-Sermet, Frédéric Laune, Julien Legrand, Mathieu Maillard, Nicolas Tallaron Eric Simon, "Optical design rules of a camera module with a liquid lens and principle of command for AF and OIS functions," in *Proc. SPIE 7849, Optical Design and Testing IV*, 2010
- [81] Yoseph Bar-Cohen, *Electroactive Polymer (EAP) Actuators as Artificial Muscles: Reality, Potential, and Challenges.*, 2004.
- [82] M. Büeler, C. Grätzel, M. Aschwanden M. Blum, "Compact optical design solutions using focus tunable lenses," in *SPIE Optical Design and Engineering IV*, 2012
- [83] Qiong-Hua Wang and Wei Jiang Lei Li, "Liquid lens with double tunable surfaces for large power tunability and improved optical performance," *Journal of Optics*, vol. 13, no. 11, 2011.

- [84] Michael J. Stephen and Joseph P. Straley, "Physics of liquid crystals," *Review of Modern Physics*, vol. 46, no. 4, pp. 617-704, 1974.
- [85] Nicolas Fraval, *Composants à cristaux liquides nématiques adressés par électrodes résistives en polymère conducteur et applications*, 2010.
- [86] Vincent Carrat, "Correction dynamique d'un SLM pour une holographie de haute fidélité. Réalisation d'un MOT-2D pour l'application de modes de Laguerre-Gauss," 2013.
- [87] N. J. Powell, A. Purvis, M. G. Clark G. Williams, "Electrically Controllable Liquid Crystal Fresnel Lens," in *33rd Annual Technical Symposium*, San Diego, 1989
- [88] Susumu Sato, "Liquid-Crystal Lens-Cells with Variable Focal Length," *The Japan Society of Applied Physics Japanese Journal of Applied Physics*, vol. 18, no. 9, 1979.
- [89] Ming-Syuan Chen, and Yi-Hsin Lin Hung-Chun Lin, "A review of electrically tunable focusing Liquid Crystal lenses," *TEEM*, vol. 12, no. 6, pp. 234-240, 2011.
- [90] S Sato, "Applications of Liquid Crystals to Variable-Focusing Lenses," *OPT REV*, vol. 6, p. 471, 1999.
- [91] Bin Wang, and Susumu Sato Mao Ye, "Liquid-crystal lens with a focal length that is variable in a wide range," *Applied Optics*, vol. 43, no. 35, pp. 6407-6412, 2004.
- [92] Paul C.-P. Chao, and Chieh-Wen Hsueh Yung-Yuan Kao, "A new low-voltage-driven GRIN liquid crystal lens with multiple ring electrodes in unequal widths," *Opt. Express*, vol. 18, no. 18, pp. 18506-18518, 2010.
- [93] Nabeel A. Riza and Michael C. DeJule, "Three-terminal adaptive nematic liquid-crystal lens device," *Opt. Letters*, vol. 19, no. 14, pp. 1013-1015, 1994.
- [94] VincentLaud, "Twisted-nematic liquid-crystal pixelated active lens," *Optics Communications*, vol. 153, no. 1-3, pp. 134-152, 1998.
- [95] Jean Louis de Bougrenet de la Tocnaye Nicolas Fraval, "Low aberrations symmetrical adaptive modal liquid crystal lens with short focal lengths," *Applied Optics*, vol. 49, no. 15, pp. 2778-278, 2009.
- [96] Michael C. DeJule Nabeel A. Riza, "Three-terminal adaptive nematic liquid-crystal lens device," *Optics Letters*, vol. 9, no. 14, pp. 1013-1015, 1994.

- [97] V. N Belopukhov, F. L Vladimirov, G. V Vdovin, G. D Love et A. F Naumov M. Yu Loktev, "Wave front control systems based on modal liquid crystal lenses," *Review of Scientific*, vol. 71, no. 9, pp. 3290-3297, 2000.
- [98] Loktev MY, Guralnik IR, Vdovin G. Naumov AF, "Liquid-crystal adaptive lenses with modal control.," *Optic Letters*, vol. 23, no. 13, pp. 992-994, 1998.
- [99] Philip B. Morgan, John H. Clamp, and Helen F Gleeson Harry E. Milton, "Electronic liquid crystal contact lenses for the correction of presbyopia," in *Opt. Express* 22, 2014
- [100] Sarabjot Kaur, Harry E. Milton, Devesh Mistry, James Bailey, Philip B. Morgan, J. Cliff Jones, and Helen F. Gleeson Ishtiaque M. Syed, "Novel switching mode in a vertically aligned liquid crystal contact lens," in *Opt. Express* 23, 2015
- [101] Y.-J. Kim, H. Milton, D. Mistry, I. M. Syed, J. Bailey, K. S. Novoselov, J. C. Jones, P. B. Morgan, J. Clamp, and H. F. Gleeson S. Kaur, "Graphene electrodes for adaptive liquid crystal contact lenses," in *Opt. Express* 24, 2016
- [102] T. Yoshida, H. Kashiyama, and T. Hamamoto A. Yokota, "High speed depth estimation by smart image sensor system," 2004.
- [103] E. Krotkov, "Focusing," *International Journal of Computer Vision*, vol. 1, pp. 223-237, 1987.
- [104] T. Yoshida, H. Kashiyama, and T. Hamamoto A. Yokota, "High speed depth estimation by smart image sensor system ," 2004.
- [105] Lorenz K. Müller, Stephen J. Carey and Piotr Dudek Julien N. P. Martel, "High-speed Depth from Focus on a Programmable," in *IEEE International symposium on circuits & systems (ISCAS)*, Baltimore, 2017
- [106] Nicolas Fraval, "Composants à cristaux liquides nématiques adressés par électrodes résistives en polymère conducteur et applications," Brest, Thèse 2010.
- [107] Shin-Tson Wu and Khoon-Cheng Lim, "Absorption and scattering measurements of nematic liquid crystals," *Applied Optics*, vol. 26, no. 9, pp. 1722-1727, 1987.
- [108] Jetson TK1. [Online]. [https://elinux.org/Jetson TK1](https://elinux.org/Jetson_TK1)
- [109] Se-Hwan Yun, Suki Kim Quoc Kien Vuong, "A New Auto Exposure and Auto White-Balance Algorithm to Detect High Dynamic Range Conditions Using CMOS Technology ," in *World Congress on Engineering and Computer Science*, San Francisco, 2008

Annexes

Annexe 1 : Etude du cercle de confusion

Afin de mieux comprendre la mécanique du flou, nous avons mené une étude optique géométrique sur l'évolution théorique du diamètre d'une tâche lumineuse, aussi appelé cercle de confusion, image d'une source ponctuelle.

Cette annexe a pour but d'étudier comment le diamètre de la tâche lumineuse dans le plan du capteur évolue en fonction de la position de la source dans le plan objet. Cette démonstration est construite sur les lois de Descartes et le théorème de Thales. Cela implique de faire l'hypothèse que les angles considérés sont petits et que la lentille est mince (condition de Gauss).

Notons L l'ouverture de la lentille, d_c la distance capteur/lentille, d_o la distance objet/lentille, f la distance focale et C le diamètre de la tâche lumineuse. Soit D_c le conjugué de d_c par la relation de Descartes. (Figure A1-1). Nous nous plaçons dans le cas d'une mono lentille où $d_o \in [f + \infty[$, c'est-à-dire lorsque l'image est réelle. Le cas d'image virtuelle n'est pas traité ici puisque très peu, voir aucun, appareil photo ne le permet. D'une part cela n'a pas vraiment d'intérêt d'un point de vue photographique, de plus les distances focales équivalentes de leurs optiques sont très courtes (inférieures à 2cm dans la plupart des cas), ce qui correspond à des distances pour lesquelles la luminosité de la scène serait très réduite par la proximité de l'appareil si $d_o < f$.

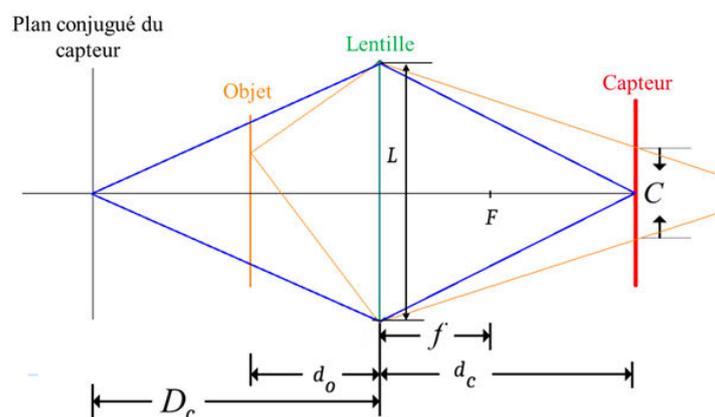


FIGURE A1-1 : SCHEMA RECAPITULATIF DES NOTATIONS

Cas d'objets lointains par rapport au plan conjugué du capteur :

Dans le cas d'objets plus lointains que le plan conjugué du capteur, d_o dans $[D_c, +\infty[$, le point E' image de E₁ se trouve entre la lentille et le capteur (Figure A1-2).

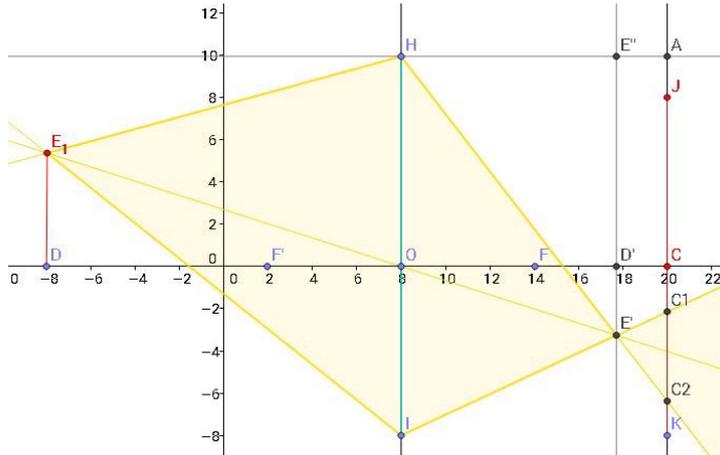


FIGURE A1-2 : OBJET FOCALISE DEVANT LE CAPTEUR (SEGMENT (KJ) ROUGE)

Dans les conditions de Gauss on peut écrire :

$$\frac{1}{OD'} - \frac{1}{OD} = \frac{1}{f} \Leftrightarrow \frac{1}{OD'} = \frac{1}{OD} + \frac{1}{f} = \frac{1}{f} - \frac{1}{d_o} \quad (\text{A1-0})$$

Le plan conjugué D_c est lié à la distance du capteur par l'équation (A1-0), il vient :

$$\frac{1}{d_c} + \frac{1}{D_c} = \frac{1}{f} \Leftrightarrow \frac{1}{D_c} = \frac{1}{f} - \frac{1}{d_c} \quad (\text{A1-1})$$

En utilisant le théorème de Thalès pour les triangles $E'HI$ et $E'C1C2$ dont les bases sont parallèles, on a :

$$\frac{C_{far}}{L} = \frac{C1C2}{HI} = \frac{C2E'}{E'H} \quad (\text{A1-2})$$

En se plaçant dans le triangle $HAC2$ avec $E'E''$ parallèle à $AC2$ par construction, par application du théorème de Thalès, on a :

$$\frac{d_c}{OD'} = \frac{HA}{HE''} = \frac{HC2}{HE'} = \frac{HE' + E'C2}{HE'} = 1 + \frac{E'C2}{HE'} \quad (\text{A1-3})$$

De (A1-2) et (A1-3) :

$$\frac{C_{far}}{L} = \frac{d_c}{OD'} - 1 \quad (\text{A1-4})$$

En remplaçant $\frac{1}{OD'}$ via l'équation (A1-0), on obtient:

$$C_{far} = L \cdot d_c \cdot \left(\frac{1}{f} - \frac{1}{d_o} - \frac{1}{d_c} \right) \quad (\text{A1-5})$$

Soit, en remplaçant par (A1-1)

$$C_{far} = L \cdot d_c \cdot \left(\frac{1}{D_c} - \frac{1}{d_o} \right) \quad \text{pour } D_c < d_o \quad (\text{A1-6})$$

Cas d'objet proche par rapport au plan conjugué du capteur :

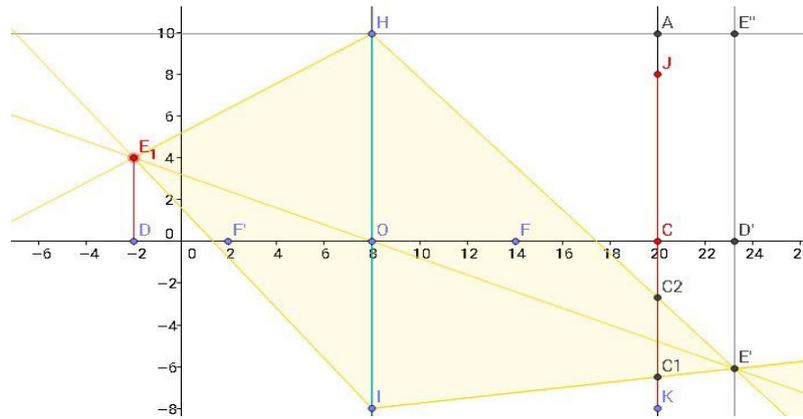


FIGURE A1-3 : OBJET FOCALISE DERRIERE LE CAPTEUR (ROUGE)

Dans le cas d'objets plus près que le plan conjugué du capteur), d_o dans $[f, D_c]$, le point focal E' se trouve derrière le capteur (Figure A1-3). En appliquant le même raisonnement que précédemment, on a :

$$\frac{c_{near}}{L} = \frac{C1C2}{HI} = \frac{C2E'}{E'H} \quad (A1-7)$$

En se plaçant dans le triangle $HE'E''$ avec $E'E''$ parallèle à $AC2$, d'après le théorème de Thalès, on a :

$$\frac{d_c}{OD'} = \frac{HA}{HE''} = \frac{HC2}{HE'} = \frac{HE' - E'C2}{HE'} = 1 - \frac{E'C2}{HE'} \quad (A1-8)$$

De (A1-7) et (A1-8) :

$$\frac{c_{near}}{L} = 1 - \frac{d_c}{OD'} \quad (A1-9)$$

Il vient :

$$C_{near} = L \cdot d_c \cdot \left(\frac{1}{d_o} - \frac{1}{f} + \frac{1}{d_c} \right) \quad (A1-10)$$

Et donc :

$$C_{near} = L \cdot d_c \cdot \left(\frac{1}{d_o} - \frac{1}{D_c} \right) \quad \text{pour } f < d_o < D_c \quad (A1-11)$$

Cas général :

Nous avons obtenu deux formules (A1-4) et (A1-8) pour caractériser le diamètre de la tache lumineuse dans le plan capteur. On en déduit le diamètre de cette tache (A1-9) valable sur l'ensemble du domaine de d_o pour lequel l'image est réelle.

$$C = L \cdot d_c \cdot \left| \frac{1}{d_o} - \frac{1}{D_c} \right| \quad d_o \in [f, +\infty[\quad (A1-12)$$

L'image est nette sur un capteur lorsque le diamètre de la tâche lumineuse C est inférieur au pas pixel pour chaque source ponctuelle de la scène. Cette équation est indépendante du positionnement de la source dans le plan perpendiculaire à l'axe optique. Pour une distance objet donnée, dans les conditions de Gauss, le flou est donc le même partout dans l'image (cela devient faux dès qu'on considère les aberrations optiques).

Il existe trois évolutions possibles pour la tâche de flou lors d'un balayage de foyers de proche à l'infini.

- Dans le cas d'un objet situé trop proche, l'objet est flou dès le départ et son niveau de flou augmente avec l'augmentation de la distance focale. C'est une situation désespérée, les seules solutions pour rendre l'objet net seront de reculer physiquement l'appareil photo ou bien de changer d'objectif.
- L'objet est flou au plus proche foyer, tend vers une image nette puis redevient flou. Cet objet est généralement dans un champ proche ou moyennement proche de l'appareil.
- L'objet est flou au départ, l'évolution de la mise au point améliore son image jusqu'à ce qu'elle soit nette quelle que soit l'augmentation de la distance focale. C'est le cas des objets situés à « l'infini ».

La Figure A1-4 montre une simulation de l'évolution de C lors d'un balayage de foyers pour 3 objets en fonction de la distance du plan conjugué du capteur. On a fixé la distance capteur lentille d_c à 2cm, l'ouverture L à 5cm, une mise au point minimum de 8cm et les 3 objets à 6cm, 30cm et 500 cm. La ligne horizontale en pointillés représente un pas pixel. Celui-ci a été choisi arbitrairement grand devant la taille réelle des pixels standards de nos jours (de l'ordre de quelques microns) afin qu'il soit visible sur la figure.

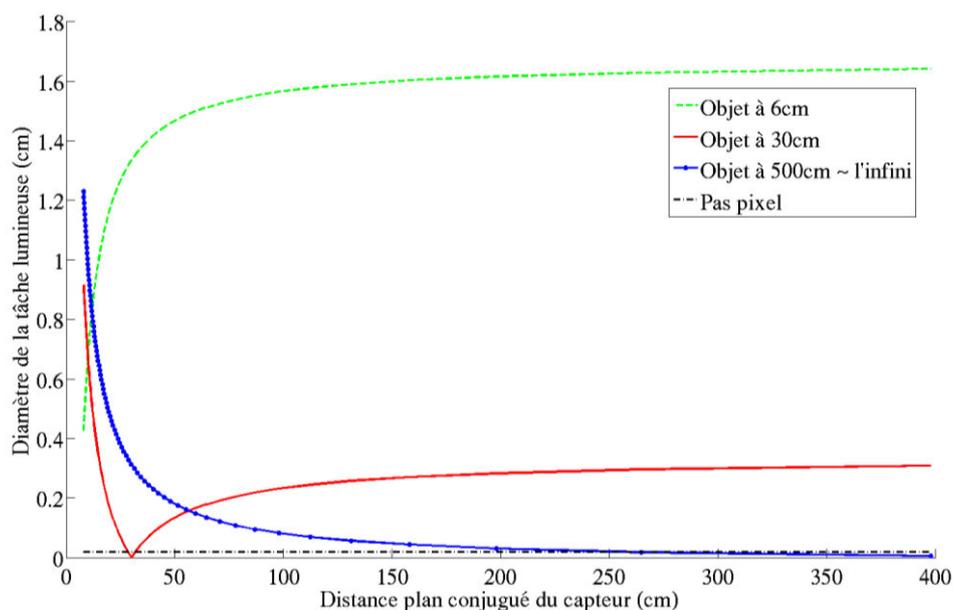


FIGURE A1-4 : EVOLUTION DE LA TACHE LUMINEUSE EN FONCTION DE LA POSITION DU PLAN CONJUGUE DU CAPTEUR

La courbe verte pointillée illustre le premier cas d'évolution de flou vu plus haut. La tâche lumineuse est au départ plus grande que le pixel, l'image est donc floue et le devient de plus en plus avec l'augmentation de la distance focale. Ce cas est source d'erreur dans toutes les méthodes d'analyse de profondeur mono capteur. Dans le cas du DfD, il est difficile à différencier d'un cas de flou évolutif correspondant à la courbe rouge sur la Figure A1-4, dans les cas du DfF, l'évaluation de profondeur sera toujours fausse ce qui peut être fatal, notamment dans le cadre d'une navigation autonome d'un robot.

La courbe bleue correspond aux objets considérés à l'infini. Il existe une distance focale unique pour laquelle l'image fait exactement la taille du pixel. Une fois cette distance dépassée, le diamètre de la tâche lumineuse reste sous la barre du pas pixel malgré l'éloignement du plan conjugué. Il s'agit du cas qui limite les performances des méthodes de mesure de profondeur avec un seul capteur. Il existe donc une distance, communément appelée distance hyperfocale, au-delà de laquelle les images de chaque élément n'évoluent plus. Ils sont tous nets sans pour autant se trouver à une même distance de la lentille.

Sur la courbe continue rouge, on voit l'évolution de la tâche lumineuse d'un objet situé dans le champ proche. Le diamètre de celle-ci passe sous le pas pixel puis à nouveau au dessus. Ce sont ces objets qui vont être les plus faciles à discriminer avec une approche DfF.

Annexe 2 : Evaluation de distance par inflexion de la luminosité

Les méthodes d'extraction de carte de profondeur présentées dans ce manuscrit sont basées sur une approche locale de l'estimation de la netteté. Dans ce nouvel algorithme de construction de carte de profondeur, nous avons exploité le cube multi-foyers suivant sa troisième dimension à l'instar de [63], c'est à dire en créant une dépendance entre les images consécutives du cube. Nous avons vu dans la section 2.2.4 sur la construction de critères de netteté, que l'évolution de l'intensité du pixel était impactée par la mise au point. L'idée exploitée ici consiste donc à chercher pour chaque pixel l'indice correspondant au point d'inflexion, c'est-à-dire à l'annulation de la dérivée de l'intensité lumineuse en fonction de la distance de focalisation.

Dans la réalité chaque évolution de pixel est une fonction continue qui a été échantillonnée. De ce fait nous n'avons quasiment plus aucune chance de tomber sur une dérivée exactement nulle. On peut alors chercher à détecter un changement de signe dans la dérivée pour repérer le passage au foyer le plus net. Ceci étant sans tenir compte du bruit de mesure. En effet dans des zones relativement uniformes pour lesquelles l'intensité varie très peu, le bruit de lecture peut amener à avoir un grand nombre de changements de signe successifs lors de la variation de la focalisation, masquant ainsi le vrai point net.

Au regard de l'évolution de la netteté, la dérivée de chaque pixel en fonction de la distance focale devrait présenter, en plus d'un changement de signe, une discontinuité relativement marquée. Nous avons présumé que ce saut de valeur devait être discernable malgré le bruit de mesure. On peut alors chercher le changement de signe le plus marquant en utilisant un seuil de détection ou rechercher un maximum de variation, mais comment choisir la valeur de ce seuil ? Détecter un saut marquant de valeur revient finalement à détecter un extremum dans la dérivée seconde des pixels à travers le cube multi-foyers.

Cela nous donne un calcul du type (A2-1) pour sélectionner l'indice de profondeur d'une scène.

$$DM(x, y) = \operatorname{argmax}_z \left| \frac{\partial^2 I_m}{\partial z^2}(x, y, z) \right| + 2 \quad (\text{A2.1})$$

Le calcul de la dérivée seconde à travers le cube d'image provoque un décalage d'indice qui doit être compensé, c'est pourquoi 2 est ajouté à la carte de profondeur.

Malgré ces précautions pour faire face au bruit de mesure, nous ne sommes toujours pas totalement à l'abri d'erreurs d'estimation. D'autre part, le zoom de focalisation qui a un

impact sur les méthodes basées sur le contraste local a ici aussi un impact prépondérant. En effet, l'évolution d'un pixel doit être observée à travers différentes coordonnées en fonction de l'indice de focalisation. L'estimation du décalage de pixel doit être adaptée à chaque système optique ce qui est problématique dans le cadre de réalisation d'une architecture multi-application.

En partie pour compenser le zoom de focalisation mais aussi pour supprimer du bruit résiduel, nous avons choisi d'effectuer un traitement par fenêtre de taille $w_h \times w_w$ pixels qui permet de résoudre ce problème de décalage au prix d'une légère augmentation de la complexité. Pour chaque fenêtre contenant $w_h \times w_w$ dérivées secondes calculées à travers le cube multi-foyers, nous calculons la somme des valeurs absolues et l'utilisons comme critère de netteté (A2.2).

$$S_{crit}(x, y, z) = \sum_{X,Y} \left| \frac{\partial^2 I_m}{\partial z^2}(X, Y, z) \right| \quad (A2.2)$$

Où (X,Y,z) représente un pixel dans un fenêtre de traitement de $w_h \times w_w$ pixels centrée sur (x,y,z) . Finalement les indices de profondeurs et de confiance sont extraits avec des équations similaires à celles pour nos algorithmes intra-image.

$$DM(x, y) = \underset{z}{\operatorname{argmax}}(S_{crit}(x, y, z)) + 2 \quad (A2.3)$$

$$C(x, y) = \underset{z}{\max}(S_{crit}(x, y, z)) - \underset{z}{\min}(S_{crit}(x, y, z)) \quad (A2.4)$$

Les cartes résultantes sont montrées sur la Figure A2-1.

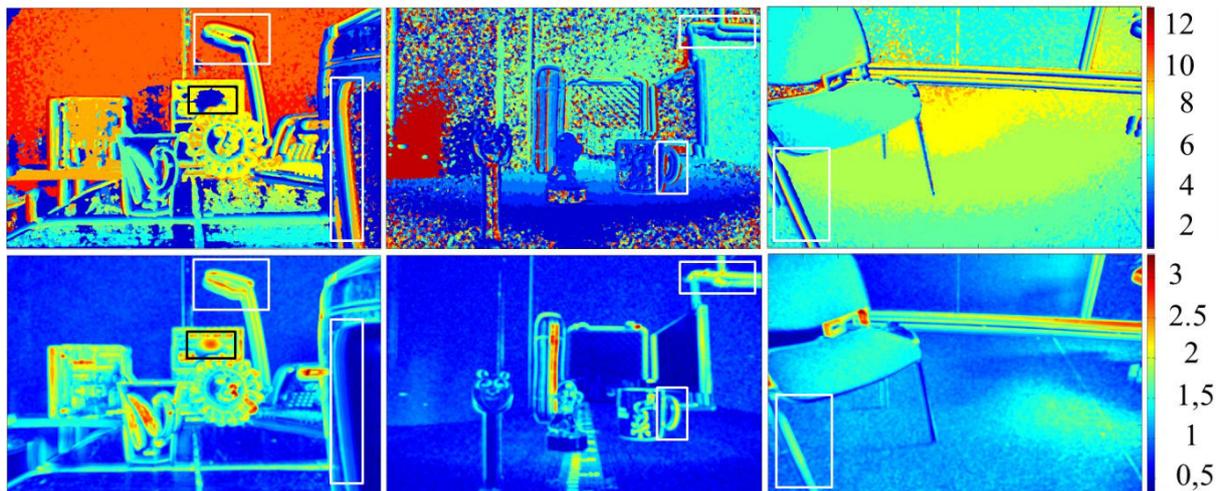


FIGURE A2-1 : CARTES DE PROFONDEURS (HAUT) ET DE CONFIANCES (BAS) POUR LES 3 SCENES ($w_x=w_y=16$)

Au regard des cartes de profondeur de la Figure A2-1, les zones micro-texturées comme la table, les murs et le sol, semblent particulièrement bien estimées par cette méthode. En revanche les détails des scènes, principalement le contour des objets sont très impactés par le zoom de focalisation comme le montre la présence de dégradés de profondeurs qui forment un halo dans le voisinage des contours (carrés blancs Figure A2-1). En outre, des zones

supposées bien estimées d'après la carte de vraisemblance sont en réalité de mauvaises estimations. Un exemple flagrant est encadré en noir sur l'image Soleil. Une tâche centrale est très vraisemblance alors qu'elle est sans ambiguïté mal évaluée.

Le tableau A2-1 montre le score de similarité τ_c entre les cartes de profondeur extraites avec diverses méthodes, présentées dans le chapitre 2, dont celle-ci et les vérités terrains. On voit que la méthode proposée ici donne des résultats moyens.

Tableau A2-1: Comparaison de différentes méthodes d'extraction de profondeur avec la carte de profondeur idéale ($s=1$)

Image	<i>Soleil</i>	<i>Panda</i>	<i>Chaise</i>
Notre méthode basse résolution blocs distincts	0.7960	0.6038	0.9256
Notre méthode intra-image en basse résolution	0.8054	0.6196	0.9332
Notre méthode intra-image en pleine résolution	0.7233	0.5283	0.9105
Notre méthode inter-image (Annexe 2)	0.6414	0.5139	0.7270
Max du module du gradient	0.6558	0.4644	0.8814
Max du Laplacien	0.6513	0.4334	0.8916
SMD	0.7014	0.5777	0.9117
Yokata (première étape)	0.3511	0.2520	0.5845

Cet algorithme montre une sensibilité très importante aux variations de luminosité comme la modulation engendrée par l'éclairage néon par exemple. Le résultat est difficilement reproductible pour différents cubes d'images d'une même scène. De plus, elle fait perdre 2 indices de résolution de profondeur à cause du calcul de la dérivée second à travers le cube d'images multi-foyers. Avec un post traitement conséquent, cette analyse d'un cube multi-foyers pourrait être intéressante mais cela va à l'encontre de l'objectif de cette thèse.

Annexe 3 : Etude de la profondeur de champ

En optique géométrique, il existe théoriquement un seul plan de l'espace qui est net pour une distance focale donnée. Mais puisque les photo-sites des capteurs d'image ont une taille minimum définie, il existe en réalité une tranche de l'espace, appelée profondeur de champ, pour laquelle la tâche lumineuse est comprise dans le pixel, et donc pour laquelle la netteté est maximale. Nous allons de nouveau nous intéresser à l'évolution du cercle de confusion afin d'étudier les paramètres qui influencent la profondeur de champ.

La profondeur de champ est problématique pour le DfF car elle est une source d'imprécision sur la mesure de la netteté de la distance. La Figure A3-1 permet de l'appréhender à partir de deux sources lumineuses ponctuelles dans les conditions de Gauss.

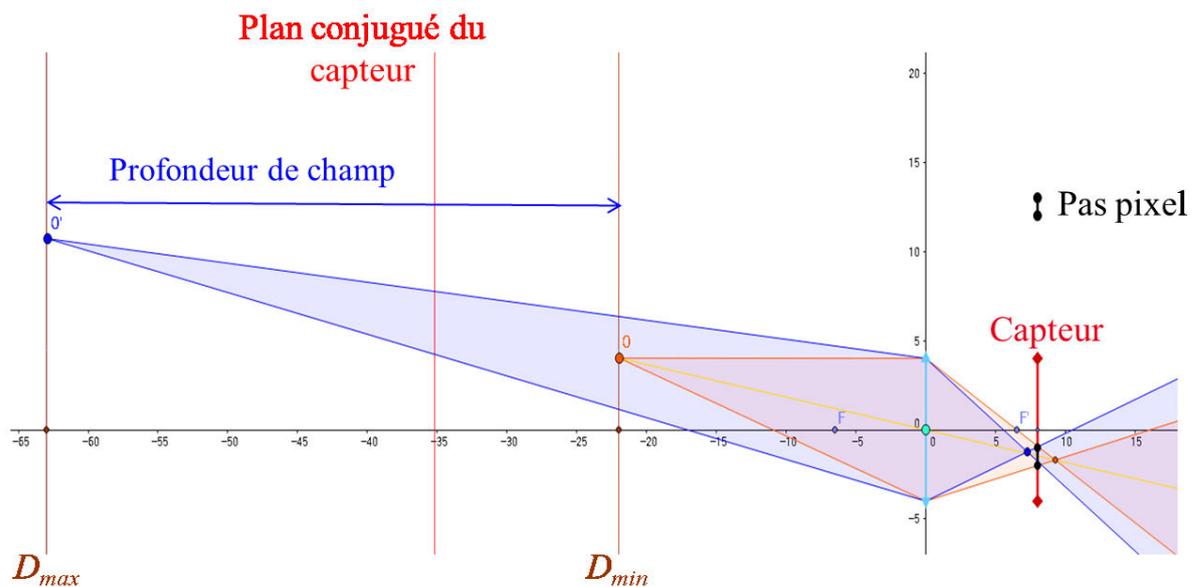


FIGURE A3-1 : ILLUSTRATION DE LA PROFONDEUR DE CHAMP

Le pas pixel (segment noir délimité par deux points noirs sur le capteur) est volontairement gros afin de faire ressortir le phénomène observé. Les sections des cônes lumineux des deux sources lumineuses avec le capteur correspondent exactement au pas pixel. Bien qu'une certaine distance les sépare, les images de chacune de ces sources sont nettes.

Avant de décrire mathématiquement la profondeur de champ, une petite réflexion en regard de la Figure A3-1 permet de prévoir les paramètres influents. L'origine de la profondeur de champ est due à la taille du pixel qui ne peut être infiniment petit. Elle est directement liée à la tâche lumineuse qui est tout simplement la section du cône lumineux issu de la lentille avec le plan capteur. Ce cône est dépendant des distances objet et image, puisqu'elles définissent

le point de focalisation de l'objet. Enfin l'ouverture joue également un rôle sur la section des cônes lumineux dans le plan de la lentille et donc aussi sur la profondeur de champ. Cette dernière est d'ailleurs bien connue des photographes car, pour un objectif donné, la variation d'ouverture constitue très souvent le seul levier permettant de faire varier l'effet de bokeh (effet de flou). Plus l'ouverture est grande, plus l'angle du cône est obtus, le diamètre de la tâche lumineuse augmente alors ce qui réduit l'intervalle de distance pour lequel la tâche lumineuse est incluse dans un pixel.

En résumé, il y aurait donc 5 paramètres influents sur la profondeur de champ : d_c , d_o , L , f et P_{pix} . On retrouve les 4 paramètres présents dans l'équation (A1.12) représentant le diamètre de la tâche lumineuse.

Reprenons l'étude de la tâche lumineuse du paragraphe précédent en effectuant un zoom autour de la distance de focalisation pour deux sources lumineuses situées à deux distances différentes (Figure A3-2).

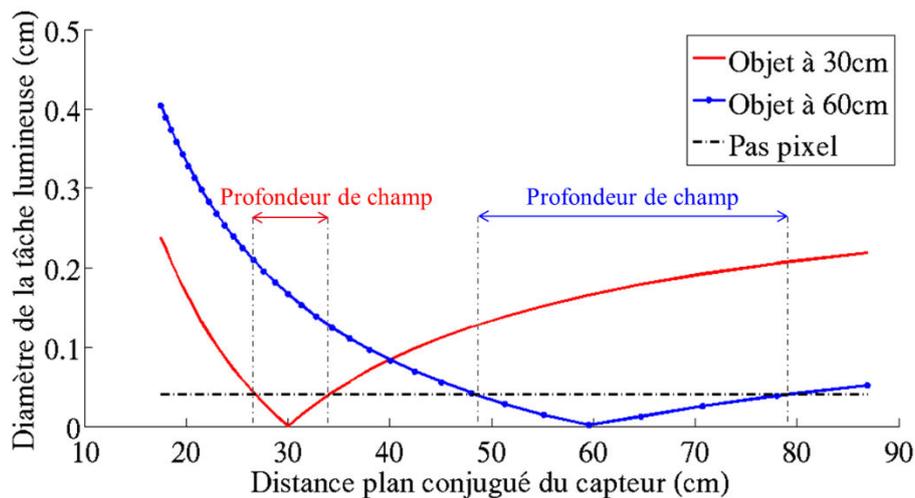


FIGURE A3-2 : INFLUENCE DU PAS PIXEL ET DE LA DISTANCE FOCALISEE SUR LA PROFONDEUR DE CHAMP

Ce graphique montre que la profondeur de champ grandit avec l'éloignement et tend vers l'infini puisqu'il existe une distance objet (lorsque $(d_c - f) \rightarrow 0$) à partir de laquelle l'ensemble des éléments est net. Avec la figure A3-2, il apparaît clairement que la profondeur de champ augmente lorsque le pas pixel augmente.

Mathématiquement, la profondeur de champ peut donc être définie comme un ensemble de 2 distances $\{D_{min}, D_{max}\}$. Celles-ci sont obtenues à partir des équations (A1.6) et (A1.11) en posant $C = P_{pix}$.

Zone post focale :

Calculons la valeur limite de D_{max} pour laquelle $C = P_{pix}$ dans le cas d'une mise au point lointaine.

$$C_{far} = L \cdot d_c \cdot \left(\frac{1}{D_c} - \frac{1}{D_{max}} \right) = P_{pix} \quad (A3.1)$$

Ceci implique que

$$\frac{Ld_c}{P_{pix}} \left(\frac{1}{D_c} - \frac{1}{D_{max}} \right) = 1 \quad (A3.2)$$

En posant K une constante résultant de la conception d'un capteur d'ouverture L , de distance capteur d_c conçu avec un pas de pixel P_{pix} :

$$K = \frac{Ld_c}{P_{pix}} \quad (A3.3)$$

On obtient

$$\frac{1}{D_c} - \frac{1}{D_{max}} = \frac{1}{K} \quad (A3.4)$$

Pour une focale donnée et donc un D_c donné, la condition sur d_o pour que le système soit dans les conditions hyperfocales est :

$$\frac{1}{D_{max}} = \frac{1}{D_c} - \frac{1}{K} \quad (A3.5)$$

et

$$D_{max} = \begin{cases} \frac{KD_c}{K-D_c} & \text{pour } f < D_c < K \\ +\infty & D_c \geq K \end{cases} \quad (A3.6)$$

Zone pré focale :

De la même manière, on s'intéresse à la distance pré focale D_{min} définie par :

$$C_{near} = L \cdot d_c \cdot \left(\frac{1}{D_{min}} - \frac{1}{D_c} \right) = P_{pix} \quad (A3.7)$$

Ceci implique que

$$\frac{Ld_c}{P_{pix}} \left(\frac{1}{D_{min}} - \frac{1}{D_c} \right) = 1 \quad (A3.8)$$

Et donc, en remplaçant par la même constante technologique K (A3-3)

$$K \left(\frac{1}{D_{min}} - \frac{1}{D_c} \right) = 1 \quad (A3.9)$$

Et

$$\frac{1}{D_{min}} = \frac{1}{K} + \frac{1}{D_c} \quad (A3.10)$$

Soit

$$D_{min} = \frac{KD_c}{K+D_c} \quad (A3.11)$$

Profondeur de champ :

La profondeur de champ est la distance $D_{max} - D_{min}$ pour le cas ou $D_c < K$

$$D_{prof} = \frac{KD_c}{K-D_c} - \frac{KD_c}{K+D_c} = KD_c \left(\frac{1}{K-D_c} - \frac{1}{K+D_c} \right) \quad (A3.12)$$

$$D_{prof} = KD_c \left(\frac{K+D_c - K+D_c}{K^2 - D_c^2} \right) \quad (A3.13)$$

$$D_{prof} = \frac{2KD_c^2}{K^2 - D_c^2} = 2K \left(\frac{1}{\left(\frac{K}{D_c}\right)^2 - 1} \right) \quad (A3.14)$$

Et finalement

$$D_{prof} = \begin{cases} 2K \left(\frac{1}{\left(\frac{K}{D_c}\right)^2 - 1} \right) & \text{pour } f < D_c < K \\ \infty & \text{pour } D_c \geq K \end{cases} \quad (A3-15)$$

Avec

$$K = \frac{Ld_c}{P_{pix}}$$

Ces résultats montrent que les paramètres clefs pour la profondeur de champ d'un capteur sont l'ouverture du capteur L , la distance du capteur à la lentille d_c et le pas pixel P_{pix} . Une fois fixés, ces paramètres définissent la profondeur de champ en fonction du plan conjugué du capteur D_c via de manière unique.

Au-delà de $D_c = K$, la distance hyperfocale est atteinte et le cercle de confusion des objets situés au-delà de cette distance sont toujours inférieurs à la taille du pixel. Le paramètre K , appelée distance hyperfocale, définit donc la distance maximale de scène pour laquelle une mesure DfF est possible.

La Figure A3-3 ci-dessous décrit la profondeur de champ normalisée D_{prof}/K en fonction du de la distance de mise au point, représenté par D_c/K , elle permet de servir d'abaque pour la définition des paramètres physiques du système.

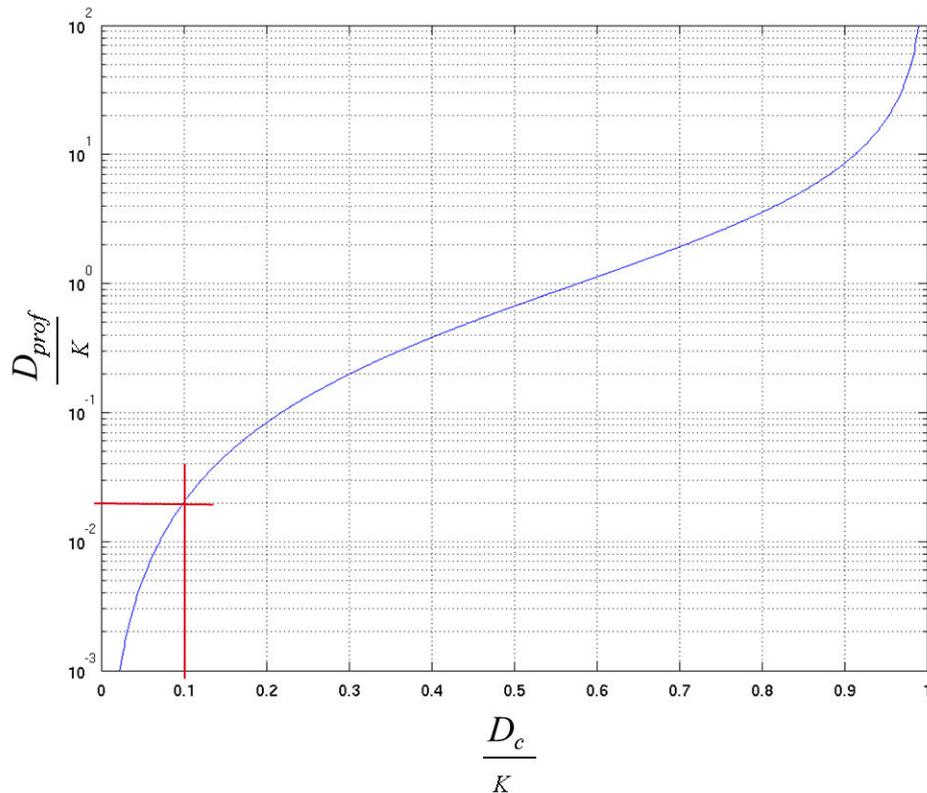


FIGURE A3-3 : EVOLUTION DE LA PROFONDEUR DE CHAMP EN FONCTION DE LA DISTANCE CONJUGUEE DU CAPTEUR

Prenons l'exemple d'une configuration typique d'un petit module imageur : pas pixel de 3.75 microns, une lentille équivalente à focale variable placée à environ 8mm du capteur et une ouverture de 6mm $K = 12.8m$. Pour $D_c = 1.28m$ (tracé rouge sur la figure A3-3) on a une profondeur de champ de $D_{prof} = 2E^{-2} \times K$ soit 25.6cm.

La profondeur de champ augmente rapidement en fonction de la distance du plan net. Cela va poser un problème de précision lors de l'évaluation de la distance pour les objets lointains. Avec la configuration de capteur précédente, l'imprécision passe d'environ 25.6cm à 1.28m de distance à environ 1.28m à 3m de distance. À ce modèle il faudrait ajouter les diverses aberrations optiques, par exemple les aberrations chromatiques, ou encore l'effet de pupille, ou tâche d'Airy, qui viendraient dégrader l'estimation.

Il existe une astuce théorique pour améliorer la résolution de profondeur malgré une profondeur de champ qui augmente : si on est capable d'extraire précisément pour chaque élément de la scène le foyer exact pour lequel celui-ci devient net lors d'un balayage de foyers croissant partant d'une faible distance focale, on peut alors estimer précisément sa profondeur qui est alors D_{max} . Il reste à savoir si un algorithme de détection de netteté est capable d'extraire avec précision cette première image nette.

Annexe 4 : Calcul de l'épaisseur d'une lentille à cristaux liquides

Dans cette annexe, nous cherchons à calculer l'épaisseur e d'une lentille à cristaux liquides de diamètre de pupille P permettant d'obtenir une distance focale minimale f_{min} .

Une lentille à cristaux liquide joue le rôle d'une lame à retard de phase dont l'indice optique dans la couche de cristaux liquides a une distribution sphérique. C'est ce gradient d'indice qui permet de générer une lentille (Figure A4-1)

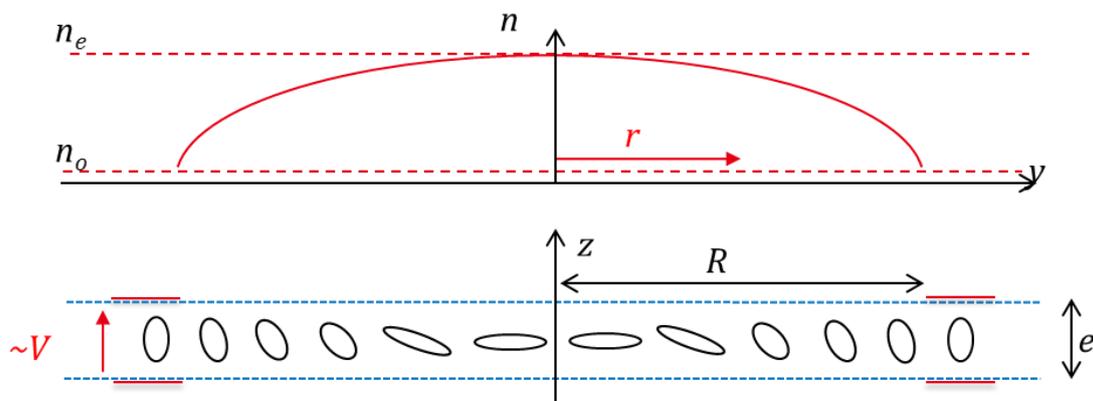


FIGURE A4-1 : REPARTITION DE L'INDICE DE REFRACTION SELON UNE DISTRIBUTION SPHERIQUE DANS UNE LENTILLE A CRISTAUX LIQUIDES

La thèse de Nicolas, [106] p25-27, présente la méthode de calcul de l'épaisseur d'une lentille à cristaux liquide en faisant une analogie en termes de déphasage avec une lentille mince biconvexe.

Pour une longueur d'onde λ , la fonction de transfert optique de cette structure et celle d'une lame à retard en considérant un indice variable $n(r)$:

$$F(r) = \exp\left(\frac{2\pi}{\lambda} \cdot e \cdot n(r)\right) \quad (\text{A4-1})$$

L'indice de réfraction de la couche de cristaux liquides s'écrit (A4-2) avec r telle que $r^2 = y^2 + x^2$ et R le rayon de la pupille.

$$n(r) = n_e + \frac{n(r) - n_e}{R^2} \cdot r^2 \quad (\text{A4-2})$$

On en déduit la fonction de transfert théorique

$$F(r) = \exp\left(\frac{2\pi}{\lambda} \cdot e \left(n_e + \frac{n(r) - n_e}{R^2} \cdot r^2\right)\right) \quad (\text{A4-3})$$

$$F(r) = \exp\left(\frac{2\pi}{\lambda} \cdot e \cdot n_e\right) \cdot \exp\left(\frac{2\pi}{\lambda} \cdot e \cdot \frac{n(r) - n_e}{R^2} \cdot r^2\right)$$

Prenons le cas d'une lentille mince biconvexe d'indice n et d'épaisseur e en son centre telle que représentée sur la figure A4-2.

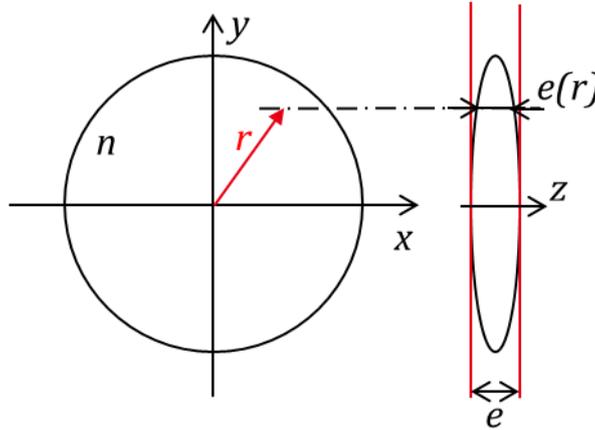


FIGURE A4-2 : SCHEMA LENTILLE MINCE BICONVEXE

Le retard de phase induit par cette lentille est une fonction de son épaisseur (A4-4).

$$\varphi_L(r) = \frac{2\pi}{\lambda} \cdot n \cdot e(r) \quad (\text{A4-4})$$

Le déphasage introduit par le milieu d'indice n' entre les plans tangents à l'entrée et à la sortie du dioptre s'exprime (A4-5).

$$\varphi_m(r) = \frac{2\pi}{\lambda} \cdot n' \cdot e - \frac{2\pi}{\lambda} \cdot n' \cdot e(r) \quad (\text{A4-5})$$

On en déduit le retard total subit par l'onde entre les plans tangents aux deux dioptres des lentilles par (A4-5).

$$\varphi_t(r) = \frac{2\pi}{\lambda} \cdot n' \cdot e + \frac{2\pi}{\lambda} \cdot (n - n') \cdot e(r) \quad (\text{A4-5})$$

La fonction de transfert optique modélisant la tranche de l'espace contenu entre les deux plans normaux au dioptre de la lentille est donc (A4-6)

$$F(r) = \exp\left(\frac{2\pi}{\lambda} \cdot e \cdot n'\right) \cdot \exp\left(\frac{2\pi}{\lambda} \cdot (n - n') \cdot e(r)\right) \quad (\text{A4-6})$$

En considérant l'approximation paraxiale, on peut exprimer l'épaisseur de la lentille par (A4-7) où R_1 et R_2 sont les rayons de courbures des faces d'entrée et de sortie de la lentille.

$$e(r) = e - \frac{r^2}{2} \cdot \left(\frac{1}{R_1} - \frac{1}{R_2}\right) \quad (\text{A4-7})$$

Avec l'approximation de Gauss et en considérant que la lentille mince et dans l'air ($n' = 1$), la distance focale est donnée par l'expression (A4-8)

$$\frac{1}{f} = (n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \quad (\text{A4-8})$$

On obtient une nouvelle fonction de transfert pour la lentille mince qui est (A4-9)

$$F(r) = \exp\left(\frac{2\pi}{\lambda} \cdot e \cdot n\right) \cdot \exp\left(-\frac{2\pi}{\lambda} \cdot \frac{r^2}{2 \cdot f}\right) \quad (\text{A4-9})$$

Par analogie entre la lentille à cristal liquide la lentille mince, on déduit (A4-10) des équations (A4-9) et (A4-3)

$$\frac{1}{2 \cdot f} = e \cdot \frac{n(r) - n_e}{R^2} \quad (\text{A4-10})$$

Ce qui mène à la distance focale minimum équivalente de la lentille à cristaux liquide f_{\min} avec l'équation (A4-11) en considérant $n(R) = n_0$ et le diamètre de la pupille $P=R^2$

$$f_{\min} = \frac{P^2}{8 \cdot \Delta n \cdot d} \quad (\text{A4-11})$$

Annexe 5 : Etapes de fabrication d'une lentille à cristaux liquides

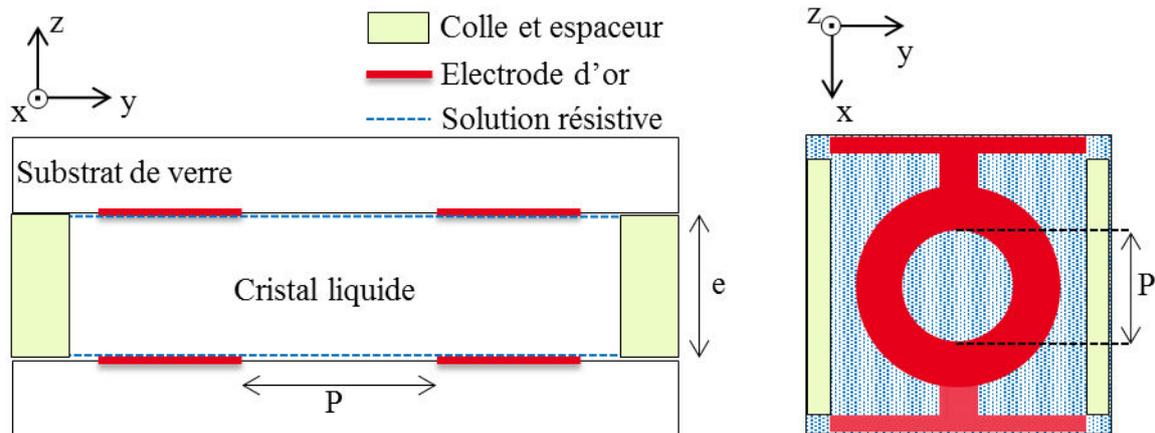


FIGURE A5-1 : SCHEMA DES LENTILLES REALISEES : VUE EN COUPE (GAUCHE) ET VU DE DESSUS (DROITE)

L'IMT Atlantique a à sa disposition une salle blanche d'optique dans laquelle nous avons réalisé des lentilles telles que présentées sur le schéma figure A5-1.

Plusieurs étapes sont nécessaires pour la réalisation d'une lentille. Une étape de préparation des substrats, une étape de gravure, une étape de dépôt de couche résistive, une étape de broyage puis d'assemblage.

La propreté des substrats est primordiale pour toutes les étapes de réalisation. Entre chacune d'elles, ils sont, au minimum nettoyés par un flux d'air sous pression, et lorsque c'est possible dans un bain à ultrason puis avec, dans cet ordre, de l'acétone, du isopropanol et de l'eau déminéralisée.

Préparation des substrats :

La première étape consiste à découper des substrats de verre recouverts d'une couche conductrice. Nous avons expérimenté deux technologies : des substrats recouverts d'une couche d'Oxyde d'Indium dopé à l'Etain, ou ITO (Oxyde d'Indium-Etain) qui est transparent en couche mince, et des substrats recouverts d'une couche de chrome. Les substrats d'ITO à notre disposition ont l'avantage d'être plus fin, 550 microns contre 1500 microns pour le chrome, ce qui permet de limiter l'épaisseur finale de la lentille en ainsi, de ne pas trop réduire le champ de vision. En revanche le dépôt d'or puis la gravure des électrodes avec l'ITO sont plus délicats qu'avec le chrome.

Pour améliorer le contact électrique entre les couches conductrices et entre les connecteurs et les électrodes, une fine couche d'or de 50nm est déposée côté conducteur. Pour cela, on a utilisé un ionisateur à argon qui permet de vaporiser une cible d'or dont les atomes vont se déposer sur les substrats. 300s d'ionisation permettent de déposer les 50 nanomètres d'or recommandés.

Gravure des électrodes :

On effectue ensuite un nettoyage, très délicat pour ne pas rayer la couche d'or, avant de faire un dépôt de résine photosensible à l'aide d'une centrifugeuse. On cherche à obtenir une épaisseur constante et connue afin de permettre une insolation optimum de la résine avant l'étape de gravure. Pour cela trois principaux paramètres sont ajustés pour la centrifugeuse : l'accélération angulaire, la vitesse de rotation et le temps de rotation. Avant application, la résine traverse un filtre de 0,2 microns afin d'éliminer tout agglomérat de matière qui aurait pu se former durant son stockage.

Pour insoler les substrats, on utilise un masque tel que celui présenté Figure A5-2. Pour une grande précision, il peut être réalisé en chrome, par photo-traçage puis gravé sur une plaque de verre. Dans notre cas, le motif étant relativement gros, une simple impression avec une haute teneur d'encre, pour maximiser l'opacité, est suffisant.



FIGURE A5-2 : EXEMPLE DE MASQUE POUR UNE LENTILLE DE 5MM DE PUPILLE

Le temps d'insolation pour une épaisseur de résine de l'ordre de la centaine de nanomètres est de 15s. On révèle ensuite la résine insolée avant de graver l'or puis le chrome ou l'ITO, avec deux solutions différentes. Le procédé a d'abord été caractérisé pour obtenir des temps optimum de gravure de 3minutes pour l'or, 1min30 pour le chrome et 6min pour l'ITO. Enfin, on rince les substrats avec de l'acétone pour supprimer toute trace de résine avant le dépôt de la couche résistive.

Dépôt de la couche résistive :

La couche résistive est constituée d'un mélange de Poly(3,4-éthylènedioxythiophène)-poly(styrène sulfonate) (PEDOT-PSS) de chez Heraeus qui apporte la résistivité, et d'alcool polyvinyl employé comme solvant. Les quantités mélangées de ces éléments ont un fort impact sur le fonctionnement de la lentille. Si la couche de PEDOT-PSS n'est pas suffisamment résistive, le potentiel se répand trop rapidement dans l'électrode et nous n'avons plus le gradient de champ électrique nécessaire pour l'émulation de lentille. Si elle est trop résistive

c'est l'effet inverse qui se produit confinant le champ seulement au niveau des électrodes. La répartition du champ, et donc de la résistance de la couche, doit être adaptée en fonction de la taille de la pupille (Voir [106]).

Le substrat est placé dans la centrifugeuse puis quelques gouttes de cette solution sont déposées après un filtrage à 0.8microns. Les paramètres de la centrifugeuse sont bien définis pour obtenir la bonne épaisseur de solution résistive.

A partir de cette étape, l'air comprimé constitue le seul moyen que nous avons pour nettoyer les substrats, il est donc primordial de les maintenir dans des boites afin de limiter les risques de contamination par une quelconque poussière.

L'étape suivante consiste à brosser les électrodes pour former des sillages longilignes dans la couche de PEDOT-PSS. Ceux-ci constitueront une surface d'ancrage pour les cristaux liquides. Les cellules devront être assemblées avec des directions de brossage opposées. Ce brossage est effectué à l'aide d'un cylindre de velours tournant à vitesse constante sous lequel les cellules vont venir se frotter à l'aide d'une table mobile. La hauteur de brossage est délicate à régler, si le brossage est trop fort, la couche résistive risque d'être arrachée et s'il est trop faible, les cristaux liquides ne s'ancreront pas correctement ce qui rendra leur mise en rotation chaotique lors de l'application du champ électrique. Après divers essais, nous avons opté pour un brossage à 150 microns en dessous de la hauteur de touche du rouleau.

Assemblage et remplissage des cellules :

Les électrodes sont prêtes à être assemblées. Pour cela, nous avons préparé des mélanges de colle avec des espaceurs (microbilles) calibrés de différentes tailles. Ils permettent de garantir une épaisseur constante de la couche de cristaux liquides. Deux joints de colle sont réalisés sur deux des bords d'une des électrodes puis la cellule est assemblée avec une force suffisante pour arriver en butée contre les espaceurs, tout en prenant garde à l'alignement des mires d'assemblage. L'ensemble est ensuite placé sous lampe UV pendant 45s afin de polymériser la colle. L'alignement est primordial pour une bonne qualité optique mais il n'est pas toujours aisé de le maintenir, lorsque la colle est très visqueuse et que la concentration d'espaceurs est trop importante, il peut arriver que les deux substrats ne soient pas parallèles créant ainsi une épaisseur non constante. Ce type de défaut peut se manifester par une difficulté accrue à maintenir alignés les deux substrats. Après un court repos de quelques minutes, les substrats se désalignent tous seuls par effet d'apesanteur. Ils redeviennent parallèles mais les deux pupilles n'ont plus le même axe. Il faut donc limiter le volume du joint de colle.

Le cristal liquide est introduit dans la cavité par simple effet de capillarité. Pour un remplissage optimum, on met une quantité non négligeable de cristaux liquides sur le bord d'une des électrodes et on observe l'avancée du front de cristal dans la cavité. Celui-ci doit être le plus droit possible et doit avancer avec une vitesse relativement constante. S'il n'y a pas assez de

cristal au départ, on prend le risque que des bulles d'air soient emportées par le cristal liquide dans la cavité ce qui engendrerait un dysfonctionnement de la lentille.

Suite au constat du vieillissement rapide des lentilles durant cette thèse, cela peut être intéressant de laisser dégazer le cristal liquide pour éviter les microbulles d'air dans la cellule. Un remplissage sous vide pourrait également être une solution.

Enfin pour éviter toute fuite des cristaux par pression sur la cellule, ou changement de pression dû à l'altitude, la cellule est scellée avec un joint de colle.

Procédé alternatif pour créer une cavité d'épaisseur calibrée :

Pour certaines épaisseurs de la couche de cristal liquide, il est assez difficile et surtout cher de se procurer des espaceurs. En effet, la plupart des applications avec des cristaux liquides ont une couche d'épaisseur comprise entre 1 et 20 microns. Il est donc courant de trouver de telles valeurs d'espaceur mais pour des épaisseurs de l'ordre de 50 à 150 microns, il faut parfois innover pour être capable de garantir l'écart entre les deux électrodes.

Nous avons donc mis au point un autre procédé de fabrication en utilisant un film résine pour créer nos propres espaceurs. Celui-ci est au départ destiné au vernissage des cartes électroniques. Nous disposons de rouleaux de résine calibrés à 40,20 et 5 microns.

Sur l'une des électrodes préalablement gravée, nous avons ajouté des couches de résines afin d'obtenir l'espacement voulu pour contenir la couche de cristal liquide. Puis nous l'avons insolée avec un masque de sorte à conserver seulement deux bandes de résine sur les bords d'une des deux électrodes après gravure.

La figure A5-3 montre à la fois une électrode avec des bandes de résines gravées (deux couches de 40 microns) et le profil des bandes de résine mesuré avec le profileur.

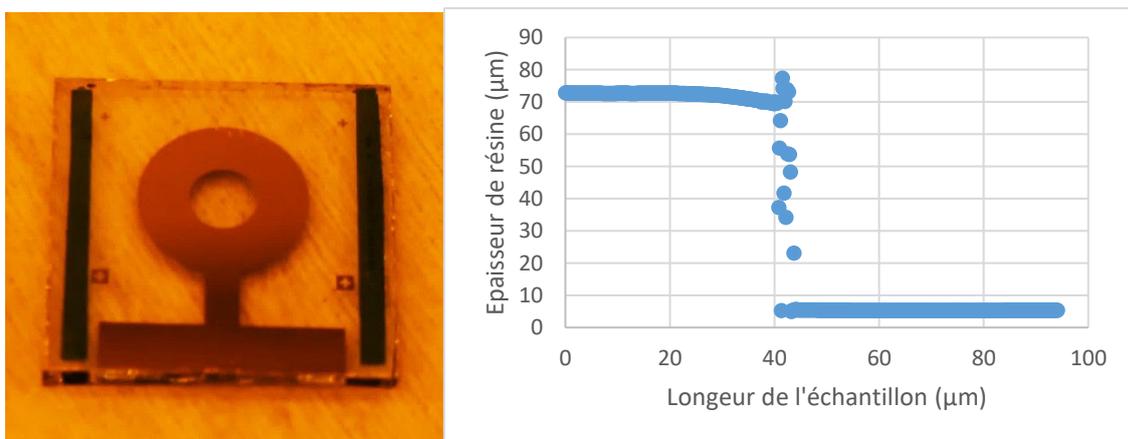


FIGURE A5-3 : PROFIL D'ÉPAISSEUR DE LA COUCHE DE RESINE PHOTOSENSIBLE DEPOSEE POUR CONTROLER L'ÉPAISSEUR DE LA COUCHE DE CRISTAUX LIQUIDES.

La figure A5-3 montre un profil relativement régulier avec une épaisseur de résine de 67,5µm. Il semblerait que l'épaisseur de la résine ne soit pas exactement de 40 microns, en revanche on a un profil relativement plat qui devrait permettre un bon maintien de l'électrode supérieure.

Si la résine ne suffit pas à obtenir l'épaisseur désirée, il reste la possibilité d'utiliser à la fois la résine et une colle contenant des espaceurs par-dessus la résine lors de l'assemblage.

Une fois la résine déposée, on peut commencer l'étape de dépôt de la couche résistive. Les bandes de résines peuvent perturber l'établissement de la couche de PEDOT-PSS à la centrifugeuse, en particulier parce qu'elle empêche la matière en surplus d'être éjectée par force centrifuge. Cependant nous avons remarqué que le centre de la pupille était homogène, ce qui devrait garantir le bon fonctionnement des lentilles ainsi réalisées.

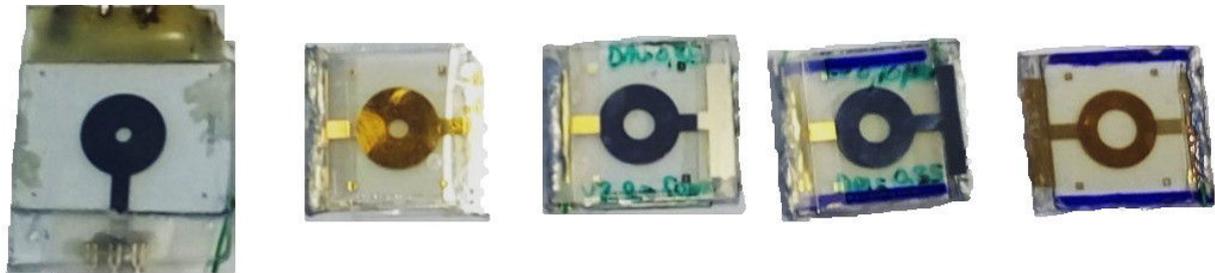


FIGURE A5-4 : EXEMPLE DE LENTILLE A CRISTAUX LIQUIDES, DE CAUCHE A DROITE, PUPILLE DE 2MM, 2.5MM, 4MM ET 5 MM

Evaluation de la qualité :

Pour contrôler le fonctionnement et la qualité de la lentille après la fermeture d'une cellule, celle-ci est observée sous loupe binoculaire entre polariseurs croisés.

Le déphasage induit par les cristaux liquides crée l'effet d'une lame à retard qui entre polariseurs croisés module l'intensité lumineuse. Nos électrodes étant circulaires les lignes d'équi-déphasage devraient être circulaires et centrées sur l'axe optique.

L'image 6 de la Figure A5-5 montre les franges d'interférences d'une lentille d'assez bonne qualité.

Notons qu'une fois la couche résistive déposée, aucun moyen de vérification du bon établissement des procédés n'est possible, le processus doit donc être complet afin de pouvoir faire une vérification du fonctionnement. Dans le cas où une lentille s'avère inexploitable, que ce soit à cause d'une poussière qui s'est glissée dans la lentille, d'une solution de PEDOT-PSS non adaptée, d'un brossage trop faible ou trop fort, de l'apparition de bulles dans la couche de cristaux liquides, tout le procédé doit être recommencé pour faire une nouvelle cellule. Il n'y a hélas aucun recyclage « possible ».

De façon générale, l'interprétation des dysfonctionnements est complexe. Et ce principalement parce que nous ne pouvons faire pratiquement aucun test avant l'assemblage final. Divers défauts ont été observés au cours de ces travaux, quelques-uns sont présentés figure A5-5.

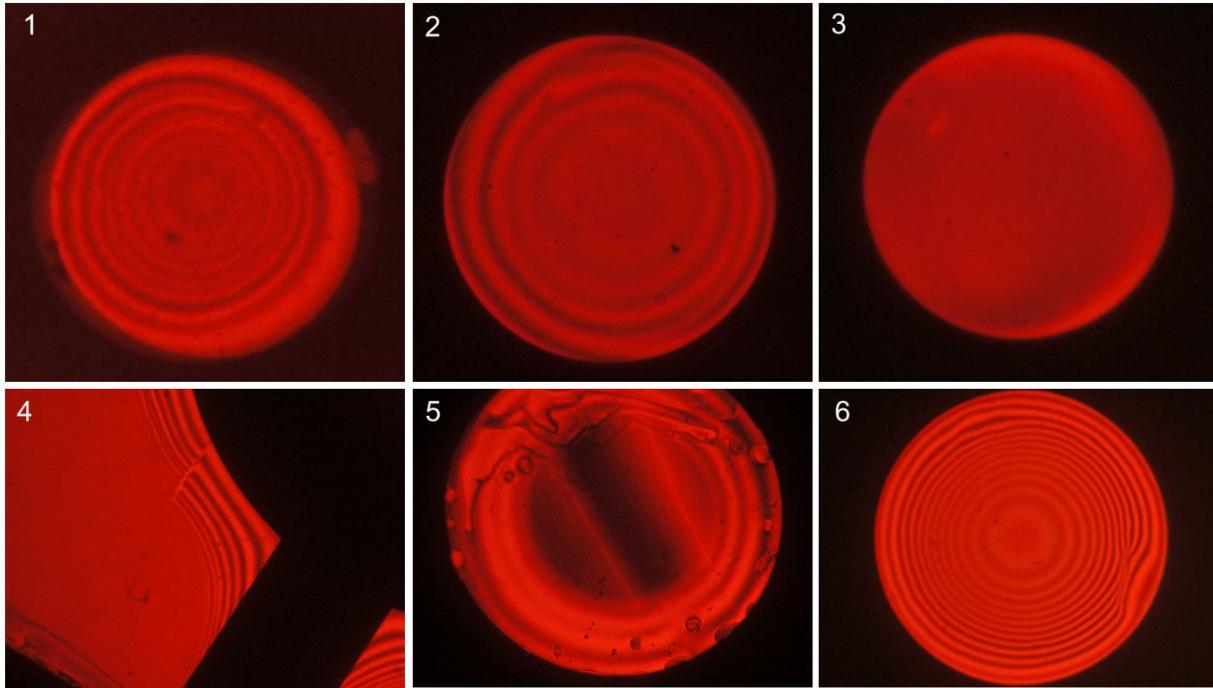


FIGURE A5-5 : 1 LENTILLE DESAXEE, 2 DEFAUT NON IDENTIFIE, 3 SOLUTION RESISTIVE NON ADAPTE, 4 ARRACHEMENT DE PEDOT-PSS, 5 BULLES + CHEVELU NON IDENTIFIE ET 6 QUALITE ACCEPTABLE

Annexe 6 : Réalisation d'un driver de lentille à cristaux liquides

Les cristaux liquides nématiques ne supportent pas d'être soumis à un champ électrique continu. Cela les électrolyse, modifiant ainsi le milieu de façon irréversible. Le champ électrique qui les contrôle doit être un champ alternatif sans composante continue. Dans le cas particulier des lentilles à cristaux liquide, à la fois l'amplitude et la fréquence du signal de contrôle permettent de faire varier la distance focale. D'un point de vue électrique, une lentille à cristaux liquides se comporte au premier ordre comme un condensateur, de capacité avoisinant le nF, en parallèle avec une résistance, de l'ordre du M Ω . Il faut donc prendre garde à conserver une faible résistance série dans le circuit de polarisation afin d'avoir une fréquence de coupure qui ne perturbe pas la génération de la fonction de lentille.

Nous souhaitons générer un signal crête à haute tension (+-100V max) et à haute fréquence (maximum 2MHz). Ces caractéristiques sont relativement difficiles à obtenir. En effet, un signal carré haute fréquence nécessite un fort courant de charge ce qui combiné à la haute tension impose d'utiliser des composants supportant de gros pics de puissance.

Pour simplifier la réalisation d'un driver nous avons seulement tenu compte du balayage de fréquence. La tension est réglable à partir d'un générateur externe.

Une manière simple de générer un signal à fréquence fixe est d'utiliser un microcontrôleur qui intègre matériellement des fonctions de génération de signal PWM (Pulse Width Modulation). Cela permet d'avoir des signaux à fréquence variable entre 0 et 3.3V qu'il faut ensuite amplifier afin d'avoir un champ suffisamment fort pour impacter l'orientation des cristaux liquides, sans oublier de filtrer pour supprimer la composante continue.

Le microcontrôleur utilisé dans le cadre de ces travaux est un STM32f4 intégré à une carte de développement STM32F4DISCOVERY. L'horloge interne est cadencée à 40MHz afin de garantir un pas relativement précis de la fréquence jusqu'à 2MHz. La communication avec l'ordinateur qui exécute les programmes d'analyse est réalisée grâce à un module FTDI. Celui-ci assure une liaison USB/série facilitant ainsi le développement de la communication. La transmission est configurée pour un débit de 115200 bauds ce qui permet de limiter le retard entre la commande et l'exécution. Le programme exécuté sur le STM32 ne fait rien de plus qu'ajuster à la demande la fréquence de la PWM. Dès qu'une requête lui parvient, une interruption de la liaison série est levée, la fréquence est immédiatement mise à jour. Cela garantit l'aspect temps réel du driver.

Pour obtenir un signal amplifié et centré, plusieurs solutions ont été expérimentées.

1- Amplificateur et filtre passe haut passif

Une première solution était d'utiliser un comparateur pour amplifier le signal PWM puis de le filtrer avec un condensateur de sorte à supprimer la composante continue (Figure A6-1).

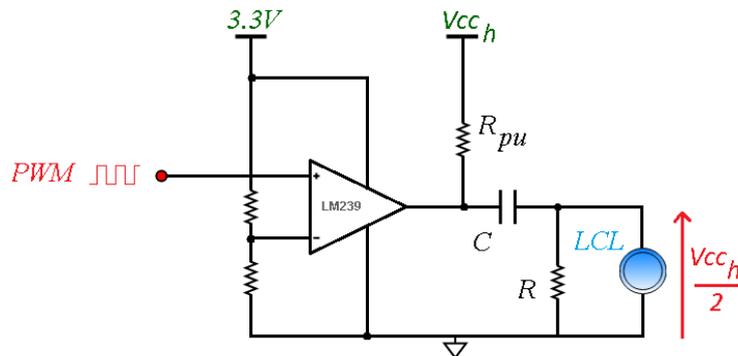


FIGURE A6-1 : DRIVER DE LENTILLE A CRISTAUX LIQUIDE PAR FILTRAGE PASSE HAUT.

Cette solution n'est pas idéale. Premièrement, la tension haute V_{cc_h} qui permet l'amplification doit être équivalente à la tension crête à crête de polarisation de la lentille, soit le double de l'amplitude nécessaire. Malheureusement ce n'est pas très adapté aux systèmes embarqués, du moins pour des amplitudes supérieures à 10V. Ensuite, même si d'un point de vue théorique la fréquence de coupure est bien placée, l'expérience nous montre que lorsqu'on tient compte des temps de charge et décharge du condensateur, le signal est très déformé en basses fréquences et il apparaît une composante continue, fatale pour les cristaux liquides.

2- Comparateur et masse fictive

L'architecture que nous avons utilisée pour les lentilles L1 et L2 est relativement simple (Figure A6-2).

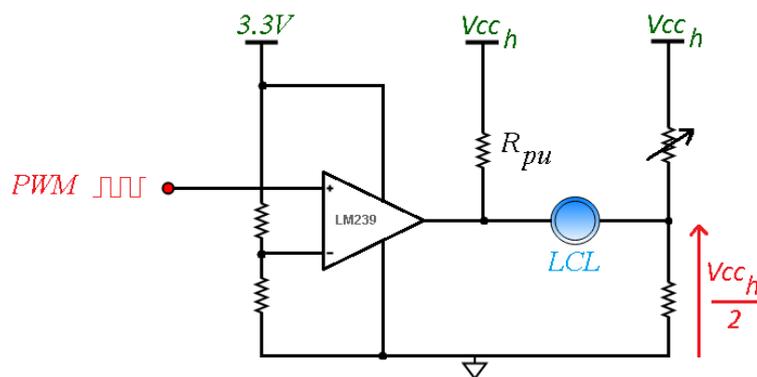


FIGURE A6-2 : DRIVER DE LENTILLE A CRISTAUX LIQUIDE

Un comparateur pur est également utilisé comme amplificateur du signal PWM. Une masse fictive ajustée à la moitié du potentiel de polarisation V_{cc_h} permet de centrer le créneau haute tension du point de vue de la lentille. L'amplificateur utilisé est le LM239. Il permet d'adapter la tension de sortie indépendamment de son alimentation grâce à un port de sortie en drain ouvert. Cela octroie la possibilité de faire des tests avec des tensions V_{cc_h} entre 1V et 36V,

tout en ayant une bande passante assez grande pour les fréquences maximales de L1 et L2. Cependant, une fois encore le carré n'est pas parfaitement conservé. En effet, les résistances de Pull up et de polarisation ralentissent le chargement de la lentille vue comme un condensateur au premier ordre. De plus, le problème de la double alimentation persiste. Cet amplificateur nous a cependant permis de faire de nombreux tests avec les lentilles L1 et L2.

3- Amplificateur à pont en H

Les lentilles contenant un cristal liquide à faible $\Delta\epsilon$ ont besoin d'une amplitude de polarisation supérieure à 20V ce qui n'est pas possible avec l'amplificateur précédent. De plus, les fréquences d'utilisation sont également bien plus grandes.

Nous avons alors envisagé l'utilisation d'un pont en H pour la polarisation de telles lentilles. C'est de loin la solution la plus élégante, elle est push-pull, garantie la symétrie du signal, tout en offrant une amplification qui n'oblige pas à doubler la tension de polarisation.

Peu de pont en H existant sont réellement fait pour des charges capacitives ce qui nécessite des courants de charges très grand pour une forte réactivité.

De plus, ils sont difficiles à contrôler. En effet, les signaux de contrôle doivent être parfaitement synchronisés et non recouvrant. En aucun cas les deux transistors d'une même branche ne doivent être passants en même temps. Cela provoquerait un court-circuit et de détériorerait le driver et potentiellement la lentille. La figure A6-3 présente le montage d'un pont en H et les signaux de contrôle associés.

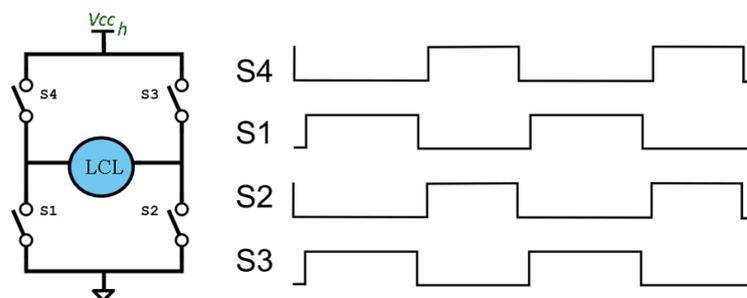


FIGURE A6-3 : MONTAGE PONT EN H ET SIGNAUX DE CONTROLE

La difficulté pour réaliser un tel montage dans le cadre de cette thèse est qu'il faut combiner vitesse et puissance. La solution retenue fut d'utiliser des demi-ponts EPC9004C intégrant des transistors Gan. Ils supportent des tensions jusqu'à 200V et ont un slew rate de l'ordre de 300V/ μ s. Ils sont donc parfaitement adaptés pour générer des signaux en haute tension et haute fréquence. Ces demi-ponts ont l'inconvénient de nécessiter deux alimentations, une comprise entre 7 et 11V pour faire fonctionner la logique intégrée, la seconde pour la haute tension de polarisation des lentilles. En revanche, il suffit de leur fournir une PWM pour que cela fonctionne. Pour créer un pont en H il faut deux demi-ponts et des signaux de commande

en opposition de phase. Afin de généré ces signaux non recouvrant, nous avons utilisé deux portes NOR et une porte inverseuse (Figure A6-4).

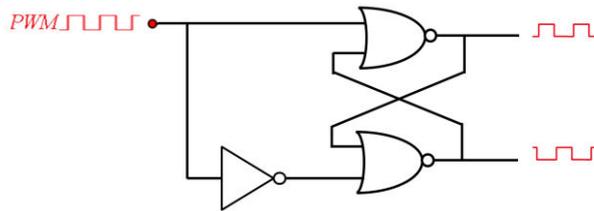


FIGURE A6-4 : LOGIQUE COMBINATOIRE POUR LA GENERATION DE DEUX PWM NON RECOUVRANTES

La figure A6-5 est une photo du driver expérimental.

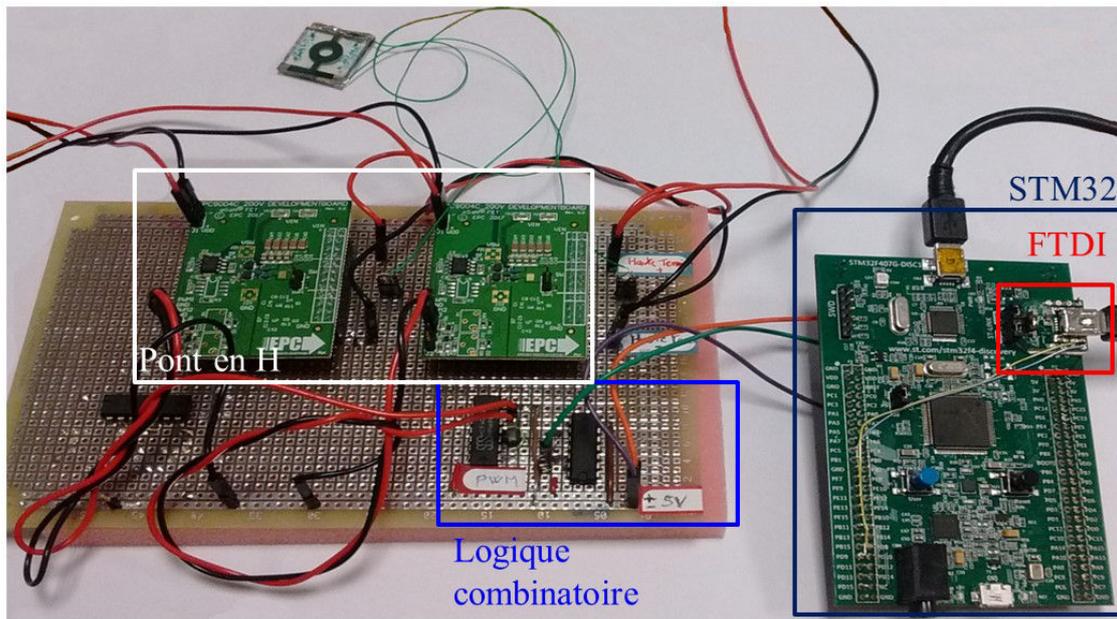


FIGURE A6-5 : PHOTO DU DRVER EXPERIMENTAL DE LENTILLE A CRISTAUX LIQUIDES SANS LES ALIMENTATIONS.

L'acquisition de la profondeur d'une scène en plus de son image est une caractéristique souhaitable pour de nombreuses applications qui dépendent de l'environnement proche. L'état de l'art dans le domaine de l'extraction de profondeur propose de nombreuses méthodes, mais très peu sont réellement adaptées aux systèmes embarqués miniaturisés. Certaines parce qu'elles sont trop encombrantes en raison de leur système optique, d'autres parce qu'elles nécessitent une calibration délicate, ou des méthodes de reconstructions difficilement implantables dans un système embarqué. Dans cette thèse nous nous concentrons sur des méthodes à faible complexité matérielle afin de proposer une solution algorithmique et optique pour réaliser un capteur permettant à la fois d'extraire la profondeur de la scène, de fournir une évaluation de pertinence de cette mesure et de proposer des images focalisées en tout point. Dans ce sens, nous montrons que les algorithmes du type Depth from Focus (DfF) sont les plus adaptés à ces contraintes. Ce procédé consiste à acquérir un cube d'images multi-foyers d'une même scène pour différentes distances de focalisation. Les images sont analysées afin d'annoter chacune des zones de la scène d'un indice relatif à sa profondeur estimée. Cet indice est utilisé pour reconstruire une image nette en tout point.

Nous avons travaillé sur la notion de netteté afin de proposer des solutions peu complexes, uniquement basées sur des additions et comparaisons, et de fait, facilement adaptables pour un portage sur une architecture matérielle. La solution proposée effectue une analyse bidirectionnelle de contraste local puis combine les meilleures estimations de profondeur en fin de traitement. Elle se décline en trois approches avec une restriction de la complexité de plus en plus forte et ainsi une aptitude de plus en plus marquée pour l'embarqué. Pour chaque méthode, des cartes de profondeurs et de confiances sont établies, ainsi qu'une image totalement focalisée constituée d'éléments issus de l'ensemble du cube multi-foyers. Ces approches sont comparées en qualité et en complexité à d'autres méthodes de l'état de l'art de complexité similaire. Une architecture est proposée pour une implantation matérielle de la solution la plus prometteuse.

La conception de ces algorithmes soulève le problème de la qualité d'image. Il est en effet primordial d'avoir une évolution remarquable du contraste ainsi qu'une invariance de la scène lors de la capture du cube multi-foyers. Un effet très souvent négligé dans ce type d'approche est le zoom parasite provoqué par la lentille responsable de la variation de la mise au point. Ce zoom de focalisation fragilise l'aspect invariance de la scène et provoque l'apparition d'artefacts sur les trois informations Profondeur, Image et Confiance. La recherche d'optiques adaptées au DfF constitue donc un second axe de ces travaux. Nous avons évalué des lentilles liquides industrielles et des lentilles modales expérimentales à cristaux liquides nématiques conçues durant cette thèse. Ces technologies ont été comparées en termes de rapidité, de qualité d'image, d'intensité de zoom de focalisation engendré, de tension d'alimentation et enfin de qualité des cartes de profondeur extraites et des images totalement focalisées reconstruites.

La lentille et l'algorithme répondant le mieux à cette problématique DfF embarqué ont ensuite été évalués via le portage sur une plateforme de développement CPU-GPU permettant l'acquisition d'images et de cartes de profondeurs et de confiances en temps réel.

Mots clef : Capteur d'image 3D, Acquisition de la profondeur, Reconstruction d'image totalement focalisée, Imagerie All-in-Focus

Acquiring the depth of a scene in addition to its image is a desirable feature for many applications which depend on the near environment. The state of the art in the field of depth extraction offers many methods, but very few are well adapted to small embedded systems. Some of them are too cumbersome because of their large optical system. Others might require a delicate calibration or processing methods which are difficult to implement in an embedded system. In this PhD thesis, we focus on methods with low hardware complexity in order to propose algorithms and optical solutions that extract the depth of the scene, provide a relevance evaluation of this measurement and produce all-in-focus images. We show that Depth from Focus (DfF) algorithms are the most adapted to embedded electronics constraints. This method consists in acquiring a cube of multi-focus images of the same scene for different focusing distances. The images are analyzed in order to annotate each zone of the scene with an index relative to its estimated depth. This index is then used to build an all in focus image.

We worked on the sharpness criterion in order to propose low complexity solutions, only based on additions and comparisons, easily adaptable on a hardware architecture. The proposed solution uses bidirectional local contrast analysis and then combines the most relevant depth estimations based on detection confidence at the end of treatment. It is declined in three approaches which need less and less processing and thus make them more and more adapted for a final embedded solution. For each method, depth and confidence maps are established, as well as an all-in-focus image composed of elements from the entire multi-focus cube. These approaches are compared in quality and complexity with other state-of-the-art methods which present similar complexity. A hardware implementation of the best solution is proposed.

The design of these algorithms raises the problem of image quality. It is indeed essential to have a remarkable contrast evolution as well as a motionless scene during the capture of the multi-focus cube. A very often neglected effect in this type of approach is the parasitic zoom caused by the lens motion during a focus variation. This "focal zoom" weakens the invariance aspect of the scene and causes artifacts on the depth and confidence maps and on the all in focus image. The search for optics adapted to DfF is thus a second line of research in this work. We have evaluated industrial liquid lenses and experimental nematic liquid crystal modal lenses designed during this thesis. These technologies were compared in terms of speed, image quality, generated focal zoom intensity, power supply voltage and finally the quality of extracted depth maps and reconstructed all in focus images.

The lens and the algorithm which best suited this embedded DfF issue were then evaluated on a CPU-GPU development platform allowing real time acquisition of depth maps, confidence maps and all in focus images.

Keywords: 3D Image-sensor, Depth acquisition, All Focus image generation, Depth Map, Confidence Map, Liquid crystal lens