



HAL
open science

plenoptic imaging: from visible light to X-rays

Charlotte Herzog

► **To cite this version:**

Charlotte Herzog. plenoptic imaging: from visible light to X-rays. Image Processing [eess.IV]. Université de Bordeaux, 2020. English. NNT : 2020BORD0116 . tel-03129863

HAL Id: tel-03129863

<https://theses.hal.science/tel-03129863>

Submitted on 3 Feb 2021

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.

THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET
D'INFORMATIQUE

par **Charlotte Herzog**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : INFORMATIQUE

Plenoptic imaging: from visible light to X-rays

Date de soutenance : 24 septembre 2020

Devant la commission d'examen composée de :

Xavier GRANIER	Professeur, Institut d'Optique	Directeur
Ombeline DE LA ROCHEFOUCAULD	Ingénieure PhD, Imagine Optic ...	Co-directrice
Céline LOSCOS	Professeure, Univ. Reims	Présidente
Daniel MENEVEAUX	Professeur, Univ. Poitiers	Rapporteur
David ROS	Maître de conf., Univ. Paris-Saclay	Rapporteur

Titre Imagerie plénoptique: de la lumière visible aux rayons X

Résumé

L'imagerie plénoptique est une technique basée sur l'acquisition des informations spatiales et angulaires des rayons lumineux provenant d'une scène. A partir d'une seule acquisition, un traitement numérique des données permet diverses applications comme la synthèse d'ouverture, le changement de point de vue, la refocalisation à différentes profondeurs, voire une reconstruction en 3D de la scène. L'imagerie plénoptique est beaucoup étudiée dans le visible. La transposition du visible dans le domaine des rayons X est un réel défi. L'imagerie plénoptique permettrait une imagerie 3D en rayons X à partir d'une seule acquisition. Cela aiderait à réduire fortement la dose absorbée par l'échantillon, par rapport à la tomographie qui nécessite une centaine de vues.

Dans cette thèse, nous considérons une caméra plénoptique constituée d'une lentille principale, d'une matrice de microlentilles et d'un détecteur. Deux configurations optiques distinctes, constituées de ces trois éléments, sont présentées dans la littérature : la caméra plénoptique *traditionnelle* et celle *focalisée*. La seule différence se trouve dans l'ajustement des distances entre les éléments optiques. L'observation d'une continuité entre ces deux configurations nous a amené à établir un système unique d'équations permettant leur conception optique, ainsi que l'expression théorique des résolutions associées. Ces résolutions ont été validées expérimentalement dans le visible. De plus, l'étude de l'évolution du contraste en fonction de la profondeur a montré que le contraste diminue quand on s'éloigne d'une position privilégiée intrinsèque à la configuration. C'est un résultat important car il pourrait affecter la qualité de l'image reconstruite et l'extraction de la profondeur.

Nous avons aussi travaillé sur les algorithmes de refocalisation préexistants, développés indépendamment pour chaque configuration. Nous avons élaboré un nouvel algorithme valide pour les deux configurations. Ce dernier est basé sur les distances physiques entre les éléments optiques, et permet une refocalisation à une distance arbitraire de la caméra. Tout d'abord, nous avons défini une nouvelle paramétrisation entre les espaces objet et capteur, en établissant la relation matricielle qui régit le trajet d'un rayon lumineux à l'intérieur de la caméra. Cette relation permet de reprojeter les données acquises par le capteur dans l'espace objet, et ainsi de reconstruire une image pixel par pixel à la profondeur choisie. En inversant les équations, nous avons montré qu'il était possible de créer des images plénoptiques synthétiques. La reconstruction de ces données synthétiques nous a permis de valider la cohérence des résultats après reconstruction, et de quantifier la précision de ce nouvel algorithme.

Cet algorithme permet de reconstruire séparément chaque plan de profondeur. Dans chacun d'entre eux, les éléments physiques qui appartiennent réellement à ce plan sont nets, alors que les objets des plans adjacents sont

floos. Nous utilisons cette propriété de contraste pour extraire l'information de profondeur dans les images refocalisées. Nous avons sélectionné plusieurs méthodes provenant du domaine de « depth from focus » et avons étudié leurs efficacités sur nos images.

Dans le cadre d'une collaboration européenne, nous avons construit la première caméra plénoptique dans les rayons X au synchrotron PETRA III. Grâce au travail réalisé pendant cette thèse, nous avons choisi les configurations optiques les plus adaptées aux optiques disponibles et aux caractéristiques du faisceau. Nous avons réalisé le montage de la caméra, acquis des images plénoptiques en rayons X, refocalisé ces images avec notre algorithme, et vérifié les résolutions optiques. Les méthodes de « depth from focus » appliquées sur les images refocalisées ont permis de retrouver la profondeur attendue. Ce travail correspond aux premières images acquises avec une caméra plénoptique en rayons X.

Mots-clés

Plénoptique, Champ lumineux, Rayons X, Imagerie 3D, Refocalisation

Laboratoire d'accueil

Laboratoire Photonique, Numérique, Nanosciences LP2N (UMR 5298) CNRS

Imagine Optic

Title Plenoptic imaging: from visible light to X-rays

Abstract

Plenoptic imaging is a technique that acquires spatial and angular information of the light rays incoming from a scene. After a single acquisition, numerical data treatment allows various applications such as synthetic aperture, changing viewpoint, refocusing at different depths, and consequently 3D reconstruction of the scene. Visible plenoptic has been widely studied. However, transposition from visible to X-rays has never been done and remains challenging. X-ray plenoptic would be beneficial to the X-ray imaging panorama. A single acquisition should be sufficient to reconstruct a volume, against hundreds for X-ray tomography that is the today reference in 3D X-ray imaging.

In this thesis, we consider plenoptic camera composed of a main lens, a microlens array and a detector. So far, two different configurations have been developed: the traditional and the focused plenoptic setups. Although these configurations are usually studied separately, they only differ by the distances between the optical elements. These two configurations were studied in details to choose the most suitable for X-ray imaging, considering the constraints of X-ray optics. We observed a full continuity between the two systems. Therefore, we extended the previous works to more general formulas about optical configuration and theoretical resolutions. Theory about resolution along the depth axis was refined, as depth reconstruction and extraction are the main interest of X-ray plenoptic. Specific study was done on the evolution of contrast along depth as being a key parameter for depth reconstruction. We realized that contrast decreases when moving away from a privileged depth. This is important to consider as it can affect image reconstruction and quality of depth extraction.

We also worked on refocusing algorithms. The refocusing algorithms are usually developed for each configuration separately. We wanted to go beyond this separation. We developed a new algorithm valid for all configurations. Moreover, our algorithm is based on real distances between the optical elements, allowing image reconstruction at any distance from the plenoptic camera. We defined a new parameterization between object and sensor spaces. Using geometrical optics, we calculated the matrix transformation between the two spaces. This allows back-projecting data from the acquired raw image to the object space, and reconstructing the pixels one by one, until the whole object. Reversing the process, we were able to simulate the process of image acquisition, and create synthetic plenoptic data. Reconstruction of these data was used to quantify the accuracy of the novel algorithm and prove its consistency.

The refocusing algorithm allows reconstructing the depth planes one by one. Each refocused plane contains information about the whole 3D scene

that has to be disentangled. In each depth plane, the elements physically present at this depth are intrinsically sharp, whereas the ones located at other depths are blurred. We used this contrast property to extract depth from the refocused stack. We tested several existing methods derived from the field of depth from focus and studied their efficiency when applied to our images.

In collaboration with European teams, we realized the first X-ray plenoptic camera that was tested at P05 beamline of PETRA III synchrotron. Based on the theoretical work developed in this thesis, we defined the best optical configuration, mounted the plenoptic camera, acquired X-ray plenoptic images, numerically refocused them using our new algorithm and verified the experimental resolutions and contrasts. Depth from focus techniques applied on the refocused stack allow to retrieve the expected depth plane. These are the first images acquired with an X-ray plenoptic camera.

Keywords Plenoptic, Light-field, X-rays, 3D imaging, Refocusing

Laboratory

Laboratoire Photonique, Numérique, Nanosciences LP2N (UMR 5298) CNRS

Imagine Optic

Acknowledgments

First of all, I wish to thank my supervisors Xavier Granier and Ombeline de La Rochefoucauld for having accompanied me during these last years. I am aware of the high quality of the supervision they both provided throughout my thesis, in terms of science and personal investment. I specially thank Xavier for having guided me in answering the algorithmic questions raised by plenoptic reconstruction. The clarity of his ideas and his mathematical way of thinking taught me a lot, as well as his kindness and patience. A special thank to Ombeline for having been my day-to-day accompaniment from my first discovery of plenoptic to the thesis defence. She introduced me to plenoptic and X-ray imaging and gave me the opportunity to experiment on synchrotron and discover the world of research. Ombeline pushed me to go beyond my limits by writing articles and going to conferences abroad. Her continuous support and remarkable pragmatism have been of great help during the full duration of my thesis.

This thesis was conducted within Imagine Optic. I wish to thank the board of Imagine Optic for having accepted to finance my research. I feel grateful to Samuel Bucourt, who first proposed me the subject of plenoptic and motivated me to start this thesis. This experience would not have been possible without meeting him four years ago. Within Imagine Optic, I benefited from scientific and technical advice from many colleagues, including Guillaume Dovillaire, Fabrice Harms, Xavier Levecq, Pauline Treimany, Laura Oudjedi and Martin Piponnier. I also thank the administrative and IT support for all the necessary work related to the good progression of this thesis. During these last years, I had the pleasure of visiting Imagine Optic Orsay several times, either to mount experiments, participate in trainings or present my work. I also took part in many festive activities, such as the "trainees' parade", the Christmas dinners or June barbecues, and the 20 years celebration of Imagine Optic. My everyday working life was shared with the Bordeaux team, Laura, Martin and Ombeline. It was a pleasure to share the office with Laura, her everyday good mood and the interesting discussions we had together. Martin taught me a lot about optics and coding in Python, as well as ecological commitment. It was an everyday pleasure to complete my thesis in such a warm and peaceful environment.

I thank the LP2N laboratory for having hosted my thesis. During my PhD,

I was part of the Manao team, located at Inria Bordeaux-Sud-Ouest. Being part of Manao was the occasion to meet other researchers and PhD students, and experimenting another working environment. I really appreciated the support and advice provided when needed, for example the scientific discussions with Loïs Mignard-Debise, Gaël Guennebaud, Corentin Cou, and many others. I will keep good memories of the team atmosphere, especially the last team seminar with beers and mölkky on the beach.

This thesis was partially conducted within the VOXEL project. Taking part of this project has been a valuable experience, as it was the occasion to meet international researchers, post-docs and PhD students, and learning from their technical knowledge and living the functioning of an international collaboration in between several European structures. I thank the principal investigators of VOXEL for having mounted this project and for all the interesting scientific discussions: Marta Fajardo (coordinator), Joost Batenburg, Alessia Cedola, Hamed Merdji, Eduardo Oliva and Philippe Zeitoun. I wish to express my gratitude to Philippe for his explanations in the field of X-ray imaging and his help in analysing the results of the X-ray plenoptic experiment. Thanks to Nicola Vigano, with who I learned about tomography and its link with plenoptic, as well as code structure and optimisation. Thanks to Elena Longo and Imke Greving, who allowed us to realize the first X-ray plenoptic experiment at synchrotron PETRA III in DESY, which represents the practical application of my theoretical work. This experiment was also the occasion to discover the working life and night shifts at a synchrotron. Thanks to Ying Li, whose collaboration allowed to gain a significant time in processing data from the DESY experiment.

I wish to also thank the jury members for having read my thesis manuscript, given fruitful comments and attended my defence in such particular conditions. I thank the thesis committee for having followed the progress of my thesis over years.

Finally I would like to thank my family, friends and all those who have supported me, helped me grew up and gain confidence during my thesis years in Bordeaux.

Funding

This thesis was partially funded by the VOXEL project. The VOXEL project is a RIA-FET-OPEN project, which received funding from European Union's Horizon 2020 research and innovation program under grant agreement N° 665207.

This work was partially funded by the Région Nouvelle-Aquitaine and the European Union FEDER under the XPULSE project: “Development of an imaging system using X-rays based on ultra-short intense laser for applications in breast cancer imaging”.

This thesis was partially funded by the Association Nationale Recherche Technologie (ANRT) under CIFRE agreement n° 2017/0401.

List of publications

- Charlotte Herzog, Xavier Granier, Fabrice Harms, Philippe Zeitoun, and Ombeline de La Rochefoucauld. Study on contrast variations with depth in focused plenoptic cameras. *Optics Letters*, 44(19):4825, 2019a. doi: 10.1364/OL.44.004825. URL <https://hal.archives-ouvertes.fr/hal-02301709>.
- Charlotte Herzog, Ombeline de La Rochefoucauld, Fabrice Harms, Philippe Zeitoun, and Xavier Granier. Unifying the refocusing algorithms and parameterizations for traditional and focused plenoptic cameras. ICCP 2019 - IEEE International Conference on Computational Photography 2019, May 2019b. URL <https://hal.inria.fr/hal-02096378>. Poster.
- Nicola Viganò, Pablo Martínez Gil, Charlotte Herzog, Ombeline de la Rochefoucauld, Robert van Liere, and Kees Joost Batenburg. Advanced light-field refocusing through tomographic modeling of the photographed scene. *Optics Express* 27(6):7834, 2019. URL <https://doi.org/10.1364/OE.27.007834>
- Nicola Viganò, Henri Der Sarkissian, Charlotte Herzog, Ombeline de La Rochefoucauld, Robert Van Liere, and Kees Joost Batenburg. Tomographic approach for the quantitative scene reconstruction from light field images. *Optics Express*, 26(18):22574, 2018. doi: 10.1364/OE.26.022574. URL <https://hal.inria.fr/hal-02042417>.
- Charlotte Herzog, Guillaume Dovillaire, Xavier Granier, Fabrice Harms, Xavier Levecq, Elena Longo, Lois Mignard-Debise, Philippe Zeitoun, and Ombeline de La Rochefoucauld. Comparison of reconstruction approaches for plenoptic imaging systems. In *Unconventional Optical Imaging*, volume 10677 of *Proceedings of SPIE*, pages 106772U:1–106772U:11, Strasbourg, France, April 2018. SPIE. doi: 10.1117/12.2306800. URL <https://hal.inria.fr/hal-01819610>.

Résumé long

Un champ lumineux (*light-field* en anglais) est une fonction représentant la luminance le long de chaque rayon lumineux dans un espace. Il est défini pour chaque position 3D et chaque direction 2D, avec parfois des composantes additionnelles telles que le temps, la longueur d'onde et la polarisation de la lumière. Dans les techniques d'imagerie classiques, comme la photographie ou la microscopie, le champ lumineux est projeté sur le plan 2D du capteur. Au lieu de cela, l'imagerie de champ lumineux cherche à capturer les cinq dimensions (ou plus) de la luminance, grâce à un système optique spécifique. Une fois l'acquisition effectuée, ce champ lumineux peut être manipulé à l'aide d'algorithmes dédiés. À partir d'une seule image de ce champ, diverses applications sont possibles. Les plus courantes incluent l'extraction en profondeur, la refocalisation à différentes profondeurs et la reconstruction 3D de la scène.

L'imagerie de champ lumineux a connu des développements importants au cours des dernières décennies. Ces avancées ont été réalisées pour un spectre allant du visible à l'infrarouge. Plusieurs types de caméras ont été développés. La plus courante est la caméra plénoptique, basée sur l'ajout d'une grille de micro-lentilles entre la lentille principale et le capteur. Cette caméra est compacte, relativement facile à réaliser, et ce à des coûts raisonnables, ce qui explique son succès dans la communauté scientifique.

La possibilité d'une reconstruction en 3 dimensions à partir d'une seule acquisition est d'un grand intérêt, et pourrait être appliquée à d'autres domaines spectraux que celui de la lumière visible. Dans le domaine des rayons X, la technique d'imagerie de référence est la tomographie, qui est utilisée par exemple dans l'imagerie médicale, le contrôle non destructif en industrie ou bien l'étude du patrimoine culturel. La tomographie permet de reconstruire une imagerie volumétrique de l'échantillon, offrant de très bonnes résolutions et une qualité d'image élevée. Son principe consiste à combiner des images 2D acquises à différents angles de vue, et donc cette technique nécessite un grand nombre d'acquisitions pour être efficace (généralement plusieurs centaines ou milliers d'acquisitions). Par conséquent, une dose importante de rayons X est absorbée par l'objet à imager, ce qui est nocif pour le corps humain mais également pour les échantillons biologiques ou les spécimens fragiles. Les différentes acquisitions nécessitent des rotations du système d'imagerie ou bien de l'échantillon, et donc un équipement coûteux pour garantir une pré-

cision angulaire suffisante. De plus, la tomographie suppose que l'échantillon reste statique pendant tout le processus d'acquisition, ce qui exclut l'imagerie d'un échantillon en mouvement. L'imagerie plénoptique adaptée aux rayons X permettraient de réduire la dose absorbée ainsi que le temps d'acquisition, puisqu'une seule acquisition serait nécessaire au lieu de quelques centaines ou milliers. L'appareil serait plus compact et la reconstruction 3D plus rapide.

Cette thèse s'est déroulée dans le contexte du projet VOXEL dont le but est de réaliser de l'imagerie plénoptique en rayons X, en construisant deux prototypes dans des domaines spectraux différents. Le premier prototype est prévu dans les rayons X mous (0,44 keV) qui sont habituellement utilisés pour l'imagerie de petits échantillons biologiques tels que les cellules. Le deuxième prototype concerne l'imagerie du petit animal dans le domaine des rayons X durs (environ 10 keV). L'aspect instantané de l'imagerie plénoptique est intéressant pour étudier des animaux vivants. Les travaux réalisés lors de cette thèse fournissent les principaux éléments nécessaires pour construire la première caméra plénoptique en rayons X.

Nous présentons ici les principaux résultats de cette thèse concernant la transposition de l'imagerie plénoptique du domaine visible aux rayons X.

Le chapitre 1 présente les notions fondamentales concernant le champ lumineux, sa paramétrisation ainsi que l'intérêt de pouvoir le capturer et le manipuler. Dans cette thèse, nous nous concentrons sur l'imagerie plénoptique, dont l'élément principal est une grille de microlentilles, positionnée entre une lentille principale et un capteur. Deux configurations différentes sont possibles, la configuration traditionnelle et la configuration focalisée, chacune ayant été présentée avec son algorithme spécifique. Dans ce document, nous mettons en évidence les points communs et différences entre ces deux configurations, mais également entre les deux algorithmes de reconstruction. Nous montrons une continuité entre les deux configurations optiques, et cette continuité se retrouve sur le plan numérique. Les deux méthodes de reconstruction se sont révélées être basées sur le même principe d'intégration des données angulaires, à position spatiale fixée. Cette comparaison permet de mieux comprendre les rôles distincts de la configuration et de l'algorithme dans les résolutions de l'image finale. Par conséquent, la configuration optique et la méthode numérique doivent être optimisées séparément afin d'améliorer les performances d'un système plénoptique.

Dans le chapitre 2, nous poursuivons la comparaison entre les deux configurations optiques. Nous reprenons d'abord les équations générales permettant la conception d'un système plénoptique, et incluons de nouvelles conditions garantissant un bon éclairage de l'échantillon. Ces nouvelles équations sont d'une importance capitale pour l'imagerie par rayons X, où l'éclairage se fait par transmission à travers l'objet. Puis nous établissons les formules de résolu-

tions d'une caméra plénoptique, en fonction des distances et des paramètres de la configuration physique (tels que les ouvertures numériques, les diamètres des lentilles ...). Ces formules permettent de prévoir la taille maximale de l'objet qui peut être imagé (champ de vue et profondeur de champ plénoptique), ainsi que la taille minimale du voxel dans l'image reconstruite (résolutions latérale et en profondeur). Nous présentons les formules de telle manière que les deux configurations traditionnelles et focalisée puissent être facilement comparées.

Ces équations sont ensuite validées expérimentalement. Nous avons construit deux caméras plénoptiques dans le visible, qui nous permettent d'acquérir nos propres données. Les expériences menées ont confirmé que nos formules de résolution latérale et de champ de vue sont appropriées. L'estimation des résolutions en profondeur est plus complexe car elle nécessite l'acquisition, la reconstruction et l'étude de plusieurs images à différentes profondeurs. Pour trouver un critère expérimental permettant de les quantifier, nous avons réalisé une étude sur le flou et sur la netteté de l'image en fonction de la profondeur. Cette validation expérimentale indique que les résolutions le long de la zone de profondeur sont affectées par des phénomènes inattendus, tels que les variations de contraste et de grandissement. En parallèle des expériences, nous avons réalisé des simulations d'imagerie plénoptique. Ces simulations nous ont aidés à comprendre que le grandissement et la diffraction interfèrent avec le flou de défocalisation. Les effets observés pourraient affecter les résolutions et le processus d'extraction de profondeur à partir des images reconstruites.

Après l'étude du design optique, le chapitre 3 s'intéresse maintenant à l'aspect numérique. Une image brute plénoptique n'est pas lisible telle quelle. Il est nécessaire d'utiliser un algorithme de refocalisation, qui permet de reconstruire une image 2D à une certaine distance de la caméra, choisie par l'utilisateur. La comparaison des deux algorithmes préexistants nous a convaincus qu'il était possible de les fusionner pour n'en former qu'un seul, valable pour les deux configurations traditionnelle et focalisée. Nous définissons un nouvel algorithme basé sur un raisonnement purement physique, permettant de refocaliser une image à n'importe quelle distance de la caméra, à une résolution choisie par l'utilisateur. Nous démontrons une nouvelle paramétrisation de la transformation d'un rayon à travers le système, conduisant à une équation qui fait le lien entre espace objet et espace capteur. Cette équation est ensuite utilisée pour reconstruire les pixels un par un, en projetant les rayons à travers les ouvertures de l'objectif principal et des micro-lentilles, et en considérant l'étendue spatiale des pixels du détecteur et ceux de l'image à reconstruire. Le processus inverse a également été implémenté, et permet de simuler des images brutes plénoptiques à partir d'un objet. Ces simulations permettent d'évaluer l'algorithme de reconstruction, en comparant les images simulées puis refocalisées avec l'objet initial. Les résultats prouvent que l'algorithme est plus précis que l'état de l'art. Il permet de reproduire numériquement l'apparition

progressive du flou lié à la défocalisation quand l'objet est hors focus. Cet algorithme permet de créer une pile d'images refocalisées à différentes profondeurs de la scène.

À partir de cette pile d'images, nous cherchons à extraire la profondeur des objets en utilisant des méthodes de "depth from focus". Dans le chapitre 4, nous nous intéressons à ce domaine qui exploite les propriétés de flou et de netteté dans l'image afin de retrouver les profondeurs des objets. Nous avons implémenté et testé différentes méthodes préexistantes de mesure de flou et d'extraction de profondeur sur des piles d'images réalisées à partir de un, deux ou trois plans superposés. Ce travail permet de sélectionner une mesure de flou efficace afin d'extraire la bonne valeur de profondeur des plans. L'ajout d'une étape de filtrage gaussien permet de lisser les résultats et de les améliorer dans le cas d'un seul plan. Lors de plusieurs plans superposés, il faut au contraire enlever ce filtrage pour détecter le plan situé à l'arrière. Les résultats sont intéressants et montrent que la capacité d'extraction de profondeur dépend des objets présents dans les images et de plusieurs paramètres: la méthode utilisée, la présence du filtrage et la taille de la fenêtre sur laquelle la mesure de flou est faite.

Enfin, le chapitre 5 présente notre expérience plénoptique par rayons X. Les travaux théoriques, expérimentaux et numériques réalisés au cours de cette thèse nous donnent les outils nécessaires pour réaliser la première caméra plénoptique par rayons X. Nous présentons les spécificités des optiques en rayons X et leurs conséquences sur la conception d'une caméra plénoptique: la source, la lentille principale, la grille de micro-lentilles, le détecteur et les échantillons utilisés. Les formules de résolutions ont été utilisées pour choisir la meilleure configuration possible, malgré les faibles ouvertures numériques des optiques. L'éclairage par transmission à travers l'échantillon, spécifique aux rayons X, a également été pris en compte. Nous avons réalisé des images de deux mires 1951 USAF positionnées l'une derrière l'autre. À notre connaissance, ces images correspondent aux premières images plénoptiques dans les rayons X réalisées à ce jour. Nous montrons les images plénoptiques expérimentales et expliquons comment les interpréter. Nous utilisons ensuite notre algorithme de refocalisation pour reconstruire les images à différentes profondeurs. Ces images reconstruites nous permettent de valider les formules de résolution latérale et en profondeur, de champ de vue et de profondeur de champ plénoptique établies au chapitre 2. Nous avons testé les techniques de "depth from focus", qui nous ont permis d'extraire la bonne profondeur de la première mire, et de la séparer de la seconde mire et de l'arrière-plan. Nous prouvons ainsi que la théorie plénoptique, initialement développée dans le visible, est également valable dans le domaine des rayons X.

Pour conclure, cette thèse a permis de définir un cadre théorique pour la conception d'une caméra plénoptique, de démontrer une approche plus précise pour la reconstruction et d'étudier la faisabilité d'une approche "depth-from-

focus" pour l'extraction de la 3D. Tout ceci a permis de transposer l'imagerie plénoptique du visible vers les rayons X, et de réaliser la première caméra plénoptique par rayons X. Ce cadre théorique, algorithmique et expérimental est une base pour des travaux futurs dans l'imagerie plénoptique s'étendant au-delà du spectre visible.

Contents

Contents	xix
List of Figures	xxiii
Introduction and Motivation	1
1 Review and comparison of previous work	5
1.1 Light-field imaging	5
1.1.1 The light-field	5
1.1.2 Two-planes parameterization	5
1.1.3 The phase-space diagram	6
1.1.4 Acquisition of the light-field	8
1.1.5 Applications of light-field imaging	9
1.1.6 Application to 3D reconstruction	9
1.2 Traditional and focused plenoptic cameras	10
1.2.1 Plenoptic cameras	10
1.2.2 The traditional plenoptic camera	11
1.2.3 The focused plenoptic camera	13
1.2.4 Conclusion on optical configurations	16
1.3 Refocusing algorithms: traditional vs focused	17
1.3.1 The traditional refocusing algorithm	17
1.3.1.1 Phase-space diagram	17
1.3.1.2 Refocusing process: shift-and-sum algorithm . .	19
1.3.1.3 Example of refocused images	21
1.3.2 The refocusing algorithm for the focused configuration .	22
1.3.2.1 Construction of the phase-space diagram	22
1.3.2.2 The different rendering methods	26
1.3.2.3 Re-expressing the rendering method	29
1.3.3 Comparison of the two algorithms	32
1.3.4 Application on image reconstruction	32
1.3.5 Conclusion on refocusing algorithms	34
1.4 Previous studies on resolutions	34
1.5 Synthesis of the comparison	35

1.6	X-ray plenoptic imaging	36
1.6.1	Published work	37
1.6.2	Challenges of X-ray imaging	37
1.7	Conclusion	38
2	Optical system: equations and resolutions	39
2.1	Optical design	39
2.1.1	Continuity between the optical setups	40
2.1.2	Aperture matching condition	41
2.1.3	Conditions on illumination	43
2.2	Theoretical studies on resolutions	45
2.2.1	Field of view (FOV)	47
2.2.2	Lateral Resolution (res_{lat})	49
2.2.3	Spatio-angular sampling (N_u and N_s)	51
2.2.4	Depth resolution (res_{depth})	52
2.2.5	Plenoptic depth of field (DOF_{pleno})	53
2.2.6	Summary of the formulas	56
2.3	Experimental validation: methodology	58
2.3.1	Presentation of the two setups	58
2.3.2	Chosen configurations and expected resolutions	61
2.3.3	Distance measurement	62
2.4	Experimental validation of the resolutions	63
2.4.1	Field of view	64
2.4.2	Lateral resolution	65
2.4.3	Depth resolution	68
2.4.3.1	Depth resolution in the HASO 1.0	69
2.4.3.2	Depth resolution in the Stingray 1.0	71
2.4.3.3	Depth resolution in the Stingray 2.0	73
2.4.4	Plenoptic depth of field	75
2.4.4.1	Methodology for drawing the contrast curve	75
2.4.4.2	Experimental validation of plenoptic depth of field	76
2.4.4.3	Study of DOF_{pleno} with a different bar width	78
2.5	Simulation of the impact of magnification and diffraction on depth of field	81
2.5.1	Simulations	81
2.5.2	Comparison with experimental data	85
2.6	Conclusion	88
3	New parameterization and refocusing algorithm	89
3.1	Introduction	89
3.2	Propagation from object-space to sensor-space	90
3.2.1	Previous parameterizations and models	90

3.2.2	2D derivation of the affine transformation of rays	91
3.2.3	Matrix equation of the 4D transformation	93
3.3	Proposed refocusing approach	93
3.3.1	Principle and comparison with previous methods	93
3.3.2	Our approach	95
3.3.3	Simulation of a plenoptic raw image	96
3.4	Implementation details	97
3.4.1	Approximations on the apertures	97
3.4.2	Faster access to contributing pixels	98
3.4.3	Correction of reconstruction vignetting	98
3.5	Numerical results	99
3.5.1	Reconstruction of a single plane	100
3.5.2	Performance analysis	102
3.5.3	Reconstruction for another optical configuration	103
3.5.4	Objects at different depths	105
3.5.5	Reconstruction of a real plenoptic image	106
3.6	Analysis of the limitations	108
3.6.1	On computation time	108
3.6.2	Limits of plenoptic reconstruction	109
3.7	Conclusion	109
4	Experimenting depth extraction with depth from focus	111
4.1	Principle of depth from focus	111
4.2	Chosen methods and implementation	113
4.2.1	Blur measures	113
4.2.2	Depth extraction	114
4.3	Tests on synthetic images	116
4.3.1	Detection of a single plane	117
4.3.1.1	Blur measures	117
4.3.1.2	Blur measures - comparison of behaviours in- side the volume	121
4.3.1.3	Depth extraction	123
4.3.2	Detection of two separated planes	124
4.3.2.1	Blur measures	125
4.3.2.2	Depth extraction	127
4.3.3	Detection of three overlapping planes	129
4.3.3.1	Blur measures	129
4.3.3.2	Depth extraction	132
4.4	Conclusion	133
5	Demonstration of X-ray plenoptic	135
5.1	The experiment and optics	135
5.1.1	The X-ray source	136

5.1.2	The main lens	137
5.1.3	The Micro-Lens Array	138
5.1.4	The imaging system	139
5.1.5	The sample	139
5.1.6	Optical configurations	139
5.2	First results and preprocessing	141
5.2.1	Analysis of a raw image	141
5.2.2	Acquisition procedure and preprocessing	143
5.2.3	Image stitching	144
5.3	Refocusing of the acquired data	146
5.3.1	Configuration 1.0	146
5.3.2	Configuration 2.0	147
5.3.2.1	Effects of the stitching	148
5.3.2.2	Refocusing with algorithm 2.0	148
5.3.2.3	Refocusing with our algorithm	149
5.3.2.4	Study of depth resolution and plenoptic depth of field	151
5.3.2.5	Refinement of depth resolution	153
5.3.2.6	Note on the distances	156
5.4	Depth extraction	157
5.5	Conclusion	160
	Conclusion	163
	Future work	165
	Bibliography	169

List of Figures

1.1	Representation of a ray in the two possible parameterizations: its radiance can be expressed as $LF(x, y, z, \theta, \phi)$ with 3 spatial and 2 angular coordinates, or $LF(s, t, u, v)$ in the two-planes parameterization.	6
1.2	Illustration of the phase-space diagram of a ray bundle converging on a single point, depending on its depth (images from [Lam15]).	7
1.3	Design of a traditional plenoptic camera, composed of a main lens, an array of micro-lenses (μ LA) and a sensor. The main lens forms the image of the object directly on the plane of the μ LA. The μ LA then distributes the different angular information over the pixels of the sensor.	11
1.4	Travel of the light from the object-plane (or world focal plane) to the sensor plane, showing how the angular distribution is transmitted through the optical system (image from [FT14]).	12
1.5	Illustration of the aperture matching condition (images from [Ng06]).	13
1.6	Example of a raw image acquired with a traditional plenoptic camera. The letter "a" was placed at depth z_0^{trad} , resulting in uniform white, gray or black sub-images in the raw image.	14
1.7	Design of a focused plenoptic camera, composed of a main lens, an array of micro-lenses (μ LA) and a detector. The main lens forms the image of the object on the intermediate image plane. The μ LA then acts as a relay imaging system, with each micro-lens imaging only a part of the object on the sensor. We introduce distance $c = z_1 + a$ as the physical distance between the main lens and μ LA.	14
1.8	Example of two raw images acquired with a focused plenoptic camera, imaging the object shown in Figure 1.6a. As expected, there is an inversion of the sub-images between the Keplerian and Galilean modes. Note that the two images were acquired with slightly different magnifications.	16

1.9	Schematic 2D representation of a plenoptic system. The spatial sampling is done by the μ LA (l or s -axis) with a step Δs , whereas the angular sampling is achieved by the sensor (p or u -axis) with a step Δp corresponding to Δu on the main lens. The refocusing process uses a new parameterization s' on the plane at distance $z'_1 = \alpha z_1$ from the main lens.	18
1.10	Schematic representation of the traditional plenoptic camera with the raw image, the phase-space diagram and the sensor- μ LA representation.	19
1.11	Traditional plenoptic reconstruction principle in the sensor- μ LA diagram. Points with a red circle correspond to pixels acquired by the detector while points circled with yellow represent missing data that has to be created by interpolation. To get the reconstructed image, the integration along the line of equation $l(p) = l' + p \frac{\alpha(1-\frac{1}{a})}{r}$ has to be performed, as illustrated by the arrows. We represent here the reconstruction with $\alpha > 0$, which corresponds to a negative slope. The lateral resolution of the reconstructed image equals the micro-lens size.	21
1.12	Refocused images from the same raw image of a rotated ruler. Changing the refocusing parameter α results in a change of the depth plane of reconstruction, which is visible through the left-to-right displacement of the in-focus areas across the images. . .	22
1.13	Representation of the successive matrices used to establish the phase-space diagram.	23
1.14	Step by step construction of the phase-space diagram for the focused plenoptic camera, showing how the light-field is sampled spatially (s -axis) and angularly (u -axis).	26
1.15	Illustration of the patch size, P , the refocusing parameter in the refocusing algorithm for the focused configuration.	27
1.16	Basic rendering algorithm. In each micro-lens, only 2 pixels are selected to form the reconstructed image ($P = 2$).	28
1.17	Rendering with blending algorithm. Each pixel of the refocused image is formed by integrating over several pixels of the raw image. All the pixels of the raw image are used.	29
1.18	Schematic representation of the focused plenoptic camera with the raw image, the phase-space diagram and the sensor- μ LA representation.	30
1.19	a) Illustration of patches composed of 3×3 pixels and their correspondence with neighboring sub-images. b) Representation of the "rendering with blending" method in the sensor- μ LA diagram, which consists in an integration along lines of equation $l(p) = (p_0 - p)/P$. We consider a reconstruction in the case $a > 0$, leading to $P > 0$ and a negative slope.	31

2.1	Using the same setup and distances, images were acquired at several depth of the object "ab". We recognize typical raw images of the focused Galilean (scheme <i>a</i> and image <i>b</i>), traditional (scheme <i>c</i> and image <i>d</i>) and focused Keplerian case (scheme <i>e</i> and image <i>f</i>).	41
2.2	Calculation of the period and size of the sub-images on the sensor.	42
2.3	Conditions on illumination of the system: both size and divergence of the source should be large enough to illuminate the whole main lens and the array of N_2 micro-lenses of pitch d_2 . . .	44
2.4	Schematic representation of the four resolutions, with the lateral resolution res_{lat} , the field of view FOV , and along the optical axis, the depth resolution res_{depth} and the plenoptic depth of field DOF_{pleno}	46
2.5	The three possible definitions of the field of view: the strictest (green), middle (red) and loosest (blue) definitions. The light rays show how the formulas of the three fov definitions are established (equations 2.14-2.15-2.16), and the reconstructed image illustrates the different levels of vignetting.	48
2.6	Schematic representation of the traditional (top) and the Keplerian focused (bottom) cameras, showing how the lateral resolutions are calculated from Δ_s in the intermediate object plane.	50
2.7	Reasoning for the geometrical term of the depth resolution: it corresponds to the depth of focus of the main lens in the intermediate image plane, i.e. the depth along which the projected rays stay inside Δ_s	53
2.8	The geometric term of the plenoptic depth of field in a traditional configuration: the ray beam coming from the main lens is divided into N_u beams acquired by different pixels of size Δ_{peff} .	54
2.9	The geometric term of the plenoptic depth of field in a focused configuration: the ray beam coming from the main lens is divided into N_u beams corresponding to the N_u illuminated micro-lenses. It also corresponds to the depth of field of a single pixel of size Δ_{peff} , projected back in intermediate image plane, and then in object plane.	54
2.10	The experimental HASO setup composed of an object (left), the main lens (middle), and the HASO wavefront sensor (right). . .	58
2.11	The experimental Stingray setup composed of the illumination system and an object (left), the main lens with a diaphragm (middle), the μLA and the Stingray camera (right).	59
2.12	The optical elements used in the two setups.	59

2.13	The 1951 USAF resolution test target. The three sets of bars studied in the experiments are highlighted, with the indication of the width of the bars.	60
2.14	Example of a raw image acquired in configuration Stingray 2.0, $a > 0$, where we can see $gr2 - el1$ (bar width = $125\mu m$).	63
2.15	Full reconstructed images used to estimate the experimental field of view in each configuration. The image from the HASO 1.0 dataset has clearly a larger FOV than the three other cases. The blue rectangle in 2.15a shows the only bars that are visible in the images from the Stingray setup because of its reduced FOV	65
2.16	Images used to estimate the lateral resolution in each of the four experimental configurations. The different images from the Stingray setup were assembled to virtually restore the whole test target, compensating for their small FOV . The hatched area in 2.16a represents the area where the contrast was measured for this bar width ($gr2 - el1$).	66
2.17	Curves of the relative contrast along bar width, allowing to quantify the lateral resolution in each configuration. The theoretical lateral resolution varies between $47\mu m$ and $50\mu m$ depending on the configuration. The reconstructed images show $gr3 - el3$ (bar width = $49.5\mu m$) in the four configurations.	67
2.18	Depth resolution for the HASO 1.0: series of refocused images at the same z_0 ($\alpha = 1$) and contrast curves extracted on elements $gr2 - el1$, $gr2 - el5$ and $gr3 - el3$	70
2.19	Depth resolution for the Stingray 1.0 configuration: series of images refocused at the same z_0^{trad} ($\alpha = 1$) for $gr2 - el1$ (a) and $gr3 - el3$ (b), and the contrast curves extracted on these elements along depth (c).	72
2.20	Depth resolutions for the four configurations, with the series of refocused images of $gr2 - el1$, and the corresponding contrast curves. This comparison allows to extrapolate the results obtained in the HASO 1.0 and Stingray 1.0 cases to the Stingray 2.0 $a > 0$ and $a < 0$	74
2.21	Methodology for building the contrast curve to study the plenoptic depth of field in the Stingray 1.0 configuration. Each image is reconstructed with various refocusing parameters then the maximal contrast is selected. The curves for the different P values can be related to the depth resolution at a depth different from z_0 . The decrease in height and width is the sign of lower lateral and depth resolutions when moving away from z_0	76
2.22	Plenoptic depth of field based on the $gr3 - el3$ in the HASO and Stingray 1.0 configurations.	77

2.23	Plenoptic depth of field based on the <i>gr2 – el1</i> in the HASO, the Stingray 1.0 and 2.0, $a > 0$ configurations.	79
2.24	A focused plenoptic system. Red rays come from the depth plane z_0^{foc} whereas green rays correspond to another depth to illustrate the Circle of Confusion (CoC) due to defocus.	82
2.25	Simulated 1D raw image of the test object versus position r on the detector with the effect of defocus only (<i>green</i> ; contrast = 1), with diffraction only (<i>red</i> ; contrast = 0.59) and with both effects combined (<i>blue</i> ; contrast = 0.52), calculated at a depth $z_0 = 240\text{mm}$, compared to the image with no effect (<i>dotted black</i> ; contrast = 1).	83
2.26	Simulated 1D raw image with the three effects versus position r on the detector for different depths z_0 : $z_0^{\text{foc}} = 233\text{mm}$ (<i>blue</i>), 231mm (<i>dotted orange</i>), 235mm (<i>dotted green</i>), and 240mm (<i>purple</i> , same depth as in Fig. 2.25).	84
2.27	Contrast curves versus depth: with defocus and magnification only (<i>green</i>), with diffraction and magnification only (<i>orange</i>) and with the three effects combined (<i>blue</i>). Depths $z_0^{\text{foc}} = 233\text{mm}$ and $z_0^{\text{trad}} = 228.6\text{mm}$. Note that the small oscillations are due to pixel discretization of the signal.	85
2.28	a) Raw sub-images acquired at planes: $z_0 = 223.5, 228.5, 233, 236, 242\text{mm}$; b) Corresponding reconstructed images; c) Comparison of the simulated contrast curve (<i>blue</i>) and experimental contrast curve measured on raw sub-images (<i>pink</i>) and on reconstructed images (<i>gray</i>). The gap located around $z_0^{\text{trad}} = 228.6\text{mm}$ is due to the large magnification that makes the bars exceeding the sub-image size.	86
3.1	Notations for equation derivation of the ray transformation though the optical system.	91
3.2	Principle of the refocusing approach: (1) Transformation of the bundle of rays \mathbf{A}_{mn} from the reconstructed pixel p_{mn} passing by the aperture of the main lens into the sensor-space bundle of rays \mathbf{A}_{mnl} through the micro-lens $\mu\mathbf{L}_{kl}$. (2) This bundle of rays is intersected with the bundle of rays \mathbf{B}_{ijkl} between the sensor's pixel p_{ij} and micro-lens $\mu\mathbf{L}_{kl}$ to compute the contribution of p_{ij} on p_{mn}	95
3.3	Phase-space diagram of Fig. 3.2.	96
3.4	Original virtual knife-edge (left) and zoom in the plenoptic raw image simulated by our algorithm in configuration 1 (right). The red squares represent the micro-images due to the micro-lenses.	101

3.5	Reconstructions of the knife-edge in configuration 1 with our algorithm (left) and using algorithm from [Ng06] (right). The two top figures are crops of the reconstructed images at $z = 106.20mm$ (size 10×40). The bottom figures show the evolution of a single row extracted from the reconstruction images, with depth from $z = 105.2mm$ to $z = 106.5mm$ in steps of $0.01mm$. The rows extracted from the images at $z = 106.20mm$ are highlighted in red.	101
3.6	Root Mean Square Error (RMSE) between original and reconstructed images of the knife-edge in configuration 1: using our algorithm (3.6a), using algorithm from [Ng06] (3.6b), and comparison of the two curves (3.6c).	103
3.7	Reconstruction of the knife-edge in configuration 2 with our algorithm: the plenoptic image simulated by our algorithm with red squares delimiting the sub-images (3.7a), evolution of a row extracted at regular steps of $0.1mm$ between depth $z = 470mm$ and $z = 510mm$ (3.7b), and the corresponding RMSE curve between original and reconstructed images (3.7c).	104
3.8	Test with two letters a and b . We use the two images on the left separately, located respectively at $z_0^a = 105.64mm$ and $z_0^b = 105.80mm$ to generate synthetic plenoptic images that we combine in the final result shown partially on the right.	105
3.9	Reconstructed images from the two letters test scene (top) with the corresponding RMSE curves (bottom) between reconstructed images and original letter a (blue), and between reconstructed images and original letter b (dotted blue).	106
3.10	Reconstruction of the pencil holder [GL10b] from the unmodified data available on-line. A partial view of the data is shown on the left. In the middle row, we show the uncorrected version of the result of our algorithm exhibiting vignetting effects (top) together with the image of the correction factors (bottom). Once divided by the correction factor, the final image on the right is obtained.	107
3.11	Illustration of the grid effects on the reconstruction of a point in configuration 1 : evolution of a row between $z = 104mm$ and $z = 108mm$ (3.11a), and the refocused image at $z = 107mm$ (3.11b).	109
4.1	Principle of depth from focus methods: 1) Blur is measured on each plane of the stack separately, then 2) using these measurements, depth is extracted for each (x,y) pixel. This process results in a depth map.	113

4.2	Relative positions of the synthetic depth planes: the Left plane (green) corresponds to plane number 14 ($z = 14$), the Right plane to $z = 30$ and the Third plane to $z = 35$. The size of a pixel is $3 \times 3 \mu m$ inside a plane, and the spacing between two planes is $10 \mu m$	116
4.3	Original, plenoptic and reconstructed images of the Left plane, due to the simulation and refocusing algorithms. The blue lines in Figure 4.3a and 4.3c show the row corresponding to the longitudinal sections in Figure 4.3d and 4.3e. The green lines in Figure 4.3d represent the positions of the profiles of Fig. 4.5 and 4.6.	118
4.4	Longitudinal sections of blur measures applied on the reconstructed stack of the Left plane.	119
4.5	Comparison of depth profiles from the original and reconstructed stacks at 3 different locations defined in Figure 4.3d, as well as blur measures with or without Gaussian filtering for the location $y=30$. Each curve was independently normalized to have a maximal value at 1. In Figure 4.5d the scale of depth plane z has been enlarged to better distinguish the different curves.	120
4.6	Comparison of the action of the <i>Energy Laplacian</i> , the <i>Ring Difference Filter</i> and the <i>Variance</i> on three different profiles: $y=25$ (dashed), $y=30$ (solid) and $y=50$ (dotted).	122
4.7	Depth extraction methods tested on the Left plane alone.	124
4.8	Original and reconstructed stack in the case of two separated planes, and the results of the Ring Difference Filter with and without filtering. The green lines in Figure 4.8a represent the positions of the profiles of Fig. 4.9.	125
4.9	Profiles for the case with both Left and Right planes: comparison of the same locations $y = 68$ and $y = 73$ taken in the original and reconstructed stacks, and the actions of the five blur measures with and without Gaussian filtering.	127
4.10	Depth extraction methods tested on the case of the Left and Right planes together.	128
4.11	Original and reconstructed stack in the case of three overlapping planes, and the results of the Ring Difference with and without filtering. The green lines in Figure 4.11a represent the positions of the profiles of Fig. 4.12.	130
4.12	Profiles for the case with three overlapping planes: comparison of the same locations $y=30$ and $y=68$ taken in the original and reconstructed stacks, and the results of the five blur measures with and without Gaussian filtering.	131
4.13	Depth extraction methods tested on the case of the three overlapping Left, Right and Third planes together.	132

5.1	Aerian view of PETRA III buildings at DESY.	136
5.2	Locations of the optical components in the mounted experiment.	136
5.3	Illustration of a <i>Fresnel Zone Plate</i> (left), and the mounted main lens just after the sample (right).	137
5.4	Image of the μ LA (top left) and its blocker (top right) acquired with a visible microscope, and how they were mounted together for the experiment (bottom).	138
5.5	The camera and its relay imaging system. We can also notice the μ LA and its blocker before it.	139
5.6	The two USAF 1951 tests patterns used in the experiments, imaged with the main lens and camera only. TP1 (left) is $1.3mm$ closer to the main lens than TP2 (right). Notice the heterogeneous illumination.	140
5.7	A raw image as acquired by the imaging system, and a zoom showing the first (blue) and zero (yellow) orders due to the <i>FZP</i>	142
5.8	Schematic representation of the preprocessing, consisting in averaging followed by normalization.	143
5.9	A raw image after preprocessing. Some sub-images are highlighted in blue. The zoomed areas show variations of magnification in the images of the lines with the same line width.	144
5.10	Principle of stitching: the μ LA is shifted to create new sub-images. We obtain 36×36 sub-images instead of 9×9 initially.	145
5.11	Datasets 1.0 and 2.0 after preprocessing, extraction of the sub-images, and combination of the 36×36 sub-images from the stitching.	146
5.12	Dataset 1.0 reconstructed using algorithm 2.0 and our algorithm.	147
5.13	Effects of the stitching process on the refocused images in configuration 2.0.	148
5.14	Set of refocused images acquired with the configuration 2.0, and using algorithm 2.0, with various patch sizes P	149
5.15	Refocused image at $z = 128.77mm$ in configuration 2.0 using our algorithm.	150
5.16	Comparison of the same dataset of configuration 2.0 refocused with our algorithm at different depths.	151
5.17	Comparison of the datasets 2.0, 2.a and 2.b refocused at the same depth $z = 128.77mm$	152
5.18	Zooms on the refocused images from Figure 5.16. The size of the extracted areas is drawn in orange in Figure 5.16g.	154
5.19	Contrast curve extracted from the refocused stack of the X-ray plenoptic experiment.	155

5.20 Illustration of two blur measures applied on the refocused stack. The blue line in Figure 5.20a shows the row that was extracted in the five other images. Figure 5.20b corresponds to the z-axis variations in the reconstructed stack. Figures 5.20c and 5.20e show the results of the Ring Difference Filter with and without filtering, whereas Figures 5.20d and 5.20f show the results for the Variance method. Figures 5.20g and 5.20h present the profiles extracted from Figures 5.20e and 5.20f at $y = 60$, whose position is shown in green in Figure 5.20b. 158

5.21 Depth extraction of the refocused stack using the Variance method as blur measure and Max as depth extraction method: without the Gaussian filtering step (Fig. 5.21b) and with the filtering step (Fig. 5.21d). TP1 is located at $z = 128.77mm$ whereas TP2 is around $\sim z = 130mm$, outside the refocused depth range. 159

Introduction and Motivation

A light-field is a function representing the radiance along each ray that is, for each 3D position and each 2D direction in space. Sometimes, it may also include time, light polarization or spectral variations. Classical imaging techniques, such as photography or microscopy, correspond to projection of a light-field on a 2D sensor plane. Instead, light-field imaging aims at capturing the whole 5D light-field using an appropriate optical design. Once acquired, a light-field could be manipulated using dedicated algorithms. From a single light-field image, a variety of applications could be addressed. The most popular include depth extraction, refocusing at various depths and 3D reconstruction of the scene from a single acquisition.

Light-field imaging has undergone important developments over the past decades. The technical advances have been done in the visible and infrared ranges. Various designs have been developed to achieve light-field acquisition, the most common one being plenoptic imaging, based on the addition of an array of micro-lenses between the main lens and the sensor. This design is compact and relatively easy to implement at reasonable costs, explaining its popularity among the scientific community.

The possibility of 3D reconstruction from a single acquisition is of great interest, and could be applied to other fields than visible light. In the X-ray range, the gold standard imaging technique is computerized tomography, whether imaging a human body, a biological or cultural heritage samples. Tomography allows volumetric imaging of the sample, providing high resolutions and image quality. However, it consists in combining 2D images acquired at various viewing angles, and thus requires a high number of acquisitions to be efficient (usually between hundreds and thousands acquisitions). This induces a high irradiation dose absorbed by the object being imaged, which could be harmful for human body and potentially damageable for biological samples or fragile specimens. This also implies rotations of either the machine or the sample, which necessitates a large equipment with precise angular accuracy. Moreover, the long acquisition time requires the sample to stay static during the whole imaging process, which excludes following a moving sample. Transposing light-field imaging techniques in X-rays would help reducing the irradiation dose and acquisition time, since only one acquisition is needed. The device would be more compact and the resulting 3D reconstruction would be

limited by the acquisition time of a single image.

X-ray plenoptic imaging is the objective of the *VOXEL* project, in which the current thesis took place. *VOXEL* is an European project, whose name is the abbreviation for Volumetric medical X-ray imaging at extremely low dose. The objective was to build two prototypes at different X-ray energies. The first setup was in soft X-rays (0.44 keV) which is known as efficient to image small biological samples such as cells. The second prototype concerns small animal imaging (around 10 keV). The instantaneous aspect of plenoptic would be interesting for dynamic studies on living animals. The work presented in this thesis provides elements that have been studied in order to achieve the first X-ray plenoptic camera.

The goal of this thesis is to study the feasibility of adapting light-field imaging to the X-ray range. X-ray imaging is a complex field, with the X-ray optics being highly technical and expensive (see Chapter 1). Adapting plenoptic to the X-ray range requires a good understanding of plenoptic techniques and being able to predict their resolutions (see Chapter 2). Refocusing of the images necessitates an accurate algorithm, based on the parameters and distances of the physical setup (see Chapter 3). This algorithm allows to generate a re-focused stack, from which depth could be extracted (see 4). The theoretical work on the physical and numerical aspects of plenoptic has led to build and test a X-ray plenoptic setup, which allows to better understand the possibility of plenoptic for biological imaging (see Chapter 5).

Chapter 1 presents fundamental notions concerning light-field imaging and its applications. It focuses on plenoptic imaging, based on a microlens array. Two different designs are possible, the traditional and the focused configurations, each of them being associated with its specific algorithm. We review them and highlight the common characteristics and differences between the two. We also present the specificities of X-ray optics and the consequences on the design of an X-ray plenoptic system.

In Chapter 2, the comparison between the two optical configurations is deepened. We first review the general equations for establishing a plenoptic system, and include additional conditions for a good illumination of the sample, of particular importance for X-ray imaging. Then the resolutions of both configurations are studied, depending on the distances and parameters of the physical setup. We establish theoretical formulas by a geometrical reasoning, and validate them using our own experiments conducted in the visible range. This experimental validation indicates that the resolutions along depth are affected by unexpected phenomena, such as contrast and magnification variations. We present a simulation of these effects in a plenoptic system to better understand them.

Whereas Chapter 2 provides elements to optimize the optical setup, Chap-

ter 3 concentrates on the numerical aspect of refocusing. A plenoptic raw image should be reconstructed with an appropriate algorithm to fully exploit the acquired information. We here present our work towards a new physically-based algorithm, that takes into account the spatial extents of the apertures of both microlenses and sensor pixels. We develop a new parameterization of the ray transformation through the system, which is the basis of our refocusing algorithm. The method is evaluated on plenoptic images simulated using the reverse counterpart of the refocusing process. The results indicate that the algorithm is accurate.

This refocusing algorithm offers an accurate method to create a refocused stack of images reconstructed at different depths in the scene. From this stack, depth could be extracted using depth from focus methods. The purpose of Chapter 4 is to study the efficiency of depth from focus techniques on numerically refocused images. We implement published several methods of blur measurement and depth extraction, and conduct a progressive analysis on simulated images of one, two, then three depth planes.

Finally, Chapter 5 presents our X-ray plenoptic experiment. The lessons learned from the previous chapters give us the necessary tools to build a plenoptic system in the X-ray range. We introduce the physical elements of the setup and how they might impact the plenoptic system: the source, the main lens, the array of micro-lenses, the sensor and the samples used. They constraint us in the choice of a plenoptic configuration. We show experimental plenoptic images and explain how to interpret them. We refocus the images and present the consequences of the optical configuration and refocusing algorithm on image resolutions and quality. We verify the coherence of plenoptic theory in the X-ray range, that was previously established in the visible. We finally generate the refocused stack and experiment with depth from focus techniques. We show that a plane could be extracted at its correct depth.

We conclude on the main contributions of this thesis and the feasibility of X-ray plenoptic imaging.

Chapter 1

Review and comparison of previous work

1.1 Light-field imaging

1.1.1 The light-field

The notion of light-field dates back to Faraday’s ideas presented in 1846 in [Far46]. In 1936, Gershun formalized this notion and introduced the term *light-field* [Ger39]. The light-field represents the radiance along each ray, i.e. for each position and direction in space. Depending on the authors, the light-field is also called plenoptic [AB91] or lumigraph [GGSC96], in the sense that it contains a complete representation of all the information carried by the light in a volume. This way, being able to capture the entire light-field coming from a scene provides the possibility to perfectly describe it and, with an appropriate algorithm, potentially reconstruct it in 3 dimensions (3D). In comparison, traditional photography captures a 2D image corresponding to a projection of the light-field on the sensor.

1.1.2 Two-planes parameterization

The positions and directions of the light-field function are described by 5 coordinates, 3 for its location in 3D space and 2 for the angular components (see Fig. 1.1). The light-field LF is thus defined for each $ray(x, y, z, \theta, \phi)$. To simplify, we express it as: $LF(ray) = LF(x, y, z, \theta, \phi)$ (see Fig. 1.1). This notion was taken up later by Adelson and Bergen in [AB91], who extended the light-field to a 7D function, including variables of time and wavelength.

A property of radiance is that it remains constant along a ray when there is no blocker or absorption [McC14]. This property implies that the choice of the parameterization of the rays has no impact on the value of the light-field function. Levoy and Hanrahan [LH96] presented a simpler model of parameter-

ization than the 5D coordinates. The light rays flowing inside a homogeneous volume (such as free-space) could be parameterized by their intersections with two planes located at two different depths in the volume (see Fig. 1.1). The 2D coordinates on both planes is sufficient to describe the geometry of each ray. When the light is monochromatic and constant over time, the light-field is reduced to a 4D function : $LF(s, t, u, v)$. This model is called the *two-planes parameterization*.

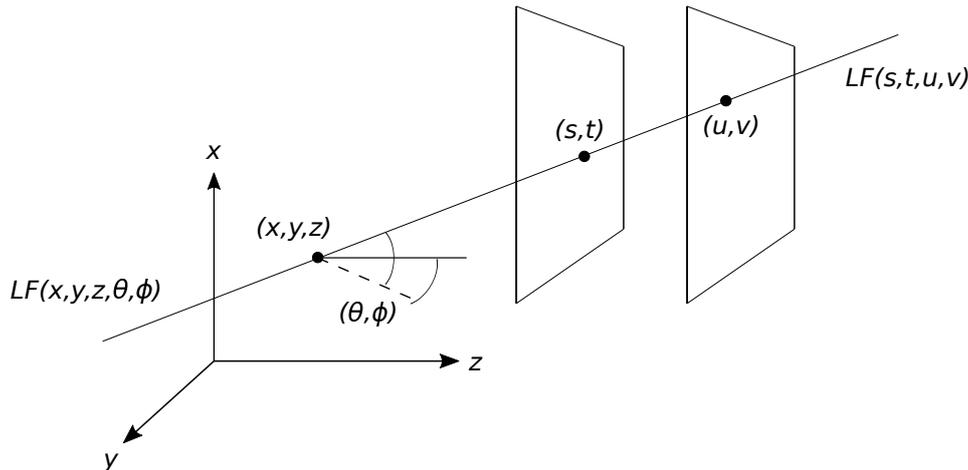


Figure 1.1 – Representation of a ray in the two possible parameterizations: its radiance can be expressed as $LF(x, y, z, \theta, \phi)$ with 3 spatial and 2 angular coordinates, or $LF(s, t, u, v)$ in the two-planes parameterization.

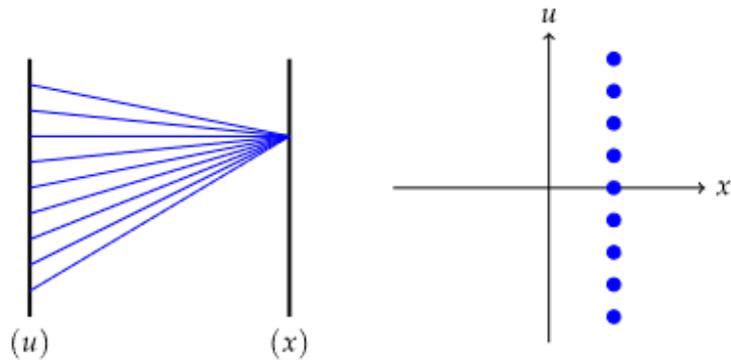
In the majority of imaging systems, the two transverse axes (x and y -axes) could be decoupled. The 4 coordinates of the light-field can be separated into 2 sets of 2 coordinates, each set describing one of the transverse axes. When considering linear propagation, the behaviour of the light rays along each axis is independent. Thus the 2 sets of coordinates could be treated separately, and the 4D light-field function could be factorized as a 2×2 D function. As a result, the parameterization and equations could be done in 2D.

1.1.3 The phase-space diagram

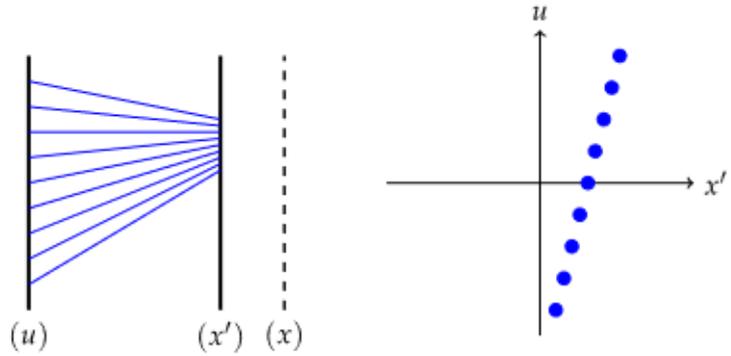
This two-planes parameterization could be visualized as a phase-space diagram, whose axes correspond to the coordinates of the two-planes parameterization [Lam15]. In the context of light-field imaging, the parameterization is usually chosen so that one plane is responsible for the spatial sampling of the light-field, when the other plane is responsible for the angular sampling. Consequently, in 2D the phase-space diagram typically displays the spatial information along the horizontal axis and the angular information along the vertical axis.

Figure 1.2 presents an example of a two-plane parameterization together with its associated phase-space diagram. The main idea is that a 2D light

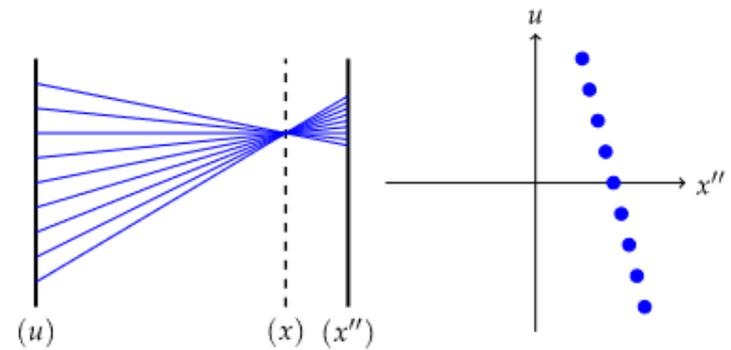
ray in the physical space of the parameterization becomes a single point in the phase-space diagram. Thus the phase-space representation reduces the description of the physical space, providing a synthetic overview of the light-field. Another interesting property is that the rays converging to the same point in the physical space happen to be aligned in the phase-space diagram [IRMD16]. This allows to easily represent a bundle of rays focusing on the same point as



(a) Ray bundle focusing on the second plane of parameterization.



(b) Ray bundle focusing after the second plane of parameterization.



(c) Ray bundle focusing before the second plane of parameterization.

Figure 1.2 – Illustration of the phase-space diagram of a ray bundle converging on a single point, depending on its depth (images from [Lam15]).

a simple line in the phase-space diagram. The line is rotated depending on the depth of the point of convergence (see Fig. 1.2). By extension, a surface area in the phase-space diagram corresponds to the ray bundle between two apertures on the two-planes parameterization. For a detailed derivation or interpretation of the phase-space diagram, the reader might refer to [IRMD16; Lam15].

The phase-space diagram is a useful tool to visualize and study how an imaging system samples the light-field. It could be used to compare different optical designs and estimate their resolutions [LGC12]. It is important to notice that the phase-space diagram depends on the positions of the two planes of parameterization. Moving one of the two planes along the optical axis results in a shift of the light-field in the phase-space diagram [Lam15]. This impacts the shape and size of the sampling.

1.1.4 Acquisition of the light-field

Light-field imaging, also called plenoptic [AW92] or integral imaging [Lip08b], corresponds to the techniques attempting to capture the light-field coming from a scene. The challenge of light-field imaging is to enable separating the spatial and angular components of each light ray, as a representation of its position and direction before the imaging system.

Actually, prior to the theoretical conceptualisation of light-field, there had already been some experimental attempts of achieving light-field imaging. The first practical implementations can be dated to 1903 [Ive03] and 1908 [Lip08a]. In [Ive03], two apertures allow to separate the light rays, acting like the left and right eyes in human vision. In [Lip08a], the author presented a prototype based on a home-made array of small lenses. But it is only more recently that light-field imaging techniques have been further developed, thanks to the technical advances in micro-lenses fabrication and in computation possibilities. In 1992, Adelson and Wang presented a plenoptic camera based on an array of micro-lenses, together with its corresponding algorithm [AW92]. They observed similar performances than with stereo imaging in terms of depth extraction, while the setup was more compact and needed less calibration.

Ng was the one that made light-field imaging popular in the more recent years. In his thesis [Ng06], he provided a complete description of the *traditional* plenoptic camera, considering both experimental considerations and refocusing algorithm. In his camera, an array of micro-lenses plays the role of separating the angular and spatial components of the incoming light rays. It has led to an interest of the scientific community in light-field imaging and since that the research on the topic has considerably increased. In 2008 Lumsdaine and Georgiev presented another plenoptic camera with a different design [LG08a; LG09]. They called it *focused* plenoptic camera in contrast to the *traditional* camera presented by Ng. These two cameras are both based on micro-lens arrays but vary in terms of distances between the opti-

cal elements. Other camera designs were also developed, based on different optical components: the camera array [WJV⁺05], various systems based on prisms [CTDL11], pinholes [CCC⁺16], coded apertures [VRA⁺07] or even mirrors [MRK⁺13]. For more details on the history of plenoptic cameras, the reader can refer to [Lam15; IRMD16]. Deeper reviews on the variety of light-field imaging systems can be found in [WILH11; CDMBV16]. Mignard-Debise et al [MDRI17] have shown that light-field imaging systems can all be modeled as camera arrays, making the link between these different systems.

1.1.5 Applications of light-field imaging

The main interest of light-field imaging is certainly the variety of its possible applications. Due to its comprehensive nature, the information contained in a single acquisition enables novel applications than in photography. After the acquisition has been made, the light-field can be manipulated using appropriate algorithms to render different effects. Among the possible applications can be mentioned synthetic aperture, changing viewpoint of the scene, depth estimation or numerical refocusing at various depths [Ng06; GL10a; FCS14; HSVD17]. In particular, 3D rendering is an interesting application. From a single acquisition it is possible to reconstruct an object [PW12; HSVD17] or to give a sense of 3D [LNA⁺06]. For example, several light-field microscopes have been developed [LNA⁺06; BGY⁺13; MDI15], and present interesting results in terms of resolutions and depth extraction. This ability could be compared to other techniques such as stereo cameras [FH15] or tomography systems [KSW02], which require more equipment, calibration and acquisition time. Light-field imaging represents an alternative in terms of compactness and acquisition time.

1.1.6 Application to 3D reconstruction

In this thesis, light-field imaging is studied as a potential technique to achieve 3D reconstruction from a single acquisition in the X-ray range. In order to extract depth from the acquired light-field image, several numerical approaches validated in the visible range can be considered to be transposed in the X-ray domain. The refocusing algorithms associated with the two configurations [Ng06; GL10a] provide a stack of 2D refocused images containing simultaneously the in-focus elements of the refocused depth and the out-of-focus elements of adjacent depths.

Some methods allows direct depth extraction from the raw light-field image, without the need for the intermediate step of refocusing. The possibility to transpose any light-field imaging system to a camera array [MDRI17] justifies the use of algorithms coming from the field of stereo-imaging. Stereo-imaging consists in combining two or more 2D images of the same scene but taken from different angles of view, in order to extract depth of the scene [FH15]. Some of

the numerical methods have already been adapted to process light-field imaging [GL10b; THMR13; WER15; HSVD17] or X-ray imaging [DCH⁺19]. These methods are adapted for surface reconstruction, and might not be appropriate for 3D volumetric data.

The similarities between light-field imaging and tomography have been demonstrated by [LNA⁺06; BGY⁺13], and experimented by [CZL⁺19]. In the context of the *VOXEL* project, Vigano et al. showed that light-field imaging can be interpreted as limited-angle cone-beam tomography. They adapted tomography algorithms to reconstruct plenoptic data [VDH⁺18; VMH⁺19]. They were tested on numerical and real-world examples, acquired on the visible setups developed during this thesis (see Chapter 2). These algorithms allowed to perform depth estimation and volumetric reconstruction.

Another approach is to consider depth from focus techniques. They allow to estimate depth from a stack of 2D images refocused at regular depths of the scene [SN17; SJP⁺17]. Depth from focus methods are classically used in the field of microscopy imaging [MDP13; BCJ⁺13]. Using refocusing algorithms would allow to generate a stack of refocused images, from which depth could be extracted. This two-steps framework corresponds to the numerical approach considered in this thesis.

1.2 Traditional and focused plenoptic cameras

This thesis focuses on the feasibility of light-field cameras to X-ray imaging. We will study the types of light-field cameras that seem transposable to X-rays: the ones based on arrays of micro-lenses.

1.2.1 Plenoptic cameras

In the literature, the word *plenoptic camera* has been used specifically for cameras made with micro-lenses [LG08b; Hog18], in contrast to *light-field cameras* referring to any of the previously described techniques enabling to acquire the light-field (based on micro-lens, pinholes, mirrors...). On the other hand, the word *plenoptic* can be understood as a broader notion than the *light-field*, because its origin comes from assembling the latin word *plenus* (complete or full) to the word *optic* [AB91]. This ambiguity shows that these notions are not fully standardised, and this explains why *plenoptic* and *light-field* are often considered synonymous. In this thesis we will use preferentially the term of *plenoptic camera* to describe the one based on a micro-lens array.

Only three optical elements are needed to build a plenoptic camera: a main lens, an array of micro-lenses (μ LA) and a detector (see Figures 1.3 and 1.7). With only three optical elements, this type of camera is compact, easy to build and model. Its simple structure and the possibility to build a plenoptic camera

either from separated elements [ZLE⁺18] or from already existing cameras or microscopes [LNA⁺06; MDI15] explain its popularity compared to other light-field systems.

The problematic of plenoptic cameras lies in the placement of the three components: main lens, μ LA and sensor. Two particular configurations are noticeable: the *traditional* plenoptic camera also called *unfocused* or *plenoptic camera 1.0* [Ng06] and the *focused* plenoptic camera also called *plenoptic camera 2.0* [LG08a; LG09]. These two configurations contains the same optical elements (main lens, μ LA and sensor), they only vary in the distances between them. In both configurations, the same two equations govern the placement of the optical elements: the thin-lens equation and the aperture matching condition.

Figures 1.3 and 1.7 show the two configurations and present the names of the parameters and distances that will be used throughout the thesis. The configurations are presented in 2D for greater readability, but the names and corresponding equations could be extended to include the third dimensions. The main lens has a diameter d_1 and focal length f_1 . The μ LA is composed of N_2 micro-lenses of diameter d_2 and focal length f_2 . For simplicity, the micro-lenses are supposed to be adjacent, otherwise it is necessary to differentiate their apertures from their pitches (or periodicity) in some equations. The sensor is made of N_p pixels of size Δ_p .

1.2.2 The traditional plenoptic camera

The thin-lens equation determines the placement of the elements to guarantee that the image on the detector is sharp. In the traditional configuration (see Figure 1.3), the object and μ LA are positioned so that the two planes are

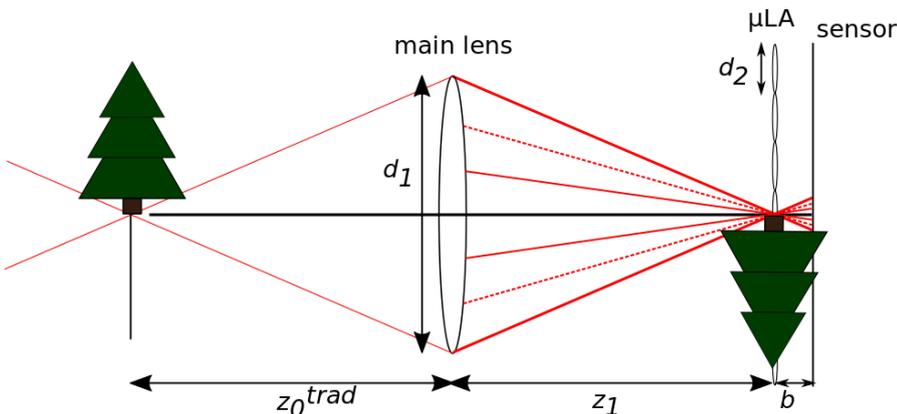


Figure 1.3 – Design of a traditional plenoptic camera, composed of a main lens, an array of micro-lenses (μ LA) and a sensor. The main lens forms the image of the object directly on the plane of the μ LA. The μ LA then distributes the different angular information over the pixels of the sensor.

optically conjugated according to:

$$\frac{1}{z_0} + \frac{1}{z_1} = \frac{1}{f_1} \quad (1.1)$$

with z_0 the distance between object and main lens, and z_1 the distance between main lens and μ LA. This way, the image of the object is located directly on the plane of the μ LA. The μ LA is responsible for separating the spatial components from the angular ones. Collecting the angular components is then done by the pixels of the sensor. To ensure correct sampling, the distance b between the μ LA and the detector is imposed to be equal to the focal length of the micro-lenses:

$$b = f_2 \quad (1.2)$$

Figure 1.4 shows the travel of light through a plenoptic system, with colors coding the directional information from the object-plane to the sensor. In the traditional configuration, the angular range of the data in object-space is split over the main lens, and the same distribution is observed across the pixels under each micro-lens on the sensor plane.

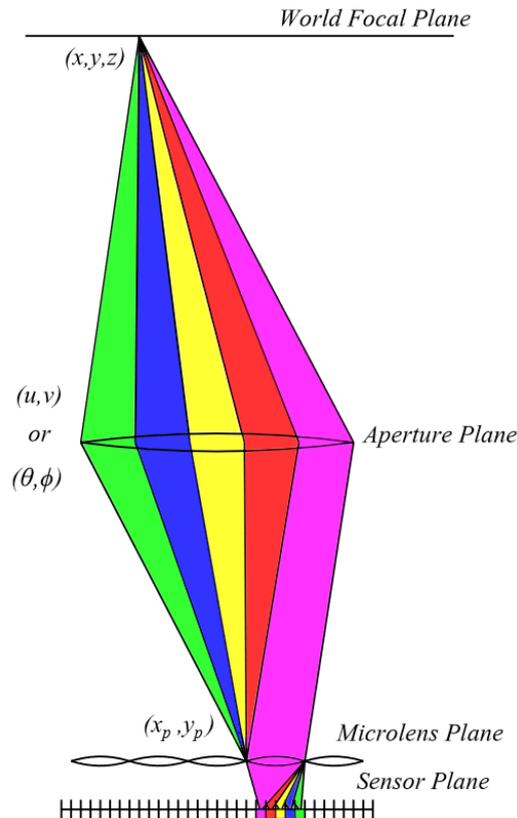
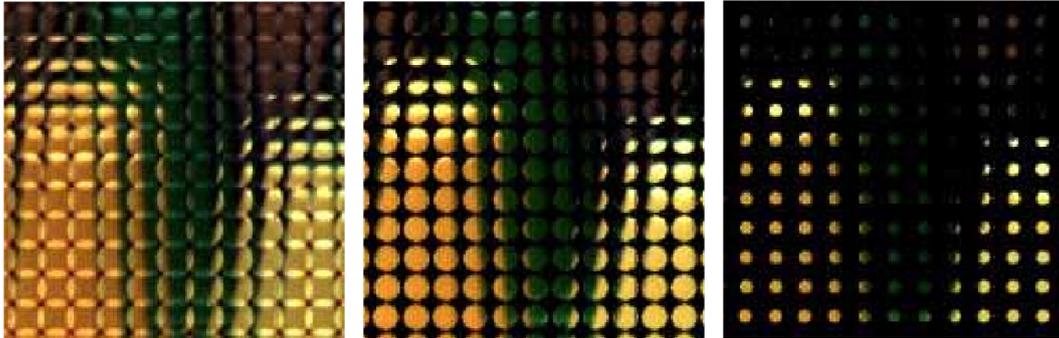


Figure 1.4 – Travel of the light from the object-plane (or world focal plane) to the sensor plane, showing how the angular distribution is transmitted through the optical system (image from [FT14]).

In addition to these two Equations 1.1-1.2, a third equation is needed, called aperture matching condition. The role of this equation is to optimize the use of the pixels on the sensor. Indeed the raw image acquired by the plenoptic camera is composed of a large number of sub-images corresponding to the micro-lenses of the μ LA. The size and periodicity of these sub-images depend on the parameters of the whole setup. In order to optimize the acquired data, the sub-images should be adjacent on the detector plane. According to [Ng06; LNA⁺06], the image-side numerical aperture of the main lens should match the image-side numerical aperture of the micro-lenses to ensure adjacent sub-images. Defining the numerical aperture as $NA = \text{aperture}/(2 \cdot \text{distance})$ [IO94], the condition can be written:

$$\frac{d_1}{z_1} = \frac{d_2}{f_2} \quad (1.3)$$

Figure 1.5 shows the impact of the aperture matching condition. Respecting this condition leads to adjacent sub-images on the sensor (Fig. 1.5b). Otherwise, the sub-images might be too small, resulting in a large number of unused pixels on the sensor (Fig. 1.5c). On the contrary, having too large sub-images generate overlaps and a loss of information (Fig. 1.5a).



(a) Overlapped sub-images. (b) Adjacent sub-images. (c) Small sub-images.

Figure 1.5 – Illustration of the aperture matching condition (images from [Ng06]).

Figure 1.6 shows an example of a raw image acquired with a traditional plenoptic camera. In this configuration, the spatial components of the light-field are sampled by the μ LA whereas the angular components are sampled by the pixels on the detector. Thus the sub-images could be interpreted as large pixels (spatial components) inside which the small intensity variations represent the angular components.

1.2.3 The focused plenoptic camera

The second alternative is the focused configuration (see Figure 1.7) [GL10a]. Contrary to the traditional one, the main lens does not project the image of the

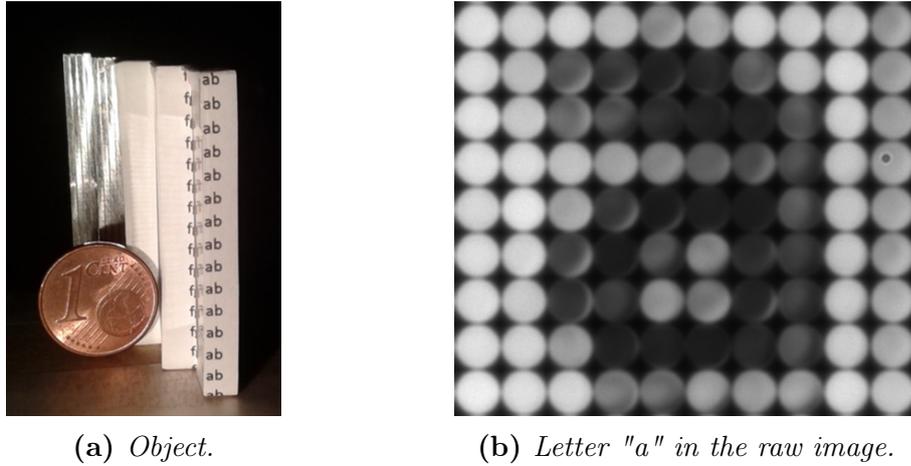


Figure 1.6 – Example of a raw image acquired with a traditional plenoptic camera. The letter "a" was placed at depth z_0^{trad} , resulting in uniform white, gray or black sub-images in the raw image.

object onto the μLA plane. The plane where the image is formed is called the intermediate image plane, located at distance z_1 from the main lens (following Eq. 1.1) and at distance a from the μLA . The micro-lenses act as a relay imaging system, each of them imaging a portion of the intermediate image plane. To acquire sharp images, the sensor should be optically conjugated with the intermediate image plane. This is done by applying the thin-lens

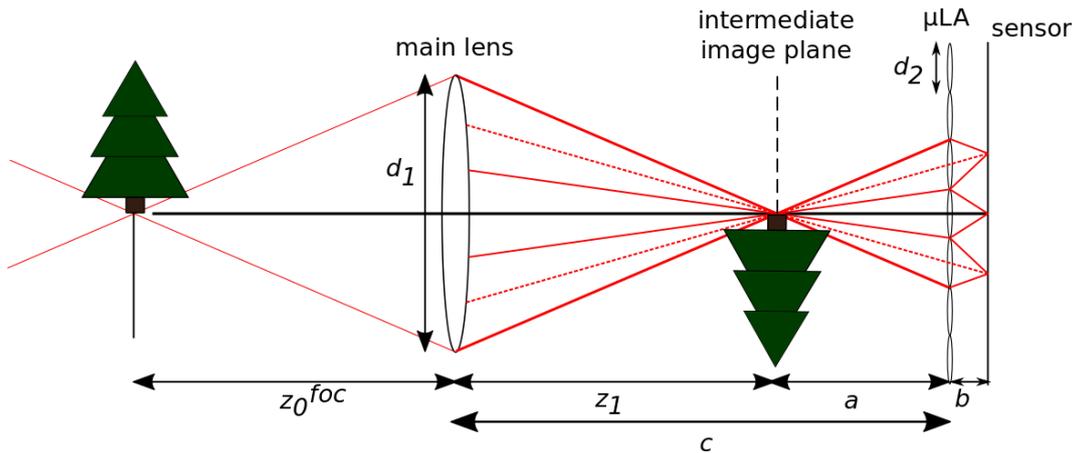


Figure 1.7 – Design of a focused plenoptic camera, composed of a main lens, an array of micro-lenses (μLA) and a detector. The main lens forms the image of the object on the intermediate image plane. The μLA then acts as a relay imaging system, with each micro-lens imaging only a part of the object on the sensor. We introduce distance $c = z_1 + a$ as the physical distance between the main lens and μLA .

equation to the micro-lenses:

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f_2} \quad (1.4)$$

Similarly to the traditional configuration, Georgiev and Lumsdaine [GL10b] established the aperture matching condition as the equality between the image-side numerical apertures of the main lens and μ LA:

$$\frac{d_1}{z_1} = \frac{d_2}{b} \quad (1.5)$$

Equations 1.1-1.4-1.5 constitute the three necessary equations to build a focused plenoptic camera.

Two different modes are possible according to the location of the intermediate image plane. If the plane is between the main lens and the μ LA as in Figure 1.7, the configuration is called Keplerian [GL09a]. The micro-lenses "see" the intermediate image plane as a *real object* because it is located before the μ LA, according to the direction of propagation of light in the system (from left to right). The optical distance a is positive ($a > 0$), which implies that the associated distance b is greater than the focal length ($b > f_2$).

On the contrary, if the main lens projects the image of the object beyond the plane of the μ LA, the configuration is said Galilean [GL09a]. The intermediate image plane becomes a *virtual object* for the micro-lenses. This time, the distances respect the relationships: $a < 0$ and $b < f_2$. According to [GL10a], the mode of a focused camera does not impact the general reasoning on experimental resolutions and data treatment. The Keplerian mode will be used for conceptual reasoning and illustration in the rest of the thesis.

However, the two different modes produce slightly different raw images. Figure 1.8 shows the acquired raw images of the same object using the two modes. In both images, the sub-images represent small parts of the acquired object. In the Keplerian mode, the sub-images are inverted compared to the Galilean mode. This is due to the *virtual* versus *real* object, from the point of view of the μ LA. In the Keplerian mode, the image of the object undergo two inversions due to the main lens and the μ LA. In the Galilean mode, only the inversion of the main lens occurs. Because the intermediate image plane is a *virtual* object for the micro-lenses, the sub-images keep the same orientation. This property allows to interpret the raw image as Galilean or Keplerian mode according to the structure of data inside the image.

In a focused configuration, the properties of the images in the sub-images vary according to depth in object-space. Magnification of the image depends on depth of the object. In Figure 1.8, the letter "a" has a larger size than the letter "f". Moreover, there is a lateral shifts of the letters when going from one sub-image to its neighbour. The size of this displacement depends on the depth of the letter in object-space. As the letter "a" is closer to the camera,

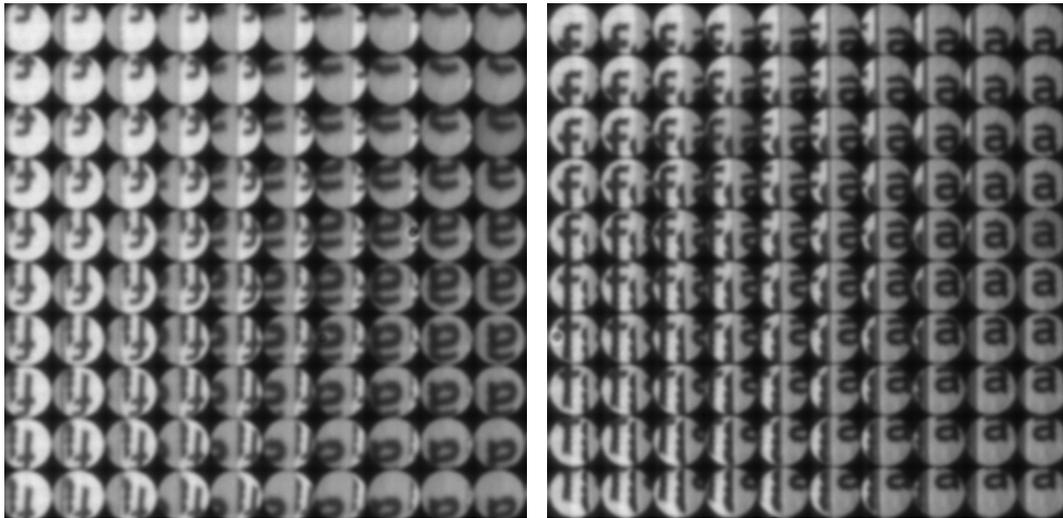
(a) *Keplerian mode ($a > 0$).*(b) *Galilean mode ($a < 0$).*

Figure 1.8 – *Example of two raw images acquired with a focused plenoptic camera, imaging the object shown in Figure 1.6a. As expected, there is an inversion of the sub-images between the Keplerian and Galilean modes. Note that the two images were acquired with slightly different magnifications.*

the displacements of its images are slightly larger than the ones of the letter "f", which is further to the camera and hence closer to the background. The background represents the optical infinity whose image position should not vary across the sub-images. These two effects are more visible on Figure 1.8a than on Figure 1.8b.

These two properties are the signs of depth information included in the raw image, as an indication of depth variation in object-space. Whereas lateral shifts according to depth are an interesting aspect that can be exploited for depth estimation, variations in magnification are an undesirable effect which might negatively affect the object reconstruction.

1.2.4 Conclusion on optical configurations

Based on their respective schemes (see Figures 1.3 and 1.7), the traditional and focused configurations appear very similar. They only differ in adapting the relative positions of the object and the μ LA, and adjusting distance b accordingly. The relationship seems obvious between the two schemes. However, when it comes to raw images, the link is not as clear, due to the different spatio-angular samplings in the two cases. This is also the case of the parameterizations and the refocusing algorithms, each of them being very specific to one or the other configuration, as presented in the following section.

1.3 Refocusing algorithms: traditional vs focused

Refocusing is the action of numerically reconstruct an image at a chosen depth. The light-field is extracted from the raw plenoptic image and back-projected in object-space. The algorithm synthetically mimics the action of optical focusing on a sensor, as if it has been acquired by a classical photography camera. The resulting image contains the in-focus elements that correspond to the depth of refocusing, but also the out-of-focus elements of adjacent depth planes. With refocusing algorithm, we observe the same properties as with classical photography: the in-focus elements are sharply reconstructed, whereas the out-of-focus ones are blurred. This numerical process allows to reconstruct an image at any chosen depth, and thus provides the possibility to change the depth of reconstruction.

In the literature, each experimental design is presented with an associated refocusing algorithm and a general framework to interpret the acquired data [Ng06; LG08a; GL09a]. We here present the two refocusing algorithms, for the traditional and focused configurations.

1.3.1 The traditional refocusing algorithm

The refocusing algorithm for traditional plenoptic cameras is based on the two-planes parameterization (see Section 1.1.2). We present the parameterization and the associated phase-space diagram before moving to the algorithm.

1.3.1.1 Phase-space diagram

In the traditional configuration, it is common to parameterize the light-field between the planes of the main lens and the μ LA (see Figure 1.9). The main lens projects the image of the object on the plane of the μ LA. Thus, the spatial information of a ray is coded by the micro-lens (s) and the angular information is coded by the pixel under this micro-lens (p). The pixel coordinate (p) is related to the coordinate (u) on the main lens aperture. This parameterization allows to draw the phase-space diagram of the light-field through the system.

A raw image captured by a traditional camera is illustrated in Figure 1.10a. From a 1D profile of the raw image, it is possible to extract the 2D light-field, by separating the different sub-images under the micro-lenses. The light-field can then be represented in the phase-space diagram, as a function of u and s (in Figure 1.10b). The dotted rectangles correspond to the sub-images, and the smaller continuous ones to the pixels inside the sub-image. The color code makes the link with the raw image in Figure 1.10a. In the phase-space diagram, the spatial sampling along s is done at a step d_2 corresponding to the pitch of

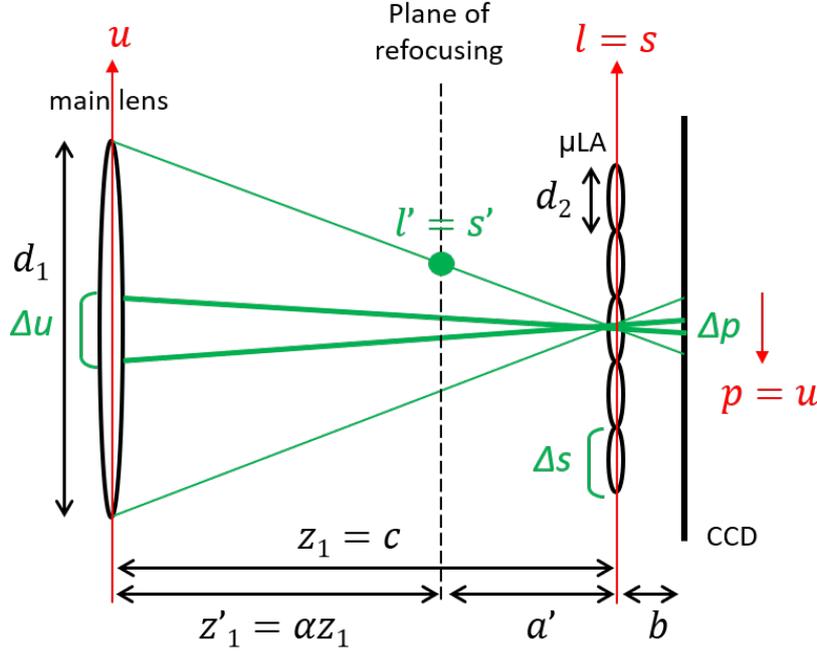
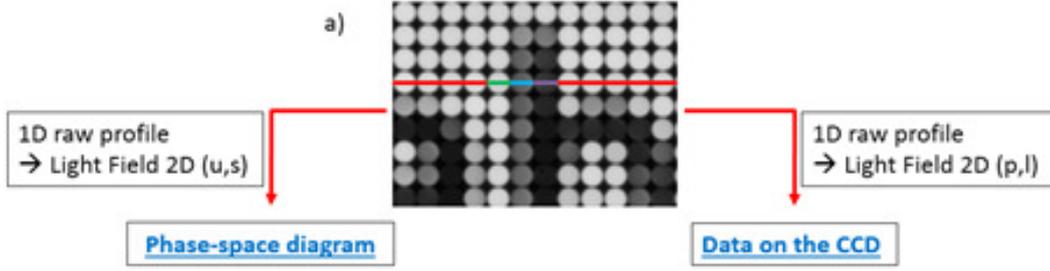


Figure 1.9 – Schematic 2D representation of a plenoptic system. The spatial sampling is done by the μ LA (l or s -axis) with a step Δs , whereas the angular sampling is achieved by the sensor (p or u -axis) with a step Δp corresponding to Δu on the main lens. The refocusing process uses a new parameterization s' on the plane at distance $z'_1 = \alpha z_1$ from the main lens.

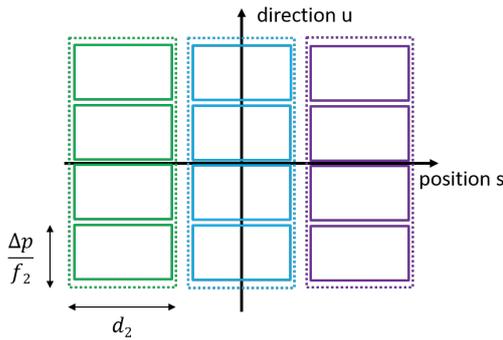
the micro-lenses. The angular sampling along u is $\Delta p/f_2$, that corresponds to the angular extent of one pixel, seen from the μ LA-plane.

In the context of this thesis, we introduce another two-planes parameterization. As explained in Section 1.1.2, the representation of the light-field in the phase-space diagram depends on the position of the two planes. In order to be independent from the parameterization, we introduce the sensor- μ LA diagram as a direct representation of the raw data on the sensor. This diagram represents the light-field as a function of pixels (p) and micro-lenses (l) (see Figure 1.10c). The l -axis corresponds to the s -axis of the phase-space diagram, and the p -axis corresponds to the projection of the u -axis on the sensor-plane. Thus, the phase-space diagram and the sensor- μ LA diagram are very close (Figures 1.10b-1.10c).

This parameterization is independent from the configuration in which the light-field was acquired, because it does not take into account the position of the optical elements before the μ LA and sensor. We will use later this sensor- μ LA diagram to compare the numerical reconstruction methods from the same light-field (see Section 1.3.3).



(a) Raw traditional plenoptic image.



(b) Phase-space diagram.

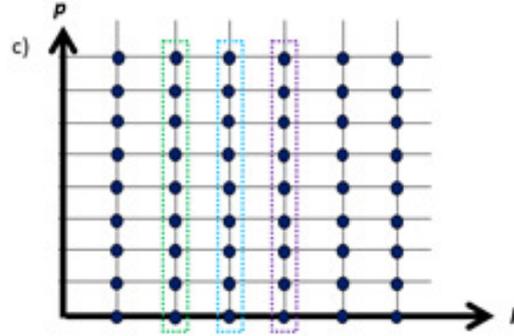

 (c) Sensor- μ LA diagram.

Figure 1.10 – Schematic representation of the traditional plenoptic camera with the raw image, the phase-space diagram and the sensor- μ LA representation.

1.3.1.2 Refocusing process: shift-and-sum algorithm

The refocusing algorithm for the traditional configuration is based on the integration of angular data in the phase-space diagram [Ng06; Lam15]. As for the phase-space diagram, the refocusing process is usually described between the planes of the main lens and the μ LA. The idea is to shift the two-plane parameterization from the main lens- μ LA to a new parameterization between the main lens and the chosen plane of refocusing.

We define the plane of refocusing located at distance z'_1 from the main lens, optically conjugated with the plane at depth z'_0 in object-space. The refocusing parameter α makes the link between the change of parameterization (see Fig. 1.9). It can be expressed as:

$$\alpha = \frac{z'_1}{z_1} = 1 - \frac{a'}{z_1} \quad (1.6)$$

After this change of parameterization, the process of refocusing consists in *focusing* the light rays to generate the image im_α on this new plane z'_1 . This is done by combining together all the angular components passing through the same spatial coordinate of this new plane of refocusing. Mathematically, it corresponds to an integration over u at fixed s' . We re-write Equation 2.3

from [Ng06] using our formalism:

$$im_\alpha(s', t') = \frac{1}{\alpha^2 z_1^2} \int \int LF(u(1 - \frac{1}{\alpha}) + \frac{s'}{\alpha}, v(1 - \frac{1}{\alpha}) + \frac{t'}{\alpha}, u, v) du dv \quad (1.7)$$

We re-express Equation 1.7 in 2D as a function of the coordinates of the sensor- μ LA diagram, replacing spatial coordinate (s) with micro-lens (l) and angular coordinate (u) with pixel (p):

$$im_\alpha(l') = \int LF(l' + p \frac{(\alpha - 1)}{r}, p) dp \quad (1.8)$$

Compared to the original Equation 1.7, we discarded the normalization term before the integral symbol. We also multiplied the spatial coordinate of the light-field by α : $\frac{l'}{\alpha} + p \frac{(1 - \frac{1}{\alpha})}{r}$ becomes $l' + p \frac{(\alpha - 1)}{r}$. This allows to avoid the zooming/dezooming effect in the reconstructed image when performing refocusing. Finally we introduced the parameter r , to compensate the physical difference in sampling frequencies between spatial and angular coordinates. According to Figure 1.9, the spatial sampling is given by the pitch of a micro-lens: $\Delta s = d_2$. The angular sampling step can be quantified on the main lens plane by: $\Delta u = \Delta p \cdot z_1 / b$. This gives the following expression:

$$r = \frac{\Delta s}{\Delta u} = \frac{d_2}{\Delta p} \frac{b}{z_1} \quad (1.9)$$

The representation of the algorithm in the sensor- μ LA diagram allows to better understand how it could be practically implemented (see Figure 1.11). As shown in Figure 1.2, changing the two-planes parameterization from the main lens- μ LA planes to the new one considering the main lens and the plane of refocusing corresponds to a rotation of the light-field in the phase-space diagram [Ng06; Lam15]. The algorithm reproduces this rotation of the sub-images (initially vertical) controlled by the parameter α , followed by an integration of the light-field along these tilted lines (dashed blue arrow in Figure 1.11). This process is called the *shift-and-sum* algorithm because it corresponds to a rotation (*shift*) of the light-field in the phase-space diagram, followed by an integration (*sum*) along the angular axis (see Figure 1.11).

The parameter α is the refocusing parameter. The case $\alpha = 1$ corresponds to a reconstruction at the exact plane where the camera was focused, i.e. the reconstruction of the object placed at a distance z_0 in front of the main lens. When $\alpha \neq 1$, the refocusing is done at a different plane. Because of the magnification of the main lens, the relative positions in object-space (z'_0 compared to z_0) are inverted from the relative positions in image-space, between main lens and μ LA (z'_1 and z_1): for $\alpha < 1$, the refocusing is done closer to the main lens in image-space, meaning further from it in object-space; on the contrary, $\alpha > 1$ means a refocusing further from the main lens in image-space and closer in object-space.

Two characteristics can be noticed from this diagram. First, the lateral resolution of the reconstructed image, which can be read as the spacing between the lines of integration, is constant and independent from α . On the plane of the μ LA, the lateral resolution corresponds to the spacing of the micro-lenses, i.e. the pitch d_2 of the micro-lenses. Second, as the lines of integration are rotated compared to the initial data, we can see that some points are missing along the lines (points surrounded with yellow in Figure 1.11). These points were not acquired by the camera, and have to be created by interpolation before performing the integration. This interpolation across pixels from adjacent sub-images can be a source of errors in the refocused image.

1.3.1.3 Example of refocused images

Figure 1.12 shows an example of the action of refocusing at various depths in the scene. A traditional plenoptic camera was used to acquire the raw plenoptic image of a ruler. The ruler was positioned with an angle of about 21 degrees from the optical axis. This rotation allows a progressive depth increase along the ruler while still being able to see the numbers.

The same raw image was then refocused using different values of the refocusing parameter α . Changing α induces a change in the depth where the image is reconstructed. The change in depth of reconstruction is visible in the

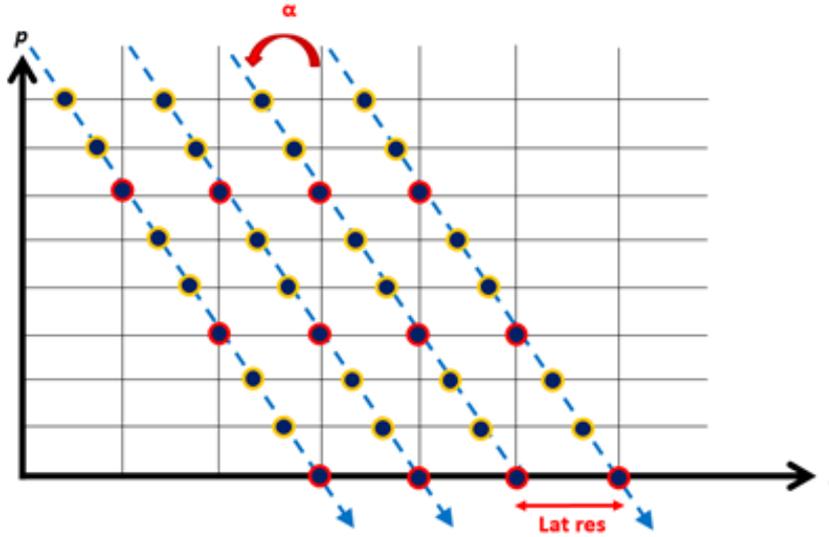


Figure 1.11 – Traditional plenoptic reconstruction principle in the sensor- μ LA diagram. Points with a red circle correspond to pixels acquired by the detector while points circled with yellow represent missing data that has to be created by interpolation. To get the reconstructed image, the integration along the line of equation $l(p) = l' + p \frac{\alpha(1-\frac{1}{\alpha})}{r}$ has to be performed, as illustrated by the arrows. We represent here the reconstruction with $\alpha > 0$, which corresponds to a negative slope. The lateral resolution of the reconstructed image equals the micro-lens size.

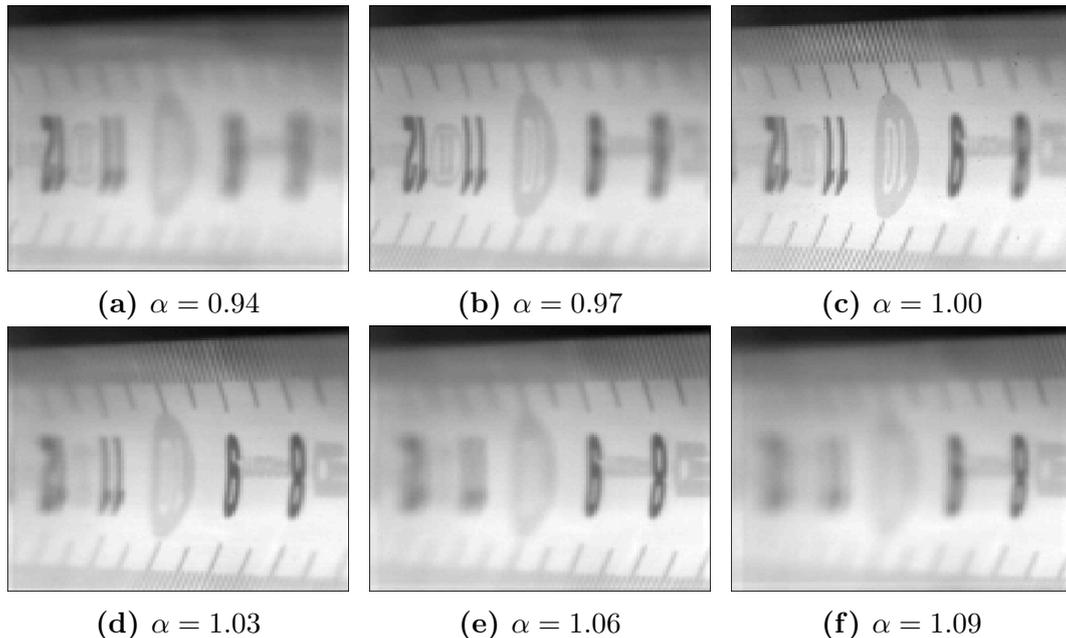


Figure 1.12 – Refocused images from the same raw image of a rotated ruler. Changing the refocusing parameter α results in a change of the depth plane of reconstruction, which is visible through the left-to-right displacement of the in-focus areas across the images.

images through the lateral displacement of the in-focus (sharp) area. At the extreme location ($\alpha = 0.94$ and $\alpha = 1.09$), the images are nearly completely blurred. The intermediate images present a sharp area whose position depends on α : between numbers 11 and 12 ($\alpha = 0.97$), between 10 and 11 ($\alpha = 1.00$), around number 9 ($\alpha = 1.03$), finally around number 8 ($\alpha = 1.09$). Such a series of images refocused at various depths will be called a *refocused stack*.

1.3.2 The refocusing algorithm for the focused configuration

The phase-space diagram and the refocusing algorithm for the focused plenoptic camera are less straightforward, because the spatial and angular information are not separated as in the traditional case. We here explain how the phase-space diagram could be built, and then present the basic refocusing approach.

1.3.2.1 Construction of the phase-space diagram

As in the traditional configuration, the phase-space diagram for the focused configuration is established considering a two-planes parameterization between the main lens and μ LA. More specifically, the phase-space diagram describes

how the intermediate image-space is projected onto the sensor and how the corresponding light-field is sampled. Contrary to the traditional setup, the spatio-angular data are mixed on both μ LA and sensor planes, thus a more complex approach is necessary to build the phase-space diagram.

Lumsdaine et al. proposed an approach based on the ABCD matrix formalism [LGC12]. ABCD matrix is a compact representation of ray tracing through an optical system, in the context of the geometrical optics and paraxial approximation [Gro05]. It allows to easily express the angle and position of the emerging ray as a function of the angle and position of the entering ray, using matrix calculus. In their article, Lumsdaine et al. use the ABCD matrix to make the link between the intermediate image plane and the sensor, as the input and output planes of an optical system.

In the formalism of ABCD matrix, the angle and position of a ray are expressed by a 2-coordinates vector. Here we call $ray = \begin{pmatrix} s \\ u \end{pmatrix}$ the input ray located at the intermediate image plane, and $ray' = \begin{pmatrix} s' \\ u' \end{pmatrix}$ the output ray at the

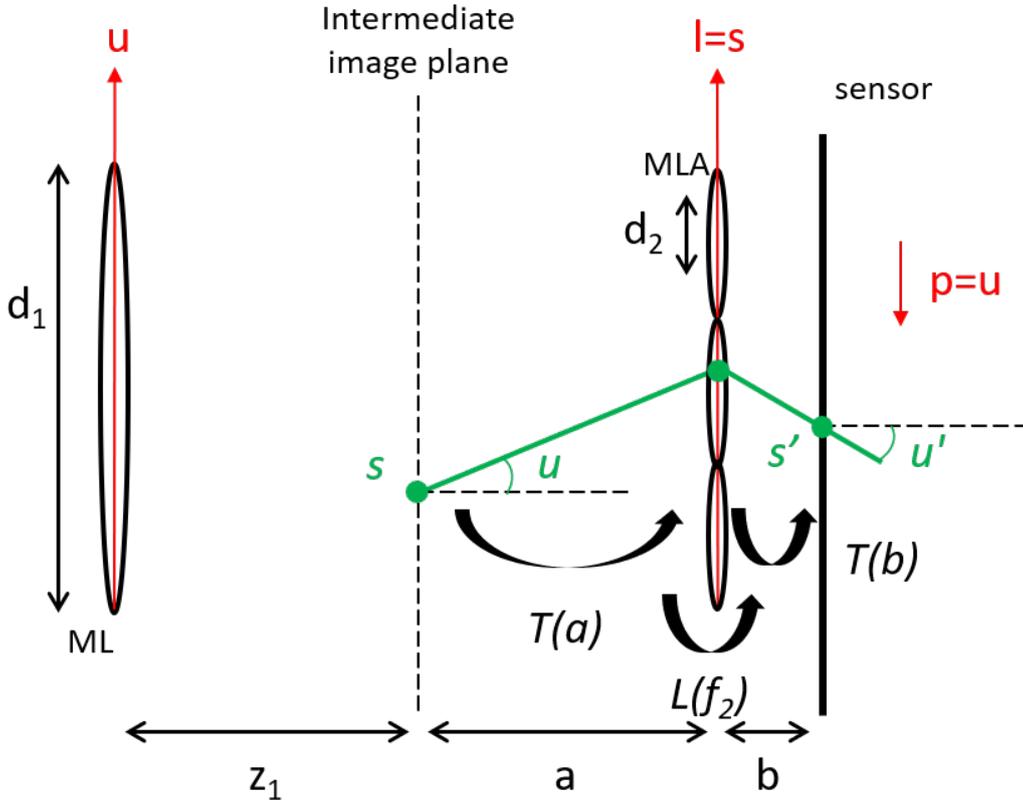


Figure 1.13 – Representation of the successive matrices used to establish the phase-space diagram.

sensor plane (see Figure 1.13). The expression of the coordinate transformation for free-space propagation on a distance z can be written as the matrix [Gro05]:

$$T(z) = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \quad (1.10)$$

Similarly, the matrix describing the coordinate transformation due to a lens of focal length f is:

$$L(f) = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \quad (1.11)$$

The matrix describing a complex optical system is given by the multiplication of the matrix of each single element. Between the intermediate image plane and the sensor, the ray goes through a free-space of length a , then through a micro-lens, and finally a free-space of distance b (see Figure 1.13). Consequently, the global matrix can be expressed as:

$$A = T(b) \circ L(f_2) \circ T(a) = \begin{pmatrix} 1 - \frac{b}{f_2} & (1 - \frac{b}{f_2}) \cdot a + b \\ -\frac{1}{f_2} & 1 - \frac{a}{f_2} \end{pmatrix} \quad (1.12)$$

In the specific case of the focused plenoptic configuration, we remind that the intermediate image plane is optically conjugated with the sensor. Thus the thin lens equation (Equation 1.4) could be used to simplify matrix A as:

$$A = \begin{pmatrix} -\frac{b}{f_2} & 0 \\ -\frac{1}{f_2} & -\frac{a}{b} \end{pmatrix} \quad (1.13)$$

Using this matrix, the coordinates of an incident ray is transformed into the output ray by: $A \begin{pmatrix} s \\ u \end{pmatrix} = \begin{pmatrix} s' \\ u' \end{pmatrix}$. Reciprocally, the coordinates from the incident ray could be calculated from the ones of the output ray by: $\begin{pmatrix} s \\ u \end{pmatrix} = A^{-1} \begin{pmatrix} s' \\ u' \end{pmatrix}$, with the inverse matrix A^{-1} being:

$$A^{-1} = \begin{pmatrix} -\frac{a}{b} & 0 \\ \frac{1}{f_2} & -\frac{b}{a} \end{pmatrix} \quad (1.14)$$

To create the phase-diagram, the sampling done by the pixels on the sensor should be expressed back on the intermediate image plane, using the transformation matrix. The underlying assumption is that the radiance of the light-field remains constant when going through the optical system [McC14]. At this time, Lumsdaine et al. reminded that a 2D image on the sensor is created from the integration of the 4D light-field along its angular components [LGC12]. The light-field integration at a spot s on the sensor can be written as:

$$\begin{aligned}
I_{sensor}(s) &= \int_{-\frac{d_2}{2b}}^{\frac{d_2}{2b}} ray_{sensor} \left[\begin{pmatrix} s \\ u \end{pmatrix} \right] .du \\
&= \int_{-\frac{d_2}{2b}}^{\frac{d_2}{2b}} ray_{intermediate} \left[A^{-1} \begin{pmatrix} s \\ u \end{pmatrix} \right] .du \\
&= \int_{-\frac{d_2}{2b}}^{\frac{d_2}{2b}} ray_{intermediate} \left[\begin{pmatrix} -\frac{a}{b}s \\ \frac{1}{f_2}s - \frac{b}{a}u \end{pmatrix} \right] .du
\end{aligned} \tag{1.15}$$

The integration is done on the whole angular range available at the sensor plane. From the sensor plane, this angular range corresponds to the numerical aperture of one single micro-lens, $NA = \frac{d_2}{2b}$, considered both upward and downward the optical axis. Equation 1.15 allows to draw the phase-space diagram of a spot s . When the variable u browses an angular range of $\frac{d_2}{b}$ in the integral, the incident ray browses an angular range of $\frac{b}{a}\frac{d_2}{b} = \frac{d_2}{a}$. This corresponds to the vertical line of length $\frac{d_2}{a}$ in Figure 1.14a.

However, the pixels on the sensor are not punctual. Their size Δp should be taken into account. In addition to the previous angular integration, it is necessary to achieve a spatial integration over the spatial extent Δp of a pixel. Equation 1.15 becomes:

$$\begin{aligned}
I_{sensor}(\Delta s) &= \int_{-\frac{\Delta p}{2}}^{\frac{\Delta p}{2}} \int_{-\frac{d_2}{2b}}^{\frac{d_2}{2b}} ray_{sensor} \left[\begin{pmatrix} s \\ u \end{pmatrix} \right] .du .ds \\
&= \int_{-\frac{\Delta p}{2}}^{\frac{\Delta p}{2}} \int_{-\frac{d_2}{2b}}^{\frac{d_2}{2b}} ray_{intermediate} \left[\begin{pmatrix} -\frac{a}{b}s \\ \frac{1}{f_2}s - \frac{b}{a}u \end{pmatrix} \right] .du .ds
\end{aligned} \tag{1.16}$$

Using Equation 1.16, the phase-space diagram could be extended from a line to a surface (see Figure 1.14b). The spatial extent Δp of the sensor pixel is responsible for extending the phase-space to a range of $\frac{a}{b}\Delta p$ along the (horizontal) s -axis, and an additional range of $\frac{\Delta p}{f_2}$ along the (vertical) u -axis. This 2D parallelogram represents the whole spatio-angular data acquired by a pair of micro-lens and pixel. Figure 1.14b is similar to Figure 5b in [LGC12] excepted the range $\frac{\Delta p}{f_2}$ replacing their $\frac{\Delta p}{b}$.

Similarly, the same reasoning allows to draw the phase-space diagram of the whole sub-image under one micro-lens, considering all the pixels (see Figure 1.14c). The adjacent micro-lenses are then added to complete the phase-space diagram of the focused plenoptic camera (see Figure 1.14d).

Lumsdaine et al. showed that the same approach using ABCD matrices allows to construct the phase-space diagram of the traditional plenoptic configuration [LGC12]. The general matrix A in Equation 1.12 could be simplified using the condition $b = f_2$ specific to the traditional configuration (see

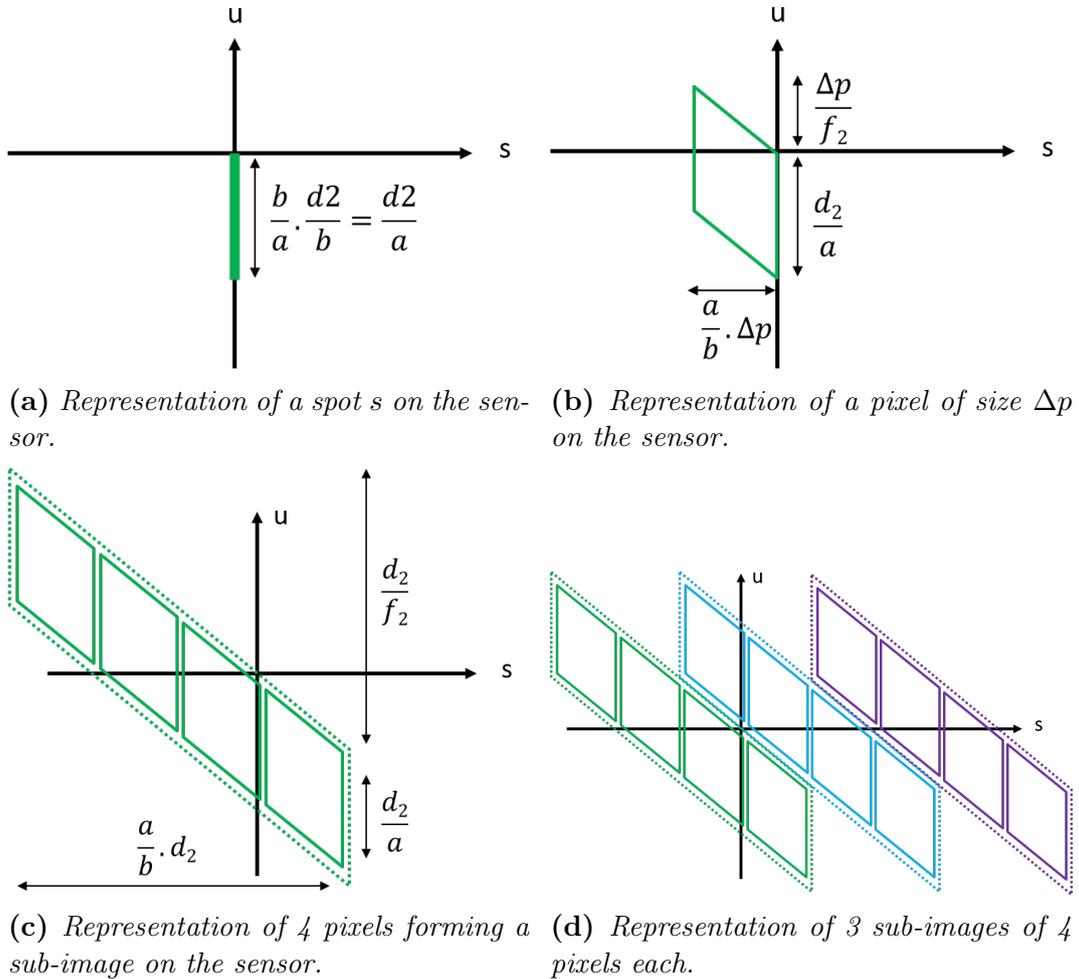


Figure 1.14 – Step by step construction of the phase-space diagram for the focused plenoptic camera, showing how the light-field is sampled spatially (s -axis) and angularly (u -axis).

Equation 1.2). The resulting diagram converges to the one presented in Figure 1.10b. This allows to specify the vertical extent of each pixel in this configuration.

1.3.2.2 The different rendering methods

To reconstruct an image at or around the intermediate image plane, Lumsdaine et al. presented a basic rendering algorithm, called the patch-tilling algorithm [LG08b]. It can be directly expressed from the raw light-field image and the geometry of the system. In the raw image, we notice that the sub-images consists in imaging small parts of the object (see Figure 1.8). This is due to the configuration acting as a relay-imaging system. However, the full size of the object is spread on several sub-images due to the distribution of the

possible to reconstruct the scene is a discrete function, while it is a continuous interval for the shift-and-sum algorithm.

This rendering algorithm could also be expressed in the phase-space diagram. Figure 1.16 shows the refocusing process on a few sub-images, with parameter $P = 2$. For each micro-lens, only 2 pixels are extracted and assembled together to form the final image. This simple example shows that the basic refocusing algorithm uses only a part of the acquired data. In each sub-image, only the central P pixels are used for the refocusing, ignoring the other acquired data.

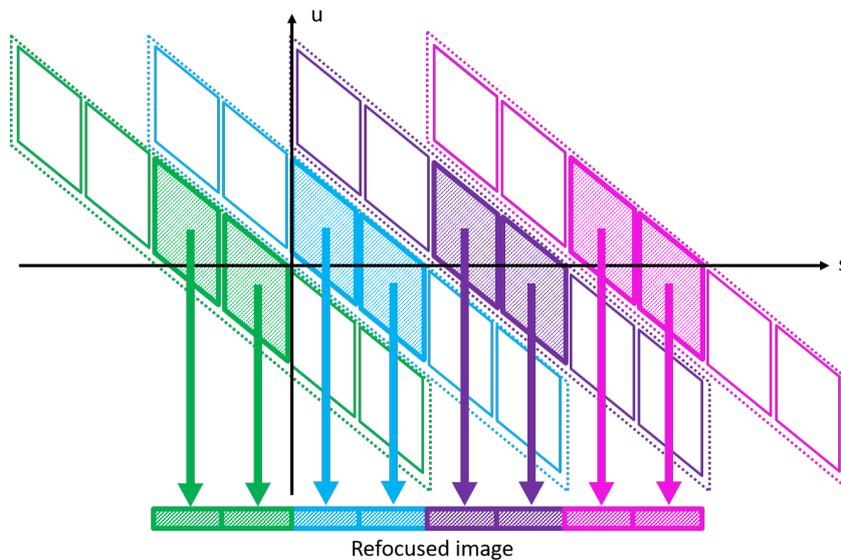


Figure 1.16 – *Basic rendering algorithm. In each micro-lens, only 2 pixels are selected to form the reconstructed image ($P = 2$).*

Another method has been presented by Georgiev and Lumsdaine, which uses all the available data [GL10a]. They called it "rendering with blending". In this method, the pixels from adjacent sub-images are combined together to form the reconstructed pixel of the final image. Actually, this is performed by extracting surrounding patches and averaging them together. In the phase-space diagram, this corresponds to the integration of data along the vertical lines (see Figure 1.17). The advantage of this method is that it uses all the data acquired by the plenoptic system. As a result, the refocused image is smoother with less noise, due to the averaging process.

A third method has been presented by Georgiev et al. in [GCL11]. This is the "superresolution" method. Instead of integrating over aligned pixels from adjacent sub-images as in the "rendering with blending", this method slightly translates along the s -axis the patterns of the sub-images in the phase-space diagram. Considering a subpixel shift between the previously aligned pixels, they could now be combined alternatively in the refocused image. The resulting

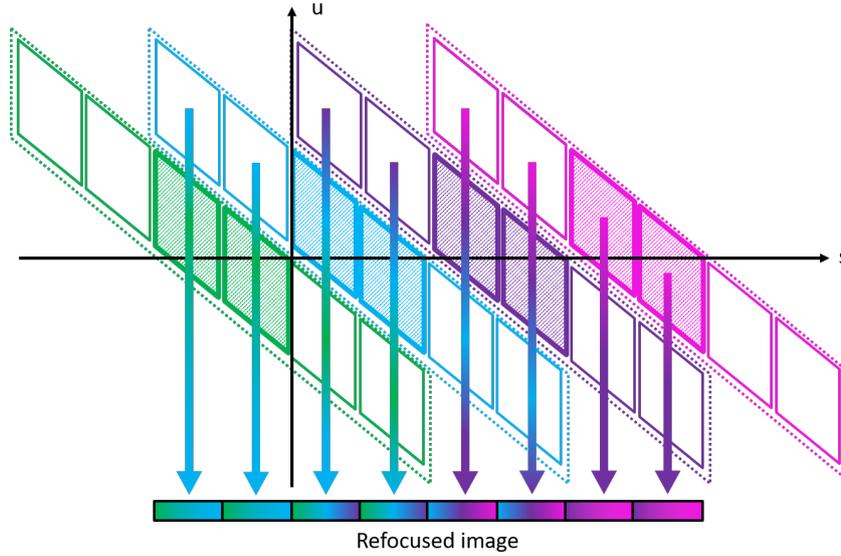


Figure 1.17 – *Rendering with blending algorithm. Each pixel of the refocused image is formed by integrating over several pixels of the raw image. All the pixels of the raw image are used.*

image contains more pixels, thus the term "superresolution".

1.3.2.3 Re-expressing the rendering method

To compare the refocusing algorithms from both the traditional and the focused configurations, we need to re-express the rendering method "with blending" in the formalism of the sensor- μ LA diagram, that was presented in Section 1.3.1.2 for the traditional configuration.

The sensor- μ LA diagram is the direct representation of the data as acquired on the sensor, after each sub-image has been extracted from the raw image. It does not assume any relationship between the distances, contrary to the phase-space diagram which was built specifically for each configuration, either traditional or focused.

Part of a raw image captured by a focused plenoptic camera is illustrated in Figure 1.18a. From a 1D profile extracted from the raw image, the acquired data could be represented in two ways: either the phase-space diagram (in Figure 1.18b) or the sensor- μ LA diagram (in Figure 1.18c). When acquired with a focused camera, the light-field has not the same shape in the phase-space and sensor- μ LA diagrams, contrary to the traditional configuration (see Figure 1.10). In the phase-space diagram (Figure 1.18b), the patterns of the sub-images are rotated, whereas they are still vertical in the sensor- μ LA diagram.

We now re-express the "rendering with blending" method in the sensor- μ LA diagram. This method is chosen over the three ones because it is the

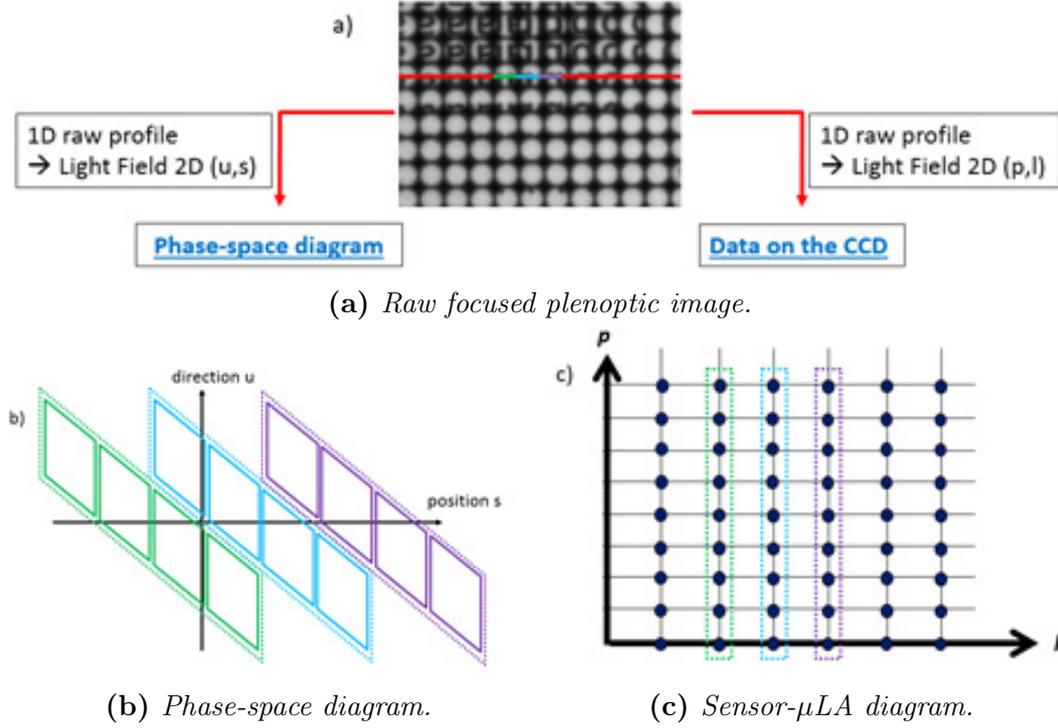


Figure 1.18 – Schematic representation of the focused plenoptic camera with the raw image, the phase-space diagram and the sensor- μ LA representation.

closest to the shift-and-sum algorithm, due to the integration process. Figure 1.19 compares the patch extraction on the sub-images with the pixels in the sensor- μ LA diagram. In the sub-images, patches of the same color represent the same spatial data distributed over several sub-images (see Fig. 1.19a). In the diagram, the patches are represented as blocks of P adjacent pixels along the p -axis (see Fig. 1.19b). When integration is performed, patches of the same color are combined together. The size P of the patches is responsible for the slopes of the lines of integration.

With this method, one red patch (in Figure 1.19a) is associated with the other red patches (taken in the neighbor sub-images by shifting the central patches by an integer number of patches, i.e. a multiple of P pixels) and then the integration is performed (represented by the blue arrows in Figure 1.19b). The equation of the line of integration can be expressed in terms of pixels (p) and micro-lens (l) as:

$$p(l) = p_0 - l.P \quad (1.18)$$

which is equivalent to

$$l(p) = \frac{p_0 - p}{P} \quad (1.19)$$

Therefore, the reconstructed pixel $im_P(p_0)$ in the output image of the "rendering with blending" can be calculated by integrating the light-field using the

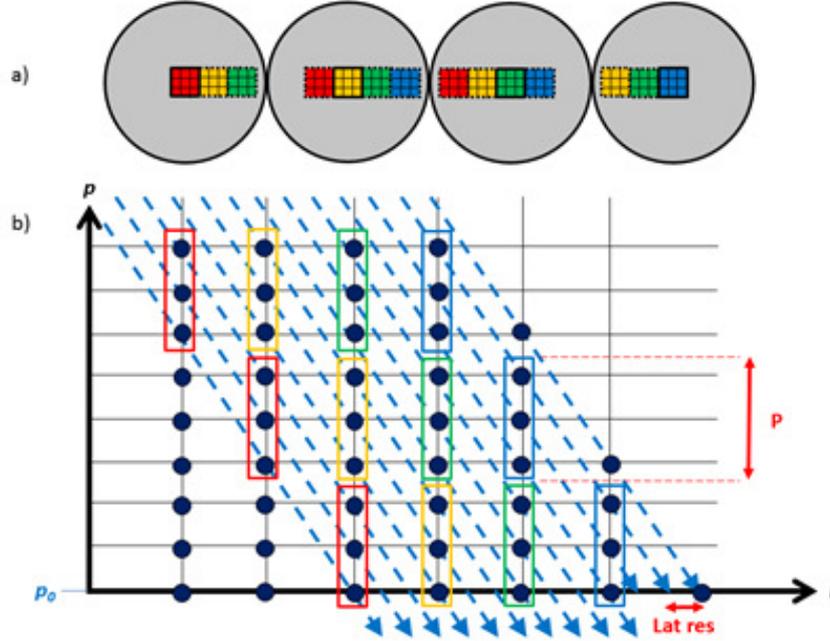


Figure 1.19 – a) Illustration of patches composed of 3×3 pixels and their correspondence with neighboring sub-images. b) Representation of the "rendering with blending" method in the sensor- μ LA diagram, which consists in an integration along lines of equation $l(p) = (p_0 - p)/P$. We consider a reconstruction in the case $a > 0$, leading to $P > 0$ and a negative slope.

following equation:

$$im_P(p_0) = \int LF\left(\frac{p_0}{P} - \frac{p}{P}, p\right) dp \quad (1.20)$$

Figure 1.19 allows to notice the characteristics of the rendering with blending method, in terms of resolution and image quality. All the points used along the lines of integration are part of the grid defined by the pixels and micro-lenses, meaning that they were acquired by the sensor. This corresponds to the pixels of the patches being extracted from the raw image, without the need of an interpolation as for the shift-and-sum algorithm (see Figure 1.11). Moreover, the numerous lines of integration result in a small lateral resolution in the reconstructed image. There are P lines of integration between two adjacent vertical lines representing the extracted sub-images, corresponding to the P pixels in each patch. In the intermediate image plane, the lateral resolution equals d_2/P . This value can then be projected back in the object-space using the magnification M of the main lens. One can note that this value depends on P thus on z'_0 and z'_1 , which means that, for the rendering with blending, the number of pixels in the output image depends on the chosen plane of refocusing.

1.3.3 Comparison of the two algorithms

Writing the shift-and-sum algorithm and the rendering with blending in the same diagram and formalism allows to better compare their mathematical equations. Both methods could be expressed as an integration over the light-field, leading to similar Equations 1.8 and 1.20. When reconstructing the same data with both algorithms independently, the refocusing parameters α and P should be adapted according to the chosen depth of reconstruction z'_0 and z'_1 in the object and image spaces. The slope of integration of the rendering with blending (Eq. 1.20) is given by $\frac{-1}{P}$, while the slope for the shift-and-sum algorithm (Eq. 1.8) equals $\frac{(\alpha-1)}{r}$. From the expressions of r , α and P (Eqs. 1.9 - 1.6 - 1.17), we can establish the equality:

$$\frac{(\alpha - 1)}{r} = \frac{-1}{P} \quad (1.21)$$

This equation means that the two slopes are equal, and that the orientation of the lines is exactly the same for both algorithms when choosing the same plane of refocusing. Even if these two algorithms were originally built for different optical systems, either the traditional or the focused configurations, this highlights a major common point between them: they both perform integration over lines *with exactly the same slope*.

1.3.4 Application on image reconstruction

Knowing that the principle of reconstruction of the two methods is basically equivalent, it is interesting to compare their outputs by applying them on two raw images, one obtained using a traditional setup and the other one using a focused one. Georgiev and Lumsdaine presented the results of the focused algorithm on traditional images [LG09], and reciprocally the shift-and-sum algorithm applied on their focused image [GL09b]. We follow the same idea and apply both the shift-and-sum algorithm and the rendering with blending on the raw images presented in Figures 1.6 and 1.8. These raw images have been acquired using the same optics. This comparison allows to focus on the differences due to the configuration or the reconstruction method independently. The results are presented in Table 1.1.

For the raw image acquired in the traditional configuration, there is nearly no visual difference between the results obtained with the two different reconstruction algorithms (Table 1.1: top row). For the image acquired with the focused plenoptic camera (Table 1.1: bottom row), one can notice a difference in resolution and image quality between the output images from the two different algorithms. The focused image reconstructed by the shift-and-sum algorithm (Table 1.1: bottom left) is of the same quality than the reconstructed image from the traditional raw image. On the contrary, the same focused image reconstructed using the rendering with blending presents better resolution

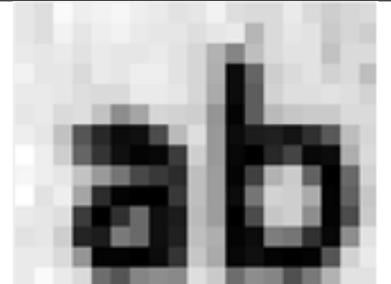
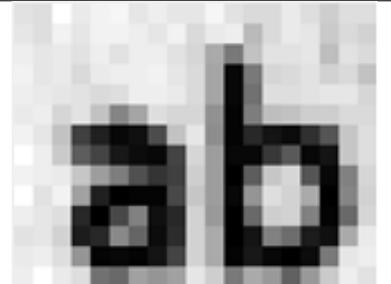
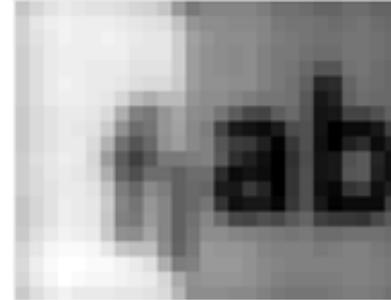
	Reconstruction with the shift-and-sum algorithm	Reconstruction with rendering with blending
Image acquired with a traditional plenoptic camera (M=2.2)		
Image acquired with a focused plenoptic camera (M=1.67)		

Table 1.1 – Comparison of images acquired in a traditional and focused configurations, and reconstructed using both refocusing algorithms.

and image quality (Table 1.1: bottom right). This Table shows that it is the combination of a specific configuration and an algorithm that determines the characteristics of the reconstructed image.

The difference in resolution can be explained mathematically by considering the representation of the algorithms in the sensor- μ LA diagram. With the shift-and-sum algorithm, the lateral resolution is given by the pitch of a microlens, d_2 , while for the rendering with blending, the lateral resolution equals d_2/P . The rendering with blending generates more lines of integration hence more pixels in the reconstructed image. A relationship could be established between the resolutions of the traditional vs focused refocusing algorithms:

$$res_{lat}(\text{rendering with blending}) = \frac{res_{lat}(\text{shift-and-sum})}{P} \quad (1.22)$$

P being an integer number of pixels, it is larger than 1. Equation 1.22 explains the finest resolution observed in the output of the rendering with blending compared to the shift-and-sum algorithm, when applied on the focused raw image (see Table 1.1: bottom row). The resolution could be even finer using a superresolution refocusing algorithm [GCL11]. Because of the difference of resolution, the impact of the interpolation performed across adjacent sub-images in the shift-and-sum algorithm (see Section 1.3.1.2) could not be observed on these images.

On the contrary, the reconstructions from the traditional raw image do not present many differences according to the refocusing method (see Table 1.1:

top row). The resolution is not improved by the focused algorithm. This illustrates that the optical configuration constraints the acquired light-field data, and limits the resolutions of the reconstructed image.

Refocusing from the traditional raw image also raises the question of refocusing at the specific depth z_0^{trad} using the rendering with blending method. The depth z_0^{trad} corresponds by definition to a distance $a = 0$, the intermediate image plane coinciding with the plane of the μ LA. According to Equation 1.17, $a \rightarrow 0$ leads to $P \rightarrow \infty$, meaning an infinite number of pixels. It is theoretically impossible to refocus at this specific depth z_0^{trad} using the rendering with blending method, which has been developed for the focused configuration. In practice, all the pixels of each sub-image are integrated together along the vertical lines in the sensor- μ LA diagram (patch covering the whole sub-image). This is the same process as done in the shift-and-sum algorithm, this is why the two images are very similar (see Table 1.1: top row). The few differences could be explained by small differences in implementation.

1.3.5 Conclusion on refocusing algorithms

As a conclusion, similarities exist between the shift-and-sum algorithm and the rendering with blending. It illustrates that all the plenoptic cameras composed of a main lens, a μ LA and a detector are a single optical concept, independent of the relative position of the optics. Although the two reconstruction methods associated to the traditional and focused configurations are often presented as independent, we have shown that their principles are in fact very close. They consist on an integration along a line, with a slope depending on the position of the plane of refocusing. The mathematical expressions of the slopes are in fact exactly the same for both algorithms. This similarity justifies that each of them can be applied independently on raw images acquired either with a traditional or focused plenoptic camera. The resulting image resolution and quality depend on the chosen algorithm.

1.4 Previous studies on resolutions

Up to now, the traditional and focused optical setups were presented, as well as the refocusing algorithms associated with each of them, allowing to reconstruct the raw images. In this section, we discuss about the optical resolutions and image quality of the reconstructed images.

The previous comparisons between traditional and focused plenoptic configurations were done in the visible. In order to compare the choice of one or the other configuration, the expected resolutions for both designs should be compared using the same formalism and validated with setups using the same optical components and parameters. The formalisms introduced by

Mignard-Debise [MD18] or Hahne [HAV⁺16] give interesting tools to compare the different setups. Lumsdaine et al. have shown that the phase-space diagram could also be used in order to predict the lateral resolutions in both configurations [LGC12]). Other authors gave formula for the traditional configuration [LNA⁺06; DZC⁺16; MCJ18] or for the focused configuration [GL09a; Tur16; ZLE⁺18].

When it comes to microscopy, the expected resolutions are a crucial issue, in order to tell which size of the sample could be accurately reconstructed. Several studies on resolutions were done in the context of visible microscopes, mainly used in a traditional configuration [LNA⁺06; MDI15; DZC⁺16]. Levoy et al. showed theoretical and experimental results in the case of the traditional light-field microscope [LNA⁺06]. They explained how both the lateral and depth resolutions could be established from the parameters of the experimental system, considering the effects of diffraction through the main lens.

Zhu et al. derived complete equations and confronted them to experimental measurements in both traditional and focused configurations [ZLE⁺18]. However, their equations correspond to depth estimation of a single point source, which allows finer precision than for imaging application. Moreover, they separate the reasoning for the traditional and focused cases, with each configuration being reconstructed with a different refocusing algorithm. As shown in Section 1.3.4, the refocusing method affects the resolution and quality of the image. Their proposed formulas result from the combination of the limits due to the optical setup and the ones imposed by the refocusing process, whereas it would be interesting to separate the optical and the numerical effects.

1.5 Synthesis of the comparison

To choose between the traditional and focused cases, not only the resolutions but also all the aspects previously presented should be taken into account. Table 1.2 provides a global overview of the characteristics of both configurations, highlighting their advantages and limitations, according to previous work achieved in the visible.

The choice of a configuration depends on the optical elements used to build the experimental setup. When a μ LA with a high number of small micro-lenses is available, a traditional configuration could be implemented. The separation of the spatio-angular data in this configuration makes it relatively easy to calibrate and reconstruct the images, while the small micro-lenses would provide interesting resolutions. On the contrary, an array of a few large micro-lenses would be more adapted in a focused configuration [GL10b]. The more complex distribution of spatio-angular data over the micro-lenses and pixels allows more flexibility in the optical design. The possibility to modify this distribution by adjusting the magnification of the μ LA enables a largest

Traditional	Focused
Optical configurations	
Desired optics: μ LA with a large number of small micro-lenses	Desired optics: potentially larger micro-lenses
Separated spatial and angular data	Mixed spatio-angular data
Fixed resolutions imposed by the number and size of micro-lenses	Flexibility in the resolutions determined by the magnification of the μ LA
Refocusing algorithms	
Same principle of integration along rotated lines, whose slopes depend on the chosen depth of reconstruction	
Continuous refocusing parameter α	Integer refocusing parameter P
Limitation: Interpolation to generate missing data before integration	Limitation: Difficulty to refocus at depth z_0^{trad}

Table 1.2 – *Global overview of the characteristics of the traditional and refocused cases, considering both optical setups and refocusing algorithms.*

choice in the resolutions of the reconstructed image, hence potentially finer resolutions. As a conclusion, the number and size of the micro-lenses appear to be the major parameters leading the choice of a plenoptic configuration, depending on the expected image resolutions and characteristics.

1.6 X-ray plenoptic imaging

In the introduction, we mentioned that the objective of this thesis is to study how plenoptic imaging could be applied in the X-ray domain. Until now, in the state of the art, we reviewed only works that have been achieved in visible light.

1.6.1 Published work

As far as we know, only one team approached X-ray plenoptic imaging. A team from the Institute of Physics of Jagiellonian University proposed an imaging system based on multi X-ray sources, created by multi-capillary focusing optics [DDWK13; SJK18; SKK20]. Although the system does not contain any main lens nor μ LA, it produces an array of sub-images very similar to those obtained with a conventional plenoptic camera. With a single acquisition, they performed 3D reconstruction using limited-angle tomography algorithms. The proposed imaging system intrinsically corresponds to multi-angle radiography, with the sub-images being formed by an array of low divergence sources. On the contrary, the word *plenoptic* is usually understood as an optical system sampling the light-field coming from a single illuminating source. Thus the Jagiellonian microscopy device, with several source locations and no optics between the sample and the sensor, is closer to tomography than to plenoptic.

1.6.2 Challenges of X-ray imaging

An X-ray plenoptic system is expected to offer significant benefits. The well known advantage of X-rays is to make biological samples transparent, opening the way to their full 3D reconstruction. The diffraction limit decreases with the wavelength, leading to expected improvements in resolutions.

However, transposing such a setup in X-rays implies several challenges. First, ray propagation through biological samples raises the question of the impact of refraction. In the X-ray domain, refraction is negligible [Att99]. Therefore, the plenoptic theory, which considers straight line propagation between optics, is still valid. The illumination of the sample is done by transmission through the volume, which implies several conditions on the size and divergence of the source.

The second challenge lies in the X-ray optics. In the soft X-ray range (between $\sim 30\text{eV}$ and $\sim 5\text{keV}$), X-rays are strongly absorbed by any material. Thus it is better to avoid refractive optics. The optics could be mirrors or diffractive optics, like a Fresnel Zone Plate. High quality mirrors (for sub-100nm resolution) are extremely expensive, while diffractive optics can reach today sub-10nm resolution at reasonable cost.

In the hard X-ray range (between $\sim 5\text{keV}$ and $\sim 100\text{keV}$), the optics constituents become very transparent. In that case, mirrors are much less efficient, they can only work under grazing incidence (less than a few degrees). To keep a high numerical aperture, the mirrors have to be very long, typically a meter or more. Such kinds of mirrors exist for synchrotrons, but are too expensive for our experiment. Refractive X-ray lenses have been demonstrated in 1996, but they have the drawback that the index of refraction is extremely close to 1 [SKSL98]. A single lens has a typical focal length of a few kilometers. Today

it is common to stack few thousands lenses to achieve a focal length in the meter range. However such lenses, generally made in beryllium, are also too expensive. Once again, the only solution is to use diffractive lenses. The numerical aperture of this kind of lenses is proportional to the wavelength [Att99]. As a consequence, it is possible to get relatively high numerical aperture in the soft X-ray range ($10^{-1} - 10^{-2}$), while for hard X-rays, numerical apertures drop to 10^{-4} or less. This property might have consequences in the design and resolutions of an X-ray plenoptic setup.

For diffractive optics, both costs and manufacturing difficulties increase with the surface, thus limiting the number of micro-lenses inside the μ LA. We have seen that the μ LA is of crucial importance, as it has a strong impact on the spatio-angular sampling in a plenoptic setup. This is expected to be a potential constraint when transposing plenoptic to X-rays.

1.7 Conclusion

The objective of the *VOXEL* project is to implement in the X-ray range a plenoptic camera composed of a main lens, a μ LA and a sensor, to achieve 3D X-ray imaging at very low dose. The aim of this thesis is to study the possibility to transpose the above qualities of light-field imaging to X-ray imaging: compactness, reduction of the absorbed dose and of the acquisition process.

In the literature, two configurations can be built from these three elements: the traditional and focused plenoptic cameras, presented with their respective refocusing algorithms. A deep analysis of the resolutions should be conducted, in order to decide the best configuration to build our X-ray plenoptic camera, depending on the available optics. It is important to study separately the experimental and numerical aspects, as it has been shown that they both contribute to the resolutions and quality of the reconstructed image. We have seen that the two refocusing algorithms are actually based on the same concept. To fully exploit the acquired data, we will propose a new refocusing algorithm, valid for any configuration.

With an X-ray plenoptic setup, we expect to be able to reconstruct and separate different depth planes in the scene, as a first step to 3D reconstruction. We have seen that stereo algorithms might not be appropriate, and adapting tomography techniques is the approach chosen by other members in the *VOXEL* project. We will follow a different approach. Using the refocusing algorithm, a stack of refocused 2D images will be generated by scanning over the whole depth range. From this refocused stack, the in-focus elements should be extracted at each depth, based on their sharpness. These treated images could then be combined to reconstruct the 3D scene. We will present the developments of our proposed approach in the following chapters.

Chapter 2

Optical system: equations and resolutions

As introduced in Chapter 1, the plenoptic systems studied in this thesis are composed of a main lens, an array of micro-lenses (μ LA) and a sensor. Two main configurations have been built from these three optical elements: a traditional plenoptic camera (also called plenoptic 1.0) and a focused plenoptic camera (or plenoptic 2.0). In previous chapter, we presented the two configurations and a general comparison between them. In this chapter, we present additional elements of comparison. We provide general equations that guide the construction of a plenoptic imaging system, including conditions on the light source in the case of illumination by transmission. These equations meet the needs of accuracy of X-ray plenoptic imaging.

We define our own theoretical formulas to estimate the expected resolutions depending on the parameters and distances of the setup. Four resolutions are studied: the lateral and depth resolutions, that correspond to the finest structures that could be detected in the reconstructed image, the field of view and the plenoptic depth of field, that delimit the volume which can be reconstructed from a single plenoptic acquisition.

These theoretical resolutions are validated with images acquired by two different visible optical setups. The resolutions were studied through a visual assessment over the reconstructed images, completed by contrast curves extracted from these images. Some simulations were performed to better understand how contrast varies with depth along the whole depth range.

2.1 Optical design

It is first important to define the theoretical framework that leads to the optical design of a plenoptic camera. The generic equations for building a plenoptic camera are studied in details, especially the condition of matching between the numerical apertures. Some underlying approximations are highlighted. We

then focus on illumination conditions in the case of transmitted light. This results in specific requirements on the shape of the source illuminating the scene.

2.1.1 Continuity between the optical setups

The two generic configurations, the traditional and focused plenoptic cameras, only differ in the distances between the optical components: main lens, μ LA and sensor. Their designs are both based on the thin lens equation and on a condition of aperture matching between the main lens and the μ LA.

These two configurations are usually presented from the point of view of a camera or a microscope being modified to build a plenoptic system [Ng06; LNA⁺06; Lam15]. In this way, it corresponds to adapting the design of an already existing imaging system to the structure of plenoptic camera. This is done by adding a μ LA either at the plane where the sensor was originally located (traditional configuration), or at a shifted location from this plane (focused configuration).

The μ LA-sensor distance determines the configuration. Depending on this distance being equal, larger or smaller than the focal lens of the micro-lenses, respectively defines a traditional or focused configuration, either in Keplerian or Galilean mode (see Chapter 1). In addition, using the thin lens equation successively on the μ LA and on the main lens, the positions of the intermediate image plane, and then of the object-plane can be calculated. The distance b between sensor and μ LA is therefore responsible for a privileged position between main lens and μ LA (the intermediate image plane). Considering also the physical distance c between the optics, this plane defines a privileged position in object space, which is the only plane whose image is perfectly on focus with the sensor plane. As a result, the configuration (either traditional or focused) of the system, and the depth of this privileged plane, are intrinsically determined by the physical positions of the different elements of the setup.

Theoretically, for each depth z_0 , the distance b should be adjusted accordingly so that the image on the sensor is perfectly in-focus. However, we experimentally notice some flexibility in the positioning of the object. Due to the depth of focus of the micro-lenses (see Section 2.2.5), the same plenoptic camera can acquire raw images typical of the three configurations, without changing the value of b . This phenomena is shown in Figure 2.1. Using one of our home-made plenoptic cameras, images were acquired for various positions of the object, without moving the optical elements (c and b fixed). Depending on the depth of the object, the image progressively presents features typical of focused Galilean, traditional and then focused Keplerian configurations (see Section 1.2). This property of acquiring images typical of different configurations using the same experimental setup illustrates the **continuity** between the plenoptic cameras. This has already been observed in the litera-

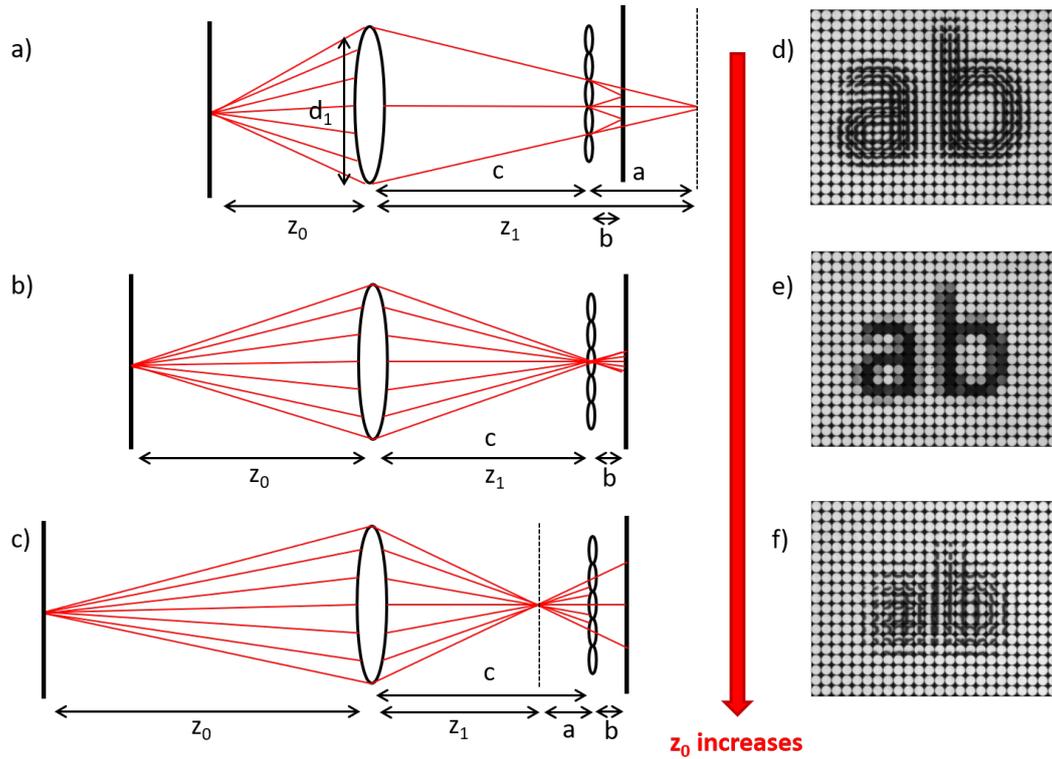


Figure 2.1 – Using the same setup and distances, images were acquired at several depth of the object "ab". We recognize typical raw images of the focused Galileian (scheme a and image b), traditional (scheme c and image d) and focused Keplerian case (scheme e and image f).

ture [LG08b; ZJW16].

2.1.2 Aperture matching condition

We now go deeper into the condition of aperture matching. The role of this condition is to optimize the use of pixels on the sensor, by maximizing the size of the sub-images until they are adjacent. For both traditional and focused configurations, previous works have specified that the image-side numerical aperture of the micro-lenses should match the image-side numerical aperture of the main lens (sometimes mentioned in terms of F-number matching) [Ng06; LNA⁺06; GCL11].

Here we propose a different approach. We calculate the size and period of the sub-images separately. Having adjacent sub-images corresponds to an equality between them. Figure 2.2 shows how these two quantities can be calculated. To estimate the periodicity of the sub-images on the sensor, we consider the green rays connecting the center of the main lens to the centers of the micro-lenses until the sensor plane. In the context of geometric optics, they are not deviated when passing through the centers of the micro-lenses. For

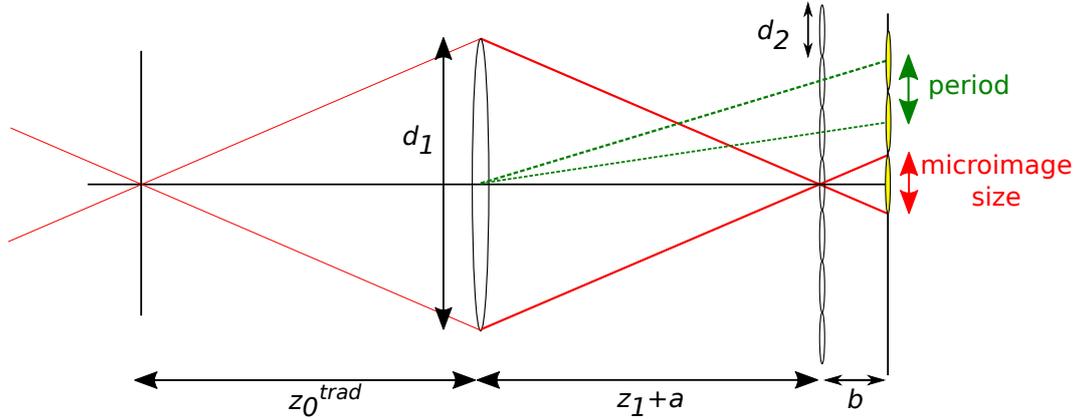


Figure 2.2 – Calculation of the period and size of the sub-images on the sensor.

both traditional and focused configurations, the shape of the sub-images on the detector corresponds to the shape of the main lens. Therefore, the periodicity of the center of the main lens projected on the sensor corresponds to the periodicity of the whole sub-images. With this scheme can be deduced that the sub-images are regularly positioned on the sensor, with a constant periodicity. Using the property of similar triangles, this period can be quantified as:

$$period_{sub-image} = d_2 \frac{z_1 + a + b}{z_1 + a} \quad (2.1)$$

We use the red rays of Figure 2.2 to estimate the maximal extent of the sub-images. This results in the following formula:

$$size_{sub-image} = d_1 \frac{b}{z_1 + a} \quad (2.2)$$

Imposing adjacent sub-images on the sensor can be written as an equality between Equations 2.1 and 2.2, leading to:

$$\boxed{\frac{d_1}{z_1 + a + b} = \frac{d_2}{b}} \quad (2.3)$$

Equation 2.3 is our proposed equation of aperture matching. The equation is slightly different from previous works. Compared to Equations 1.3 and 1.5 in Chapter 1, the distance z_1 between main lens and μ LA has increased by the distance $a + b$. This difference can be explained by an underlying assumption between the focal lengths of both main lens and micro-lenses. In most works, the authors consider that the focal length of the micro-lenses is smaller than the one of the main lens ($f_2 \ll f_1$) [GL10b; Ng06]. The consequence of this assumption is that the distances related to the micro-lenses (distances a and b) can be neglected compared to the distances related to the main lens (distances

z_0 and z_1). When applying $a, b \ll z_1$ in Equation 2.3, we retrieve the previous Equations 1.3 and 1.5.

In the cited works, this assumption is justified as $f_2 \ll f_1$. This is not always the case, for example in our X-ray plenoptic camera (see chapter 5). During our studies in the visible, we experimentally noticed that our proposed equation 2.3 was more accurate in order to obtain adjacent sub-images on the sensor. This shows that this assumption and the resulting difference in the aperture matching equation can impact the design of the optical setup. In section 2.2.6, we will show the consequences on the formulas of resolutions.

2.1.3 Conditions on illumination

Finally, we present here additional conditions concerning the illumination of the plenoptic system. In photography, illumination is done directly on the scene [Ng06; GL10b]. The scene is enlightened by ambient light, which reflects over the surfaces of the objects. These surfaces can be considered as secondary sources, whose extent is the size of the scene and whose angular range depends on the reflection properties of the objects. In the case of Lambertian surfaces, the light is considered to be equally reflected in all angular directions [Kop14]. This illumination **by reflection** guarantees that the light-field entering the plenoptic camera contains the needed spatio-angular information to reconstruct the scene.

In plenoptic microscopy, the illumination is usually done **by transmission** through the sample [LNA⁺06; MDI15]. In visible microscopy, a diffuser [LNA⁺06] or Köhler illumination lens [Tur16] could be added to ensure a correct illumination of the sample. This guarantees that the sample is well illuminated, with a sufficient spatial extent and angular range. This way, the spatio-angular information acquired by the camera is only limited by the geometry of the plenoptic system, and is not restricted by the illumination.

In our X-ray plenoptic experiment, the illumination is also done by transmitted light through the sample. However, the characteristics of the source and optics limit the available spatio-angular data that can be acquired by the system (see Chapter 5). We need to ensure that the main lens and the array of micro-lenses are fully illuminated by the source. We establish two conditions on the size and divergence of the source. These conditions guarantee a good illumination of the plenoptic system.

In practice, it is sufficient to concentrate on the extreme rays between main lens and μ LA. In Figure 2.3, the light rays reaching the bottom of the μ LA are drawn. They correspond to the maximal possible angles and positions inside the volume between main lens and μ LA. According to the thin lens equation, the three rays intersect on the plane located at z_0^{trad} , optically conjugated with the plane of the μ LA. The quantity δd_1 can be calculated as:

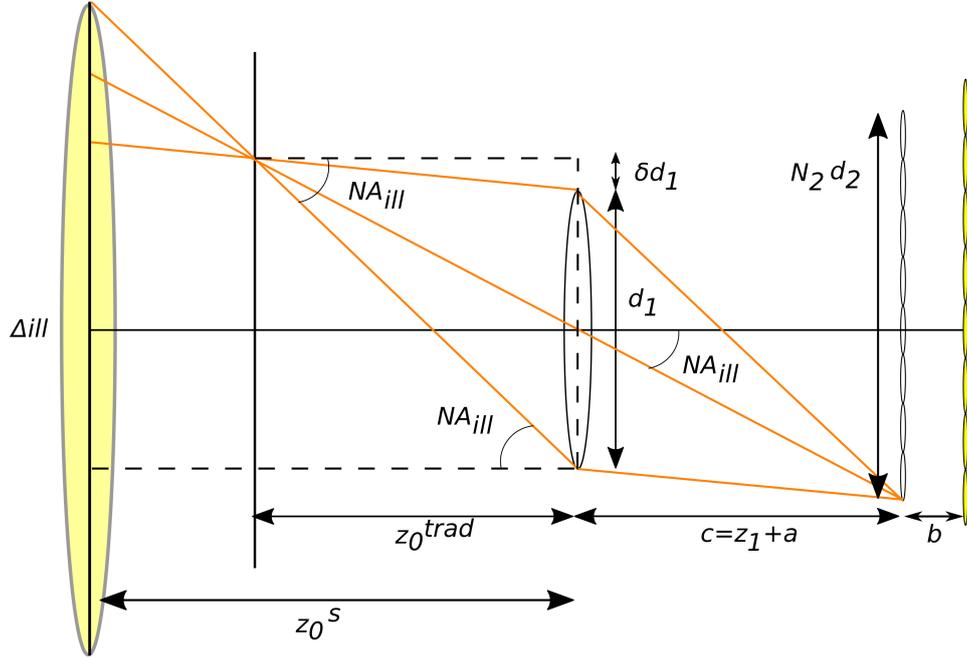


Figure 2.3 – Conditions on illumination of the system: both size and divergence of the source should be large enough to illuminate the whole main lens and the array of N_2 micro-lenses of pitch d_2 .

$$\frac{\left(\frac{d_1}{2} + \delta d_1\right)}{z_0^{\text{trad}}} = \frac{N_2 d_2}{2c} \quad (2.4)$$

$$\delta d_1 = \frac{z_0^{\text{trad}}}{c} \frac{N_2 d_2}{2} - \frac{d_1}{2} \quad (2.5)$$

These rays allow to determine the maximum angles in object-space that could be acquired. It imposes that the numerical aperture of the illuminating source covers this angular extent. We define NA_{ill} , which is the minimal angular range that the source should have in order to illuminate the whole main lens and the successive μ LA. Its expression is:

$$NA_{ill} = \frac{(d_1 + \delta d_1)}{z_0^{\text{trad}}} \quad (2.6)$$

$$NA_{ill} = \frac{d_1}{2z_0^{\text{trad}}} + \frac{N_2 d_2}{2c} \quad (2.7)$$

This equation is the angular condition for the source. If the half divergence of the source is lower than NA_{ill} , less data might be acquired by the plenoptic camera, resulting in lower performance and resolutions. It is interesting to notice that NA_{ill} can be written as the sum of two angular components: the

object-side numerical apertures of the main lens ($\frac{d_1}{2z_0^{trad}}$) and of the μ LA ($\frac{N_2d_2}{2c}$). NA_{ill} can be interpreted as a combination of two angular ranges, the first one imposed by the main lens and the second one by the μ LA. This equation can be considered as an extension of the aperture matching condition, considering the illuminating source.

Moreover, the spatial extent of the source should be large enough to include the orange rays in Figure 2.3. The source of minimum size Δ_{ill} is located at z_0^s from the main lens. The condition on the size of the source can be written:

$$NA_{ill} = \frac{(\frac{\Delta_{ill}}{2} + \frac{d_1}{2})}{z_0^s} \quad (2.8)$$

$$\boxed{\Delta_{ill} = N_2d_2\frac{z_0^s}{c} + d_1(\frac{z_0^s}{z_0^{trad}} - 1)} \quad (2.9)$$

We notice that the condition on the size of the source depends on its position, which was not the case for NA_{ill} (see Eq. 2.7).

To conclude, three sets of equations should be respected when designing a plenoptic camera: the thin lens equation (Eq. 1.1 and either Eq. 1.2 or 1.4), the aperture matching condition (Eq. 2.3) and the illumination conditions (Eqs. 2.7 and 2.9). Respecting these equations guarantees that the plenoptic system is optimally illuminated, with adjacent micro-lenses occupying the whole sensor. In this context, the resolutions are only limited by the geometry imposed by the optics and the sensor.

2.2 Theoretical studies on resolutions

In this section, we define the theoretical resolutions of a plenoptic imaging system.

We wish to warn the reader for potential ambiguity around the different meanings of the word **resolution**. The sensor resolution usually corresponds to the number and size of the pixels on the sensor, describing its spatial precision. An image resolution corresponds to the size of the smallest details that are visible in the image. Usually the image resolution relates to the number of pixels over the whole image (which can be called pixel resolution), but it is sometimes used to describe how the image has been affected by the acquisition process (sometimes called spatial resolution). The optical resolution is the ability of an imaging system to acquire the finest possible details of the scene [Bal08]. It can be defined as the ability to separate two incident rays (angular resolution) or two points on the object plane (spatial resolution). The projection of the rays on a plane links the angular and the spatial definitions together. In the context of plenoptic imaging, we could also mention the nu-

merical resolutions, which would describe the performance of an algorithm to render finer details in a refocused image.

In this section we concentrate on optical resolutions, constrained by the acquisition process and defined as spatial quantities in object-space. We wish to evaluate the resolution limits due to the optical setup, independently from the refocusing algorithm used. The proposed formulas aim at describing the intrinsic resolutions contained in a single plenoptic raw image. They indicate how the spatio-angular data are distributed during the acquisition process.

Four resolutions will be studied in this section: the lateral (res_{lat}) and depth (res_{depth}) resolutions describe the smallest structure that could be reconstructed from a single plenoptic image. On the other hand, the field of view (FOV) and the plenoptic depth of field (DOF_{pleno}) represent the lateral and depth limits of the volume that could be reconstructed from a single acquisition (see Fig. 2.4). Spatial (N_s) and angular (N_u) samplings are also defined. In that way, a volume of size $FOV \times FOV \times DOF_{pleno}$ could be reconstructed, containing $N_s \times N_s \times N_u$ effective voxels of size $res_{lat} \times res_{lat} \times res_{depth}$.

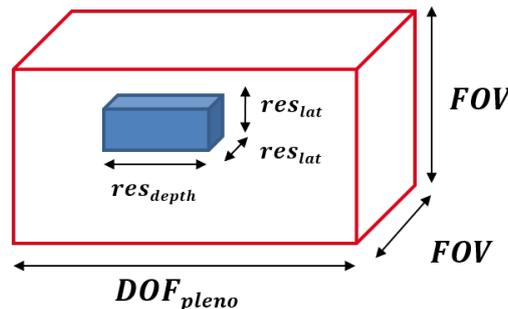


Figure 2.4 – Schematic representation of the four resolutions, with the lateral resolution res_{lat} , the field of view FOV , and along the optical axis, the depth resolution res_{depth} and the plenoptic depth of field DOF_{pleno} .

Actually, the image quality along this volume is heterogeneous. Because of the angular variety among the light rays, the spatio-angular sampling varies through the optical system. This affects the four resolutions across the whole volume. As a consequence, the real shape of the reconstructed volume, as well as the effective voxels, are not perfectly rectangular boxes. Their shapes correspond more to diamonds, with resolutions and magnification decreasing when depth moves away from z_0 (see Figure 4 in [IRMD16] and Figure 5 in [MDRI17]). As a result, the resolutions defined here are valid at depth z_0 , where they are assumed to be maximal. Besides this depth, the field of view, the lateral and depth resolutions are expected to decrease as a function of depth. The theoretical formulas of each resolution will be established, taking into account parameters and distances of the setup. The proposed equations are independent from whether or not the aperture matching condition is respected.

2.2.1 Field of view (FOV)

The field of view corresponds to the lateral width of the reconstructed volume at depth z_0 (see Fig. 2.4). It is sometimes interchangeably used instead of the *angle of view*, which is the angular extent of the scene that is visible by the camera [Dob12]. Actually, the field of view corresponds to the projection of the angle of view on a plane in object-space.

In a plenoptic system, the field of view can be limited either by the size of the μ LA or by the size of the sensor, defined by $N_p \cdot \Delta_p$, with N_p the number of pixels of size Δ_p . Using the period of the sub-images on the sensor (see Eq. 2.1), it is possible to determine which of the two is the limiting factor. The maximal number of sub-images that fits onto the sensor can be calculated by:

$$N_{period} = \frac{N_p \Delta_p}{period_{sub-image}} \quad (2.10)$$

The minimum between N_{period} and the number N_2 of micro-lenses on the μ LA gives the total number of visible micro-lenses that limits the field of view. We define N_{2eff} as the effective number of usable micro-lenses:

$$N_{2eff} = \min(N_2, N_{period}) \quad (2.11)$$

In a traditional plenoptic configuration, the image is formed directly on the μ LA. The field of view corresponds to the spatial extent of these N_{2eff} visible micro-lenses of pitch d_2 . On the plane of the μ LA, it equals:

$$fov_{trad} = d_2 N_{2eff} \quad (2.12)$$

Going back in object-space, and considering the magnification $M = z_1/z_0$ of the main lens results in :

$$FOV_{trad} = \frac{z_0 d_2 N_{2eff}}{z_1} \quad (2.13)$$

In a focused plenoptic camera, the reasoning is more complex. The field of view corresponds to the lateral area where a sufficient spatio-angular data has been acquired. Due to the angular variations of the rays, several definitions can be considered, as illustrated in Figure 2.5. In the central part (green rectangle), each spatial position has been acquired by a maximal number of rays. This relates to the angular sampling N_u that specifies the number of angles per spatial position (see Section 2.2.3). On the contrary, the outer parts (dark rectangle) located at the borders of the field of view were acquired by fewer light rays, and less angular data are available in the raw plenoptic image.

According to the chosen criterion, three different formulas can be derived. In a strict definition, each spatial position should be imaged by exactly N_u angles (green rays). On the contrary, the loosest definition defines the field of

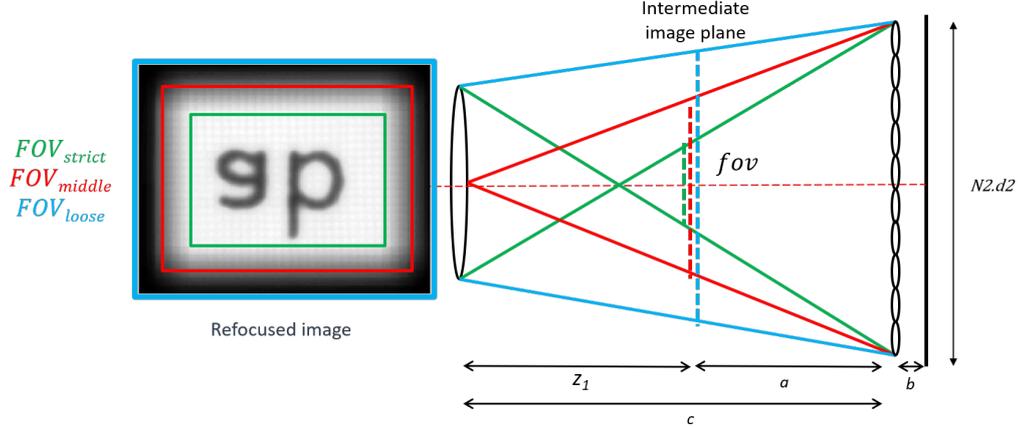


Figure 2.5 – The three possible definitions of the field of view: the strictest (green), middle (red) and loosest (blue) definitions. The light rays show how the formulas of the three fov definitions are established (equations 2.14-2.15-2.16), and the reconstructed image illustrates the different levels of vignetting.

view as the area where each position has been acquired by at least one angle (blue rays). In between these two definitions, we also define a middle case where we consider the positions acquired by at least $N_u/2$ angles (red rays).

In Figure 2.5 are drawn the extreme rays that delimit the spatio-angular data acquired by the system. With simple geometry, it is possible to derive the following formulas at the intermediate image plane:

$$fov_{strict} = \frac{z_1 d_2 N_{2eff} - d_1 a}{c} \quad (2.14)$$

$$fov_{middle} = \frac{z_1 d_2 N_{2eff}}{c} \quad (2.15)$$

$$fov_{loose} = \frac{z_1 d_2 N_{2eff} + d_1 a}{c} \quad (2.16)$$

The two extreme formulas define a zone where some vignetting occurs in the reconstructed image. The intermediate formula is the one that we decide to keep for the experimental configurations. To express it in the object-space, the magnification M of the main lens should be taken into account.

$$FOV_{focused} = \frac{z_0 d_2 N_{2eff}}{c} \quad (2.17)$$

In the end, we realize that the two formulas for a traditional (Equation 2.13) and for a focused configuration (Equation 2.17) converge to a single formula. This shows the continuity between the two configurations.

This equation is valid provided that the conditions on illumination are respected. If the size and divergence of the source are large enough (see Equations 2.7 and 2.9), the plenoptic system allows to study an object of lateral size

FOV. Otherwise, the observed field of view is restricted by the illumination itself.

2.2.2 Lateral Resolution (r_{estlat})

Apart from the field of view, the other three resolutions of the system are established through a geometric reasoning, taking into account diffraction. As illustrated in Fig. 2.6, the lateral resolution corresponds to the smallest lateral extent in object-space and can be determined from Δ_s , the smallest pixel in the intermediate image plane.

An optical imaging system is limited by diffraction. Following the work of [LNA+06], we consider the Sparrow criterion [Spa16; Hec17], that defines the resolution limit as the smallest distance for which an intensity decrease could be observed at the midpoint between the images of two point sources. We define R_{obj} at the intermediate image plane as the minimal distance necessary to be able to separate the images of these two points. R_{obj} depends on the object side numerical aperture ($NA_{obj\ ML} = d_1/2.z_0$) and on the wavelength (λ) of light [IO94; LNA+06]. When expressed in the intermediate image plane, its formula is:

$$R_{obj} = \frac{0.47\lambda}{NA_{obj\ ML}}M \quad (2.18)$$

The size of the spot due to diffraction needs to be compared to the physical size of the pixel on the detector. We define the effective pixel size Δ_{peff} as the smallest distance that can be resolved on the sensor plane, which is limited either by the sensor pixel or the diffraction spot.

In the case of a traditional plenoptic camera, the image is formed on the plane of the μ LA after going through the main lens. The effective pixel size is calculated as the maximal between the size of the pixel of the detector (Δ_p) and R_{obj} ([LNA+06]):

$$\Delta_{peff-trad} = \max(\Delta_p, R_{obj}) \quad (2.19)$$

In a focused camera, the main lens forms an image at the intermediate image plane, which is then relayed by the μ LA on the sensor. Diffraction occurs at these two stages. The diffraction spot due to the main lens needs to be projected to the sensor plane to be compared to the pixel size. This is done by multiplying R_{obj} by the magnification $m = b/a$ of the μ LA. The micro-lenses are also responsible for diffraction. Following the same approach, we define a limiting distance r_{obj} that describes the diffraction spot due to each micro-lens. This time, the magnification and the numerical aperture used are the ones of the μ LA:

$$r_{obj} = \frac{0.47\lambda}{NA_{obj\ \mu LA}}m \quad (2.20)$$

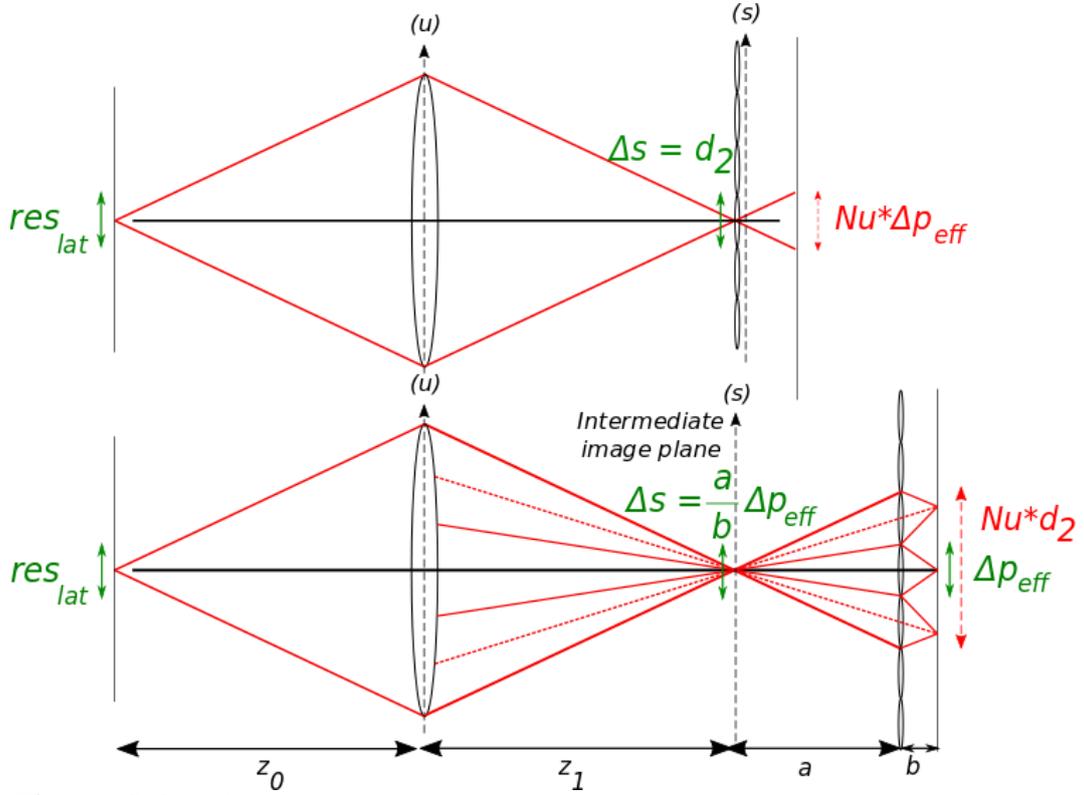


Figure 2.6 – Schematic representation of the traditional (top) and the Keplerian focused (bottom) cameras, showing how the lateral resolutions are calculated from Δ_s in the intermediate object plane.

with $NA_{obj \mu LA} = d_2/2a$. The effective pixel size corresponds to the maximum between the sensor pixel and these two diffraction criteria:

$$\Delta_{peff-foc} = \max(\Delta_p, mR_{obj}, r_{obj}) \quad (2.21)$$

We define the pixel size Δ_s on the intermediate image plane in both configurations (see Figure 2.6). The pixel on this plane is used to estimate the lateral resolutions, as well as the depth resolutions and plenoptic depth of field.

In the traditional camera, the pitch d_2 of the micro-lenses is responsible for the spatial sampling of this image. The pixel size Δ_s on this plane is :

$$\Delta_{s-trad} = d_2 \quad (2.22)$$

In the focused camera, the image is formed by the main lens on the intermediate image plane located at a distance a from the μLA . In this configuration, the spatial and angular information are mixed together, distributed across the micro-lenses and the pixels in the sub-images. Expressed on the intermediate image plane located between the main lens and μLA , Δ_s corresponds to the effective pixel size $\Delta_{peff-foc}$ divided by the magnification $m = b/a$ of the μLA :

$$\Delta_{s-foc} = \Delta_{peff-foc}/m \quad (2.23)$$

The lateral resolution of both configurations corresponds to Δ_s projected back in the object-space:

$$Res_{lat} = \frac{\Delta_s}{M} \quad (2.24)$$

with $M = z_1/z_0$, the magnification of the main lens.

2.2.3 Spatio-angular sampling (N_u and N_s)

In a focused plenoptic configuration, we have seen that the field of view defines a zone where the spatio-angular sampling is maximal and considered homogeneous. This spatio-angular sampling can be quantified from the parameters of the plenoptic system. N_s is the spatial sampling, also called spatial resolution in [Ng06]. It corresponds to the number of different spatial positions acquired by the system. On the opposite, N_u is the angular sampling, also called the directional resolution [Ng06]. N_u corresponds to the number of different angular components collected for each spatial position.

In the traditional camera (see Fig. 1.4), the spatial sampling is done by the micro-lenses, with each micro-lens acquiring a different position. Therefore the spatial sampling N_{strad} equals N_{2eff} , the effective number of micro-lenses:

$$N_{strad} = N_{2eff} \quad (2.25)$$

On the contrary, the angular sampling is done by the sensor. All the light rays reaching the same sub-image belong to the same spatial position (they went through the same micro-lens). The angular sampling is done by each pixel of the sub-image acquiring a different direction, as illustrated by Fig. 1.4. However, due to the effect of diffraction, we need to consider the effective pixel $\Delta_{peff-trad}$ instead of the physical pixel of the sensor [LNA+06]. Nu_{trad} is therefore given by the total number of effective pixels per sub-image of size d_1b/z_1 (see Fig. 2.6):

$$Nu_{trad} = \frac{d_1b}{z_1\Delta_{peff-trad}} \quad (2.26)$$

For the focused plenoptic camera, the spatio-angular sampling is considered from the point of view of a pixel located on the intermediate image plane. The spatial sampling corresponds to the maximal number of pixels of size Δ_{s-foc} that could be acquired by a well illuminated system. In the intermediate image plane, the field of view (see Eq. 2.15) corresponds to $N_{s_{foc}}$ different pixels of size Δ_{s-foc} that were acquired. We obtain the following formula:

$$N_{s_{foc}} = \frac{z_1d_2N_{2eff}}{c.\Delta_{s-foc}} \quad (2.27)$$

Using the established equations of N_s , res_{lat} and FOV in both configurations, we notice a general relationship:

$$FOV = N_s.res_{lat} \quad (2.28)$$

This equation is coherent with the meaning of the spatial sampling N_s . As introduced at the beginning of this Section 2.2, at depth z_0 it is possible to reconstruct N_s pixels of size res_{lat} , for a total extent of size FOV .

The angular sampling N_u corresponds to the number of angular data acquired for one single spatial position s on the intermediate image plane (see Fig. 2.6). For a given spatial position, the angular range is spread over several micro-lenses. To estimate the angular sampling N_u , we need to count the number of different micro-lenses that acquire part of this angular range. The spatial position s is illuminated by the aperture of the main lens of size d_1 , which then spreads over N_u micro-lenses of pitch d_2 . Considering the distances z_1 and a , the intercept theorem gives:

$$N_{u_{foc}} = \frac{d_1|a|}{d_2z_1} \quad (2.29)$$

2.2.4 Depth resolution (res_{depth})

The resolution along the depth axis is now examined. The depth resolution corresponds to the ability to separate features of different planes along the depth axis. This resolution is more difficult to establish because it depends on both the angular and spatial samplings.

Actually, the depth resolution corresponds to the notion of depth of field in classical imaging. In classical imaging, a single lens (or several successive lenses) forms an image directly on the detector, as in photography or microscopy. In this context, the depth of field describes the depth range where the structures are considered as sharp. The general relationship $DOF = res_{lat}/NA$ can be established from drawing the rays fitting inside a maximal acceptable blur limit, generally the size of the pixel on the sensor [Gro05].

In plenoptic imaging, the same reasoning is applied considering Δ_s at the intermediate image plane [LNA⁺06]. In figure 2.7, the blue and green rays correspond to the extreme front and back depths, for which the rays still fit inside the spatial extent Δ_s . This reasoning is valid for both traditional and focused configurations, considering the appropriate formula of Δ_s . This results in the geometrical term of the depth resolution:

$$res_{depth-geom} = \frac{\Delta_s}{M.NA_{obj} ML} \quad (2.30)$$

Based on the work of [LNA⁺06], the depth resolution is defined by the sum of two terms. The first one is the geometrical depth resolution and the second

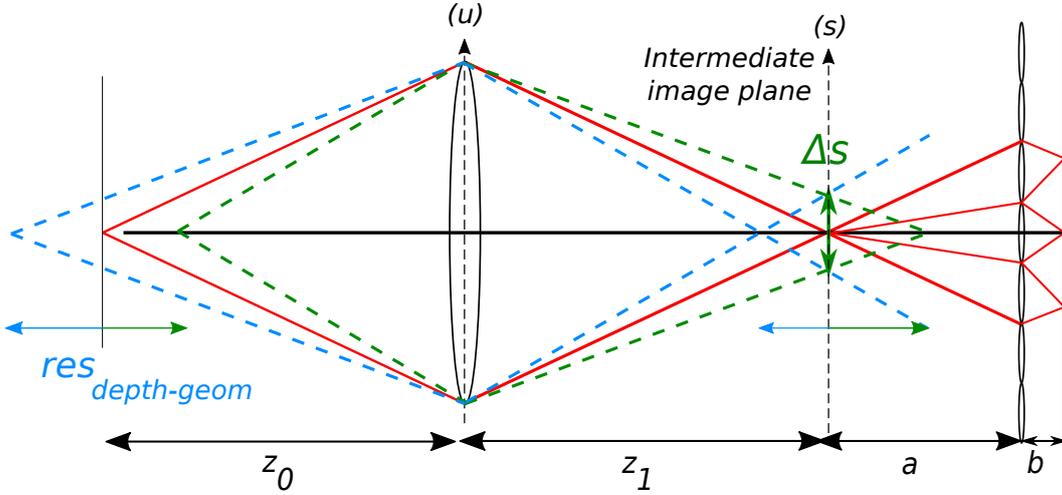


Figure 2.7 – Reasoning for the geometrical term of the depth resolution: it corresponds to the depth of focus of the main lens in the intermediate image plane, i.e. the depth along which the projected rays stay inside Δ_s .

one is a pure wave optics component. It corresponds to the spatial extent of the diffraction pattern along the depth axis [IO94]. The final equation of depth resolution combines the geometric and wave terms together:

$$res_{depth} = \frac{res_{lat}}{NA_{obj\ ML}} + \frac{\lambda}{NA_{obj\ ML}^2} \quad (2.31)$$

2.2.5 Plenoptic depth of field (DOF_{pleno})

The depth of field of a plenoptic system corresponds to the depth range over which it is possible to reconstruct a sharp image from a single acquisition. Similarly to the depth resolution, the plenoptic depth of field is defined as the sum of two terms, representing the geometric and wave components. As explained in [LNA+06], the wave term remains the same as for depth resolution, because it corresponds to the diffraction limit due to the numerical aperture of the main lens. We here present how to estimate the geometric term in both configurations.

The case of a traditional plenoptic configuration has already been discussed in previous works [LNA+06; PW12; DZC+16]. The whole angular range of the ray bundle coming from the main lens is separated into N_u different ray beams, each of them being acquired by a different Δ_{peff} (see Figure 2.8). Considering the same $\Delta_s = d_2$ and the reduced angular range $NA_{obj\ ML} / N_u$, Equation 2.30 becomes:

$$DOF_{pleno-geom-trad} = \frac{N_u \Delta_s}{M \cdot NA_{obj\ ML}} \quad (2.32)$$

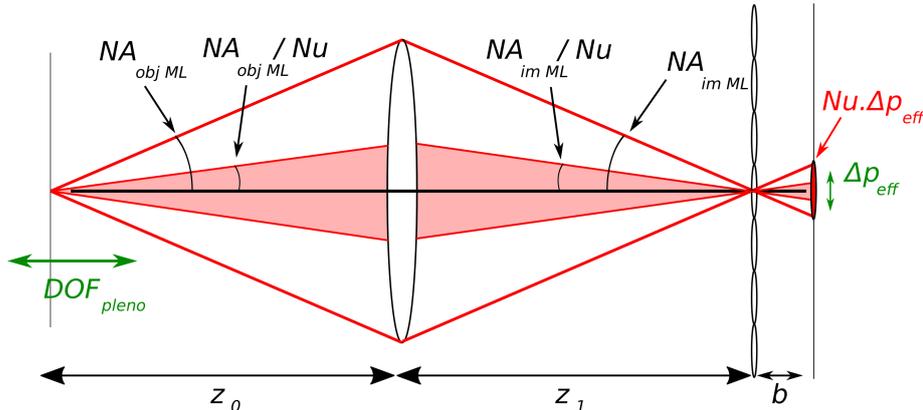


Figure 2.8 – The geometric term of the plenoptic depth of field in a traditional configuration: the ray beam coming from the main lens is divided into N_u beams acquired by different pixels of size Δp_{eff} .

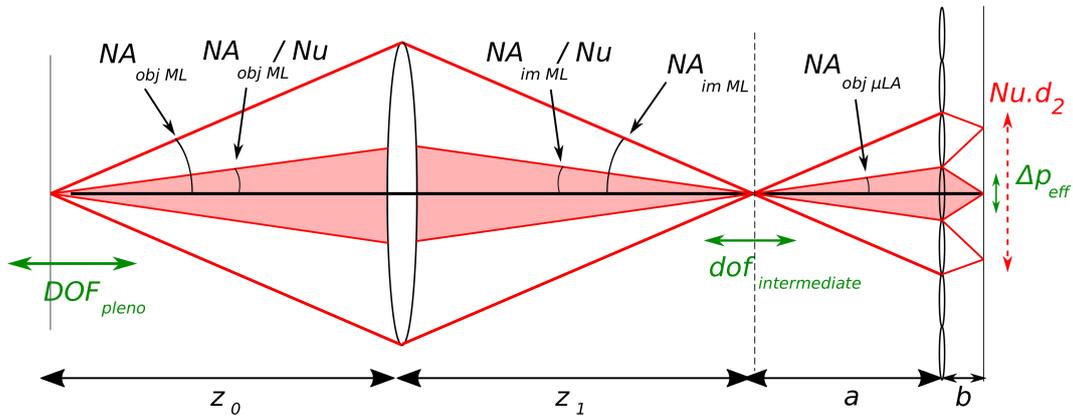


Figure 2.9 – The geometric term of the plenoptic depth of field in a focused configuration: the ray beam coming from the main lens is divided into N_u beams corresponding to the N_u illuminated micro-lenses. It also corresponds to the depth of field of a single pixel of size Δp_{eff} , projected back in intermediate image plane, and then in object plane.

We now express the geometric term for the focused system. Similarly to the traditional configuration, the angular range coming from the main lens is separated into N_u ray beams, each of them being acquired by a different micro-lens and then projected on Δp_{eff} (see Figure 2.9). The same reasoning as for the traditional case could be done, reducing $NA_{obj ML}$ by N_u [Tur16]. This leads to the same Equation 2.32, adapting the formula of N_u and Δ_s to the focused case (see Sections 2.2.2-2.2.3).

Another demonstration of the above result could be performed by the explicit calculation of the depth of field in the system. According to [MCJ18], DOF_{pleno} is the depth of field expressed in the object plane associated to Δp_{eff} . Based on the same reasoning as for res_{depth} , we express the depth of field of Δp_{eff} in the intermediate image plane. In Equation 2.30, we replace Δ_s by

Δ_{peff} and the main lens by the μLA :

$$dof_{intermediate} = \frac{\Delta_{peff}}{m.NA_{obj \mu LA}} \quad (2.33)$$

Finally, to get the plenoptic depth of field, Equation 2.33 needs to be transposed back in object-space. Contrary to lateral resolutions, where equations should be divided by the magnification, the square of the magnification M^2 should be used in the case of resolutions along depth [Gro05]. We obtain:

$$DOF_{pleno-geom-trad} = \frac{\Delta_{peff}}{m.M^2.NA_{obj \mu LA}} \quad (2.34)$$

The numerical aperture used in Equation 2.34 is the object-side numerical aperture of a single micro-lens of the μLA . Based on Figure 2.9, it is N_u times smaller than the image-side numerical aperture of the main lens. To switch from image-side to object-side of the main lens, a division by the magnification of the main lens should be done. We obtain the following relationship:

$$NA_{obj \mu LA} = \frac{NA_{im ML}}{N_u} = \frac{NA_{obj ML}}{N_u.M} \quad (2.35)$$

Considering $\Delta_s = \Delta_{peff}/m$ (see Eq. 2.23), the equation could be simplified and we obtain the same formula as for the traditional case:

$$DOF_{pleno-geom-foc} = \frac{N_u \Delta_s}{M.NA_{obj ML}} \quad (2.36)$$

Considering the wave component, we deduce the following formula, valid for both focused and traditional systems:

$$DOF_{pleno} = \frac{N_u res_{lat}}{NA_{obj ML}} + \frac{\lambda}{NA_{obj ML}^2} \quad (2.37)$$

Lastly, we observe an interesting property between the depth resolution and the plenoptic depth of field. Considering their geometric terms only, we can write:

$$DOF_{pleno-geom} = N_u.res_{depth-geom} \quad (2.38)$$

This equation is the equivalent of Equation 2.28, but this time along the depth axis. This means that inside a volume of depth range DOF_{pleno} , N_u different planes of depth precision res_{depth} could be reconstructed and differentiated. However, this relationship may be impacted by diffraction. In the case of low diffraction, the resolutions follow the following scheme: from a single plenoptic acquisition, we expect to reconstruct a volume of dimensions $FOV \times FOV \times DOF_{pleno}$ made of $N_s \times N_s \times N_u$ effective voxels of size $res_{lat} \times res_{lat} \times res_{depth}$.

2.2.6 Summary of the formulas

The parameters that characterize a plenoptic camera are assembled together in Table 2.1. Some formulas are common between the traditional and focused setups, while others are specific to each geometry. The proposed equations are valid independently from the aperture matching condition.

	Traditional $a = 0, b = f_2$	Focused $m = b/a$
$NA_{obj\ ML}$	$\frac{d_1}{2z_0}$	
M	$\frac{z_1}{z_0}$	
R_{obj}	$\frac{0.47\lambda\ M}{NA_{obj\ ML}}$	
r_{obj}	$\frac{0.47\lambda\ 2b}{d_2}$	
Δ_{peff}	$max(\Delta_p, R_{obj})$	$max(\Delta_p, mR_{obj}, r_{obj})$
Δ_s	d_2	$\frac{\Delta_{peff}}{m}$
N_{2eff}	$min(N_2, \frac{N_p\Delta_p(z_1 + a)}{d_2(z_1 + a + b)})$	
res_{lat}	$\frac{\Delta_s}{M}$	
res_{depth}	$\frac{res_{lat}}{NA_{obj\ ML}} + \frac{\lambda}{NA_{obj\ ML}^2}$	
DOF_{pleno}	$\frac{N_u res_{lat}}{NA_{obj\ ML}} + \frac{\lambda}{NA_{obj\ ML}^2}$	
FOV	$\frac{z_0 d_2 N_{2eff}}{z_1 + a}$	
N_u	$\frac{d_1 b}{z_1 \Delta_{peff}}$	$\frac{d_1 a }{z_1 d_2}$
N_s	N_{2eff}	$\frac{z_1 d_2 N_{2eff}}{(z_1 + a) \cdot \Delta_s}$

Table 2.1 – Overview of the proposed theoretical resolutions.

Some simplifications can be done by assuming that $f_2 \ll f_1$ and $a, b \ll z_1$, as was done in other works (see Section 2.1.2, [GL10b; Ng06]). Considering these approximations, the aperture matching condition becomes $d_1 b = d_2 z_1$,

which is the aperture condition presented in previous works (see Eq. 1.3 and 1.5).

With the condition $d_1 b = d_2 z_1$, the formulas of the spatial and angular samplings can be simplified. In the traditional configuration, the angular sampling becomes:

$$N_{utrad} = \frac{d_2}{\Delta_{peff-trad}} \quad (2.39)$$

In the focused configuration, additionally considering $c \sim z_1$ leads to the simplifications:

$$N_{ufoc} = \frac{|a|}{b} \quad (2.40)$$

$$N_{sfoc} = \frac{N_{2eff} d_2}{\Delta_{s-foc}} \quad (2.41)$$

With Equation 2.40, we retrieve the formula that the angular sampling is related to the magnification of the micro-lenses: $N_u = 1/m$ [GL10a; Tur16]. In the case of aperture matching, the combination of expressions for the spatial and angular samplings, either in the focused case (Equations 2.40 and 2.41) or in the traditional case (Equation 2.39 and Table 2.1), results in an interesting property:

$$N_u \cdot N_s = \frac{N_{2eff} d_2}{\Delta_{peff}} \quad (2.42)$$

We remind that N_{2eff} corresponds to the effective number of micro-lenses imaged on the sensor (see Section 2.2.1), and thus $N_{2eff} d_2$ is the effective area used on the sensor. In other words, $\frac{N_{2eff} d_2}{\Delta_{peff}}$ represents the number of effective pixels that fit in the effective sensor size. Equation 2.42 means that the product of the spatial and angular samplings is constrained by the total number of acquired pixels. This translates the spatio-angular trade-off usually described for focused configurations [GL10a; Tur16].

To summarize, the equations of resolutions (res_{lat} , res_{depth} , FOV and DOF_{pleno}) were established based on geometrical optics while taking into account the limit due to diffraction. The spatio-angular sampling (N_u , N_s), together with the numerical aperture $NA_{obj\ ML}$ and effective pixel size Δ_{peff} due to diffraction, define the optical performances of a configuration. In Chapter 5, we will see how the low NA impacts the resolutions when designing our X-ray plenoptic camera.

Table 2.1 is expected to represent a useful toolbox for anyone wishing to build their own plenoptic camera. In the next section, we will show how these theoretical equations can be validated experimentally.

2.3 Experimental validation: methodology

2.3.1 Presentation of the two setups

To validate the theoretical resolutions, two optical setups were built from separate optical elements that were assembled in the laboratories of Imagine Optic, Bordeaux. This leaves an important freedom in the choice of distances between the different elements, allowing to test different configurations and study the influence of each distance over the final resolutions.

A first set of images was acquired using a wavefront sensor (WFS) from Imagine Optic: the HASO3 128 GE2 Shack Hartmann wavefront sensor (see Fig. 2.12d). The WFS is composed of a μ LA attached to a detector, exactly positioned at the focal plane of the micro-lenses. This corresponds to the traditional configuration ($b = f_2$) (see Eq. 1.2). Combined with a main lens, the WFS is adapted to validate the resolutions of a traditional plenoptic camera. The WFS also has the specificity of providing a μ LA composed of a large number of small micro-lenses. It is made of 128×128 micro-lenses of pitch $d_2 \sim 110\mu m$ and a camera with large aperture ($14.6 \times 14.6 mm^2$). Because of their confidential aspect, the other technical information of the HASO WFS are not given in this thesis. An overview of the mounted setup is presented in Figure 2.10.

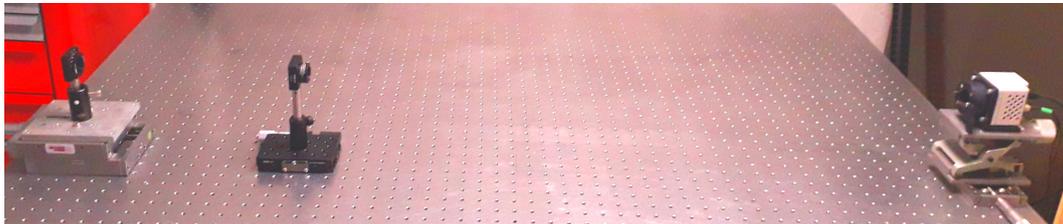


Figure 2.10 – *The experimental HASO setup composed of an object (left), the main lens (middle), and the HASO wavefront sensor (right).*

The second setup is more flexible. It is composed of independent main lens, μ LA and a Stingray F-145B camera from Allied Vision. It is possible to change all the distances in the configuration (see Fig. 2.11). With this setup, three configurations were tested: traditional (or 1.0) and focused (or 2.0) in both Keplerian ($a > 0$) and Galilean ($a < 0$) modes. Contrary to the first setup, this one is composed of a μ LA with fewer micro-lenses that are also larger. Figure 2.12 shows the separate optical elements used in both setups. Comparing the two setups will help in choosing the best configuration adapted to the experimental conditions in X-rays.

Table 2.2 presents the parameters used with the Stingray camera. The housing cover of the front end of the camera was removed, so that its entrance pupil could be positioned as close as needed to the μ LA (see Fig. 2.12e). The μ LA is composed of square micro-lenses of size $300\mu m$ and focal length

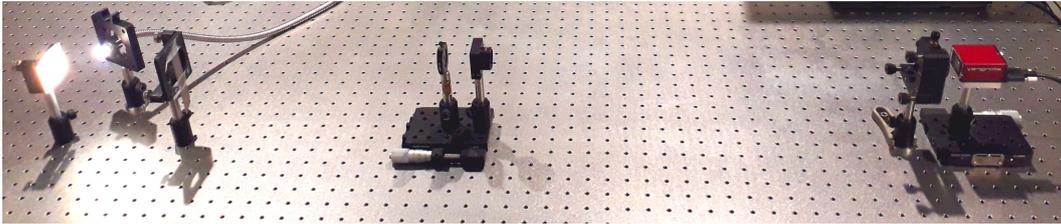
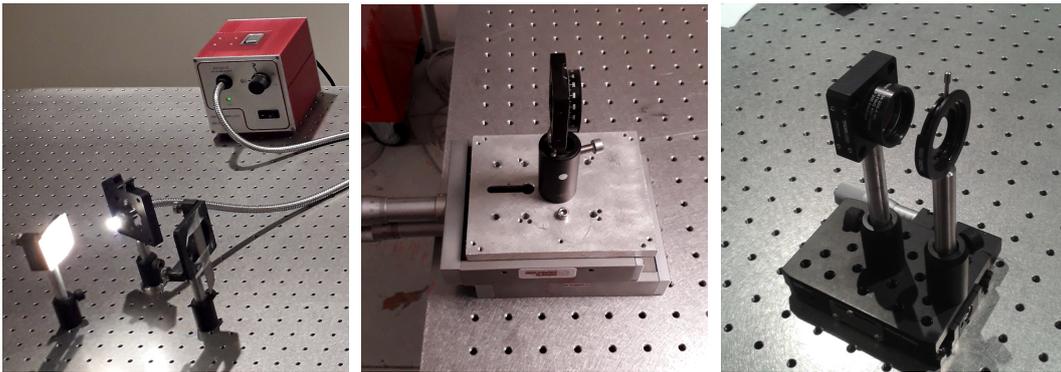


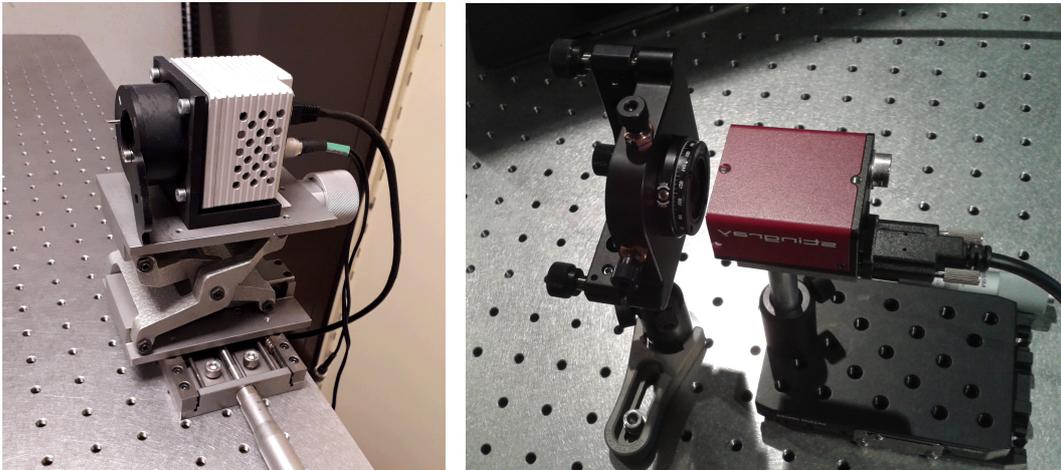
Figure 2.11 – The experimental Stingray setup composed of the illumination system and an object (left), the main lens with a diaphragm (middle), the μ LA and the Stingray camera (right).



(a) Illumination system.

(b) Test target.

(c) Main lens+diaphragm.



(d) HASO wavefront sensor.

(e) Microlens array and Stingray camera.

Figure 2.12 – The optical elements used in the two setups.

18.6mm. The micro-lenses diameter is nearly 3 times larger compared to the WFS ($d_2 = 300\mu\text{m}$ instead of $\sim 110\mu\text{m}$). The number of micro-lenses is also much less: only 20×25 of them were visible on the raw images in the tested configurations. It means that the size of the camera limits the number of micro-lenses used to capture the light-field in the Stingray setup. This will be the limiting factor for the *FOV*.

For both experimental set-ups, the main lens is a doublet with a 200mm

Main Lens	focal f_1	200mm
	Maximal aperture d_1	25mm
Microlens Array	focal f_2	18.6mm
	aperture d_2	300 μ m
	# visible micro-lenses	20 \times 25
Stingray camera	pixel size Δ_p	6.45 μ m
	# pixels	1280 \times 960
Monochromatic approximation	wavelength used λ	520nm

Table 2.2 – Parameters of the optics, sensor and wavelength used in the Stingray setup.

focal length. A diaphragm positioned in front of it allows to adjust its aperture d_1 , with a maximal value of 25mm (see Fig. 2.12c). The optics were fixed on a breadboard on an optical table.

The theoretical resolutions were validated by taking images of a 1951 USAF Resolution Test Target (see Fig. 2.12b). The test target is composed of opaque elements of various sizes on a transparent substrate (see Fig. 2.13). A screen was placed behind the test target, illuminated by a white light lamp (see Fig. 2.12a). The screen plays the role of the illuminating source for the test target, which is thus illuminated by transmission. This system allows to respect the illumination conditions presented in Subsection 2.1.3. Its size and NA are large enough so that they do not limit the resolutions of the acquired images. The polychromatic white source was approximated by a monochromatic source

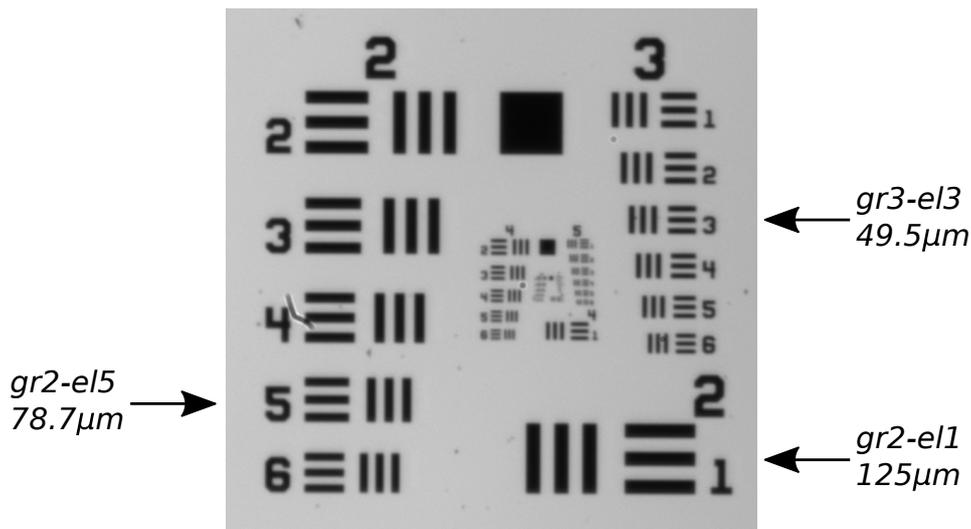


Figure 2.13 – The 1951 USAF resolution test target. The three sets of bars studied in the experiments are highlighted, with the indication of the width of the bars.

of wavelength $520nm$ to calculate the resolutions.

2.3.2 Chosen configurations and expected resolutions

Four different datasets were acquired with these two setups. Thanks to the flexibility of the Stingray setup, it was possible to test three configurations: traditional, focused Keplerian ($a > 0$) and focused Galilean ($a < 0$). On the contrary, the HASO setup allows only to test a traditional configuration. These four datasets give the possibility to experimentally compare two different optical setups in the same traditional configuration, and three different configurations using the same optical setup. The configurations are named as follows:

- Traditional (or 1.0 [LG09]) plenoptic camera:
 - HASO 1.0
 - Stingray 1.0
- Focused (or 2.0 [LG09]) plenoptic camera:
 - Stingray 2.0, $a > 0$
 - Stingray 2.0, $a < 0$

Table 2.3 presents the four configurations. The theoretical resolutions were determined using equations from Table 2.1. Distances z_0 , z_1 , a and b were adjusted to obtain similar theoretical resolutions for all the configurations. The res_{lat} and DOF_{pleno} are almost equal among the four cases. The three Stingray configurations shows similar res_{depth} and FOV , but they are different from the ones in the HASO setup, due to the difference in the optics used. This choice of similar theoretical resolutions allows to verify easily that their experimental resolutions correspond. Only the distances for the Stingray configuration were given to protect the confidential data of the HASO WFS.

The optical system was manually aligned and the distances were adjusted according to the desired configurations. The distances were measured as precisely as possible. The error margin is estimated to a precision of a few millimeters, even if a low error can easily impact the final resolutions, as will be explained in Subsection 2.3.3.

After the installation of each configuration, images were acquired for different positions of the test target, moving it forwards or backwards along the optical axis, with a step of $0.2mm$ over a range of a few centimeters. The positions of the optics were not moved during the series of acquisitions. This results in four datasets containing images with progressive depth locations of the test target. The dataset was then refocused at various depths. The traditional (1.0) or focused (2.0) algorithms were used depending on the setup and the configuration, as will be detailed in Section 2.4.

	HASO	Stingray	Stingray	Stingray
	1.0	1.0	2.0, a>0	2.0, a<0
$z_0(mm)$	-	233	233	231
$c(mm)$	-	1412	1597	1293
$b(mm)$	-	18.6	20.7	17
res_{lat} (μm)	47.1	49.5	49.2	49.9
res_{depth} (mm)	1.85	1.10	1.10	0.74
FOV (mm)	6.04	0.99	0.88	1.07
DOF_{pleno} (mm)	9.56	10.21	10.10	10.37
N_u	6.51	10.87	10.81	11.05
N_s	128	20	17.78	21.48

Table 2.3 – Distances and resolutions of the four different configurations. Due to confidentiality reasons, the distances for the HASO 1.0 configuration are not given.

These refocused images were used to determine the experimental resolutions. They were compared to the expected ones presented in Table 2.3. The patterns of varying sizes over the test pattern allow to verify the lateral resolutions and estimate the field of view. The series of images with different depths of the test target enable to measure experimental depth resolutions and plenoptic depth of field for these configurations.

2.3.3 Distance measurement

In the experimental setups, we realized that the question of the distances was a real issue when building a plenoptic system. When measuring the distances, some practical problems arise. The theory of the optical design and resolutions was made considering the optical distances between the elements, in the context of geometrical optics. It means that it is necessary to locate the optical centers of each lenses, in order to measure the effective distances related to the ones in the equations. For example, in the case of the doublet used as main lens, its thickness was around 20mm and it was difficult to locate exactly its optical center. Moreover, the long distances in the configurations make it even more difficult to measure them with precision, explaining the remaining error of few mm in both setups. This will also be the case in the X-ray experiment that will be presented in Chapter 5.

This imprecision in the distances is problematic in a plenoptic system. Because of its complex geometry and the two combined magnifications of the main lens and the μLA in the acquired image, a small error in the distances may have a huge impact on the resolutions. This is especially the case for such

small resolutions, compared to the large distances of the setup. This can affect the optical properties of the setup and the acquired data, such as difficulties to physical mount the configurations, lower resolutions, or blurred images in the case of out-of-focus data. This will also have consequences in the refocusing process and final output, as the distances will be used as input parameters in our refocusing algorithm (see chapter 3).

2.4 Experimental validation of the resolutions

Figure 2.14 shows an example of a cropped raw image in configuration Stingray 2.0, $a > 0$. In the sub-images we can see the horizontal bars of $gr2 - el1$, which is the largest element of the test target (see Fig. 2.13).

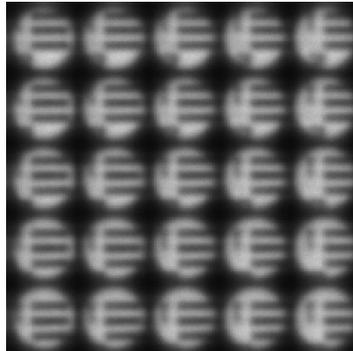


Figure 2.14 – Example of a raw image acquired in configuration Stingray 2.0, $a > 0$, where we can see $gr2 - el1$ (bar width = $125\mu m$).

The acquired datasets from the four configurations were reconstructed with either algorithm 1.0 or 2.0 according to the setup and configuration. The HASO 1.0 dataset was fully reconstructed with algorithm 1.0. Indeed, the refocusing parameter α allows fine choice of refocusing depth, contrary to the integer parameter P in algorithm 2.0. On the contrary, the Stingray 1.0 dataset was reconstructed using both algorithms depending on depth. For the exact 1.0 depth plane at $z_0^{trad} = 233mm$, algorithm 1.0 was used because algorithm 2.0 is not adapted to refocus at this specific depth (see Chapter 1). Except this depth, algorithm 2.0 was used instead. It produces a larger number of pixels in the refocused images, allowing more precision when measuring the resolutions. It is possible to use algorithm 2.0 for the 1.0 dataset from the Stingray setup because of the larger size of the micro-lenses that creates larger sub-images, contrary to the HASO setup. The 2.0 datasets from the Stingray setup (both $a > 0$ and $a < 0$) were also reconstructed with algorithm 2.0. The reconstructed images were then used to estimate the four experimental resolutions in each configuration: field of view, lateral resolution, depth resolution and plenoptic depth of field.

A criterion is needed to quantify the experimental resolutions and compare them to the theoretical ones. Contrast has already been used in other works on plenoptic cameras [BGY⁺13; LR15; Tur16]. It is a simple measure that could be implemented even on a few pixels. Indeed, the varying magnifications and different pattern sizes on the test target may result in bar width covering only a few pixels (for example in Section 2.4.2), making it difficult to draw profiles or extract more complex measures of resolution. In our experimental studies, the contrast used as our criterion is defined by:

$$\text{contrast} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (2.43)$$

where I_{max} and I_{min} are the maximal and minimal intensity values over a chosen area in the reconstructed image.

2.4.1 Field of view

To estimate the *FOV*, only a single image is needed for each configuration. The refocused image at depth z_0 is taken from the four series of images, with z_0 determined by Table 2.3.

Figure 2.15 shows the refocused images corresponding to the four configurations. The whole images are presented allowing to visually evaluate their full extent. The image for the HASO 1.0 case is clearly larger than the images from the Stingray setup. In blue is highlighted the part that is visible in the Stingray setup, which is clearly smaller than the whole HASO image. This is consistent with the *FOV* of the HASO configuration being larger than the other ones (around $6mm$ compared to $1mm$). Between the three configurations from the Stingray setup, the perceived size of the bars is slightly varying, whereas they were displayed with the same size. This is an indication that their *FOV* are slightly different.

To quantify the *FOV* on these images, we concentrate on the three bars of the *gr2 - el1* that is the only group visible in the four configurations. The width of a single bar is known as $125\mu m$, which is also the width of the spacing between two bars. By measuring the number of pixels occupied by the three bars, the *FOV* of the whole image can be evaluated by interpolation. Using this approach, the *FOV* of the four configurations were evaluated as:

- HASO 1.0: $FOV = 6.40mm$
- Stingray 1.0: $FOV = 0.95mm$
- Stingray 2.0, $a > 0$: $FOV = 1.18mm$
- Stingray 2.0, $a < 0$: $FOV = 1.47mm$

The experimental values roughly correspond to the expected ones from Table 2.3. The orders of magnitude match the theory, with the *FOV* in the HASO setup around $6mm$ and the *FOV* in the Stingray setup around $1mm$.

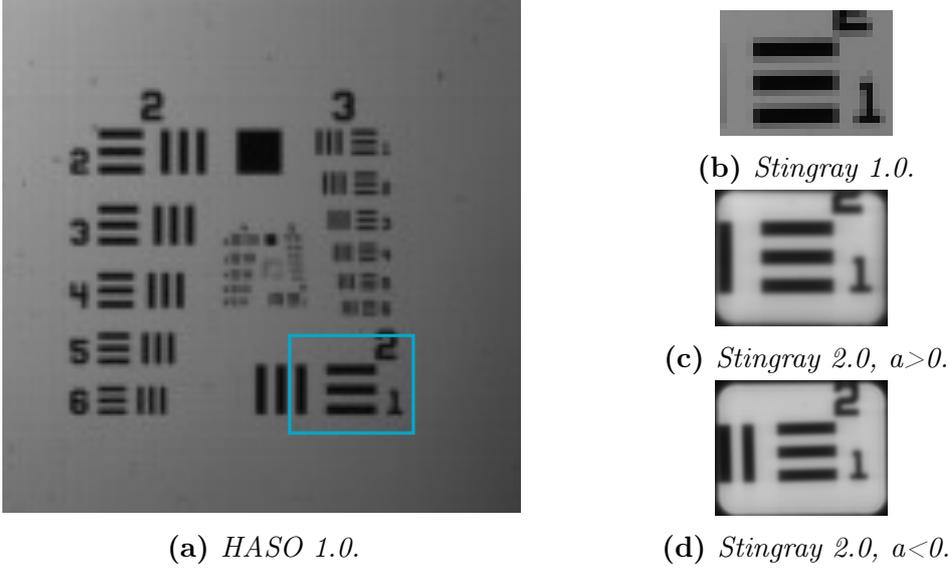


Figure 2.15 – Full reconstructed images used to estimate the experimental field of view in each configuration. The image from the HASO 1.0 dataset has clearly a larger FOV than the three other cases. The blue rectangle in 2.15a shows the only bars that are visible in the images from the Stingray setup because of its reduced FOV.

The relative error is calculated around 5% for the two 1.0 configurations and around 35% for the two 2.0 configurations. This high difference in the Stingray 2.0 configurations comes from the chosen algorithm 2.0 with integration. This algorithm allows to enlarge the FOV of the reconstructed image, until the largest formula of the field of view (see Equation 2.16). It is responsible for the vignetting effect that is visible in Figures 2.15b and 2.15c. On the contrary, the FOV in Table 2.3 were calculated using the medium formula (see Equation 2.17).

The FOV was recalculated for these two cases, removing these extra vignettted borders to match this medium formula. The new values are $0.92mm$ and $1.14mm$ for the Stingray 2.0 $a > 0$ and $a < 0$ configurations. The values now corresponding to errors around 5%, similar to the 1.0 cases. These remaining errors can be easily explained by the measurement uncertainties described in Section 2.3.3.

2.4.2 Lateral resolution

To estimate the lateral resolution res_{lat} , it is necessary to see bars of various sizes on the test target, in order to compare how they are resolved by the optical system. In the Stingray setup, as the field of view was too small, only one group of bars was visible for each acquisition. To overcome this issue, several images were acquired after moving the test target laterally inside the z_0 depth plane. The scan is then used to virtually restore the whole test target.

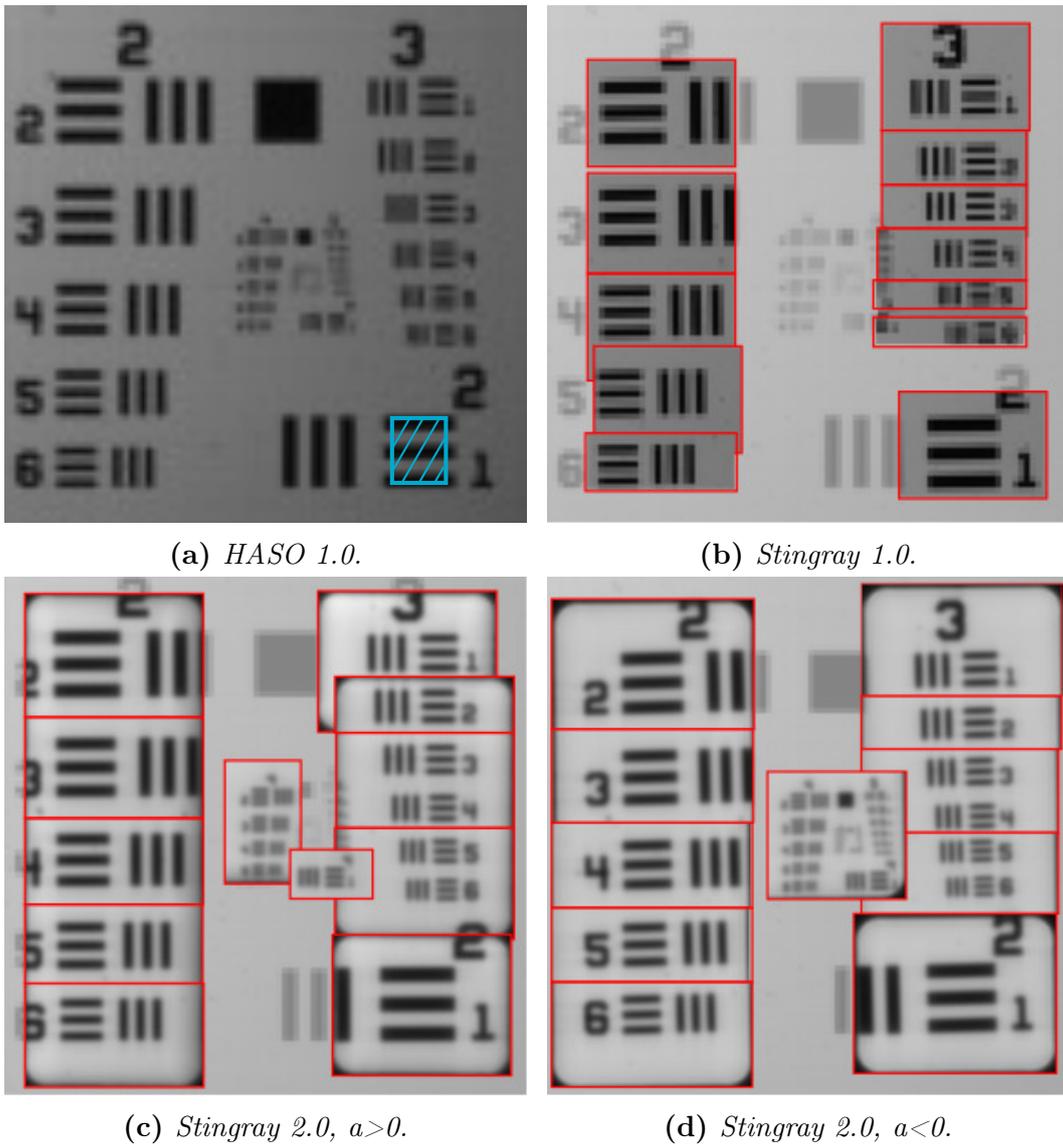


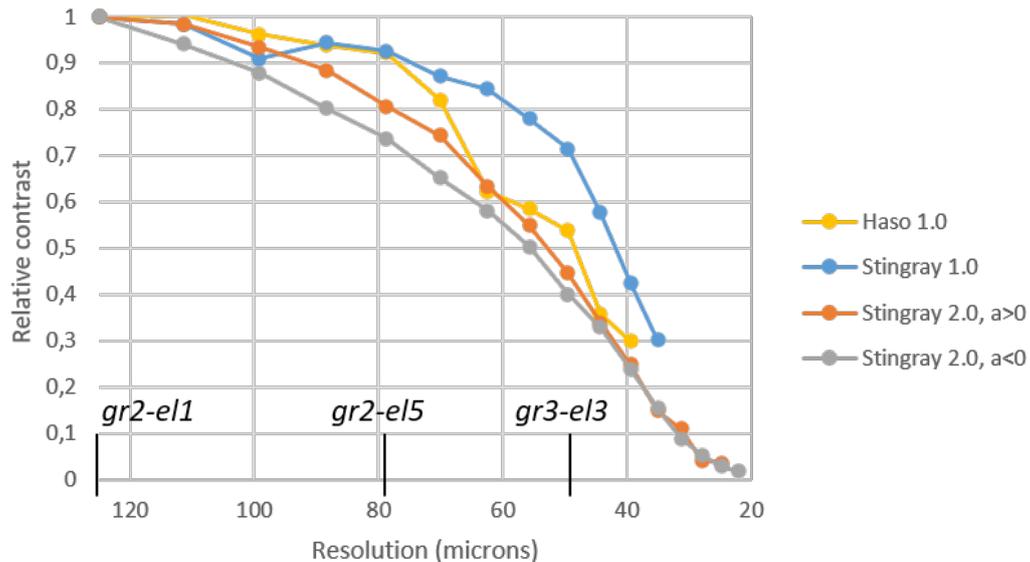
Figure 2.16 – Images used to estimate the lateral resolution in each of the four experimental configurations. The different images from the Stingray setup were assembled to virtually restore the whole test target, compensating for their small FOV. The hatched area in 2.16a represents the area where the contrast was measured for this bar width ($gr2 - el1$).

Figure 2.16 presents the assembled images in the three Stingray cases, to be compared to the image from the HASO 1.0.

Figure 2.17 presents the reconstructed images showing $gr3 - el3$ in each configuration. The reconstructed images of the same $gr3 - el3$ do not provide the same image quality depending on the configuration. In the two 1.0 images (Figs. 2.17a and 2.17b), the horizontal bars of $gr3 - el3$ are sharp and well contrasted. On the contrary, the other bars visible in the same two



(a) HASO 1.0. (b) Stingray 1.0. (c) Sting. 2.0, $a > 0$. (d) Sting. 2.0, $a < 0$.



(e) Contrast curves as a function of bar width, normalized by the value for $gr2 - el1$. **Figure 2.17** – Curves of the relative contrast along bar width, allowing to quantify the lateral resolution in each configuration. The theoretical lateral resolution varies between $47\mu\text{m}$ and $50\mu\text{m}$ depending on the configuration. The reconstructed images show $gr3 - el3$ (bar width = $49.5\mu\text{m}$) in the four configurations.

images (either the vertical bars of $gr3 - el3$, or the bars of $gr3 - el2$) look blurred. In a traditional configuration, the spatial positions of the test target are sampled by the micro-lenses, resulting in large pixels in the reconstructed image (see Chapter 1). The limit of resolution is determined by the highest frequency that could be acquired by the spacing of the micro-lenses, following the Nyquist–Shannon sampling theorem [Sha49]. The theoretical value of res_{lat} is close to the Nyquist frequency, mainly limited by the pitch d_2 of the μLA . Thus the correct sampling of the bars of $gr3 - el3$ by the μLA strongly depends on the spatial positioning of the bars compared to the micro-lenses. In order to compensate this effect, we carefully aligned the images of the bars on the μLA when performing the lateral scan of the test target in the Stingray 1.0. This effect is not visible in the images acquired in a focused configuration (Figs. 2.17c and 2.17d). The different spatio-angular sampling, combined with the "rendering with blending" algorithm, provide a higher number of pixels in

the reconstructed images.

The width of the bars in $gr3 - el3$ ($49.5\mu m$) corresponds to the theoretical value of res_{lat} , which is similar for the four cases (see Table 2.3). By definition, we should be able to see this group sharply. By eyes, $gr3 - el3$ seems resolved in the four cases (see Figs. 2.17a-2.17b-2.17c-2.17d). In order to quantify the lateral resolutions for each configuration, we study the decrease of contrast as a function of the width of the bars (see Fig. 2.17e).

For each group of bars, contrast was measured using equation 2.43, over a large area, as illustrated by the blue rectangle in Figure 2.16a. This allows to average the measured contrast along the length of the bars, and get more accurate results. The curves were normalized so that contrast on the largest bars ($gr2 - el1$) equals 1, in order to compensate for differences in illumination and refocusing algorithms. These curves can be compared to the modulation transfer function (MTF) curve classically measured on optical systems [BGY⁺13; Hec17].

The four curves follow the same pattern: a slow decrease as a function of the width of the bars (see Fig. 2.17e). The two curves for the Stingray 2.0 configurations are very close because they correspond to two similar configurations acquired on the same setup. They are also very regular, due to the higher number of pixels in the images reconstructed with algorithm 2.0. The two curves for the 1.0 cases are less regular because of the lower number of pixels and the sensitivity to the lateral positioning of the test target, as previously discussed.

The curves can be used to estimate the experimental values of res_{lat} for each configuration. A first measure of this limit is done by retrieving the resolution when the contrast reaches a value of 50%. With this criteria we obtained experimental values of: $48\mu m$ for HASO 1.0, $42\mu m$ for Stingray 1.0, $52\mu m$ for Stingray 2.0 $a>0$ and $56\mu m$ for $a<0$. These values are quite close to the expected values between $47\mu m$ and $50\mu m$ (see Table 2.3). The extracted resolutions depend on the normalization of the curves. A normalization by a different value would have given different results.

2.4.3 Depth resolution

To measure the resolutions along the depth axis, we follow the approach presented in [LNA⁺06]. The depth resolution res_{depth} is the depth range over which it is not possible to distinguish different depths. To determine res_{depth} , it is necessary to refocus the images using the same refocusing parameter, either $\alpha = 1$ in algorithm 1.0, or with the same P in algorithm 2.0. The depth resolution should be determined from the element $gr3 - el3$ to ensure precision in the measure, as the width of these bars correspond to the lateral resolution. res_{depth} corresponds to the depth range over which the refocused images of $gr3 - el3$ are not substantially affected by defocus blur. Inside this interval,

the images are the sharpest and look identical, so that it is barely possible to tell that they actually correspond to different depths of the scene.

2.4.3.1 Depth resolution in the HASO 1.0

Figure 2.18 presents the HASO case, in which sufficient data is available, thanks to the wider *FOV* in this setup. The light-field was acquired for many depth positions of the test target. Figure 2.18a-2.18b-2.18c shows the series of refocused images between $-1.2mm$ and $+1.2mm$ around depth z_0^{trad} , in steps of $0.4mm$. Three different elements are shown: $gr3 - el3$ ($49.5\mu m$) that corresponds to the theoretical lateral resolution, the largest visible element $gr2 - el1$ ($125\mu m$), and an intermediate element $gr2 - el5$ ($78.7\mu m$) (see Fig. 2.13). All the images were reconstructed with algorithm 1.0 using the same parameter $\alpha = 1$.

The images of $gr2 - el1$ and the ones of $gr2 - el5$ appear very similar along the depth range. In the displayed images, they are not affected by defocus blur due to their large sizes. On the contrary, the thinner bars of $gr3 - el3$ are progressively blurred when moving away from z_0^{trad} (red square). By eyes, the refocused images of $gr3 - el3$ seem identical from $z_0^{trad} - 0.4mm$ to $z_0^{trad} + 0.4mm$, resulting in a depth resolution of $0.8mm$, which is much finer than the $1.8mm$ theoretical limit.

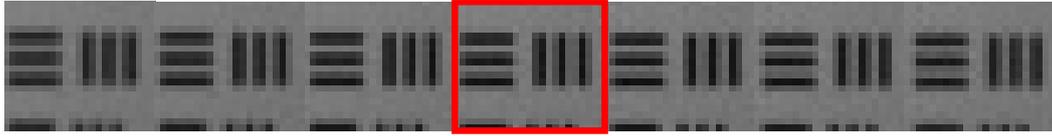
To be more quantitative, contrast is extracted along depth for these three elements. As for the lateral resolutions, the contrasts were extracted from an area covering each set of bars, allowing to average the contrast measured along the length of the bars and reduce the measurement error. The resulting contrast curves for these three elements are plotted in Figure 2.18d. The curves are centered around depth z_0^{trad} , and their values are normalized by the maximal contrast of the curve for $gr2 - el1$.

The curve for $gr3 - el3$ presents a surprising shape with rebounds on both side of the maximal contrast at z_0^{trad} . Contrast reaches 0 at $z_0^{trad} + 1.2mm$ and symmetrically a dip can be observed around $z_0^{trad} - 1.6mm$. This can be explained by a misalignment between the bars and the micro-lenses, as discussed in Section 2.4.2. When moving the test target closer or further from depth z_0^{trad} , the size of the *FOV* can change (see Section 2.2.1). This causes a progressive misalignment of the bars during the defocusing process. The test target was laterally positioned so that the contrast is maximal at depth z_0^{trad} , and this misalignment is responsible for a progressive decrease in contrast until the value 0 when the image of the bars is shifted by exactly $d_2/2$ compared to the micro-lenses frequency.

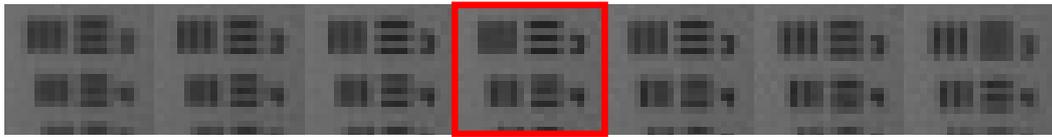
The envelope of the contrast curve is drawn in dotted black. The envelope gives an idea of the contrast curve that could be measured if the bars were aligned with the μLA at each depth. The curves for the bars $gr2 - el1$ and $gr2 - el5$ are less affected by this misalignment along depth. A slight dip can



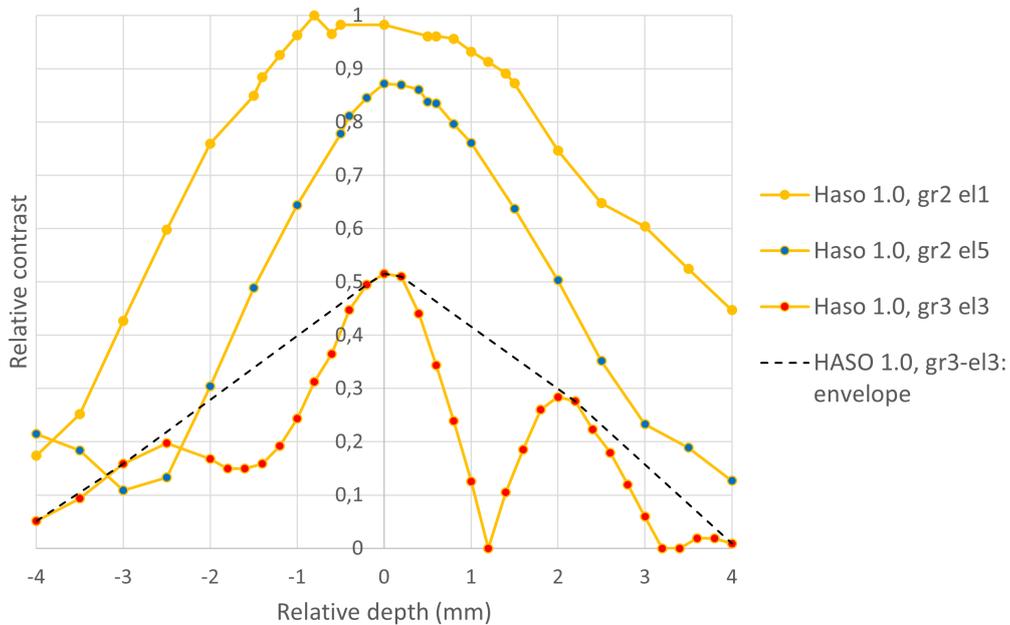
(a) HASO 1.0: $gr2-el1$, from $z = z_0^{trad} - 1.2mm$ to $z_0^{trad} + 1.2mm$, steps of $0.4mm$.



(b) HASO 1.0: $gr2-el5$, from $z = z_0^{trad} - 1.2mm$ to $z_0^{trad} + 1.2mm$, steps of $0.4mm$.



(c) HASO 1.0: $gr3-el3$, from $z = z_0^{trad} - 1.2mm$ to $z_0^{trad} + 1.2mm$, steps of $0.4mm$.



(d) Contrast curves normalized by the maximal contrast value for $gr2-el1$.

Figure 2.18 – Depth resolution for the HASO 1.0: series of refocused images at the same z_0 ($\alpha = 1$) and contrast curves extracted on elements $gr2-el1$, $gr2-el5$ and $gr3-el3$.

still be noticed around $z_0^{trad} - 3mm$ for $gr2-el5$.

These three curves together allow to better show the effect of defocus depending on the size of the bars. Their amplitudes (highest values) can be directly linked to the lateral resolution studied in previous section. Their widths are related to depth resolution. The finer the bar width on the test target,

the faster the decrease of the contrast curve. This quantifies what was already visible in the images, that the smaller elements are more rapidly affected by defocus blur.

Using these contrast curves, we would like to quantify depth resolution in the HASO 1.0 configuration. As res_{depth} is defined as the depth range over which the reconstructed images look identical (it is not possible to separate their depths), we expect to find a criterion to measure it on the contrast curves. According to Table 2.3, the theoretical res_{depth} equals $1.85mm$. This depth interval can be retrieved from the envelope of the contrast curve of $gr3 - el3$, which is the limit in lateral resolution. Considering the envelope at 80% of the maximal value, the width goes from $-0.8mm$ to $1.1mm$, corresponding to a range of $1.9mm$. By comparison, the criterion taken at 50% instead corresponds to the envelope going from $-2mm$ to $2.3mm$, hence a range of $4.3mm$, which is more than twice the expected value of res_{depth} . The same measures taken directly on the contrast curve instead of the envelope give lower values. Using the criterion of 80%, the measured depth range is $1mm$ on the contrast curve (between $-0.5mm$ and $0.5mm$). We retrieve a similar value to the previous resolution of $0.8mm$ estimated by eyes.

The same measures are done on the other two contrast curves. Concerning $gr2 - el5$, the widths at 80% and 50% respectively correspond to $2mm$ and $3.8mm$. They roughly fit with the ones measured for $gr3 - el3$. The bars of $gr2 - el5$ have a fine width and are thus rapidly affected by defocus blur. On the contrary, the contrast curve of $gr2 - el1$ presents a large plateau centered on z_0 , where defocus of the test target is not strong enough to impact the larger bars. The measured widths at 80% and 50% are $3.6mm$ and $6.5mm$, which are completely different from the previous ones. When it comes to resolution along depth, the size of the bars considered for the measures highly affects the experimental results.

We have found the criterion of 80% as an experimental tool allowing to retrieve res_{depth} from the $gr3 - el3$ contrast curve. We will see in the next sections if this criterion can be applied to the other configurations.

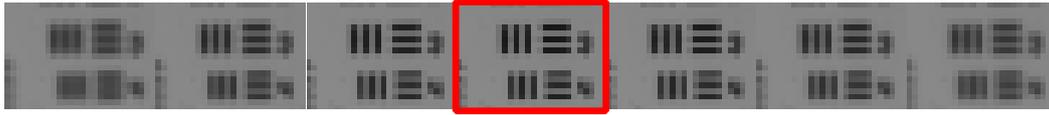
2.4.3.2 Depth resolution in the Stingray 1.0

The same study is done for the Stingray 1.0 configuration. Because of the small FOV , two different series of images were acquired. The first one shows the evolution of the $gr2 - el1$ along depth (see Fig. 2.19a), while the second one presents the $gr3 - el3$ as the limit of lateral resolution (see Fig. 2.19b). The two series of images were refocused at $\alpha = 1$ to study depth resolution. The images at z_0^{trad} are highlighted by the red squares. Contrast was extracted on these two series and the corresponding contrast curves are plotted in Figure 2.19c. The two curves were normalized by the maximal contrast of the $gr2 - el1$ curve, similarly to what was done in the HASO 1.0 case. They are centered

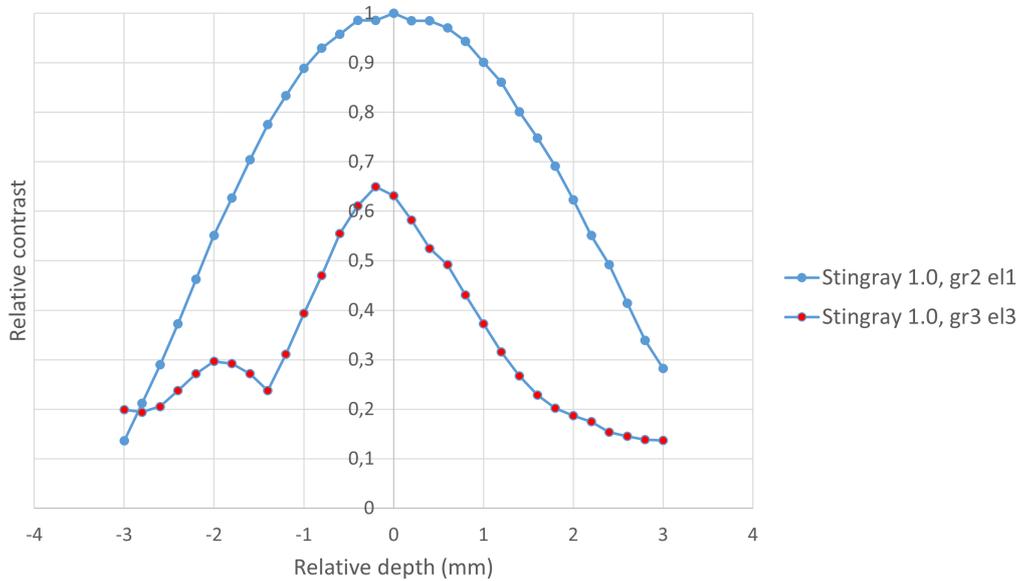
around $z_0^{trad} = 233mm$ to enable comparison with the other cases.



(a) *Stingray 1.0: gr2 – el1, from $z = 230.6mm$ to $z = 235.4mm$, steps of $0.8mm$.*



(b) *Stingray 1.0: gr3 – el3, from $z = 231.8mm$ to $z = 234.2mm$, steps of $0.4mm$.*



(c) *Contrast curves normalized by the maximal contrast value for gr2 – el1.*

Figure 2.19 – *Depth resolution for the Stingray 1.0 configuration: series of images refocused at the same z_0^{trad} ($\alpha = 1$) for gr2–el1 (a) and gr3–el3 (b), and the contrast curves extracted on these elements along depth (c).*

The curve for *gr2 – el1* is smooth and presents a small plateau around z_0^{trad} . This plateau can be related to the series of refocused images, where several images around z_0^{trad} look sharp and identical. By eyes, we evaluate this depth range from $z_0^{trad} - 0.8mm$ to $z_0^{trad} + 0.8mm$, which gives an interval of size $1.6mm$. In the *gr3 – el3*, only the two adjacent images seem as sharp as the one at depth z_0^{trad} . With a step of $0.4mm$, the size of this interval is $0.8mm$.

These qualitative results are now quantified using the contrast curves. Concerning the *gr3 – el3*, the width of the curves measured at 80% of its maximum goes from $z_0^{trad} - 0.7mm$ to $z_0^{trad} + 0.4mm$, for a total extent of $1.1mm$. We here recover the theoretical value of $res_{depth} = 1.10mm$. The same process done at 50% of the maximum gives a depth range of $2.4mm$. For comparison,

the same process applied on the $gr2 - el1$ curve resulted in values of $2.7mm$ for the 80% criterion and $3.5mm$ for the 50% one. As in the previous case, we notice that the curve for $gr3 - el3$ presents a dip at $z_0 - 1.4mm$, corresponding to misalignment artifacts. The lateral placement during the acquisition process of $gr3 - el3$ explains that the artifacts are reduced compared to the HASO case (see Fig. 2.18).

These experimental values are coherent with the ones measured in the HASO 1.0. In both cases, the value measured at 80% on $gr3 - el3$ allows to recover the theoretical res_{depth} , whereas the same criterion on $gr2 - el1$ gives approximately the double. We have found an experimental way of measuring res_{depth} based on contrast curves, and we obtained coherent results on two completely different setups.

2.4.3.3 Depth resolution in the Stingray 2.0

Unfortunately, for the two Stingray 2.0 configurations, only series of images of the $gr2 - el1$ have been acquired. Therefore, the criterion of 80% applied on $gr2 - el1$ will be used to estimate twice the experimental depth resolution.

Figure 2.20 presents the four configurations together, to compare their re-focused images as well as their contrast curves. The images of the two Stingray 2.0 datasets were reconstructed using $P = 5$. We remind here the theoretical res_{depth} for each of them: $1.85mm$ for HASO 1.0, $1.10mm$ for Stingray 1.0 and Stingray 2.0, $a > 0$, and only $0.74mm$ for Stingray 2.0, $a < 0$ (see Table 2.3). The contrast curves were normalized. They were shifted so that z_0^{trad} or z_0^{foc} corresponds to the relative depth 0 for each configuration on the graph.

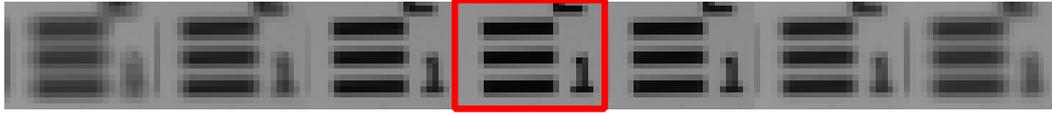
The series of images allow the reader to visually appreciate the progressive defocus blur when moving away from depth z_0 , depending on the configuration. By eyes, the bars look in focus over adjacent images on both sides of z_0 . Again the depth range runs over $\sim 1.6mm$, corresponding to roughly double the value of res_{depth} . This is consistent with the results obtained using our criterion of 80%.

In Figure 2.20e, the three contrast curves for the Stingray configurations overlap very closely. The top of the HASO curve merges with the Stingray configurations, but the HASO curve is a bit larger at the bottom. This larger size can be explained by the larger expected value of res_{depth} for this configuration. According to their theoretical values, the Stingray 2.0, $a < 0$ curve should be the thinner one, with the two others being very similar. This is not the case, this curve seems slightly wider than the two other Stingray curves.

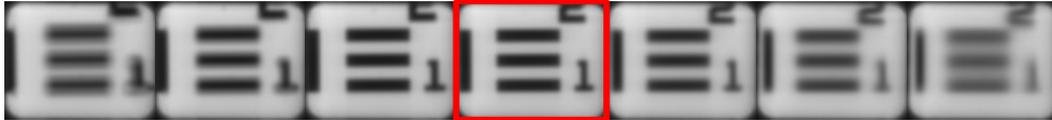
The width of the curves was measured using the criterion of 80% of the maximum. The results are the following: $3.6mm$ for HASO, $2.7mm$ for Stingray 1.0, $3.2mm$ for Stingray 2.0, $a > 0$, and $3.3mm$ for the $a < 0$ case. As expected, we stay in the same order of magnitude of twice the theoretical res_{depth} . This allows to approximate the experimental value of res_{depth} , overcoming the issue



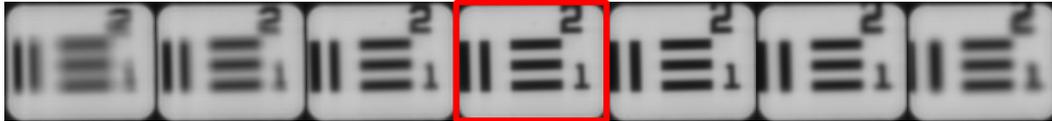
(a) *HASO 1.0: gr2 – el1, from $z = z_0^{trad} - 2.4mm$ to $z_0^{trad} + 2.4mm$, steps of 0.8mm.*



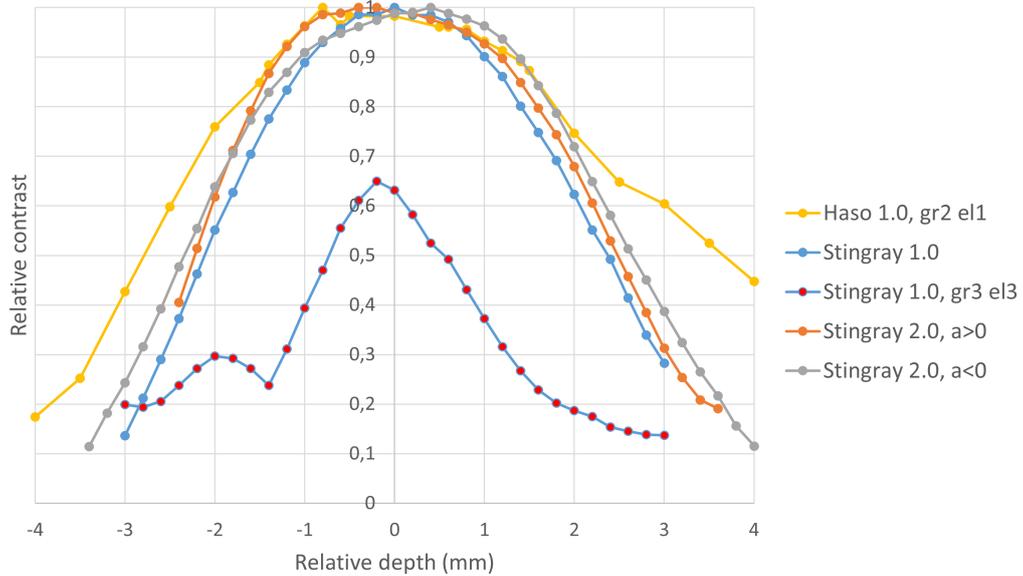
(b) *Stingray 1.0: gr2 – el1, from $z = 230.6mm$ to $z = 235.4mm$, steps of 0.8mm.*



(c) *Stingray 2.0, $a > 0$: gr2 – el1, from $z = 230.6mm$ to $235.4mm$, steps of 0.8mm.*



(d) *Stingray 2.0, $a < 0$: gr2 – el1, from $z = 228.6mm$ to $233.4mm$, steps of 0.8mm.*



(e) *Contrast curves normalized by the maximal contrast value for gr2 – el1.*

Figure 2.20 – *Depth resolutions for the four configurations, with the series of refocused images of gr2 – el1, and the corresponding contrast curves. This comparison allows to extrapolate the results obtained in the HASO 1.0 and Stingray 1.0 cases to the Stingray 2.0 $a > 0$ and $a < 0$.*

of the missing gr3 – el3 contrast curves.

To conclude, the criterion of 80% on the $gr3-el3$ curve allowed to precisely estimate the theoretical values in the HASO 1.0 (Fig. 2.18) and the Stingray 1.0 configurations. In the Stingray 2.0 $a>0$ and $a<0$ cases, these curves were missing. With the same criterion of 80%, the contrast curve for $gr2-el1$ enabled to approach twice the depth resolution. We have found a consistent criterion to evaluate depth resolution from the experimental images.

2.4.4 Plenoptic depth of field

The plenoptic depth of field is the largest depth range over which it is possible to accurately refocus the test target from a single acquisition (see Section 2.2.5). Following the approach of [LNA⁺06], each image needs to be reconstructed using the most appropriate refocusing parameter (P or α), so that the refocused image is as sharp as possible.

2.4.4.1 Methodology for drawing the contrast curve

In the Stingray 1.0 configuration, images of $gr2-el1$ were acquired at different depths. Each raw image was reconstructed using different values of P , then the contrast was measured. In Figure 2.21 are drawn the contrast curves corresponding to different values of P , showing how this parameter impacts the refocusing quality along depth. For example at depth $z_0 = 5mm$, the reconstruction with $P = -4$ (orange) has a higher contrast than the reconstruction with $P = -3$ (blue) or $P = -6$ (yellow). Taking the maximal contrast for each z allows to choose the best refocusing parameter for this depth. This results in the dotted black curve, that will be used to study the plenoptic depth of field in this configuration. This curve allows to select the images with the adapted refocusing parameter that will be put together in the series of images in Figure 2.23b.

Actually these different P curves can be related to the notion of depth resolution studied in previous section. It corresponds to reconstruction of several depths with the same refocusing parameter, as was done in Figures 2.18-2.19-2.20. Each P curve corresponds to the study of depth resolution at another depth than z_0 . The series of curves shows how the depth resolution varies with depth, as was mentioned in Section 2.2. When depth is further from z_0 , the P curve has a lower height and its shape changes. The lower contrast is the sign of potential higher difficulties in detecting small structures on the image. This is related to the decrease in resolutions when moving away from z_0^{trad} in the 1.0 configuration. Figure 2.21 gives also a visual comparison of depth resolution versus plenoptic depth of field.

The same process of taking the maximum of the P curves is performed for

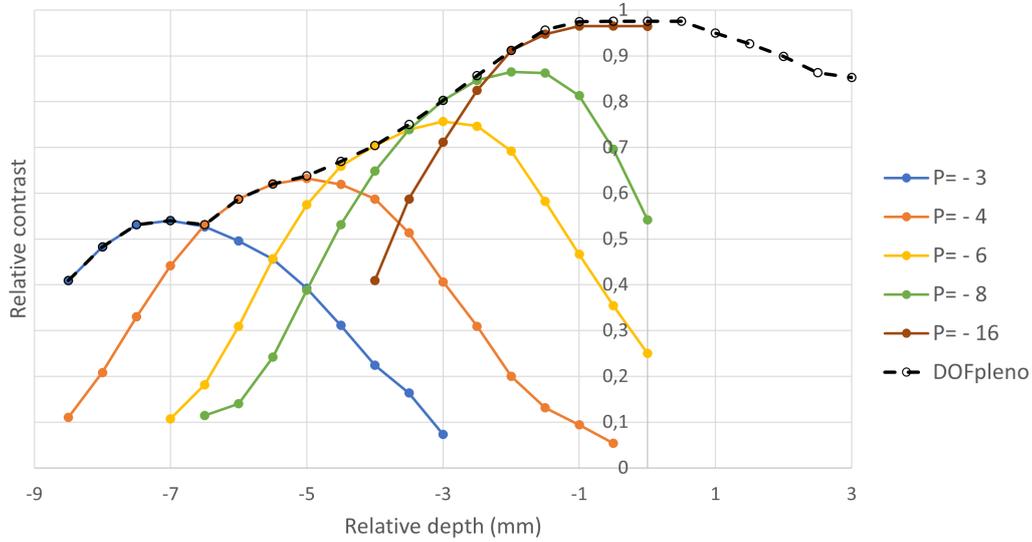


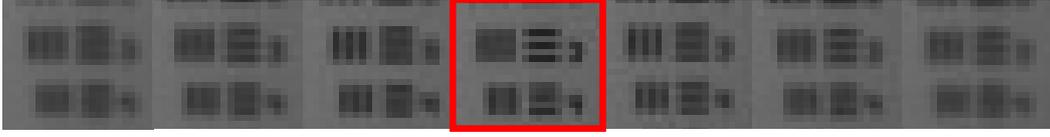
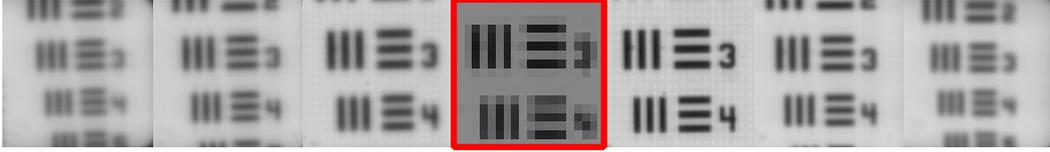
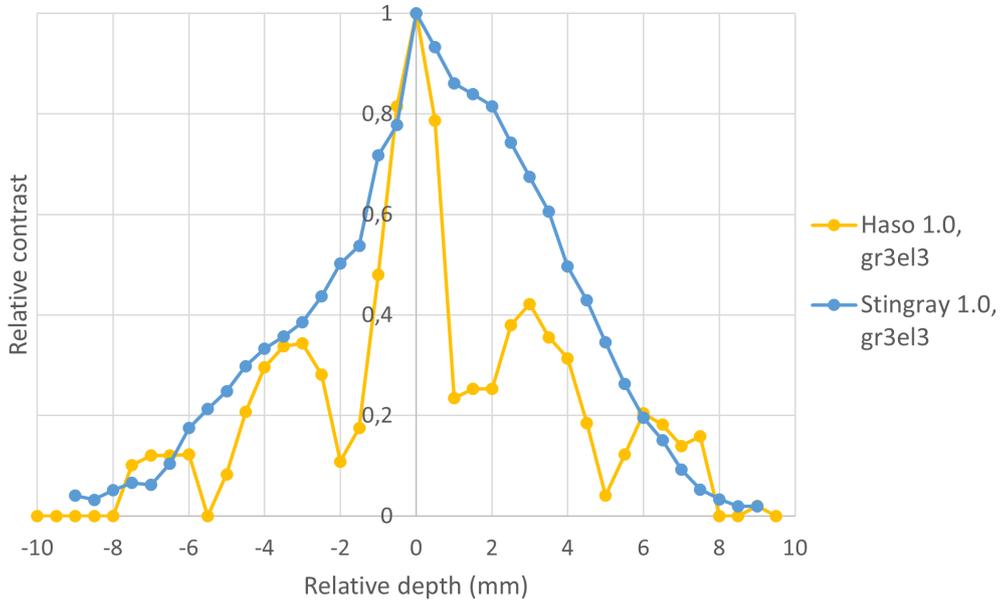
Figure 2.21 – Methodology for building the contrast curve to study the plenoptic depth of field in the Stingray 1.0 configuration. Each image is reconstructed with various refocusing parameters then the maximal contrast is selected. The curves for the different P values can be related to the depth resolution at a depth different from z_0 . The decrease in height and width is the sign of lower lateral and depth resolutions when moving away from z_0 .

each different configuration, resulting in the curves and series of images used in the following sections. We obtain a series of refocused images with a depth-dependant refocusing parameter, over which we expect to recover DOF_{pleno} . As for depth resolution, a criterion has been defined to correctly measure the depth range over which the images are considered well refocused.

2.4.4.2 Experimental validation of plenoptic depth of field

The plenoptic depth of field is evaluated on the series of images of *gr3-el3*, as their bar width represents the lateral resolution limit. Only images for HASO 1.0 and Stingray 1.0 are available for this bar width. Figure 2.22 presents the study done on these two cases.

The images for HASO 1.0 were all refocused using algorithm 1.0, while the images for Stingray 1.0 were refocused with the algorithm 2.0 outside of depth z_0 , because of its ability to produce images with a higher number of pixels (see Section 2.4). The parameter of refocusing (either α or P) was adjusted for each image of the two series, to get the best focused images. We observe that the refocused images look sharp on a few images around depth z_0 , and then suffer from defocus blur on both sides of the series (see Fig. 2.22a and 2.22b). We can also notice a variation in the size of the bars along depth in the images. This effect comes from a difference in magnification when the test target is moved

(a) HASO 1.0: $gr3 - el3$, from $z = z_0^{trad} - 6mm$ to $z_0^{trad} + 6mm$, steps of $2mm$.(b) Stingray 1.0: $gr3 - el3$, from $z = 227mm$ to $z = 239mm$, steps of $2mm$.(c) Normalized contrast curves for $gr3 - el3$.**Figure 2.22** – Plenoptic depth of field based on the $gr3 - el3$ in the HASO and Stingray 1.0 configurations.

on a different depth plane. This will be studied in more details in Section 2.5.

We would like to quantify the plenoptic depth of field on these series of refocused images. By definition, if the plane of the reconstructed image belongs to the plenoptic depth of field, the reconstructed image should be well refocused, in the sense that the bars should be visible with sufficient contrast. By eyes, the bars of $gr3 - el3$ are visible between $z_0^{trad} - 4mm$ and $z_0^{trad} + 4mm$ for the HASO 1.0 and between $z_0^{trad} - 6mm$ and $z_0^{trad} + 6mm$ in the Stingray 1.0 case. This results in values of $DOF_{pleno} = 8mm$ and $DOF_{pleno} = 12mm$, compared to the theoretical value $\sim 10mm$ according to Table 2.3.

Figure 2.22c presents the contrast curves measured on these series of images. Each curve was independently normalized by its maximal value located

at z_0 . As for the lateral and depth resolutions, some regular dips can be observed on the HASO curve, due to the progressive misalignment between the bars and the μ LA when defocusing the test target (see Section 2.4.2). The curve for the Stingray case does not present such kind of dips, due to the use of algorithm 2.0 for the refocusing. What is interesting is that the Stingray curve seems to be the envelope of the HASO curve. This means that despite their different setups and algorithms, the ability of refocusing images shows the same evolution when moving away from z_0 . This is consistent with their similar theoretical DOF_{pleno} around $\sim 10mm$.

As for the previous resolutions, we found a criterion to read the plenoptic depth of field on the contrast curves. In the Stingray case, the curve reaches a width of $10mm$ for a contrast of 30% of the maximal value at depth z_0 . This value of 30% is much lower than the 80% for the depth resolution. This reflects the different natures of res_{depth} and DOF_{pleno} . For res_{depth} the criterion was chosen in order to quantify the depth range over which the images are the sharpest, whereas for DOF_{pleno} , the goal is to determine until where it is still possible to refocus the bars, even with a lower contrast. This value of 30% defines a threshold that corresponds to the theoretical DOF_{pleno} and also to the visual assessment based on the refocused images.

2.4.4.3 Study of DOF_{pleno} with a different bar width

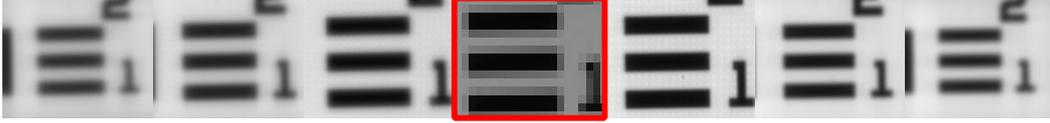
The plenoptic depth of field is now studied on the $gr2 - el1$. As for the depth resolution, the difference in bar width is expected to impact the lateral slopes of the contrast curves, and affects the depth range over which the bars are visible. Studying a different bar size is the opportunity to examine the effects of the size on the experimental DOF_{pleno} .

Figures 2.23a-2.23b-2.23c show the series of refocused images in three configurations: the HASO 1.0, the Stingray 1.0 and the Stingray 2.0, $a > 0$. For each image, the refocusing parameter was adapted to the depth in order to generate a sharp reconstruction. Around z_0 (in red), the refocused images are sharp, then blur progressively appears when moving away on both sides of the series. However, the bars are visible along the whole depth range, because of the largest width of $gr2 - el1$ compared to $gr3 - el3$. Whereas the displayed depths largely exceed the theoretical value of $DOF_{pleno} \sim 10mm$, the bars are still visible at both ends of the interval, making it impossible to detect a limit for plenoptic depth of field based on $gr2 - el1$.

As for the series of $gr3 - el3$ (see Fig. 2.22), we observe variations in the magnification of the bars according to the depth of the test target. In the HASO case, these variations are mild. In the Stingray cases, the magnification changes more quickly. This is due to a higher magnification in this setup, related to the lower field of view. Indeed, the axial magnification can be approximation by M^2 [Gro05], and therefore a higher magnification results



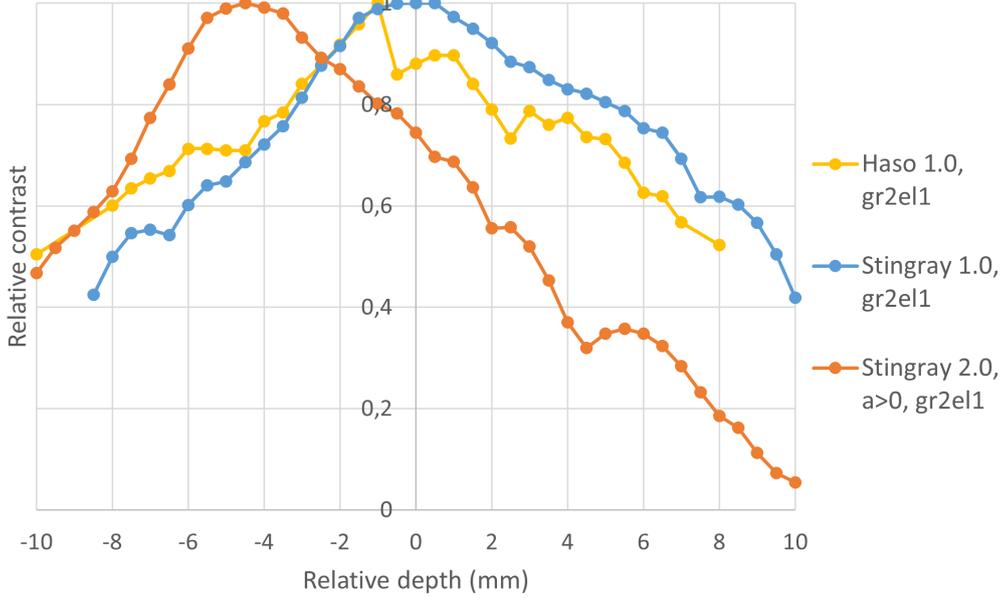
(a) HASO 1.0: $gr2 - el1$, from $z = z_0^{trad} - 6mm$ to $z_0^{trad} + 6mm$, steps of 2mm.



(b) Stingray 1.0: $gr2 - el1$, from $z = 227mm$ to $z = 239mm$, steps of 2mm.



(c) Stingray 2.0, $a>0$: $gr2 - el1$, from $z = 227mm$ to $z = 239mm$, steps of 2mm.



(d) Normalized contrast curves for $gr2 - el1$. The relative depth $z = 0$ corresponds to z_0^{trad} for the HASO 1.0 and Stingray 1.0 and to z_0^{foc} for the Stingray 2.0, $a>0$.

Figure 2.23 – Plenoptic depth of field based on the $gr2 - el1$ in the HASO, the Stingray 1.0 and 2.0, $a>0$ configurations.

in higher variations when changing depth z_0 . In the Stingray 1.0, the magnification is maximal at depth z_0 and decreases similarly on both sides. On the contrary, the series for Stingray 2.0, $a>0$ shows a maximal magnification around $z_0 - 4mm$ instead of z_0 . This depth corresponds to z_0^{trad} , which is the depth plane optically conjugated with the μLA , as in the 1.0 configuration.

This higher magnification and the lack of symmetry around z_0 will have impact on the measured contrast and estimated plenoptic depth of field. This will be studied in details in Section 2.5.

This difference in magnification and image quality can be directly related to the measured contrast curves. For each series, contrast was extracted on the images and plotted against depth (see Fig. 2.23). The measured contrast was normalized by the maximal value of the curve, and laterally shifted so that z_0 corresponds to a relative depth of 0 for each configuration. The curves of the HASO and Stingray 1.0 cases are symmetric around their reference depth z_0^{trad} . However, this is not the case of Stingray 2.0, $a > 0$, whose curve is completely decentered a few mm on the left. Its maximal value is located at $z_0 - 4.5mm$, which is the same depth where the magnification is maximal in the refocused images. This means that in the 2.0 case, the maximal contrast is located at z_0^{trad} instead of z_0^{foc} . This is a problem as there could be a confusion between z_0^{trad} and z_0^{foc} based on this curve. It could also have an impact on the measure of DOF_{pleno} , as the contrast decreases from its maximal value located at z_0^{trad} instead of z_0^{foc} . The highest values are thus centered around z_0^{trad} . The formula of DOF_{pleno} does not seem valid for the focused configuration. Simulations are necessary to understand the reasons, and will be presented in section 2.5.

Except the shift in the Stingray 2.0 case, the three curves present a regular decrease on both sides, with similar slopes values. This represents the similar expected value of $DOF_{pleno} \sim 10mm$ in the three cases. Contrary to $gr3 - el3$, it is not possible to measure DOF_{pleno} on the contrast curve of $gr2 - el1$, due to its larger size. The bars can be visually separated along the whole series, whereas the depth range of the displayed images largely exceeds the theoretical value of DOF_{pleno} . The criterion of 30% on the contrast curve of $gr3 - el3$ confirms that the bars can be considered as visible along the whole depth range.

As a conclusion, it appears that the theoretical DOF_{pleno} is a good approximation of the reconstruction distances allowed by a plenoptic camera. In both 1.0 configurations, DOF_{pleno} gives reliable limits for the depth range that could be correctly reconstructed. This way the plenoptic depth of field, together with the field of view, delimits a volume where the lateral and depth resolutions can be considered as homogeneous.

However, it is not as straightforward in a 2.0 configuration. The maximal contrast value is shifted from z_0^{foc} to z_0^{trad} on the contrast curve. This indicates that some additional effects, like magnification, should be considered in the plenoptic depth of field theory. Consequently, even if the width of the 2.0 curve seems identical to 1.0 ones, in agreement with the expected values, the proposed formula should be used with special care.

2.5 Simulation of the impact of magnification and diffraction on depth of field

In order to better understand the impact of magnification on plenoptic depth of field, a deeper study as been done on the Stingray 2.0, $a > 0$ case. A simulation framework was implemented to study the combined effects of magnification, diffraction and defocus. The objective of this simulation is to examine how the three phenomena impact the contrast variations with depth.

Up to this point, the different resolutions were studied on reconstructed images. However, the reconstruction algorithm can impact image quality and contrast. In the simulation, the plenoptic depth of field is studied using contrast estimation on a single sub-image, instead of a reconstructed image. The condition of 3D reconstruction at a chosen depth is that the plane at this depth has been imaged by sharp sub-images: a depth belongs to DOF_{pleno} if and only if it is included in the depth of field of a single sub-image [MCJ18]. As explained in Section 2.2.5, the plenoptic depth of field is the depth of field of a single sub-image. Therefore, DOF_{pleno} can be directly studied on the raw plenoptic image, without the need for reconstruction.

2.5.1 Simulations

The simulation reproduces the configuration of the Stingray 2.0, $a > 0$ (see tables 2.2 and 2.3). It models all the transformation that occur between the test target in the object space until the formation of a sub-image on the detector.

A raw plenoptic sub-image is simulated on the detector's plane, positioned at a fixed distance b^{foc} from the μ LA (Fig. 2.24). The simulated object is the three bars of *gr2 - el1* of the test target (bar width = $125\mu m$, see Fig. 2.13), virtually placed at a variable distance z_0 from the main lens. The distance between the main lens and the μ LA is kept fixed to $c^{foc} = z_1^{foc} + a^{foc} = 1597mm$. We remind that z_0^{foc} is the privileged depth for this focused configuration, whereas z_0^{trad} is the depth optically conjugated with the μ LA in this context, ($a = 0$). We consider three phenomena affecting contrast of the raw plenoptic sub-image: diffraction, defocus and magnification, introduced by Eq. 2.37.

Diffraction. To simulate diffraction, we performed the convolution of the object with a cardinal sine function, representing the diffraction pattern of our square microlenses [Hec17]:

$$I(r) = I_0 \operatorname{sinc}^2\left(\frac{\pi \cdot r \cdot d_2}{\lambda \cdot b^{foc}}\right) \quad (2.44)$$

with r the position on the detector, $\operatorname{sinc}(x) = \sin x/x$ and I_0 a normalization constant such that $\int I(r) dr = 1$. Diffraction from the main lens is ignored,

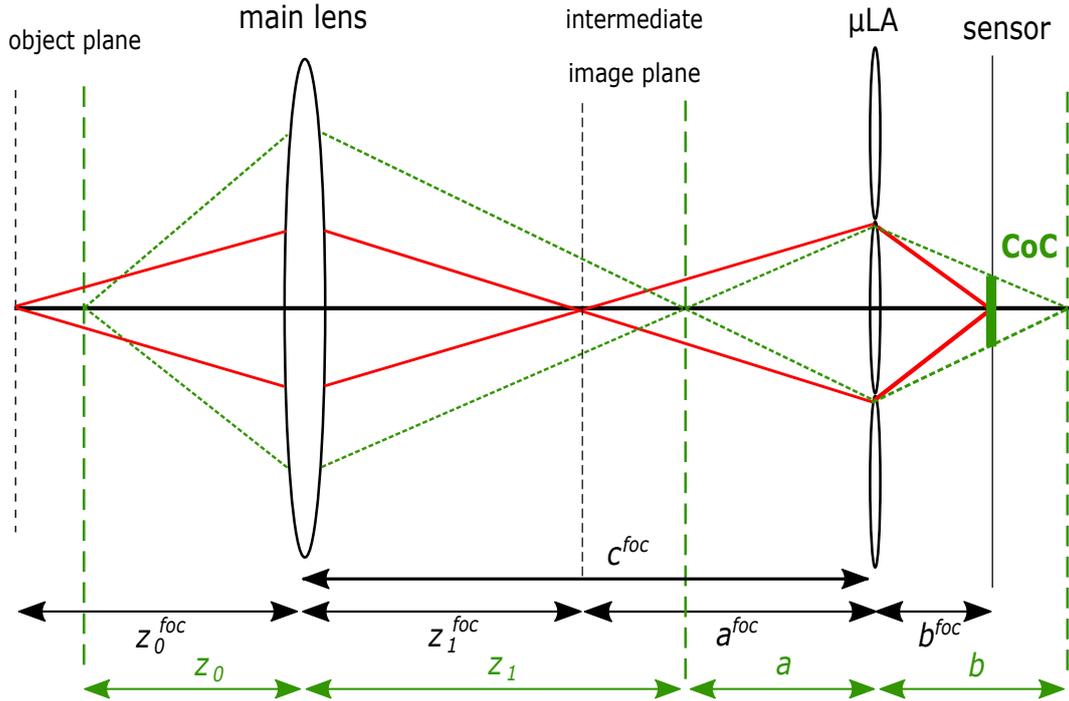


Figure 2.24 – A focused plenoptic system. Red rays come from the depth plane z_0^{foc} whereas green rays correspond to another depth to illustrate the Circle of Confusion (CoC) due to defocus.

because it is reduced by the magnification of the μLA compared to the sub-image and pixel sizes.

Defocus. The Circle of Confusion corresponds to the region over which a point is blurred in the case of defocus (Fig. 2.24) [ZLE⁺18]. Its diameter CoC is defined as:

$$CoC = d_2 \left| 1 - \frac{b^{foc}}{b} \right| \quad (2.45)$$

The object is in focus and the image is sharp when the image plane is close to the detector (i.e., $|1 - b^{foc}/b| \ll 1$).

Convolutions of the object with $I(r)$ from Eq. (2.44) and a circle of size CoC from Eq. (2.45) simulate the effects of diffraction and defocus. They strongly depend on the depth z_0 .

Magnification. The third phenomenon is the total magnification M_{tot} of the system, combining the effects of the main lens and the μLA . It is a hyperbolic

2.5. SIMULATION OF THE IMPACT OF MAGNIFICATION AND DIFFRACTION ON DEPT

function of depth z_0 :

$$M_{\text{tot}} = \frac{z_1}{z_0} \frac{b^{\text{foc}}}{a} = \frac{f_1 b^{\text{foc}}}{c^{\text{foc}} - f_1} \frac{1}{z_0 - z_0^{\text{trad}}} \quad (2.46)$$

$$\text{with } z_0^{\text{trad}} = \frac{f_1 c^{\text{foc}}}{c^{\text{foc}} - f_1}$$

Impact of diffraction and defocus. Fig. 2.25 shows the effects of diffraction and defocus on a plenoptic raw sub-image for an object at $z_0 = 240\text{mm}$, whereas the optical system is optimized for $z_0^{\text{foc}} = 233\text{mm}$. The black dotted curve represents the shape of the original signal without the effects of diffraction or defocus. At $z_0 = 240\text{mm}$, the defocus (green curve) only affects the shape and width of the bars, because the size of CoC is too small compared to the width of the bars on the screen to affect the amplitude and contrast of the signal. On the contrary, considering the diffraction alone (red curve) results in a drop of contrast from 1 to 0.59. With the combination of both diffraction and defocus (blue curve), the contrast is even lower with a value of 0.52.

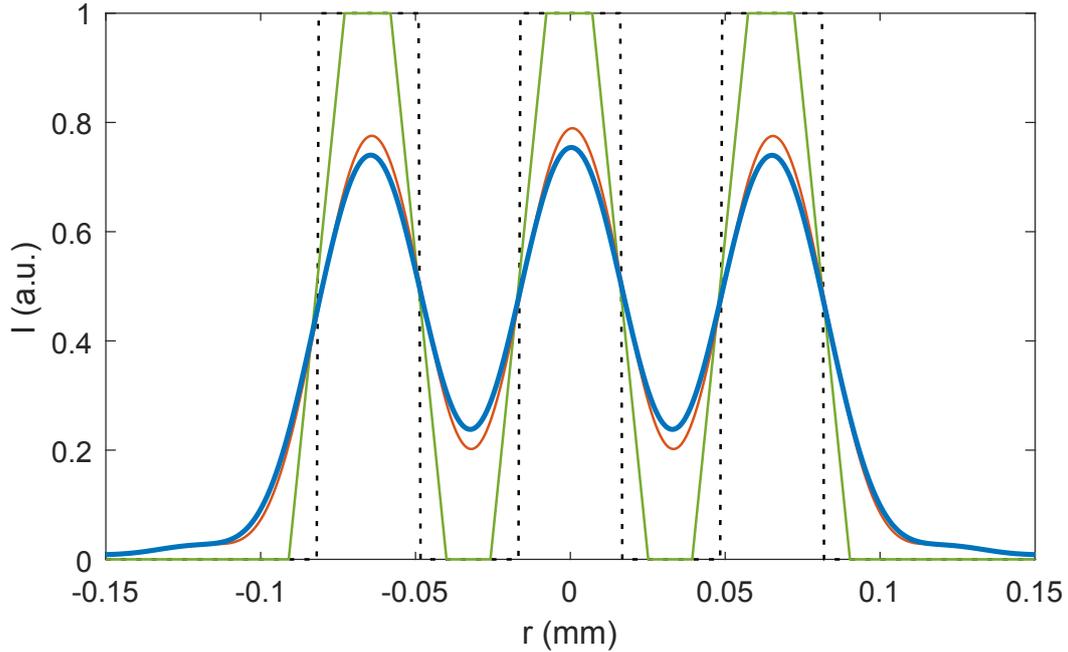


Figure 2.25 – Simulated 1D raw image of the test object versus position r on the detector with the effect of defocus only (green; contrast = 1), with diffraction only (red; contrast = 0.59) and with both effects combined (blue; contrast = 0.52), calculated at a depth $z_0 = 240\text{mm}$, compared to the image with no effect (dotted black; contrast = 1).

Impact of diffraction, defocus and magnification. The cumulative effects of the three factors are illustrated in Fig. 2.26 at four different depths.

Magnification governs the spatial frequency of the test object. On the detector, we measure a spatial frequency of 3.20cy/mm, 5.89cy/mm, 8.59cy/mm and 15.34cy/mm for $z_0 = 231\text{mm}$, 233mm, 235mm and 240mm respectively. In comparison, the physical size of the test target of bar width of $125\mu\text{m}$ corresponds to a frequency of 4.00cy/mm.

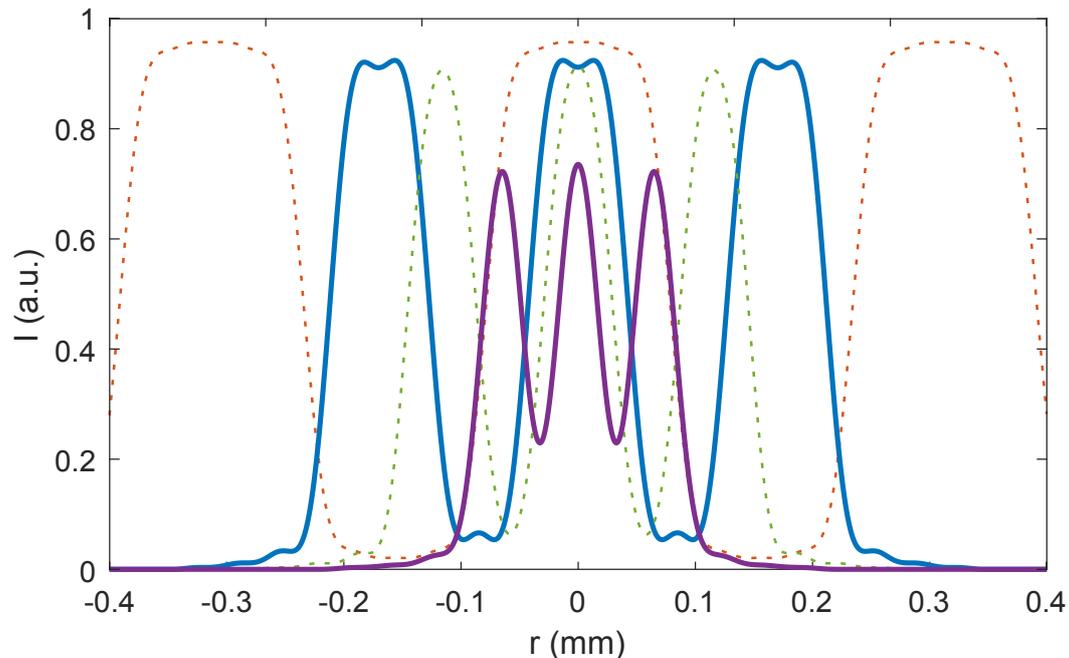


Figure 2.26 – Simulated 1D raw image with the three effects versus position r on the detector for different depths z_0 : $z_0^{\text{foc}} = 233\text{mm}$ (blue), 231mm (dotted orange), 235mm (dotted green), and 240mm (purple, same depth as in Fig. 2.25).

The difference in amplitude is mainly due to diffraction and defocus. Magnification plays a role in the amplitude variations, but effects are not symmetric: the contrast is higher close to z_0^{rad} (0.96 at 231mm), and decreases as the depth increases (0.89 at 233mm, 0.86 at 235mm and 0.5 at 240mm). Amplitude at z_0^{foc} is 11% lower in respect to the test target at object plane.

Simulated contrast curves as a function of depth. Parameters given in tables 2.2 and 2.3 were used for the simulations, except that the size of the sub-image has been artificially enlarged to fully include the magnified image. Then the contrast in the simulated raw sub-images was extracted at different depths to study the consequence of the three phenomena (diffraction, defocus and magnification) on the plenoptic image. Fig. 2.27 shows their impacts on contrast curves as a function of depth position z_0 .

In Fig. 2.27, defocus (green curve) does not significantly affect the contrast as depth varies: a plateau of contrast=1 nearly centered around the z_0^{foc} position is observed, as already seen in Fig. 2.25. The curve with diffraction and

2.5. SIMULATION OF THE IMPACT OF MAGNIFICATION AND DIFFRACTION ON DEPT

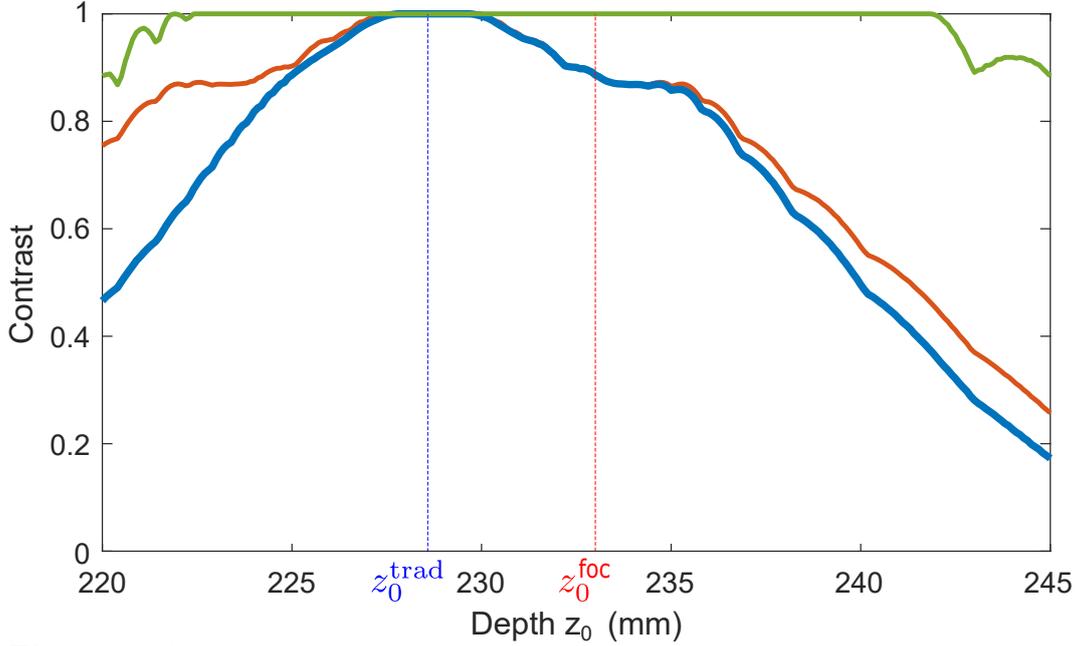


Figure 2.27 – Contrast curves versus depth: with defocus and magnification only (green), with diffraction and magnification only (orange) and with the three effects combined (blue). Depths $z_0^{\text{foc}} = 233\text{mm}$ and $z_0^{\text{trad}} = 228.6\text{mm}$. Note that the small oscillations are due to pixel discretization of the signal.

magnification (orange) is maximal at and symmetric around the z_0^{trad} position, and not at the expected z_0^{foc} position. When the three effects are combined (blue curve), the maximum contrast is located around z_0^{trad} . Defocus is responsible for an asymmetry, with a slow decrease of contrast from z_0^{trad} to a plateau around z_0^{foc} . The fast variations in magnification, combined with diffraction and defocus, explain the fast decrease in the contrast curve: higher contrast is measured for larger magnification (around z_0^{trad}), whereas for smaller magnification the image is more affected by defocus and mainly by diffraction. These curves have to be compared to the theoretical value: a region of maximal contrast over 16.3mm around the reference position z_0^{foc} .

The shapes of the curves can be generalized to other sets of parameters. When considering magnification and defocus only, the contrast curve is centered around z_0^{foc} . On the other hand, the curve combining magnification and diffraction is centered at z_0^{trad} . When the three effects are combined, the shape of the contrast curve will depend on the respective magnitudes of magnification, diffraction and defocus.

2.5.2 Comparison with experimental data

The results from the simulation are now compared to the experimental data (see Fig. 2.23). Fig. 2.28a shows sub-images extracted from raw plenoptic

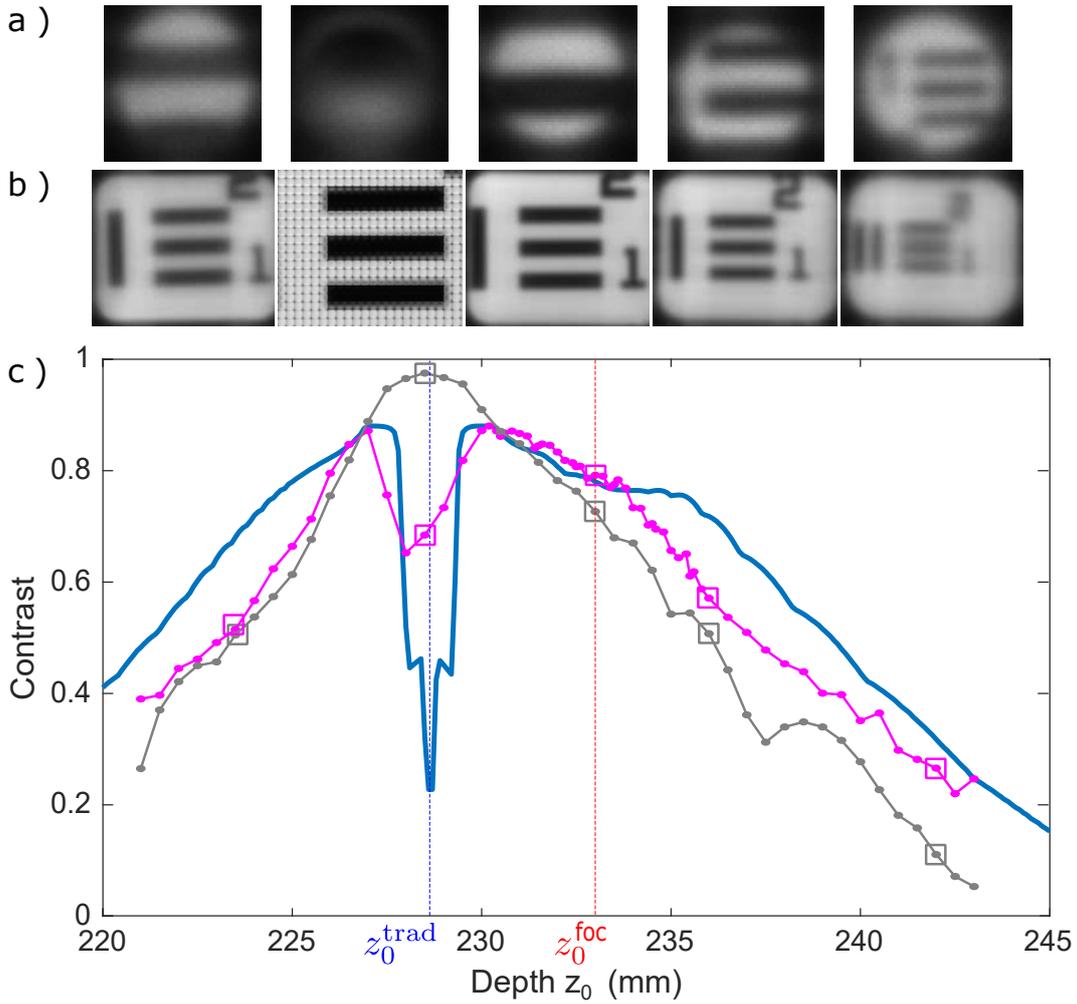


Figure 2.28 – a) Raw sub-images acquired at planes: $z_0 = 223.5, 228.5, 233, 236, 242$ mm; b) Corresponding reconstructed images; c) Comparison of the simulated contrast curve (blue) and experimental contrast curve measured on raw sub-images (pink) and on reconstructed images (gray). The gap located around $z_0^{\text{trad}} = 228.6$ mm is due to the large magnification that makes the bars exceeding the sub-image size.

images acquired at different depths. In this configuration, total magnification is maximum at the traditional plenoptic plane $z_0^{\text{trad}} = 228.6$ mm, and decreases as depth increases (Eq. 2.46). Fig. 2.28a confirms this effect: at 228.5 mm, only a part of one bar appears while as the depth increases, three bars can slowly be seen on the sub-images.

For each experimental raw plenoptic image we measured the contrast of one sub-image taken in the middle of the image, and plot it against depth (pink curve in Fig. 2.28c). It can be compared to the simulated curve (blue, Fig. 2.28c), obtained using the parameters of the setup (see tables 2.2 and 2.3).

Both curves present the same shape, with maximal values at $z_0 = 227$ and 230 mm surrounding a large gap located at z_0^{trad} . The gap is due to the

2.5. SIMULATION OF THE IMPACT OF MAGNIFICATION AND DIFFRACTION ON DEPT

magnification being so large that the image of the bars does not fit inside the sub-image (as illustrated by the 2nd figure from the left in Fig. 2.28a). This leads to unreliable contrast values measured on only a part of the total magnified pattern, similarly to what happens in the experimental setup.

On both sides of this gap, the two curves slowly decrease following the decrease in magnification. Near the position z_0^{trad} , magnification is large enough so that diffraction and defocus have little effect on the measured contrast. When the bars are smaller on the detector (far from z_0^{trad}), defocus and especially diffraction reduce the contrast. At z_0^{foc} , defocus is minimal. Effect of defocus increases when z_0 moves further from z_0^{foc} which explains the asymmetry of the curves around the z_0^{foc} position compared to the left part of the curve.

Although the simulated (blue) and measured (pink) curves are close, they do not overlap. It may be explained by a difference of dynamic ranges: the experimental images are affected by noise and light adjustments that will affect the intensity distribution and thus the measured contrasts in the image.

The quality of a light-field imaging system depends on the combination of an optical systems with a reconstruction algorithm to compute the final images. In our work, we chose to estimate the contrast on raw sub-images instead of the reconstructed ones. To validate such a choice, we compare the contrast curve from the simulation with the one measured on the reconstructed, which was previously presented in Figure 2.23. As expected, the resulting contrast curve (gray) has a similar shape to the one from raw images (pink) except around z_0^{trad} , where it reaches a local maximum instead of a gap, due to the refocusing process.

To conclude, the simulation confirms that the plenoptic depth of field is more complex than expected by theoretical equation 2.37. This equation was established considering only defocus and diffraction. Without the magnification, the resulting contrast curve is expected to present a plateau centered around z_0^{foc} . The simulation allows to deeper analyze the combined effects of diffraction, defocus and magnification on the resulting image. Magnification together with diffraction is responsible for important unexpected effects. The contrast curve reaches a maximum at the traditional plenoptic configuration z_0^{trad} instead of the reference depth plane z_0^{foc} . Defocus is responsible for an asymmetry with a slow decrease in-between these two depths and a plateau around z_0^{foc} . Beyond this interval, the contrast decreases quickly due to the fast variations of magnification.

These results will have consequence when building our X-ray plenoptic system. The maximal contrast at z_0^{trad} instead of the expected z_0^{foc} is an obstacle for depth extraction based on contrast measurement. The variations of magnification over the whole depth range is a potential problem for 3D reconstruction: not only is it responsible for contrast reduction, but it will also lead to large variations of size of the image depending on its location. As a

solution to this issue, a telecentric plenoptic system can be built, meaning the magnification is constant across depth planes [LNA⁺06; WN96]. Furthermore, appropriate algorithms should be used in order to correct for potential size errors [DW88].

2.6 Conclusion

In this chapter, we presented a theoretical and experimental study on the resolutions of a plenoptic system. First, the equations that govern the design of a plenoptic camera were revised. A new version of the aperture matching condition was given. The case of illumination by transmission of light through the sample was studied in details. Discarding these illumination conditions involves the risk of restricting the total amount of acquired data, and reducing the performances of the system.

Four optical resolutions were defined, allowing to evaluate the optical performance of the plenoptic system, prior to any refocusing. The theoretical resolutions were then checked experimentally in the visible. Three of the four resolutions (res_{lat} , res_{depth} and FOV) were validated by the experiments. Simulations were necessary to better understand the differences between theoretical and experimental plenoptic depth of field.

From our study, it appears that Δ_{peff} plays an important role in most of the equations. Its size restricts the amount of acquired data, which concretely shows how diffraction affects the resolutions. The second important parameter is the numerical aperture of the main lens $NA_{obj\ ML}$. A larger numerical aperture means smaller depth resolution and plenoptic depth of field. The spatio-angular sampling (N_u and N_s) is the third important parameter to consider in order to optimize the data acquired by a plenoptic system.

The equations presented in table 2.1 should be used when designing a plenoptic system, as was done to design and optimize the X-ray plenoptic system in Chapter 5.

Chapter 3

New parameterization and refocusing algorithm

3.1 Introduction

Plenoptic raw data have to be processed to render 2D images or a 3D volume. In Chapter 1, we highlighted that the final resolutions of the reconstructed image depend on both the optical configuration in which data have been acquired, and the algorithm used to refocus them. The present chapter describes the developments achieved during this thesis, in order to improve the quality of the rendered images. The refocusing algorithm should be able to well exploit the information contained in the plenoptic raw image, in the perspective of an accurate 3D reconstruction.

In Chapter 1, we presented the two refocusing algorithms that are commonly used to reconstruct plenoptic data of the traditional or focused configurations. We highlighted that they are actually based on the same principle of integration of lines in the pixel-micro-lens diagram, with exactly the same slope of integration. The two methods differ in the use of the pixels acquired by the sensor: the "shift-and-sum" algorithm performs interpolation over pixels of adjacent sub-images, whereas the "rendering with blending" method extracts an integer number of pixels, leading to a discrete refocusing parameter P and a discontinuous range of possible depths of reconstruction. Some raw images from both traditional and focused configurations were reconstructed using the two algorithms (see Chapter 1). The results demonstrate that the choice of the algorithm impacts the resulting resolutions and image quality.

In this chapter, we present a new refocusing algorithm that overcomes the distinction between traditional and focused algorithms. The idea is to propose an algorithm usable on any plenoptic data, regardless of the configuration in which they were acquired. Unifying both approaches allows to remove the previously mentioned difference in the use of the pixels. Moreover, the link between the two optical setups have been pointed out, as it was shown that

it is possible to switch from one configuration to another one inside the same physical setup (see Section 2.1.1). Having a unique method valid for any configuration allows to homogenize the result of the refocusing process.

We first propose a unique set of equations describing the propagation of the light rays in a plenoptic setup (Section 3.2). Based on this equation, we introduce a new refocusing approach valid for any plenoptic configuration (Section 3.3). With this method, we aim at refocusing images with arbitrary resolution at any distances from the camera, with good depth accuracy and without the discontinuity due to change of optical configuration. Our approach may also be used to generate synthetic data. Based on the first implementation (see Section 3.4), we demonstrate on synthetic data that our approach accurately simulate refocusing: the minimal blur is closely located nearby the expected depth (maximum error of 0.24% in our tests from Section 3.5). We also demonstrate the ability of our approach to work on real captured data. Finally, we point some current limitations and possible improvements for the implementation (Section 3.6).

3.2 Propagation from object-space to sensor-space

To model the propagation of a ray through the plenoptic system, a parameterization is needed in both the object- and sensor-spaces. The link should be made between these two spaces, as a transformation of the coordinates of a ray. We review some previously published models before establishing our own equations. Our approach is physically-based, in the sense that it describes the ray propagation as a function of the physical distances of the setup, without assuming a specific configuration or any relationship between the distances and optical parameters.

3.2.1 Previous parameterizations and models

A light field representation depends on the way the rays are parameterized. In Chapter 1, we introduced the two-planes parameterization which is the basis of many works. A property of the light-field is that it remains constant along a ray in free-space propagation [McC14]. In the context of geometric optics, the ray \mathbf{r}_{sen} reaching the sensor can be related to the incident ray \mathbf{r}_{rec} on the system (see Fig. 3.1). Due to the conservation of radiance, the same light field is the same in the different parameterizations: $\text{LF}(\mathbf{r}_{\text{rec}}) = \text{LF}(\mathbf{r}_{\text{sen}})$. In order to be independent from the choice of the parameterization planes (either object- or sensor-spaces), the transformation between \mathbf{r}_{sen} and \mathbf{r}_{rec} has to be modeled.

Previous transformations are usually dependent on the experimental configuration and the choice of the planes of parameterization [Ng06; GL10a]. Dansereau et al. [DPW13] have proposed a 5×5 matrix to simulate the transformation of the rays. However, the μ LA is modelled as an array of pinholes, ignoring the spatial extent of the apertures of the micro-lenses. Mignard-Debise et al. [MDRI17] take into account the aperture of the micro-lenses in their equivalent camera array framework. However, they do not explicitly propose a mathematical model of transformation in-between the ray parameterizations. In this chapter, we demonstrate that this transformation is affine and can be expressed with a 4×4 separable matrix, and that it does not introduce any discontinuity when moving from an optical configuration to another (e.g., from traditional to focused plenoptic configuration).

3.2.2 2D derivation of the affine transformation of rays

The first step of our approach is to model what happens to a single light ray when passing through the optical system. As explained in Chapter 1, the 4D light-field coordinates could be decoupled into 2×2 D. As illustrated in Figure 3.1, we express the parameterization change when a 2D ray \mathbf{r}_{rec} expressed in the object-space (between the reconstruction plane and the main lens) is transformed into a 2D ray \mathbf{r}_{sen} in sensor-space (between the μ LA and the detector planes). In this chapter, we use the term sensor-space instead of the classical image-space to prevent the reader from being confused with the multiple use of the word image. The distance between the main lens and the μ LA is written c .

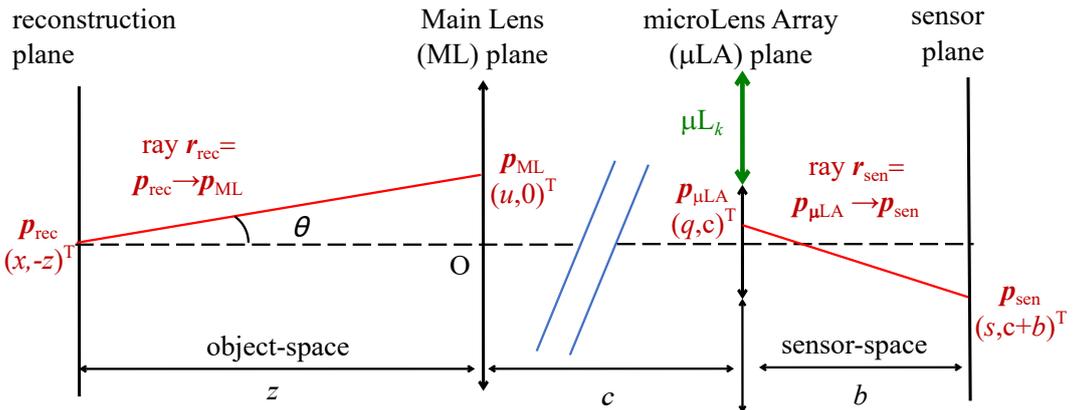


Figure 3.1 – Notations for equation derivation of the ray transformation through the optical system.

Using the coordinates of rays as defined in Figure 3.1, and assuming a micro-lens μL_k whose optical center is located at (q_k) , we model the successive effects of the main lens and the micro-lens μL_k using the thin lens equations.

We consider the successive intersections between the ray and the different planes in the system. We model these intersections by two coordinates, corresponding to the position of the intersection point on the plane, and the signed distance between the plane and the main lens. The center of the main lens is chosen as the origin of the coordinate system.

The incident ray on the system is parameterized by a pair of points: $\mathbf{p}_{\text{rec}} = (x, -z)^T$ on the reconstruction plane and, $\mathbf{p}_{\text{ML}} = (u, 0)^T$ on the main lens plane. We define \mathbf{p}_1 as the image of \mathbf{p}_{rec} on the intermediate image plane after going through the main lens of focal f_1 :

$$\mathbf{p}_1 = \frac{f_1}{z - f_1} \begin{bmatrix} -x \\ z \end{bmatrix}. \quad (3.1)$$

The ray $\mathbf{p}_{\text{rec}} \rightarrow \mathbf{p}_{\text{ML}}$ is thus transformed into the ray $\mathbf{p}_{\text{ML}} \rightarrow \mathbf{p}_1$ which intersects the μLA plane located at $(0, c)^T$ at the position

$$\mathbf{p}_{\mu\text{LA}} = \begin{bmatrix} q \\ c \end{bmatrix} \text{ with } q = u - \frac{c(z - f_1)}{f_1 z} u - \frac{c}{z} x. \quad (3.2)$$

Similarly to \mathbf{p}_1 , we define \mathbf{p}_2^k as the image of \mathbf{p}_{ML} through the micro-lens μL_k of focal f_2 , centered at the position $(q_k, c)^T$:

$$\mathbf{p}_2^k = \frac{f_2}{c - f_2} \begin{bmatrix} -u + q_k \\ c \end{bmatrix}. \quad (3.3)$$

The ray $\mathbf{p}_{\text{ML}} \rightarrow \mathbf{p}_{\mu\text{LA}}$ is thus transformed into the ray $\mathbf{p}_{\mu\text{LA}} \rightarrow \mathbf{p}_2^k$ which intersect the sensor plane located at $(0, c + b)^T$ at the position

$$\mathbf{p}_{\text{sen}} = \begin{bmatrix} s \\ c + b \end{bmatrix} \text{ with } s = \frac{bcx}{f_2 z} - \frac{(c + b)x}{z} + \frac{bc(z - f_1)u}{f_2 f_1 z} - \frac{bu}{f_2} - \frac{(c + b)(z - f_1)u}{f_1 z} + u + \frac{bq_k}{f_2}. \quad (3.4)$$

These equations make the link between the ray $\mathbf{r}_{\text{rec}} = (x, u)$ between \mathbf{p}_{rec} and \mathbf{p}_{ML} , and the ray $\mathbf{r}_{\text{sen}} = (q, s)$ between $\mathbf{p}_{\mu\text{LA}}$ and \mathbf{p}_{sen} . Considering the vertical coordinates of these points on the successive planes (Eqs. 3.2 and 3.4), we establish the relationship:

$$\begin{bmatrix} q \\ s \end{bmatrix} = \frac{1}{z} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \frac{b}{f_2} \begin{bmatrix} 0 \\ q_k \end{bmatrix} \quad (3.5)$$

or

$$\mathbf{r}_{\text{sen}} = \frac{1}{z} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \mathbf{r}_{\text{rec}} + \frac{b}{f_2} \begin{bmatrix} 0 \\ q_k \end{bmatrix} \quad (3.6)$$

with

$$m_{11} = -c, \quad m_{12} = z + c - \frac{cz}{f_1}, \quad m_{21} = \frac{cb}{f_2} - (b + c)$$

$$\text{and } m_{22} = b + z + c - \frac{(b+c)z}{f_1} - \frac{(z+c)b}{f_2} + \frac{czb}{f_1 f_2}.$$

3.2.3 Matrix equation of the 4D transformation

Having established the 2D transformation of the ray through the optical system, the equation can be easily extended to 4D adding the missing 2 coordinates. This leads to the following relation:

$$\mathbf{r}_{\text{sen}} = \mathbf{M}\mathbf{r}_{\text{rec}} + \frac{b}{f_2}(0, q_k, 0, r_l)^T \quad (3.7)$$

$$\text{with } \mathbf{M} = \frac{1}{z} \begin{bmatrix} m_{11} & m_{12} & 0 & 0 \\ m_{21} & m_{22} & 0 & 0 \\ 0 & 0 & m_{11} & m_{12} \\ 0 & 0 & m_{21} & m_{22} \end{bmatrix}$$

Thus, knowing the 4D coordinates of a ray in object-space, we can calculate the exact coordinates of the corresponding ray in sensor-space. Since \mathbf{M} is invertible, the inverse affine transformation is possible from \mathbf{r}_{sen} to \mathbf{r}_{rec} .

This equation depends on the focal lengths of the lenses (f_1 and f_2) and on the distances between the optical elements (c and b). There is no imposed relationship between these distances, contrary to some previous models that are dependent on the optical configuration [Ng06; GL10a]. Therefore it is continuous for all $z > 0$ and valid for any light field camera composed of a main lens, a μ LA and a detector, such as traditional and focused setups. The formulas in Equation 3.7 allow to calculate the matrix coefficients directly from the parameters of the physical system.

3.3 Proposed refocusing approach

This matrix transformation makes the link between object- and sensor-space. Taking its inverse allows to go from sensor- to object-space, according to the principle of reversibility of light. This is the main idea of our refocusing approach: rays from sensor-space \mathbf{r}_{sen} are transformed into object-space \mathbf{r}_{rec} and the refocusing is done on this new parameterization.

3.3.1 Principle and comparison with previous methods

The captured raw image corresponds to a discretization of the incident light field on the plenoptic camera. For each ray \mathbf{r}_{rec} as in Fig. 3.1, the reconstructed light field LF (\mathbf{r}_{rec}) can be written as a combination using basis functions $\Phi_{ijkl}(\mathbf{r})$ weighted by p_{ijkl} (as introduced by [GGSC96]):

$$\text{LF}(\mathbf{r}_{\text{rec}}) = \sum_{i,j,k,l} p_{ijkl} \Phi_{ijkl}(\mathbf{r}_{\text{rec}}) \quad (3.8)$$

Based on the work of Goesele et al. [GGHS03], we express $\Phi_{ijkl}(\mathbf{r})$, when adjusted to the context of light field cameras based on micro-lens arrays, as:

$$\Phi_{ijkl}(\mathbf{r}_{\text{rec}}) = \frac{z^2}{\cos^4 \theta(\mathbf{r}_{\text{rec}})} \phi^{\text{ML}}(\mathbf{u}) \phi_{kl}^{\mu \text{L}}(\mathbf{r}_{\text{rec}}) \phi_{ij}^{\text{sen}}(\mathbf{r}_{\text{rec}}), \quad (3.9)$$

where $\phi^{\text{ML}}(\mathbf{u})$, $\phi_{kl}^{\mu \text{L}}(\mathbf{r}_{\text{rec}})$ and $\phi_{ij}^{\text{sen}}(\mathbf{r}_{\text{rec}})$ are mathematical expressions of respectively the apertures of the main lens, the micro-lens μL_{kl} and the pixel p_{ij} on the sensor (we consider the surface of the pixel as an aperture as in [MDRI17]). They optically correspond to filters which equal 1 only if \mathbf{r} is going through the corresponding aperture in the optical system and 0 elsewhere. The weights p_{ijkl} thus corresponds to the flux going through the apertures of μL_{kl} and p_{ij} . The term $z^2 / \cos^4 \theta(\mathbf{r}_{\text{rec}})$ is a normalization term which expresses the energy conservation. Such a reconstruction function is based on the fact that the filtering process due to the different apertures is the optical equivalent of a numerical orthogonal projection of LF onto the basis functions Φ_{ijkl} (in the sense of min-square error).

Our goal is to reconstruct an image located at depth z . We want to calculate the value of pixel p_{mn} of this new image. Optically, it corresponds to the flux going through the aperture of the pixel. Using the same parameterization of the light field as in Fig. 3.1, the reconstruction of p_{mn} corresponds to the integration on the two parameterization planes of all the rays from the light-field which go through the new pair of apertures (main lens aperture $\psi^{\text{ML}}(\mathbf{u})$ and equivalent aperture of the pixel in object-space $\psi_{mn}^{\text{pixel}}(\mathbf{r}_{\text{rec}})$):

$$p_{mn} = \int_{\mathbf{x}} \int_{\mathbf{u}} \text{LF}(\mathbf{r}_{\text{rec}}) \psi^{\text{ML}}(\mathbf{u}) \psi_{mn}^{\text{pixel}}(\mathbf{r}_{\text{rec}}) \frac{\cos^4 \theta(\mathbf{r}_{\text{rec}})}{z^2} \quad (3.10)$$

Combining Equations 3.8, 3.9 and 3.10 results in the following interpolation scheme of the recorded values:

$$p_{mn} = \sum_{i,j,k,l} p_{ijkl} \int_{\mathbf{x}} \int_{\mathbf{u}} \phi^{\text{ML}} \psi^{\text{ML}} \phi_{kl}^{\mu \text{L}} \phi_{ij}^{\text{sen}} \psi_{mn}^{\text{pixel}} \quad (3.11)$$

Thanks to the definition of the reconstruction function introduced by Goesele et al. [GGHS03], the term $\cos^4 \theta / z^2$ disappears.

When using the same aperture of main lens for acquisition and refocusing ($\phi^{\text{ML}} = \psi^{\text{ML}}$), and reminding that ϕ^{ML} equals either 1 or 0, we obtain the relationship $\phi^{\text{ML}} \psi^{\text{ML}} = \phi^{\text{ML}}$. Equation 3.11 might be simplified to:

$$p_{mn} = \sum_{i,j,k,l} p_{ijkl} \int_{\mathbf{x}} \int_{\mathbf{u}} \phi^{\text{ML}} \phi_{kl}^{\mu \text{L}} \phi_{ij}^{\text{sen}} \psi_{mn}^{\text{pixel}} \quad (3.12)$$

When the sub-images are not overlapping, each pixel p_{ij} is illuminated by maximum one single micro-lens μL_{kl} . The flux p_{ijkl} can be simplified to p_{ij} .

Therefore, the new pixel p_{mn} to reconstruct is a function of the pixels p_{ij} of the sensor, each of them being weighted by an integral factor. The key point of our algorithm consists in being able to precisely estimate this integral factor.

3.3.2 Our approach

We use this principle to reconstruct an image at any chosen distance z from the main lens. This is schematized in Fig. 3.2. For each pixel p_{mn} of the reconstructed image located at depth z , we consider the bundle of rays \mathbf{A}_{mn} defined by the pixel and the aperture of the main lens (in yellow in object-space). This corresponds to the set of rays that have been (potentially) acquired by the plenoptic system. We then transform the bundle of rays in sensor-space for each micro-lens μL_{kl} using Eq. 3.7: this gives the bundle of rays \mathbf{A}_{mnkl} (in yellow in sensor-space). Meanwhile we consider the bundles of rays \mathbf{B}_{ijkl} (in green) intrinsically defined in sensor-space by each pair of sensor pixel p_{ij} – micro-lens aperture μL_{kl} . The contribution of the sensor pixel p_{ij} to the reconstructed pixel p_{mn} corresponds to the intersection between the bundles \mathbf{A}_{mnkl} and \mathbf{B}_{ijkl} (in orange).

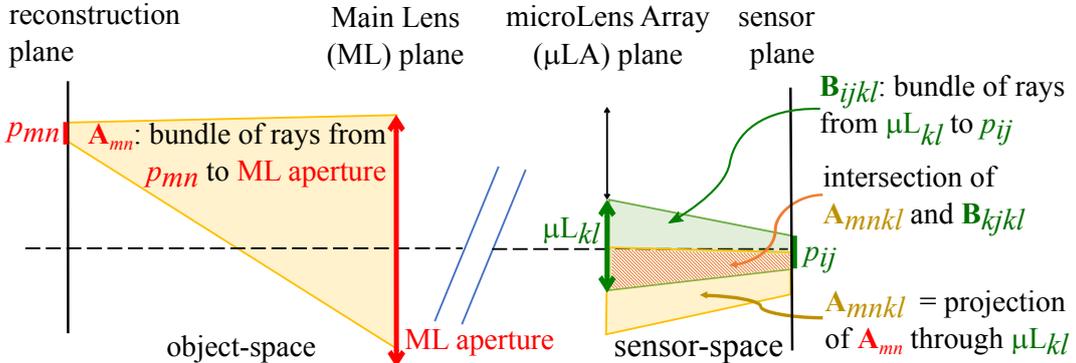


Figure 3.2 – Principle of the refocusing approach: (1) Transformation of the bundle of rays \mathbf{A}_{mn} from the reconstructed pixel p_{mn} passing by the aperture of the main lens into the sensor-space bundle of rays \mathbf{A}_{mnkl} through the micro-lens μL_{kl} . (2) This bundle of rays is intersected with the bundle of rays \mathbf{B}_{ijkl} between the sensor’s pixel p_{ij} and micro-lens μL_{kl} to compute the contribution of p_{ij} on p_{mn} .

Figure 3.3 shows the phase-space diagram of the setup. The light-field is represented in two different ways, considering a parameterization between the μ LA and the sensor planes (in black), and the parameterization between the reconstructed pixel and the main lens, projected from object-space to sensor-space (in blue). The bundle of rays \mathbf{A}_{mnkl} projected from object-space to sensor-space corresponds to a distorted parallelogram (in orange). The bundle \mathbf{B}_{ijkl} is represented by a rectangle (in green). The intersection between them

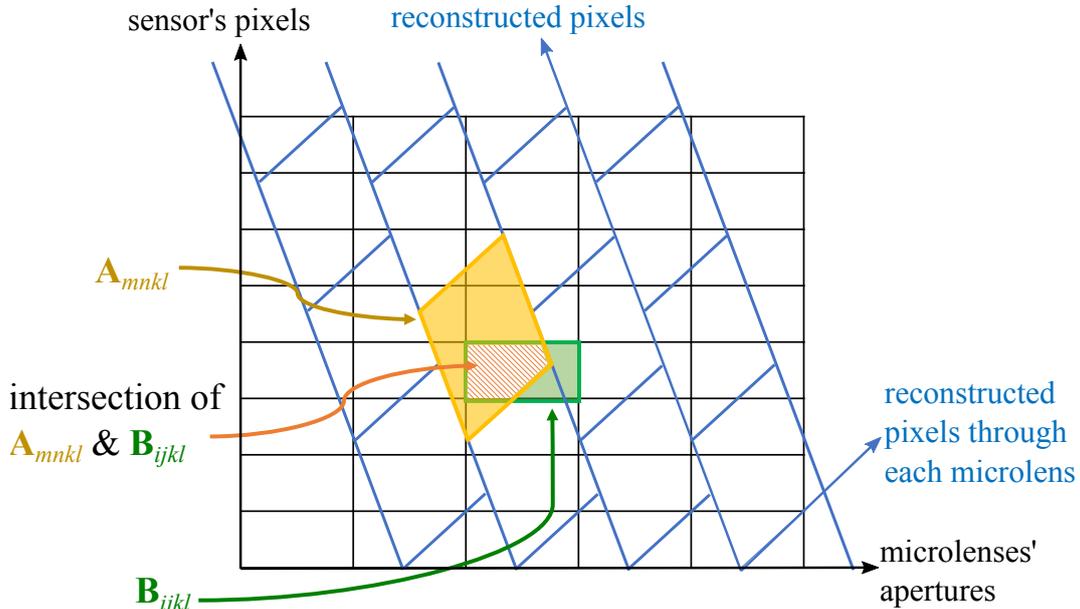


Figure 3.3 – Phase-space diagram of Fig. 3.2.

has a complex shape, whose surface (or volume in 4D) should be evaluated as the contribution of p_{ij} to p_{mn} .

To retrieve the entire value of the reconstructed pixel p_{mn} , the contribution of each sensor pixel p_{ij} should be estimated. The final value of p_{mn} is the sum of all the values of p_{ij} weighted by the volume of the intersection $\mathbf{A}_{mnkl} \cap \mathbf{B}_{ijkl}$. The approach is summarized in Algorithm 1. More formally, this corresponds to the estimation of the integral factor in Eq. 3.12. Our approach differs from the classical refocusing algorithms [Ng06; GL10a] since we consider sensor's pixels and apertures as surfaces and not as points, thus mimicking accurately the setup.

3.3.3 Simulation of a plenoptic raw image

Eq. 3.7 allows us to build the reconstruction method, and for each pixel on the reconstruction plane to fetch the useful information from the pixels on the sensor plane, for a pixel-wise reconstruction of the whole image. Likewise, the process can be reversed, and for each pixel on the sensor, it is possible to determine the pixels on the reconstruction plane from which the light is coming. Following this idea, we implemented a reversed algorithm that, for an object plane at depth z , projects the corresponding light field from the object-space onto the detector. This gives us a tool to simulate a light field image that would have been acquired from a plenoptic camera.

Algorithm 1 Reconstruction at depth z - principle

```

1: for each pixel  $p_{mn}$  do
2:    $p_{mn} := 0$ 
3:   Consider the bundle of rays  $\mathbf{A}_{mn}$ 
4:   for each micro-lens  $\mu\mathbf{L}_{kl}$  do
5:     Project  $\mathbf{A}_{mn}$  through  $\mu\mathbf{L}_{kl}$  to obtain  $\mathbf{A}_{mnkl}$ 
6:     for each sensor's pixel  $p_{ij}$  do
7:       Consider the bundle of rays  $\mathbf{B}_{ijkl}$ 
8:       Compute the volume  $V$  of  $\mathbf{A}_{mnkl} \cap \mathbf{B}_{ijkl}$ 
9:       Add the contribution of  $p_{ij}$  to  $p_{mn}$ :
10:       $p_{mn} := p_{mn} + V \times \text{value}(p_{ij})$ 
11:     end for
12:   end for
13: end for

```

3.4 Implementation details

Behind its apparent simplicity, the formula expressed in Eq. 3.12 (simple integral on apertures) is complex to evaluate. Each single ray goes through a unique combination of apertures (reconstruction plane, main lens, $\mu\mathbf{L}$, and sensor planes), and the resulting bundle of rays considered in the integral often results in a complex geometrical shape (such as the orange shape in Fig. 3.3). It is not possible to simplify the final geometrical shape by separating the contribution of each optical element, as the integration is done at once in the space of rays. The algorithm thus necessitates a high number of complex volume estimations. In order to reduce the complexity of the algorithm, we first introduce approximations on the apertures and then develop the algorithm in order to prevent un-needed computation.

3.4.1 Approximations on the apertures

Apertures of lenses could be of various shapes: hexagonal, circular... In our implementation, we approximate the circular apertures of our lenses by squares. To be conservative and to prevent introducing any use of sensor's pixel that does not contribute to the reconstructed pixel, we consider the inscribe square (and not the outer square). Even if this is a limitation of our current implementation, this approximation still takes into account the size of the aperture, which is more precise and realistic than approximating them by their center position only [DPW13].

Moreover, we assume that the apertures of our lenses are axis-aligned with the pixels of the sensor. These assumptions (square and axis-aligned) allow to keep the separability between the two pairs of coordinates (x, u) and (y, v)

that exists in the Eq. 3.7. Working on separated coordinates allows to pool some parts of code and to fasten the intersection (cf. Section 3.4.2).

3.4.2 Faster access to contributing pixels

We want to evaluate the computational complexity of our algorithm, defined by the computation time as a function of input data. The intrinsic complexity to reconstruct each pixel p_{mn} is $O(I \times J \times K \times L)$ where $I \times J$ is the number of pixels in each sub-image of the plenoptic image and, $K \times L$ is the number of micro-lenses. This can be optimized taking into account that the useful information for a single pixel p_{mn} only comes from a small subset of all the bundles of rays \mathbf{B}_{ijkl} acquired by the system. Therefore, we compute closer minimal and maximal bounds to more precisely restrict the loops over the micro-lenses and sensor's pixels (lines 4 and 6 in Algorithm 1) to effective ones.

For each p_{mn} , we use Eq. 3.7 to compute the bounding coordinates $\mathbf{q} = (q, r)^T$ that are common to all the bundles of rays \mathbf{A}_{mnkl} : they do not depend on the choice of the microlens. We thus only iterate over the micro-lenses that intersect the corresponding geometry. This corresponds to projecting \mathbf{A}_{mn} through the main lens only, and checking which micro-lenses are illuminated by the bundle of rays. Depending on the parameters, this can greatly reduce the number of micro-lenses in the loop.

Similarly, we now compute the coordinates $\mathbf{s} = (s, t)^T$ of the remaining bundle of rays \mathbf{A}_{mnkl} . This gives us the illuminated sensor's pixels over which we need to iterate. This can again highly reduce the number of considered sensor's pixels.

Finally, we use the separability property defined in Section 3.4.1 to optimize the calculation. Thanks to this property, the set of coordinates (x, u, q, s) and (y, v, r, t) are independent. For the intersection, this means that, instead of evaluating a volume intersection in 4D (coordinates (q, s, r, t) of $\mathbf{A}_{mnkl} \cap \mathbf{B}_{ijkl}$), we can calculate separately the coordinates (q, s) and (r, t) . The resulting 2D intersections are simple and the 4D volume size of the intersection is simply the product of the 2D volumes.

We present these changes in Algorithm 2, where the orange lines correspond to the new elements after optimization.

3.4.3 Correction of reconstruction vignetting

The captured light field partially represents the original light field. On the border of the image, we may lack micro-lens - sensor's pixels to fully cover the bundle of rays for the reconstructed pixels through the aperture of the main lens. This results in some vignetting effects, i.e. the decrease of intensity in the border as will be shown later in Figures 3.10 and 3.11.

In order to correct this effect, we simply compute the sum of the intersected volumes and scale down the pixel value by this sum. This way, we consider only the effective part of the bundle of rays. We obtain the final algorithm presented in Algorithm 2, where the blue lines correspond to the new elements due to this correction.

Algorithm 2 Reconstruction at depth z - optimized version

```

1: for each pixel  $p_{mn}$  do
2:    $p_{mn} := 0$ 
3:    $w_{mn} := 0$ 
4:   Consider the bundle of rays  $\mathbf{A}_{mn}$ 
5:   Compute coordinates  $(q, r)$  of  $\mathbf{A}_{mnkl}$ 
6:   Derive bounds for  $k, l$ 
7:   for each micro-lens  $\mu L_{kl}$  do
8:     Compute other coordinates  $(s, t)$  of  $\mathbf{A}_{mnkl}$ 
9:     Derive bounds for  $i, j$ 
10:    for each sensor's pixel  $p_{ij}$  do
11:      Consider the bundle of rays  $\mathbf{B}_{ijkl}$ 
12:      Compute the volume  $V$  of  $\mathbf{A}_{mnkl} \cap \mathbf{B}_{ijkl}$ 
13:      Compute  $v_{qs}$  for the  $(q, s)$  coordinates
14:      Compute  $v_{rt}$  for the  $(r, t)$  coordinates
15:       $V := v_{qs} \times v_{rt}$ 
16:      Add the contribution of  $p_{ij}$  to  $p_{mn}$ :
17:       $p_{mn} := p_{mn} + V \times \text{value}(p_{ij})$ 
18:      Add the corresponding weight to  $w_{mn}$ :
19:       $w_{mn} := w_{mn} + V$ 
20:    end for
21:  end for
22:  Correct  $p_{mn}$  by the total weight:
23:   $p_{mn} := p_{mn} / w_{mn}$ 
24: end for

```

3.5 Numerical results

We implemented the Algorithm 2 in Python, version 3.6.8, using the numpy package for matrix computation. All the computations were done on a 3.4 GHz 64-bits Intel Core i7-6700 CPU workstation with 16 GB of RAM.

To evaluate our reconstruction method, we run our algorithm on different representative cases: first, a synthetic knife-edge, then synthetic images of letters "a" and "b" at different depths, finally, a real image acquired by Georgiev et al. [GL10b].

3.5.1 Reconstruction of a single plane

We first study the ability of the algorithm to compute the refocused images of a single plane with a good depth accuracy. For this purpose, we choose to work with a simple pattern, a knife-edge, which is a sharp and straight edge along one direction. This pattern allows us to study the defocusing effects due to the reconstruction process of our algorithm along this single direction.

We choose a configuration based on real optics easily found on the market, while trying to optimize the acquired angular and spatial data. The parameters and distances of this configuration are displayed in the first column in Table 3.1.

		configuration 1	configuration 2
Main Lens	focal f_1	100mm	200mm
	aperture d_1	60mm	12mm
Micro-lens Array	focal f_2	4.1mm	18.6mm
	aperture d_2	150 μ m	600 μ m
	# micro-lenses	40x40	4x4
Sensor	pixel size Δ_p	3.225 μ m	3.548 μ m
	# pixels	2064x2064	707x707
Distances	object-ML or z_0	105.64mm	490mm
	ML- μ LA c	2086mm	450mm
	μ LA-detector b	4.18mm	22.3mm

Table 3.1 – Parameters used for the simulations and the reconstructions of the knife-edge and the letters examples.

We model a knife-edge by a virtual black and white image (cf. Fig. 3.4-left), that is located at depth $z_0 = 105.64mm$. As detailed in Section 3.3.3, we simulate the corresponding plenoptic image that would have been acquired by a light field camera in configuration 1. This results in the plenoptic raw image presented in Fig. 3.4-right. Then, we use this synthetic plenoptic image as an input to reconstruct the knife-edge pattern at different depths. We reconstruct images regularly spaced from $z = 105.2mm$ to $z = 106.5mm$, in steps of 0.01mm. In Fig. 3.5-left (top), we show the reconstructed image at depth $z = 106.20mm$. This position is far from the original depth $z_0 = 105.64mm$, hence the blur which occupies nearly the whole width of the image. For each reconstructed image, we extract the same single row (in red in Fig. 3.5-left (top)). We present this row according to depth in Fig. 3.5-left (bottom).

In Fig. 3.5-left (bottom), we observe that the sharpest reconstruction corresponds to depth $z_0 = 105.64mm$, in the sense that the transition between white and black is done over the fewest pixels. This depth is the one where

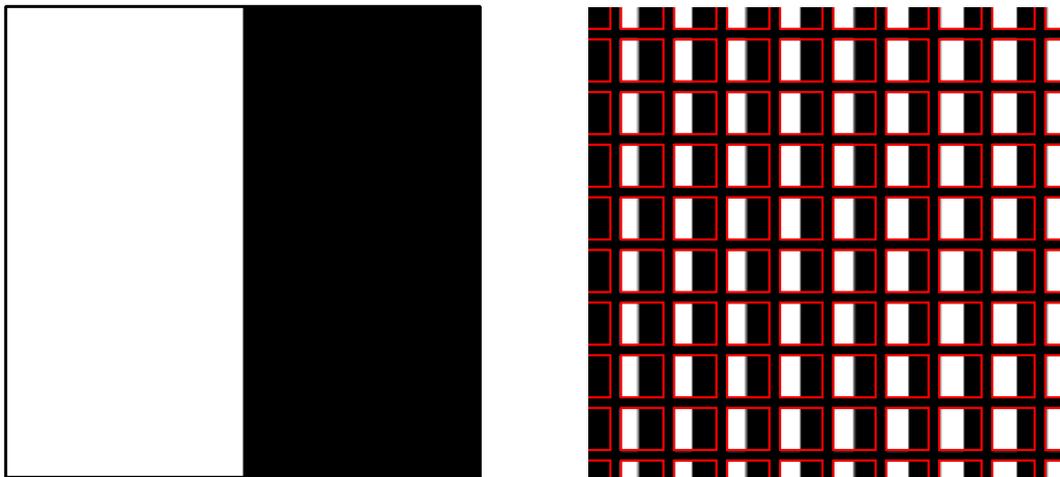


Figure 3.4 – Original virtual knife-edge (left) and zoom in the plenoptic raw image simulated by our algorithm in configuration 1 (right). The red squares represent the micro-images due to the micro-lenses.

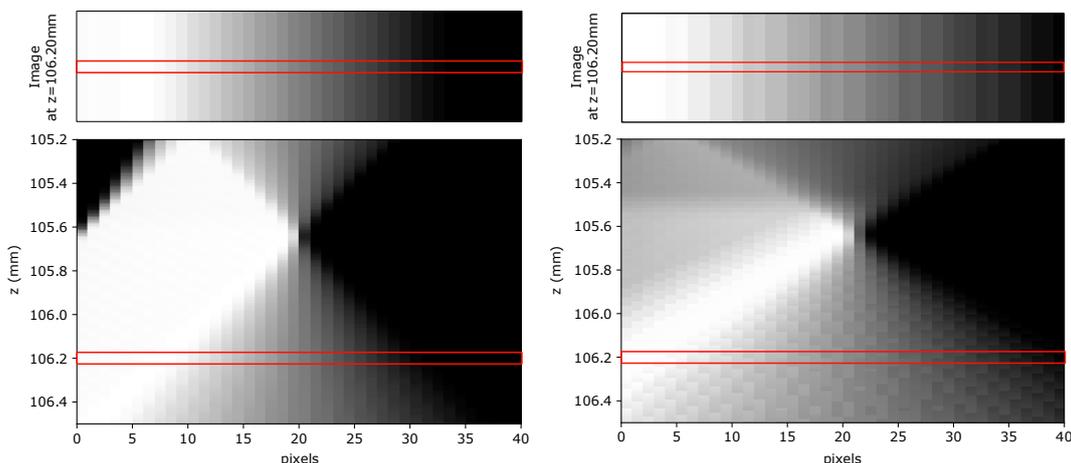


Figure 3.5 – Reconstructions of the knife-edge in configuration 1 with our algorithm (left) and using algorithm from [Ng06] (right). The two top figures are crops of the reconstructed images at $z = 106.20\text{mm}$ (size 10×40). The bottom figures show the evolution of a single row extracted from the reconstruction images, with depth from $z = 105.2\text{mm}$ to $z = 106.5\text{mm}$ in steps of 0.01mm . The rows extracted from the images at $z = 106.20\text{mm}$ are highlighted in red.

the knife-edge was originally located. For smaller or larger values of z , the edge is progressively blurred due to the defocus from the original depth z_0 of the knife-edge. For smaller z , we can also notice edge effects due to the lack of information in the virtually acquired light field (top-left black corner in Fig. 3.5-left). This corresponds to the *FOV* reduction according to the depth of the reconstruction (see Equation 2.17 in Section 2.2.1).

As a comparison, we reconstruct the same synthetic plenoptic raw image (Fig. 3.4) using the shift-and-sum algorithm that was presented for traditional

plenoptic cameras (see Chapter 1, [Ng06]). In Fig. 3.5-right (top), we show the reconstructed image at depth $z = 106.20mm$ for comparison with our reconstruction. We extract the same rows of the reconstructed images and display them in Fig. 3.5-right (bottom) as a function of depth. The rows follow the same hourglass shape as in Fig. 3.5-left (bottom), with a minimum of white-to-black transition located at depth $z_0 = 105.64mm$, and progressive blur on both sides. However, the stack reconstructed using the shift-and-sum algorithm exhibits some artifacts at high depths. They are responsible for the vertical lines of various gray tones in Fig. 3.5-right (top). We also notice a large shadow on the left of the row evolution, which exceeds the black corner visible in our reconstructions. This artifact affects the reconstructed images over nearly the whole depth range. The reconstructions took 2min13s with our algorithm and 0.06s with [Ng06] on average.

We do not compare with the patch method for focused plenoptic cameras (see Chapter 1, [GL10a]) due to its lack of ability to continuously reconstruct an image at any depth, as pointed in Chapter 1.

3.5.2 Performance analysis

To quantify the accuracy of the reconstruction, we use the criteria of the root mean square error (RMSE) (as in [DPW13; VMH⁺19]). It quantifies the mean error between the reconstructed versus original images:

$$RMSE(z) = \sqrt{\frac{\sum_{m,n}^{M,N} (p_{mn}^z - p_{mn}^{\text{orig}})^2}{M \times N}},$$

with p_{mn}^z the value of the pixel p_{mn} in the image reconstructed at depth z , p_{mn}^{orig} the value of the pixel p_{mn} in the original image, and $M \times N$ the number of pixels in both images. The depth of the minimum RMSE corresponds to the depth of the best reconstructed image (with less error compared to the original image).

The variations of RMSE against the depth z from the reconstruction (cf. Fig. 3.5-left) is presented in Fig. 3.6a. We observe a fast decrease of the RMSE to a minimum located at $z = 105.64mm$, that corresponds to the depth z_0 where the original image was localized. The regular slopes on both sides of the minimal value show the ability of our approach to reconstruct a sharp and focused image at its exact original position, here with a precision of $0.01mm$. It also creates the appropriate defocus blur on both sides of this position.

We also compute the RMSE curve for the shift-and-sum algorithm (cf. Fig. 3.6b). In Fig. 3.6c, we display on the same diagram the two RMSE curves for our algorithm (blue) and the shift-and-sum refocusing method (orange). We observe the following:

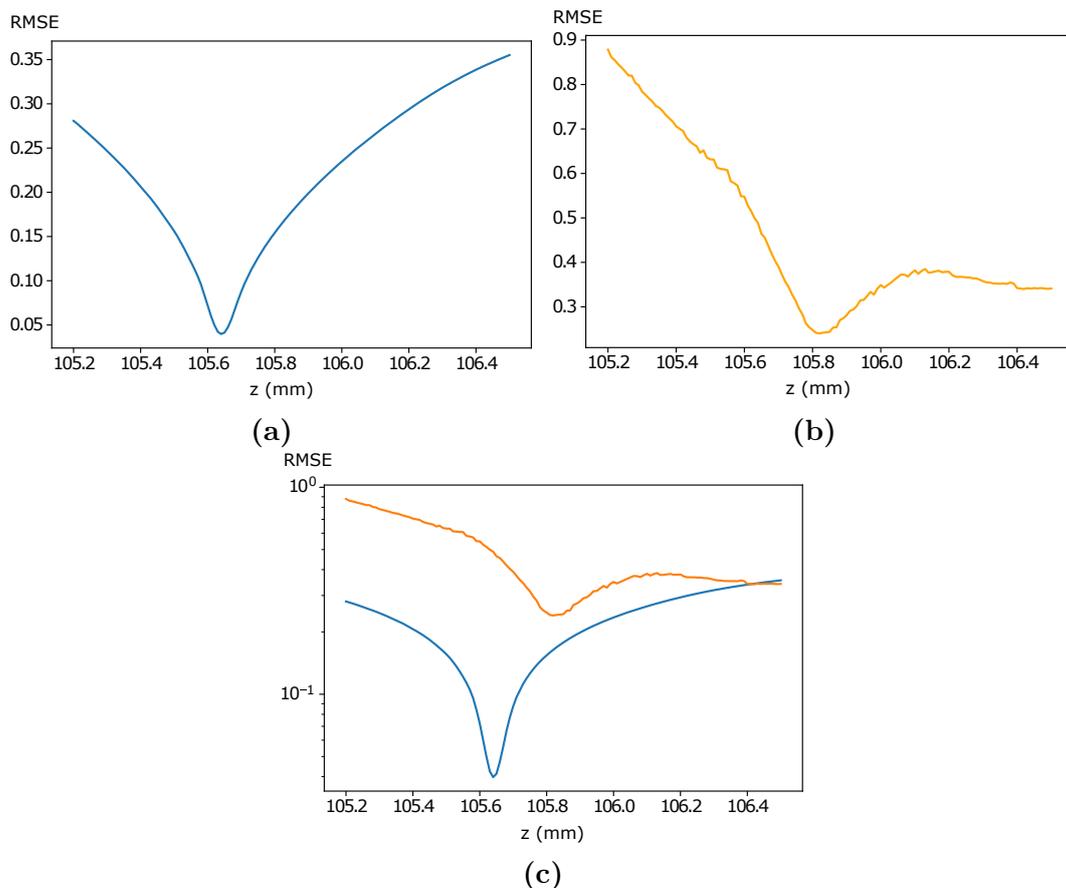


Figure 3.6 – Root Mean Square Error (RMSE) between original and reconstructed images of the knife-edge in configuration 1: using our algorithm (3.6a), using algorithm from [Ng06] (3.6b), and comparison of the two curves (3.6c).

- the orange curve from the shift-and-sum refocusing method has higher values than the ones from our approach on the whole range of refocusing depths. This shows that, at each depth, our algorithm provides a better reconstruction, with less error compared to the original image;
- the orange curve has its minimum value at $z = 105.82\text{mm}$, slightly shifted from the correct value $z_0 = 105.64\text{mm}$ that our approach is able to detect;
- on the right side, the rise of the orange curve is slower than the blue one. This shows the ability of our algorithm to reconstruct a regular defocus blur on both sides of the correct depth of the object.

3.5.3 Reconstruction for another optical configuration

We also test our simulation and reconstruction approaches with a more extreme optical configuration with only 4×4 micro-lenses (named configuration 2 in Table 3.1) We also use the knife-edge but located at depth $z_0 = 490\text{mm}$

to compute a new simulated raw image, composed of 4×4 sub-images (cf. Fig. 3.7a).

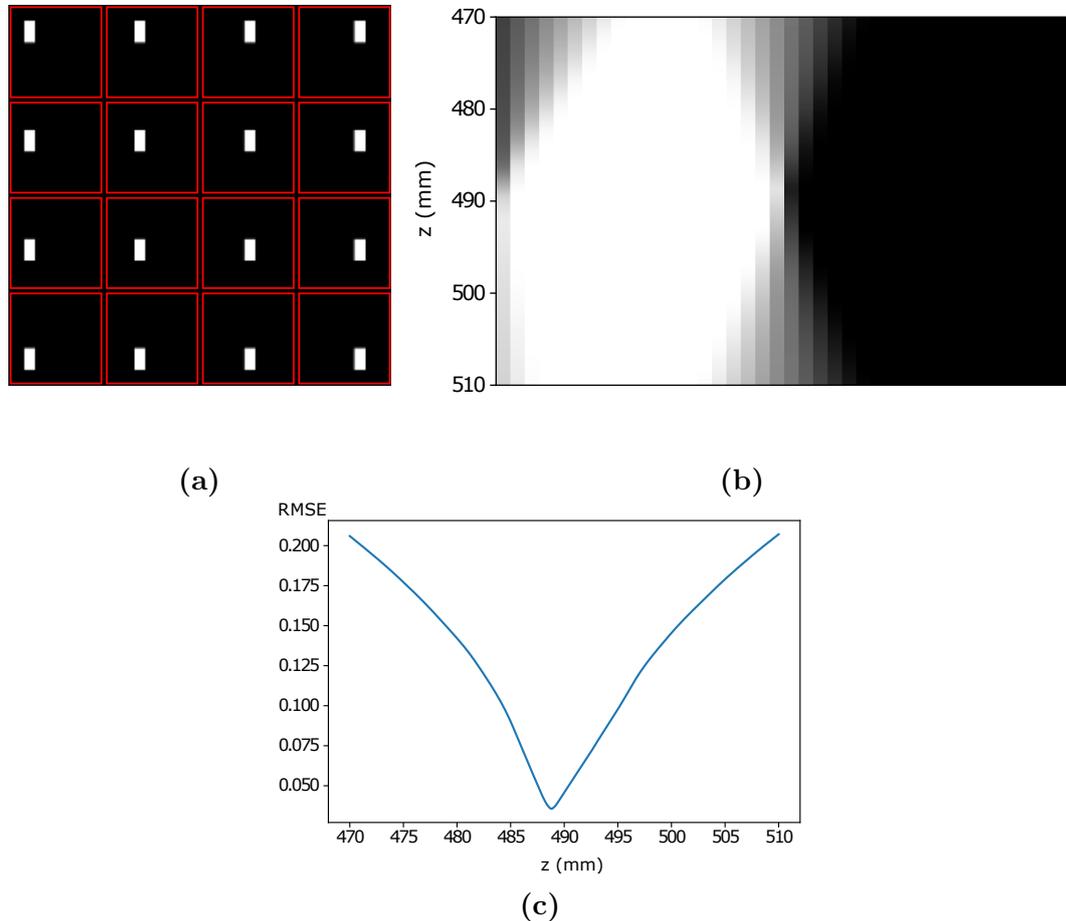


Figure 3.7 – Reconstruction of the knife-edge in configuration 2 with our algorithm: the plenoptic image simulated by our algorithm with red squares delimiting the sub-images (3.7a), evolution of a row extracted at regular steps of 0.1mm between depth $z = 470\text{mm}$ and $z = 510\text{mm}$ (3.7b), and the corresponding RMSE curve between original and reconstructed images (3.7c).

We use the similar methodology to test our approach in this configuration. As shown in Fig. 3.7b and 3.7c, we obtain similar results with a minimum RMSE located at $z = 488.8\text{mm}$ instead of the original position $z_0 = 490\text{mm}$ (error of 0.24%). This shifted depth can be explained by the low number of micro-lenses.

By definition, the reconstructed image by the shift-and-sum algorithm [Ng06] has a number of pixels that corresponds to the number of micro-lenses. In this configuration, it would produce an image with only 4×4 pixels. This demonstrates the flexibility of our approach to deal with a low number of micro-lenses, as we were able to generate an image with a sufficient number

of pixels.

3.5.4 Objects at different depths

We now go back to configuration 1 (cf. Table 3.1) to study the behaviour of the algorithm with two objects placed at different depths (cf. Fig 3.8-left). We choose as objects the images of a letter "a" and "b" placed at depths $z_0^a = 105.64mm$ and $z_0^b = 105.80mm$ respectively. We have chosen the position of the two letters in order to ensure that they are not hiding each other in any sub-image of the plenoptic image. We thus simulate separately the corresponding plenoptic images for the two letters and sum them to obtain the final one (cf. Fig 3.8-right).

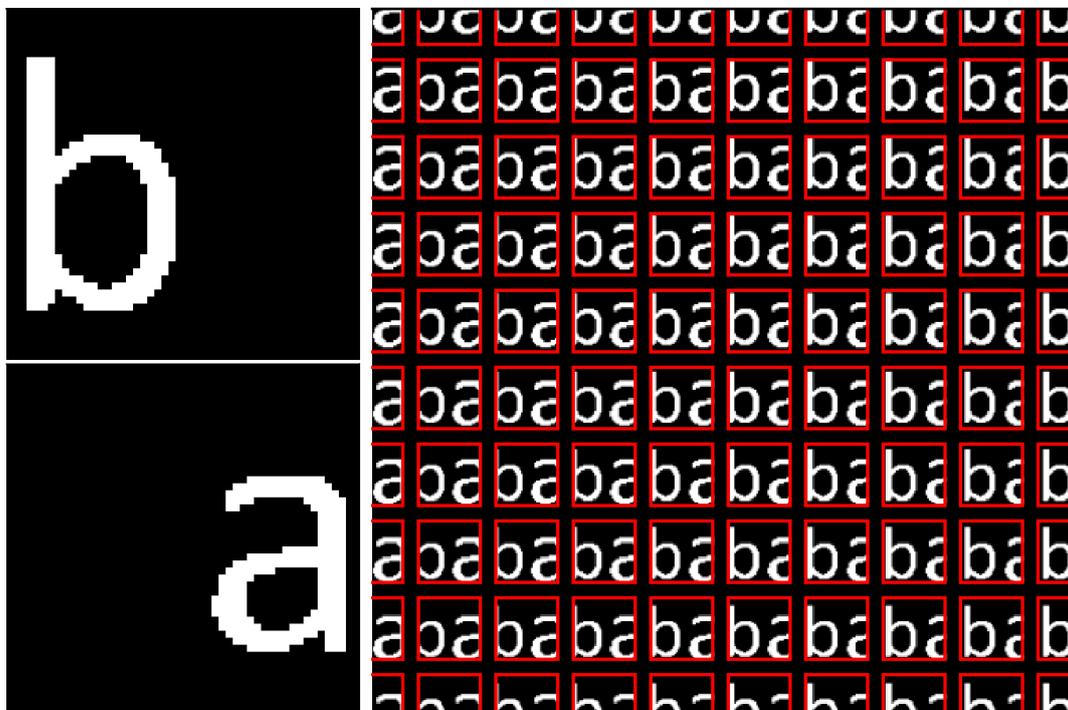


Figure 3.8 – *Test with two letters a and b. We use the two images on the left separately, located respectively at $z_0^a = 105.64mm$ and $z_0^b = 105.80mm$ to generate synthetic plenoptic images that we combine in the final result shown partially on the right.*

Fig. 3.9-top shows reconstructed images at the original depths where the two letters were placed. The sharpest reconstruction is obtained for each letter for a refocusing at their respective original depths. When looking at the in-focus letters, the transition between black and white pixels is done by only one or two blurred pixels. On the contrary, when looking at the out-of-focus letters, we observe a large blur over more pixels. Despite the small distance between the two positions, the defocus blur strongly affects the reconstruction.

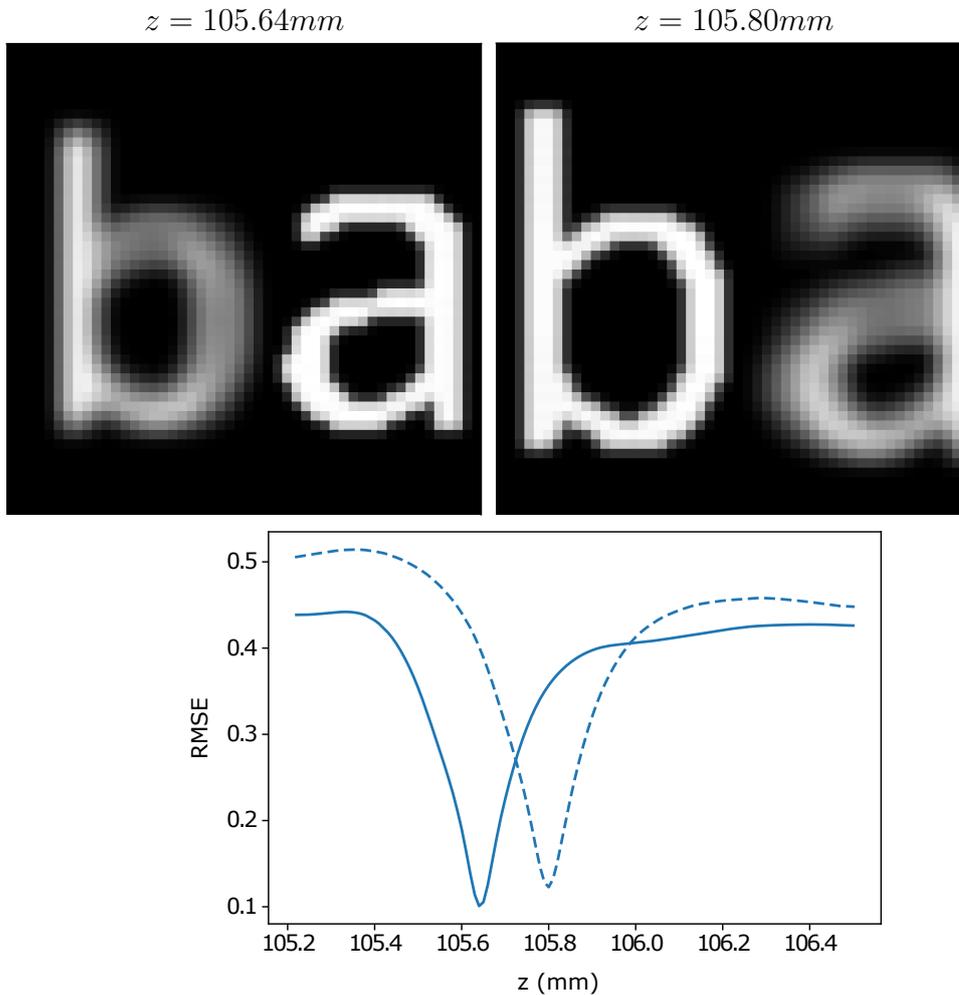


Figure 3.9 – Reconstructed images from the two letters test scene (top) with the corresponding RMSE curves (bottom) between reconstructed images and original letter **a** (blue), and between reconstructed images and original letter **b** (dotted blue).

Again, we display the RMSE curves along the whole depth range (cf. Fig. 3.9-bottom). We extract the RMSE values between the reconstructed images and the original images of letter **a** and **b** separately, in order to verify the accuracy of the reconstruction for each position. We notice that each RMSE curve has its minimum at the exact position where the corresponding letter was originally put: $z_0^a = 105.64mm$ for letter **a** and $z_0^b = 105.80mm$ for letter **b**, again with a precision of $0.01mm$. This shows the possibility to refocus accurately even when there is multiple objects at different depths.

3.5.5 Reconstruction of a real plenoptic image

We now use our algorithm to reconstruct an image acquired by an experimental plenoptic setup. We choose an experimental image acquired by Georgiev et

al. [GL10b] available on-line. This image represents small objects and smooth variations of distances: a pencil holder with a small chain in the front. A part of the original plenoptic image is shown in Fig. 3.10-left.

We use the parameters provided in [GL10b] and estimate the missing values of distances from the given magnifications (see Table 3.2). We reconstruct an image of size 1000x1000 pixels, and display the result in Fig. 3.10-right. The computation time is 1h41minutes.

Main Lens	focal f_1	80mm
	aperture d_1	26.7mm
micro-lens Array	focal f_2	1.5mm
	aperture d_2	0.5mm
	# micro-lenses	22x23
Sensor	pixel size Δ_p	6.8 μ m
	# pixels	1629x1713
Distances	object-ML or z_0	242mm
	ML- μ LA c	133mm
	μ LA-detector b	1.7mm

Table 3.2 – Parameters used for the reconstructions of the pencil holder. The distances, missing in [GL10b], were estimated from the provided magnifications.

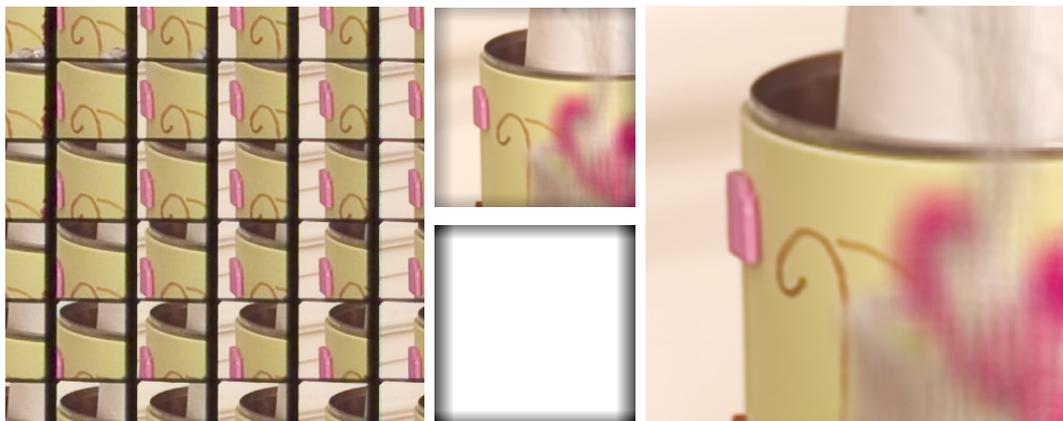


Figure 3.10 – Reconstruction of the pencil holder [GL10b] from the unmodified data available on-line. A partial view of the data is shown on the left. In the middle row, we show the uncorrected version of the result of our algorithm exhibiting vignetting effects (top) together with the image of the correction factors (bottom). Once divided by the correction factor, the final image on the right is obtained.

The reconstructed image exhibits all the expected effects, with the sharpest edges on the focused part. The chain in the front of the holder is blurred, with

patterns probably due to the intrinsic sparsity of the acquired light field as shown in [BGY⁺13]. The in-focus details are sharp and continuous (see for example the pink structure on the left, or the upper edge of the pencil holder). This image shows the ability of our algorithm to reconstruct images from real captured plenoptic images.

3.6 Analysis of the limitations

In this section, we study the limits of our new algorithm.

3.6.1 On computation time

Despite the approximation we use and the optimization of our first implementation of our approach, the remaining complexity of the evaluation of the integral makes the computation very slow. This is the major challenge we have to address.

In the short term, we may optimize more our reconstruction to take full advantage of the existing separability. Indeed, thanks to the separability property, the 2D intersections for one dimension could be used for the reconstruction of several $\mathbf{A}_{mnkl} \cap \mathbf{B}_{ijkl}$. For example, for the same values of m , k , and i , the (q, s) coordinates of $\mathbf{A}_{mnkl} \cap \mathbf{B}_{ijkl}$ will be identical, no matter the values of n , l , and j . This remark should allow to pool some calculations and to reduce the reconstruction complexity from quadratic $O(I \times J \times K \times L)$ to linear $O(I \times K + J \times L)$.

The low memory footprint of the algorithm (we only need to store the raw plenoptic and the resulting images) and the intrinsic parallelism lead to some potential large gains if we implement it on GPU. This will not change intrinsically the complexity of the algorithm, but will definitively result in a very large speed-up.

The cost is mainly due to the complex shape of the bundles of rays when projected. There is potentially two ways to address this problem. The first one is to find if it exists a combination of two planes for the parameterization that exhibits a better suited shape of the bundles to intersect and to compute the resulting volume.

The second solution is to precompute as much as possible the weight of each sensor's pixel, and store their contributions for each reconstructed pixel for different refocusing distances. This may result in a very large data set. However, the symmetries that exist in the optical systems together with the continuity in the variation of the weights lead to potential size reduction using compression schemes or fitting techniques.

3.6.2 Limits of plenoptic reconstruction

The quality of our algorithm highlights some intrinsic limitations of plenoptic imaging. When refocusing at depths outside the plenoptic depth of field (before $z = 105\text{mm}$ and after $z = 106.5\text{mm}$), some oscillations appear. They generate artifacts in the reconstructed images, visible as vertical lines in Figure 3.11a and grid patterns in Figure 3.11b).

These artifacts occur when the algorithm reaches a limit due to the sparsity of the acquired light-field. The available data correspond to a discrete spatio-angular sampling of the light-field, made by the μLA and pixels on the sensor. This spatio-angular sampling is optimal at depth z_0^{foc} , but for distant depths, the sampling becomes sparse [BGY⁺13]. This results in variations in the number of angular data across spatial positions in the image, hence the grid artifacts. These artifacts show that the refocusable depth range is limited.

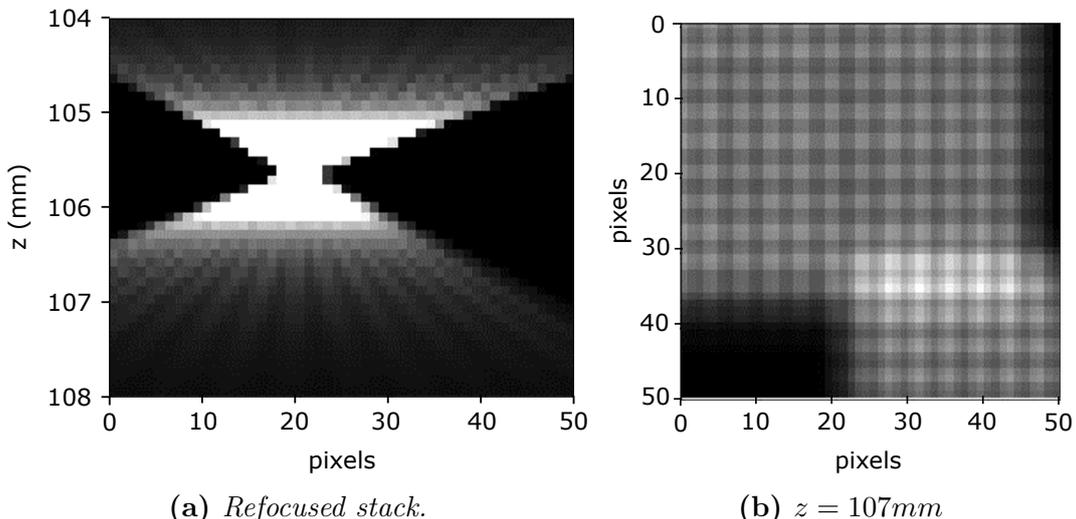


Figure 3.11 – Illustration of the grid effects on the reconstruction of a point in configuration 1 : evolution of a row between $z = 104\text{mm}$ and $z = 108\text{mm}$ (3.11a), and the refocused image at $z = 107\text{mm}$ (3.11b).

3.7 Conclusion

In this chapter, we have introduced a new approach to compute refocused images from plenoptic data acquired with optical setups based on a micro-lens array. Our first contribution is a transformation between different two-planes parameterizations of the rays flowing through the light field imaging system. Compared to the state of the art, this transformation is continuous according to the different parameters of the system, because it is independent from any relationship between the distances in the setup.

Our second contribution is the refocusing approach itself. Compared to previous works, it can be used to reconstruct images for an arbitrary number of pixels at any depth from the main lens of the system. Furthermore, it is fully generic and can be applied for any plenoptic images acquired with a system based on a micro-lens array. Finally, it provides accurate simulation of blur due to defocus with a good depth accuracy. The main limitation of the current implementation is its large computation time.

Our approach can also be inverted to simulate acquired plenoptic images. We used this feature to create synthetic data in order to demonstrate the previously mentioned quality of our approach. We also demonstrated that it works on real plenoptic images from the literature.

This algorithm will be used to reconstruct images acquired during the X-rays plenoptic experiment (see Chapter 5). Its accuracy and its ability to reproduce defocus blur are interesting properties that encourage to apply depth from focus techniques on the refocused stack (see Chapter 4).

Chapter 4

Experimenting depth extraction with depth from focus

In this chapter, we study the possibility to extract the depth of an object from a stack reconstructed by our refocusing algorithm, presented in the previous Chapter 3. In the context of X-ray plenoptic imaging, the acquired raw images are expected to present specific features and a low quality differentiating them from visible plenoptic images. In our case, due to the characteristics of the illuminating source, the sub-images have a low number of pixels and a low signal-to-noise ratio. This prevents from using multi-view stereo-imaging techniques to extract depth directly from the sub-images [DCH⁺19].

Tomography algorithms are an alternative for direct reconstruction from the raw images. They consist in back-propagating the information from the images to the object scene and optimizing the reconstruction using an iterative algorithm [KSW02; Buz11]. These methods have been experimented during the *VOXEL* project by Vigano et.al., on experimental images we acquired on our visible plenoptic setup [VDH⁺18; VMH⁺19].

Here we explore another approach. Our refocusing algorithm allows to accurately refocus the scene at different depths within a chosen interval. In this chapter, we explore depth from focus techniques to extract depth from the stack reconstructed with our algorithm.

4.1 Principle of depth from focus

The algorithm proposed in Chapter 3 is a refocusing algorithm. It outputs a stack of images focused at different depths. In photography or microscopy a stack can be physically acquired by either changing the distance between the scene and the camera, or by changing the focal distance of the optics [WH02]. In such a focal stack, each image contains the whole scene, but only the in-focus parts are sharp whereas the out-of-focus parts are blurred. The in-focus and out-of-focus parts depend on the depths of the structures in the scene and

the depth at which the image is refocused. Depth from focus techniques aim to exploit this sharp versus blur property to determine depth for each pixel in the scene.

Various methods of depth from focus were proposed in the literature [SDN04; SN17]. Some methods were designed for a single or a few number of images. In our case, the algorithm is capable of outputting a large number of images therefore we will focus on techniques designed for several images. The objective is to assign a single depth value z to each (x,y) pixel of the 2D images. This process creates a depth map, which is a 2D representation of the whole scene where depth is coded by colors, giving information on where the objects are located in the whole 3D scene.

Some commercial softwares for depth extraction are available, including for example *Enfused*, *Keyence* or *Phaseview*. We tested two free plugins on Image J: the *Stack Focuser* and the *Extended Depth of Field* [AVDVU08; FVDVB⁺04]. It was not possible to obtain the expected result as the plugins do not allow enough degrees of freedom to refine the result. Our numerically reconstructed stack has specific focused structures and blur, compared to focal stacks acquired by a microscope. For example the discrete nature of the acquired data (for both pixels and micro-lenses) is responsible for the grid artifacts when refocusing out of the plenoptic depth of field (see Section 3.6.2). Therefore, we studied and implemented a few methods of depth from focus to apply them on our refocused stack.

To create the depth map of the scene, the pipeline usually consists in two steps: blur measurement and depth extraction, as shown in Figure 4.1. Blur measurement consists in quantifying local blur (or sharpness) inside a 2D window around a pixel (blue square on the left of Fig. 4.1). The blur is extracted on the same window for each 2D plane of the refocused stack. Then the blur measures of the same pixel in each image of the stack are combined to form a curve along depth. Depth extraction compares the obtained 2D measures along the depth axis (middle in Fig. 4.1). The resulting curve allows to select the depth corresponding to the lowest blur (or highest sharpness). The output of depth from focus is represented as a depth map, representing the estimated depth for each pixel in the image by a color (right in Fig. 4.1). The depth map combined with the refocused stack allows to reassign each pixel to a depth and reconstruct the volume. Both blur measurement and depth extraction can be affected by noise or image quality. Consequently some filtering can be added after each step to smooth the results and avoid outliers. Otherwise, it is also possible to combine different methods in order to improve the robustness of the result.

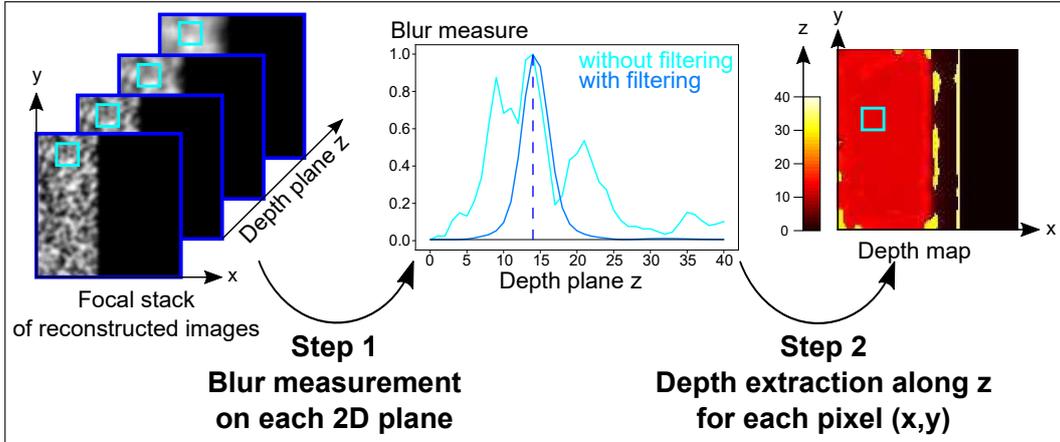


Figure 4.1 – Principle of depth from focus methods: 1) Blur is measured on each plane of the stack separately, then 2) using these measurements, depth is extracted for each (x,y) pixel. This process results in a depth map.

4.2 Chosen methods and implementation

In order to explore the pertinence of depth from focus on a stack of images reconstructed by our algorithm, we have tested a selection of blur measurements and depth extraction methods. Between this two steps, we apply a Gaussian filter to reduce the noise in the blur measurements before extracting depth.

4.2.1 Blur measures

The idea of blur measurement is to highlight contrasted parts of the image. That is why some blur measures are based on the same principles as edge detection algorithms [SDN04]. As for edge detection, a large variety of blur measures are available, based on the idea of either derivative, gradients, Laplacians or wavelets. Some approaches compute local statistics on small regions in the image, when others are based on calculations of the histogram [SDN04]. Following the work of Cou and Guennebaud [CG19], we implemented five of these methods to test them on our images.

The first method *Square Gradient* calculates the first derivatives in the 2D images to compute the squared norm of the gradient in an image I [SOdSV+97]:

$$BM_{Sq-Grad} = \left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 \quad (4.1)$$

The *Tenenbaum Gradient* uses the same idea with Sobel filters which operate on the image like derivatives [YOJS93; Kro88]. This allows to extend the contrast detection on neighboring pixels:

$$BM_{Ten-Grad} = (S_x * I)^2 + (S_y * I)^2 \quad (4.2)$$

$$\text{with } S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \text{ and } S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

However, the methods based on the first derivatives might be sensitive to noise. This is why we also implemented and tested methods based on second derivatives of the images. For example, the *Energy Laplacian* consists in squaring the second derivative components of the Energy Laplacian as defined in [SCN93]:

$$BM_{En-Lapl} = (EL * I)^2 \quad (4.3)$$

$$\text{with } EL = \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix}$$

The *Ring Difference Filter* is an alternative presented in [SJP+17]. It uses a larger convolution kernel that compares the current pixel to a larger number of neighbours that are selected two pixels away. Thus we expect it to be less sensitive to noise:

$$BM_{RDF} = (RDF * I)^2 \quad (4.4)$$

$$\text{with } RDF = \begin{bmatrix} 0 & -1 & -1 & -1 & 0 \\ -1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 12 & 0 & -1 \\ -1 & 0 & 0 & 0 & -1 \\ 0 & -1 & -1 & -1 & 0 \end{bmatrix}$$

Finally, we also implemented a statistical method. The *Variance* method is the variance of the image measured inside a window W of a chosen size. The difficulty here is to choose the appropriate window size that allows a representative measure (large enough) at a fine granularity (small enough). After performing few tests, we empirically chose the value of 5 pixels.

$$BM_{Var} = \sum_{x_i \in W} \sum_{y_j \in W} (I(x_i, y_j) - \overline{I_W})^2 \quad (4.5)$$

with $\overline{I_W}$ the average value of I inside W

4.2.2 Depth extraction

The depth extraction methods consist in selecting the best depth based on the blur measures that were evaluated along the depth axis. The presented blur measures are supposed to highlight the sharpest structures of the stack. The different elements of the volume are thus expected to present a maximum in this blur measures at the depth where they are in focus. Based on [CG19], we implemented the following three methods: *Max*, *Mean* and *Symmetry*. The *Max* method consists in taking the global maximum along the blur measure.

For each pixel, the curve of blur measure along depth is drawn and its maximal value is taken as the correct depth:

$$DE_{Max} = \underset{z}{\operatorname{argmax}}(BM(z)) \quad (4.6)$$

The *Mean* and *Symmetry* methods are based on more local aspects of the curves. The *Mean* consists in calculating local mean values over a sliding window along the depth axis. First, a confidence factor (*Conf*) allows to estimate which window contains the higher values of the blur measure. The highest confidence factor defines which window is likely to contain the correct depth. The calculated mean allows to refine the estimated depth inside this window. The used formulas are the followings:

$$Conf = \sum_{w \in W} \frac{BM(w)^2}{\sum_{w \in W} BM(w)} \quad (4.7)$$

$$Mean = \sum_{w \in W} \frac{w.BM(w)}{\sum_{w \in W} BM(w)} \quad (4.8)$$

$$DE_{Mean} = \underset{z}{\operatorname{Mean}}(\underset{z}{\operatorname{argmax}} Conf(z)) \quad (4.9)$$

The *Symmetry* method also uses the principle of sliding window to extract local property [CG19]. This time, the objective is to quantify the assumed symmetry of the blur due to the action of refocusing. This method is based on the assumption of a local symmetry around a focused structure: the sharp details are surrounded by progressive blur. The blur measure is multiplied by its symmetry centered in z , and the result is summed inside the window W :

$$DE_{Symm} = \underset{z}{\operatorname{argmax}} \left(\sum_{w \in W} BM(z-w).BM(z+w) \right) \quad (4.10)$$

After the blur measure and before depth extraction, the optional filtering consists in a Gaussian filter along the x and y axis only. The choice of the standard deviation is difficult. As for the *Variance* method, the appropriate value should be determined. A large value removes the noise efficiently, but with the drawback of losing the useful contrast of small structures located at different planes. For our tests we used a standard deviation of 3.

Similarly, it is also possible to add a z filtering after the depth extraction step. This is a filtering along the depth axis that allows to smoothen the data in order to have a more continuous result. This filtering is interesting in the case of surfaces, because they are usually expected to be continuous in the real world. However, as our aim is to work on volumes, this second filtering is not interesting in our application.

4.3 Tests on synthetic images

These methods are tested to determine the abilities of existing algorithms to extract depth from plenoptic refocused stacks. The test case should be a plenoptic stack representative from a realistic plenoptic configuration. Using configuration 1 in Table 3.1, synthetic data is created by our simulation algorithm, resulting in a plenoptic raw image. The plenoptic image is then used as input for our refocusing algorithm to generate a stack of images refocused at regular depths. With this refocused stack, the blur measures, the influence of Gaussian filtering and the depth extraction methods can be alternatively tested.

The synthetic object should be simple, so that the result could be easily analysed. We thus simulated 2D planes, composed of salt and pepper noise. This noise creates highly contrasted structures over the planes, that helps depth extraction methods in finding the correct depths.

In Figure 4.2 are presented the three different planes that will be used in the following tests. The size of the whole volume is $108 \times 108 \times 41$ pixels and voxel size corresponds to $3 \times 3 \times 10 \mu\text{m}^3$ in object-space. In comparison, the lateral and depth resolutions equal $9.3 \times 9.3 \mu\text{m}^2$ and $32.8 \mu\text{m}$ in configuration 1. The Left plane (green) is positioned at depth $z = 105.64 \text{mm}$ which is the reference image plane of the chosen configuration. It corresponds to depth plane number 14 in the volume. The Right plane (blue) is placed at depth $z = 105.80 \text{mm}$ or depth plane number 30, far enough so that the two planes should not interfere in their detections. The Third plane (purple) is placed

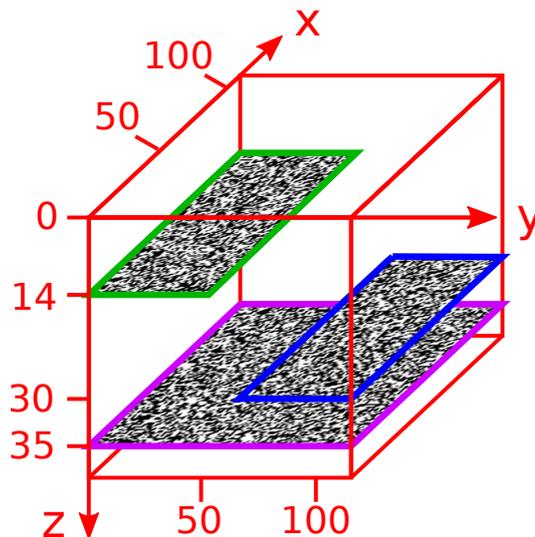


Figure 4.2 – Relative positions of the synthetic depth planes: the Left plane (green) corresponds to plane number 14 ($z = 14$), the Right plane to $z = 30$ and the Third plane to $z = 35$. The size of a pixel is $3 \times 3 \mu\text{m}$ inside a plane, and the spacing between two planes is $10 \mu\text{m}$.

at depth $z = 105.85\text{mm}$ or depth plane number 35. It is placed at the limit of DOF_{pleno} , which is 0.52mm for this configuration. As it is 0.05mm away from the Right plane, it is still expected to be detected separately from it, as res_{depth} is calculated to be 0.033mm . The working interval is chosen between depth $z = 105.50\text{mm}$ and $z = 105.90\text{mm}$ roughly corresponding to DOF_{pleno} . Beyond these limits the algorithm generates some grid artifacts according to Section 3.6.2.

In the rest of the chapter, the locations inside this volume will be given as pixel number instead of real distances in mm . The planes will be studied either alone (see Subsection 4.3.1) or together (in Subsections 4.3.2, 4.3.3).

4.3.1 Detection of a single plane

First, the pipeline of depth from focus is tested on the simple case of a single plane simulated and then refocused. This allows to confirm the possibility to build a depth map from a plenoptic refocused stack and to study its accuracy depending on the methods used, both in step 1 (blur measure) or step 2 (depth extraction). Figure 4.3 presents the process of simulation and reconstruction on the Left plane alone (in green in Fig. 4.2). From the original Left plane placed at depth $z = 105.64\text{mm}$ (or depth plane 14) (see Fig. 4.3a), the plenoptic raw image is simulated (Fig. 4.3b) and then used to reconstruct the image at the same depth (see Fig. 4.3c). The image of the original plane 4.3a shows highly contrasted structures whereas they are blurred in the reconstructed image 4.3c.

For both the original and reconstructed stacks, we extract the same blue row in each image of the stack and display them as a function of depth (Figs. 4.3d and 4.3e). In the original volume 4.3d, the plane occupies only a single depth at $z = 14$ and there is no signal on the other planes. Notice that the plane appears as a dotted line, due to the use of the salt and pepper noise, that generates 0 (black) or 1 (white) values. The refocused stack 4.3e presents signal over the whole depth interval (z -axis), due to the blur propagating along depth. We observe high variations of colors (i.e. of contrast) at depth $z = 14$, which follow the same structures as the dotted line in Fig. 4.3d. The objective of depth from focus methods is to separate such peaks of contrast from the blur of adjacent planes.

4.3.1.1 Blur measures

This dataset is used to test the blur measurement methods. The reconstructed focal of the Left plane alone is taken as input, and the position of the original Left plane corresponds to the expected output. Fig. 4.4 shows the results of the five blur measures presented in Section 4.2 applied on the Left plane alone. The results are presented as in Figs. 4.3d and 4.3e, to allow comparison with the original and refocused volumes. Blur measurement meets the expectations

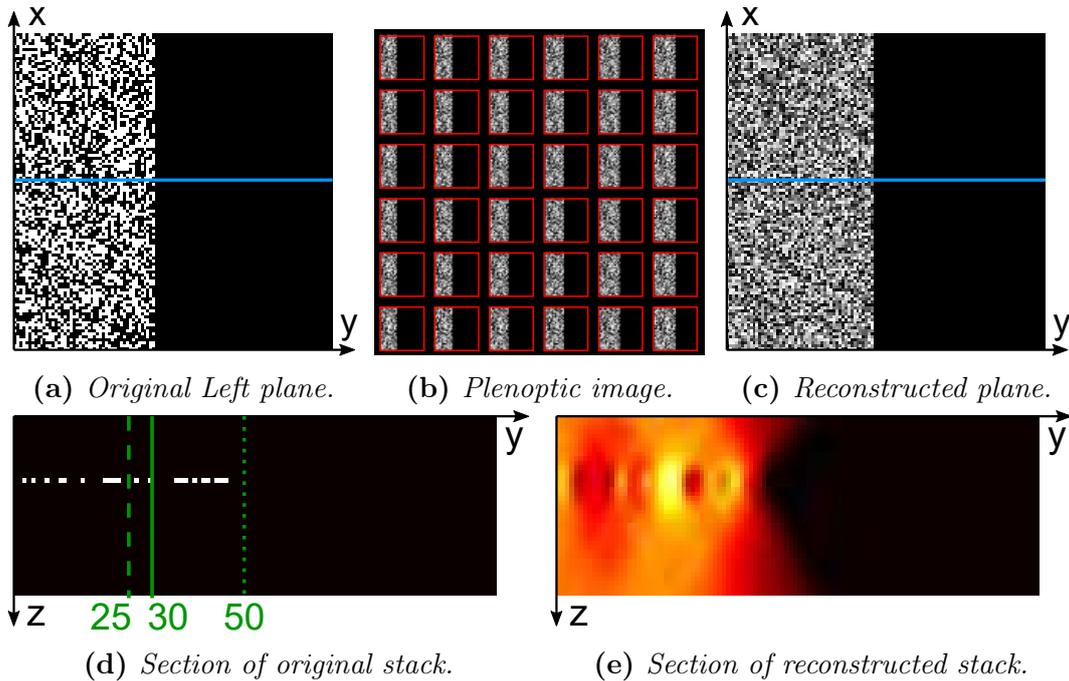


Figure 4.3 – Original, plenoptic and reconstructed images of the Left plane, due to the simulation and refocusing algorithms. The blue lines in Figure 4.3a and 4.3c show the row corresponding to the longitudinal sections in Figure 4.3d and 4.3e. The green lines in Figure 4.3d represent the positions of the profiles of Fig. 4.5 and 4.6.

of highlighting contrast at depth $z = 14$, while ignoring blur propagation to adjacent planes. However, the response is different according to the blur measure used.

With first order derivative methods (*Square Gradient* and *Tenenbaum Gradient*), the detected contrast is positioned on but also around the right plane location (see Figs. 4.4a and 4.4b). It generates a halo around the structures of the plane, especially on the extreme left and right borders of the plane. It means the blur measures detect the border of the image and the border of the plane. It is the principle of edge detection that inspired these methods but it is not our goal here.

The second order derivative based methods are more accurate (*Energy Laplacian* and *Ring Difference Filter*). Both detect contrast located on the good depth plane, without leaking on the neighbour depths (see Figs. 4.4c and 4.4d). Figure 4.4f shows the action of Gaussian filtering applied after the *Ring Difference Filter*. The action of filtering smoothens the data, and highlights a clear and continuous plane.

Finally, the *Variance* method is the only statistical method we tested. Its result is similar to the ones from *Square Gradient* and *Tenenbaum Gradient* in the sense that the heat regions look blurred inside the volume. Yet, there are several differences. In the *Variance* result the left and right borders of the plane

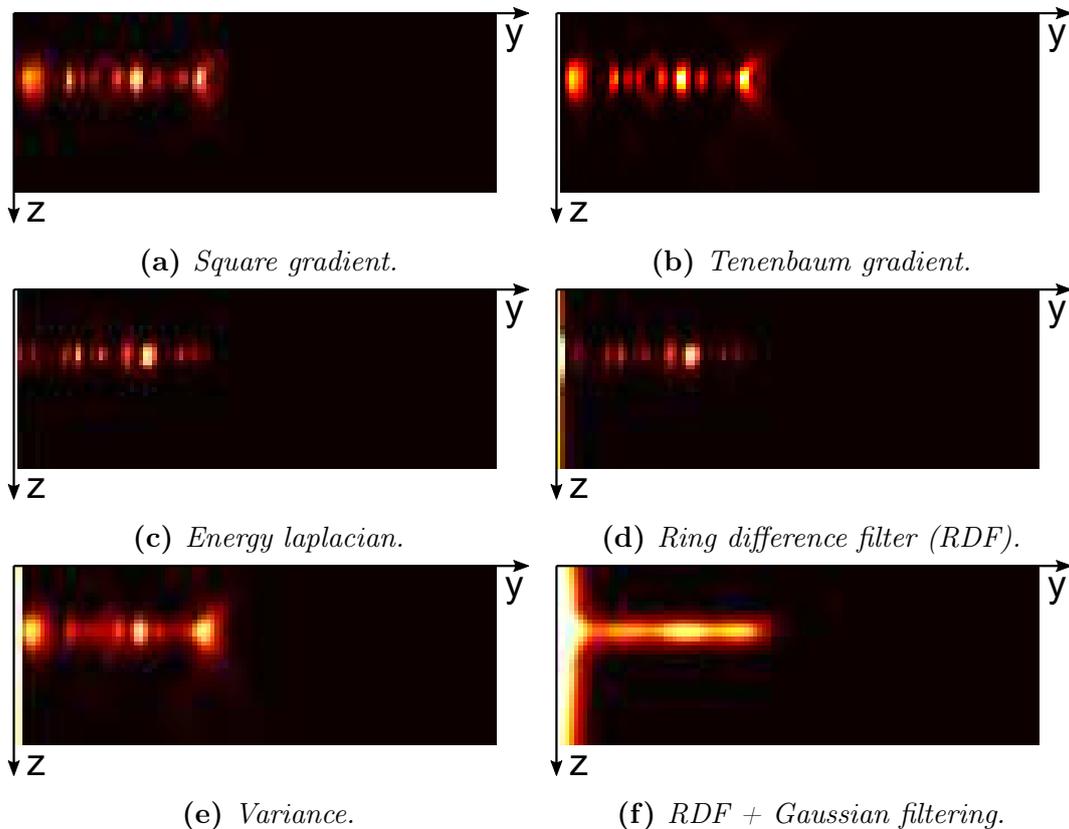


Figure 4.4 – Longitudinal sections of blur measures applied on the reconstructed stack of the Left plane.

are even more highlighted, which is again not a desired effect. They present large leaks on the neighbouring depths. A positive effect is that the structures inside the plane are more detected. Indeed, this blur measure is a statistical measure which has to be defined on a given window. Consequently, the measure itself already contains a x-y filtering effect, that extends the detected spots to the whole depth plane. The measure is efficient to detect such a depth plane, but would not in the case of more diffuse data.

To better study the action of each blur measure, we draw profiles along depth for different locations (see Fig. 4.5). Profiles at $y = 25$ and $y = 30$ correspond to two positions inside the Left plane whereas $y = 50$ is outside of the Left plane, as illustrated in Fig. 4.3d. This can be seen in the profiles from the original stack where a peak is observed at $z = 14$ for $y = 30$, whereas the two other positions correspond to flat curves (Fig. 4.5a). In the case of the refocused stack, the two profiles for $y = 25$ and $y = 30$ present mirror behaviours, with either a maximum or a minimum at $z = 14$ (Fig. 4.5b). They both have a baseline around 150, corresponding to the blur when moving away from the plane $z = 14$. The extremum corresponds to a mix between the reconstructed signal from a specific location and its neighbours, that results

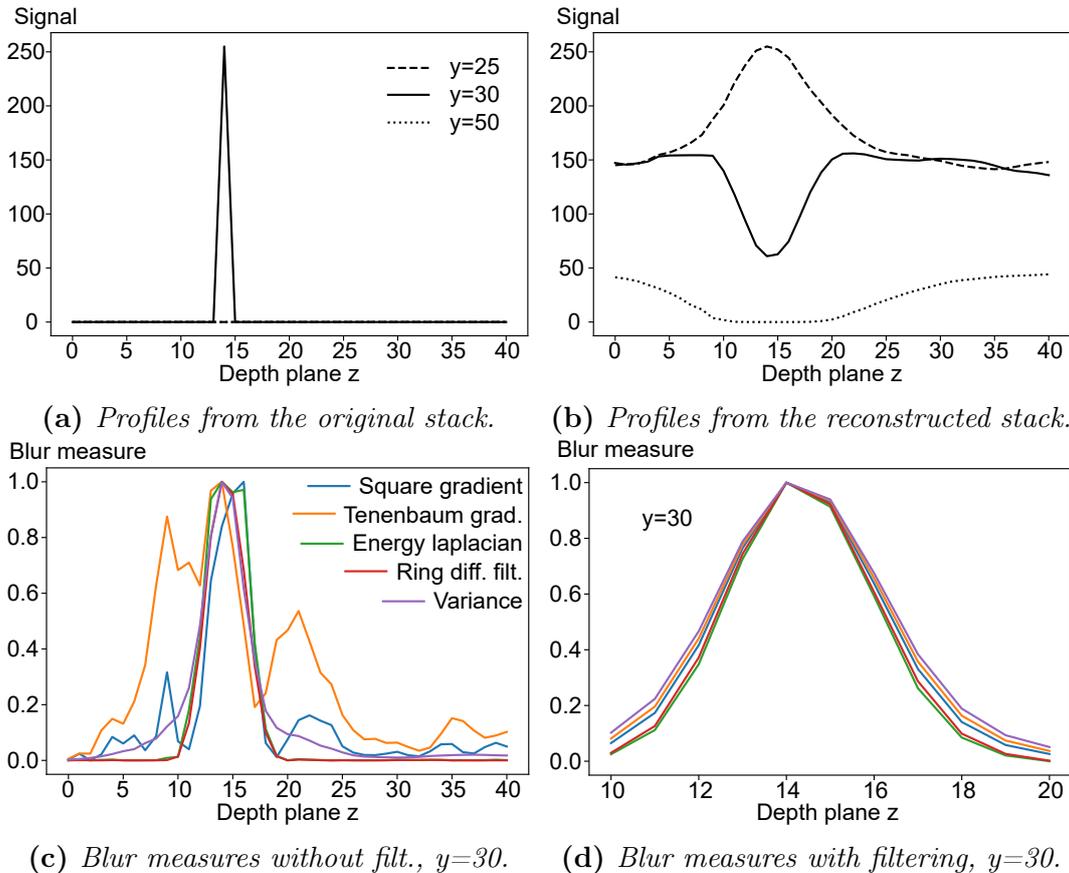


Figure 4.5 – Comparison of depth profiles from the original and reconstructed stacks at 3 different locations defined in Figure 4.3d, as well as blur measures with or without Gaussian filtering for the location $y=30$. Each curve was independently normalized to have a maximal value at 1. In Figure 4.5d the scale of depth plane z has been enlarged to better distinguish the different curves.

either in an upper or a lower value compared to its reconstructed neighbours values. The important point is that this singularity occurs at the correct depth $z = 14$. On the contrary, the behaviour of the profile for $y = 50$ is smooth. It corresponds to a location outside of the Left plane in the original image, and unsurprisingly the reconstructed signal is quite low. It reaches its minimal value around $z = 14$, but on both sides the curve increases due to the blur coming from the reconstructed Left plane. This is an undesirable effect that could possibly cause detection of ghost structures where there should be none.

Figure 4.5c compares the profiles extracted at $y = 30$ for each blur measure. This location was chosen because the resulting curves are representative from the behaviours of the five methods in the whole volume. The five curves all show maximal values around $z = 14$. Note that the curves were all normalized. The *Square Gradient* and especially the *Tenenbaum Gradient* result in noisy curves. This corresponds to contrast detected in adjacent planes as already

noticed in Figure 4.4a and 4.4b. The profile due to the *Tenenbaum Gradient* is so noisy that a high peak occurs at $z = 9$ that could be detected as the maximal contrast instead of the correct $z = 14$. The three other methods show clear peaks without noise. The profile of the *Variance* method is a bit broader at its base. This corresponds to a blur effect because the *Variance* is calculated over a window which makes this measure intrinsically less local and more extended than the other ones. The *Energy Laplacian* presents a small plateau at the top at the curve, which might be a problem for the precision of depth extraction. Finally, the *Ring Difference Filter* method presents a thin peak without any noise or plateau. For this location as for others randomly chosen across the whole plane, this method seems the most appropriate blur measure.

Figure 4.5d presents the result of the Gaussian filtering applied on each blur measure. Again the curves were all normalized. The corresponding curves are very similar, and the depth z axis has been enlarged to be able to distinguish the five curves. The action of lateral filtering is very strong because it removes the noise and plateau. This is because the denoising is done in the x - y plane, which improves the signal when plotted along z . Moreover, the size of the blur is important compared to the size of the original noise (standard deviation of 3 compared to fluctuations of 1).

The *Ring Difference Filter* method seems to be the best blur measure method. This method will be kept as the best blur measurement method, and used in the rest of the chapter to test the depth extraction methods.

4.3.1.2 Blur measures - comparison of behaviours inside the volume

A good blur measure should be able to highlight contrast in a volume, while having a consistent behaviour between different locations. In order to confirm the precedent conclusion, we compare the three methods *Energy Laplacian*, *Ring Difference Filter* and *Variance* at different y in the refocused stack, displaying structures with different characteristics. Figure 4.6 presents detailed profiles comparing blur measures applied on locations used in Figures 4.3d: $y = 25$, $y = 30$ and $y = 50$. Normalization is done for each blur measures (with or without filtering). Profiles $y = 25$ and $y = 30$ were both taken from inside the Left plane, so they are expected to detect high contrast at depth $z = 14$. This is the case for the three methods, but the profiles for $y = 25$ presents lower values than for $y = 30$. The *Energy Laplacian* method shows plateaus for several profiles, that might be a problem for precise depth localization. After filtering, the two profiles for the three methods have nearly the same height, which shows that filtering has an positive action of homogenizing the measured blur through the plane. However, *Energy Laplacian* shows some artifacts around $z = 7$, which might cause error in depth detection. This is not the case with the two other methods.

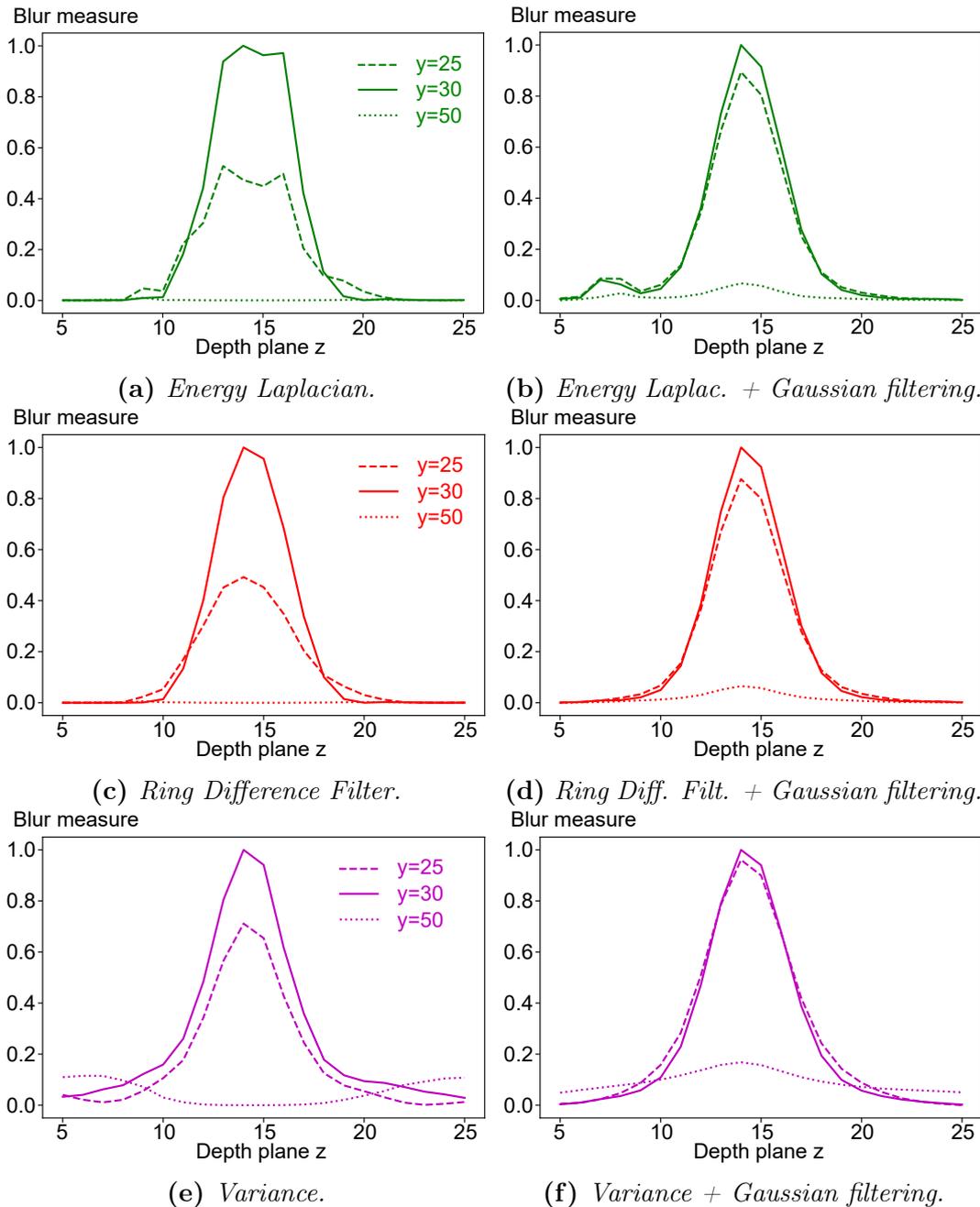


Figure 4.6 – Comparison of the action of the *Energy Laplacian*, the *Ring Difference Filter* and the *Variance* on three different profiles: $y=25$ (dashed), $y=30$ (solid) and $y=50$ (dotted).

The behaviour of the profile at $y = 50$ is also interesting (in Fig. 4.6). This location was taken outside of the Left plane, but Figure 4.5b has shown that blur from the Left plane raises the reconstructed values for depth planes far from $z = 14$. The *Variance method* conserves this trend for z below 10 or over

20. On the contrary, the *Energy Laplacian* and *Ring Difference Filter* result in very flat curves. After the Gaussian filtering, the three methods present a light peak due to this blur propagation. The peak is slightly more important in the *Variance + filtering* than in the two others, which might causes incorrect depth detection.

This new comparison confirms the choice of the *Ring Difference Filter* followed by Gaussian filtering as the best blur measure for this plenoptic dataset. This method will be used as input for testing depth extraction, that is the second step of the pipeline of depth from focus.

4.3.1.3 Depth extraction

The result of the *Ring Difference Filter* followed by the Gaussian filtering is used as input to test the depth extraction methods. Figure 4.7 presents the depth maps generated by the *Max*, *Mean* and *Symmetry* methods without (top row) and with (bottom row) the filtering step, compared to the expected depth map.

The expected depth map (in Fig. 4.7a) is composed of a red rectangle on the left, corresponding to the plane $z = 14$, and a black rectangle corresponding to the background. The depth maps generated by the different methods globally show the same pattern, with red on the left and black on the right. The depth value on the left corresponds to the correct depth of the Left plane. In the middle of the image, the lighter colors correspond to the edge of the plane that is detected at extreme depths, due to blur propagation of the Left plane through the volume, as illustrated by the profile measured at $y = 50$ in Fig. 4.5c.

Figure 4.7i shows the three profiles of the *Max* (blue), *Mean* (purple) and *Symmetry* (green) methods taken at the same location (shown in green in Fig. 4.7a). Figure 4.7j shows the same profiles but with the Gaussian filtering. In both figures, the three curves globally follow the expected curve (in dotted black) with the Left plane detected at depth $z = 14$ and the background at $z = 0$. As shown in Fig. 4.7j, filtering clearly denoises the signal.

In the middle of the profiles, the large oscillations correspond to artifacts between the Left plane and the background. With filtering, the artifacts have become large plateaus for the *Max* and *Symmetry* methods. These plateaus could be misinterpreted as a third plane present in the scene, which might be a problem for plane detection. This effect is less important with the *Mean* method, for which the artifacts remain two thin peaks when the filtering is applied. These smaller artifacts could be more easily detected and removed in a post-processing step. These profiles could be related to the depth maps, where the lighter colors correspond to the high values of the artifacts.

Notice also that the three curves seem to extend the Left plane, as their plateaus at $z = 14$ extend until $y = 54$ instead of the correct value of $y = 48$ for

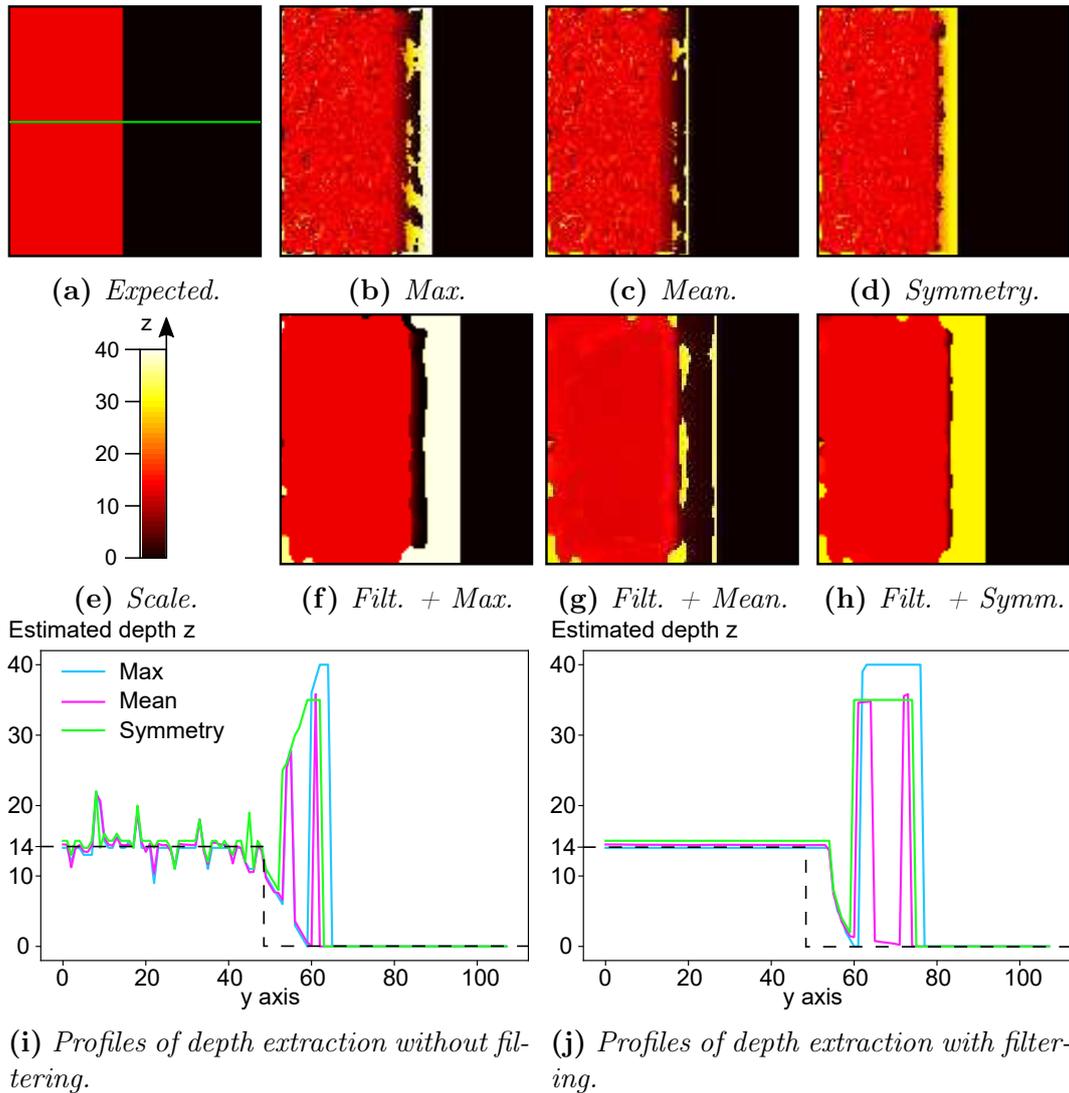


Figure 4.7 – Depth extraction methods tested on the Left plane alone.

the original Left plane. This can also create an imprecision in the extraction of the correct depth.

Even if the results are similar, the *Mean* method with filtering provides the best results in the case of a single plane. In the next sections will be presented what happened when several planes are positioned together in the same volume.

4.3.2 Detection of two separated planes

In this section, the methods are tested when applied in the case of two planes positioned at two different depths in the volume. The Left plane is still placed at depth $z = 14$ (or $z = 105.64\text{mm}$) and the Right plane is at depth $z = 30$

(or $z = 105.80\text{mm}$) as was shown in Figure 4.2.

Our simulation algorithm allows to generate plenoptic images for planar objects located at a unique depth. Thus, for this configuration with two planes, we generate two plenoptic images and combine them afterwards. The combination is done following the principle of alpha blending [PD84], which consists in the multiplication of contributing factors to simulate transparency of the traversed objects.

To combine the two plenoptic images, we multiply the inverse of the individual plenoptic image of each plane. This simulates the action of cumulative absorption when light rays go through several successive planes. We then take the inverse of the multiplication result and use it as input for the refocusing algorithm to generate the refocused stack.

4.3.2.1 Blur measures

The blur measures were studied on the corresponding reconstructed stack. Figure 4.8 shows the positions of the two planes in the original volume and the refocused stack, and the result after the *Ring Difference Filter* with and without Gaussian filtering. Only this blur measure is presented here, as it was the one selected in section 4.3.1.

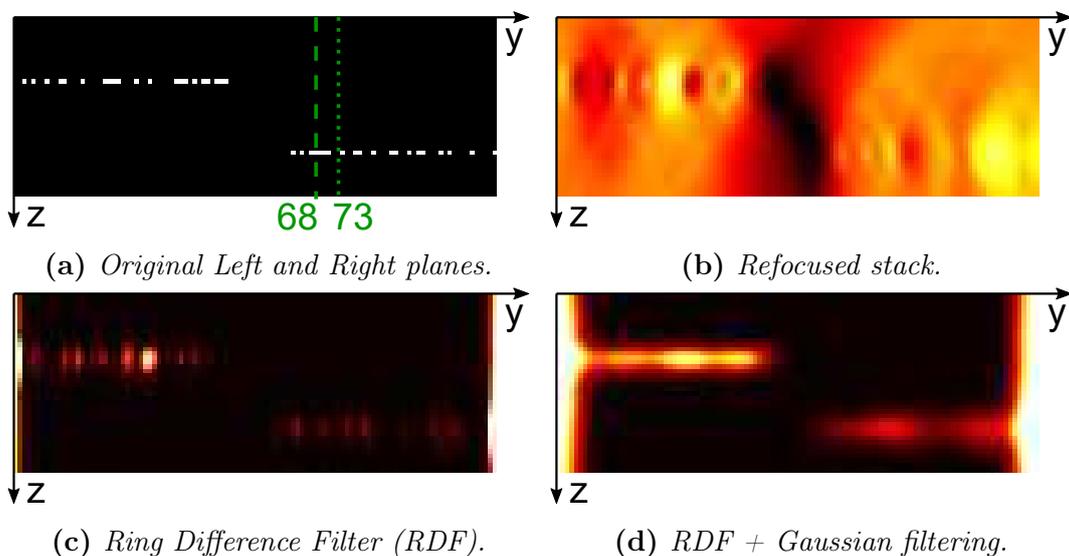


Figure 4.8 – Original and reconstructed stack in the case of two separated planes, and the results of the *Ring Difference Filter* with and without filtering. The green lines in Figure 4.8a represent the positions of the profiles of Fig. 4.9.

The refocused stack shows the blurring effect due to the refocusing process, this time with two planes. As the Left and Right planes were positioned on two different parts of the volumes, their propagating blurs do not overlap. The *Ring Difference Filter* is then applied with (Fig. 4.8c) and without filtering

(Fig. 4.8d). In both cases, the two planes are well detected, with contrasted structured detected by the *Ring Difference Filter* at the correct depths of the planes with little detection errors on the surrounding depths. In Fig. 4.8d, the filtering allows to homogenize the two planes, while increasing the response of the blur measure.

However, we notice a difference of colors between the Left and Right planes, which represents a difference of blur quantified by the measure. This is not due to the blur measure itself but to the refocused stack. In the stack we can see by eyes that the Left plane is refocused with a higher contrast than the Right plane. This is because the two planes were placed at two different depths in the volume and can be explained by the evolution of contrast with depth along the plenoptic Depth of Field, as was explained in Chapter 2. As a consequence, the two planes are not reconstructed with the same contrast, as measured by the *Ring Difference Filter*. Here, the contrast variations have concrete consequences on depth detection.

In Figure 4.9 the actions of the *Ring Difference Filter* and the other blur measures are studied more in details at two different positions in the volume. This time the positions are selected inside the Right plane, at $y = 68$ and $y = 73$. The Right plane corresponds to the peak at $z = 30$ in Figure 4.9a, which is the depth at which the Right plane was virtually placed. In the reconstructed stack (in Fig. 4.9b) the curves are more difficult to interpret than in the previous case of the Left plane alone: there is a light peak at $z = 30$ but it is not as frank as in Fig. 4.5b. This is due to the low contrasted signal of the reconstructed Right plane that does not clearly emerge from its reconstructed blur, due to the varying contrast through the depth planes.

Blur measures applied along $y = 68$ result in noisy curves (in Fig. 4.9c) while having a clear peak at $z = 30$. This noise reflects the lack of contrast of the reconstructed Right plane. The *Square Gradient* is particularly noisy, followed by the *Tenenbaum Gradient* and *Variance*, whereas the *Energy Laplacian* and *Ring Difference Filter* show very few noise. As for the single Left plane, the addition of the Gaussian filtering clearly improves the result (in Fig. 4.9d). The five curves are now overlapping at the peak at $z = 30$, but their baseline signals reflect the noise that has been removed by the filtering.

Concerning the profiles at $y = 73$ (Figs. 4.9e and 4.9f), the five blur measures exhibit clear peaks at $z = 30$ with very low noise. However, the peaks of the methods *Energy Laplacian* and *Ring Difference Filter* are centered on $z = 29$ and $z = 28$ instead of the correct value $z = 30$. The Gaussian filtering allows to recenter the peaks at the correct location.

One again these examples show the importance of filtering for noise removal and peak location. The *Ring Difference Filter* stays an efficient blur measure, although less obvious in this more complex scene. It will be the one used to test the depth extraction methods.

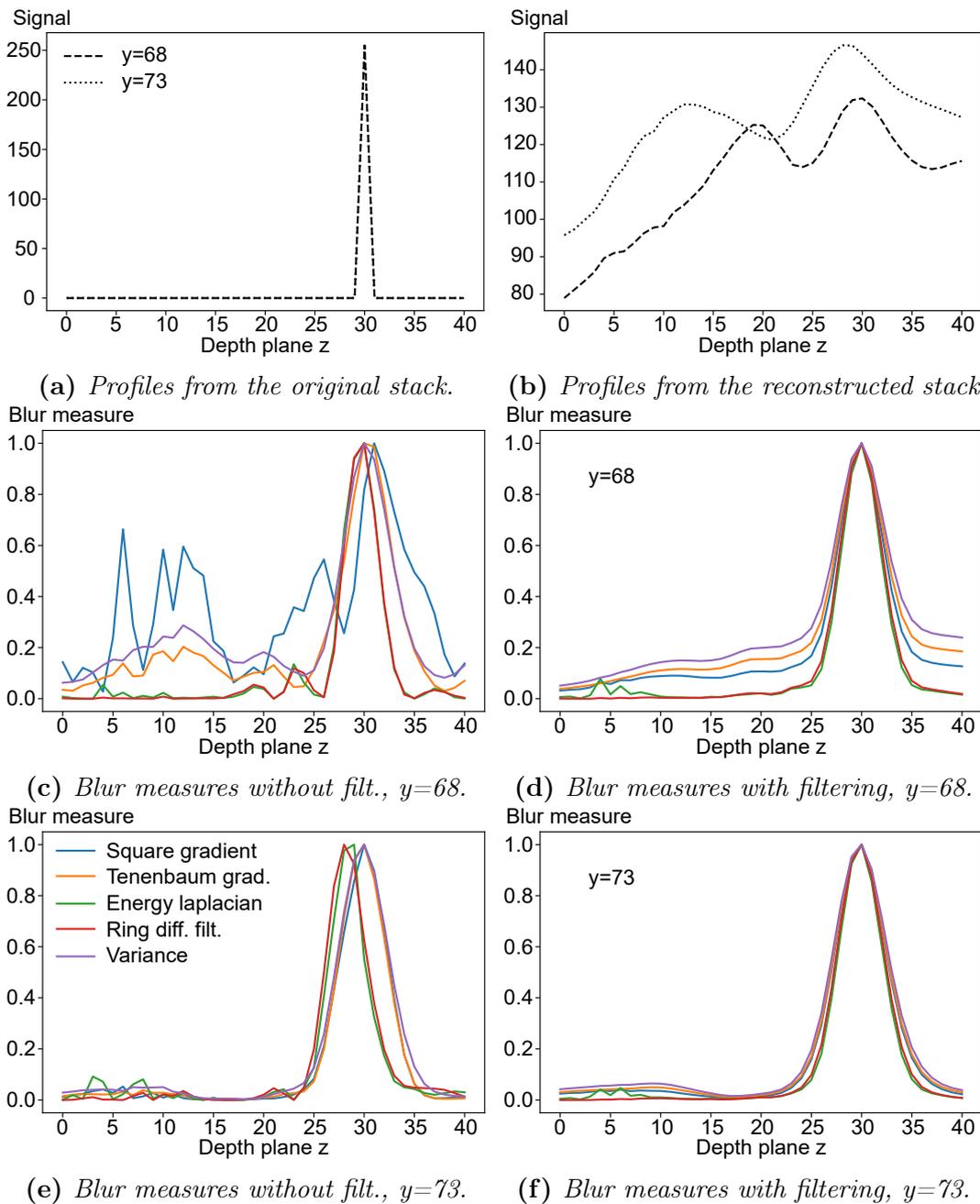


Figure 4.9 – Profiles for the case with both Left and Right planes: comparison of the same locations $y = 68$ and $y = 73$ taken in the original and reconstructed stacks, and the actions of the five blur measures with and without Gaussian filtering.

4.3.2.2 Depth extraction

The depth extraction methods were tested on this two-planes cases and the corresponding depth maps are presented in Fig. 4.10. In this case, the expected depth map consists in a red part on the left (depth $z = 14$), a yellow part on

the right (depth $z = 30$), with a small black line in the center, corresponding to the background (depth $z = 0$). As for the single plane, the depth extraction methods are able to detect the presence of the two planes and to retrieve their correct depths. Without the filtering step, the depth maps are noisy, but able to detect the gap between the planes, where background is expected. With filtering, the two planes are more homogeneous but the small gap has completely disappeared. Note also the artifacts at the borders of the depth maps. The filtering extends the size of the artifacts created by the blur measures.

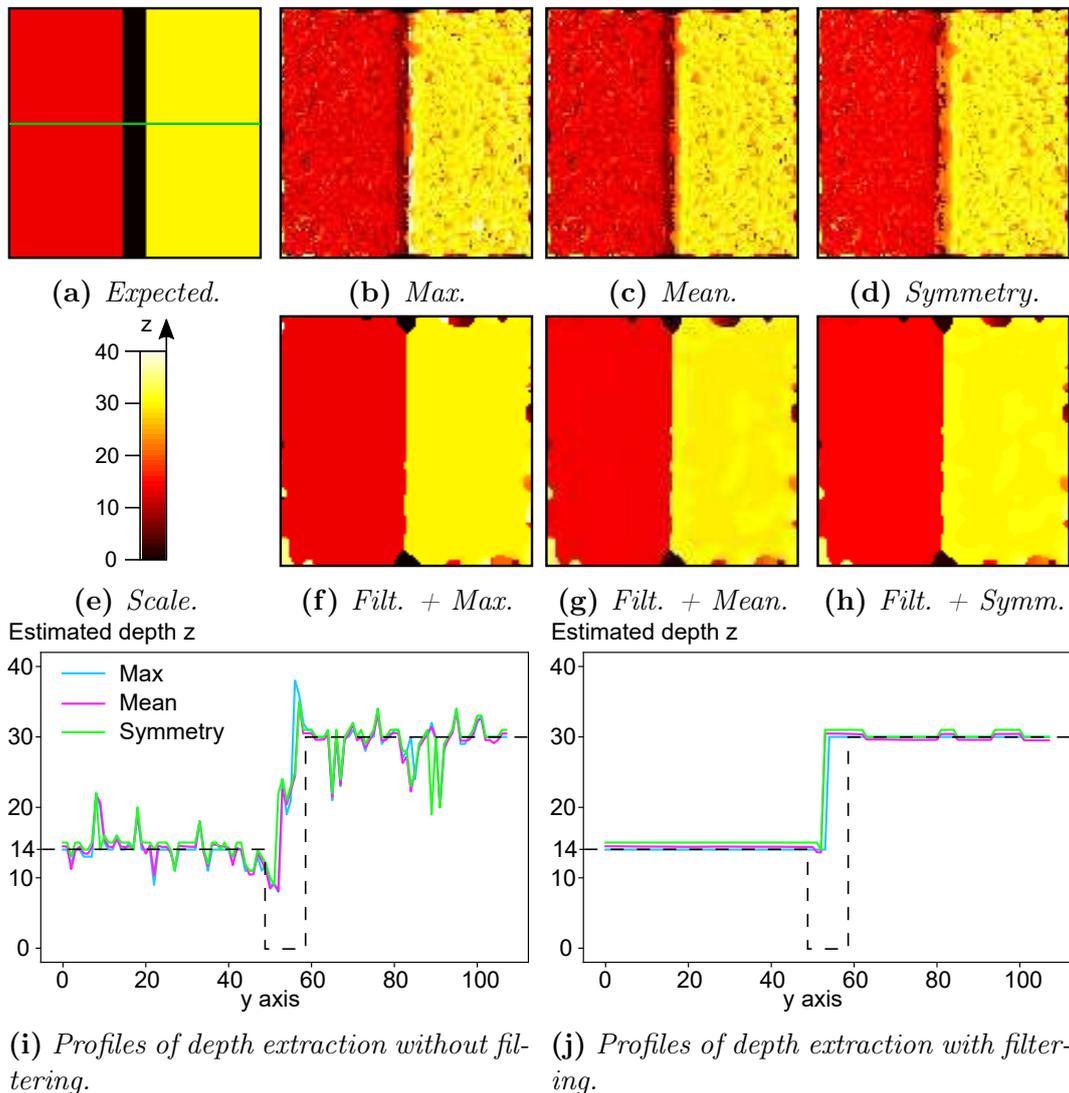


Figure 4.10 – Depth extraction methods tested on the case of the Left and Right planes together.

Figures 4.10i and 4.10j show profiles extracted from these depth extraction methods at the same location (in green in Fig. 4.10a). The results without filtering are noisy but globally follow the expected profile. The Left and Right

planes are well detected, but the methods fail at detecting the background between the two: the signal forms a smooth transition between the two planes. There is a small decrease around $y = 50$ but it does not reach the value $z = 0$ of the background. It could be interpreted as noise, preventing the detection of the gap between the two planes.

After the addition of the Gaussian filtering, the Left and Right planes become completely flat, with a sharp transition between them (Fig. 4.10j). This is due to the filtering being performed on window, that extends the planes more than expected, as was already shown in Subsection 4.3.1. The gap between the two planes has completely disappeared. In the case of two separated planes, the Gaussian filtering can bring benefits or drawbacks depending on what effect should be highlighted, either the continuity of the planes or the small structures at different distant depths. The impact of the Gaussian filtering can be modified through the change of its standard deviation. However, this effect should be adapted carefully as it depends on the elements in the volume and the desired result. There is no real difference between the outputs of the three methods *Max*, *Mean* and *Symmetry*.

4.3.3 Detection of three overlapping planes

4.3.3.1 Blur measures

Finally, the methods are run on a case with three planes located at different depths. The interest is to have planes overlapping along the x-y axis and to study whether the methods are still able to detect and locate them. For this case the Left, Right and Third planes are first used separately to generate the corresponding plenoptic images. Then they are combined together and refocused as explained in Section 4.3.2. The planes were placed at depths $z = 14$, $z = 30$ and $z = 35$ as shown in Figure 4.2. This corresponds to a variety of depth locations and depth differences, as we know that it impacts the contrast and quality of the refocused planes.

In Figure 4.11 we show the initial location of the three planes in the original image. The Right and Third planes are close but still separated by more than the res_{depth} (distance of $0.05mm$ compared to $res_{depth} = 0.033mm$). This means that the corresponding plenoptic image contains barely enough information to reconstruct them separately. In the refocused stack, these two planes look mixed together and it is difficult to distinguish one from the other. Even on the left part of the stack, the blurs due to the refocusing of both Left and Third planes are added together and decrease the contrast of these two planes. However, the *Ring Difference Filter* is still able to detect contrast coming from the three planes. The method highlight structures inside the Right and Third planes separately at their correct depths, even if they were placed very close. The Gaussian filtering step seems to blur the two planes together.

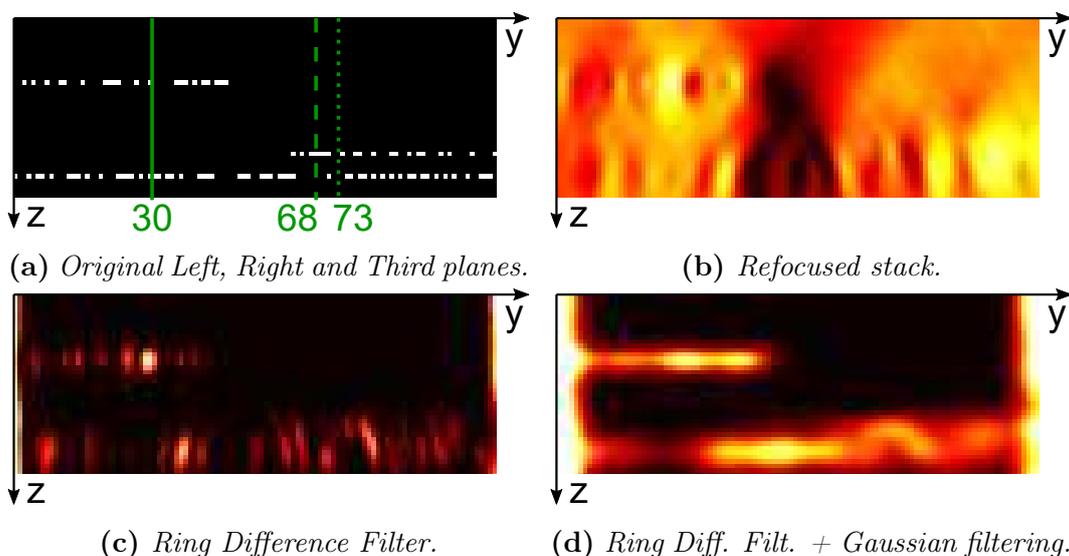


Figure 4.11 – Original and reconstructed stack in the case of three overlapping planes, and the results of the Ring Difference with and without filtering. The green lines in Figure 4.11a represent the positions of the profiles of Fig. 4.12.

We draw the profiles in Figure 4.12 corresponding to $y = 30$, $y = 68$ and $y = 73$ (see Fig. 4.11a). With the Third plane, the profile taken at $y = 30$ goes through both Left and Third planes as shown in Figure 4.12a. Both profiles for $y = 68$ and $y = 73$ go across the Right and the Third planes. The profiles from the reconstructed stack are chaotic (see Fig. 4.12b) because the propagated blurs due to refocusing from the three different planes influence each other. The Left plane at $z = 14$ can be identified by the downward peak of the $y = 30$ curve, but it is not possible to separate the two other planes from these profiles. Note also the signal value that is different from the reconstructed signal in the case of the Left plane alone in Fig. 4.5b. The signal is higher but with a lower SNR (between 127 and 180 instead of between 61 and 154) due to the reconstructed signal from the different planes adding together.

The blur measures create noisy curves for both positions $y = 30$ and $y = 68$. Depending on the method, the two expected peaks can be correctly detected, whether Left and Third for $y = 30$ or Right and Third for $y = 68$. The *Energy Laplacian* and *Ring Difference Filter* follow the same behaviour: contrary to other methods, they miss the Third plane for $y = 30$, but they are able to detect both Right and Third planes for $y = 68$. This again shows that these two methods have close behaviours.

The Gaussian filtering smooths the behaviours of the curves. For $y = 30$, the two peaks of the Left and Third planes are detected. Again there is a difference in the amplitude of the detected contrast because of their depth positions. Regarding $y = 68$, the two peaks for the Right and Third planes are now merged together and it is not possible to distinguish the two planes

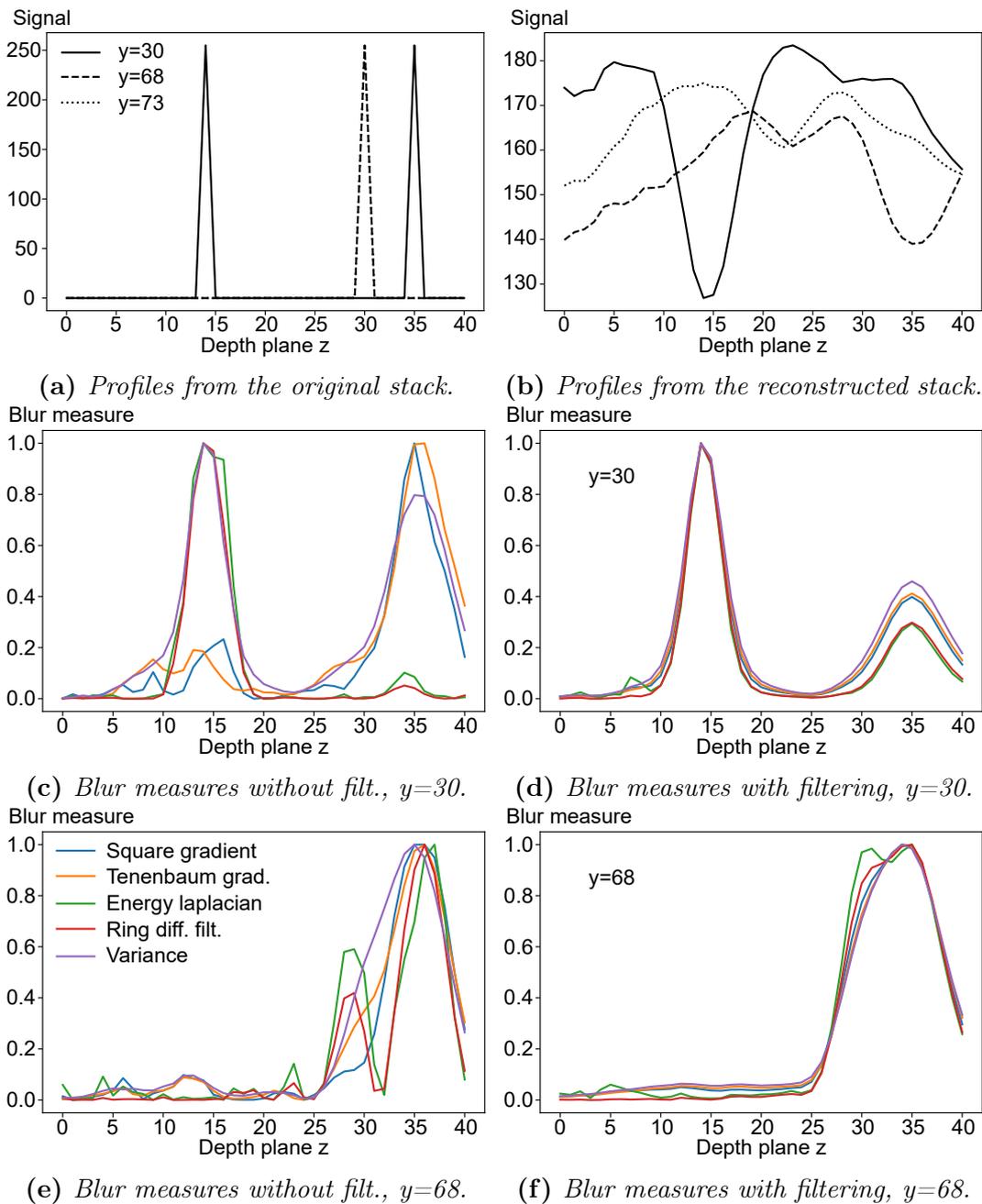


Figure 4.12 – Profiles for the case with three overlapping planes: comparison of the same locations $y=30$ and $y=68$ taken in the original and reconstructed stacks, and the results of the five blur measures with and without Gaussian filtering.

on this single position.

This case confirms the interest of the *Ring Difference Filter*, but questions the use of filtering for detecting planes close to each others.

4.3.3.2 Depth extraction

As for the previous cases, the depth extraction methods are applied on the result from the *Ring Difference Filter* followed or not by a Gaussian filtering step. In Figure 4.13 we present the result of depth extraction beside the expected depth map. Contrary to previous cases where we wanted to detect planes on separated parts of the image, this time the objective is to detect the presence of overlapping planes, while being able to locate them in depth. The expected depth map (Fig. 4.13a) was built by reproducing the noise of the original planes and assigning the correct color depths to each of them. The Third plane (off-white) is visible through the Left (red) and Right (yellow)

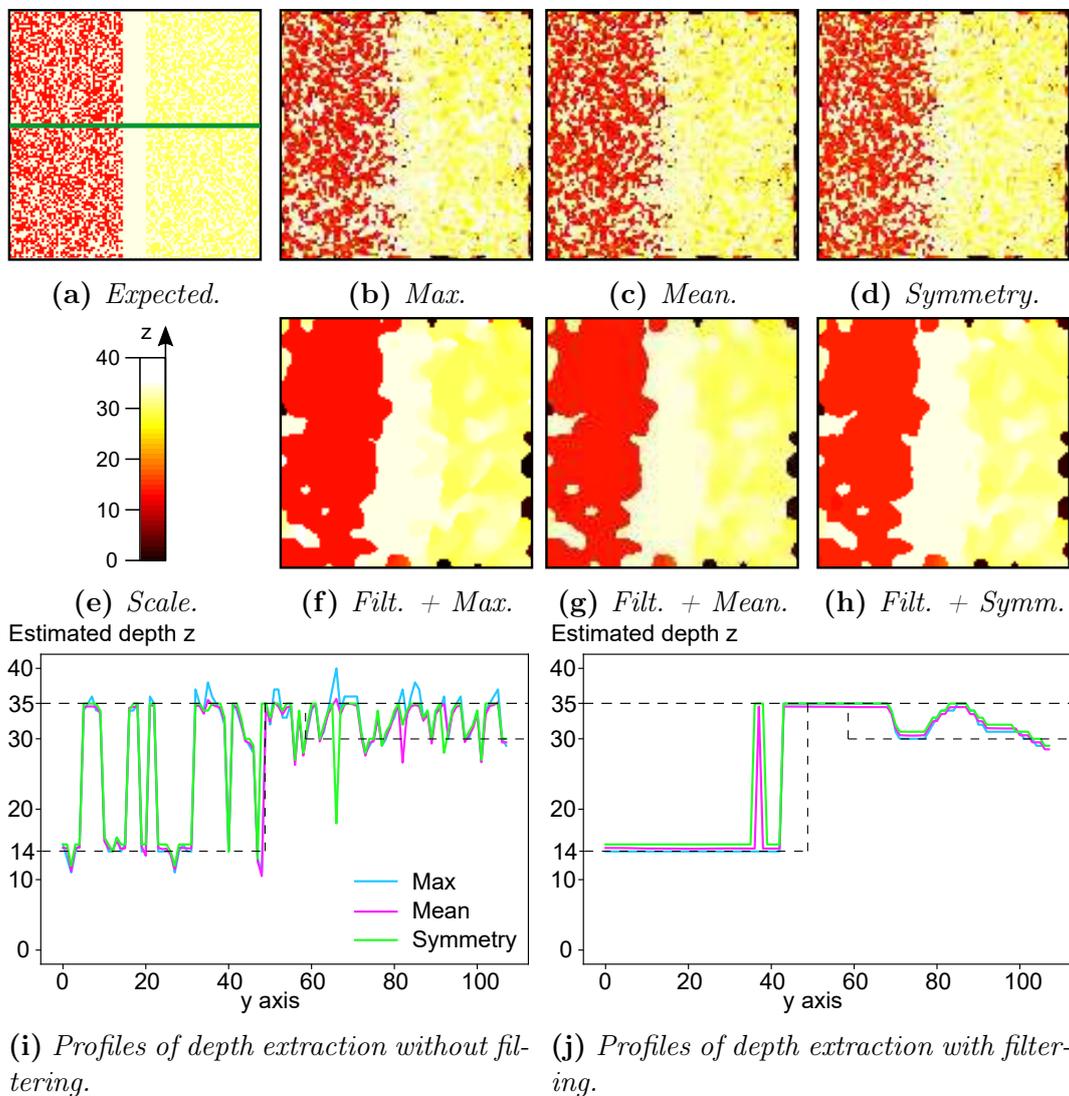


Figure 4.13 – Depth extraction methods tested on the case of the three overlapping Left, Right and Third planes together.

planes.

The same granularity is obtained in the depth maps created by the depth extraction methods without filtering. Some noise is recreated with the colors corresponding to the depths of the planes. Contrary to the previous cases, the noisy areas here represent the overlapping planes that the methods are able to detect. The presence of the Third plane that occupies the whole surface helps in obtaining coherent results. It prevents undesired noise as well as artifacts at the edge of the plane as in the previous cases (Fig. 4.7).

On the contrary, the filtering step results in a blurring where the Left and Right planes occupies large areas. They are better detected and spread over the depth map because they are in the front compared to the Third plane. The Third plane is only visible in the small area between the two Left and Right planes of the foreground. The gap between the two Left and Right planes is better detected compared to the previous two-planes case (Fig. 4.10).

In Figures 4.13i and 4.13j, the profiles confirm what was observed on the depth maps. The methods applied on the *Ring Difference Filter* without filtering shows interesting profiles, with curves alternating between two bounds corresponding to the Left/Right and Third planes. For low y values, the Left and Third planes can be distinguished because the signal alternates between two distant depths z . But for high y , the signal stays between the two close Right and Third planes located at $z = 30$ and $z = 35$. Without the expected signal in dotted black, these quick variations could be interpreted as noise around a plane located around $z = 32/33$. This is the consequence of the choice of two planes very close to the limit of res_{depth} . The profiles with filtering show flattered curves with remaining artifacts. This corresponds to the depth maps composed of large flat areas with irregular contours.

4.4 Conclusion

In this chapter we tested the ability of several depth from focus methods to extract depth maps from refocused stacks. We used our simulation and reconstruction algorithms developed in Chapter 3 to generate stacks of one, two adjacent or three overlapping planes. Depth from focus is a two-steps technique, blur measurement followed by depth extraction, which outputs a depth map. The five tested blur measures were able to detect a single plane at its correct depth and location. A deeper study allowed to select the *Ring Difference Filter* as the best measure. Applying depth extraction has shown similar results for the three tested methods, but with less artifacts with the *Mean*. A Gaussian filtering between the two steps helped reducing noise, but increased the artifacts.

In the two-planes case, we observed a difference of contrast depending on depth, as previously studied in Chapter 2. This variation of contrast might

complicate depth extraction. The *Ring Difference Filter* associated with Gaussian filtering succeeded at extracting the planes at their correct depths.

With the three-planes case, the separation of the two close overlapping planes was difficult, due to their proximity (inter-plane distance $\sim 1.5 \text{ res}_{depth}$). Removing the filtering enabled detection of the third plane located behind the other ones.

These methods were accurate in this specific case of 2D planes detection. However, depending on the scene, some methods might be more efficient than others, especially concerning the use of filtering and the choice of window size. This is a drawback of depth from focus.

The notion of sharpness and contrast is very relative in the refocused stack. It depends on many parameters including the composition of the scene, the chosen depth of reconstruction and defocus blur coming from adjacent structures. These three different effects compete one against each other and affect the defocus blur needed for depth from focus methods. In the next chapter, we will see the application of these methods on experimental X-ray plenoptic images.

Chapter 5

Demonstration of X-ray plenoptic

In the previous chapters, we studied plenoptic from a theoretical and algorithmic point of view. We validated the configuration and resolutions formulas on experimental images taken in the visible spectral range, and tested the refocusing and Depth From Focus methods on both simulated and visible light data. However, we oriented our work in the objective of performing X-ray plenoptic microscopy and took into account the specificities of X-rays: transmission of light through the sample instead of reflexion on its surface, low number of micro-lenses.

In this chapter, we present the application of the theoretical and algorithmic work on an experiment of X-ray plenoptic microscopy. In the context of the *VOXEL* project, we accessed a beamline at the synchrotron PETRA III of DESY, in Hambourg, Germany, and performed several experiments. The aim was to test the possibility to achieve X-ray plenoptic microscopy, to acquire images in various configurations, to study their resolutions, refocus them with our algorithms and finally test some Depth From Focus methods.

5.1 The experiment and optics

The experiments took place in the experimental hutch 1 at the beamline P05 at PETRA III in DESY, which is the beamline dedicated for nanotomography and X-ray imaging techniques (see Fig. 5.1). We benefited from material and technical support from the Helmholtz Centre Geesthacht team working at DESY. The beamline corresponds to an X-ray microscope, composed of an illuminating system, a main lens ensuring a high magnification of ~ 100 , and a pixelated detector. To achieve X-ray plenoptic, we modified the existing configuration by adding our X-ray micro-lens array in front of the detector. Figure 5.2 shows a broad view of the mounted experiment with the different optical components. In the following subsections, we will present in details the optics used and how they limit the choice of the optical configuration and impact the procedure of image acquisition.



Figure 5.1 – Aerial view of PETRA III buildings at DESY.

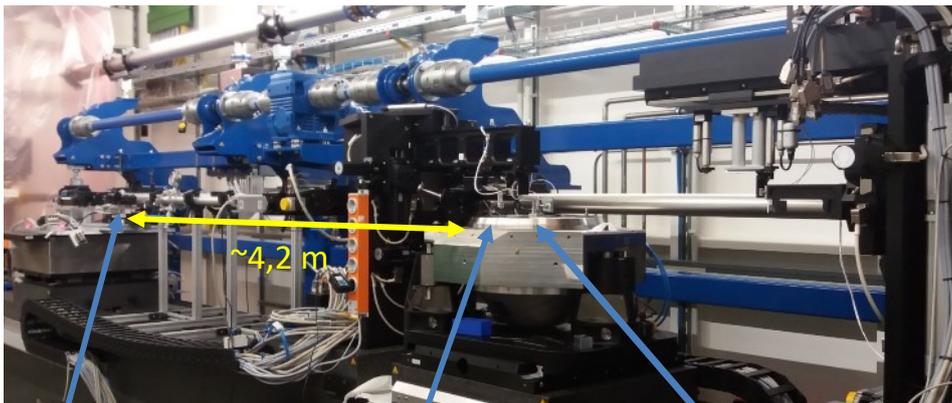


Figure 5.2 – Locations of the optical components in the mounted experiment.

5.1.1 The X-ray source

X-rays are electromagnetic wave with photon energy between typically 100eV to 100keV , which correspond to wavelengths between 3nm to 0.03nm [Att99]. Synchrotron PETRA III generates a white beam with a cut-off around 100keV . The beam has been monochromatized at an energy of 11keV , which corresponds to a wavelength of $1.127 \times 10^{-10}\text{m}$. At this energy, the high penetrating properties in biological tissues allows to perform biological imaging. The rays are quasi parallel when entering the experimental hutch. They are then focused on the sample zone with the use of a *beam shaper*. This optics was

specifically designed for the pre-existing X-ray microscope, so that the numerical aperture (NA) of the beam shaper matches the NA of the main lens. The focal spot has a fixed diameter of $50\mu m$. For more details about the beamline, the reader might refer to [PI].

5.1.2 The main lens

X-rays optics are difficult and costly to fabricate. As mentioned in Chapter 1, refractive lenses are not common in hard X-rays. Instead, we use *Fresnel Zone Plates* (*FZP*) (see Fig. 5.3). *FZP* are diffractive lenses, which behave here like focusing lenses [Att99]. They are made of circumscribed circular lines, also called Fresnel zones, that alternatively dephase the incoming light by 0 or π [MKV⁺14]. These structures correspond to variable spacing diffraction gratings that condensate the diffracted light on a focal spot. The circular symmetry provides a circular spot. With a *FZP*, it is well known that the spatial resolution equals the size of the last line. Also, the size of this line and the X-ray energy determine the focal distance, which can then be used in the thin lens equation [TV01]. This justifies the use of the plenoptic theory and algorithm, initially designed for visible light.

In our experiment, the main lens is a *FZP* with a focal length of $124.21mm$ at $11keV$ and a diameter of $280\mu m$. Figure 5.3 shows the mounted main lens placed just after the sample location in the experiment. Its characteristics correspond to a numerical aperture of $NA = \frac{d}{2f} = 1.127 \times 10^{-3}$. This very low numerical aperture is the main limit in our experiment. This is a drawback for a plenoptic experiment, as larger sources and larger NA allow better resolutions (see Chapter 2). The first consequence of this low NA is a very low depth resolution compared to the lateral resolution ($res_{depth} / res_{lat} = 1/NA = 787$), predicting difficulties for 3D reconstruction. This will be discussed in Section 5.3.2.4. The second consequence of this low NA is large working distances for the main lens. Combined with high magnification, this

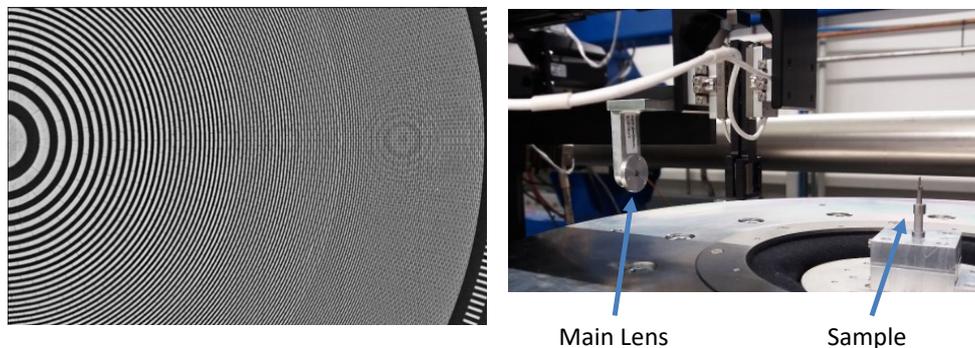


Figure 5.3 – Illustration of a *Fresnel Zone Plate* (left), and the mounted main lens just after the sample (right).

implies very long distances for the experiment, inducing additional technical difficulties (see Section 5.3.2.6).

5.1.3 The Micro-Lens Array

The μ LA was specifically fabricated in the context of the *VOXEL* project (see Fig. 5.4 left). It is composed of 9×9 similar *FZP* micro-lenses. This very low number is due to the technical difficulties and cost of fabricating a large number of very small *FZP*. Their focal distance is 88.7mm at 11keV and their diameter is $100\mu\text{m}$. Two adjacent micro-lenses are separated by a distance of $10\mu\text{m}$, which gives a center-to-center distances of $110\mu\text{m}$, also called *pitch*.

In the case of diffractive lenses, part of the radiation is transmitted through the *FZPs* without being diffracted. This is called the zero order. The useful signal is contained in the first order, which corresponds to the focused signal. Thus, it is necessary to block the zero order with a central disk of diameter $30\mu\text{m}$. This blocker is made of Si_3N_4 covered with gold on necessary parts. Indeed, Si_3N_4 is transparent and gold is opaque to 11keV X-rays. The blocker was carefully aligned with the μ LA. They were mounted on the same structure so that they can be manipulated together (see Fig. 5.4 bottom).

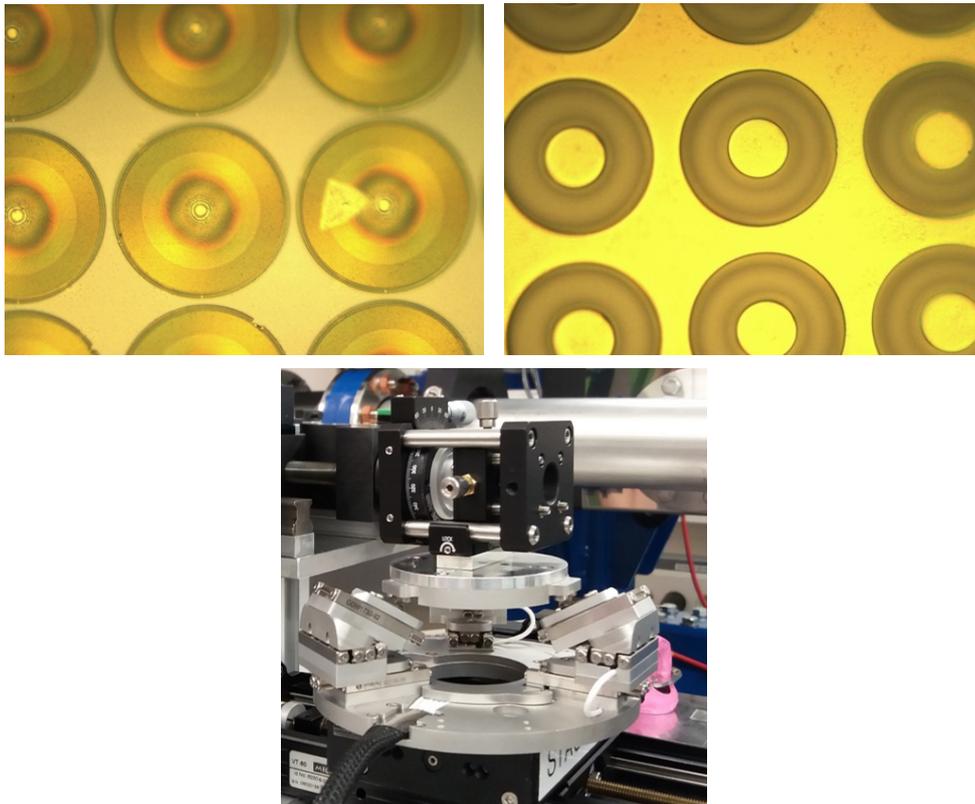


Figure 5.4 – Image of the μ LA (top left) and its blocker (top right) acquired with a visible microscope, and how they were mounted together for the experiment (bottom).

5.1.4 The imaging system

The imaging system is composed of a scintillator that transforms X-rays into visible light, then acquired by a PCO camera (see Fig. 5.5). A relay imaging system guides the light to the camera which is slightly disaligned from the X-rays optical axis, in order to prevent potential damage of the camera. The entire system can be considered as a camera of 2048×2048 pixels of size $0.93\mu m$.

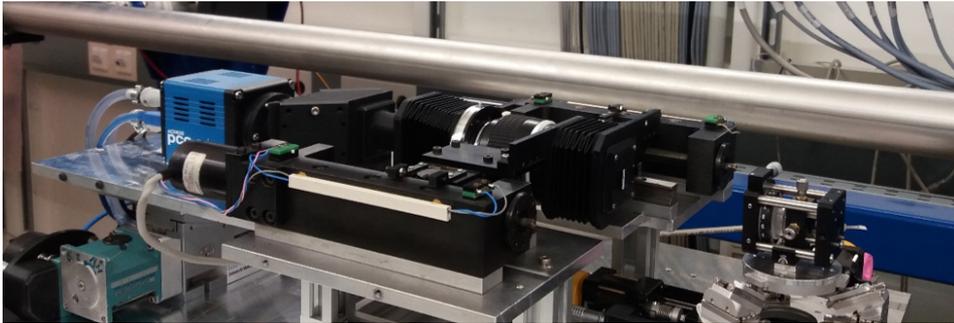


Figure 5.5 – *The camera and its relay imaging system. We can also notice the μLA and its blocker before it.*

5.1.5 The sample

Finally, the sample used in the experiments are two USAF 1951 test patterns built for X-rays imaging. They contain lines of various sizes allowing to study lateral and depth resolutions. Before building the plenoptic setup, the sample is aligned with the optics and is imaged by the main lens and the camera only (see Fig. 5.6). The first test pattern is located on the bottom half of the image and shows lines of width $1\mu m$ and $0.5\mu m$; we call it TP1. The second test pattern shows thinner lines until $250nm$; we call it TP2. We acquired a focal stack that allows to verify that the two test patterns are parallel. We measured a distance of $1.3mm$ between the two, with TP1 closer to the main lens than TP2.

5.1.6 Optical configurations

The characteristics of the experimental components are summarized in Table 5.1. All these components (the two test patterns, the main lens, the μLA and its blocker, and the camera) were carefully aligned with the beam. Some of them were mounted on micrometric motors so that we could precisely change the distances and test different configurations.

The equations established in Chapter 2 are used to choose the configuration for our X-ray plenoptic camera. Unfortunately, after testing different designs,

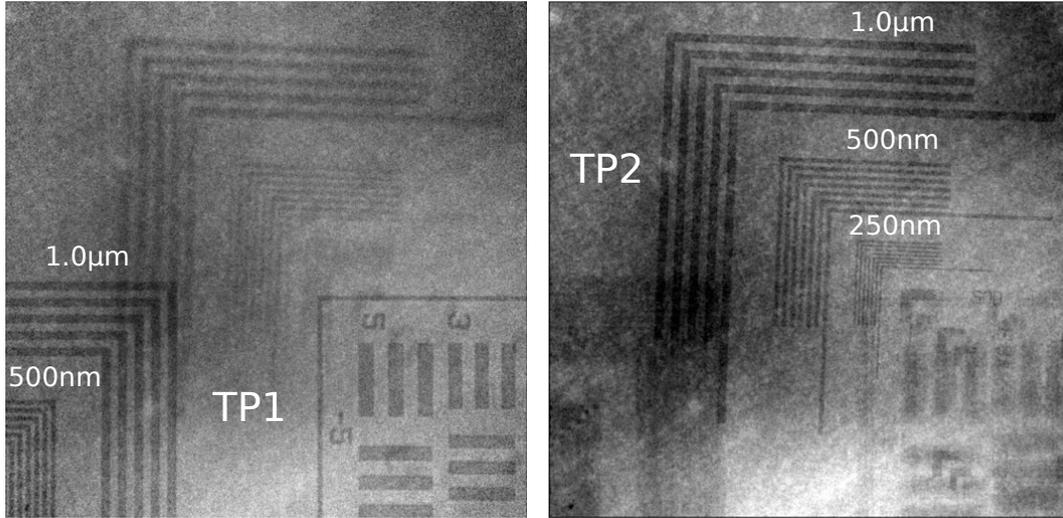


Figure 5.6 – The two USAF 1951 tests patterns used in the experiments, imaged with the main lens and camera only. TP1 (left) is 1.3mm closer to the main lens than TP2 (right). Notice the heterogeneous illumination.

Beam shaper	spot size Δ_{beam}	$50\mu m$
	divergence NA_{beam}	$1.1 \times 10^{-3} rad$
Main lens	focal f_1	$124.21mm$
	aperture d_1	$280\mu m$
Microlens array	focal f_2	$88.7mm$
	aperture	$100\mu m$
	pitch d_2	$110\mu m$
	# micro-lenses N_2	9×9
Sensor	pixel size Δp	$0.93\mu m$
	# pixels	2048×2048

Table 5.1 – Parameters of the X-rays plenoptic experiment at DESY.

we realized that it was not possible to respect the aperture matching condition (see Eq. 2.3) with the available optics. On the one hand, a high magnification is needed in order to project the small illumination spot ($\Delta_{beam} = 50\mu m$) on the large μLA ($N_2 \times d_2 = 990\mu m$). This requires a magnification $M \geq 20$, which reduces accordingly the divergence of the beam. Combined with the focal length of the main lens $f_1 = 124.21mm$, it implies a large distance c between ML and μLA (more than $2m$). On the other hand, the characteristics of the μLA impose a high divergence of the incoming light, thus a small distance c (around $150mm$). The unique solution was to work without respecting the aperture matching condition. We introduce a factor k to quantify the deviation

from Equation 2.3:

$$\frac{d_1}{z_1 + a + b} = k \frac{d_2}{b} \quad (5.1)$$

We wanted to implement both 1.0 and 2.0 types of configurations. The selected configurations are presented in Table 5.2 with the expected resolutions. They were optimized to achieve the best possible resolutions in all directions. The very low factors k indicate that we are far from the aperture matching condition. This implies that sub-images are expected to be very small on the sensor. We also consider the illumination conditions imposed by the optical configuration (see Section 2.1.3). The minimum necessary source size $\Delta_{ill} = 33\mu m$ is smaller than the physical spot size $\Delta_{beam} = 50\mu m$ due to the beam shaper, and both NA are equivalent: $NA_{ill} \sim NA_{beam}$. This indicates that illumination will not restrict the resolutions presented in Table 5.2.

		configuration 1.0	configuration 2.0
Distances	object-ML or z_0	128.4mm	128.9mm
	ML- μ LA c	3845.8mm	3845.8mm
	μ LA-detector b	91.8mm	111.8mm
Theoretical resolutions	res_{lat}	3.67 μm	0.13 μm
	FOV	33.05 μm	33.18 μm
	res_{depth}	3.46mm	219.58 μm
	DOF_{pleno}	254.71mm	135.27 μm
Spatio-angular samplings and magnifications	N_u	4.59	0.32
	N_s	9	246.3
	M	29.95	26.5
	m	-	0.26
Aperture matching	k	0.065	0.079
Illumination	Δ_{ill}	33.04 μm	
	NA_{ill}	$1.2 \times 10^{-3} rad$	

Table 5.2 – Configurations used in the experiment, with their distances and their expected resolutions.

5.2 First results and preprocessing

5.2.1 Analysis of a raw image

Figure 5.7 shows an example of a raw image acquired by the plenoptic camera in configuration 2.0. The image is composed of 9x9 patterns corresponding to

the 9×9 FZP of the μLA . Each pattern contains a bright disk inside a gray ring. The central bright disk is the first order of the FZP and contains the useful information. It corresponds to the sub-image that will be extracted and used in the refocusing process. The larger gray ring around it is the zero order of the FZP . It contains useless information in this context. The important aspect is that it does not affect the sub-image as the gray ring does not overlap the bright disk, thanks to the blocker.

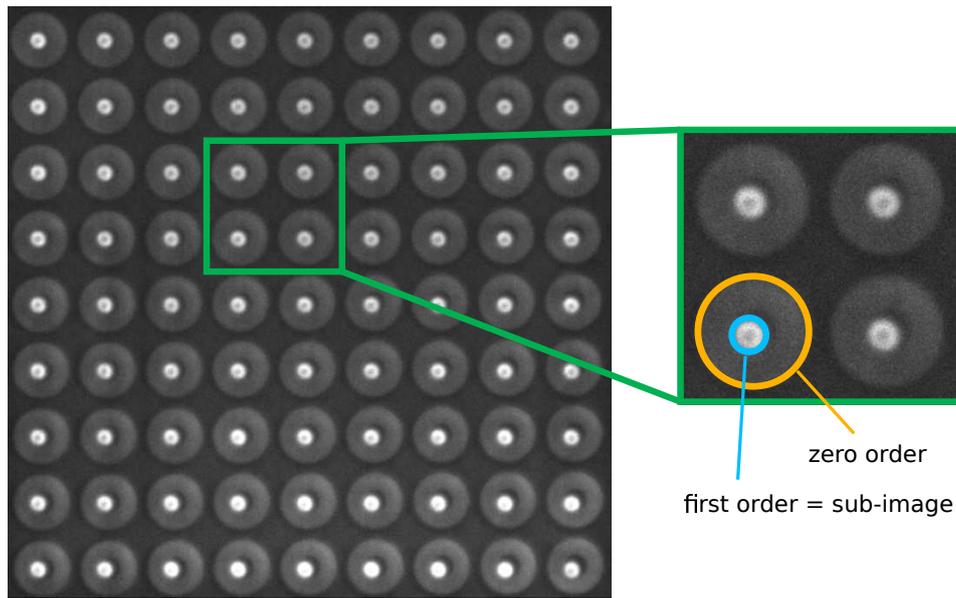


Figure 5.7 – A raw image as acquired by the imaging system, and a zoom showing the first (blue) and zero (yellow) orders due to the FZP .

On the bottom right of the raw image, one may notice that the two orders (bright disks and gray rings) of the FZP are slightly shifted. This corresponds to the divergence of the beam, which projects the zero and first orders with different angles. In the zoomed portion, the sub-image (or first order) corresponds to a ring with a darker point at the center, creating a *donut shape*. This is a characteristic pattern of a zero order blocker (see Fig. 5.4). The main lens was a diffractive lens provided with an integrated blocker, and this *donut shape* probably corresponds to the blocker being imaged on the sensor. In some configurations, the sub-image is very small and this effect is negligible: the sub-images can be considered as full disks. Larger sub-images are preferred to maximize the spatial sampling of the sample, but they create this *donut shape*. We tested several configurations leading to different sizes of the sub-images. We ended with configuration of Table 5.2, which produces the largest possible sub-images without being too much affected by the *donut shape*.

5.2.2 Acquisition procedure and preprocessing

We set the distances to one of the chosen configuration and began the acquisition. Because of the strong attenuation of the different elements constituting the whole system, the acquisition time was set to 5 seconds per image. For each configuration, ten images were acquired and then averaged to create an image with better *SNR*. The beam was also slightly heterogeneous and varying with time. Therefore, it was necessary to acquire reference (blank-field) images by removing the sample from the *FOV*. The average of four reference images is then used for a normalization step, to correct for beam inconsistencies and motor errors. This normalization allows to extract the useful signal from the background (see Fig. 5.8).

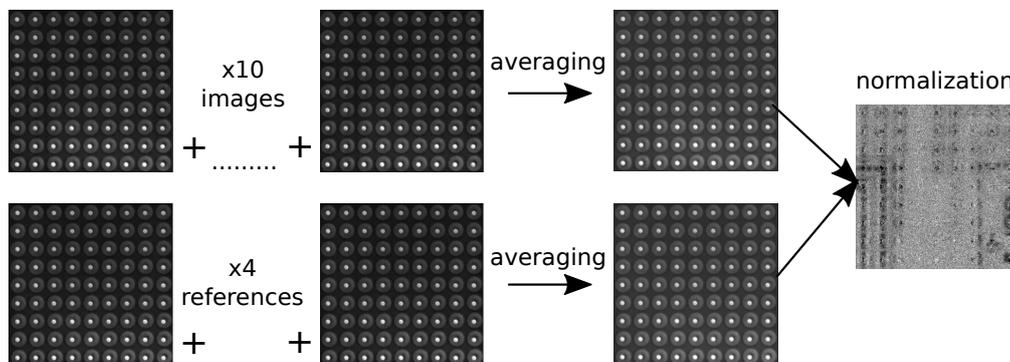


Figure 5.8 – Schematic representation of the preprocessing, consisting in averaging followed by normalization.

An image after preprocessing is presented in Figure 5.9. Thanks to the preprocessing and an adapted contrast display, the two test patterns are now visible. Similarly to Figure 5.7, we observe the 9x9 sub-images surrounded by the rings of the zero orders. Around the sub-images we can see the shadow of the test patterns that is created by transmission of the rays through the μ LA. The signal in the sub-images is more contrasted (inside the blue circles in the zoomed parts). We can see the lines of the test patterns with various magnifications. The zoomed part on the right shows lines from TP2, with original width of $0.5\mu\text{m}$ in object-space. We can recognize three lines in each sub-image. The middle zoomed area represents some lines of TP2 but this time with a larger width of $1.0\mu\text{m}$. In these sub-images, two lines are visible each time. The left zoomed part shows the lines of TP1 with exactly the same line width of $1.0\mu\text{m}$. Only one line is visible in each sub-image. This corresponds to a difference of magnification of the two test patterns, due to their different depth locations. This shows that the raw image contains depth information that could be exploited for reconstruction.

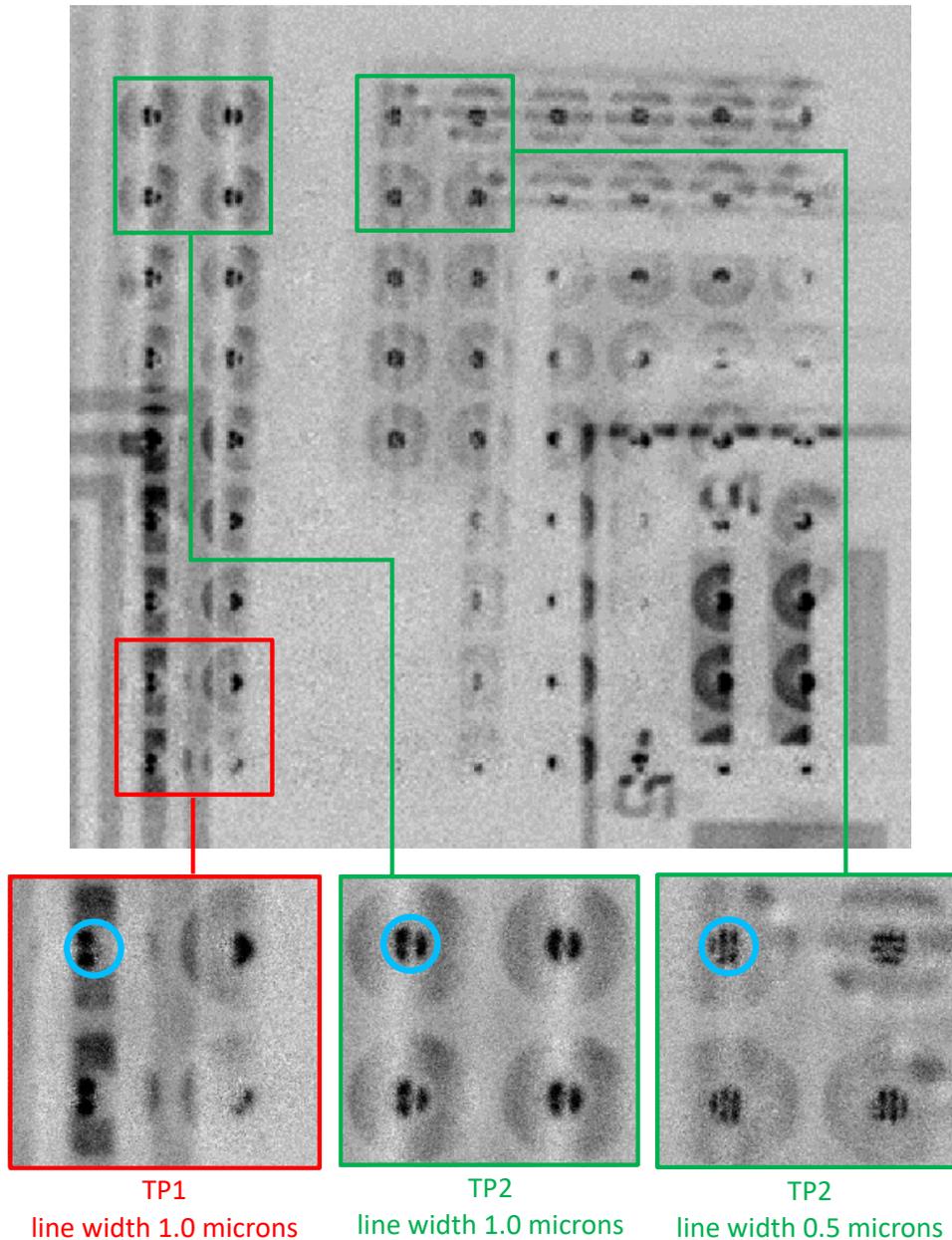


Figure 5.9 – A raw image after preprocessing. Some sub-images are highlighted in blue. The zoomed areas show variations of magnification in the images of the lines with the same line width.

5.2.3 Image stitching

As explained above, the small size of the sub-images is a drawback, but unfortunately they cannot be increased due to the *donut shape*. To compensate this, we performed a technique of image stitching. We acquired several images by changing the lateral position of the μ LA. Between each set of acquisition (10 images and 4 references each time), we moved the μ LA by a distance of

$d_2/4 = 27.5\mu m$, so that the centers of the sub-images are shifted by one fourth of their center-to-center distance. The number of 4 was chosen experimentally to optimize the space on the detector, so that sub-images acquired with different shifted positions of the μLA are adjacent without overlapping. We repeated these shifts in both lateral directions. The process is schematized in Figure 5.10-left.

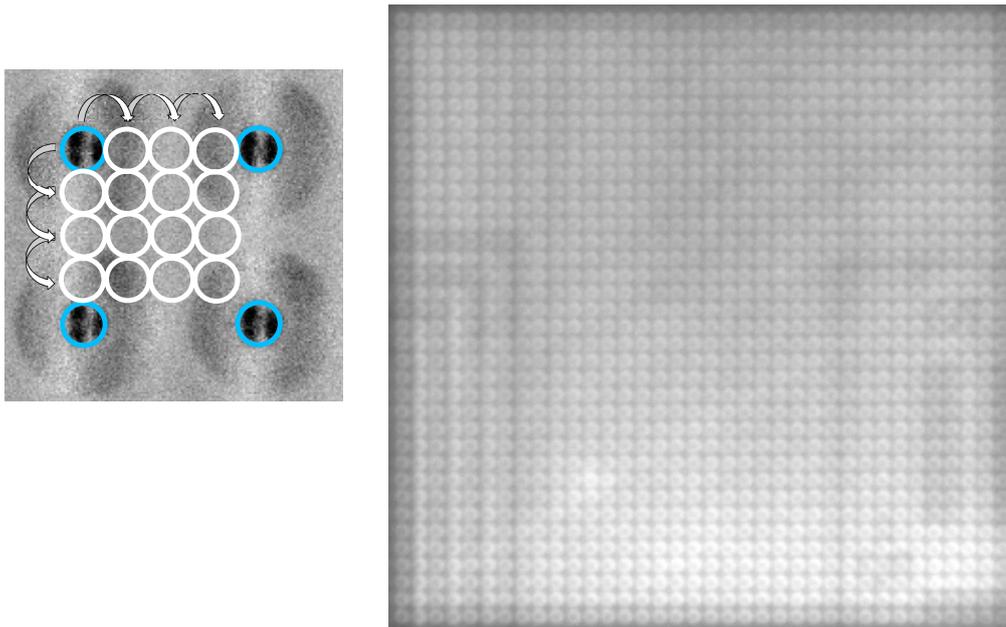
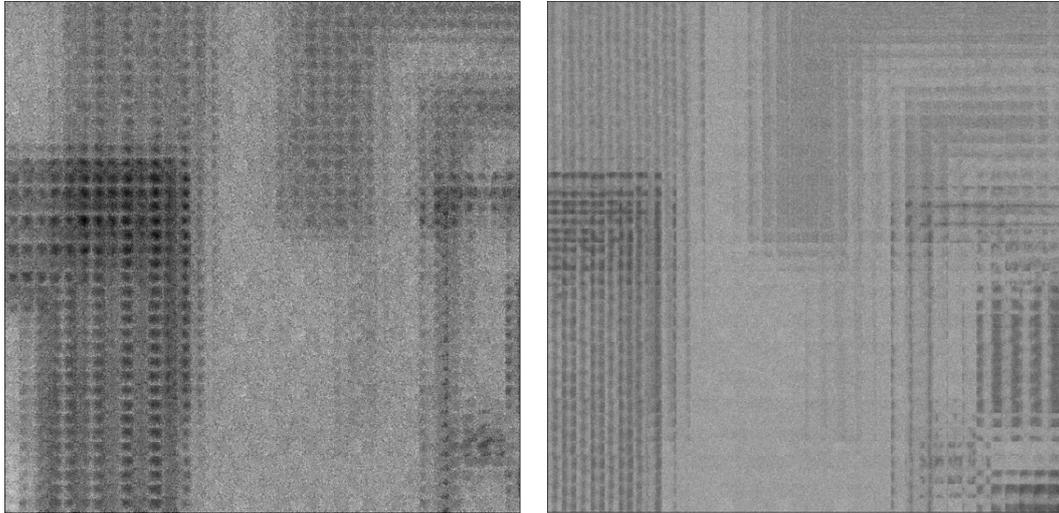


Figure 5.10 – *Principle of stitching: the μLA is shifted to create new sub-images. We obtain 36×36 sub-images instead of 9×9 initially.*

The right part of this figure shows an image combining all the sub-images together, to illustrate how the stitching complete the data from a single acquisition. The number of acquired sub-images is multiplied by 4 in both directions, thus we now have 36×36 sub-images. However this is not equivalent to a system with physically 36×36 micro-lenses. During the stitching, the light-field is sampled differently by the different lateral position of the μLA . Thus, stitching is expected to complete missing data, but will not improve the resolutions of the reconstructed image, as the optical configuration has not changed (same distances and optics). In Section 5.3 we will study the effect of stitching on image quality. This stitching process also increases the acquisition time by a factor of 4×4 . It takes more than one hour to acquire each dataset.

The preprocessing step is executed for each position of the μLA separately. Finally, the exact center of each sub-image has to be found manually, allowing to extract all the sub-images from the different raw images. The sub-images are then combined together to form a dataset with 36×36 sub-images. Figure 5.11 presents the result of this combination in the case of configurations 1.0 and 2.0. The datasets are now ready to be used as inputs for the different refocusing

algorithms.



(a) *Dataset 1.0.*

(b) *Dataset 2.0.*

Figure 5.11 – *Datasets 1.0 and 2.0 after preprocessing, extraction of the sub-images, and combination of the 36×36 sub-images from the stitching.*

5.3 Refocusing of the acquired data

We first present the results from configuration 1.0 before moving to configuration 2.0. We refocused using both 2.0 and our algorithms. When using our algorithm, an additional calibration step is mandatory. It consists in calculating the size and period of sub-images from the parameters of the system (see Eqs. 2.1 and 2.2), and verifying that they match the physical size and period of the experimental sub-images. This is done by small adjustments of the distances used as input in the algorithm.

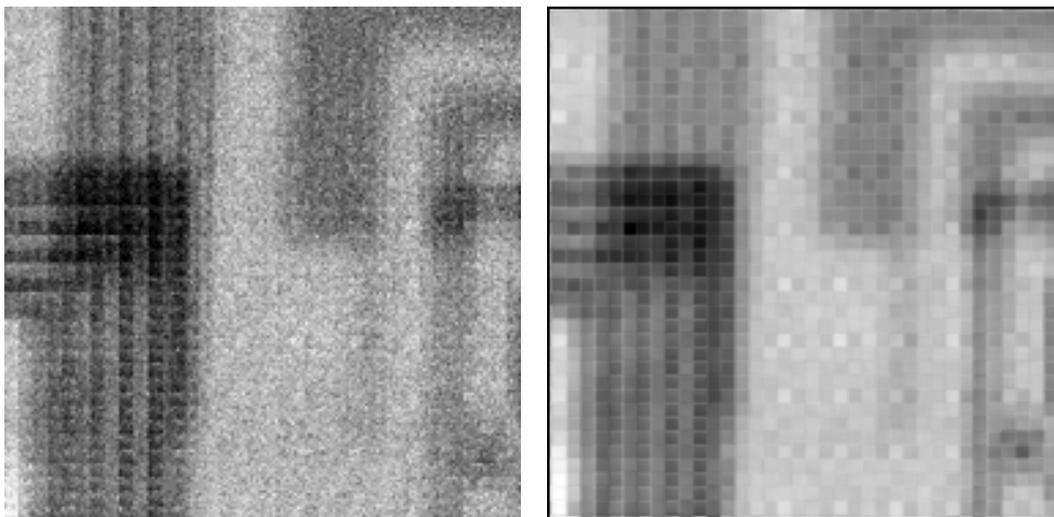
5.3.1 Configuration 1.0

The configuration 1.0 was mounted and the raw and reference images were acquired, using the two test patterns as samples (see Fig. 5.6). The dataset then followed the whole preprocessing and sub-images extraction as described in previous section. The corresponding preprocessed image is shown in Figure 5.11a.

Refocusing this dataset using algorithm 1.0 provides an image with only 36×36 pixels (not shown), corresponding to the number of sub-images after the stitching. Figure 5.12 presents the reconstructions using algorithm 2.0 (left) and our algorithm (right), which allow a larger number of pixels (respectively 189×189 and 400×400). In both refocused images, we can recognize the two

test patterns by their general shapes. On the left of the images, some of the lines of width $1.0\mu\text{m}$ are visible. This is below the theoretical value of $3.67\mu\text{m}$ (see Table 5.2), which indicates that it might be due to artifacts, similar to misalignment artifacts in Chapter 2.

The lines of TP1 are more contrasted than TP2, probably due to their different locations, but the quality of the images prevents from extracting reliable depth. The reconstruction using our algorithm (see Fig. 5.12b) is less noisy than with algorithm 2.0 (see Fig. 5.12a), but it does not improve the visibility of the lines. The square patterns over the image are due to the spatial sampling done by the micro-lenses in a configuration 1.0. As explained in Chapter 2, the image quality and resolution are related to the number and size of the micro-lenses. On this beamline and using the available X-ray optics, a configuration 1.0 is not suitable for achieving high quality X-ray plenoptic images.



(a) *Reconstruction with algorithm 2.0* ($P = 10$). (b) *Reconstruction with our algorithm* ($z = 128.40\text{mm}$).

Figure 5.12 – *Dataset 1.0 reconstructed using algorithm 2.0 and our algorithm.*

5.3.2 Configuration 2.0

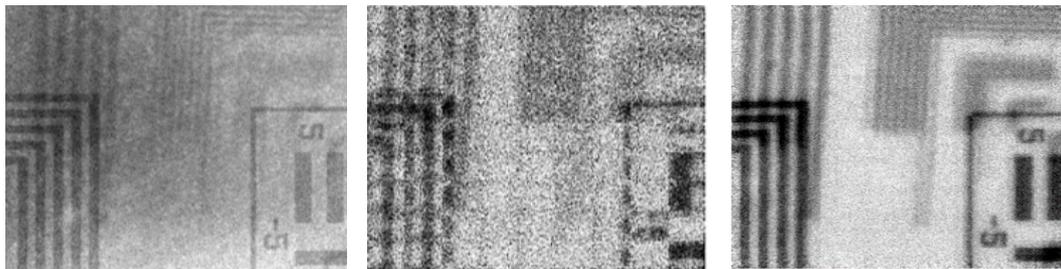
We now present the results obtained in configuration 2.0 (see Table 5.2). Figure 5.11b presents the dataset after the stitching, preprocessing and combination of the sub-images. Using this dataset, we study in detail the action of stitching, the lateral and depth resolutions, as well as the plenoptic depth of field.

5.3.2.1 Effects of the stitching

Figure 5.13 compares refocusing using the 9x9 sub-images given by a single position of the μ LA to the refocusing using the 36x36 sub-images from the stitching. We also show the image of the test patterns acquired without the μ LA to compare to a non-plenoptic setup (in Fig. 5.13a). The first important difference between refocusing with 9x9 and 36x36 sub-images is that the dotted lines in Figure 5.13b are now continuous. Stitching allows to complete the missing data and the refocused image is fully reconstructed.

Moreover, stitching reduces the noise in the final image. Figure 5.13b is very noisy and the lines of TP2 are poorly visible. On the contrary, the *SNR* of Figure 5.13c is clearly better with also a higher contrast. As explained in Section 5.2, the stitching process consists in acquiring 16 times more images than using a single μ LA position, combined together in the refocusing process. This can be considered as signal averaging of 16 images, which is known to improve *SNR* by a factor of $\sqrt{16} = 4$ [HA10].

By improving image quality and contrast and completing missing data, stitching improves perceived resolutions of the test patterns. We use the 36x36 sub-images in the rest of this section.



(a) Reference image (see Fig. 5.6). (b) Refocused image using 9x9 sub-images. (c) Refocused image using 36x36 sub-images.

Figure 5.13 – Effects of the stitching process on the refocused images in configuration 2.0.

5.3.2.2 Refocusing with algorithm 2.0

We study the refocusing done with the "rendering with blending" method (as defined in Section 1.3.2). Figure 5.14 shows refocused images for different patch sizes P . The depth values corresponding to these P were calculated using to the formula 1.17. We see the evolution of the refocusing sharpness according to the value of P , for example looking at the large line of TP1 (circles in red) or the thin lines of TP2 (circles in green). For low and high values of P , the test patterns are blurred, whereas for intermediate values, the lines are sharp enough so that we can separate them. The lines seem to be on focus in the image with patch size $P = 8$ (Fig. 5.14c). Both TP1 and TP2 seem on focus in

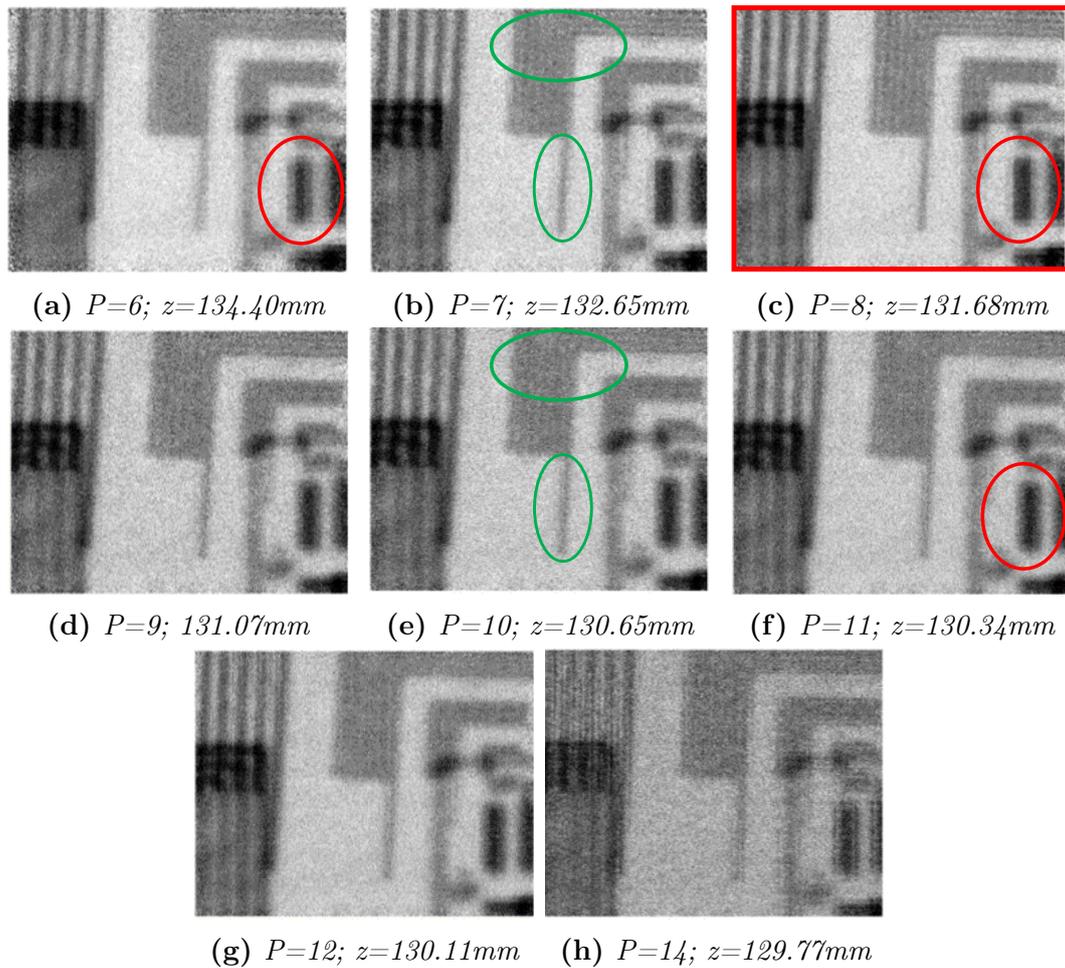


Figure 5.14 – Set of refocused images acquired with the configuration 2.0, and using algorithm 2.0, with various patch sizes P .

this same image. With this refocusing algorithm, it is not possible to separate the two test patterns distant of 1.3mm from each other.

5.3.2.3 Refocusing with our algorithm

We then refocus the same images of configuration 2.0 with our algorithm (see Chapter 3). Figure 5.15 shows the refocused image at depth $z = 128.77\text{mm}$, which was the image with the sharpest reconstruction for TP1. The quality of the reconstruction is better than with algorithm 2.0, especially for TP1, on which we can now separate the lines on the bottom left of the image, and read the number 5 that was unreadable in Figure 5.14. This algorithm allows to better separate the lines from the two test patterns when they overlap on the left of the image. Neither the contrast nor the noise are particularly improved by using our algorithm. On this image, the lines are visible until a thickness of $0.5\mu\text{m}$. The lines of smaller thickness are indistinguishable. This can be

compared to the theoretical value of $res_{lat} = 0.13\mu m$ for this configuration 2.0. There is a factor of 4 between the observed and the expected res_{lat} , but we stay in the same order of magnitude. Using this image, we estimate an experimental FOV of $33.73\mu m$, consistent with the theoretical value $33.18\mu m$ (see Table 5.2).

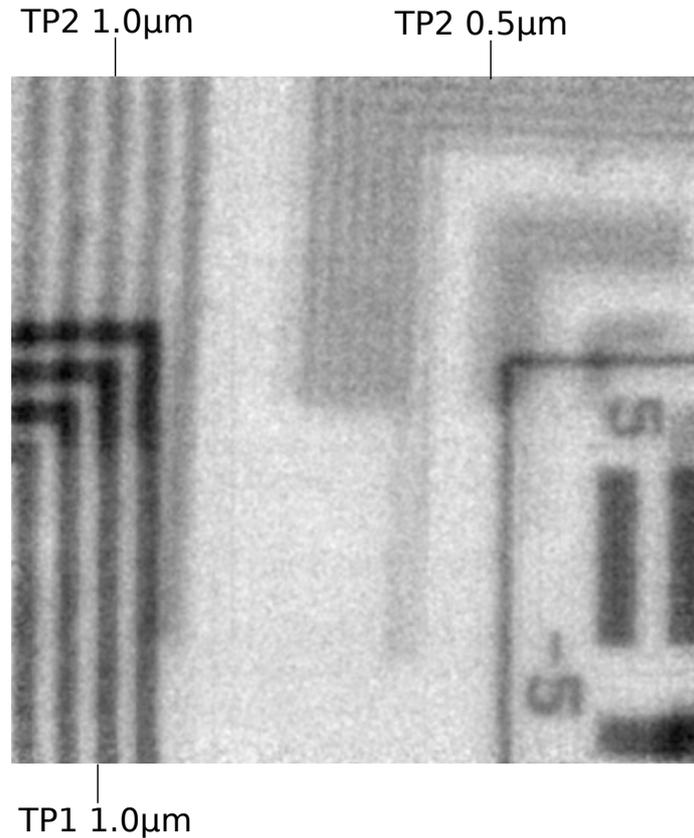


Figure 5.15 – Refocused image at $z = 128.77mm$ in configuration 2.0 using our algorithm.

In Figure 5.16 we compare refocused images at various depths. As for the refocusing with algorithm 2.0, the image of the test patterns is sharp for intermediate depths (around $z = 128.77mm$) but is blurred for highest and lowest depth values. In Figure 5.16d we surrounded areas that the reader can use to follow the evolution of image blur with depth: the number 5 of TP1 (red), the thin lines (green) and the vertical line (pink) of TP2, and overlapping lines of the two test patterns (blue). The lines of both test targets are sharp at the same depth $z = 128.77mm$ (see Figure 5.16e). TP1 is more contrasted and seems sharper than TP2, which suggests that it is close to perfect focus. TP2 is sharper in this figure compared to the other images, as thin lines (in green in Fig. 5.16d) can be separated. However, the vertical line (in pink) is duplicated compared to previous reference image (see Fig. 5.6), which indicates that TP2 is not perfectly in focus. The two test patterns separated by $1.3mm$ should be

refocused on two different planes. Next section will explain why they are both the sharpest at the same depth $z = 128.77mm$.

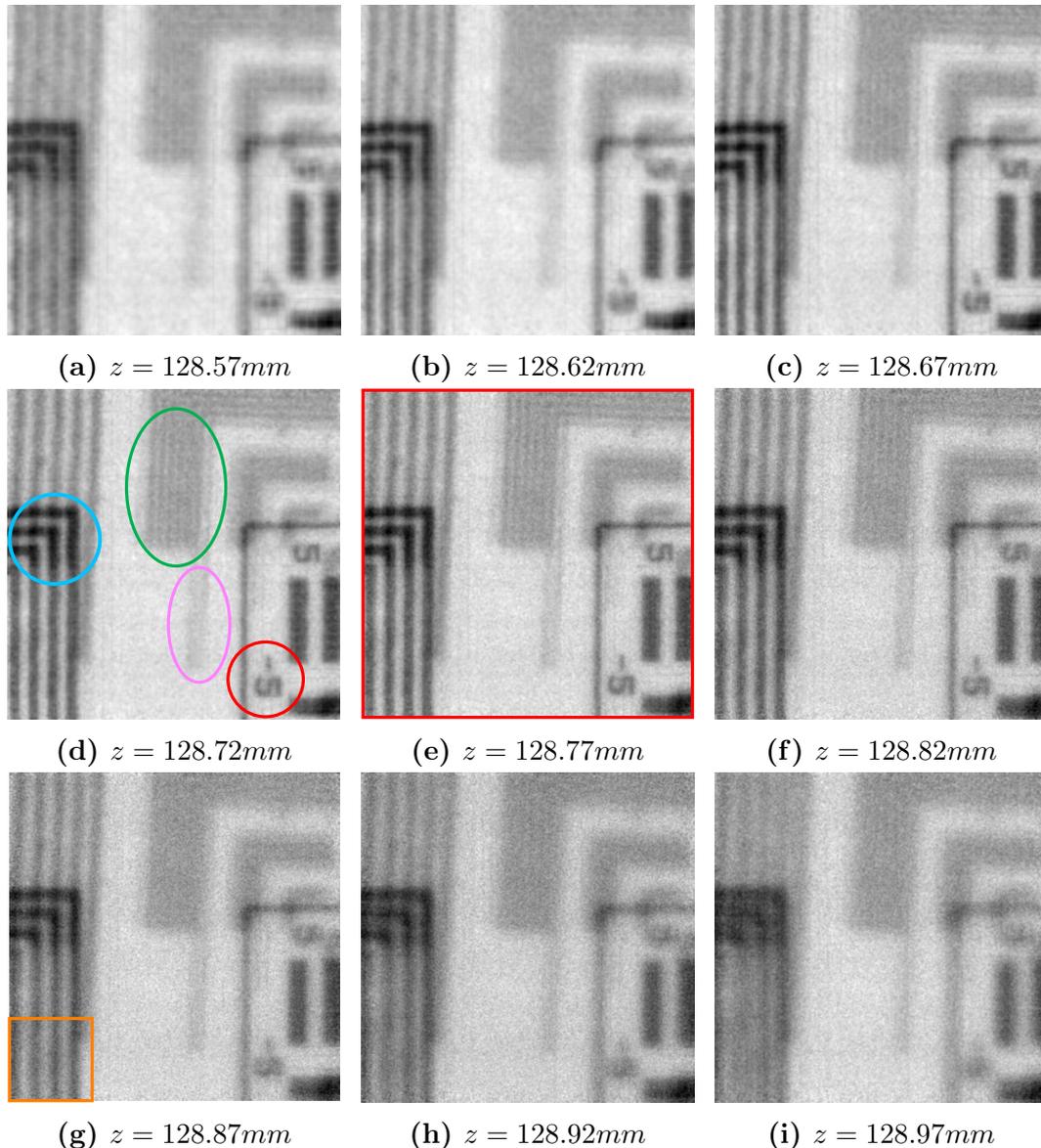
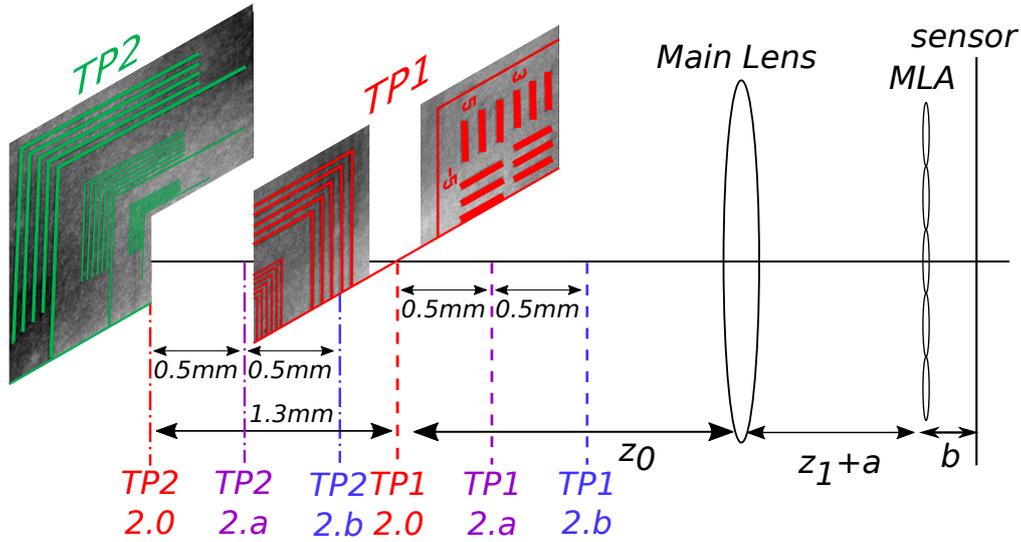


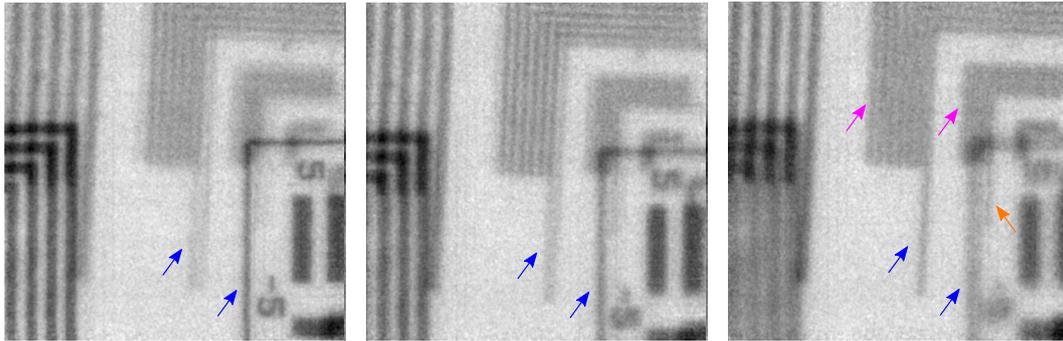
Figure 5.16 – Comparison of the same dataset of configuration 2.0 refocused with our algorithm at different depths.

5.3.2.4 Study of depth resolution and plenoptic depth of field

With both algorithms (2.0 and ours), it is not possible to separate the two test patterns, despite the distance of $1.3mm$ between them. This behaviour raises the question of the cause: is it due to the refocusing process or to the intrinsic properties of the acquired data ? To better study the impact of the geometry,



(a) Scheme of the displacement of the test patterns between the datasets in configuration 2.0.



(b) Dataset 2.0.

(c) Dataset 2.a.

(d) Dataset 2.b.

Figure 5.17 – Comparison of the datasets 2.0, 2.a and 2.b refocused at the same depth $z = 128.77\text{mm}$.

we acquired new data with different positions of the test patterns. We moved the platform that holds the test patterns 0.5mm closer to the main lens for a second dataset (called 2.a), and again 0.5mm even closer for a third dataset (called 2.b). This is schematized in Figure 5.17a. As the two test patterns are attached on the same platform, the distance between them does not change. According to Sections 2.4.3 and 2.4.4, using the appropriate refocusing methods, it would be possible to retrieve both res_{depth} and DOF_{pleno} from the three datasets 2.0, 2.a and 2.b.

We refocused the new datasets 2.a and 2.b at different depths, similarly to dataset 2.0 in Figure 5.16. We observed the same effects for both datasets (not shown): despite the new positions of the sample, the sharpest (not necessarily in focus) images of both test patterns are always at depth $z = 128.77\text{mm}$ and blur progressively appears on adjacent depths. As $z = 128.77\text{mm}$ is indepen-

dent from the position of the object or the algorithm used, we concluded that this privileged depth is imposed by the *geometry* of the optical configuration.

Figure 5.17 proposes a comparison of the refocused images at the privileged position $z = 128.77mm$ for the three datasets. The sharpness and focus properties are well visible on the vertical lines (blue arrows). In dataset 2.0 (Fig. 5.17b, same as Fig. 5.16e), TP1 is in focus, whereas TP2 is sharp but not in focus. In the intermediate dataset 2.a, lines of both test patterns are still sharp but not in focus. In dataset 2.b (Fig. 5.17d), TP1 is completely blurred, whereas TP2 approaches its best focus: the vertical line of width $0.5\mu m$ is thin, and the vertical line of width $0.25\mu m$ is now visible (orange arrow). Misalignment effects are responsible for contrast reduction in the group of seven bars (pink). An accurate image inspection allows to distinguish the group of $0.25\mu m$ lines (pink). This gives a new estimation for the lateral resolution, closer to the theoretical value $0.13\mu m$.

As illustrated in Figure 5.17, the three datasets correspond to a series of six images taken with six different physical depths of test patterns. Following the method presented in Section 2.4.3, this series can be used to estimate res_{depth} . Based on Figure 5.17, two planes seem to be in focus: TP1 in dataset 2.0 and TP2 in dataset 2.b. Their relative distance gives an estimation of res_{depth} around $300\mu m$, to be compared to the theoretical value of $res_{depth} = 219.58\mu m$ (see Table 5.2). We retrieve the same order of magnitude. However, the smallest step between two positions was $300\mu m$, preventing to refine the experimental value.

The same series of images can be used to estimate DOF_{pleno} . As shown in Figure 2.21, the methodology to experimentally measure DOF_{pleno} consists in scanning and selecting the best refocusing parameter for each depth of the test target. In the X-ray case, the same refocusing parameter ($z = 128.77mm$) is always selected as the one generating the best image quality, therefore the same series of refocused images can be used alternatively for res_{depth} and DOF_{pleno} . As a result, the experimental value of DOF_{pleno} is also $300\mu m$, to be compared to the theoretical value of $DOF_{pleno} = 135.27\mu m$ (see Table 5.2).

5.3.2.5 Refinement of depth resolution

To refine the experimental value of res_{depth} , we study an alternative method. Instead of physically moving the test targets and refocusing using the same parameter, we do the opposite: refocusing the same raw image (i.e. same depth of the test pattern) using different parameters of the algorithm. This corresponds to the images previously presented in Figure 5.16. To study the evolution of contrast with depth, we concentrate on the $1.0\mu m$ -lines of TP1, as framed in orange in Figure 5.16g. This same area is extracted from the refocused images, and presented in Figure 5.18 at different depths z .

The zoomed images allow to perceive finer details. The sharpest images

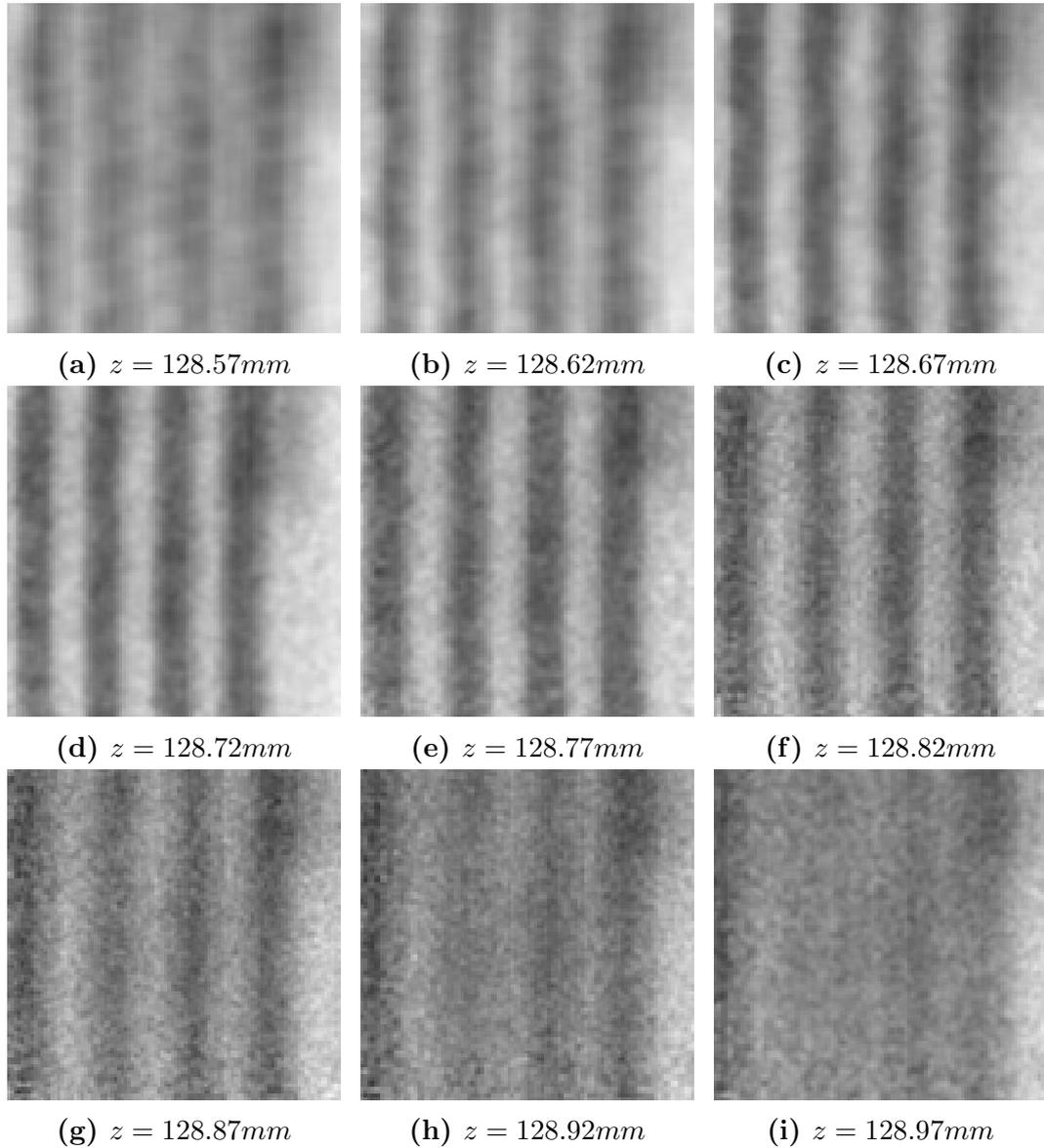


Figure 5.18 – Zooms on the refocused images from Figure 5.16. The size of the extracted areas is drawn in orange in Figure 5.16g.

are for $z = 128.72mm$ and $z = 128.77mm$. On both sides of these values the signal of the bars is degraded asymmetrically.

For lower depth values, the bars are progressively blurred as the distance from $z = 128.72mm$ increases. The blur is uniform over the image, but some square artifacts appear, starting from depth $z = 128.62mm$ and increasing at $z = 128.57mm$. These square artifacts appear when approaching the traditional position $z_0^{trad} = 128.36mm$, where the spatial sampling is done by the micro-lenses, resulting in larger pixels of size d_2/M .

For higher depth values, the refocused images suffer from a high level of

noise. From depth $z = 128.82\text{mm}$, several phenomena occur together: the bars are progressively blurred, the images become darker, and the noise increases. This can be explained by both the configuration and the refocusing process. When the depth increases, the amount of angular data for a single spatial position slowly decreases, resulting in fewer available pixels for the integration process. In our specific conditions, the experiment imposes a low SNR in the raw images and the configuration is responsible for a low N_u . When combining fewer pixels in the refocusing process at higher depths, the resulting images are more sensitive to noise.

From the images of Figure 5.18, the contrast is extracted over a large area in order to average the signal, and plotted against depth in Figure 5.19. The resulting curve presents a smooth profile, with a peak at $z = 128.78\text{mm}$ (the corresponding image was not shown in Fig. 5.18). This is close to depth $z = 128.77\text{mm}$ which is the one that has been presented as the sharpest image throughout this chapter. Based on our previous studies performed in the visible (see Section 2.4.3), we extract res_{depth} considering the criterion of 80% of maximum contrast. The limits of this depth range are at 128.66mm and 128.88mm , which give an experimental measure of $220\mu\text{m}$. It agrees well with the theoretical value of $res_{depth} = 219.58\mu\text{m}$ (see Table 5.2).

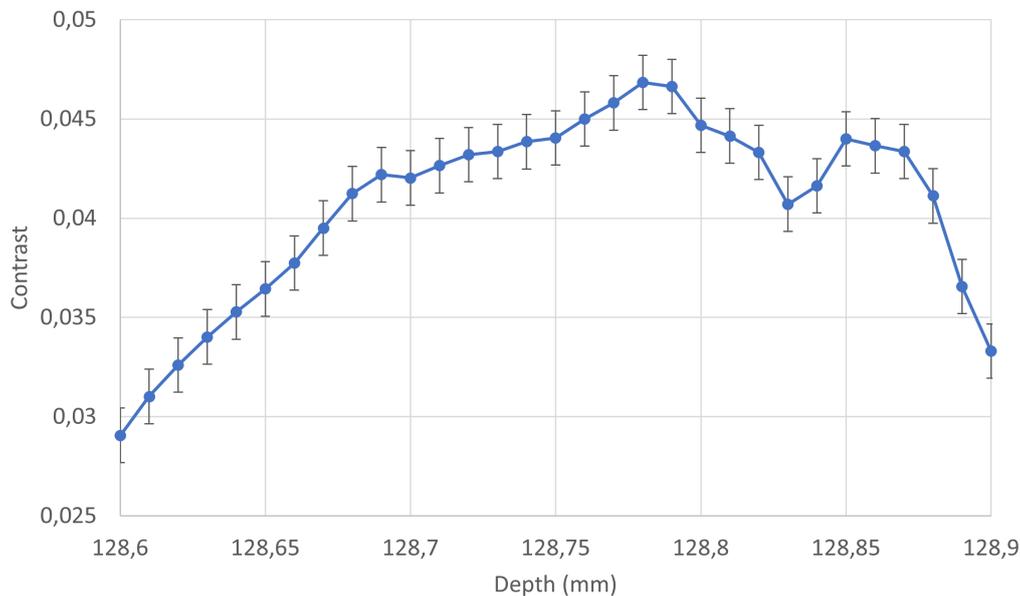


Figure 5.19 – Contrast curve extracted from the refocused stack of the X-ray plenoptic experiment.

The curve presents a plateau between $z = 128.69\text{mm}$ and $z = 128.87\text{mm}$, where the contrast value is roughly constant. The dip at $z = 128.83\text{mm}$ might be explained by misalignment effects that were already noticed in the visible (see Section 2.4.3). On both sides of the plateau, the curve decreases rapidly.

The two slopes are different, with the left slope being smoother than the right one. As studied in Section 2.5, the contrast curve is the result of combined effects of magnification, diffraction and defocus, which imply highest contrasts around both z_0^{trad} and z_0^{foc} . Contrary to previous curves obtained in the visible which peaked at z_0^{trad} (in Section 2.5), the maximal contrast in Figure 5.19 is found at z_0^{foc} , far from $z_0^{trad} = 128.36mm$. In this specific design, this is due to the effects of defocus being more important than magnification variations.

Finally, one can notice that the absolute values of the contrast are very low, with a peak at 4.68%. As a comparison, the same contrast measured in the reference images without μ LA gives a peak value of 7.47% (see Figure 5.6). The addition of the (absorbing) μ LA followed by the refocusing process are responsible for a decrease of only 37.3% of the initial signal. Thus, it is more the experimental conditions than the plenoptic and refocusing process that are responsible for the low contrast, especially the source, camera, and absorption of the main lens. These values of contrast are extremely low compared to the ones usually encountered in the visible light (around 80%).

5.3.2.6 Note on the distances

In this chapter, we presented all the experimental and refocused images using distances and configurations from Table 5.2. However, these distances have to be taken with caution. It was indeed quite difficult to measure distances during the experiment. On the one hand, the sample has to be placed at z_0^{foc} within $DOF_{pleno} = 135.27\mu m$. On the other hand, the distances are very large (more than 4 meters in total) and could not be measured at that desired precision with the available tools. As explained in Section 5.1.6, the sample, optics and camera were fixed on micrometric motors, which allow to change the distances with very fine precision. We were thus able to control very precisely the distance variations when moving the optics. The distances can be considered as very precise relative measures, but approximate absolute measures.

We tried to re-estimate the distances a posteriori. We studied the refocused stack of the two test patterns without μ LA (see Fig. 5.6) and extracted the magnification and the depths of best focus. This allows to better estimate the object-to-main lens and the main lens-to-sensor distances. We did the same with the dataset acquired with the μ LA in place, which allows to re-evaluate the main lens-to- μ LA and the μ LA-to-sensor distances. The distances presented in Table 5.2 are the results of this work.

This process allows to improve the accuracy of the measured distances but some errors remain. However in plenoptic systems, small errors in the distances can seriously impact the whole configuration, as explained in Section 2.3.3. This can be noticed by the variability in the object-to-main lens distance in configuration 2.0 throughout the document. In Table 5.2, $z_0 = 128.90mm$ is

the distance at which we wanted to place the test patterns. In Figure 5.14, the sharpest image corresponds to depth $z = 131.68mm$. With our algorithm, the focus was measured at depth $z = 128.77mm$ in Figure 5.16, much closer to the expected one. This variability in z_0 is the sign of remaining inaccuracies in the other distances of the setup. This might explain the remaining differences between experimental and theoretical values.

5.4 Depth extraction

It is now interesting to study whether depth could be extracted from the refocused stack in configuration 2.0. We apply the techniques of depth from focus studied in Chapter 4 on this stack. The experimental values of $res_{depth} = 220\mu m$ and $DOF_{pleno} = 300\mu m$ are very close, so only one depth plane is expected to be observable along DOF_{pleno} . The images were refocused along a depth range of $300\mu m$, over which the depth plane of TP1 should be localized around depth $z = 128.77mm$.

In Figure 5.20 are presented the results of two blur measures: the *Ring Difference Filter* and the *Variance*. For comparison, Figure 5.20b shows the z -axis variations in the refocused stack, extracted at the blue line in Figure 5.20a. In this figure, the different colors corresponds to the intensity in the refocused images. The large $res_{depth} = 219.58\mu m$ and $DOF_{pleno} = 135.27\mu m$ explain why the bars seem slowly affected by the defocus blur along this depth range of $300\mu m$.

Figures 5.20c - 5.20e present the results of the Ring Difference Filter without or with applying the Gaussian filtering. The Ring Difference Filter detects mainly the presence of the noise in the high depth values and fails at highlighting the correct depth $z = 128.77mm$. The Gaussian filtering emphasises the noise instead of the expected depth. A profile of Figure 5.20e is extracted at $y = 60$ and drawn in Figure 5.20g. In this profile one can notice a slight peak at the expected depth $z = 128.78mm$, but largely dominated by the rise at high depths due to the noise. This method does not seem appropriate to localize TP1.

Figures 5.20d - 5.20f - 5.20h present the same images and profile for the Variance method. A clear peak is observed at $z = 128.73mm$ in Fig. 5.20h, which corresponds to the yellow colors in Figs. 5.20d - 5.20f. The Gaussian filtering between Fig. 5.20d and Fig. 5.20f improves even more the detection of the peak for each bar, without increasing the noise. Some remaining noise is observed at high depths, but does not overpass the values at the intermediate depths. At 80% of the maximal value (as in Section 2.4.3), the curve presents a width of $80\mu m$, smaller than the expected $res_{depth} = 220\mu m$.

Compared to Chapter 4, the efficiency of the two blur measures is reversed, with the Variance method providing better results than the Ring Difference

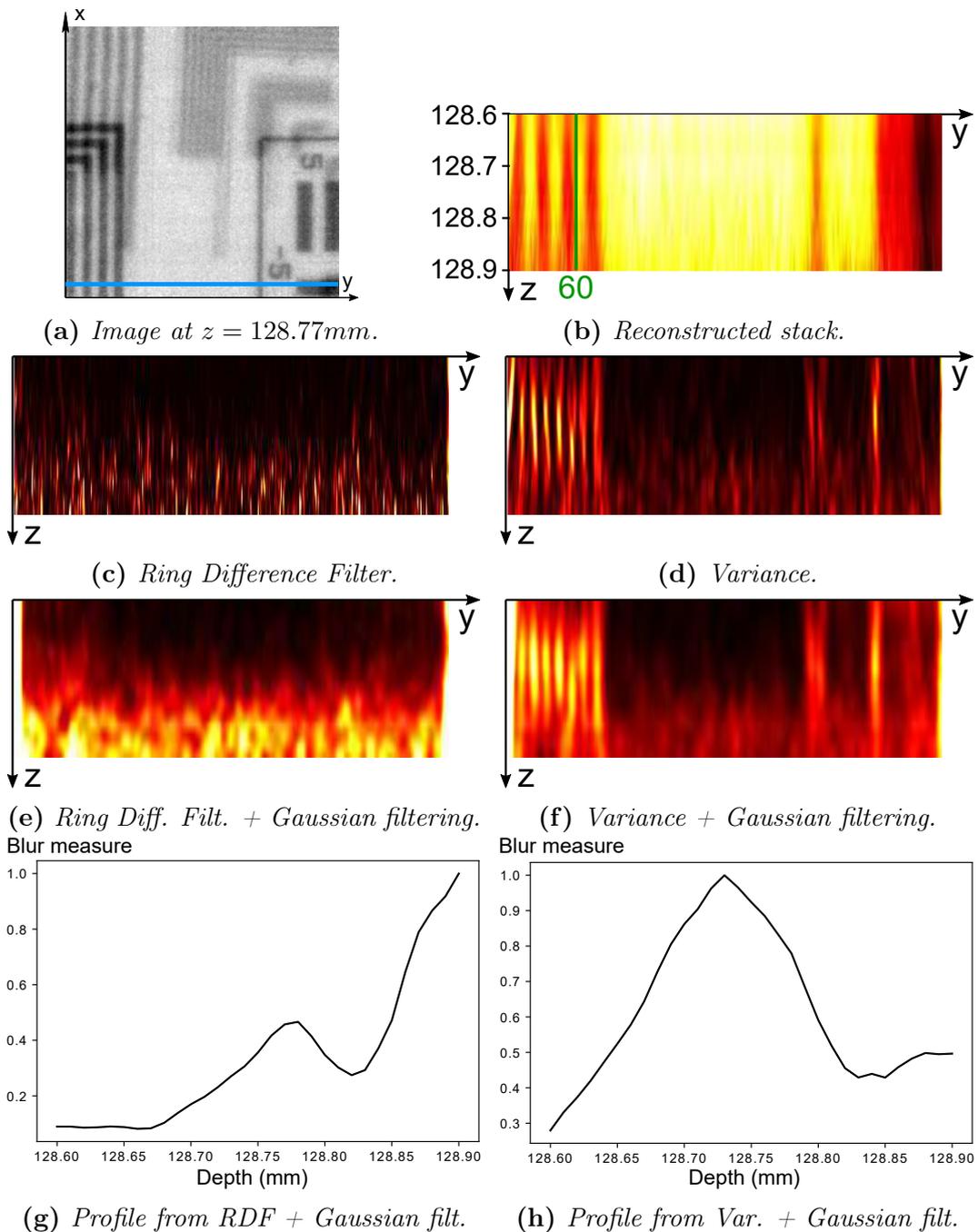


Figure 5.20 – Illustration of two blur measures applied on the refocused stack. The blue line in Figure 5.20a shows the row that was extracted in the five other images. Figure 5.20b corresponds to the z -axis variations in the reconstructed stack. Figures 5.20c and 5.20e show the results of the Ring Difference Filter with and without filtering, whereas Figures 5.20d and 5.20f show the results for the Variance method. Figures 5.20g and 5.20h present the profiles extracted from Figures 5.20e and 5.20f at $y = 60$, whose position is shown in green in Figure 5.20b.

Filter. In Chapter 4, the test case was a synthetic refocused stack, generated from several depth planes made of highly contrasted patterns, without considering noise in adjacent planes. The Ring Difference Filter succeeded at detecting the correct depth planes, due to the local aspect of this method. On the contrary, the X-ray plenoptic stack presents a high level of noise and smooth variations with depth, due to the experimental conditions, in particular the low NA . For this stack, the more global aspect of the Variance is an advantage, as it provides more reliable results less affected by the noise.

The Variance is applied on the whole image, followed by the depth extraction step. The results are presented in Figure 5.21. The displayed depth extraction method is the Max, but the Mean and Symmetry methods provided similar results. The addition of the depth extraction step allows to accurately extract the features of TP1, while removing the contribution of TP2.

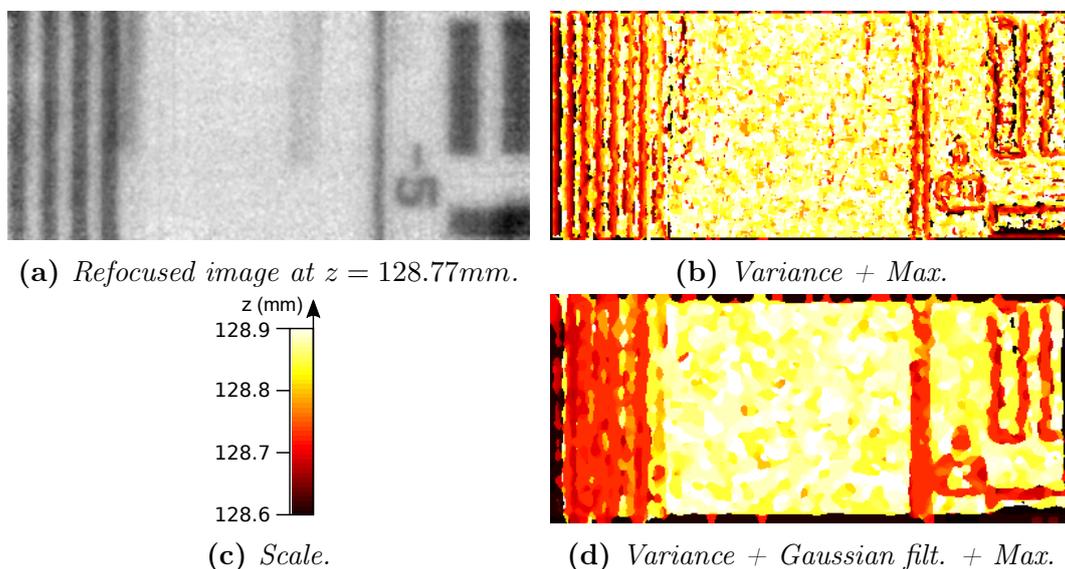


Figure 5.21 – Depth extraction of the refocused stack using the Variance method as blur measure and Max as depth extraction method: without the Gaussian filtering step (Fig. 5.21b) and with the filtering step (Fig. 5.21d). TP1 is located at $z = 128.77\text{mm}$ whereas TP2 is around $\sim z = 130\text{mm}$, outside the refocused depth range.

In Figure 5.21b, the Variance and Max methods detect the contours of the bars of the test target. It is typical of a contrast detection technique. With this method, it is still possible to recognize the number 5, and to count the number of bars on the left of the image. Some noise is still present all over the image, and especially in the middle area where no features are supposed to be detected.

Adding the Gaussian filtering helps reducing the noise in Figure 5.21d. The features of TP1 are even more clearly detected in contrast to the background. The bars, lines and number of TP1 are highlighted in orange-red, corresponding to the intermediate depths, whereas the background remains yellow and white,

which are the colors of the highest depths. However, this framework slightly decreases the resolutions, as it is not possible to visually separate the different bars on the left side. Figure 5.21d is a depth map that should be combined with the refocused stack (Figure 5.21a) to generate a 3D reconstruction with finer resolutions. These results have shown that it is possible to extract reliable depths from experimental X-ray plenoptic data. It corresponds to a significant improvement compared to the theoretical expectations ($res_{depth} > DOF_{pleno}$) and in regards to the low contrasted and noisy data.

5.5 Conclusion

In this chapter was presented the first X-ray plenoptic experiment and the reconstruction of the corresponding images, which was the final objective of the thesis. Thanks to the theoretical work presented in Chapter 2, we were able to define both 1.0 and 2.0 configurations for the plenoptic camera, taking into account limitations of both optics and X-ray source. The algorithm presented in Chapter 3 allowed to refocus images. Thanks to reliable preprocessing, the test patterns were clearly visible in the output images. Being able to see the test patterns from X-ray plenoptic data is already an important result. It shows that we overcame the difficulties of alignment and manipulation, and managed to make the X-rays go through the whole optical system, composed of a main lens *FZP*, a μ LA *FZP* and a camera. To our knowledge, such a plenoptic system with X-ray optics has not been presented in the literature yet.

We then studied the resolutions of the refocused images. The field of view, plenoptic depth of field, lateral and depth resolutions were close to the expected values. This indicates that the theoretical formulas previously validated using visible data are still relevant in X-ray plenoptic.

In this experiment, the main limiting factor was the numerical aperture of the main lens affecting the whole construction of the system. This low *NA* makes impossible to respect the aperture matching condition. It induces a low angular sampling N_u and a bad $res_{depth} / DOF_{pleno}$ ratio, which reduces the possibility to reconstruct and separate different depths from only one acquisition. These optical constraints have consequences in both experimental and refocusing processes, and can impact the optical resolutions and the quality of the refocused images.

Finally, two depth from focus techniques were applied on the refocused stack. The significant presence of noise prevented from applying the Ring Difference Filter. The Variance method allowed to retrieve the correct location of the first test pattern, thanks to its more global approach. Applying depth extraction, we succeeded in completely extracting the plane of TP1 from the background and TP2.

To summarize, this experiment has shown the possibility to acquire plenoptic images with X-ray optics and provided coherent results in this very specific context. This is a proof a concept that the plenoptic theory as well as the refocusing algorithm presented in this thesis are consistent for X-ray plenoptic imaging, even if they were originally developed in the visible. The knowledge gained during this experiment convinced us that multi-planes X-ray plenoptic and 3D depth extraction are feasible on short terms. Small modifications of the experimental setup would allow reaching this new objective in a future experiments.

Conclusion

In this thesis were presented the theoretical, numerical and experimental works that lead up to the achievement of the first X-ray plenoptic camera. More specifically, the objective of this thesis in the context of the *VOXEL* project was to deepen the knowledge on plenoptic imaging, in order to transpose it to the X-ray range. We here review the main ideas and contributions of this work to the novel field of X-ray plenoptic imaging.

Plenoptic cameras are composed of three elements: a main lens, an array of micro-lenses and a sensor. The specificity of plenoptic imaging is the ability to record the spatio-angular components of the light-field. The acquired image should be processed using dedicated algorithms to generate a refocused image at a chosen depth in object-space. Two different plenoptic cameras, the traditional and the focused ones, have been reported in the literature, each of them being associated with a specific configuration and refocusing algorithm.

In this thesis we highlighted the links between traditional and focused cameras, regarding both their optical setups and their algorithms (see Chapter 1). The two numerical methods were shown to be based on the same principle of integration of angular data, following the continuity of the optical designs. This comparison provided elements to better understand the role of the configuration and the algorithm in the resolutions of the final image. To improve the performances of a plenoptic system, we concluded that the physical setup and the numerical method should be optimized separately.

Studying both traditional and focused configurations resulted in generalized formulas to quantify the expected resolutions of a plenoptic system, depending on the distances and parameters of the physical setup (such as numerical apertures, lens diameters...) (see Chapter 2). They are useful to anticipate the finest achievable voxel and the largest observable object when building an X-ray plenoptic camera. Two visible plenoptic cameras were built to validate our theoretical work. The experiments confirmed that our geometrical reasoning was appropriate for lateral resolutions and fields of view. Estimating resolutions along depth was more complex since it necessitated a study of image blur and sharpness over the depth range. In parallel to the experiments, we implemented simulations of plenoptic imaging. They helped us understand that magnification and diffraction interfere with defocus blur. This might affect the resolutions and the process of depth extraction from the reconstructed

stack.

The previous comparison on refocusing algorithm convinced us to implement our own method (see Chapter 3). The objective was to overcome the specificities of the traditional and focused algorithms, especially their different uses of sensor pixels. We developed a *physically-based* algorithm that could be applied indifferently on each configuration, and could render an image of any chosen size. Our method is based on a full parameterization of a single ray throughout the whole plenoptic setup, leading to an equation that describes the transformation from object-space to sensor-space. Our refocused approach then uses this equation to reconstruct pixels one by one, considering the apertures of both main lens and micro-lenses, and the spatial extent of pixels. The reverse framework was also implemented to simulate raw plenoptic images. The refocusing algorithm was evaluated on the output of these simulations. The refocused images indicated that the method was accurate, producing sharp and blur at the expected depths.

Techniques of depth from focus were tested based on this blur properties (see Chapter 4). The accuracy of our algorithm was important to generate refocused stacks with consistent defocused patterns. We implemented and tested various published methods of blur measurement and depth extraction. The methods were efficient on a single plane, and give interesting results in the case of successive depth planes.

The theoretical, experimental and numerical works achieved during this thesis allowed to reach our objective of adapting plenoptic to X-rays (see Chapter 5). The formulas for resolutions were used to choose the appropriate configuration, in order to make the best use of the available X-ray optics, despite their low numerical apertures. The illumination by transmission through the sample, which is specific to X-ray, was also taken into account. Our X-ray plenoptic setup allowed to acquire the first raw plenoptic images of two overlapping test targets. Our algorithm was used to reconstruct the image at various depths. We validated the formulas of lateral and depth resolutions, field of view and plenoptic depth of field. Depth from focus techniques applied on the refocused stack allow to extract the correct depth of the first test pattern.

This thesis presented the first plenoptic images in the X-ray range, and validated the transposition of plenoptic imaging from visible to X-rays.

Future work

The work presented in this thesis could be continued and deepened. We here present a list of remaining tasks and potential improvements.

Refocusing algorithm

One of the issues raised during this thesis is the measurement errors concerning the distances of the physical setup. Our algorithm necessitates precise parameters as inputs, including the distances of the setup. In microscopy, because of the large magnifications, small errors in the distances might result in inaccuracy in the reconstruction process. This problem could be reduced by a calibration step before reconstruction. However, the calibration methods previously proposed are based on calibration grids and could hardly be transposed to X-ray imaging [DPW13; ZJW16]. In our X-ray plenoptic experiment, we tried to retrospectively refine the distances using focal stacks acquired with only the main lens and the sensor (see Chapter 5). This calibration step deserves to be improved and possibly automated for future experiments.

Another main remaining task consists in reducing the computation time of our algorithm (see Chapter 3). The current version of the code is a prototype that has been implemented in order to test the feasibility of our ideas, but is not intended for direct real-time application. The large computation time is an important drawback of the method, as it is currently 2000 times slower than the comparable methods. It is not possible to use the current implementation for real-time rendering, with several hours needed to reconstruct an image of 1000×1000 pixels. The large computation time is due to each pair of pixel-main lens aperture in object-space been projected onto each pair of pixel-micro-lens aperture in sensor-space. This results in a high number of ray bundles that needs to be projected and intersected in 4D.

A first improvement would be parallelization of the code. The algorithm presents features that justifies code parallelization, such as separability of the 4D into 2×2 D, and the independence of ray tracing through adjacent pixels or micro-lenses. The current implementation has been parallelized on CPU, but a further parallelization could be achieved using GPU. Moreover, Python was the programming language used for the implementation, which is an interpreted language. A compiled programming language such as C or C++ would help improving the execution time of the algorithm. These technical modifications are pure implementation aspects but we believe that they would have a huge impact on the execution time.

Another line of improvement is the pre-computation of the 4D intersections needed to reconstruct the image. The shapes of the ray bundles in 4D only depend on the plenoptic camera, i.e. the main lens, μ LA and sensor. The object to be imaged affects the intensity of the ray bundles, but not their 4D

patterns. When using the same plenoptic configuration for several successive acquisitions, it would be interesting to store these patterns, or more specifically the areas of their intersections that are used as weights in the algorithm. This would prevent from re-computing them several times when successive acquisitions are made using the same configuration.

Another possibility is to explore a deeper factorization of the computations. The loops have already been restricted to the absolute minimum by the calculation of precise bounds. An idea would be to study the redundancy over the whole 4D space, corresponding to redundancy in the 4D intersections to be computed. The pixels-lens apertures form a regular grid of parallelograms over a 2D plane. We thus have two overlapping regular grids, one for the object-space and one for the sensor-space. For some specific configurations, it is possible that the two grids form repeated similar intersection patterns, that could be exploited to fasten the calculation. A deeper study should be performed to determine how it is possible to detect such a redundancy directly from the configuration, through a frequency analysis for example. This idea needs to be experimentally tried to estimate its benefit, as it concerns the structure of the algorithm instead of its technical implementation.

Depth extraction

The work achieved on depth from focus techniques should be continued. Some other blur measures and depth extraction methods could be implemented and tested on our refocused stack, as only a few of them were selected for our study.

Alternatives to depth from focus approaches could be experimented, such as deconvolution of the refocused stack. Indeed, the shape of the blur in the stack is very specific because it has been generated by an algorithm with specific parameters and properties. Deconvolution would exploit the fact that the underlying algorithm is known, by using a deconvolution kernel specifically calibrated for each refocused stack. This numerical deconvolution kernel would intrinsically take into account the discrete aspect of the data, due to the pixels and the micro-lenses. We expect it to reduce the grid artifacts described in Chapter 3.

We also had the idea of a stochastic approach: rays would be randomly chosen through the scene, and then propagated until the sensor plane to retrieve the needed information. The scene would be progressively reconstructed as the number of selected rays increases. This approach would have the advantage of allowing the users to themselves choose their desired trade-off between computation time and precision of the reconstruction.

In Chapter 1, we mentioned stereo and tomography algorithms which could be adapted to extract depth directly from the raw plenoptic image. Stereo algorithms are more adapted to surfaces and require sub-images with a high number of pixels. Alternatively, tomography algorithms seem worth exploring,

as they could allow to reconstruct directly in 3D. A promising approach would be to combine our work to the one performed by Vigano et. al. during the *VOXEL* project [VDH⁺18; VMH⁺19]. A tomography algorithm was adapted to the context of plenoptic. It is based on an iterative framework that allows 2D refocusing at a chosen depth, similarly to our algorithm. A projection operator is needed to make the link between object- and sensor-spaces, similarly to matrix \mathbf{M} in our algorithm (see Chapter 3). It would be interesting to replace their projection operator by our matrix \mathbf{M} . The result is expected to combine the accuracy of our parameterization with the efficiency of their framework.

X-ray plenoptic imaging

The most exciting work lies in the advances achieved in the field of X-ray plenoptic imaging. To pursue exploration of X-ray plenoptic imaging, refocusing and extraction of several depths from one single plenoptic dataset should be performed. It would be necessary to acquire new experimental data with a different setup, either with another configuration or with other optics. During our experiment, we have seen that the low numerical apertures of X-ray optics largely restrict the possibilities of plenoptic (see Chapter 5). An experimental work on these optical components would open up possibilities for testing new configurations.

The flexibility of plenoptic, especially the focused camera (2.0), allows a variety of possible configurations, even using the same optics. Various designs could be explored to find the best trade-off between the constraints of illumination conditions, aperture matching equation, and desired resolutions. In a future experiment, the depth resolution should be smaller to allow refocusing at separate depth planes. We believe that implementing a focused Galilean configuration ($a < 0$) would help reducing the distances and increasing the sub-images size on the sensor, to approach the aperture matching condition. Larger sub-images would allow acquisition of more data, and possibly better resolutions in the reconstructed images. Noise removal is also an interesting field that could improve data treatment. It would possibly improve the results of the depth from focus methods applied on X-ray plenoptic data, as a step closer to 3D reconstruction.

The formulas of resolutions presented in Chapter 2 were used to build our X-ray plenoptic camera. However, the large number of physical parameters and distances make them difficult to manipulate. It would be very useful to create a tool guiding scientists when designing a plenoptic configuration. This tool should include both aspects of the expected resolutions and the physical limitations imposed by the setup.

The interesting results of this thesis demonstrate that X-ray plenoptic is a

promising field. Lessons learnt during this work will allow to optimize the configuration and processing in a future experiment. We expect to fully demonstrate refocusing and depth extraction of several planes on X-ray plenoptic images.

Bibliography

- [AB91] Edward H Adelson and J R Bergen. The plenoptic function and the elements of early vision. *Computational Models of Visual Processing*, 1991.
- [Att99] D. Attwood. *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*. Cambridge University Press., first edit edition, 1999.
- [AVDVU08] F. Aguet, D. Van De Ville, and M. Unser. Model-based 2.5-D deconvolution for extended depth of field in brightfield microscopy. *IEEE Transactions on Image Processing*, 17(7):1144–1153, July 2008.
- [AW92] Edward H Adelson and John Y A Wang. Single Lens Stereo with a Plenoptic Camera. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 14(2), 1992.
- [Bal08] Anna Balazs. International vocabulary of metrology-basic and general concepts and associated terms. *Chemistry International*, pages 20–1, 2008.
- [BCJ⁺13] Bastien Billiot, Frédéric Cointault, Ludovic Journaux, Jean-Claude Simon, and Pierre Gouton. 3d image acquisition system based on shape from focus technique. *Sensors*, 13(4):5040–5053, 2013.
- [BGY⁺13] Michael Broxton, Logan Grosenick, Samuel Yang, Noy Cohen, Aaron Andalman, Karl Deisseroth, and Marc Levoy. Wave optics theory and 3-D deconvolution for the light field microscope. *Optics Express*, 21(21):25418, 2013.
- [Buz11] Thorsten M Buzug. Computed tomography. In *Springer Handbook of Medical Technology*, pages 311–342. Springer, 2011.
- [CCC⁺16] Che-Wei Chang, Min-Hung Chen, Kuan-Chang Chen, Chi-Ming Yeh, and Yi-Chang Lu. Mask design for pinhole-array-based hand-held light field cameras with applications in depth estimation. In *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, pages 1–4, 12 2016.

- [CDMBV16] Kevin Cossu, Guillaume Druart, Aurélie Montmerle Bonnefois, and Marie-Thérèse Velluet. Caméras plénoptiques pour l'imagerie tridimensionnelle. *Techniques de l'ingénieur Mesures tridimensionnelles et états de surface*, base docum(ref. article : r1393), 2016.
- [CG19] Corentin Cou and Gael Guennebaud. Algorithmes de “Depth from Focus and Defocus” pour la numérisation 3D au micron. Master’s thesis, Institut d’Optique Graduate School, 2019.
- [CTDL11] Xun Cao, Xin Tong, Qionghai Dai, and Stephen Lin. High resolution multispectral video capture with a hybrid camera system. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 297 – 304, 07 2011.
- [CZL⁺19] Lixia Cao, Biao Zhang, Jian Li, Xianglei Song, Zhiyong Tang, and Chuanlong Xu. Characteristics of tomographic reconstruction of light-field tomo-piv. *Optics Communications*, 442:132 – 147, 2019.
- [DCH⁺19] J. Duarte, R. Cassin, J. Huijts, B. Iwan, F. Fortuna, L. Delbecq, H. Chapman, M. Fajardo, M. Kovacev, W. Boutu, and H. Merdji. Computed stereo lensless X-ray imaging. *Nature Photonics*, 13:449–453, 2019.
- [DDWK13] K. M. Dąbrowski, D. T. Dul, A. Wróbel, and P. Korecki. X-ray microlaminography with polycapillary optics. *Applied Physics Letters*, 102(22):224104, 2013.
- [Dob12] T. Dobbert. *Matchmoving: The Invisible Art of Camera Tracking*. Serious skills. Wiley, 2012.
- [DPW13] Donald G Dansereau, Oscar Pizarro, and Stefan B Williams. Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras. In *IEEE conference on computer vision and pattern recognition*, 2013.
- [DW88] T. Darrell and K. Wohn. Pyramid based depth from focus. In *Proceedings of IEEE: Conference on Computer Vision and Pattern Recognition*, pages 504–509, 1988.
- [DZC⁺16] Eric A Deem, Yang Zhang, Louis N Cattafesta, Timothy W Fahringer, and Brian S Thurow. On the resolution of plenoptic PIV. *Measurement Science and Technology*, 27:084003, 2016.
- [Far46] Michael Faraday. Thoughts on ray-vibrations. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 28(188):345–350, 1846.

- [FCS14] Juliet Fiss, Brian Curless, and Richard Szeliski. Refocusing Plenoptic Images using Depth-Adaptive Splatting. In *2014 IEEE International Conference on Computational Photography (ICCP)*, pages 1–9, 1-9, 2014. IEEE.
- [FH15] Yasutaka Furukawa and Carlos Hernández. Multi-view stereo: A tutorial. *Foundations and Trends® in Computer Graphics and Vision*, 9(1-2):1–148, 2015.
- [FT14] Timothy W. Fahringer and Brian S. Thurow. 3d particle position reconstruction accuracy in plenoptic piv. In *52nd Aerospace Sciences Meeting*, 01 2014.
- [FVDVB⁺04] B. Forster, D. Van De Ville, J. Berent, D. Sage, and M. Unser. Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images. *Microscopy Research and Technique*, 65(1-2):33–42, September 2004.
- [GCL11] Todor Georgiev, Georgi Chunev, and Andrew Lumsdaine. Superresolution with the focused plenoptic camera. *SPIE Electronic Imaging*, 7873:1–13, 2011.
- [Ger39] Andrey Gershun. The Light Field. *Journal of Mathematics and Physics*, 18:51–151, 1939. Originally published in Moscow in 1936.
- [GGHS03] Michael Goesele, Xavier Granier, Wolfgang Heidrich, and Hans-Peter Seidel. Accurate Light Source Acquisition and Rendering. *ACM Transactions on Graphics*, 22(3):621 – 630, 2003.
- [GGSC96] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. In *Proc. 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, pages 43–54. ACM, 1996.
- [GL09a] Todor Georgiev and Andrew Lumsdaine. Depth of Field in Plenoptic Cameras. *Eurographics 2009*, pages 5–8, 2009.
- [GL09b] Todor Georgiev and Andrew Lumsdaine. Superresolution with Plenoptic Camera 2.0. *Adobe Technical Report*, 2009.
- [GL10a] Todor Georgiev and Andrew Lumsdaine. Focused plenoptic camera and rendering. *Journal of Electronic Imaging*, 19(2):021106, 2010.
- [GL10b] Todor Georgiev and Andrew Lumsdaine. Reducing plenoptic camera artifacts. *Computer Graphics Forum*, 29(6):1955–1968, 2010.
- [Gro05] Herbert Gross. *Handbook of Optical Systems, Volume 1, Fundamentals of Technical Optics*, volume 1. Wiley-VCH, 2005.

- [HA10] Umer Hassan and Sabieh Anwar. Reducing noise by repetition: Introduction to signal averaging. *European Journal of Physics*, 31:453, 03 2010.
- [HAV⁺16] Christopher Hahne, Amar Aggoun, Vladan Velisavljevic, Susanne Fiebig, and Matthias Pesch. Refocusing distance of a standard plenoptic camera. *Optics Express*, 24(19):21521, 9 2016.
- [Hec17] E. Hecht. *Optics*. Pearson Education, fifth edit edition, 2017.
- [Hog18] Matthieu Hog. *Light Field Editing and Rendering*. Thesis, Rennes 1, November 2018.
- [HSVD17] Matthieu Hog, Neus Sabater, Benoit Vandame, and Valter Drazic. An Image Rendering Pipeline for Focused Plenoptic Cameras. *IEEE Transactions on Computational Imaging*, 3(4):811–821, 12 2017.
- [IO94] S Inoué and R Oldenbourg. Optical instruments: microscopes handbook of optics 2nd edn, vol 2 ed m bass, 1994.
- [IRMD16] Ivo Ihrke, John Restrepo, and Lois Mignard-Debise. Principles of Light Field Imaging: Briefly revisiting 25 years of research. *IEEE Signal Processing Magazine, Institute of Electrical and Electronics Engineers*, 33(5):59–69, 2016.
- [Ive03] Frederic E Ives. Parallax stereogram and process of making same, 1903. US Patent 725567.
- [Kop14] Sanjeev J. Koppal. *Lambertian Reflectance*, pages 441–443. Springer US, Boston, MA, 2014.
- [Kro88] Eric Krotkov. Focusing. *International Journal of Computer Vision*, 1(3):223–237, 1988.
- [KSW02] Avinash C Kak, Malcolm Slaney, and Ge Wang. Principles of computerized tomographic imaging. *Medical Physics*, 29(1):107–107, 2002.
- [Lam15] Edmund Y. Lam. Computational photography with plenoptic camera and light field capture: tutorial. *Journal of the Optical Society of America A*, 32(11):2021–2032, 11 2015.
- [LG08a] Andrew Lumsdaine and Todor Georgiev. Full resolution lightfield rendering. Technical report, Indiana University and Adobe Systems, 2008.
- [LG08b] Andrew Lumsdaine and Todor Georgiev. Full Resolution Lightfield Rendering. In *Adobe Tech Report*, pages 1–12, 2008.

- [LG09] Andrew Lumsdaine and Todor Georgiev. The focused plenoptic camera. *IEEE International Conference on Computational Photography (ICCP)*, pages 1–8, 2009.
- [LGC12] Andrew Lumsdaine, Todor Georgiev, and Georgi Chunev. Spatial analysis of discrete plenoptic sampling. In *SPIE Electronic Imaging*, volume 8299, page 829909, 2012.
- [LH96] Marc Levoy and Pat Hanrahan. Light Field Rendering. *Proceedings of SIGGRAPH*, pages 31–42, 1996.
- [Lip08a] G. Lippmann. Épreuves réversibles donnant la sensation du relief. *J. Phys. Theor. Appl.*, 7(4):821–825, 1908.
- [Lip08b] G. Lippmann. Epreuves réversibles photographies intégrales. *CR Acad. Sci*, 146:446–451, 1908.
- [LNA⁺06] Marc Levoy, Ren Ng, Andrew Adams, Matthew Footer, and Mark Horowitz. Light Field Microscopy. *ACM Transactions on Graphics*, 25(3):924–934, 2006.
- [LR15] Chia-Kai Liang and Ravi Ramamoorthi. A Light Transport Framework for Lenslet Light Field Cameras. *ACM Transactions on Graphics*, 34(2):1–19, 3 2015.
- [McC14] William Ross McCluney. *Introduction to radiometry and photometry*. Artech House, 2014.
- [MCJ18] Manuel Martínez-Corral and Bahram Javidi. Fundamentals of 3D imaging and displays: a tutorial on integral imaging, light-field, and plenoptic systems. *Advances in Optics and Photonics*, 10(3):512–566, 2018.
- [MD18] Loïs Mignard-Debise. Tools for the paraxial optical design of light field imaging systems. *Ph.D. thesis, Université de Bordeaux*, 2018.
- [MDI15] Lois Mignard-Debise and Ivo Ihrke. Light-field Microscopy with a Consumer Light-field Camera. In *3D Vision 2015*, pages 335–343, oct 2015.
- [MDP13] N. Marturi, S. Dembélé, and N. Piat. Depth and shape estimation from focus in scanning electron microscope for micromanipulation. In *2013 International Conference on Control, Automation, Robotics and Embedded Systems (CARE)*, pages 1–6, 2013.
- [MDRI17] Lois Mignard-Debise, John Restrepo, and Ivo Ihrke. A Unifying First-Order Model for Light-Field Cameras: The Equivalent Camera Array. *IEEE Transactions on Computational Imaging*, 3(4):798–810, 2017.

- [MKV⁺14] Istvan Mohacsi, Petri Karvinen, Ismo Vartiainen, Vitaliy A Guzenko, Andrea Somogyi, Cameron M Kewish, Pascal Mercere, and Christian David. High-efficiency zone-plate optics for multi-keV x-ray focusing. *Journal of synchrotron radiation*, 21(3):497–501, 2014.
- [MRK⁺13] Alkhazur Manakov, John Restrepo, Oliver Klehm, Ramon Hegedüs, Elmar Eisemann, Hans-Peter Seidel, and Ivo Ihrke. A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging. *ACM Transactions on Graphics (TOG)*, 32(4):1–47, 07 2013.
- [Ng06] Ren Ng. *Digital Light Field Photography*. PhD thesis, Stanford University, CA, USA, 2006. AAI3219345.
- [PD84] Thomas Porter and Tom Duff. Compositing digital images. *ACM SIGGRAPH Computer Graphics*, 18(3):253–259, 1984.
- [PI] PETRA-III. P05 - imaging beamline. https://photon-science.desy.de/facilities/petra_iii/beamlines/p05_imaging_beamline/unified_data_sheet_p05/index_eng.html. Accessed: 2020-07-12.
- [PW12] Christian Perwaß and Lennart Wietzke. Single Lens 3D-Camera with Extended Depth-of-Field. *Proc. SPIE*, 8291:4–, 2012.
- [SCN93] M Subbarao, T S Choi, and A Nikzad. Focusing techniques. *J. Optical Engineering*, 32(11):2824–2836, 1993.
- [SDN04] Yu Sun, Stefan Duthaler, and Bradley J Nelson. Autofocusing Algorithm Selection in Computer Microscopy. *Microscopy research and technique*, 65(3):139–49, 2004.
- [Sha49] C. E. Shannon. Communication in the presence of noise. *Proceedings of the Institute of Radio Engineers*, 37(1):10–21, 1949.
- [SJK18] Katarzyna M. Sowa, Benedykt R. Jany, and Paweł Korecki. Multipoint-projection x-ray microscopy. *Optica*, 5(5):577–582, May 2018.
- [SJP⁺17] Jaeheung Surh, Hae-Gon Jeon, Yunwon Park, Sunghoon Im, Hyowon Ha, and In So Kweon. Noise Robust Depth from Focus using a Ring Difference Filter. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6328–6337, 2017.
- [SKK20] Katarzyna M. Sowa, Marcin P. Kujda, and Paweł Korecki. Plenoptic x-ray microscopy. *Applied Physics Letters*, 116(1):014103, 2020.
- [SKSL98] A. Snigirev, Victor Kohn, Irina Snigireva, and B. Lengeler. A compound refractive lens for focusing high-energy x-rays. *Nature*, 384, 01 1998.

- [SN17] Parikshit Sakurikar and P J Narayanan. Composite Focus Measure for High Quality Depth Maps. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1614–1622, 2017.
- [SOdSV⁺97] Andrés Santos, C Ortiz de Solórzano, Juan José Vaquero, Javier Márquez Pena, Norberto Malpica, and Francisco del Pozo. Evaluation of autofocus functions in molecular cytogenetic analysis. *Journal of microscopy*, 188(3):264–272, 1997.
- [Spa16] C. M. Sparrow. On Spectroscopic Resolving Power. , 44:76, September 1916.
- [THMR13] Michael W. Tao, Sunil Hadap, Jitendra Malik, and Ravi Ramamoorthi. Depth from Combining Defocus and Correspondence Using Light-Field Cameras. *2013 IEEE International Conference on Computer Vision*, pages 673–680, 12 2013.
- [Tur16] Massimo Turola. Investigation of plenoptic imaging systems: a wave optics approach. *Ph.D. thesis, City University London*, 2016.
- [TV01] A. C. Thompson and D. Vaughan. *X-ray data booklet*, volume 8. Lawrence Berkeley National Laboratory, University of California., 2001.
- [VDH⁺18] Nicola Viganò, Henri Der Sarkissian, Charlotte Herzog, Ombeline de La Rochefoucauld, Robert van Liere, and Kees Joost Batenburg. Tomographic approach for the quantitative scene reconstruction from light field images. *Optics Express*, 26(18):22574–22602, sep 2018.
- [VMH⁺19] Nicola Viganò, Pablo Martínez Gil, Charlotte Herzog, Ombeline de la Rochefoucauld, Robert van Liere, and Kees Joost Batenburg. Advanced light-field refocusing through tomographic modeling of the photographed scene. *Optics Express*, 27(6):7834–7856, mar 2019.
- [VRA⁺07] Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, and Jack Tumblin. Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing. *ACM Transactions on Graphics*, 26(3), 2007.
- [WER15] Ting-Chun Wang, Alexei A Efros, and Ravi Ramamoorthi. Occlusion-aware Depth Estimation Using Light-field Cameras. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3487–3495, 2015.
- [WH02] Taufiq Widjanarko and Russell C Hardie. A post-processing technique for extending depth of focus in conventional optical microscopy. *Optics & Laser Technology*, 34(4):299–305, 2002.

- [WILH11] Gordon Wetzstein, Ivo Ihrke, Douglas Lanman, and Wolfgang Heidrich. Computational Plenoptic Imaging. *Computer Graphics Forum*, 2011.
- [WJV⁺05] Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, and Marc Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics*, 24(3):765–776, 2005.
- [WN96] Masahiro Watanabe and Shree K Nayar. Telecentric optics for computational vision. In *European Conference on Computer Vision*, pages 439–451. Springer, 1996.
- [YOJS93] TTE Yeo, SH Ong, Jayasooriah, and R Sinniah. Autofocusing for tissue microscopy. *Image and Vision Computing*, 11(10):629 – 639, 1993.
- [ZJW16] Chunping Zhang, Zhe Ji, and Qing Wang. Decoding and calibration method on focused plenoptic camera. *Computational Visual Media*, 2(1):57–69, 2016.
- [ZLE⁺18] Shuaishuai Zhu, Andy Lai, Katherine Eaton, Peng Jin, and Liang Gao. On the fundamental comparison between unfocused and focused light field cameras. *Applied Optics*, 57(1):A1–A11, jan 2018.