Tools for the paraxial optical design of light field imaging systems

Lois Mignard-Debise

THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET D'INFORMATIQUE

par **Loïs Mignard--Debise**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : INFORMATIQUE

Outils de Conception en Optique Paraxiale pour les Systèmes d'Imagerie Plénoptique

---

Date de soutenance : 5 Février 2018

Devant la commission d'examen composée de :

Hendrik LENSCH	Professeur, Université de Tübingen	Rapporteur
Céline LOSCOS .	Professeur, Université de Reims	Présidente
Ivo IHRKE	Professeur, Inria	Encadrant
Patrick REUTER .	Maître de conférences, Université de Bordeaux	Encadrant
Xavier GRANIER	Professeur, Institut d'Optique Graduate School	Directeur

– 2018 –

Acknowledgments

I would like to express my very great appreciation to my research supervisor, Ivo Ihrke, for his great quality as a mentor. He asked unexpected but pertinent questions during fruitful discussions, encouraged and guided me to be more autonomous in my work and participated in the solving of the hardships faced during experimenting in the laboratory. I would like to offer my special thanks to Patrick Reuter, my second research supervisor, who let me the freedom to develop my own research path even when it strayed away from his own focus.

I am thankful to Prof. François Goudail and Prof. Henrik Lensch to have accepted to be the reviewers for the manuscript of this thesis and I am grateful to Prof. Céline Loscos to have agreed to be the president of the committee of my thesis defense despite the short notice.

I wish to express my gratitude to Prof. Xavier Granier, the director of the MANAO team, for the considerable dedication he has exhibited for the team. I would also like to thank all my colleagues at the MANAO team, permanent and non-permanent members, for the heartwarming welcome they showed me and for the deep interest they have manifested about my work. It was a pleasure to have such passionate, scientific or not, debates that animated the conversations during my stay. My thanks also go to Anne-Laure Gautier, the team assistant, for her great help concerning all the administrative matters.

I wish to thank the Ph.D. students that I met through all these years whether they have already graduated or are yet to be: Boris Raymond, Brett Ridel, Carlos Zubiaga Peña, Arthur Dufay, Thibaud Lambert, Antoine Lucat, Thomas Crespel, David Murray, Charlotte Herzog and the ones I forget to mention. I strongly appreciated your support, united in the adversity that a doctorate sometimes is and I loved sharing all of the good moments that counterbalanced this struggle with you. I am particularly grateful to John Restrepo-Aguilero for the great assistance he provided in the laboratory. In general, I would like to thank all the people I met at Inria Bordeaux Sud-Ouest and the LP2N.

I also wish to acknowledge the precedent researchers that worked on the development of light field imaging systems for the bright ideas they had at their time. And I wish that my work will help other fellow researchers create even more innovative systems.

My Ph.D. work was made possible thanks to the support of the German Research Foundation (DFG) through the Emmy-Noether fellowship IH 114/1-1 and the Agence National de la Recherche (ANR) through the Interactive Spatial Augmented Reality (ISAR) project.

At last, I would like to thank my family and my friends for their unwavering support during these past years.

Résumé L'imagerie plénoptique est souvent présentée comme une révolution par rapport à l'imagerie standard. En effet, elle apporte plus de contrôle à l'utilisateur sur l'image finale puisque les dimensions spatiales et angulaires du champ de lumière offrent la possibilité de changer le point de vue ou de refaire la mise au point après coup ainsi que de calculer la carte de profondeur de la scène. Cependant, cela complique le travail du concepteur optique du système pour deux raisons. La première est qu'il existe une multitude d'appareils de capture plénoptique différents, chacun avec sa propre spécificité. La deuxième est qu'il n'existe pas de modèle qui relie le design de la caméra à ses propriétés optiques d'acquisition et qui puisse guider le concepteur dans sa tâche.

Cette thèse répond à ces observations en proposant un modèle optique du premier ordre pour représenter n'importe quel appareil d'acquisition plénoptique. Ce modèle abstrait une caméra plénoptique par un réseau équivalent de caméras virtuelles existant en espace objet et qui effectue un échantillonnage identique de la scène. Ce modèle est utilisé pour étudier et comparer plusieurs caméras plénoptiques ainsi qu'un microscope plénoptique monté en laboratoire, ce qui révèle des lignes directrices pour la conception de systèmes plénoptiques. Les simulations du modèle sont aussi validées par l'expérimentation avec une caméra et le microscope plénoptique.

Mots-clés conception optique, imagerie plénoptique, caméra plénoptique, microscope plénoptique

Laboratoire d'accueil Équipe Manao, Inria Bordeaux Sud-Ouest, 200 avenue de la Vieille Tour, 33405, Talence. LP2N, Institut d'Optique d'Aquitaine, Rue François Mitterrand, 33400, Talence.

Résumé détaillé

Introduction

L'imagerie plénoptique (*light field* en anglais) est un domaine qui se situe à la frontière de l'optique et de l'informatique. Elle dépasse l'imagerie classique et redéfinit la conception optique. Cette dissertation propose et explore des outils pour faciliter la conception de systèmes optiques tels que les caméras et les microscopes plénoptiques.

Contexte

En imagerie classique, le système optique réalise l'image en deux dimensions d'une scène sur un détecteur qui est ensuite améliorée par des algorithmes de traitement d'images pour produire l'image finale à l'utilisateur.

Avec l'imagerie plénoptique, le rôle des optiques est de guider la lumière jusqu'au détecteur alors que celui des algorithmes est d'interpréter la valeur des pixels. D'un côté, les algorithmes nécessitent un modèle précis des propriétés d'imagerie du système optique pour générer correctement les nouvelles images. De l'autre côté, les optiques doivent être conçues avec les limitations des algorithmes de reconstruction à l'esprit.

Concevoir une caméra plénoptique est essentiellement un travail de co-conception simultanée de la partie matérielle et logicielle. La différence majeure avec l'imagerie conventionnelle est que l'image brute intermédiaire est destinée seulement à être la donnée d'entrée pour un algorithme spécifique et que l'utilisateur n'a accès qu'à l'image résultante du calcul de cet algorithme. Par conséquent, les contraintes sur la partie matérielle sont amoindries de telle façon qu'elle peut enregistrer des données à plus de dimensions que le simple échantillonnage spatial de la scène, comme la direction d'incidence de la lumière sur le capteur. La charge de calcul sur la partie logicielle du système est plus importante qu'avant car il est nécessaire d'extraire et de présenter des données compréhensibles à l'utilisateur mais son rôle est crucial. Au final, l'utilisateur peut naviguer à travers ce jeu de données multidimensionnel et synthétiser des images qu'aucun système d'imagerie traditionnelle ne pourrait produire seul.

De fait, l'imagerie plénoptique est un sous-groupe du domaine plus large qu'est l'imagerie computationnelle regroupant toutes les façons qui existent d'effectuer l'association entre un système d'acquisition optique et du calcul algorithmique. Elle englobe de nombreuses et diverses techniques telles que l'imagerie par temps de vol, l'illumination structurée et l'imagerie plénoptique, pour n'en nommer que quelques-unes.

Motivation

L'imagerie plénoptique enregistre sur le détecteur la position et la direction des rayons lumineux. Grâce à cette nouvelle information directionnelle, la synthèse d'images après capture permet de simuler différents paramètres de lentille et d'effectuer plusieurs actions telles que: la mise au point de l'image, l'extraction de la profondeur des objets de la scène, la modification du point de vue, de la profondeur de champ et de la longueur focale entre autres.

Les caméras plénoptiques offrent de nouvelles possibilités dans un large panel de situations. Des produits commerciaux sont destinés à des photographes professionnelles ou amateurs et les encouragent à travailler différemment puisqu'ils peuvent alors composer des photos interactives. Le contrôle de la mise au point et du point de vue est laissé à l'utilisateur et leur permet d'atteindre un niveau d'expérience similaire à ce que le public aime à associer au terme «holographie».

Dans l'industrie du cinéma, les caméras plénoptiques simplifient la production d'effets spéciaux grâce à l'information de profondeur qu'elles apportent à la scène. Le rendu stéréo est aussi possible sans effort additionnel, que ce soit pour le cinéma 3D ou encore pour la création de contenu pour la réalité augmentée et virtuelle.

Dans le reste de l'industrie, l'avantage principal des caméras plénoptiques sur les systèmes d'acquisition stéréoscopique et multivue est leur taille réduite qui les rend intéressantes dans des conditions difficiles d'utilisation telles que celles rencontrées en robotique.

Lorsqu'elle est appliquée au domaine de la microscopie où les contraintes optiques du système ne produisent que l'image d'une fine tranche des échantillons observés, l'imagerie plénoptique étend la profondeur de champ capturée par l'instrument et permet de reconstruire la structure tridimensionnelle de ces échantillons.

Dans le but de faire face à la diversité des applications, il existe indubitablement le besoin d'une meilleure compréhension des aspects de conception des systèmes d'imagerie plénoptique que cette dissertation vise à satisfaire.

Problématiques et contributions

Il existe une grande variété d'architectures de caméras plénoptiques, basées sur l'utilisation d'un composant optique spécifique tel que la matrice de lentilles. Cependant, la conception de ces architectures est parfois tellement différente qu'il est difficile d'évaluer laquelle correspond le mieux à une certaine application. Bien que les spécificités de chaque caméra soient connues, leur compréhension est restreinte à leur système particulier.

Nous avons observé qu'un modèle global qui permettrait l'évaluation et la comparaison de différentes combinaisons optiques manque. Idéalement, ce modèle serait complet, c'est-à-dire qu'il intégrerait à la fois les aspects de conception optique et algorithmique, son objectif étant d'optimiser l'équilibre entre l'échantillonnage spatial et angulaire du système. Ce modèle devrait être aussi capable de simuler physiquement la disposition de n'importe quel appareil d'acquisition plénoptique et

de produire un certain nombre de propriétés relatives au dispositif afin d'aider et de guider le travail de conception.

Dans cette dissertation, nous proposons de réaliser un premier pas dans la direction de ce modèle idéal. Les contributions majeures peuvent se formuler comme il suit:

- le développement d'un formalisme mathématique, alternatif à l'optique matricielle, basé sur l'image de points et non de rayons;
- un modèle d'optique paraxiale qui abstrait un système plénoptique à son réseau équivalent de caméras virtuelles (*Equivalent Camera Array* en anglais, *ECA* pour l'acronyme) en espace objet;
- la description des propriétés optiques de ce réseau de caméras;
- la dérivation d'expressions analytiques pour ces propriétés dans le cas des caméras plénoptiques utilisant un réseau de microlentilles;
- l'étude comparative entre deux modèles de caméra plénoptique utilisant un réseau de microlentilles;
- la validation du modèle ECA pour un modèle de caméra plénoptique utilisant un réseau de microlentilles;
- l'étude de l'emploi d'une caméra plénoptique pour la microscopie plénoptique et le développement d'un montage de macrophotographie plénoptique;
- la réalisation d'un montage de microscope plénoptique et de son analyse avec le modèle ECA;
- une implémentation du modèle ECA avec MATLAB.

Le modèle que nous proposons considère l'imagerie dans son approximation paraxiale et ne considère pas les aberrations introduites par la forme réelle des surfaces des composants optiques. De plus, notre modèle ne considère pas non plus les effets de diffraction dus à la nature ondulatoire de la lumière.

Une partie des travaux présentés dans cette dissertation a été publié dans des conférences ou des journaux [Mignard-Debise and Ihrke, 2015; Ihrke *et al.*, 2016; Mignard-Debise *et al.*, 2017]. Ils ont aussi été présentés à l'occasion de deux GdR ISIS (Groupement de Recherche Information, Signal, Image et ViSion): «Systèmes de Vision Grand Angle, Multi-Caméra et Plénoptique» et «Co-conception de systèmes hybrides : quand l'instrumentation et les traitements numériques se rencontrent». Cette dissertation inclut également des travaux non encore publiés qui détaillent et prolongent ces publications.

Vue d'ensemble de la dissertation

La dissertation est organisée en quatre parties. La première partie incluant les Chapitres 1 et 2, est dédiée à une présentation des connaissances requises à la compréhension des notions utilisées dans les parties suivantes. Le Chapitre 1 regroupe le savoir théorique à propos de fonction plénoptique, de conception optique et de microscopie. Le Chapitre 1 détaille l'état de l'art sur l'imagerie plénoptique, de la capture à la synthèse d'images en passant par la microscopie plénoptique, ainsi que

les nombreuses architectures de caméras plénoptiques présentes dans la littérature.

La deuxième partie se concentre sur le modèle théorique proposé pour la conception de systèmes plénoptiques dans les Chapitres 3 et 4. Le Chapitre 3 introduit un outil matriciel reposant sur l'imagerie de points comme une approche alternative aux matrices ABCD basées sur l'image de rayons lumineux et utilisées pour décrire la propagation de la lumière dans un système optique. Cet outil est ensuite utilisé dans le Chapitre 4 qui développe les aspects théoriques du modèle de réseau équivalent de caméras. Ces aspects comprennent la procédure à suivre pour construire l'ECA correspondant à un système plénoptique, la description des propriétés optiques importantes de ce réseau, ainsi que des considérations de vignettage.

La troisième partie couvre des applications du modèle de réseau équivalent de caméras à deux types d'appareils plénoptiques: les caméras plénoptiques employant des microlentilles dans le Chapitre 5 et la microscopie plénoptique dans le Chapitre 6. Le Chapitre 5 examine les similarités et les différences entre deux configurations et dérive des formules analytiques pour les propriétés de leur ECA. Ce chapitre valide également le modèle ECA par l'expérience avec la caméra plénoptique afocale. Le Chapitre 6 est séparé en deux parties principales. La première partie explore la possibilité d'associer une caméra plénoptique du marché grand public avec un objectif de microscope et présente des solutions alternatives pour la macrographie plénoptique. La deuxième partie applique le modèle ECA à un autre montage de microscopie plénoptique pour évaluer ses performances d'imagerie.

La quatrième et dernière partie est constituée seulement du Chapitre 7 et propose une implémentation du modèle ECA avec MATLAB en s'appuyant sur les outils d'imagerie du Chapitre 3. Il présente aussi un algorithme pour calculer l'étendue d'un système optique en considérant le vignettage par les ouvertures des composants optiques.

En matière de chronologie et de publications, la première section du Chapitre 6 est une version révisée de [Mignard-Debise and Ihrke, 2015]. Ce travail examine les effets de vignettage lorsque deux systèmes optiques différents sont associés. Il fut le leitmotiv du développement du modèle ECA.

Le contenu des Chapitres 3, 4, et 5 est une version révisée de [Mignard-Debise *et al.*, 2017]. La partie expérimentale du Chapitre 5 fut réalisée avec l'aide de John Restrepo, un autre doctorant.

La seconde partie du Chapitre 6 et la totalité du Chapitre 7 sont des travaux non publiés à l'heure de la rédaction de cette dissertation.

Conclusion

L'imagerie plénoptique propose un nouveau paradigme par rapport à l'imagerie traditionnelle. Par conséquent, un grand nombre d'appareils d'acquisition plénoptique existe, avec autant de façons différentes de capturer les composantes spatiales et angulaires des rayons lumineux. Cela étant, la littérature suggère implicitement

que ces appareils partagent tous un aspect similaire. En un sens, l'échantillonnage de la lumière qu'ils effectuent est équivalent à celui d'une matrice de caméras. Cette dissertation a démontré cette affirmation en supposant un régime paraxial. Un modèle théorique en a découlé ainsi que le moyen de l'utiliser pour en déduire des lignes de conduite quant à la conception optique d'appareils plénoptiques. Ce modèle ouvre la porte à la conception de meilleurs systèmes optiques, optimisés pour un grand nombre d'applications depuis le marché du grand public à celui de l'industrie, et à des domaines allant de la robotique à la biologie.

Récapitulatif

Dans cette dissertation, nous avons fourni une solution pour l'analyse de la conception optique des systèmes d'acquisition plénoptique tels que les caméras et les microscopes plénoptiques. Cette solution, le modèle de réseau équivalent de caméras, fonctionne dans un régime d'imagerie paraxiale. Le système est décomposé à son plus bas niveau en un ensemble de primitives formées de l'ouverture d'un pixel et de l'ouverture d'un élément de l'unité de multiplexage directionnel (UMD). Ces «éléments de deux ouvertures» sont ensuite imagés jusqu'à l'espace objet du système optique et réorganisés en un réseau de caméras virtuelles. L'avantage principal de ce modèle comparé aux précédents est qu'il conserve les surfaces des pixels et des éléments de l'UMD qui définissent véritablement le noyau élémentaire d'échantillonnage plénoptique.

Du fait de la nécessité d'imager un grand nombre de ces primitives (les éléments de deux ouvertures) à travers les composants du système, nous avons développé un formalisme mathématique fondé sur des matrices imageantes de points. Ce formalisme matriciel exploite le fait qu'imager à travers un composant optique tel qu'une lentille connecte les points objets et leurs images conjuguées par une projection linéaire de l'espace 3D. Ce système bénéficie, en sus, du concept des matrices homogènes provenant du domaine du rendu graphique pour représenter la projection imageante avec une matrice 4 par 4 ainsi que les décentrements des composants optiques par une composition directe avec des matrices de transformation géométriques. Bien que ces outils ignorent la propagation de rayons lumineux, leurs résultats sont identiques à ceux des techniques standards telles que les matrices ABCD agissant sur les rayons dans les conditions de l'approximation paraxiale. La conversion entre points et rayons est aisément effectuée.

Le modèle ECA considère aussi les effets de vignettage causés par les autres ouvertures dans le système optique et leurs influences sur le nombre et la position des pupilles et des pixels des caméras virtuelles. Nous avons présenté un moyen d'évaluer le vignettage de tous les éléments de deux ouvertures grâce aux propriétés de l'espace de phase optique. Par la suite, le modèle établit un ensemble de propriétés

optiques en espace objet pour chacune des caméras virtuelles ainsi que pour l'ensemble du réseau équivalent. Nous avons choisi et défini ces propriétés car elles servent d'indicateurs de performance du système optique et permettent de guider le concepteur vers l'application visée.

Nous avons appliqué le modèle ECA à deux configurations de caméras plénoptiques basées sur l'emploi d'un réseau de microlentilles. Ces configurations diffèrent uniquement de par la distance qui sépare le détecteur du réseau de microlentilles. Nous avons dérivé les formules analytiques de leurs propriétés optiques afin de les comparer et d'analyser les conséquences de cette légère modification de conception sur leurs performances optiques. Nous avons également validé le modèle avec une caméra plénoptique commerciale. Cette validation a été réalisée en récupérant les paramètres optiques de l'objectif principal de la caméra à partir d'une régression de plusieurs propriétés des caméras virtuelles mesurées expérimentalement. La régression a résulté en une bonne estimation des paramètres de l'objectif bien que le système soit sujet à de nombreuses imprécisions et aberrations qui dévient de notre modèle paraxial.

Cette dissertation a également examiné l'imagerie plénoptique pour des applications en microscopie. Nous avons d'abord étudié l'association d'une caméra plénoptique commerciale et d'un objectif de microscope. Nous avons montré que l'incompatibilité entre leur nombre d'ouverture causait un important vignettage spatial et angulaire. Bien que nous ayons réussi à utiliser la caméra plénoptique dans un mode d'imagerie inattendu qui a permis de réduire la perte de champ, le vignettage angulaire empêche l'emploi du système à un grand facteur de grandissement. Par conséquent, nous avons réalisé et testé plusieurs systèmes en remplaçant l'objectif de microscope par celui d'un appareil photomacrographique SLR. Avec ces systèmes, nous avons prouvé que la macrographie plénoptique était possible en reconstruisant la profondeur dans un volume cubique de quelques millimètres de longueur. Cependant, notre montage produit certainement de moins bon résultat qu'un système spécifiquement conçu pour cette application macrographique. Puis, nous avons construit et analysé notre propre microscope plénoptique. Nous avons débattu des règles qui sous-tendent sa conception et nous avons utilisé ce système pour valider une fois de plus le modèle ECA.

En dernier, nous avons proposé une implémentation logicielle du modèle ECA avec la boîte à outils GeODe. Elle regroupe des aspects piochés dans les logiciels standards de design optique tels que l'imagerie séquentielle, et dans le domaine du rendu graphique tels que les graphes de scène pour gérer le positionnement des composants optiques. Ces aspects sont combinés entre eux afin de représenter l'imagerie dans les systèmes plénoptiques et de simuler son réseau équivalent de caméras et ses propriétés. Nous avons également décrit en détail comment calculer la fonction

de vignettage associée à chaque élément de deux ouvertures, et fourni des façons d'optimiser ce calcul.

Travaux futurs

Le modèle de réseau équivalent de caméras que nous avons présenté est un modèle optique au premier ordre et, en tant que tel, il ignore deux aspects principaux d'imagerie: les aberrations et la diffraction. En conception optique, l'analyse au premier ordre sert de brouillon pour définir les fonctions majeures du système optique. Puis, le système est raffiné en utilisant de véritables surfaces réflectives ou réfractives, analysé avec un tracé de rayon et optimisé pour réduire les aberrations. Les aberrations et la diffraction sont présentes dans tout système optique. Dans la continuité du modèle ECA, l'objectif premier serait de développer une théorie des aberrations pour les systèmes plénoptiques qui conserve, si possible, la structure du réseau de caméras et attribue à chacune sa propre fonction d'aberrations. Le second objectif serait d'inclure en plus un modèle de diffraction aussi en lien avec le modèle ECA. Avec ces deux extensions, le modèle ECA décrirait de façon précise l'imagerie des systèmes d'acquisition plénoptique.

Nous avons vu comment le vignettage à l'intérieur des systèmes plénoptiques modifiait la disposition des caméras virtuelles ainsi que l'image blanche servant à la calibration. Un axe de recherche intéressant serait d'utiliser cette information afin de mieux extraire la fonction plénoptique du détecteur et d'améliorer les algorithmes de synthèse d'images et de reconstruction de profondeur.

Dans cette dissertation, nous avons étudié en profondeur l'utilisation du modèle ECA pour les systèmes d'acquisition plénoptique. Cependant, nous croyons que ce modèle pourrait aussi être employé pour l'étude des systèmes d'affichage plénoptique comme les écrans auto-stéréoscopiques. Les travaux futurs pourraient se concentrer sur la détermination de quels systèmes sont aptes à être modélisés par un réseau virtuel de projecteurs, et sur la réplication de nos découvertes en utilisant des méthodes similaires d'analyse et de validation. Dans les faits, nous voulions mettre cette idée à l'épreuve avec le microscope plénoptique en ajoutant un chemin d'illumination utilisant le même réseau de microlentilles que pour l'acquisition. Malheureusement, un manque de temps nous a empêchés de mener ce projet à terme.

En matière d'applications, la mesure de la fonction de distribution de réflectance bidirectionnelle (BRDF) à l'échelle microscopique et mésoscopique pour sa mise en relation avec les modèles macroscopiques de BRDF semble une piste intéressante à creuser. Ce type de mesures propose de nouveaux défis puisque l'ensemble de l'hémisphère des directions depuis la surface a besoin d'être échantillonné à la fois pour l'acquisition et l'illumination. Ces contraintes pourraient certainement

être simulées et testées avec notre modèle. Les données obtenues serviraient à établir un pont entre les différentes échelles d'observation et offriraient une meilleure compréhension quant à leur lien.

Title Tools for the Paraxial Optical Design of Light Field Imaging Systems

Abstract Light field imaging is often presented as a revolution of standard imaging. Indeed, it does bring more control to the user over the final image as the spatio-angular dimensions of the light field offer the possibility to change the viewpoint and refocus after the shot and compute the scene depth map. However, it complicates the work of the optical designer of the system for two reasons. The first is that there exist a multitude of different light field acquisition devices, each with its own specific design. The second is that there is no model that relates the camera design to its optical properties of acquisition and that would guide the designer in his task.

This thesis addresses these observations by proposing a first-order optical model to represent any light field acquisition device. This model abstracts a light field camera as en equivalent array of virtual cameras that exists in object space and that performs the same sampling of the scene. The model is used to study and compare several light field cameras as well as a light field microscope setup which reveals guidelines for the conception of light field optical systems. The simulations of the model are also validated through experimentation with a light field camera and a light field microscope that was constructed in our laboratory.

Keywords optical design, light field imaging, light field cameras, light field microscopy

Contents

List of Figures

List of Tables

Acronyms

BRDF	Bidirectional Reflectance Distribution Function
CCD	Charge Coupled Device
CMOS	Complementary Metal Oxide Semi-conductor
DMU	Directional Multiplexing Unit
DSLR	Digital Single-Lens Reflex
DVF	Discrete Vignetting Function
ECA	Equivalent Camera Array
LCD	Liquid Crystal Display
MLA	Microlens Array
PSF	Point Spread Function
SLR	Single-Lens Reflex
SNR	Signal-to-Noise Ratio
SPC	Sampling Pattern Cube

Introduction

Light field imaging is a subject area that lies at the frontier between optics and computer science. It improves imaging and redefines optical design with computational power. This dissertation proposes and explores tools that help with the design of optical systems such as light field cameras and microscopes.

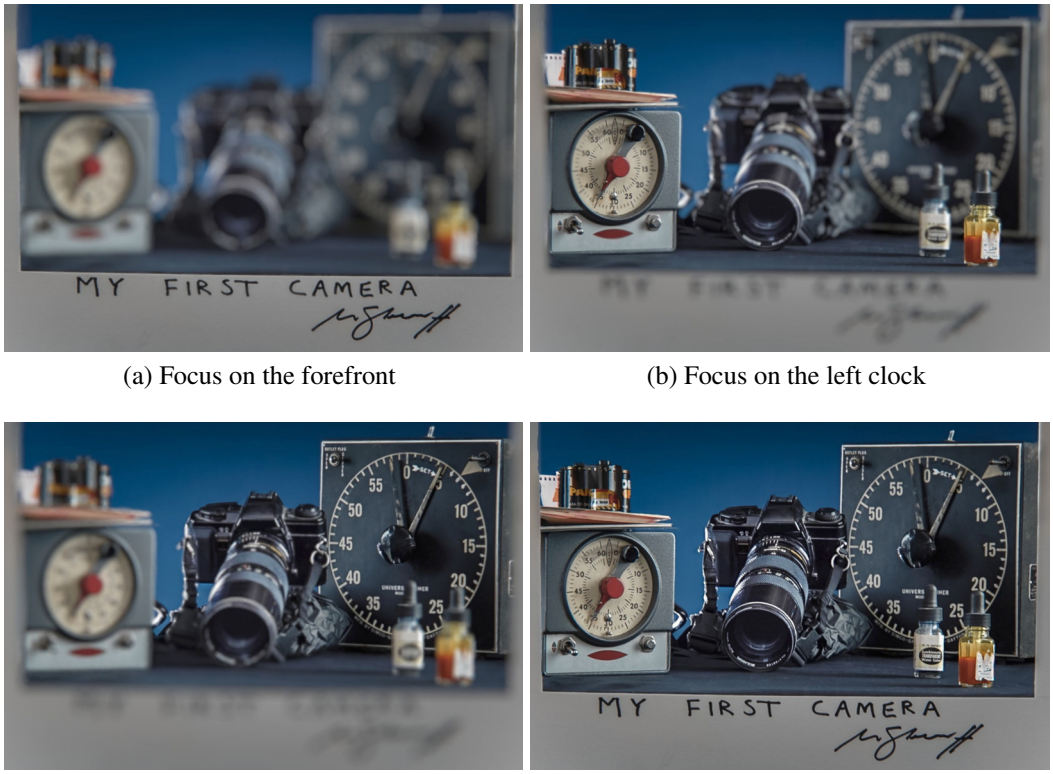
Context

In classical imaging, the image formation done by the optics results in a two-dimensional regular picture of a scene which is later improved through image processing algorithms to produce the final image for the user.

In light field imaging, the role of the optics is to guide the light to the sensor whereas the role of the algorithms is to interpret the value of the pixels. On the one hand, the algorithms need an accurate model of the imaging properties of the optical system and the sampling pattern of the sensor to synthesize new images correctly. On the other hand, the optics need to be conceived with the limits of the reconstruction algorithms in mind.

Designing a light field camera is essentially a co-design task of the hardware part and the software part at the same time. The difference with classical imaging is that the intermediate raw image is intended only as an input for a specific algorithm and that the end user only sees the output image of the computation. Consequently, the constraints on the hardware part are relaxed so it can record higher dimensional data than the regular spatial sampling of the scene such as the light direction of incidence on the sensor. The load on the computational part of the system is greater than before because it needs to extract and show understandable data to the user, but its role is crucial. In the end, the user can navigate through this high dimensional data set and synthesize images that no regular imaging system could produce on its own.

Actually, light field imaging is a subset of the larger field of computational imaging which regroups all kind of ways to perform this association between the optical acquisition system and the computational processing. It encompasses several diverse techniques such as time of flight imaging, structured lighting, and plenoptic imaging, just to name a few.



(a) Focus on the forefront

(b) Focus on the left clock

(c) Focus on the right clock

(d) All-in-focus image

Figure 1: Synthetic refocusing with the Lytro Illum. Copyright from LYTRO, INC..

Motivation

Light field imaging records the position and direction of light rays. With this new directional information, post-capture image synthesis allows to simulate a lens with different parameters and perform several actions such as to refocus the image, to extract depth information, to change the view point, the depth of field, and the focal length among other things.

Light field cameras offer new possibilities in a large variety of environments. Consumer products are aimed at amateur and professional photographers and encourage them to work differently as they can compose dynamic pictures that are interactive as illustrated in Figure 1. The focus and the viewpoint are controllable by the viewer and the experience reaches an equivalent level to what is possible with what people like to associate with the term “holography”.

In the movie industry, light field cameras simplify post-processing as the scene of action can be reconstructed in three dimensions. The light field data can be edited to remove an object using depth cues and can be filled with true data, circumventing occlusion issues. As examples, a blue or green background for scene matting is no longer needed and small objects like cables for special effects can be removed as if

they did not exist in the first place. Moreover, stereo rendering is directly possible with no additional effort. As an extension, light field cameras can also be used to create content for augmented and virtual reality scenarios.

In the rest of the industry, the main advantage of light field cameras over stereo and other multi-view acquisition systems is its reduced size that makes it an interesting option in possibly harsh use conditions like robotics.

When applied to the field of microscopy where the optical constraints of the system produce a thin 2D slice of a sample, light field imaging extends the depth of field captured by the instrument and allows to reconstruct the 3D structure of the observed specimen in one shot.

In order to face this diversity of applications, there is undoubtedly a need for a good understanding of the design aspects of these light field cameras which we address in this dissertation.

Problem Statement & Contributions

Light field cameras come in a wide variety of architectures based on specific optical components such as lens arrays for instance. However, the design of these various architectures is sometimes so different that it is hard to assess which architecture would address a particular application the best. Although the specificity of each architecture is known, the understanding is restricted to their particular design.

We have observed that a global model that allows the evaluation and comparison of different optical arrangements is missing. Ideally, this model would be complete *i.e.,* it would integrate both optical design and computer vision aspects as it aims at optimizing the spatial-angular sampling trade-off. This model would also simulate the architecture of any light field device and output meaningful properties that would help and guide the designing task.

In this dissertation, we propose a first step into the direction of this ideal model. The major contributions can be stated as follows:

- The development of an alternative mathematical scheme to ray matrices based on point imaging matrices;
- A first-order optical model that abstracts light field cameras as an equivalent camera array (ECA) in object space;
- The description of the optical properties of this equivalent camera array;
- The derivation of analytical expressions of the properties for microlens-based light field cameras;
- The comparison between the focused and afocal configurations of microlens-based light field cameras;
- The validation of the ECA model with a microlens-based light field camera;
- The study of the matching between a consumer light field camera and a microscope for light field microscopy applications and the development of alternate light field macrophotography setups;

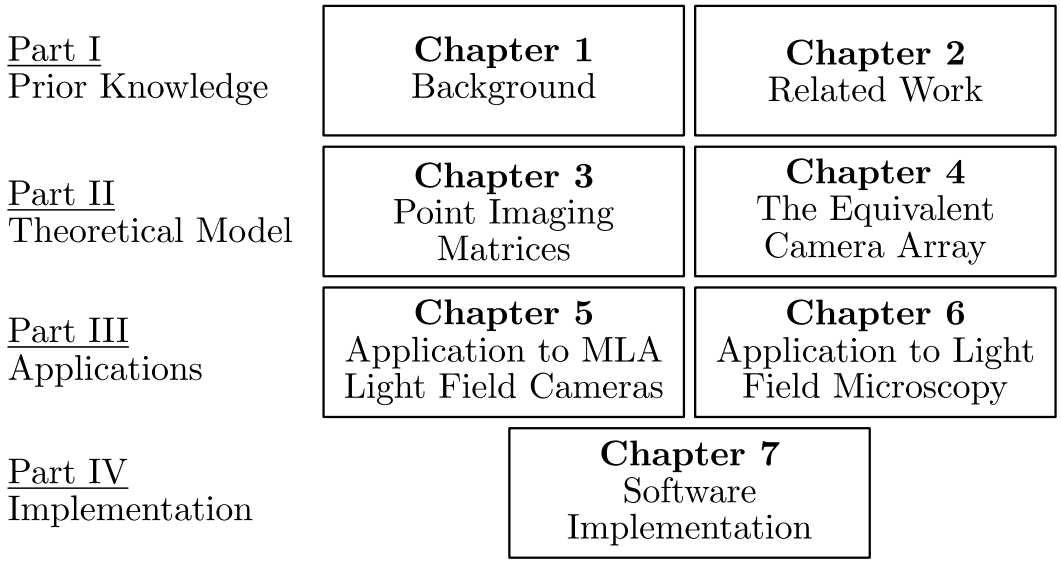


Figure 2: Dissertation overview.

- The realization of a light field microscope setup and its optical analysis with the ECA model;
- An implementation of the ECA model with MATLAB.

The model we propose considers optics as perfect imaging components and does not address aberrations introduced by the real shape of their surfaces. Moreover, our model does not consider diffraction caused by the wave nature of light.

Parts of the work presented in this dissertation have already appeared on several conferences and journals [Mignard-Debise and Ihrke, 2015; Ihrke *et al.*, 2016; Mignard-Debise *et al.*, 2017]. They were also presented at two meetings of the GdR ISIS (Groupement de Recherche Information, Signal, Image et ViSion): "Systèmes de Vision Grand Angle, Multi-Caméra et Plénoptique" and "Co-conception de systèmes hybrides : quand l'instrumentation et les traitements numériques se rencontrent". This dissertation also includes yet unpublished findings that detail and extend these publications.

Dissertation Overview

The dissertation is organized into four parts, see Figure 2. The first part, Chapters 1 and 2, is dedicated to an exposition of the prior knowledge required to understand the notions utilized in the following parts. Chapter 1 presents theoretical background knowledge about light fields, optical design, and microscopy. Chapter 2 details the state of the art in light field imaging from acquisition to image synthesis, as well as the various light field camera architectures, and the literature on light field microscopy.

The second part focuses on the proposed theoretical model for light field camera design in Chapters 3 and 4. Chapter 3 introduces a point imaging matrix tool as an alternate approach to the usual ABCD ray matrix tool used for describing light propagation in an optical system. This tool is then used in Chapter 4 which develops the theoretical aspects of the equivalent camera array model. These aspects include the procedure to construct the ECA corresponding to a light field camera, the description of the meaningful optical properties of this array, as well as, vignetting considerations.

The third part covers applications of the equivalent camera array model to two kinds of light field devices: microlens-based light field cameras in Chapter 5 and light field microscopy in Chapter 6. Chapter 5 investigates the similarities and differences of the two configurations: focused and afocal, and derives analytical expressions for the properties of their ECA. It also validates the ECA model through experimentation with an afocal light field camera. Chapter 6 is separated into two main parts. The first part explores the feasibility of associating a consumer light field camera with a microscope objective and presents alternative solutions for light field macrography. The second part applies the ECA model to a light field microscope to evaluate its imaging performance and compares it with the actual setup that has been set up in the laboratory.

The fourth part consists of Chapter 7 only and proposes an implementation of the ECA model in MATLAB which rests on the imaging tools of Chapter 3. It also presents an algorithm to compute the etendue of the optical system considering the blocking of light by the apertures of the optics.

In terms of chronology and publications, the first section of Chapter 6 is a revised version of [Mignard-Debise and Ihrke, 2015]. This work investigates the vignetting effects when associating two different optical systems. It was a leitmotiv to develop the ECA model.

The contents of Chapter 3, 4, and 5 are a revised version of [Mignard-Debise *et al.*, 2017] and the experimental part of the validation work in Chapter 5 was achieved together with John Restrepo, another Ph.D. student.

The second part of Chapter 6 and the entirety of Chapter 7 are unpublished works as of the writing of this dissertation.

Part I

Prior Knowledge

Chapter 1

Background

The current chapter provides the background for several subjects that are necessary to understand the rest of the work presented in this dissertation. These subjects encompass the fields of light field imaging, radiometry, paraxial imaging, optical design, wave optics and microscopy.

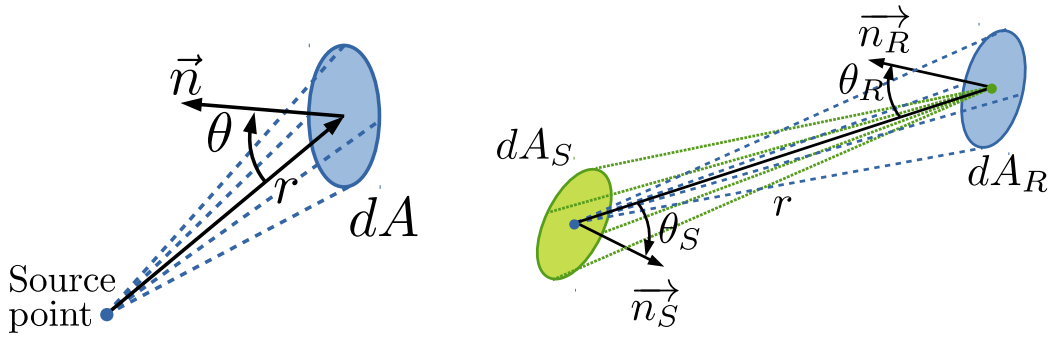
1.1 Light Field

The light field [Levoy and Hanrahan, 1996] is a ray based model that is derived from the plenoptic function [Adelson and Bergen, 1991] where rays are infinitely small beams of light radiance. In this section, we present the main aspects of the light field model, how it relates to physics, and present a tool to represent it: the phase space [Torre, 2005].

1.1.1 Radiometry

Radiometry or photometry is the field of physics that is focused on describing the different aspects of energy transport of light between an emitter and a receiver, see [Singer *et al.*, 2006]. The light emission can come from a principal source that emits light by itself such as the sun, a halogen lamp, or a light emitting diode or by a secondary source that reflects or diffuses light such as a spoon, or a piece of paper. The geometrical conditions of the light emitter or receiver such as its area size, direction and solid angle of emission or reception determine multiple radiometric quantities. All of these quantities integrate over the spectrum of the light source. The corresponding quantities per unit of wavelength exist and are referred to by adding "spectral" to the name of the quantity.

Solid angle The solid angle is a generalization of the usual angle in a 2D plane to 3D space, see Figure 1.1a. The solid angle of a cone of light emitted by a source point P through the surface element dA at a distance r, the normal $\overrightarrow{n}$ of which



(a) (b)

Figure 1.1: Geometric definitions for radiometry. (a) Solid angle of a source point. (b) Etendue between a source surface S and a receiver surface R.

makes an angle θ with the axis of the cone is defined as:

$$d\Omega = \frac{\cos\theta dA}{r^2}\,. \tag{1.1}$$

The unit of the solid angle is the steradian $[sr]$ and a sphere has a solid angle of 4π steradians.

Radiant flux The radiant flux or power Φ of a source is the radiant energy emitted or received per unit of time. It is expressed in Watts $[W]$.

Radiance The radiance L is the density of power per unit of surface area $[m^2]$ and solid angle of the light source. So we have

$$L = \frac{d^2\Phi}{\cos\theta\; d\Omega\; dA} \tag{1.2}$$

measured in $[W \cdot m^{-2} \cdot sr^{-1}]$.

Etendue From the previous definition of radiance, the flux emitted by a source S of infinitesimal surface dA_S and received by a detector R of infinitesimal surface dA_R, see Figure 1.1b, is

$$d^2\Phi_{S\to R} = L_R \cos\theta_R\; dA_R\; d\Omega_R\,, \tag{1.3}$$

where the subscript $S \to R$ denotes that the quantity is transported from the source to the receiver and with $d\Omega_R$ the solid angle of the emitting surface seen from the detector defined as:

$$d\Omega_R = \frac{\cos\theta_S\; dA_S}{r^2}\,. \tag{1.4}$$

Then, from the previous equations, the flux can be written as:

$$d^2\Phi_{S\to R} = L_R \frac{\cos\theta_R \, dA_R \, \cos\theta_S \, dA_S}{r^2} \,. \tag{1.5}$$

The geometrical part of the formula without the radiance is called the etendue of the source-detector combination:

$$d^2G = \frac{\cos\theta_R \, dA_R \, \cos\theta_S \, dA_S}{r^2} \,. \tag{1.6}$$

The reciprocal flux from the receiver to the source $d^2\Phi_{R\to S}$ is the product of the source radiance L_S with the same etendue term. The conservation of radiation transfer states that the flux received or emitted between the two surfaces is equal so $d^2\Phi_{S\to R} = d^2\Phi_{R\to S}$. Consequently, we have $L_R = L_S = L$, *i.e.,* radiance is constant along a ray.

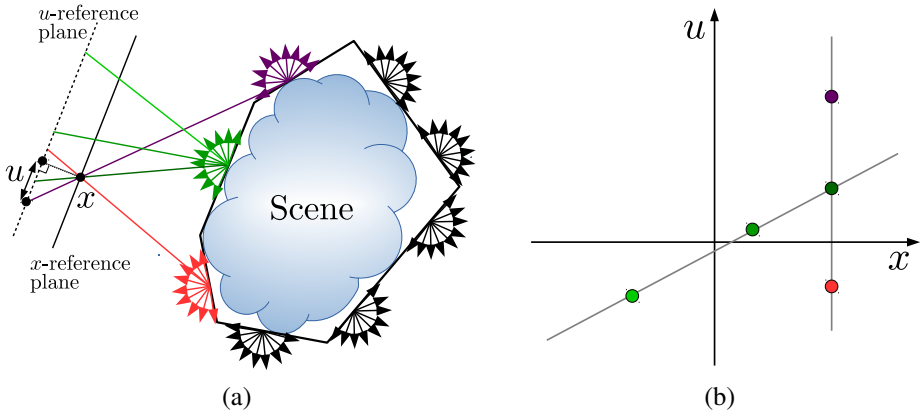
As we will see later, the etendue is an important part of optical systems because it is only related to the geometrical part of the radiation transfer and not to the property of the light source which is characterized uniquely by the radiance L.

1.1.2 Plenoptic Function

The plenoptic function as introduced by [Adelson and Bergen, 1991] is a ray based concept that attributes radiance to every light ray emitted from a physical space. This model is not based on any physical phenomenon but simply links the optical properties of the 3D objects of the scene to the observer through a dense set of light rays transmitting these properties. The observer takes samples of the plenoptic function to form an image.

The radiance of a ray is characterized by its position (3D), direction (2D), wavelength (1D), and instant in time (1D) making the plenoptic function a 7D function. A color picture from a regular camera is a 3D slice (2D for position + 1D for color) of this function where the other parameters are integrated over a range of values. The camera aperture delimits the extent of the cone of rays hitting the sensor and the exposure time averages light in a certain time frame.

However, this concept does not account for several aspects of physics. Firstly, it considers a ray as an infinitely thin beam of light. This assumption becomes largely invalid when the size of the pencil of rays approaches the order of the wavelength of light. Diffraction and interference, due to the wave nature of light then limit the applicability of the ray model. Secondly, the measure of light intensity or flux relies inherently on an integration time due to the photonic nature of light that distributes energy in packets of photons. Thirdly, some light sources such as lasers exhibit spatial and temporal coherence which manifest as interference effects. Consequently, the plenoptic function is of interest at a macroscopic scale where the structures observed are large so that wave effects can be ignored. Moreover, time scales are considered sufficiently long ($> 10\mu s$) so that the light energy can be temporally



(a) (b)

Figure 1.2: Definition of the light field in physical space (a) and its representation in phase space (b).

averaged on a detector. And finally, only incoherent light sources that emit wave packets randomly such as black body emission, fluorescence or phosphorescence are valid.

1.1.3 Light Field Definition

The light field, introduced by [Levoy and Hanrahan, 1996], is a subset of the plenoptic function at a particular instant of time and a single wavelength so that only the geometrical aspect of the ray remains. It therefore inherits the limitations of the plenoptic function. It is also assumed that the radiance of the light ray does not change along its trajectory meaning that the light ray is traversing non-diffusive free space so it does not loose energy. In order to verify this last assumption, one can consider a convex volume that marks the boundary between the studied scene and free space, see Figure 1.2a. The light field is then defined correctly at the surface and outside this volume where there is no occlusion and where an acquisition device can be located. If the surface of this convex shape matches the surface of an object of the physical world, the light field function interestingly reports the reflective properties of the material convolved with the incident lighting. From this reduction of dimensions, the light field is a four-dimensional function with two spatial and two directional dimensions.

For the moment, a light field is monochromatic. As a result of assuming that the different wavelengths are independent, three-channel RGB light fields are recorded using the same technique that RGB cameras use. A Bayer filter is placed on top of the sensor, and a colored image is produced by interpolation after the capture. However, a few light field acquisition devices mix the wavelength information with the directional information, associating each direction with a different color filter, leading to a finer chromatic resolution, see Chapter 2.

1.1.4 Representation in Phase Space

Optical Phase Space

Hamiltonian formalism for optics describes the way rays pass through optical systems, see *e.g.,* [Torre, 2005]. In this formalism, a ray is represented by its 2D position (x, y) and 2D momenta (p_x, p_y) at an evaluation plane perpendicular to the optical axis of the system. These ray coordinates (x, y, p_x, p_y) constitute a 4D space called the geometric-optical phase space. A ray in this space is represented as a point and the ray trajectory in the 3D space is represented as a trajectory of the representative point of the ray in the phase space by moving the evaluation plane along the optical axis. This representation is similar to that of mechanics where the trajectory of a point in 3D space is also represented in the 6D phase space of its 3D position and 3D momenta but the optical axis is replaced with the axis of time.

Obviously, the ray momenta differ from a particle momentum, they are the optical cosine directions of the ray scaled by the refractive index of the medium that the ray traverses. Let us set the optical axis as the z-axis of a reference frame, the x and y axis of the frame being the frame of the evaluation plane. Then, we can write:

$$p_x = n \sin(\alpha_x),\ p_y = n \sin(\alpha_y) \tag{1.7}$$

where n is the refractive index at the position of intersection (x, y, z) of the ray with the evaluation plane. The angles α_x and α_y are the angles between the ray and the planes yz and xz, respectively.

Another difference between the optical phase space and the mechanical phase space is caused by a restriction on the optical momenta imposed by this relation : $p_x^2 + p_y^2 \leq n^2$. The momenta coordinates are limited to the inside of a circle so the support of the optical phase space is $\mathcal{C}_2 \times \mathbb{R}^2$ and not $\mathbb{R}^6$ as for the mechanical phase space.

In the case of first-order optics that we address in Section 1.2, the ray momenta remain well below this limit and it is usually ignored. In the rest of the dissertation, we refer multiple time to the phase space but we always imply this linearized version.

Light Field Parametrization

There exist several ways to parametrize the light field function. The one most commonly used is the two-plane parametrization from [Levoy and Hanrahan, 1996] where a ray is parameterized by its intersection coordinates with two reference planes usually parallel to each other at a distance of 1. The planes are set and known, so the intersection of a ray with the first plane gives the spatial components x and y and the relative shift at the intersection of the second plane gives the directional component u and v of the light ray. This two-plane parametrization of a ray is actually the same as the linearized phase space representation as in the paraxial domain, $\tan\theta \approx \sin\theta \approx \theta$.

The phase space is four-dimensional and so, it is hard to represent directly. Therefore, it is often represented as a 2D diagram with the first spatial coordinate x and the first directional coordinate u only. This 2D representation makes the interpretation and the explanation of its properties easier without loss of generality.

As Figure 1.2b illustrates, a bundle of light converging or diverging from a point in space is represented as a line in phase space (and as a plane in 4D phase space for a 3D bundle). When this point lies in the reference plane, this line is parallel to the u-axis. Bundles originating from points at the same distance from the reference plane result in parallel phase space lines. Translation of the phase space reference plane along its normal results in a shear along the x-axis of the phase space diagram.

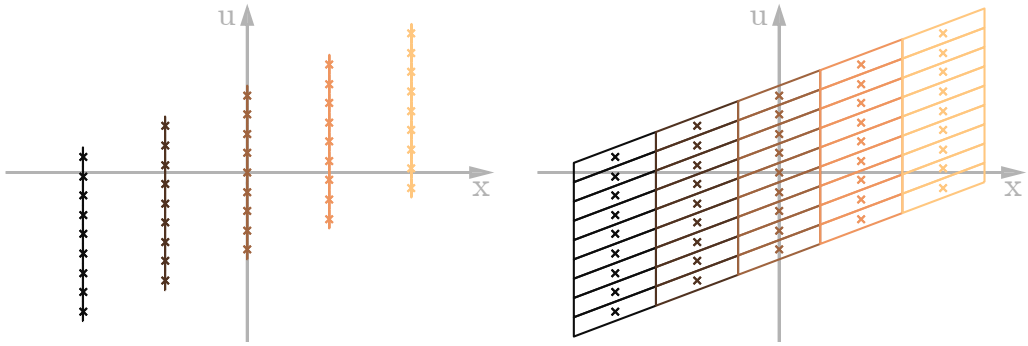
These properties make the phase space diagram a useful tool to represent ray bundles because they correspond to finite regions. As such, the light field function can be represented in the phase space by a radiance function where the radiance of a ray is associated to the corresponding phase space point. Any acquisition device integrates over the radiance of a bundle of rays. It can be viewed as sampling the phase space and consequently sampling the light field function.

1.1.5 Sampling

By placing a pinhole in the reference x-plane and having a sensor a unit away in the u-plane, a simple pinhole camera is made. Considering that the pinhole is infinitesimal but still letting light go through and that, in our model, light is transported by rays and not waves, pixels are purely directional samples represented in phase space as a series of vertically aligned and joined segments, see Figure 1.3a. The spatial dimension of the phase space is then sampled spatially by regularly translating the pinhole in the x-plane by a finite amount.

A more realistic physical model attributes a finite aperture to the pinhole so that each pixel collects a ray bundle and the phase space lines are transformed into a series of stripes of vertically aligned cells as described by [Ng, 2006]. Making the size of the pinhole equal to the spatial sampling fills the gaps between the stripes and the phase space is fully sampled, see Figure 1.3b. Each phase space cell is a single measurement of the light field function, this measurement is associated to the point at its center indicating the sampling position. The cell is the spatio-angular kernel of the pinhole/pixel pair.

From a hardware perspective, an array of pinholes in front of a sensor is an easy way to implement a device that records a light field, though it suffers from low light throughput inherent to the small aperture of the pinholes. Many different optical systems have been designed to fulfill this task with better imaging performance. Light fields can also be produced using a parallax barrier in what is known as autostereoscopic displays. These systems are presented in Chapter 2.



(a) (b)

Figure 1.3: Phase space in 2D of the light field produced by an array of 5 holes in front of a sensor with 9 pixels. (a) Using pinholes. (b) Using holes with a diameter equal to their pitch so that they touch. The phase space is evaluated at the holes plane and each hole is given a different color. The cross represents the sampling position for the cell.

1.2 Optical Design

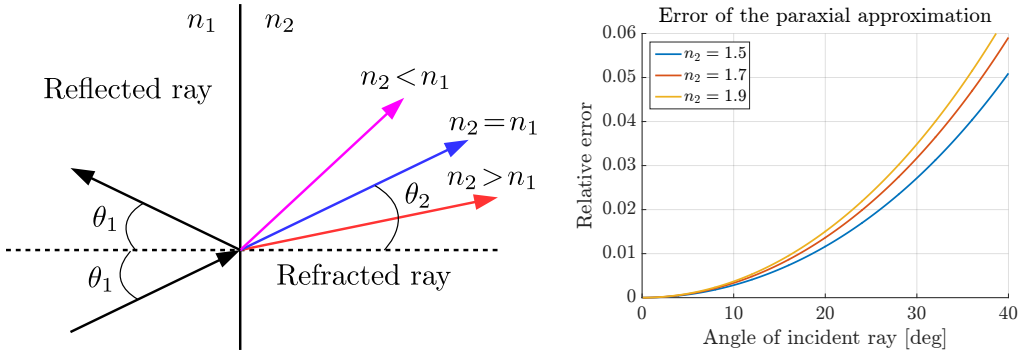
All of the optical concepts described in this section are presented in [Born and Wolf, 1980].

Designing an optical system is a task that requires to know how image formation occurs in an optical system. Image formation is possible as cones of rays are transformed by an optical element such as a refractive or reflective interface into other cones of rays that are required to be converging to an image point. Unfortunately, the output cones are not perfect and the image of a point source is most of the time a fuzzy spot. The task of the optical designer is to reduce the size of this spot as much as possible and still achieve the specific optical properties such as the light throughput or the field of view that the optical instrument requires. For instance, a zoom objective requires to stay efficient over a wide range of focal lengths by moving different lens groups which greatly affects the light transport.

Optical design relies on raytracing, *i.e.,* the propagation of light rays through each optical component of an optical system by applying Snell's law of refraction at the interface between two different media (of optical indices n_1 and n_2) or the reflection law if the interface is a mirror. As illustrated in Figure 1.4a, the refraction law links the incident ray angle θ_1 at the point of incidence to the refracted ray angle θ_2 with the following formula:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{1.8}$$

The reflection law states that the angle of the incident and reflected rays are equal. Moreover, the refracted or reflected ray belongs to the plane defined by the incident ray and the normal of the surface at the point of incidence but on the other side of



(a) (b)

Figure 1.4: Snell's refraction law. (a) Incident, reflected and refracted rays at an interface between two media. A higher index in the second medium means that the ray is refracted towards the surface normal. (b) Relative error of the angle of the refracted ray. The first-order approximation is considered acceptable for incident angles less than 20 degrees for which the error is less than 1%.

the normal and the angles are defined relative to this normal.

At first order, in what is called the paraxial domain, $\sin\theta \approx \theta$ which introduces a small amount of error for small angles of incidence in the linearized Snell's refraction law, see next Section. Light cones are transformed into perfect light cones by optical components in this approximation. Rays converging or diverging from an object point on one side of an optical element converge or diverge on the other side of the component. This is called stigmatism and the point of convergence of the rays is called the conjugate image point.

At higher orders, rays deviate slightly from the paraxial rays and do not converge onto a single image point anymore but make a blurry spot. The system is subject to what is called aberrations and the task of the optical designer is to determine the shape and position of the surfaces a well as the optical index of the glass of the optical components to minimize the aberrations of the system *i.e.,* to minimize the size of the blurry spot.

1.2.1 Paraxial Imaging

For values close to zero, the sine function is known to be well approximated by a linear function. Consequently, the formula of Equation 1.8 becomes linear:

$$n_1\theta_1 = n_2\theta_2 \tag{1.9}$$

This relation remains reasonably valid as long as the ray considered follows what is known as the Gauss conditions:

- the rays stay close to the optical axis

- the ray direction deviates only slightly from that of the optical axis

The purpose of the conditions is to have the rays encounter quasi-flat interfaces and share the same surface normal so that the approximation is as good as possible. This is possible by restraining the pupil of the system and the field of view. In practice, there is no set limit to the maximum angle the rays can reach. This limit must be defined individually in accordance with the desired accuracy, see Figure 1.4b. Paraxial optics is implicit throughout the rest of this dissertation except where it is explicitly stated otherwise.

1.2.2 Thin Lens Model

A real lens in the paraxial domain is often considered as a thin lens based on the condition that the radius of curvature of its refractive surfaces is large in comparison to its thickness. The double refraction is then considered to occur in a single plane. The optical axis of a thin lens is the axis perpendicular to this plane and is passing through its optical center. This axis is the axis of reference for angles and heights. Because most real optics are made of spherical surfaces, they are rotationally symmetric around their optical axis. Aligning many such optical components such that their optical axes coincide makes the whole system rotationally symmetric. However, in a real system, any small deviation, be it a shift or a tilt, breaks this property so the alignment of an optical system is usually an important task.

Imaging through a thin lens follows the thin lens equation that relates the position of the object and image points with the focal length of the lens. The conjugate image point $\mathbf{A}' = (x_{\mathbf{A}'}, y_{\mathbf{A}'}, z_{\mathbf{A}'})$ of an object point $\mathbf{A} = (x_{\mathbf{A}}, y_{\mathbf{A}}, z_{\mathbf{A}})$ through a thin lens with optical axis z and focal distance f, centered at the origin, see Figure 1.5, is provided by:

$$\frac{1}{z_{\mathbf{A}'}} - \frac{1}{z_{\mathbf{A}}} = \frac{1}{f}\,. \tag{1.10}$$

A point $\mathbf{A}$ at infinity emits a collimated beam of light that gets focused in the back focal plane of the lens where $z_{\mathbf{A}'} = f$. Similarly, a point $\mathbf{A}$ in the front focal plane where $z_{\mathbf{A}} = -f$ is imaged to a point $\mathbf{A}'$ at infinity. Note that by having the origin at the center of the lens, the points in front of the lens have a negative z coordinate and the points behind the lens have a positive z coordinate.

For a convex or positive lens, the focal length f is positive and cones of rays are converted into converging light cones behind the lens forming a real image if the distance between the object point and the lens is larger than the focal length. A concave or negative lens has a negative focal length and does the opposite: cones of rays are converted into diverging light cones after the lens forming a virtual image.

The magnification $M_{\mathbf{A}\mathbf{A}'}$ between the two points is defined as:

$$M_{\mathbf{A}\mathbf{A}'} = \frac{z_{\mathbf{A}'}}{z_{\mathbf{A}}} = \frac{x_{\mathbf{A}'}}{x_{\mathbf{A}}} = \frac{y_{\mathbf{A}'}}{y_{\mathbf{A}}}\,. \tag{1.11}$$

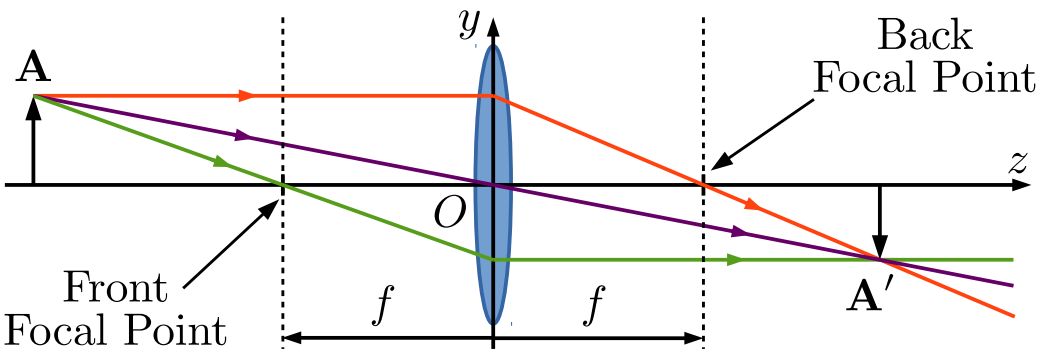


Figure 1.5: Convex thin lens. Three rays are easy to propagate. A ray (in orange) parallel to the optical axis is refracted to pass through the back focal point. A ray (in purple) passing through the lens center at O is not deviated. A ray (in green) passing through the front focal point is refracted parallel to the optical axis. The dashed lines represent the focal planes.

The magnification reports the ratio of the image size to the object size and is independent of the state (real or virtual) of the object or the image. It is constant between two conjugate planes orthogonal to the optical axis. However, it varies between different pairs of planes. The closer an object approaches the focal point of a lens, the further from it moves its image. At the same time, the magnification increases linearly. This implies that objects are three-dimensionally deformed through the lens following a perspective transformation of the space.

Real and virtual points The object point belongs to the object space of the lens and the image point to its image space. In other words, the object space is the collection of points in space before the imaging is done. The lens plane separates the object space into two regions, the one in front of the lens where light rays actually pass through the object point and the second region behind the lens where light rays cannot reach the object point as they are deviated by the lens. Object points are labeled as real in the first region and as virtual in the second region. The image space is the collection of the image points. Similarly to the object space, the lens plane separates the image space into two regions, in front of and behind the lens, depending on whether light rays actually converge to the image point or not. They cannot effectively converge before the lens so the image points are virtual but they can after the lens so the image points are real. Object and image space may overlap so that an object point and its image point may lie on the same side of the lens. In this case, one is real and the other is virtual.

This distinction between real and virtual points is important as an optical system usually observes true points from the real world that are also real optical object points. Moreover, their final image counterparts are usually real and as a result, a sensor needs to be placed in the image plane to integrate the light beam. A mag-

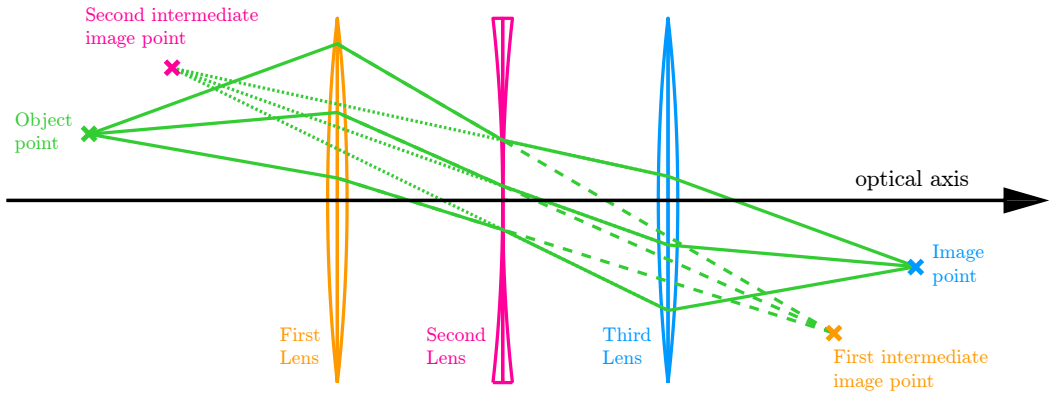


Figure 1.6: Ray diagram of a three-lens system with a second negative lens. Two intermediate image points are produced, they are virtual images. Virtual rays passing through them are indicated with a dashed line.

nifying glass, on the contrary, makes a large virtual image of a real object point so that the eye can actually see the object larger than it is.

1.2.3 Ray Propagation

The study of an optical system is usually performed by tracing pencils of rays coming from an object either on the optical axis, off the optical axis in the object plane, or at infinity sequentially through the various components of the optical system in order to compute the output pencils. Figure 1.5 describes how rays passing through the object point are deviated by a single thin lens to pass through the image point. In a system composed of more than one lens, this process can be repeated till reaching the final image space after the last lens.

Figure 1.6 illustrates the procedure. The object point is imaged to the first intermediate image point by the first lens. Any ray passing through the object point and a point on the lens plane is deviated from this plane point to pass through the image point. The first intermediate image point is both in the image space of the first lens and in the object space of the second lens. The ray passing through the first intermediate image point is extended to hit the plane of the second lens where it is deviated to pass through the image of this point which is called the second intermediate point. Then, the ray is extended to the third lens plane and is deviated again to pass through the final image point.

However, there exist quantities that are kept constant throughout the ray propagation. They relate rays from the first object space to the last image space and facilitate the ray tracing process.

Helmholtz-Lagrange invariant The Helmholtz-Lagrange invariant H relates the size of the object y, the angle of the emission cone from this object θ, and the index

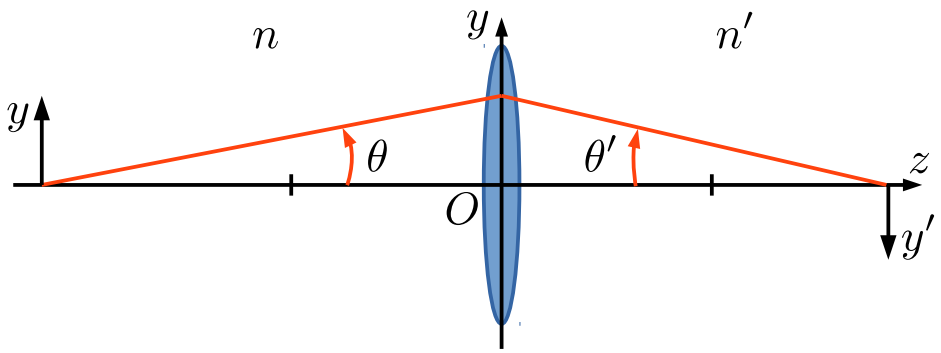


Figure 1.7: Helmoltz-Lagrange invariant.

of the medium n in the object space to their image counterparts y', θ' and n' (see Figure 1.7):

$$H = ny\theta = n'y'\theta' \,. \tag{1.12}$$

H is a measure of the optical capability of the system to transport light. It corresponds to the conservation of etendue and thus a conservation of energy. In effect, it is assumed that the refracting and reflecting surfaces have a transmission and reflection coefficient equal to one, and that a bundle of rays emerging in image space has not been obstructed by any of the optical components, so the total energy of the bundle must be conserved even if its shape changes.

Paraxial invariant The paraxial invariant P generalizes the Helmholtz-Lagrange invariant. It is defined for two rays of height y_1, y_2 and angle u_1 and u_2 in the same space of optical index n as:

$$P_{12} = ny_1u_2 - ny_2u_1 \,. \tag{1.13}$$

Two rays emitted from the object point intersect at the image point in image space because the system is linear. Therefore, the knowledge of these two rays is enough to trace any third ray without having to trace the ray in every intermediate space in between the first object space and the last image space. Tracing two rays through the full system gives the parameters of these two rays in object and image space. The parameters of a third ray in image space can be obtained from its parameter in object space and the following formulas:

$$y'_3 = -\frac{P_{23}}{P_{12}}y'_1 + \frac{P_{31}}{P_{12}}y'_2 \,, \tag{1.14}$$

$$u'_3 = -\frac{P_{23}}{P_{12}}u'_1 + \frac{P_{31}}{P_{12}}u'_2 \,. \tag{1.15}$$

Ray-transfer matrices The paraxial invariant is a good way of easily tracing rays through the system but it still relies on the prior knowledge of two rays that have to be traced completely through the system. There exists another powerful tool for ray tracing based on matrix calculus. As a result of paraxial optics being linear and of optical systems being rotationally symmetric, the height y' and angle u' of an image ray is linked to the object ray height y and angle u by the following relations:

$$\begin{aligned} y' &= Ay + Bu\,, \\ u' &= Cy + Du\,. \end{aligned} \tag{1.16}$$

These can be rewritten in matrix form as:

$$\begin{pmatrix} y' \\ u' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} y \\ u \end{pmatrix} = M \begin{pmatrix} y \\ u \end{pmatrix}\,, \tag{1.17}$$

where M is the ray-transfer matrix of the system also known as the ABCD-matrix.

The ray-transfer matrix of an optical system can be decomposed into successive single operations performed by each part of the system, each having an associated specific ray-transfer matrix. The final matrix is a product of its elementary matrices:

$$M = M_k M_{k-1} \cdots M_2 M_1\,. \tag{1.18}$$

1.2.4 Pupils

The ray tracing techniques that were presented in the previous sections assume that the optical components have an infinite aperture meaning that rays with an arbitrary large angle can propagate without being blocked. However, this is not the case in a real optical system because the cutting and mounting of the optical elements limits the extent of a bundle of rays. The finite size of the apertures also fixes the Helmholtz-Lagrange invariant to a finite value indicating the amount of light that can pass through the system.

The aperture stop, the entrance pupil, and the exit pupil The largest cone of rays emitted from a point at the intersection of the object plane and the optical axis that can pass through a centered rotationally symmetric system is limited by one of the apertures of the system which is then called the aperture stop of the system, see Figure 1.8. The main method to find this particular aperture is to trace an arbitrary ray from the object point and to record the height of this ray at every aperture plane relative to the aperture size. The aperture with the largest relative height is the stop, see [O'Shea, 1985].

The aperture that is selected as the stop highly depends on the position on the object plane. One of the goals of the optical designer is to keep the aperture stop the same even when refocusing, which changes the object plane, or zooming, which modifies both the focal length and the object plane. For purposes of exposure control, in photography, the aperture stop is often an iris diaphragm with a controllable

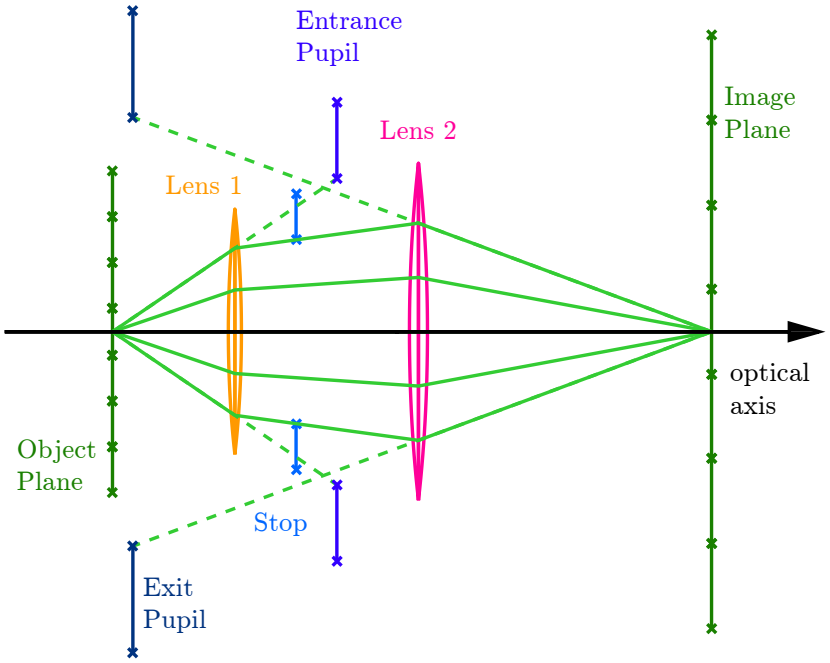


Figure 1.8: Pupils in a two-lens system.

aperture that can be reduced or increased at wish by the user. The shape of the aperture stop is usually circular as most lens mountings are circular but iris diaphragms can be made in the shape of straight or rounded polygons with the use of a multiple blades apparatus, for instance.

The entrance and the exit pupil of the system are defined as the images of the aperture stop in the object and the image space respectively, see Figure 1.8. Once the aperture stop is known, many of the object space properties of the optical system can be evaluated.

Numerical aperture and F-number There are two particular rays that turn out to yield the most useful information about the system. The first is called the marginal ray and goes from the object point on the optical axis to the border of the entrance pupil. It defines the aperture angle of the object in object space θ and image space θ' (see Figure 1.9). The numerical aperture NA and NA' of the system in object and image space, respectively, are:

$$\begin{aligned} NA &= n \sin\theta \,, \\ NA' &= n' \sin\theta' \,. \end{aligned} \tag{1.19}$$

The numerical aperture is an important aspect of an optical system as it determines its resolution limit. A larger NA means a better resolving power *i.e.,* smaller structures of the object can be imaged. The numerical aperture is also related to the

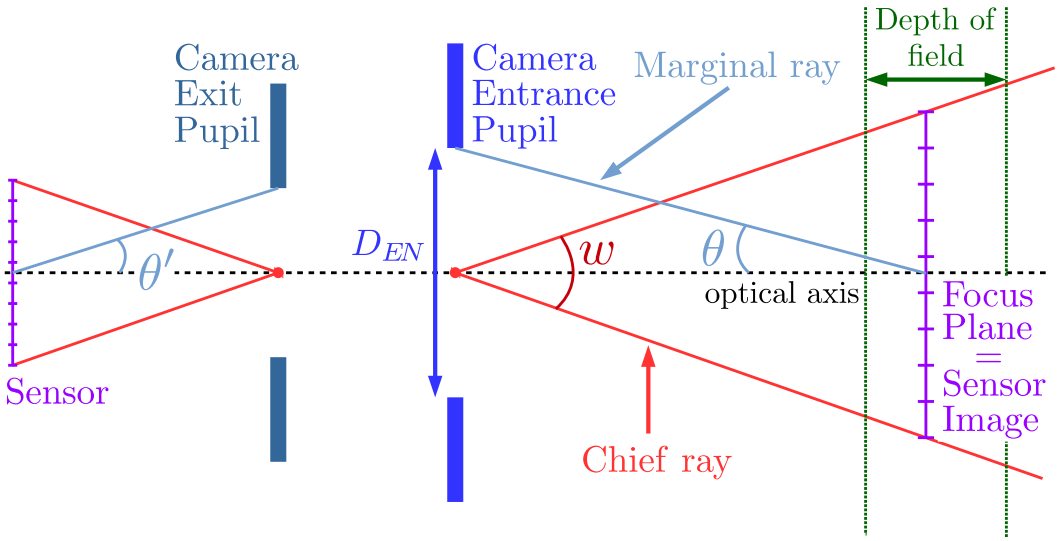


Figure 1.9: Properties of an optical system.

amount of light reaching the sensor. A larger NA means an increase in the signal-to-noise ratio (SNR) and the possibility to image in an environment with poor illumination conditions or weak emitting objects. The other way to increase the SNR is to use the sensor at a larger integration time. However, motion blur caused by fast moving objects is more likely to occur, so a larger NA is necessary for high speed cameras for instance.

Another parameter closely related to the numerical aperture is the f-number F defined as:

$$F = \frac{f}{D_{EN}} \,. \tag{1.20}$$

with D_{EN}, the diameter of the entrance pupil. The f-number is an alternative indicator of the aperture of an optical system imaging at infinity. The larger the f-number, the smaller the aperture. The f-number is specific to an objective configuration (focal length and pupil aperture) and not to the imaging conditions.

The *effective* f-number is a generalization for imaging an object at a finite distance:

$$F_{eff} = \frac{1}{2n' \sin\theta'} = F(1 - M) \,. \tag{1.21}$$

It links the characteristics of the optical system through the f-number F and the imaging conditions through the magnification M.

Resolution limit Even in a perfect optical system, free of all aberrations, diffraction by the aperture of the pupil limits the resolution power of the system. The image of an object point is thus a spot with the shape of an Airy disk. Considering that the aperture is circular, the angular separation between the central maximum

and first zero of the diffraction pattern is:

$$\theta = 1.22\frac{\lambda}{D_{EX}} \tag{1.22}$$

with λ the wavelength of the monochromatic light beam. On the sensor, this angular separation translates to the "radius" r of the spot which is expressed as:

$$r = 1.22\frac{\lambda}{2NA'}\,. \tag{1.23}$$

The Rayleigh criterion, for instance, states that two object points are just resolved by the optical instrument if the distance between the center of their spots is larger than the radius r of their spot. If the distance is larger, the two points are well-resolved, if it is smaller, they are not resolved. Other criteria such as the Schuster and the Sparrow criteria exist. The difference in the formula is a change of the 1.22 multiplying factor.

Field of view and depth of field The second interesting ray is the chief ray. This ray goes through the center of the aperture stop and comes from an off-axis point on the object plane. The pupils are the conjugate images of the aperture stop, as a result, the prolongation of the ray in object and image space goes through the center of both the entrance and the exit pupil see, respectively, Figure 1.9.

The field of view is the part of the object space that can be imaged by the optical system. Geometrically, the field of view is a cone with its vertex at the center of the entrance pupil. This cone is bounded by the maximal chief rays that are not vignetted by the system. The angular extent ω between the maximal chief rays defines the angle of view.

The sensor characteristics (number of pixels and pixel pitch) usually determine the effective field of view that is recorded. By cause of the sensor being rectangular, the field of view is a frustum which defines a vertical and a horizontal field of view.

The image of the sensor in object space sets the focus plane where any object is imaged sharply onto the sensor. Objects located outside of the focus plane appear blurry because the cone of light emitted from a point on their surface is cut by the focus plane and results in a spot and not a single point.

Due to the finite size of a sensor pixel, the focus plane has a depth called the depth of field. The depth of field is the range of distances for which the size of the blurred spot is smaller or equal than the size of a pixel. Consequently, it allows objects close to the focus plane to appear sharp as well.

The size of the aperture stop controls the extent of the depth of field. A larger aperture means a larger light cone and a smaller depth of field. This effect is well illustrated in macro-photography as it is used to emphasize the subject by blurring the foreground and the background.

Bokeh The shape of the aperture stop can be observed by placing a small light source or an object highlight outside of the depth of field. The image of this out-of-focus point results in a projection of the aperture shape onto the sensor and is called the *bokeh effect*. This image can be considered as a blur kernel with the kernel size depending on the depth of the object creating it. The further the object from the focal plane, the larger the kernel.

Hyperfocal distance The depth of field also depends on the distance of the focus plane to the optical system. It increases when the imaged objects move further away. The depth of field is not symmetric around the focal plane, so the closest and furthest sharp plane are at different distances from it. When imaging at infinity, objects close to the camera appear blurry but objects located after a finite distance still appear sharp. In this case, the closest limit of the depth of field is called the hyperfocal distance.

Aberrations As stated previously, the paraxial domain is only valid for small angles and as such the aperture of the stop is the way to control the validity of the paraxial approximation. Optical systems with large f-number, *i.e.,* a small aperture, easily satisfy the Gauss conditions but smaller f-number systems which are sought for their large numerical aperture are prone to aberrations. A theory of aberrations [Born and Wolf, 1980] was developed to predict and evaluate the effect of the aberrations on the imaging quality of optical systems. In this theory, the paraxial exit pupil plane is chosen as the reference plane to compute aberrations. Aberrations are measured in terms of positional and angular shift between the paraxial non-aberrated ray and the real ray.

Perspective and telecentricity The center of the entrance pupil is the center of view of the camera. As the aperture stop is closed down to a pinhole, only rays passing through the center of the entrance pupil can reach the sensor. The position of the entrance pupil matters for the perspective visual impression of an object. There exist three types of perspective depending on the relative position of the object and the entrance pupil: entocentric, telecentric and hypercentric.

Entocentric perspective is the normal perspective of the human eye where further objects appear smaller as the entrance pupil is located between the object and the imaging system.

Telecentric perspective is used mainly in telescopes and for microscopy. The entrance pupil is located at infinity and the effective magnification is constant. This perspective is also called orthocentric. An optical system can also be made telecentric on the image side by having the exit pupil at infinity.

Hypercentric perspective has the opposite effect of entocentric perspective. It makes further objects appear larger. The pupil is located at a finite distance behind the object. It is used in the industry to control the quality of the surface of cylindrical

objects such as bottles or barrels with a single camera instead of using many cameras around the object.

1.2.5 Vignetting

The sensor does not set the field of view of the optical system but it sets the effective field of view that is recorded. Any optical system has a maximum field of view that is limited by vignetting effects making the light irradiance non-uniform on the image. The vignetting effects are of two different sources: mechanical and natural.

Field stops and mechanical vignetting For an off-axis object point, the aperture stop is not the only aperture limiting the bundle of rays passing through the optical system. Other apertures may limit the bundle of rays. These other limiting apertures are called field stops and there can be more than one, see Figure 1.10a. The image of a field stop in object and image space are called the entrance and exit windows, respectively. The image plane of the optical system can be separated into three different field areas.

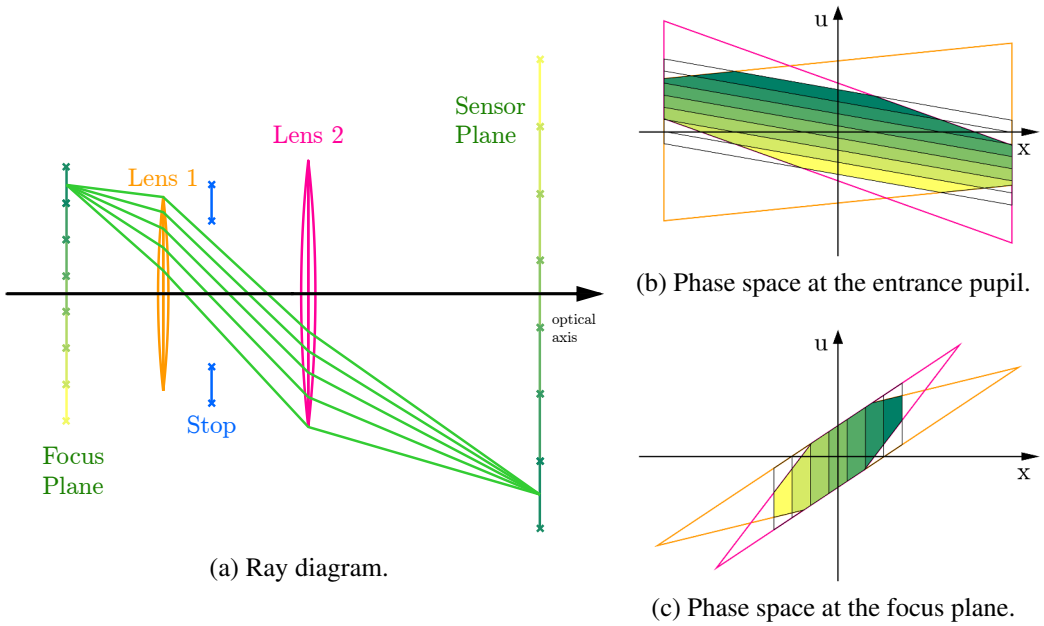
The first one exists close to the optical axis when the only light limiting element for a point in the field is the stop. This field is not subjected to geometrical vignetting and is often called the clear field.

The second area encompasses the first one and the cross-section of the cone of rays going through the stop is partially cut by the field stops. The shape of this cross-section is usually complex as multiple field stops can affect it but most of the time there is only a single window producing a biconvex shape which has the shape of a cat's eye and the reason why geometrical vignetting is also called the "cat's eye effect". The irradiance on the field decreases radially in a direction from the optical axis for centered optical systems. This area is called the vignetted field.

The third area encompasses the second and is made of the image field points that do not receive light through the optical system. The frontier between the second and the third area is the limit of the whole field which is consequently the union of the clear and vignetted field.

This vignetting can be reduced by closing down the aperture stop. In effect, the clear aperture region becomes larger, the full field becoming smaller so the vignetted region is thinner. The task of the optical designer is to match the sensor size and the clear field as not to waste pixels in a part of the field where the SNR is low.

Radiometry and natural vignetting The second source of vignetting is related to the radiometric aspects of the system. An off-axis image point even if located in the clear aperture region suffers from a decrease in received light flux. The cone of light coming through the stop aperture is tilted and the stop does not appear as a circle as compared to an on-axis point but as an ellipse of smaller area. Actually, the solid angle of this point looking at the stop decreases as it moves further away in the field, except for object-side telecentric systems.



(a) Ray diagram.

(b) Phase space at the entrance pupil.

(c) Phase space at the focus plane.

Figure 1.10: Mechanical vignetting of an off-axis object point. (a) The field stops are the rims of the two lenses. (b) and (c) The phase space regions of these field stops (in orange and pink) cut the phase space cells of the pixel (outlined in black). The intersections are the yellow to green colored regions that represent the part of the ray bundle from each pixel that passes through the system.

The pixel plane and the exit pupil plane are usually perpendicular to the optical axis. Thus, the etendue of the radiation transfer between a small area in each plane linked by a ray of length r can be computed from Equation 1.6 as:

$$\begin{aligned} d^2G &= \frac{\cos\theta_{pix}\cos\theta_{EX}\,dA_{pix}dA_{EX}}{r^2} = \frac{\cos^4\theta\, dA_{pix}dA_{EX}}{z^2} \\ &= d^2G(\theta=0)\cos^4\theta \end{aligned} \tag{1.24}$$

because $\theta_{pix} = \theta_{EX} = \theta$ and with $z = r\cos\theta$ being the distance between the two planes. The intensity variation follows the well-known $\cos^4$ law with the angle of the ray which is directly related to the position of the point on the field. This vignetting is only dependent on the distance between the exit pupil and the sensor, as well as the sensor size.

Often, an optical system suffering from vignetting is calibrated using an image of a uniform lambertian white scene. Raw pictures can then be devignetted using this radiometric calibration picture.

1.2.6 Analyzing an Optical System

Given a certain optical system, one of the main task of the designer is to analyze it and report on its several properties like: the image location and magnification, the

effective focal length of the system, its f-number, its angular field of view, its depth of field, etc. The process to compute these information is to use a table regrouping the optical surfaces in columns and their properties (position, thickness, aperture size) in lines, see [O'Shea, 1985] for some more detailed information. The major steps of the procedure are the following:

- set the system data in the table,
- run an axial ray and find the aperture stop,
- run a chief ray and find the field stops,
- compute the positions and sizes of the (entrance and exit) pupils and windows,
- compute the desired properties.

Running rays and finding the position of the pupils and windows is done employing the paraxial invariant or the ray transfer matrices.

1.2.7 Phase Space

We have seen in Section 1.1.5 that a bundle of rays is limited by two apertures, one being the pixel aperture and the other being the pinhole aperture . The bundle such defined delimits a region of phase space, a cell, the volume of which is the etendue of the bundle. The phase space diagram is also an interesting tool to study the ray propagation in an optical system and to derive the optical properties and even to study aberrations.

Propagation By choosing a reference plane to evaluate the phase space perpendicular to the optical axis and moving it in the direction of the optical axis, it is simple to visualize the propagation of light rays in the optical system. For free space propagation, we have seen that the transformation of the phase space is a simple horizontal shear the amplitude of which is given by the translation shift of the reference plane. At the plane of a thin lens, light rays abruptly change direction without changing their position. The transformation of the phase space is a vertical shear with an amplitude given by the focal length of the thin lens.

For an unvignetted pixel, the limiting aperture is the stop, so the ray bundle integrated by the pixel corresponds to a phase space cell. Consecutive pixels are represented in phase space as a series of phase space cells, one for each pixel, that are contiguous. Starting at the sensor plane and back propagating light rays through the phase space plane to the object space of the camera transforms the phase space by a sequence of horizontal and vertical shears.

When vignetting occurs on the ray bundle of an off-axis pixel, the original cell shape shrinks as the ray bundle is limited by the individual optical elements, see Figure 1.10. At some point, the cell disappears and the volume of the cell is reduced to zero. The corresponding pixel is out of the camera field of view.

Properties Once in object space, the different optical properties of the camera can be either visualized directly or extracted from the phase space diagram. When the reference plane is at the entrance pupil position, see Figure 1.10b, all the cells are vertically aligned in a contiguous region and the pupil size is the horizontal extent of this region.

The extent along the x-axis is the angle of view. When the reference plane is the focus plane, see Figure 1.10c, all the cells are next to each other with their sides parallel to the u-axis. The extent along the u-axis of a cell gives the angle to compute the numerical aperture and the horizontal extent gives the field size.

Aberrations When a system is subject to aberrations, rays deviate non-linearly causing the paraxial shape of the cell to deform, see [Ng and Hanrahan, 2007]. Cells close to the origin of the diagram are less affected as they correspond to ray bundles close to the optical axis where the paraxial conditions are respected the best. The deformation is worse for cells far from the origin. The shape of the cells is also symptomatic of the types of aberrations present in the optical system.

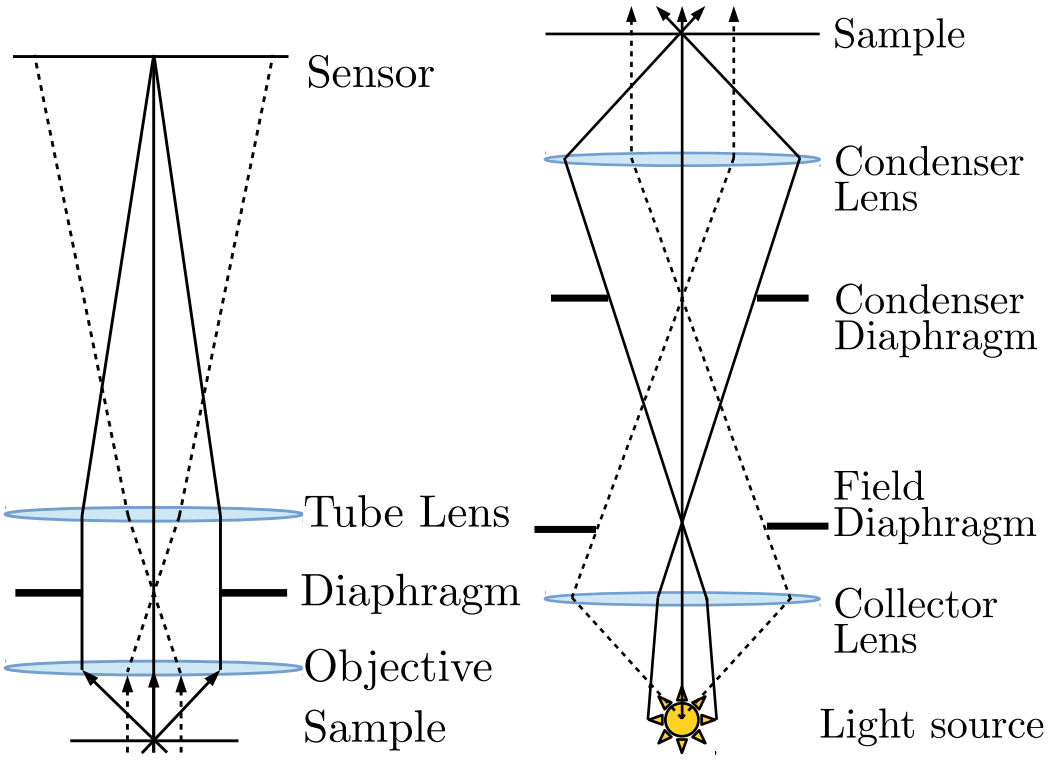
1.3 Microscopy

Microscopes are optical instruments designed to image small objects *e.g.,* geological samples. In biology, for instance, the range of the sample's scale can vary from millimeters when studying living or dead cell tissue to micrometers when locating specific parts inside a cell. Imaging such tiny objects is difficult because the required magnification of the microscope objective used to observe them is large (with a factor from 5 to 100) and because the resolving ability of microscopes is fundamentally limited by the wavelength of the light (between $400nm$ and $800nm$ for visible light).

1.3.1 Properties

Microscopes can be complex systems but the fundamental optical recipe is to have a microscope objective, working in a high magnification configuration, enlarging the object and projecting it either for the eye of the observer through an eyepiece or onto a sensor with or without the use of relay optics, see Figure 1.11a.

Modern microscope objectives are infinity-corrected, *i.e.,* they are designed to image a sample placed in their front focal plane and they need a tube lens to form the image. They can also correct for aberrations introduced by the glass plate covering the sample. The focal length of a microscope objective is a few millimeters and the focal length of the tube lens is usually between $180mm$ and $200mm$, the ratio of the focal lengths gives the overall magnification. As explained earlier, microscope objectives are typically object-side telecentric in order to have a constant magnification over the depth of field.



(a) Imaging part (b) Illumination part

Figure 1.11: Paraxial ray diagram in a microscope. The Köhler illumination is placed below the microscope objective so that the sensor is at the top, the light source is at the bottom and the sample plane is in-between.

The characteristics of a microscope objective are its magnification, its numerical aperture and its working distance. These values are usually printed or carved onto the objective case. The working distance is the distance between the object and the first surface of the objective. The choice of the magnification is to be made considering the size of the sample that is under study. Larger magnification is necessary to image smaller objects but the field is also reduced. Placing the sample on a motorized stage allows to extend easily the field by moving the region of interest.

The numerical aperture controls the amount of light captured by the objective but it also indicates the resolution of the microscope. Following Abbe's theory [Singer *et al.*, 2006] the lateral resolution $r_{x,y}$ of the microscope is limited by the diffraction limit and is given by the formula:

$$r_{x,y} = \frac{\lambda}{2NA}\,, \tag{1.25}$$

with λ the wavelength of the light used for illumination. A high numerical aperture is sought for good resolution but it hits a physical limit in air. In order to increase the numerical aperture, microscope objectives can be designed to work specifically

in immersion in another medium such as water or oil.

The axial resolution or depth of field r_z of the microscope is given by:

$$r_z = \frac{2\lambda}{NA^2} \,. \tag{1.26}$$

The relation with the numerical aperture is even stronger for the axial resolution. The consequence is that only a tiny slice of the sample can be imaged sharply. The reconstruction of a 3D sample is done by scanning the sample in depth with a motorized stage.

1.3.2 Illumination Techniques

Microscopes are most often used to study samples in transmission mode where light comes from behind the sample. The dynamic range of the image or its contrast is also an important aspect of the image quality. The contrast is controlled by the illumination system which uses different interactions between the sample and the light. Various techniques have been developed in order to increase the contrast. Even though we use only the bright field and reflection techniques in Chapter 6, we also describe other techniques for completeness.

Bright field and Köhler illumination Many samples especially biological ones are transparent. Bright field microscopy uses the difference in absorption of the different parts of the samples to reveal their structures. One of the main problems with this absorption method is that the parts of the sample that are out of focus also absorb light which decreases the contrast of the image. Some samples are also too transparent to produce a good contrast. Contrast is usually enhanced by staining the parts of interest to increase their absorption. This requires additional preparation and can eventually modify the structure of the samples.

Moreover, a specific way to illuminate the sample from the back, called Köhler illumination, was developed to diminish the impact of the loss of contrast due to the non-uniformity of the light source. It relies on two lenses: a field lens and a condenser lens, usually combined with their own diaphragm, see Figure 1.11b. The field lens images the light source to the plane of the condenser lens so that the condenser diaphragm controls the amount of light reaching the sensor. The condenser lens, on its end, images the field lens to the sample plane so the diaphragm of the field lens controls the extent of the field that is illuminated. This way, the position of the image of the light source and the location of the sample are not conjugated and they are independent. The condenser diaphragm also controls the numerical aperture of the illumination cone and, consequently, the effective numerical aperture of the entire microscope. It is often adjusted to match the numerical aperture of the microscope objective. The field diaphragm is also adjusted to match the actual field restricted by the microscope objective in order to decrease the amount of stray light in the system.

Dark field Dark field microscopy enhances the contrast by using a ring-shaped cone illumination at a higher angle than what the microscope objective can accept. Most of the time, this configuration is achieved by placing a ring diaphragm at the condenser lens. Only the light that is scattered by parts of the sample is sent to the sensor where they appear as bright spots on a dark background. The main limitation of this technique is the low level of light intensity reaching the sensor as most of the light is directly transmitted and cut out by the objective.

Polarization Some samples such as crystals or cell tissues modify the light polarization. This property is called birefringence and is highly dependent on the wavelength of the light. The two orthogonal directions of polarization propagate in the same medium but they are effected as if two media with a slight difference in optical index were present. The optical path difference between these two polarizations is the product of the index variation and the thickness of the sample.

A microscope can be modified to exploit this phenomenon. A polarizer and an analyzer are placed before and after the sample respectively in a cross configuration. As a result of both components only letting pass through a certain polarization direction, the light going through a non-birefringent sample is cut out by the analyzer. Light that passes through a birefringent medium thick enough to rotate the polarization direction by a half turn is also cut. However, when the sample is generally birefringent, the light polarization changes direction and passes through the analyzer. By using a white light source and monitoring the wavelengths that disappear, the birefringence or the thickness of the sample can be computed if the other parameter is known.

Phase contrast Phase contrast microscopy also uses the variation of the optical index of different parts of the sample but it converts the shift in the light phase onto a change in the light brightness. The operating principle is close to dark field illumination. The idea is to treat the background and sample light differently. Firstly, a ring is used as a condenser diaphragm but at a numerical aperture below that of the objective. Secondly, the sample transmits some of the light directly and the rest is scattered. The background light goes straight while the sample light has different directions. Finally, a phase shift ring in a plate is inserted after the objective to generate constructive interference between the background and the sample light. The contrast is even more enhanced by using a ring-shaped dimmer in a plate after the first one. The sizes of the rings of the phase shift and dimmer plates are chosen to affect only the background light. Phase contrast microscopy is a successful technique that does not require staining.

Reflection The surface of opaque objects can be studied at a microscopic scale using a reflection technique. The light is brought from the light source to the sample through the microscope objective by using a beam splitter positioned at $45°$ in the

light path.

1.3.3 Limitations

Apart from the diffraction limit, there also exist microscopy techniques that exploit the wave nature of the light to extract different types of data (contrast phase microscopy, ...). This is possible by using the fact that the back focal plane of an infinity-corrected microscope objective is also called the Fourier space of the objective as it is the plane where the Fourier transform of the object lies. Fourier filtering can be applied physically in this plane by using specific amplitude or phase plates. These effects cannot be described with a geometrical description employed in this thesis.

1.3.4 Light Field Microscopy

Light field microscopy was first devised in [Levoy *et al.*, 2006]. The light field capability is obtained by inserting a microlens array in the image plane of the microscope objective and by moving the sensor plane to the back focal plane of the microlens array, see Figure 1.12. This configuration reproduces the design of [Ng *et al.*, 2005a] where it was employed to extend a standard camera for light field imaging. This modification may require the use of additional optics as a relay group to image the back focal plane of the microlens to the sensor to make the alignment between the various components easier. This relay group can also serve as an additional small magnifier group. Due to the location of the microlens array, it samples the object plane spatially. The pixels under each microlens sample the cone of light reaching a microlens, they act as directional samples. The microlens array allows to multiplex the 2D spatial and 2D directional information onto the 2D sensor.

Design Levoy and colleagues [Levoy *et al.*, 2006] detail the several constraints that govern the design of the different components on the various parameters: the magnification and numerical aperture of the microscope objective, the pixel size and number, the focal length and pitch of the microlenses as well as their overall number.

The first important constraint is matching the effective f-number of the microscope objective and that of the microlenses. The effective f-number of a microscope can be rewritten from Equation 1.21 with the Helmholtz-Lagrange invariant as:

$$F = \frac{1}{2NA'} = \frac{M}{2NA} \,. \tag{1.27}$$

In order for the microlens image not to be too large which would create overlap with the neighboring microlens images, nor too small, which would decrease the angular sampling and leave unused pixels, the effective f-number of the two components

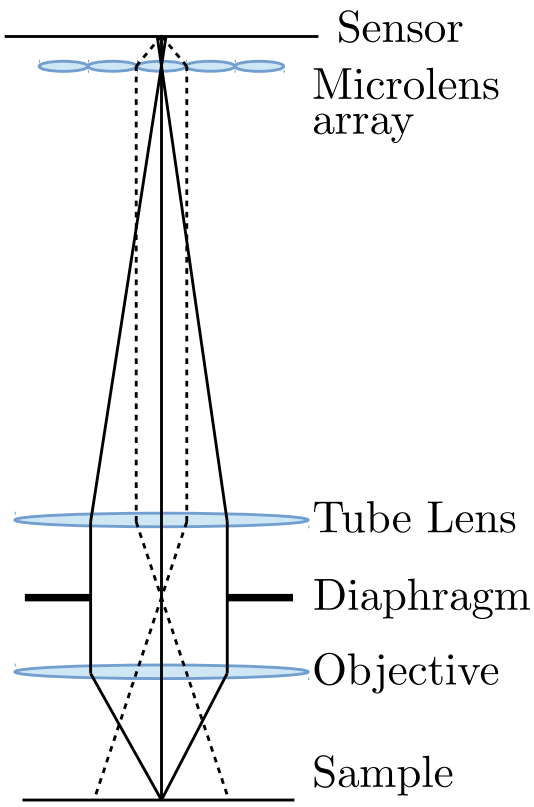


Figure 1.12: Ray diagram in the light field microscope as designed by [Levoy *et al.*, 2006]. The pitch and focal length of the of the microlenses has been exagerated for clarity.

has to be equal. The f-number of a microlens is simply the ratio of its focal length to its diameter (see Equation 1.20).

The second constraint is the diffraction limit as it sets the minimum pixel size. Consequently, the only remaining free parameter is the microlens pitch or equivalently, the choice of the trade-off between the spatial and angular resolution or equivalently again, the number of pixels under a microlens.

Applications The synthetic images that can be generated from the light field extend what the microscopist observes usually: an orthographic view of a 2D slice of the sample. Levoy and colleagues [Levoy *et al.*, 2006] show that, from a single light field capture, perspective views of the sample can be synthesized. Moreover, a refocus stack is typically obtained by translating the stage vertically and taking multiple images of the sample which takes time. Light field reconstruction techniques produce a refocus stack from a single image but a the cost of a lower spatial resolution.

Furthermore, 3D reconstruction can be performed by deconvolution using the 3D point spread function (PSF) of the system. This 3D PSF is acquired by imaging

a tiny fluorescent bead with the light field microscope and the volumetric model is obtained using standard iterative algorithms.

Light field illumination In the continuity of this work, another paper [Levoy *et al.*, 2009] replaces the Köhler illumination with light field illumination using again a microlens array and a projector. The whole system is described as well as the design choice regarding the parameters of the microlens array. Having control over the illumination allows to recreate several techniques of microscopy such as bright or dark field but also to simulate oblique directional lighting of different inclinations and extents. Generating an incident light field and recording the reflected light field allows to study the reflection and transmission properties of materials but also to correct for the aberrations introduced by the system.

Chapter 2

Related Work

Having established the optical principles of light field imaging, this chapter gives an overview of the current literature on light fields, covering aspects from acquisition to rendering. The contributions of this dissertation concern mainly the development of a theoretical first-order model for light field cameras as well as its applications to actual systems. Therefore, this chapter focusses on related works that concentrate on the description of light field acquisition systems as well as the theoretical models that were developed to characterize these systems.

2.1 Light Fields

The previous chapter gave a good understanding of what the light field function is: the radiance of a set of light bundles simplified as rays. Many state of the art, survey and review papers about computational imaging exist, see [Levoy, 2006; Zhou and Nayar, 2011; Wetzstein *et al.*, 2011; Lam, 2015; Ihrke *et al.*, 2016] for a good overview of the topics addressed in this section. These topics encompass measuring, calibrating, processing and rendering from a light field in order to generate 2D pictures from the 4D data.

2.1.1 Integral Imaging

More than a century ago, stereograms were the main items that would reconstitute the depth dimensionality of a scene and they were produced with a two-camera setup. Ives [Ives, 1903] describes how to make stereograms with a single camera. The device is made of a main lens with an opaque mask containing two small holes separated by the average human eye-to-eye distance on its back, see Figure 2.1. The photographic plate is covered with a parallax barrier (a dense set of opaque and transparent lines regularly spaced) so that the image formed through both holes are interlaced. Once the photograph is developed, the stereogram is viewed using the same parallax barrier.

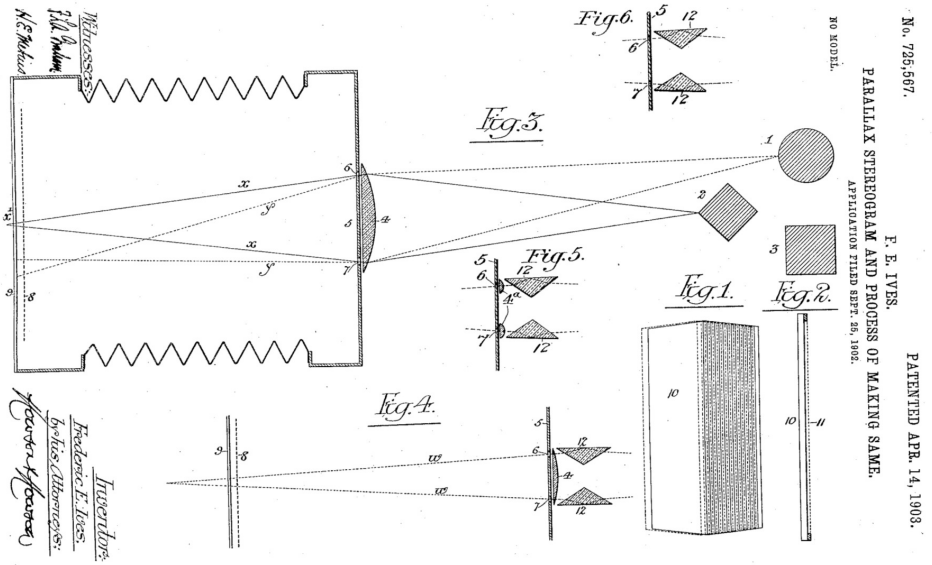


Figure 2.1: Patent sketch for the parallax stereogram making camera from [Ives, 1903]

Later, but long before the term "light field" was invented, an invention by Lippmann [Lippmann, 1908] marks the beginning of integral imaging. The device is made of an array of lenslets placed next to each other onto a photographic plate. Then, it is put in a dark chamber and exposed to the scene to be captured as a regular photograph would be taken. Each lenslet produces an image of the scene so that n lenslet produce n images from different positions. The plate is treated and reveals an array of images. By placing the array of lenslets in front of this photograph, a viewer sees a 3D continuous still image of the scene.

These two implementations set the principles of light field imaging: separating light rays by their incoming direction on the sensor using special optics and reconstructing the 3D image. In the classical implementations, the reconstruction is done using the same special optics in reverse.

2.1.2 The Plenoptic Camera

Adelson and Wang [Adelson and Wang, 1992] combine the principles of integral imaging into *the plenoptic camera*. This camera is made of a main lens which images the scene onto a pinhole array behind which a CCD sensor is located. Each pinhole creates a micro-image of the aperture of the main lens on the sensor. The pinhole array may be replaced by a lens array for better light efficiency, see Figure 2.2a. An improvement of the system is also envisioned. It involves placing a field lens just in front of the lens array and using a relay lens to separate the lens array from the sensor, see Figure 2.2b.

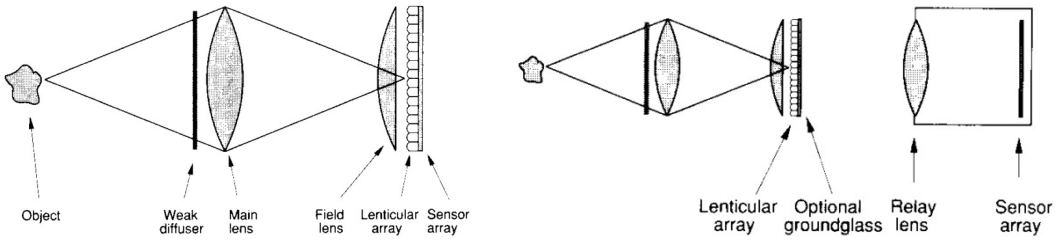


Figure 2.2: Plenoptic camera [Adelson and Wang, 1992] (a) Original setup. (b) A relay system is added after the microlens array.

The paper also discusses several design considerations that are important for many different light field capturing devices.

The first is the **trade-off between the spatial and the angular resolution**. For a fixed number of sensor pixels, the number of pixels of a micro-image defines the angular resolution and the number of pinholes or microlenses defines the spatial resolution. If the micro-images are made small, then the array can be more densely packed, angular resolution is lost in favor of an increase in spatial resolution. Generally, a sensor with a high definition and small pixels is preferred.

The second is the **f-number matching rule**. The f-number of the main lens should be equal to the f-number of the microlenses so that the micro-images are neither too large nor too small. In the first case, the pixels belonging to the overlap region between neighboring micro-images receive light from both so they cannot be used for the reconstruction. In the second case, some pixels do not receive light.

The third relates to **the system aperture**. The n pixels (in 1D) behind a pinhole or microlens, sample the system aperture. A subview can be generated by associating the same pixel of each micro-image. The effective size of the aperture of this subview is $1/n$-th the size of the full aperture. Consequently, the depth of field of the subview is increased in proportion to n. Moreover, the baseline *i.e.,* the maximum shift between two subviews is limited by the aperture size. The baseline sets the parallax sensitivity of the subviews and so, it sets the depth disparity of the optical system. Hence, a small baseline limits the accuracy of the depth disparity. Generally, the plenoptic camera is designed to have a large aperture *i.e.,* a small f-number. The loss of depth of field due to the large aperture can be compensated by the number of pixels n (in 1D) in the micro-image.

The main improvement of [Adelson and Wang, 1992] over the cameras from [Lippmann, 1908] and [Ives, 1903] is to decouple the recording part and the synthesis part of the system. Unlike the previous systems, the reconstruction of the image is not done with the same optical element that separates light rays by their direction, the pinhole or lens array, but it is done computationally. Subviews are synthesized by selecting the appropriate pixels. The paper also describes a method to recover

depth disparity from the parallax shift between different subviews.

Considering the definitions of the plenoptic function and the light field function in Chapter 1, the name *plenoptic* for the camera from [Adelson and Wang, 1992] is not appropriate. The device belongs more to the group of *light field cameras*. This naming regroups other camera designs using a single sensor. Several implementations are presented in Section 2.2.

2.1.3 Alternative to Light Field Cameras

The main constraint of the plenoptic camera or more generally, of the light field cameras, is the spatio-angular resolution trade-off as a result of all the 4D light field information being multiplexed onto a single 2D sensor. Its main issue is the low spatial resolution of the final images. There exist other ways than using a plenoptic camera to capture light fields that overcome this problem.

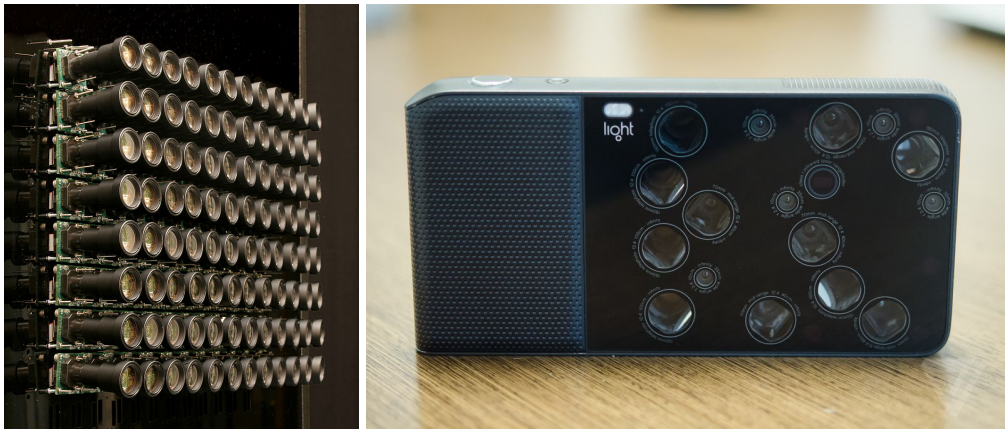
Camera arrays Stereo-camera systems use a pair of cameras, mimicking the binocular vision of humans and many other animal species. They record two images from the scene which are assembled in a stereogram to infer a depth illusion to the viewer. The most recent and popular applications of this concept are 3D movie technology and augmented and virtual reality headsets.

A more thorough sampling of the light field of a scene is done by taking more than 2 views from multiple angles and positions. This is simply implemented by using an array of cameras, see Figure 2.3, either in the form of a dome surrounding the object of interest [Rander *et al.*, 1997], or with a flat array capturing a large part of the scene but from a single side [Yang *et al.*, 2002].

In practice, this approach necessitates a good calibration procedure to determine the ray-to-pixel correspondences. Moreover, a large number of cameras is required in order to increase the angular resolution which is limited by the available bandwidth of the cameras and the computer hardware. Even so, the angular component of the light field is still sparsely sampled as there always is a gap between neighboring cameras.

However, the spatial resolution of the light field is determined by the resolution of the cameras in the array which only depends on the quality of the camera hardware. Such a setup also offers the possibility to record video light field of moving scenes. Wilburn and colleagues [Wilburn *et al.*, 2004, 2005] implemented the first light field video system.

Camera arrays are costly and cumbersome which prevents their use in many applications where data storage, physical extent, and computational power are limited. See [Wilburn *et al.*, 2004, 2005; Xu *et al.*, 2015] for more details on the major features and applications of camera arrays. There exist a few commercial applications like the soon to be available Lytro Immerge for the movie and virtual reality industries, see [Lytro, 2017].



(a) (b)

Figure 2.3: Camera arrays. (a) Camera array from Stanford University [Wilburn *et al.*, 2005] with 12 by 8 cameras. (b) Light L16 camera [Light, 2017]. It is made of 16 cameras arranged in an irregular configuration: 5 with a 18mm focal length, 5 with a 70mm focal length, and 6 with a 150mm focal length. Copyright for the picture: https://www.theverge.com.

Moving the viewpoint The other main strategy for acquiring a light field is to take several pictures with a single camera over time (or to record a video) and to move the camera between each shot, see [Gortler *et al.*, 1996; Levoy and Hanrahan, 1996]. This method is restricted to still objects but it removes the need for multiple cameras. The change of viewpoints can be done either by tracking the camera position and orientation (or by using a robotic arm) or by moving a mirror that generates multiple virtual viewpoints [Ihrke *et al.*, 2008; Taguchi *et al.*, 2010a]. The angular resolution is determined by the granularity of the movement between consecutive shots so temporal multiplexing methods are able to produce dense light fields. The additional storage capacity required for a dense light field is balanced by the high compression ratio of the data by cause of the consecutive views differing slightly and most of them being redundant, see [Magnor and Girod, 2000; Levoy and Hanrahan, 1996].

2.1.4 Analysis, Processing and Rendering

Whatever light field imaging device is used, a light field is produced after a step of proper calibration and preprocessing specific to the device configuration. For light field cameras, for instance, naive straightforward rendering from the light field raw data produces aliased low-resolution images [Adelson and Wang, 1992]. Then, in order to render high quality 2D images, the light field has to be analyzed and processed to prevent aliasing when interpolating the missing data and to remove noise [Levoy and Hanrahan, 1996; Gortler *et al.*, 1996].

The light field has an inherent specific structure determined by the geometry and reflectance properties of the scene, see [Wanner and Goldluecke, 2012]. A 2D spatio-angular cross section of the light field also known as an epipolar plane image (EPI) shows a linear structure in case of a Lambertian scene. Each line segment corresponds to a single object point on an object surface and the slope of this line indicates the relative depth of the object in the scene. The EPIs serve as constraints for depth reconstruction [Levin *et al.*, 2008]. Depth estimation is linked to refocusing as was shown in Section 1.1.5. In an EPI which is a radiance function on a slice of the phase space, vertical lines indicate that the object is in the focal plane. Refocusing shears the EPI, changing which lines are vertical and brings objects at other depths into focus.

The reconstruction of the light field was also considered as a signal processing task that can be performed in the frequency or Fourier space. Anti-aliasing, interpolation and denoising are fundamentally treated the best in the frequency domain, the novelty of light field filtering being its four-dimensional nature. Previous works have addressed several topics such as sampling rates [Chai *et al.*, 2000], sparsity, upper limits to the reconstruction quality, fast rendering [Ng, 2005; Pérez *et al.*, 2012], and denoising [Dansereau *et al.*, 2013a].

Some works in the literature explore how to take advantage of the light field cameras image formation specificity to improve the reconstruction task. Their main target is to achieve light field super-resolution. Such techniques [Bishop *et al.*, 2009] in conjunction with better depth estimation algorithms [Bishop and Favaro, 2012] have been studied using filtering in the Fourier domain [Liang and Ramamoorthi, 2015] to improve the spatial resolution and quality of the output images. Levin *et al.* [2008] and Georgiev *et al.* [2008] compare the performance and trade-offs of different camera designs in sampling the light field.

One of the main attractions towards light field imaging is its application to synthetic aperture photography [Ng *et al.*, 2005b]. A single light field acquisition allows to synthesize new images as if they were taken with a different lens. The possibilities are numerous and go beyond what a standard camera can achieve. It includes for instance, producing a refocus stack, changing the viewpoint or focal length continuously, and setting the size of the aperture. The last aspect controls the extent of depth of field so that an all-in-focus image can be produced. The synthetic images can also go beyond what is feasible with traditional imaging , for instance, both the foreground and background can be made sharp but not the objects in-between. In the movie industry, a split focus diopter can reproduce this effect but it simply split the image in two parts focused at different depths using a convex half lens. In Computer Graphics, lens effects can be simulated in real time [Lee *et al.*, 2010] from a physically-based model. However, control can also be given to the user for expressive rendering where the focus and depth of field are specified manually across the scene.

The optical model developed in this dissertation describes the light field sampling properties of a given light field camera design. It makes the connection be-

tween the optical conditions of the light field capture and the data produced by the reconstruction procedures. Moreover, it determines values for the physical limits to the absolute performance of the algorithms, like the maximum range of refocusing.

2.2 Light Field Cameras

The common trait of the light field cameras from [Ives, 1903; Lippmann, 1908; Adelson and Wang, 1992; Georgiev and Intwala, 2003; Ng *et al.*, 2005b; Veeraraghavan *et al.*, 2007; Manakov *et al.*, 2013] is to separate light rays by the angle at which they hit the sensor. In the following, several designs are presented along with their operating principles and imaging peculiarities. The large diversity of these designs makes it complicated to get a good understanding of their strengths and weaknesses and to determine how well they would perform for a specific application. The goal of the model presented in this dissertation is to set a common ground for these acquisition systems as well as some insights on their limitations and potential improvements. At the same time, each of these designs can serve as a validation system for our model through the measurement of its optical properties.

2.2.1 Monolithic Camera Array

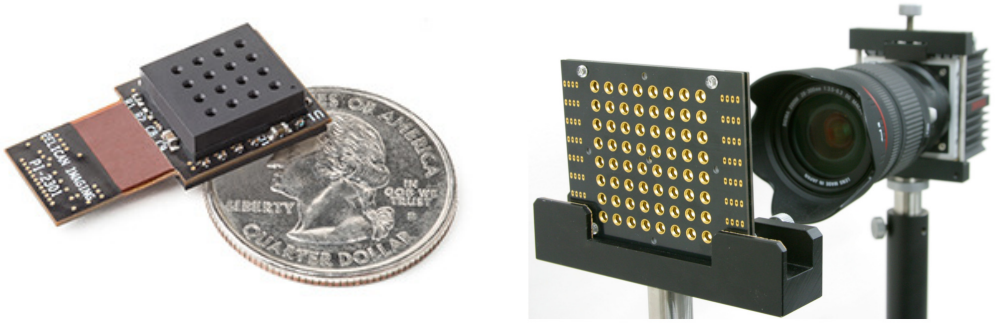
Small camera arrays can be made using a single sensor behind a lens array where each lens and the pixels behind it act as one camera. This system simplifies the calibration of multiple independent cameras, facilitates its miniaturization [Venkataraman *et al.*, 2013] and is physically very thin, see Figure 2.4a. The spatial resolution is controlled by the number of pixels per lens and the angular resolution by the number of lenses. A drawback of this design is that the baseline is quite small. A variation of this system employs a lens array that covers a flatbed scanner [Yang, 2000].

On a side note, Shack-Hartmann sensors [Platt and Shack, 2001] are built of the same hardware but made for a different purpose. They are commonly used to measure optical wavefronts in adaptive optics systems, see [Singer *et al.*, 2006]. Their spatio-angular resolution trade-off is shifted strongly towards the angular domain.

2.2.2 External Lens Array

In order to get a large baseline together with a single sensor, one possibility is to shoot a picture with a regular camera through an external large lens array, see Figure 2.4b.

Georgiev and Intwala [Georgiev and Intwala, 2006] investigate different arrangements of the optical components: a large lens or prism array placed before or after a large main lens. They implement an external lens and prism array without the large lens as their best trade-off.



(a) (b)

Figure 2.4: Devices using an external lens array. (a) Miniature monolithic camera array from [Venkataraman *et al.*, 2013], possibly for mobile phone applications. (b) External tunable lens array from [Ueda *et al.*, 2008b].

Yamamoto and Naemura [Yamamoto and Naemura, 2004] include a large Fresnel main lens in front of a gradient index lens array. Ueda and colleagues [Ueda *et al.*, 2008a] improve this setup by having a variable-focus lens array where the focal length of each lens is independently controlled. They call this technique Adaptive Integral Photography (AIP) and use it to optimally sample the light field depending on the scene.

2.2.3 Microlens Array

Placing a microlens array after the main lens and close to the sensor makes the system less bulky compared to the ones with an external lens array. As we have seen, the idea to use a pinhole array in front of the sensor is also the historical way to make a light field camera and is still valid [Chen *et al.*, 2010] but a microlens array is more light efficient. The concept was refined and two different designs were studied: the afocal and focused light field cameras, see Figure 2.5.

The Afocal Light Field Camera

In the camera devised by [Adelson and Wang, 1992], each microlens images the main lens aperture and the microlens array spatially samples its conjugate plane in object space which is the focal plane of the camera. [Ng *et al.*, 2005b; Ng, 2006] improve this design by explaining that making the distance between the microlens array and the sensor equal to the focal length of the microlens almost ensures the imaging of the lens aperture. In effect, the distance from the main lens aperture to a microlens is greatly larger than its focal length. The microlenses are focused at infinity, so we call this system "afocal". This configuration also eliminates the need for the field lens that is present in the original system.

(a) Lytro camera (b) Lytro Illum (c) Raytrix R11

Figure 2.5: Commercial light field cameras. The first (a) and second (b) generation Lytro cameras implement the afocal configuration for everyday consumer use [Lytro, 2017]. The Raytrix camera (c) implements the focused configuration and has applications in the industry [Raytrix, 2017].

As a result, the sensor pixels sample the lens aperture and a point in the aperture plane is imaged to the same position on the micro-image of every microlens. A viewpoint is directly extracted by selecting the same pixel behind each microlens.

The paper also clarifies the f-number condition for the afocal camera. It is the image side f-number of the main lens, the ratio of its diameter to its distance to the microlens array that must be equal to the f-number of the microlenses.

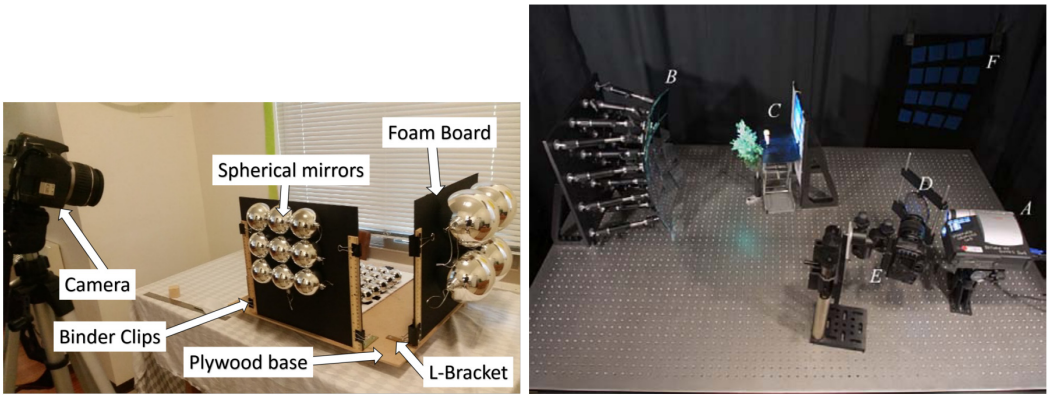
The Focused Light Field Camera

The major downside of the afocal camera is its low spatial resolution in the plane of focus because it is where the microlens array is located. The focused light field camera [Lumsdaine and Georgiev, 2008, 2009] separates the image plane from the array plane and, consequently, changes the imaging task of the microlenses. They do not image the main lens aperture onto the sensor but they image the image plane of the main lens. Therefore, the distance separating the microlens array and the sensor is not equal to the focal length of the microlenses anymore.

Two configurations exist depending on whether the separation between the microlens array and the sensor is smaller or larger than the focal length of the microlenses. In the former case, the image plane of the main lens is situated behind the array, it is virtual for the microlenses. In the latter case, the plane is in front of the array and real.

This modification of the design of the afocal configuration, also known as the standard light field camera results, in a decrease of the angular sampling of the light field in proportion to an increase of the spatial sampling. The trade-off is controlled by the magnification performed by the microlenses. The spatial sampling is not dictated by the number of microlenses but it is a function of the micro-image resolution and the amount of shared information between neighboring micro-images. Because of this improvement, the focused light field camera is also called the plenoptic camera 2.0.

The rule for not having an overlap of the micro-images is to match the effective



(a) (b)

Figure 2.6: Catadioptric setups. (a) using an array of metallic spheres by [Krishnan, 2016]. (b) using a flat mirror array from [Levoy *et al.*, 2004] for confocal light field imaging.

f-number of both the main lens and the microlenses.

In order to extend the depth of field of the camera, Perwass and Wietzke [Perwass and Wietzke, 2012] use a hexagonal arrangement of microlenses with three different focal lengths. It produces three focal planes at different depths and the focal lengths are chosen so that the depth of field of each focal plane touches the other ones forming a large sharp region.

A generalized light field camera in [Ng, 2006] extends the afocal design where the distance between the sensor and the microlens array is tunable. The study concludes that it is best to have this distance be lower than the focal length of the microlenses to allow for a control over the spatio-angular resolution trade-off.

This dissertation provides an analytical formulation of the properties of the afocal and focused configurations as well as an additional comparison study of these two light field cameras, see Chapter 5. Additionally, an experimental setup with an afocal light field camera is realized and serves as a validation procedure of the ECA model developed in this thesis.

2.2.4 Catadioptric Arrangement

Lens arrays are not the only optical components used to capture a lightfield. A combination of cameras and external spherical mirror arrays widens the field of view of the camera, see [Krishnan, 2016] illustrated in Figure 2.6a. Unger and colleagues [Unger *et al.*, 2003] capture the light field of the illumination conditions of a scene for further rendering of 3D objects in the same lighting environment. Lanman and associates [Lanman *et al.*, 2006] discuss the construction and calibration of the system taking into account the differences between a first surface mirror and

a second surface mirror. Taguchi and colleagues [Taguchi *et al.*, 2010b] present a model for rotationally symmetric mirrors, they justify that spherical mirrors are better than other mirrors because they are rotationally symmetric in all directions. A planar array of cameras generates a virtual array of viewpoints, images of the camera perspective center. Levoy and associates [Levoy *et al.*, 2004] build such a system with a camera and a projector for both light field capture and illumination to achieve synthetic aperture confocal imaging, see Figure 2.6b. Fuchs and colleagues [Fuchs *et al.*, 2013] investigates a pipeline for designing, fabricating, and utilizing faceted mirror arrays for light field capture. The final mirror design is optimized automatically from the user's input constraints.

Exploiting inter-reflections in mirrors systems, the number of virtual viewpoints can be increased considerably up to the point of realizing hemispherical [Reshetouski *et al.*, 2011] or even spherical [Ihrke *et al.*, 2012] viewpoint coverage.

2.2.5 Light Pipe

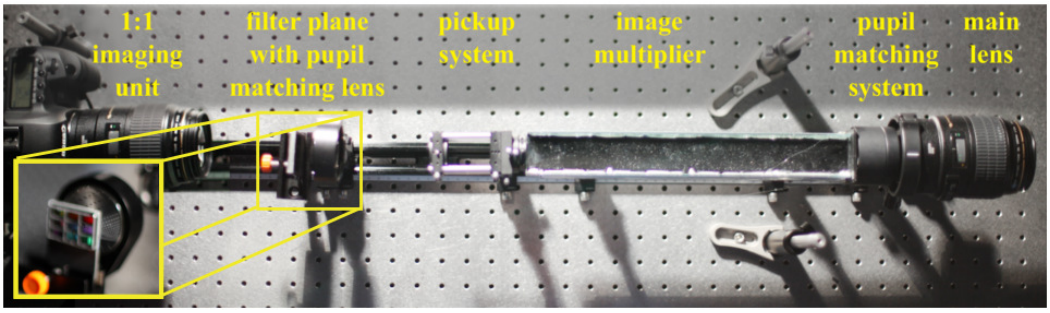
Manakov and associates [Manakov *et al.*, 2013] propose a whole different optical design for light field cameras that is inspired by kaleidoscopes. It uses a light pipe in the shape of a rectangular tube made of two pairs of parallel mirrors facing inwards and places it behind the main lens of a regular DSLR camera, see Figure 2.7a. Two additional lenses are added at each end of the tube to reduce vignetting effects, see Figure 2.7b. The first end of the pipe on the main lens side is located at the image plane of the main lens. The lens at the sensor side is chosen to conjugate the sensor plane to the first end of the pipe.

The system produces a direct image of the scene in the center of the sensor, surrounded by several copies of the images created by the consecutive reflections by each of the mirrors. As a result of the light pipe being rectangular, the sensor is tiled with the different reflected copies organized in rings around the direct image. The copies produced by a single reflection share a common side with the direct image and form the first ring. The copies produced by two reflections share a common side with the first copies forming the second ring and so on.

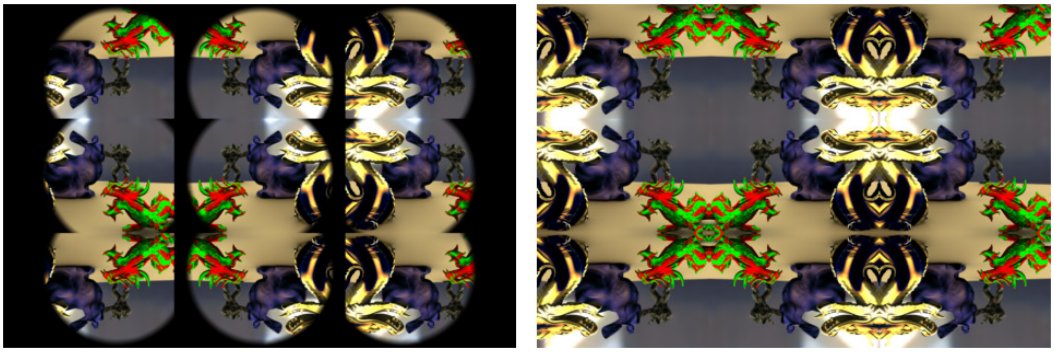
These copies are not perfect copies of the direct image but are different views of the same scene showing a parallax shift between objects at different depths. The different parameters of the system (length and size of the pipe, focal lengths of the lenses) determine the number of rings on the sensor and, consequently, the angular resolution.

One of the differences of this design over the light field cameras previously presented is that the aperture of the lens at the first end of the pipe controls the field of view of the views and that the aperture of the lens at the second end controls the numerical aperture of the views and thus their depth of field.

Manakov *et al.* [2013] and Pacheco and Liang [2014] also show the application of the camera to multi-spectral, polarization or high-dynamic range imaging with the use of an appropriate filter array so that each view is coded differently. More-



(a)



(b)

Figure 2.7: The KaleidoCamera from [Manakov *et al.*, 2013]. (a) Setup for using filters. The image multiplier is a rectangular light pipe. (b) Image on the sensor. Without the pupil-matching lens (left), the image shows strong vignetting effects. These effects are corrected with a pupil-matching lens (right).

over, Pacheco and Liang [Pacheco and Liang, 2014] study a variant of the setup using telecentric optics.

2.2.6 Hybrid Design

A simple way to have both a large spatial and angular resolution is to combine a regular sensor and a light field sensor in the same camera [Baek *et al.*, 2013]. An image is captured through the main objective but it is directed to the two sensors by a beam splitter. The light field sensor is made of a sensor and a microlens array as in the afocal camera. The high-resolution image compensates the loss of spatial resolution at the original focal plane that is limiting the light field sensor.

2.2.7 Miscellaneous Designs

More designs to capture light fields with a single camera exist. We have chosen to group here those that are not yet supported by the model we present in Chapter 4.

(a) (b) (c)

Figure 2.8: Miscellaneous designs. (a) Cathedral glass shuffles the light rays breaking the standard pinhole camera model [Wender *et al.*, 2015]. (b) Droplets act as simili-lenses of various shapes and focal lengths [Iseringhausen *et al.*, 2017]. (c) A programmable aperture using an LCD screen creates specific bokeh [Liang *et al.*, 2008].

Unstructured optics Wender and colleagues [Wender *et al.*, 2015] image with a camera through household optics such as cathedral glass or spoons, see Figure 2.8a. They develop a calibration procedure for almost arbitrary reflective or refractive objects to recover the ray-space mapping. Even though the imaging objects are prone to strong blurring, scattering and irregular sampling of the light field, a layer-based depth estimation scheme allows rendering effects such as refocusing and viewpoint change.

In the work of [Iseringhausen *et al.*, 2017], a regular camera is focused onto water droplets sprayed on a horizontal clear window, see Figure 2.8b. The outline of the droplets are segmented and their shapes are estimated and refined by ray tracing through a physical model considering the volume of the droplets. The resulting light field contains one view per droplet and amounts to one or two hundred views per image.

Coded Masks Attenuating masks can be used to encode information in the 4D Fourier domain of the light field. The heterodyne light field camera [Veeraraghavan *et al.*, 2007] uses a mask with a high frequency sinusoidal pattern in front of the sensor. The mask does not bend rays but simply attenuates the light intensity. The optimal mask pattern and position are derived. Full resolution images are obtained by decoding the Fourier space. Masks can also be placed in the aperture of the main lens [Veeraraghavan *et al.*, 2007]. In this case, the mask acts as a convolution broadband filter and creates a specific blur pattern. The resolution of out of focus parts of the scene is retrieved using appropriate deconvolution allowing refocusing.

Liang and associates [Liang *et al.*, 2008] make a programmable aperture with an

LCD panel also placed in the aperture of the main lens but they use it in a different manner, see Figure 2.8c. It allows to control the angular integration of the camera but requires the exposure of multiple patterns. The patterns are chosen to multiplex the directional information and minimize the total number of required shots. Reddy and colleagues [Reddy *et al.*, 2013] propose a similar design using an array of circular holes in front of the camera. Several pictures are taken with a different hole not being obstructed at a time. Its main advantage is its simplicity. These two systems trade temporal resolution for angular resolution.

2.3 Optical Models for Light Field Cameras

The processing techniques presented in Section 2.1 are independent of the acquisition system used to measure the light field. However, knowing the specifics of the imaging system employed in a light field camera and describing it with a proper model can help to design devices for concrete applications and to improve the calibration, processing and rendering algorithms. As an example, Feng and Shum [2000] and Takahashi *et al.* [2003] explore the depth of field limits in light field rendering. They propose criteria for optimal depth segmentation as well as minimum sampling requirements.

2.3.1 First-Order Evaluation of Optical Properties

The imaging description and analysis of the light field cameras from Section 2.2 is based on first-order optics ray-based models. Some of the optical properties of the systems are discussed in the original publications such as spatial and angular resolution, but these are typically targeted towards the specific design proposed in the paper. Some studies have focussed on describing more comprehensive models that can describe more than one device.

Ray-based model The majority of the works in the literature are focused on both the afocal and focused configurations of microlens-based light field cameras. Their goal is to derive analytical expressions for the optical properties of the light field cameras but apart from [Liang and Ramamoorthi, 2015] that consider a global framework, the two systems are often treated independently of each other.

The expressions of the properties are based on the geometric and optical parameters of the three main components of the system: the main thin lens, the microlens array, and the sensor. Particularly, the studied properties are the following: the depth of field [Damghanian *et al.*, 2014], the depth resolution [Drazic, 2010], and the lateral resolution [Damghanian *et al.*, 2014].

Recently, a first-order ray-based model for the afocal light field camera [Hahne *et al.*, 2014a] details additional properties. It is the first work that introduces the concept of a virtual camera array. It details how to compute the position of the

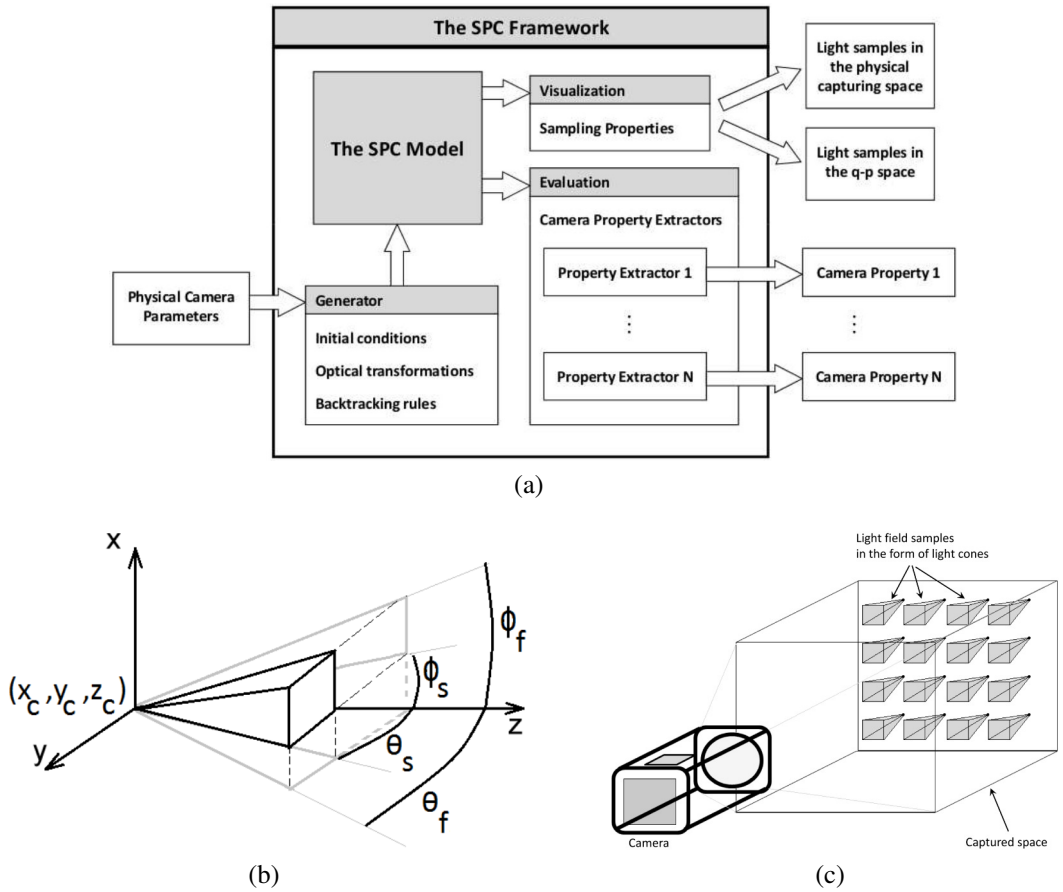


Figure 2.9: SPC model from [Damghanian, 2013]. (a) Details of the SPC framework. (b) Elementary cone of the model. It is defined by 7 parameters: its position (x_c, y_c, z_c) and the limits of its angular extent $(\theta_f, \theta_s, \phi_f, \phi_s)$. (c) The sampling pattern of a camera is decomposed into elementary cones that are imaged outside of the camera forming the SPC.

center of the virtual cameras as well as their baselines [Hahne *et al.*, 2014b]. The model also estimates the absolute position of the refocusing planes [Hahne, 2016]. The model is verified through experiments and simulations.

The sampling pattern cube model An improvement over ray-based models is the sampling pattern cube (SPC) [Damghanian, 2013]. It is a more general framework for plenoptic capturing systems. Its goal is to visualize the sampling characteristics of the system and evaluate its properties, see Figure 2.9.

The model defines a light pyramidal cone as the elementary light field sampler. The light pyramidal cone is characterized by its summit and its angular boundaries (two boundaries for two directions for a cone in 3D space). The sampling pattern cube is the set of light cones with summits at the center of sensor pixels. Elementary operations on a light cone are defined to apply the effects of the optical components

(free-space transport, aperture, lenses) to the light cone parameters. The SPC is transported from the pixels to the object space through the various elements of the optical system.

The properties extracted by the SPC model and validated through experimentation are three variations of the lateral resolution [Damghanian *et al.*, 2012] at a chosen depth plane. The first is defined as the inverse maximum distance between the center point of the base of neighboring light cones. The second is defined as the inverse of the cone base length and the third is linked to the projected image of a pixel in the plane of interest.

Compared to the ray-based model, the SPC model accounts for the size of pixels as well as the reduction of the cone extent due to vignetting by the optics apertures. However, it only derives the lateral resolution of the optical system and misses many other interesting optical properties. Moreover, the model does not give a good comprehension of the relation between the arrangement of optical components in the system and the structure of its corresponding SPC.

This dissertation proposes a more complete first-order optical model in continuation of the ray-based and light cone models, see Chapter 4, that can simulate and evaluate the light field cameras presented in Section 2.2.

2.3.2 Calibration & Computer Vision Model

The calibration task is critical to extracting light field data from the sensor. The literature is again most complete for microlens-based cameras.

The first goal of the calibration is to associate a ray to each pixel. This step requires to compensate for the misalignment between the sensor and the microlens array [Thomason *et al.*, 2014; Cho *et al.*, 2013] and to parameterize and extract the light field data.

The second goal is to find the intrinsic and extrinsic parameters of the camera and potentially its views to get an accurate description of the acquired light field. The actual calibration methods differ between the afocal [Dansereau *et al.*, 2013b] and the focused case [Johannsen *et al.*, 2013; Heinze *et al.*, 2015].

Our common model predicts the location of the perspective cameras given by the calibration. The calibration is still necessary as it accounts for variations of the parameters of the real system as well as the distortions due to aberrations. We experimentally validate the ECA model by recovering the parameters of the main lens of the camera from the fit of the measured properties predicted by our model.

2.3.3 Wave Optics Model

Microlens-based light field cameras have also been analyzed with a first order wave optics model [Turola and Gruppetta, 2014; Turola, 2016]. A wave optics simulation platform was developed to simulate light transport through the system and to study

the effects of diffraction on the performance of the camera. This work proposes and compares several methods to compute the free space propagator and the lens operator under Fresnel approximation. It applies its model to the focused and afocal light field camera.

The scope of this work is is complementary to the model developed in this thesis as diffraction effects are not targeted by the developed model.

2.3.4 Optical Aberrations

Little work has been done on analyzing the effects of aberrations in light field cameras. Ray tracing through an afocal system was introduced by [Ng and Hanrahan, 2007]. It was shown that digital correction of the aberrations can improve the quality of rendered images. Moreover, a small number of directional samples is already sufficient to significantly improve the effective resolution. The effect of irregularities in the microlens array and main lens spherical aberration were studied by [Wei *et al.*, 2015]. They have shown that these variations are beneficial to the sampling of light and image reconstruction at any depth. However, there is not yet an aberration model for light field systems on par with classical optics. This dissertation proposes a first-order model for light field cameras that may serve as a building block for a theory of light field aberrations.

2.4 Light Field Microscopy

Microscopy is a vast subject area and many different illumination and observation schemes have been developed in the past. A general overview is given in [Murphy and Davidson, 2012]; a comprehensive review of microscopy techniques, including light field microscopy, for the neuro-sciences is found in [Wilt *et al.*, 2009].

As seen in Section 1.3, light field microscopy was introduced by [Levoy *et al.*, 2006], see Figure 2.10, and later augmented with light field illumination [Levoy *et al.*, 2009]. Recently, addressing the large spatial resolution loss implicit in light field microscopy, the group has shown that computational super-resolution is achieved outside the focal plane of the microscope [Broxton *et al.*, 2013; Cohen *et al.*, 2014]. Another super-resolution scheme is combining a Shack-Hartmann wavefront sensor and a standard 2D image to compute a high-resolution microscopic light field [Lu *et al.*, 2013]. Light field microscopy has been applied to polarization studies of mineral samples [Oldenbourg, 2008] and initial studies for extracting depth maps from the light field data have been performed in microscopic contexts [Lee *et al.*, 2012; Thomas *et al.*, 2013].

Most of the work today uses the same optical configuration that was introduced in the original implementation [Levoy *et al.*, 2006]. An exception is [Lin *et al.*, 2015] where a 5-by-5 camera array is linked to the microscope objective with a two-lens relay system. An implementation of the light field microscope in the focused

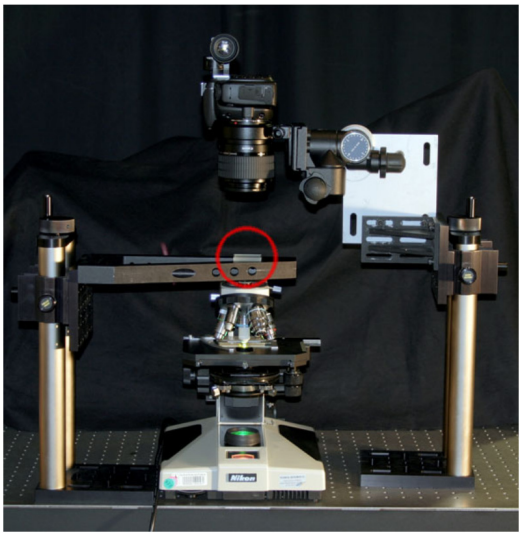


Figure 2.10: Light field microscope from [Levoy *et al.*, 2006].

configuration was studied with the wave optics model of [Turola, 2016].

This dissertation contributes to light field microscopy for two reasons, see Chapter 6. The first is an investigation of the optical matching between a commercial light field camera and a microscope to achieve an inexpensive light field microscope. The second is the realization of a setup similar to the one of [Levoy *et al.*, 2006] and its characterization with the ECA model developed in this thesis to provide a better insight of its design and to validate the model.

Part II

Theoretical Model

Chapter 3

Point Imaging Matrices

In this chapter, we introduce alternate tools to the traditional first-order ray tracing methods that use ABCD matrices to propagate light rays in optical systems [O'Shea, 1985]. However, ABCD matrices are not suitable for the decentered systems that light field cameras are. So, instead of a matrix system that outputs the deflected ray corresponding to an input ray, we propose a point matrix system that maps input points to output points. We show the different elementary matrices corresponding to basic optical components and the way to combine them. Then, we demonstrate their use through several applications. Finally, we discuss the advantages and disadvantages of this new tool and compare it to the regular ray matrices. The new matrix tools are later used as a foundation for the equivalent camera array model in Chapter 4 and are implemented in a simulation software in Chapter 7.

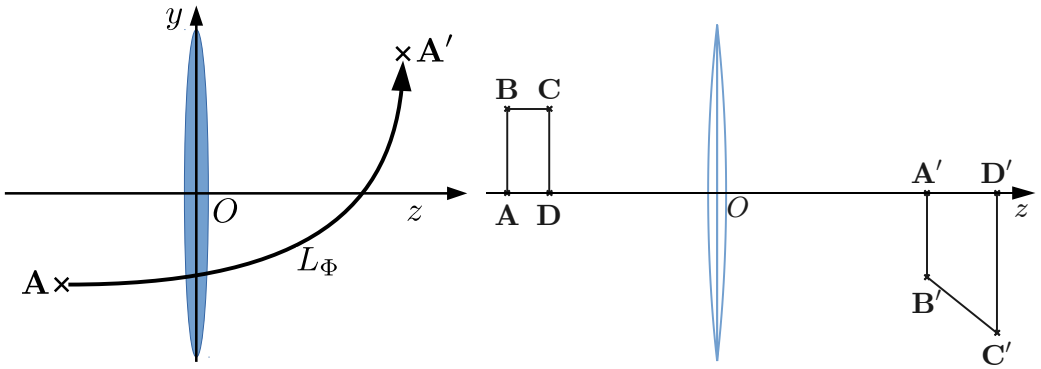
3.1 Definition and Operations

In the paraxial approximation of optical imaging, the optical laws are linear, see Section 1.2. As such, the effect of an optical system is described by a general linear transformation and an object point $\mathbf{A}$ and its image point $\mathbf{A}'$ form a conjugated pair of points. In linear algebra, this projective transformation is represented as a matrix L and the conjugated points are related by:

$$\widehat{\mathbf{A}}' = L\widehat{\mathbf{A}}\,. \tag{3.1}$$

The matrix L is actually a homogeneous matrix and the points $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{A}}'$ are the respective counterparts of the real points $\mathbf{A}$ and $\mathbf{A}'$ in homogeneous coordinates. Passing from the real coordinates to the corresponding homogeneous coordinates and inversely is done as follows:

$$\mathbf{A} = \begin{pmatrix} x_{\mathbf{A}} \\ y_{\mathbf{A}} \\ z_{\mathbf{A}} \end{pmatrix} \rightarrow \widehat{\mathbf{A}} = \begin{pmatrix} x_{\mathbf{A}} \\ y_{\mathbf{A}} \\ z_{\mathbf{A}} \\ 1 \end{pmatrix}, \quad \widehat{\mathbf{A}} = \begin{pmatrix} \widehat{x}_{\mathbf{A}} \\ \widehat{y}_{\mathbf{A}} \\ \widehat{z}_{\mathbf{A}} \\ \widehat{w}_{\mathbf{A}} \end{pmatrix} \rightarrow \mathbf{A} = \begin{pmatrix} \frac{\widehat{x}_{\mathbf{A}}}{\widehat{w}_{\mathbf{A}}} \\ \frac{\widehat{y}_{\mathbf{A}}}{\widehat{w}_{\mathbf{A}}} \\ \frac{\widehat{z}_{\mathbf{A}}}{\widehat{w}_{\mathbf{A}}} \end{pmatrix}. \tag{3.2}$$



(a) Single point mapping (b) Imaging of a scene

Figure 3.1: Imaging with a matrix through a thin lens.

3.1.1 Thin Lens

The simplest optical system is the thin lens, see Figure 3.1. By expressing the coordinates of the image point as a function of the object point coordinates from Equations 1.10 and 1.11, the point imaging matrix L_Φ of a thin lens with optical axis z and focal distance f, centered at the origin O is:

$$L_\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \Phi & 1 \end{pmatrix} \tag{3.3}$$

where $\Phi = \frac{1}{f}$ is the power of the thin lens. This matrix is not the matrix directly obtained from the equations but after division by the focal length f. The two matrices are equivalent because homogeneous coordinates are not modified by a common multiplying factor. The advantage of the matrix L_Φ is that its determinant is 1, otherwise, it would be f^4.

The second advantage lies in the inverse matrix. The matrix L_Φ corresponds to forward imaging by the lens where object points from the object space are imaged to image points in the image space. The reverse image by the same lens from the image space to the object space is then represented by the matrix inverse L_Φ^{-1}. It is easy to show that $L_\Phi^{-1} = L_{-\Phi}$ or, stated differently, reverse imaging corresponds to forward imaging with a lens of inverse focal length.

3.1.2 Thick Lens

When a lens is too thick, it cannot be considered as a thin lens anymore. Instead of having the refraction occurring on the same lens plane, it is split into two principal planes. The distance e between these two planes accounts for the effect of the thickness of the lens. The front and back principal planes are perpendicular to the

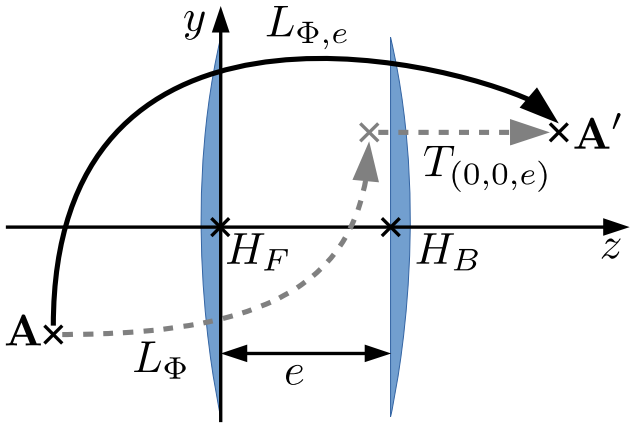


Figure 3.2: Thick lens model.

optical axis and their intersections with the optical axis are, respectively, the front and back principal points, H_F and H_B. Considering a thick lens with its front principal plane at the origin, the image space is simply translated along the optical axis by e, see Figure 3.2.

The homogeneous matrix for a translation by a vector $\mathbf{u} = (\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z)$ is:

$$T_{\mathbf{u}} = \begin{pmatrix} 1 & 0 & 0 & \mathbf{u}_x \\ 0 & 1 & 0 & \mathbf{u}_y \\ 0 & 0 & 1 & \mathbf{u}_z \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{3.4}$$

Its determinant is 1 and its inverse is: $T_{\mathbf{u}}^{-1} = T_{-\mathbf{u}}$, the translation by the opposite vector.

Consequently, the imaging matrix of a thick lens of power Φ and thickness e is:

$$L_{\Phi,e} = T_{(0,0,e)} L_\Phi . \tag{3.5}$$

The matrix of the reverse imaging is: $L_{\Phi,e}^{-1} = L_{-\Phi} T_{(0,0,-e)}$.

3.1.3 Decentered Optics

So far, the lens position is still restricted to the origin of the coordinate system with the optical axis being the z-axis. We want to model imaging through decentered optics, therefore, we show the effect of a rigid body transformation such as a translation or a rotation represented by a matrix P on a lens operation. This matrix P represents the transformation from the original coordinate frame to the coordinate frame at the optical center (or front principal point) of the lens with its z-axis aligned with the optical axis of the lens. Then the imaging process is decomposed into three steps, see Figure 3.3. First, the transformation P is reverted moving the lens to the original coordinate frame. Second, the lens projection is applied. Third,

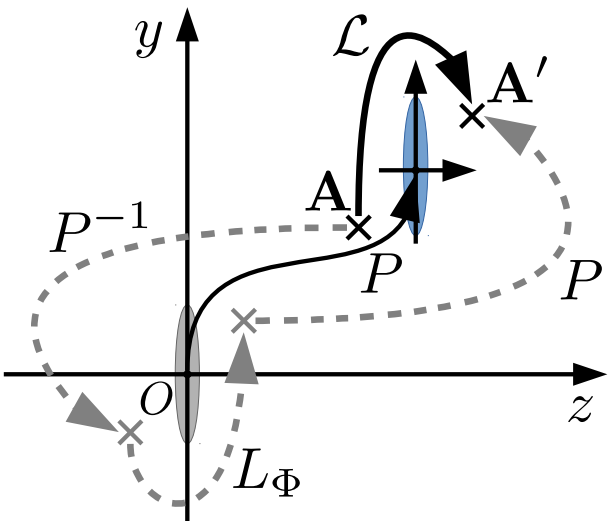


Figure 3.3: Decentering operation. The lens at the origin is virtual.

the operation of the first step needs to be canceled by applying the transformation P. The final imaging matrix $\mathcal{L}$ of these operations is:

$$\mathcal{L} = PLP^{-1}\,. \tag{3.6}$$

The matrix L can be any imaging matrix and the transformation P can be any combination of translations and rotations corresponding to the component decentering.

Reverse imaging through the same lens with the same transformation matrix gives: $\mathcal{L}^{-1} = (PLP^{-1})^{-1} = PL^{-1}P^{-1}$. Only the optical part is affected and not the transformation part.

3.1.4 Sequence Imaging

In an optical system consisting of k optical components in a sequence, the final imaging matrix is obtained by multiplying the imaging matrix of each optical component with its own decentering as follows:

$$\mathcal{L} = \left(P_k\mathcal{L}_kP_k^{-1}\right)\cdots\left(P_1\mathcal{L}_1P_1^{-1}\right)\,. \tag{3.7}$$

The determinant of the product of matrices is the product of the determinants of the individual matrices and the determinant of the inverse of a matrix is the inverse of the determinant of this matrix. Therefore, the determinant of the transformation matrices cancels so that $\det(\mathcal{L}) = \det(\mathcal{L}_1)\cdots\det(\mathcal{L}_k)$. The determinant of $\mathcal{L}$ is equal to the ratio of the image and object medium indices. For an image and object in the same medium, the optical index is the same so the determinant is equal to 1.

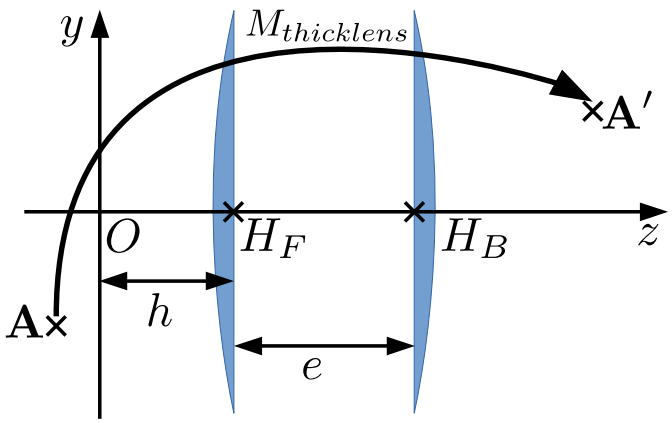


Figure 3.4: Notations for the equivalent thick lens.

3.1.5 Imaging at/to Infinity

In optics, the object and the image points can be either at a finite distance or at infinity. For instance, an afocal optical system transforms an object point at infinity to an image point at infinity. The infinity case is often an intermediate state in an optical system where the light beam is neither converging nor diverging.

Even though our matrix tools are not mapping rays, they still provide the correct information for points at infinity through the homogeneous coordinates. When placing a point at the front focal plane of a thin lens for instance, its image is at infinity. In this case, the homogeneous coordinate of the image point is zero, see Equation 3.3. The coordinates of the corresponding actual image point do not exist as they would be divided by zero but the coordinates of the homogeneous image point store the interesting data. The slope in x and y of any ray of the collimated beam is retrieved by dividing by the z coordinate and the angular direction is obtained by taking the arc tangent of the slope.

An object point at infinity is set directly in homogeneous coordinates with the tangent of its angular direction for the x and y components, 1 for the z component and 0 for the w component. If the vector is normalized, the ray direction cosines are obtained but generally this is not necessary as homogeneous vectors of different scales are describing the same quantity.

3.2 Applications to Common Cases

In the following section, we demonstrate the use of the previously defined matrix tools to common cases from the literature such as the combination of two thin lenses or a thick lens defined by two spherical surfaces. We also verify the thick lens equivalence and derive its parameters in the cited cases. We assume that the input and output media are made of the same optical index.

A thick lens is parameterized by its optical power Φ, its thickness e and the relative position along the optical axis h to a certain frame of reference, see Figure 3.4. This relative position h is necessary as the reference frame of an optical system is different from the one of its equivalent thick lens. The general matrix for this lens is

$$M_{thicklens} = T_{(0,0,h)} T_{(0,0,e)} L_{\Phi} T_{(0,0,h)}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 + (e+h)\Phi & e - h(e+h)\Phi \\ 0 & 0 & \Phi & 1 - h\Phi \end{pmatrix}. \tag{3.8}$$

The determinant of this matrix is equal to 1.

3.2.1 Association of Two Thin Lenses

Let us consider an optical system made of two aligned thin lenses with optical powers Φ_1 and Φ_2, separated by a distance D. The origin of the frame of reference is the center of the first lens and the optical axis is the axis passing through the center of both lenses. The matrix of this system is:

$$\begin{aligned} M_{2thinlenses} &= \left(T_{(0,0,D)} L_{\Phi_2} T_{(0,0,D)}^{-1} \right) L_{\Phi_1} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \Phi_1(D - D(D\Phi_2 + 1)) + D\Phi_2 + 1 & D - D(D\Phi_2 + 1) \\ 0 & 0 & \Phi_2 - \Phi_1(D\Phi_2 - 1) & 1 - D\Phi_2 \end{pmatrix}. \end{aligned} \tag{3.9}$$

Its determinant is the product of each thin lens matrix and is again equal to 1.

By identifying the elements of the matrices $M_{2thinlenses}$ and $M_{thicklens}$, four equations are produced, one being redundant. This equation system is solved to retrieve the unknown parameters of the equivalent thick lens:

$$\begin{cases} \Phi = \Phi_1 + \Phi_2 - D\Phi_1\Phi_2 \\ e = -\frac{D^2\Phi_1\Phi_2}{\Phi} \\ h = \frac{D\Phi_2}{\Phi} \end{cases} \tag{3.10}$$

The formula for the power Φ is also known as the Gullstrand formula [Singer *et al.*, 2006].

3.2.2 Thick Lens Defined by Two Surfaces

Before studying the case of a thick lens delimited by two spherical surfaces, we first need to derive the imaging matrix of a spherical interface between two media of different optical indices.

Interface between Two Media

Let us consider a spherical interface of curvature radius R which delimits two media of optical indices n_1 and n_2. The optical axis of this optical component passes through the center of the sphere C. It intersects the surface at a point S so we have $R = \overline{SC}$. We use the direction of the optical axis to define the algebraic value of R so that it is positive in the direction of the axis and negative in the opposite direction. For a reflective surface, the index of the second medium is the negative of the first medium. For a flat surface, R is infinite.

The equivalent of the thin lens equation for this configuration with the origin at S and with the z-axis as the optical axis is the following:

$$\frac{n_2}{z_{\mathbf{A}'}} - \frac{n_1}{z_{\mathbf{A}}} = \frac{n_2 - n_1}{R}\,. \tag{3.11}$$

In this reference frame, the magnification is defined as:

$$M_{\mathbf{A}\mathbf{A}'} = \frac{n_1}{n_2}\frac{z_{\mathbf{A}'}}{z_{\mathbf{A}}} = \frac{x_{\mathbf{A}'}}{x_{\mathbf{A}}} = \frac{y_{\mathbf{A}'}}{y_{\mathbf{A}}}\,. \tag{3.12}$$

So the imaging matrix L_{R,n_1,n_2} for the interface is:

$$L_{R,n_1,n_2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{n_2}{n_1} & 0 \\ 0 & 0 & \frac{n_2-n_1}{n_1 R} & 1 \end{pmatrix}. \tag{3.13}$$

As for the thin lens matrix, this matrix is not the one directly derived from the equations but the one obtained after dividing by $n_1 R$. The determinant of the matrix is the ratio of the medium indices: $\frac{n_2}{n_1}$. The matrix of reverse imaging is: $L^{-1}_{R,n_1,n_2} = L_{-R,n_2,n_1}$.

General Thick Lens

Let us consider the general case of a thick lens of optical index n_2 placed in a medium of optical index n_1. The lens is made of two spherical surfaces of radius R_1 and R_2 separated by the distance D, see Figure 3.5. The optical axis is the line passing through the center of both surfaces, C_1 and C_2. The origin of the frame of reference is the intersection of the first surface with the optical axis S_1. The imaging matrix of this optical system is:

$$M_{2surf} = \left(T_{(0,0,D)} L_{R_2,n_2,n_1} T^{-1}_{(0,0,D)}\right) L_{R_1,n_1,n_2}\,. \tag{3.14}$$

Its determinant is $\det(M_{2surf}) = \det(L_{R_2,n_2,n_1}) \det(L_{R_1,n_1,n_2}) = \frac{n_1}{n_2}\frac{n_2}{n_1} = 1$.

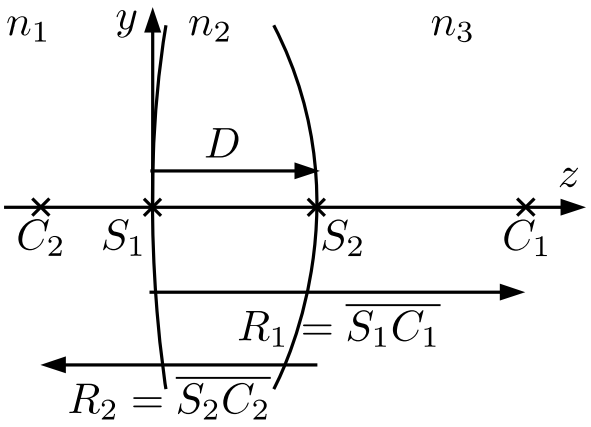


Figure 3.5: Thick lens constituted of two spherical surfaces at the transition between two media. For a lens in the same medium, n_3 is equal to n_1.

As for the case of the two thin lenses, we set the matrices M_{2surf} and $M_{thicklens}$ equal to obtain a system of equations. Then, we solve for the unknown parameters of the thick lens with the parameters of the surfaces and media.

$$\begin{cases} \Phi = (n_2 - n_1)\left(\frac{D(n_2-n_1)}{n_1 n_2 R_1 R_2} + \frac{1}{n_1}\left(\frac{1}{R_1} - \frac{1}{R_2}\right)\right) \\ e = \frac{D(n_2-n_1)(D+R_2-R_1)}{D(n_2-n_1)+n_2(R_2-R_1)} \\ h = \frac{-DR_1n_1}{D(n_2-n_1)+n_2(R_2-R_1)} \end{cases} \tag{3.15}$$

For a lens in air, the index n_1 is 1 and the solutions can be simplified further. The solution for the power Φ is the lensmaker's formula for $D = 0$. The more general formula can be found in [O'Shea, 1985].

An even more general case would consider having different input and output optical indices. In order to get the parameters of the equivalent thick lens, the matrix $M_{thicklens}$ needs to incorporate this change of indices as well as the matrix of the second spherical surface but the principle stays the same. The determinant of M_{2surf} is then equal to the ratio of the output index to the input index.

3.2.3 Association of Multiple Components

In a more general application, we derive the parameters of the equivalent thick lens for a system composed of the association of centered and not rotated multiple optical components in the same medium, sharing the same optical axis. The matrix of such a system is of the following form:

$$\mathcal{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{pmatrix}. \tag{3.16}$$

The determinant of $\mathcal{L}$ is $det(\mathcal{L}) = a_{34}a_{43} - a_{44}a_{33}$ and is either equal to 1 or -1 depending on the number of reflective surfaces, respectively even or odd. As a result of the determinant of $M_{thicklens}$ being equal to one, we consider only an even number of mirrors. The system of equation obtained by identifying the elements of $\mathcal{L}$ and $M_{thicklens}$ is:

$$\begin{cases} a_{43} = \Phi \\ a_{44} = 1 - h\Phi \\ a_{33} = 1 + (e+h)\Phi \\ a_{34} = e + h - h\left(1 + (e+h)\Phi\right) \end{cases} \tag{3.17}$$

The equation obtained from the determinant: $a_{34}a_{43} - a_{44}a_{33} = 1$ makes the fourth equation of the system redundant. Consequently, the parameters of the thick lens are:

$$\begin{cases} \Phi = a_{43} \\ h = \frac{1-a_{44}}{a_{43}} \\ e = \frac{(a_{44}+a_{33}-2)}{a_{43}} \end{cases} \tag{3.18}$$

The system has solutions under the condition that the optical power Φ is not equal to zero which means that the optical component is a lens with no refractive power (if the determinant is -1, it is a flat mirror).

More generally, for an optical system with a change of the optical indices of the input and output media including an odd number of reflective surfaces, the same procedure can be applied to retrieve the optical parameters of the equivalent simplified system. The condition on the determinant of $\mathcal{L}$ is to be equal to the ratio of the output and input medium optical indices.

3.3 Analysis

The traditional ray matrices have been studied and used as basic tools in the ray tracing of paraxial optical systems because they offer many advantages. The optical quantities of the system can be derived from the ABCD terms of the corresponding matrix. In addition, they also profit from some properties of matrix calculus and the matrices of many simple optical components are well known. In order to deal with optical components in decentered systems, the 2-by-2 matrix model is usually extended to a 3-by-3 one for 2D systems and a 5-by-5 matrix representation in three dimensions [Singer *et al.*, 2006]..

Our model is similar in its operating principles to the ray matrix model. It uses 4-by-4 matrices for imaging through a decentered 3D system instead of the 5-by-5 matrix. It benefits from the same power of matrix calculus. The optical quantities can also be derived from the matrix terms as we have seen in the previous section. Even though only a small number of matrices of optical components were presented,

there exist the corresponding point matrices to each optical component that has a ray matrix. This exhaustive work is not the main goal of this dissertation.

The main difference with the ray matrix model is its ability to obtain directly the image of a point without the additional cost of ray intersection which is a computational advantage that particularly fits our model of light field cameras, see Chapter 4. Nonetheless, a ray is easily obtained from the homogeneous coordinates of two points defining the ray. If the homogeneous coordinate of the points is non-zero, the points exist at a finite position and their subtraction gives the direction vector of the ray. If the homogeneous coordinate of one point is null, this point is at infinity and it already defines the direction of the ray. In the case where both points are at infinity, the corresponding ray is undefined. This configuration may happen when taking the two points in the front focal plane of a lens, the images of both points are then at infinity. The physical interpretation is that the ray cannot strike the lens as it is parallel to the lens plane. Generally, the resulting ray or point corresponding to the same inputs of the system parameters are the same whichever model is used.

3.4 Conclusion

We have demonstrated the use of an alternative tool to first-order ray matrices for Gaussian imaging also based on matrices. This tool considers the linear transformation of the object space into the image space. At a high level, it directly exploits the first-order approximation that implies that the images of object points are perfect points. We show that the same optical knowledge can be retrieved with this tool.

We have studied several basic cases with this model: using thin lenses as well as ideal and real thick lenses. We also re-derived well known formulas of the parameters of the thick lens equivalence of these systems.

An important result we have not presented is the Scheimpflug principle where the plane of the lens and the sensor are not parallel to each other but tilted. This principle states that the image plane and the object plane then intersect at the lens plane. This statement is respected with our model by taking the image of a point on the lens plane but we have not yet derived the relation between the angle of the object plane and the image plane. The Scheimpflug principle is an important aspect of imaging in the case of a light field setup using an array of planar mirrors oriented in different directions for instance.

Chapter 4

The Equivalent Camera Array

A light field camera spatially multiplexes directional light information onto its sensor, thereby realizing a four-dimensional light field sensor with a two-dimensional pixel array. In order to do so, it uses a component that directs the light rays that are incident at a particular spatial position from different directions to different pixels. We call this component the *directional multiplexing unit (DMU)*. The DMU is often implemented by a lens array. However, multiple designs exist, see Section 2.2, so, we prefer the more general terminology.

Our goal in the current chapter is to abstract a real light field camera by a virtual equivalent camera array (ECA) that is observing the object space. This abstraction is possible for most existing light field camera designs. We first define the elementary building stone of the light field sampling by the optical system and use it to decompose the light field device into a set of these elements that are then imaged to the object space and grouped together to form the equivalent camera array. We then describe the properties of the virtual cameras as introduced in Section 1.2, as well as additional properties that are derived from coupled information between different virtual cameras. These properties are interpreted later in Chapter 5 to analyze two light field camera designs and configurations. In the last part, we introduce a vignetting function to determine the influence of the vignetting by the system apertures on the structure of the ECA.

4.1 Principle

Individually, each element of the DMU and the pixels associated with it act as a small camera looking at the world. In a light field camera, a main lens is often added in front of this array of small cameras. From the point of view of a small camera, instead of looking directly at the object space, it is looking at the in-lens space. The in-lens space is a perspective mapping of the object space following the equations of Section 1.2. In order to retrieve the properties of an equivalent camera array, we would like to transform this in-camera array of small cameras into the object space, where it could be treated much like a normal camera array consisting

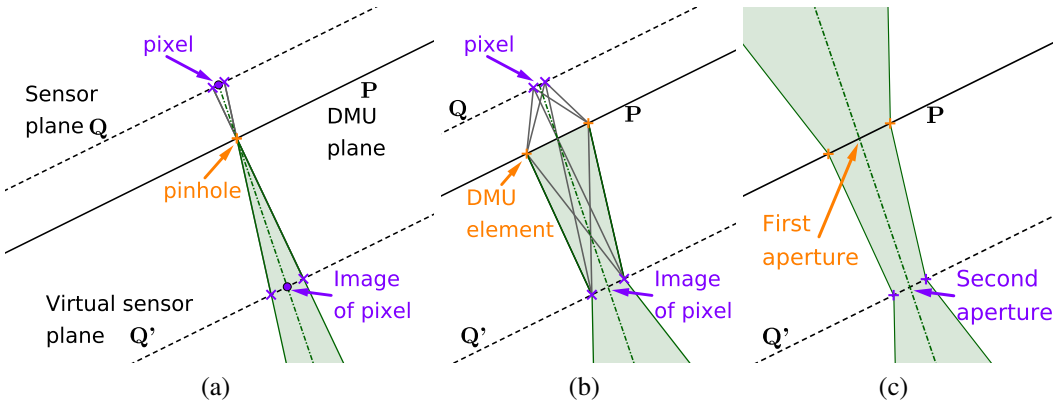


Figure 4.1: Imaging kernel. (a) Pinhole model. (b) Pixel+DMU element model. Every ray contained in the green area will be integrated by the pixel. (c) Two-aperture model.

of physical cameras. However, mapping a camera through the main lens has no obvious physical solution due to the distortion of space affected by the lens. To arrive at a solution, we need to consider an abstraction that describes the *effect* of the combined small camera-main lens system. In the process, we will lose some physical properties of the system, in particular, its image side properties will only be defined up to a one-parameter family of solutions.

4.1.1 Two-Aperture Model

We resort to the two-aperture model introduced in the pioneering work [Levoy and Hanrahan, 1996] for the analysis of light field sampling properties for real camera arrays. In light field cameras, every pixel is assigned to the main ray passing through the center of that pixel and the center of an attributed element of the directional multiplexing unit.

Let us consider a two-plane parameterization of the light field in the case of a small camera array with the sensor plane and the DMU plane as the support planes. We label them plane Q and plane P, respectively. We consider at first that the DMU element is implemented by a pinhole as illustrated in Figure 4.1a.

Any light ray hitting a certain position on the sensor also passes through the image of this particular position in the object space of the pinhole. However, a pinhole is a focus free imaging element so there is an infinite number of image planes Q' for the sensor plane Q. In terms of parameterization of the light field, the hit position of the light ray with the planes Q and Q' is a relative distance to the optical axis, so these planes are equivalent. The pair of planes (Q, P) of the parameterization can be replaced by the pair (Q', P).

This replacement allows to abstract the effect of the DMU element but still conserves the relation between the parameterization and the sampling of the scene that is implemented by the camera. This model is often used for the calibration of light field cameras [Dansereau *et al.*, 2013b] but it neglects the physical focusing aspect of light field cameras.

In a more realistic system, as shown in Figure 4.1b, the DMU element is now a first-order optical element and as such it has focusing properties. A finite-sized pixel on the sensor integrates all light that is passing through the surface of the associated DMU element and that is hitting its finite surface.

Observe that this pencil of light is also passing through the complete surface of the image of the pixel outside the DMU/ small camera, *i.e.,* the unique plane Q' which is the optically conjugate plane of Q, and intersects the same surface area on the DMU element.

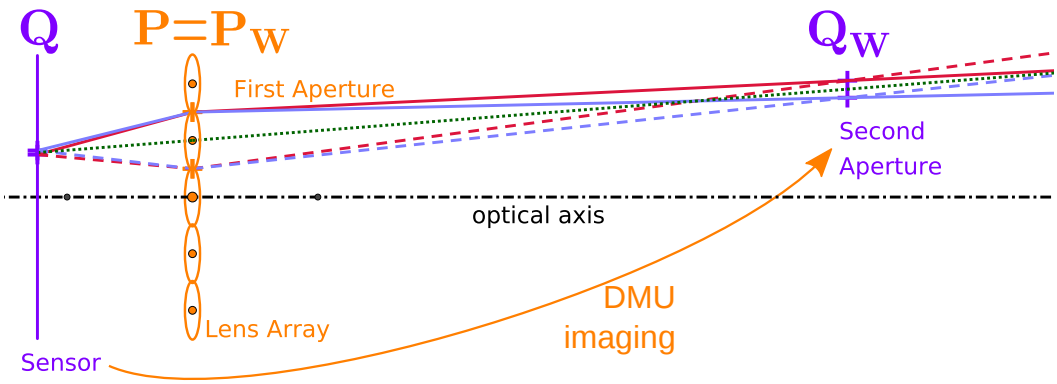
Therefore, it suffices to know the positions and the surfaces of the pixel image and the DMU element to predict the light rays that are integrated by the corresponding sensor pixel. Note that these two positions are now located in the object space of the small camera, *i.e.,* that the optical effect of the DMU can now be ignored.

As illustrated in Figure 4.1c, we abstract a pixel/DMU element combination by a two-aperture system with the positions and extents as described above. This system is consistently imaged through an optical component and it preserves the information on the focusing properties of the light field subviews. As mentioned previously, the disadvantage of this abstraction is that the apertures loose their physical properties. We discuss the consequence of this loss on the ECA in Section 4.4.

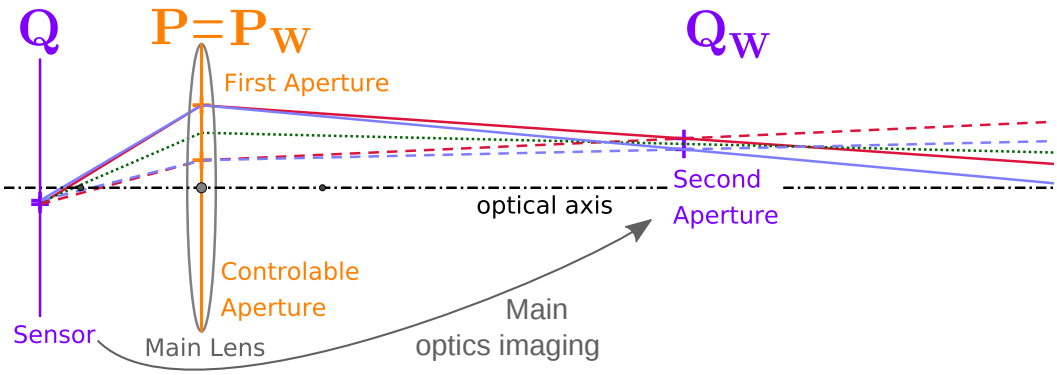
4.1.2 Building the ECA

This procedure was applied to several setups from the literature in Figure 4.2 and Figure 4.3 showing a variety of light field camera designs that can be analyzed with our model. Note that the main lens may be missing and that the directional multiplexing unit may have additional relay optics. In order to be more specific, we use the relatively complex KaleidoCamera design [Manakov *et al.*, 2013] as an illustrative example. The system is made of a main lens and a sensor with an in-between directional multiplexing unit that consists of two lenses, a field lens and a pickup lens, that are at the entrance and the exit of a mirroring light pipe. In Figure 4.3a, the pixel area of the sensor serves as the pixel aperture, whereas the pickup lens generates the aperture of the DMU. The light pipe generates a virtual DMU lens array through mirroring. The field lens is a relay system that images the plane of the DMU into the exit pupil of the main lens.

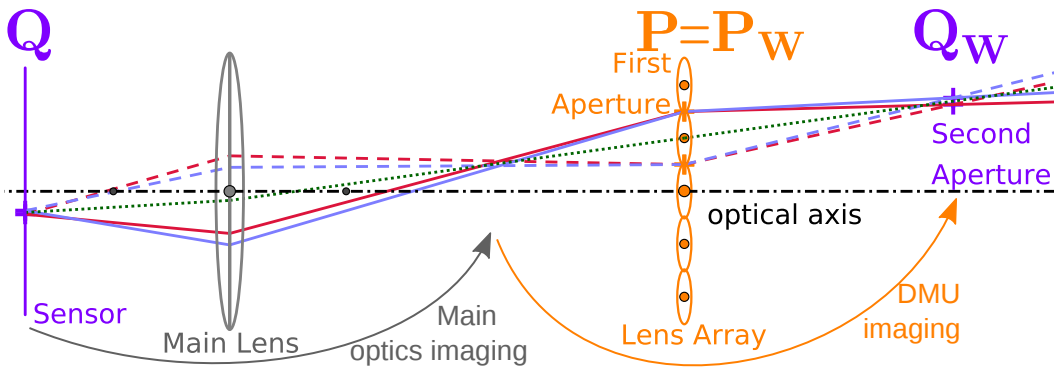
Rays starting from the pixel aperture pass through the pickup lens and are imaged and reflected through the different system components. Finally, they pass through the images of the pixel and pickup lens aperture in the object space.



(a) The monolithic camera array from [Venkataraman *et al.*, 2013].

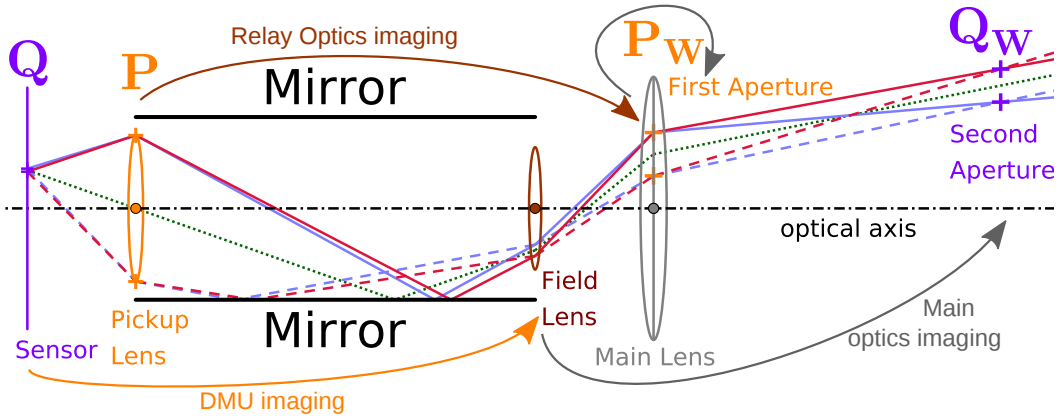


(b) The programmable aperture from [Liang *et al.*, 2008].

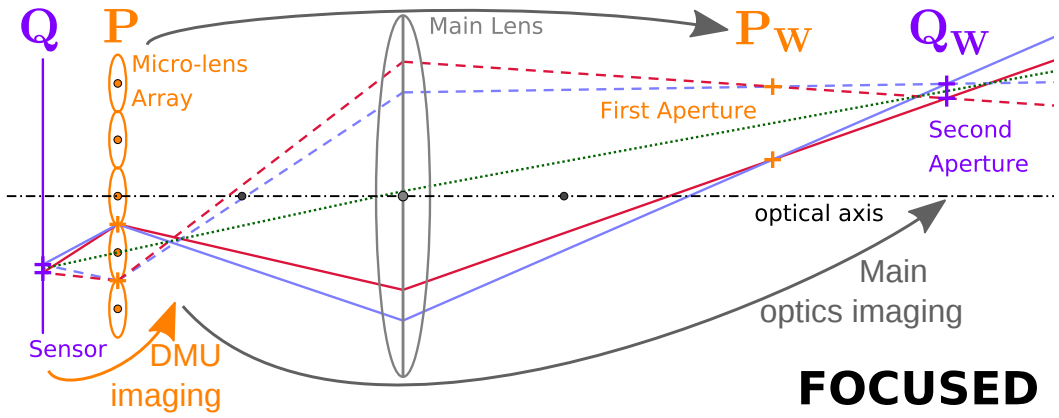


(c) The external lens array from [Georgiev and Intwala, 2003].

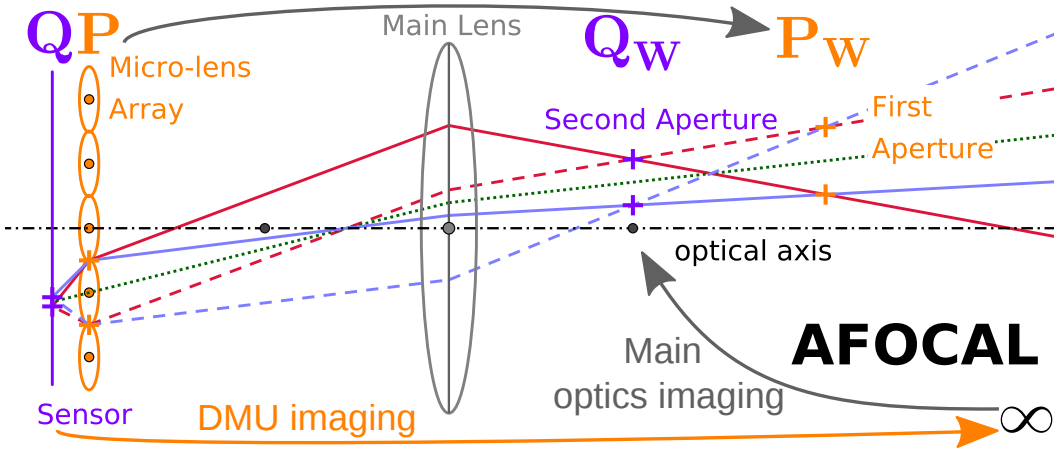
Figure 4.2: ECA applied to light field camera designs from the literature (first part).



(a) The KaleidoCamera from [Manakov *et al.*, 2013].



(b) The focused light field camera from [Lumsdaine and Georgiev, 2009].



(c) The afocal light field camera from [Ng *et al.*, 2005a].

Figure 4.3: ECA applied to light field camera designs from the literature (second part).

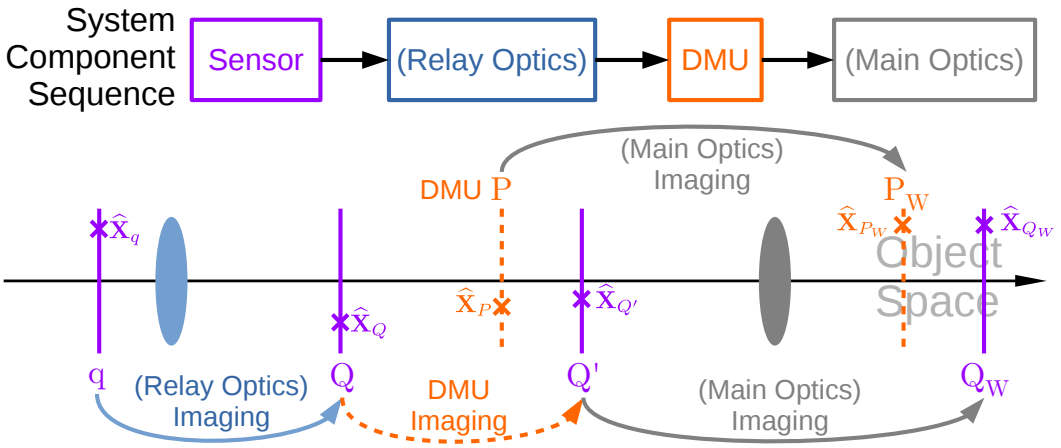


Figure 4.4: General model for the mapping of the sensor and the directional multiplexing unit to the object space. The main optics and the optics introduced between the sensor and the DMU planes, such as a relay system, for example, are optional. In this case, the imaging matrix L of the optics is replaced by the identity matrix.

The general procedure is illustrated in Figure 4.4 and is detailed as follows. Every combination of pixel/DMU elements is decomposed into two-aperture elements. They are imaged through the sequence of the optical components of the light field camera, DMU element included, to the object space of the camera. From the equations of Chapter 3, a pixel image is given by: $\widehat{\mathbf{X}}_{Q_W} = \mathcal{L}_{Main}\mathcal{L}_{DMU}\mathcal{L}_{Relay}\widehat{\mathbf{X}}_q$, and a DMU image is given by: $\widehat{\mathbf{X}}_{P_W} = \mathcal{L}_{Main}\widehat{\mathbf{X}}_P$. The two-aperture elements that share the same aperture imaged from the same DMU element are selected to form a virtual camera. There is one virtual camera per DMU element resulting in a virtual camera array alias the ECA that is equivalent to the in-camera array in the sense that it integrates the same ray bundle as the physical light field camera.

We investigate more thoroughly the microlens-based light field camera designs illustrated in Figure 4.3c and Figure 4.3b as well as the properties of their ECA in Chapter 5.

4.2 Properties

The abstraction of the pixels and DMU elements as apertures is sufficient to define similar properties as those described in Section 1.2 for a standard camera as most of them characterize the object side of the camera. Given an equivalent camera array, we can compute the view direction, the field of view, the depth of field and the resolution for each of the individual cameras. As a result of the ECA being made of several cameras, additional information of two or more cameras are used to derive new properties of the system. The following properties are only valid for the sharp region of the scene space. This region is delimited by the limits of the depth of field

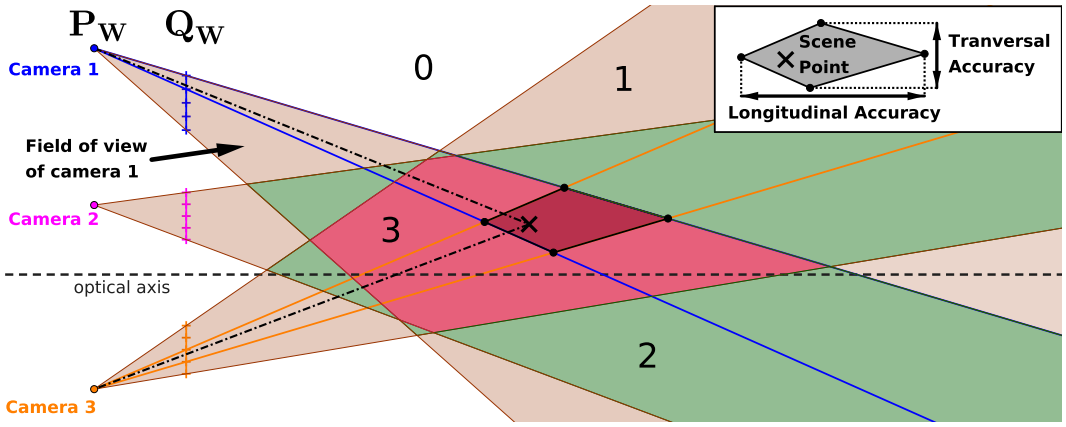


Figure 4.5: Definition of the properties for the field. The whole field is separated into different regions depending on the number of cameras observing it. Any point in the same diamond shaped region would project onto the same pixels as the defined scene point. The largest extent of this region along and orthogonal to the optical axis of the two cameras define the transversal and longitudinal accuracy at the position of a point in this region, respectively. Every pair of pixels, one per camera, defines a unique region of space.

of the cameras of the ECA. Objects outside of this region are out of focus so they appear blurry.

Disparity In stereo vision, the disparity is the difference in pixels of the position of the image of the same feature on two different cameras. The disparity is zero at the plane where the view directions intersect. For a scene point behind the no-parallax plane, the disparity is positive and it is negative when the point is closer than this plane.

Baseline The baseline is the distance between the centers of projection of two cameras looking at the scene. It is an indicator of the ability of the system to measure the disparity of a point in the scene. In the case of a camera array, multiple cameras can see the same point so the interesting value is the maximum baseline of all pairs of cameras. A baseline map of the whole object space is computed by intersecting the fields of view of every possible pair of virtual cameras as shown in Figure 4.5. It is possible that the baseline is null for scene points that are seen by only one virtual camera or undefined for scene points outside of the field of view of every camera of the array.

Accuracy A point in space can be projected onto the virtual sensor of a camera (plane Q_W). This projected point is defined as the intersection between the line formed by the point and the center of the camera, and the plane of the virtual sensor.

When a point in object space is projected onto the virtual sensor of two cameras, it falls onto one pixel in each camera. Similar to the field of view of a camera, the field of view of a certain pixel is the cone defined by the camera center and the edges of the pixel. As can be seen in Figure 4.5, the intersection of the field of view of two pixels from two cameras results in a region in space. The points belonging to this region cannot be differentiated by only utilizing the two cameras. This region is reduced as more pixels from different cameras image a scene point. We define the transversal and longitudinal accuracies as the largest lateral and depth extents of this region, respectively.

4.3 Vignetting

At the moment, the ECA model does not include vignetting effects even though it considers apertures to evaluate first-order properties. The imaging of a two-aperture element (pair of a pixel and a DMU aperture) is done without considering whether the light passing through it is blocked by another component of the system. However, vignetting is especially important in a light field camera because:

- it affects the pixel/DMU pairing,
- it modifies the properties of the ECA,
- it gives insight on some design rules,
- and it helps with the calibration of the system.

In this section, we address these topics and we present a method to evaluate the fraction of etendue of each pixel/DMU pair in our model. The main questions we answer are: considering a light field camera, from which DMU does a pixel receive light ? Is it none, one or many ? And what is its approximate irradiance ?

4.3.1 Pixel/DMU Pairing

The pairing between the pixels and the DMU elements is crucial because it has to obey a few rules in order to build a consistent camera array. The main rule is that there should be no more than one DMU element paired up to each pixel. In microlens-based light field cameras, this condition is known as the f-number matching rule. In order to prevent the overlapping or the gaps between the images made by two neighboring DMU elements (the microlenses), the effective f-number of the DMU elements and that of the main lens should match. In the KaleidoCamera, for instance, it is the field lens aperture projected onto the sensor through the DMU element center after zero, one or many reflections, that is used to select the pairings appropriately. In this case, the main lens also plays a role as it limits the extent of the ray bundle entering the camera, thus limiting the number of possible reflections.

A simple but naive approach to the problem of selecting which pixels are imaged through which DMU to form a virtual camera is to determine for each pixel which DMU is the closest. The criterion of evaluation is the distance between the center of the pixel aperture and the center of the DMU. This method ensures that no pixel is imaged by two DMUs at the same time. This is justified when the studied systems already validate the "f-number matching" rule but it is not satisfactory as it is limited to microlens-based light field cameras and does not take vignetting into account.

4.3.2 Vignetting Function

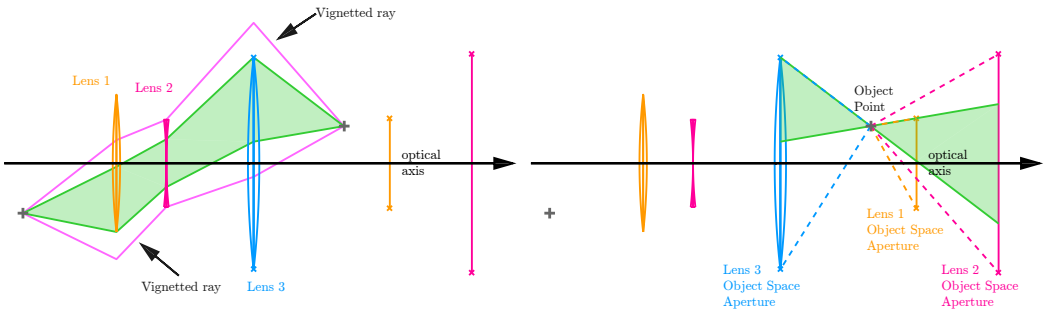
Before the selection and pairing occur, the ECA is a collection of two-aperture elements, one for each possible pair of pixel and DMU. The principle behind a two-aperture element is to abstract from how the imaging is done by the DMU and how the light is integrated by the pixel by only considering the bundle of rays going through their two apertures. In a sense, the aperture of the DMU acts as the aperture stop as it is usually the smallest aperture as seen from the pixel point of view. The pixel aperture acts as the field stop.

In order to compute vignetting, the traditional method is to trace rays sampling the field and the pupil at different positions, to propagate them through the system and check if they are intersected by the aperture of an optical component or a diaphragm, see Figure 4.6a. The ECA exists in object space, hence, there is a different approach to this computation which benefits us.

The alternate approach consists in imaging the apertures of every single optical component into object space. We call these apertures in object space, the system apertures, see Figure 4.6b. Then, vignetting can be evaluated without having to trace rays through the system. The angle of the system apertures boundaries is directly determined from the object point of interest and the ones with the smallest values are the windows of the system. This method is particularly facilitated using the point matrix tools of Chapter 3.

We introduce the discrete vignetting function (DVF) as the fraction of the initial ray bundle defined by a pixel/DMU two-aperture element that passes through the system. The DVF is a table of coefficients between 0 and 1 and there is one coefficient for each possible pixel/DMU pair. In the case where the bundle of rays of the pair is not cut by a system aperture which may happen for an element close to the optical axis, the coefficient of this pair is equal to one. When the bundle of rays is partially cut, it is not defined by the DMU and pixel apertures only but also by the apertures limiting it. The coefficient of the pair is now less than one but still larger than zero. The coefficient is zero when the ray bundle of the two-aperture element is completely blocked by one or more system apertures.

The two-aperture elements selected for pairing are the ones with a non-zero value from the vignetting function as they integrate some light. The pairing and its consequence on the ECA properties are discussed in Section 4.3.4. The next section addresses the topic of evaluating the DVF.



(a) Ray tracing method (b) Aperture mapping method

Figure 4.6: Vignetting computation. Lens 2 is the stop of the system. (a) Ray diagram. The pink ray bundle passes through the full stop. It is limited by the bottom of the aperture of lens 1 and by the top of the aperture of lens 3. The green ray bundle is the largest unvignetted ray bundle. (b) Aperture diagram. Tracing lines from the object point to the boundaries of the apertures mapped in object space visually indicates which apertures affect vignetting.

4.3.3 Evaluation

We propose two different methods to evaluate the vignetting function.

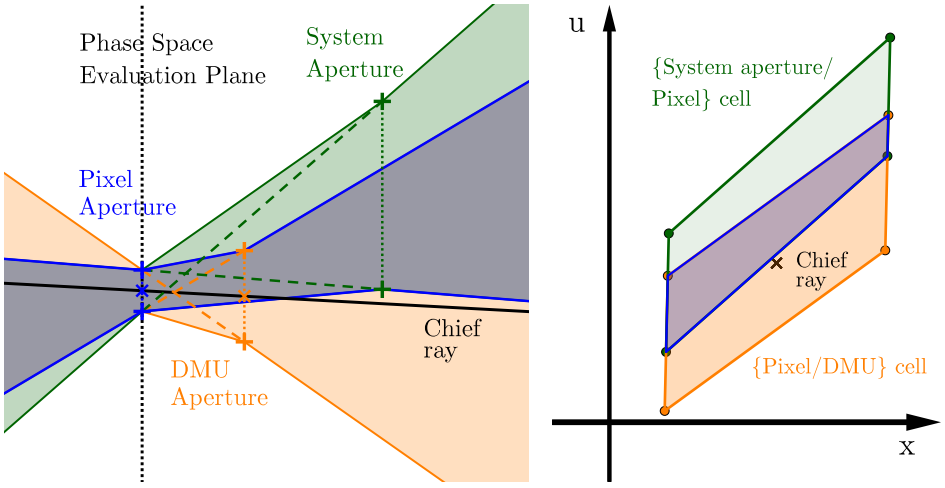
Evaluation with a Chief Ray

The first approach is to use a single chief ray. This chief ray is the ray going through the center of both pixel and DMU apertures. With this ray as the representation of the full ray bundle, we check if it passes through each system aperture. The DVF is then a binary function as the ray is either blocked by at least one system aperture which corresponds to a 0 or not blocked by any aperture which corresponds to a 1.

This method is better at selecting pixels than the naive approach, Section 4.3.1, but forces a threshold onto partially lit pixels. Moreover, it considers only the chief ray as the important ray defining the two-aperture elements, which ignores completely the aperture aspect of the elements, see Figure 4.7. However, it is easy to compute and it records well the off-axis projection effect where the images of the DMU on the sensor are not located directly behind it but are shifted toward the outside for DMUs far off the optical axis, see [Dansereau *et al.*, 2013b].

Evaluation in Phase Space

The second method determines the exact value of the DVF of a pixel/DMU pair as the ratio of the etendue of the ray bundle that passes through the pixel, the DMU and the systems apertures, $G_{Opt.Sys.}$, to the etendue of the ray bundle of the pixel/DMU



(a) Object space (b) Phase Space

Figure 4.7: Vignetting of a two-aperture element (a pixel/DMU pair) by a system aperture. The orange ray bundle delimited by the pixel and DMU apertures is partially blocked by the system aperture. In this sketch, the chief ray does not pass through the system aperture. In phase space, the corresponding point for this chief ray lies outside of the intersection region of the two cells in blue. As a result of the intersection region not being empty, the pixel still receives a partial amount of light.

pair, $G_{pix./DMU}$.

$$DVF(pixel, DMU) = \frac{G_{Opt.Sys.}}{G_{pix./DMU}} \tag{4.1}$$

Computing these values of etendue relies on a representation of the ray bundle in the phase space diagram.

Ray set A ray bundle is a set of rays in the 3D space. The phase space point $\mathbf{p}$ of a ray is a parametrization of this ray's position and direction in space. It depends on the choice of an evaluation plane. Let us define an indicator function χ over this phase space that associates a binary value for each ray, 1 if it is part of the ray bundle and 0 if it is not. This function is of the form $\chi(\mathbf{p})$ where $\mathbf{p} = (x, y, u, v)$ with x, y being the coordinates of a ray intersection with the main plane of the two-plane parametrization and u, v being the intersection coordinates with the second plane of the parametrization. A ray bundle is represented in phase space as a set of points $\mathcal{R}$ for which the indicator function is equal to 1. In a more formal manner, $\mathcal{R}$ is defined as $\mathcal{R} = \{\mathbf{p} \mid \chi(\mathbf{p}) = 1\}$.

Apertures Apertures in an optical system are usually convex 2D shapes. Most of the time, the rims of lenses are circular and their apertures as well as diaphragms are usually convex polygons, regular or not, with straight or curved edges. Let

us assume in the following arguments that every aperture in an optical system is convex. Let us consider the ray bundle delimited by such a convex aperture Ap, the aperture plane being the evaluation plane of the phase space. Before being cut by the aperture, the set $\mathcal{R}$ is equal to the whole phase space but only a subset $\mathcal{R}'$ of $\mathcal{R}$ passes through the aperture, the rays that hit the open part of the aperture. $\mathcal{R}'$ is the outer product of the set $\mathcal{A}$ of points (x, y) lying in the aperture plane within the aperture boundaries and of the set $\mathcal{D}$ of all (u, v) directions. Due to the aperture letting rays with any direction pass through, $\mathcal{D} = \mathbb{R}^2$ so $\mathcal{D}$ is convex. The aperture Ap is convex so $\mathcal{A}$ is convex too. Consequently, $\mathcal{R}'$ is convex.

Change of the phase space evaluation plane Let us now define a transport operation on the phase space corresponding to the transport of the evaluation plane from one place to another. The transport operation is a geometrical transformation applied to the evaluation plane such as a translation or a rotation. So this transformation is reversible. The change in the evaluation plane does not affect the set of rays defined by $\mathcal{R}'$ but only modifies the mapping between a ray and its corresponding phase space point such that the phase point in the new evaluation plane $\mathbf{p}'$ is defined as $\mathbf{p}' = f(\mathbf{p})$ with f the transport function between the old and new evaluation plane. Therefore, $\mathcal{R}'$ is expressed as $\mathcal{R}' = \{\mathbf{p}' \mid \chi(f^{-1}(\mathbf{p}')) = 1\}$. The transport operation does not change the convex property of $\mathcal{R}'$.

Ray set of many apertures Let us apply this scheme to the object space system apertures of an optical system. Consider at first a simple system made of two systems apertures only: a first convex aperture at the evaluation plane and a second convex aperture at another position. The goal is to determine how the ray set going through both apertures is expressed. The first aperture defines a convex set $\mathcal{R}'_1$ of the rays passing through its aperture, then the phase space evaluation plane is transported to the second aperture plane. At this plane, the second aperture also defines a convex set $\mathcal{R}'_2$ of the rays passing through its aperture. The rays passing through both apertures form the set $\widetilde{\mathcal{R}}$ defined as the intersection of both convex sets $\mathcal{R}'_1$ and $\mathcal{R}'_2$, $\widetilde{\mathcal{R}} = \mathcal{R}'_1 \cap \mathcal{R}'_2$. This set is also convex. The corresponding indicator function $\widetilde{\chi}$ for the evaluation plane being the second aperture plane is :

$$\widetilde{\chi}(\mathbf{p}) = \begin{cases} 1, & \text{if } \mathbf{p} \in \widetilde{\mathcal{R}} \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

For more than two object space system apertures, the process is repeated, moving the evaluation plane from one aperture plane to the next to define the ray set of the current aperture and intersect it with the previous ray set, see Figure 4.7 for an example with three apertures in 2D. An alternative method is to set multiple evaluation planes for the phase space, one at each aperture plane, define the ray set of the aperture there and then transport it to a common arbitrary evaluation plane to perform the intersection.

Etendue The etendue of a ray set $\mathcal{R}'$ is defined as the integral of its indicator function χ.

$$G = \int_{\mathcal{R}}' \chi(\mathbf{p}) d\mathbf{p} \tag{4.3}$$

The convexity of $\mathcal{R}'$ is important for two main reasons. The first one is that convex apertures simplify the algorithmic computation of the ray bundle etendue. We address the algorithmic computation of the vignetting function in Chapter 7. The second reason is that it guarantees that the ray bundles do not split during the process of transport through the system apertures.

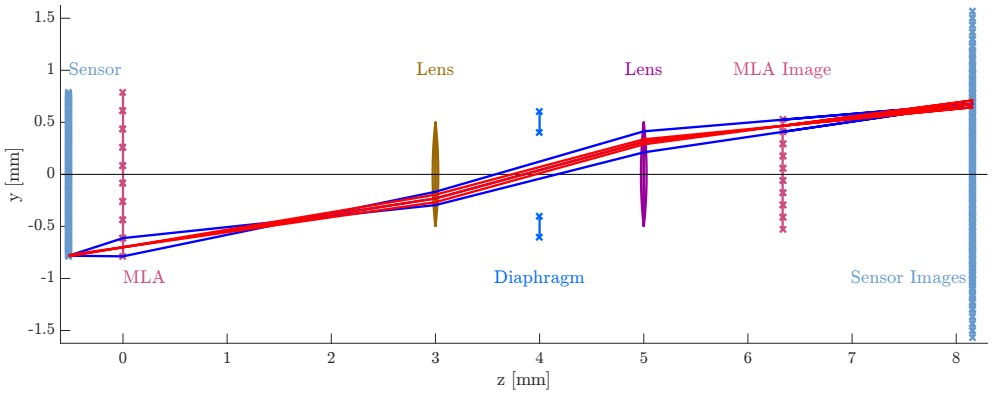
4.3.4 Effect on the Properties of the ECA

Once the vignetting function is known for every pixel/DMU pair, the selected pairs are the ones with a non-zero value. The pairing rule is the same as without vignetting considerations: the two-aperture elements that share the same aperture imaged from the same DMU element make a virtual camera.

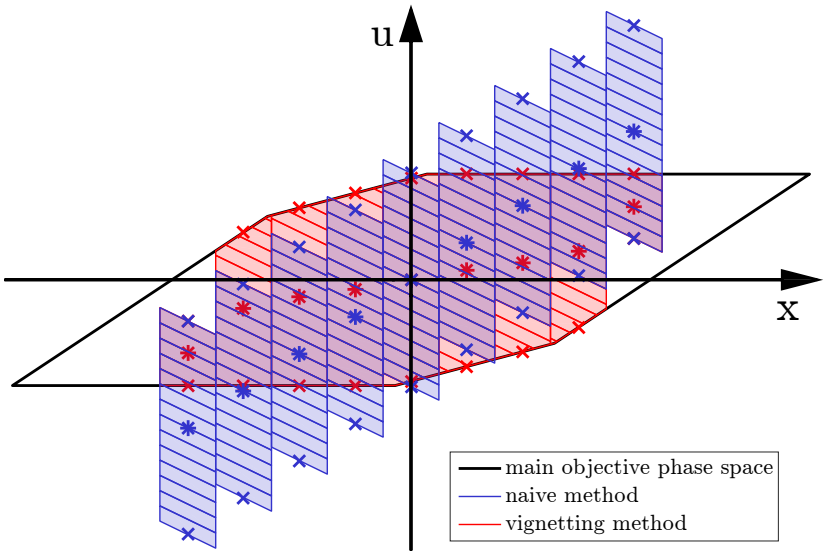
Figure 4.8b illustrates the phase space of a 2D vignetted system similar to the focused light field camera with a main objective made of two lenses with a diaphragm in the middle. The DMU of this system is a microlens array (MLA). The phase space evaluation plane is at the virtual aperture planes (the MLA image plane) so that the virtual cameras are represented by a column of cells. The naive method, Section 4.3.1, has a constant number of virtual pixels, for each camera (in blue) whereas, with the vignetted method (in red), the system apertures cut out a different set of cells. Unvignetted cells, on the center, have the shape of a rhombus and vignetted cells, on the border, are clearly distinguishable with their various smaller polygonal shapes. The figure also shows that some cameras seem to extent their field beyond what the naive method simulates. This effect is due to a shift in the position of the micro-images on the sensor. We explain this effect in more detail in Section 4.3.5. Moreover, the number of pixels of the side cameras decreases naturally. However, for the two extremal cameras, the sensor edges are the cause for the discontinuity in the red area causing the first and last cameras to be associated with fewer pixels than what it could with a wider sensor.

This change of pixel assignment affects most of the properties of the virtual cameras. As mentioned earlier, the optical properties can be retrieved directly from the phase space but this time, instead of considering the phase space made of the unvignetted pixel/DMU cells, the space contains the set of vignetted cells,

View direction and angle of view A camera is a xy column of cells in the phase space evaluated at the virtual camera aperture plane. By cause of the cells being possibly vignetted, they do not keep their regular shape. The view direction and the angle of view are changed. These properties are obtained by considering that the virtual aperture of a camera is a pinhole at its center. In the phase space, this



(a) Setup



(b) Phase space

Figure 4.8: Vignetting in a light field camera. (a) Setup of a light field camera. The main objective is made of two lenses and a diaphragm. Blue and red rays originate from the center of a pixel and a microlens, respectively. (b) Phase space with the naive vignetting method in blue and the method using the vignetting function in red. The $*$-shaped points indicate the view direction and the $\times$-shaped points indicate the limits of the angle of view of each virtual camera. The black polygon delimits the phase space region of the main objective of the camera.

is equivalent as slicing the cell stack with a plane of constant xy coordinates, these constant coordinates being the xy coordinates of the pinhole in the evaluation plane. The centroid of this surface of intersection gives the view direction. For the angle of view, it is the edge of the surface that matters and two directions are usually defined, horizontal and vertical. The two points on this edge with the same u coordinate as the centroid give the angle of view in one direction and the two points on the same

edge with the same v coordinate as the centroid indicate the angle of view in the other direction.

The visualization is easier to understand in a 2D system, see Figure 4.8. In the 2D phase space, reducing a virtual camera to a pinhole gives a line, the view direction is given by the center of this line and the angle of view by the edges of the line.

Other properties The depth of field is unaffected by vignetting as it only depends on the aperture sizes and distance along the optical axis. The rest of the properties like the baseline, transversal and longitudinal accuracies change too because they depend on the pixel assignment, view direction, and field of view of every virtual camera in the ECA.

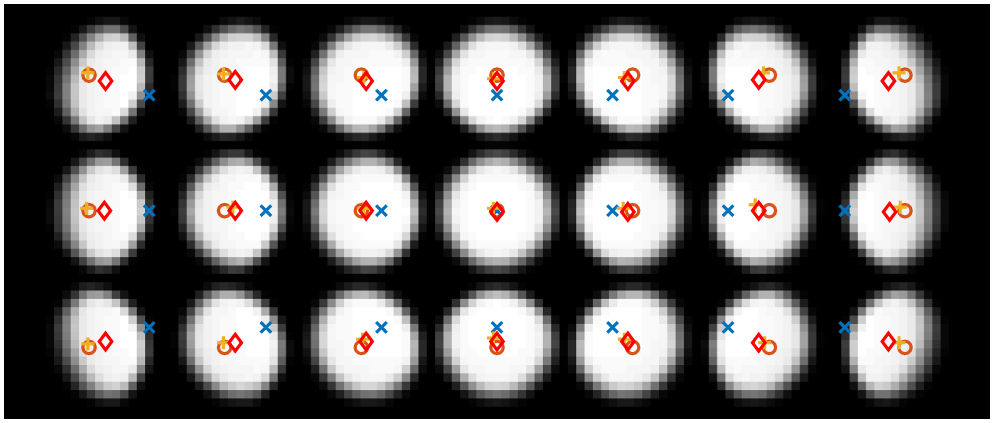
Moreover, it is also quite possible, that some pixels receive partial light from several DMUs. For such a pixel, the DVF is non-null for different DMU entries. In order to visualize this, the phase space needs to be evaluated at the sensor plane in its original space. Pixels are then represented as vertical stacks of cells, a vertical stack contains as many cells as the number of DMUs a pixel receives light from.

4.3.5 Calibration Image

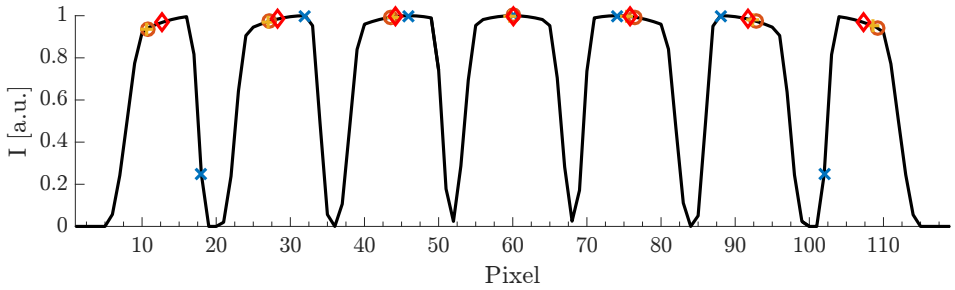
In traditional photography or in microscopy, the picture of a white lambertian scene reveals the effects of the mechanical and natural vignetting. In light field cameras, a picture in these conditions is necessary for calibration purposes. It is used to retrieve the parametrization of the 4D light field on the 2D sensor. The goal is to associate the correct angular and spatial parameter to each pixel for further analysis and reconstruction of the light field. The ECA model allows to simulate this calibration image with the vignetting function. The first step is to compute the effective etendue of each pixel/DMU pair by multiplying the unvignetted etendue of the pair by the coefficient of the DVF. The second step is to integrate the resulting vignetted etendue function over the DMUs. Then, each pixel has a single value of etendue.

$$G(pixel) = \sum_{DMU} DVF(pixel, DMU)\, G_{pix./DMU} \tag{4.4}$$

Figure 4.9 illustrates the white image of a light field camera in the focused configuration from the setup of Figure 4.8a. The DMU is a microlens array. The parameters of the DMU and sensor are exaggerated for better visualization. Its calibration image and the profile present no superposition nor large gap between adjacent micro-images. the center of these micro-images is the centroid of the pixels weighted by their brightness (red diamond-shaped markers). We observe many effects in this figure.



(a) White Image



(b) Profile of the center row

Figure 4.9: White image of the system from Figure 4.8.

1. The orthogonal projection of the microlens centers onto the sensor (blue +-shaped markers) is not the micro-image center. This orthogonal projection is also where the brightest point of the micro-image lies because this is the position that suffers the less from natural vignetting. The direction of the light rays that hit this position is close to the direction normal to the sensor. However, this position can suffer from mechanical vignetting as is the case for the extreme right and left micro-images, see Figure 4.9b.

2. The perspective projection of the microlens centers (yellow ×-shaped markers) onto the sensor is the reference point for the light field parametrization. The projection center is the center of the exit pupil of the main objective (two lenses and a diaphragm in this case) of the optical system. The reason is that even though the microlenses does not conjugate the exit pupil with the center, the exit pupil primarily defines the light cone that reaches each of the microlenses. For the micro-images close to the center of the sensor that suffer the less from mechanical vignetting, the perspective projection matches perfectly their centroids. But for microlenses far from the center, mechanical vignetting in the objective changes the shape of the micro-images from a circle to a cat's eye shape resulting in shifted centroids.

3. The position of the exit pupil is important. It directly influences the position of the perspective projection centers and the width of the micro-images. With an exit pupil closer to the microlens array, the micro-images would be larger and their perspective center further apart. Considering that the sensor parameters are fixed, it means that fewer micro-images would fit onto it but their resolution would increase. This phenomenon exhibits a spatio-angular trade-off linked to the position of the exit pupil. In the figure, for instance, the distance between neighboring perspective centers is 16 pixels whereas the distance between neighboring microlenses center is equal to 14 pixels. Moreover, in the case where the exit pupil is at infinity or sufficiently far away from the microlens array), the offset between the perspective projection center and the orthogonal projection center is null or greatly reduced.

4. The perspective projection centers are necessary to extract the correct light field parametrization but this information may not always be available. With only the white raw image and no knowledge of the microlens array parameters (distance to sensor, pitch of the elements), it would be difficult to determine accurately the correct parametrization of the pixel grid. Many computer vision algorithms attempt to recover this reference points from the white image brightness distribution. Some methods relies on taking the brightest points of each micro-image [Dansereau *et al.*, 2013b] but they would fail if the exit pupil is not far enough from the microlens array. Other methods compute the centroids or an equivalent reference point [Cho *et al.*, 2013] but would fail if the mechanical vignetting of the objective is too strong.

   We propose to estimate the true position from a simple observation. The cat's eye shape of the micro-images on the side is usually the result the stop aperture and another field aperture blocking the light rays. Since both apertures are circular, it produces this particular shape. By using a circle search algorithm on each micro-image, the center and radius of the circles can be computed. One center is most likely to be the perspective projection of the exit pupil center. This computation was performed on the white image from the figure resulting into the (orange circle-shaped markers). The points obtained with this method match the true perspective centers the best.

4.4 Discussion

4.4.1 Camera Array Equivalence

If a real camera array were to be constructed with the characteristics of the ECA (position of the centers and shape and size of the apertures) corresponding to a light field camera, the light field measured by this real camera array would be the same to the first-order as the one measured by the light field camera. One difference between the real and the virtual array is that the main lens front plane is possibly at a

different position than the virtual lens plane where the real array must be placed. In case where the array position is in front of the main lens as in Chapter 5, a physical array would not be able to see an object lying between the main lens front plane and itself, whereas a light field camera can also measure this part of the object space.

The other notable difference with a real array is that the abstraction made with the two-aperture model implies a loss of the optical properties of the apertures. As such, the effect of refraction is ignored and consequently, the focal length of the cameras of the ECA is unknown. Moreover, the real pixel pitch and sensor position are also unknown. These parameters characterize the image side of the cameras and cannot be predicted by the ECA model. Actually, all of these parameters are linked together and are parameterized by the focal length value which is free to choose. The constraints are fixed by the position and pitch of the pixels in object space given by the ECA. In the end, multiple camera arrays can be made having the same sampling as the ECA.

Finally, the main condition to obtain an equivalent camera array is to create a virtual camera from the grouping of several two aperture elements sharing the same aperture, see Section 4.3.1. The center of this common aperture is considered to be the center of projection of the virtual camera. For imaging systems that do not maintain the condition of having a common aperture, the ECA of the system does not exist. As an example, imaging systems that use components that create non-perspective views [Lanman *et al.*, 2006] or that are too disordered [Wender *et al.*, 2015] break the condition.

4.4.2 Limitations of a First-Order Model

The equivalent camera array model is a first-order model that can accurately predict the properties of a light field camera if the Gauss conditions are respected. The apertures of the optical elements must remain small as well as the angles of the rays with respect to the optical axis. Using wider apertures goes with an increase of the effect of aberrations and degrades the quality of the measured light field. This implies that light rays that in the ideal setting pass through the center of perspective of a given virtual camera, now do not converge to a single point. The effect of these aberrations are pointed out in the experiment of Section 5.2

4.5 Conclusion

We proposed a model based on constructing an virtual equivalent camera array (ECA) to describe the characteristics of a light field camera. Our model abstracts the physical elements of the sensor and the directional multiplexing unit into a pair of apertures that can be imaged independently and grouped to define an individual virtual camera. We can quantify most imaging properties for each virtual camera with the exception of focal length, for which we can only determine a family of

solutions.

Moreover, on top of the imaging configuration, the model takes into account vignetting considerations that are critical in the design process as well as the necessary calibration step that produces the pixel to ray correspondence.

We envision that it would be possible to retrieve some of the camera's parameters directly from the properties assuming a particular architecture of the system. This kind of process would be necessary to retro-engineer a device from a limited set of information or to optimize a system for a specific application.

Part III

Applications

Chapter 5

Application to MLA-based Light Field Cameras

In order to illustrate the ECA model presented in the previous chapter, we apply it to light field cameras based on the use of a microlens array, see Section 2.2.3. This chapter details the construction of the ECA from the parameters of the light field cameras and derives analytical expressions for their optical properties. The focused and afocal configurations are analyzed and compared through the spectrum of these properties. Moreover, we perform an experimental validation where the main lens parameters of an afocal light field camera are retrieved from its measured optical properties.

5.1 Simulation Based on the ECA Model

The afocal and focused configurations of light field cameras were deeply investigated in the literature. In terms of optical design, these systems are simple as there are only three main components: the sensor, the microlens array placed close to the sensor that serves as the directional multiplexing unit and a main lens and can be modeled with a few parameters. We consider that the microlenses are thin lenses and are placed at the origin of the coordinate system perpendicularly to the z-axis which is the optical axis of the system. The main lens is modeled as a thick lens. The notations for the system parameters are defined in Figure 5.1. The two configurations differ only slightly: in the focused case, the distance between the microlens array and the sensor is the distance b but for the afocal case, it is equal to the focal length of the microlens f_p.

5.1.1 Construction of the ECA

The two-aperture system we need to image to the object space of the camera is made of a pixel and a microlens. We respectively denote the position of the center

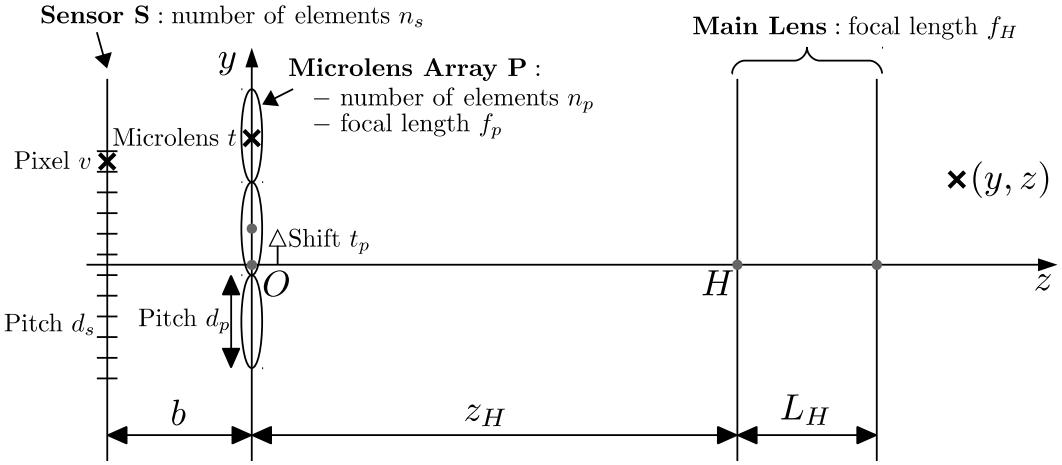


Figure 5.1: Sketch of the light field camera system indicating the used notations.

of these apertures by $\mathbf{C}_Q$ and $\mathbf{C}_P$. Their positions in the camera object space $\mathbf{C}_{Q_W}$ and $\mathbf{C}_{P_W}$ are obtained by applying the equations from Chapter 3 as follows:

$$\widehat{\mathbf{C}}_{Q_W} = \mathcal{L}_{Main}\mathcal{L}_{DMU}\widehat{\mathbf{C}}_Q\,, \tag{5.1}$$

$$\widehat{\mathbf{C}}_{P_W} = \mathcal{L}_{Main}\widehat{\mathbf{C}}_P\,. \tag{5.2}$$

The edges of the apertures $\mathcal{A}_{Q_W}$ and $\mathcal{A}_{P_W}$ in the camera object space can be imaged in the same way from $\mathcal{A}_Q$ and $\mathcal{A}_P$, the apertures of the pixel and microlens. Figure 5.2 illustrates the position and the field of view of the virtual cameras of the ECA.

Imaging Matrices

In order to derive the analytical expressions for the ECA properties, we detail the imaging matrices for the 2D case[1] with our notations starting with the matrix $\mathcal{L}_{Main}$ of the main lens:

$$\mathcal{L}_{Main} = T_H L_{\Phi_H} T_{(0,L_H)} T_H^{-1}\,, \tag{5.3}$$

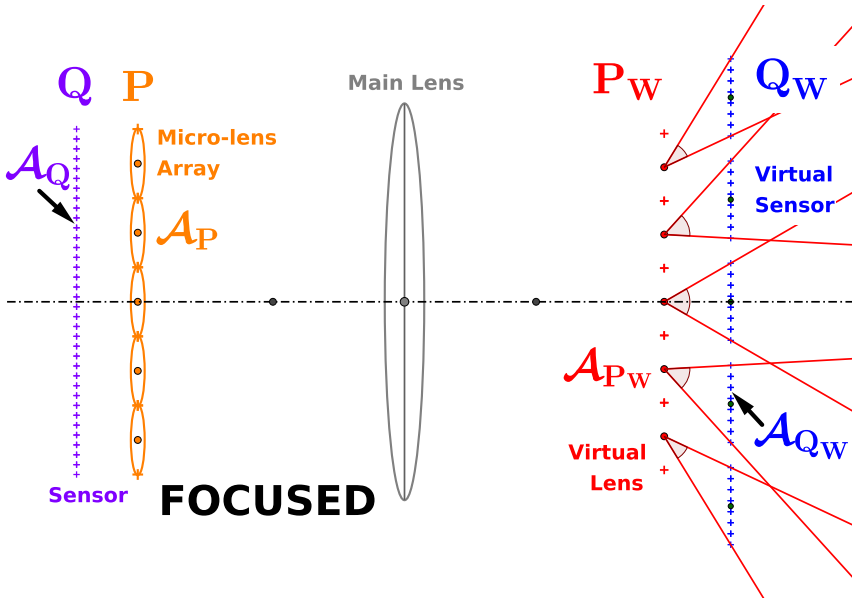
with L_{Φ_H} the lens matrix, $T_{(0,L_H)}$ the matrix accounting for the thickness of the lens and T_H the position matrix for the lens, defined as follows:

$$L_{\Phi_H} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \Phi_H & 1 \end{pmatrix},\ T_{(0,L_H)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & L_H \\ 0 & 0 & 1 \end{pmatrix},\ T_H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & z_H \\ 0 & 0 & 1 \end{pmatrix}, \tag{5.4}$$

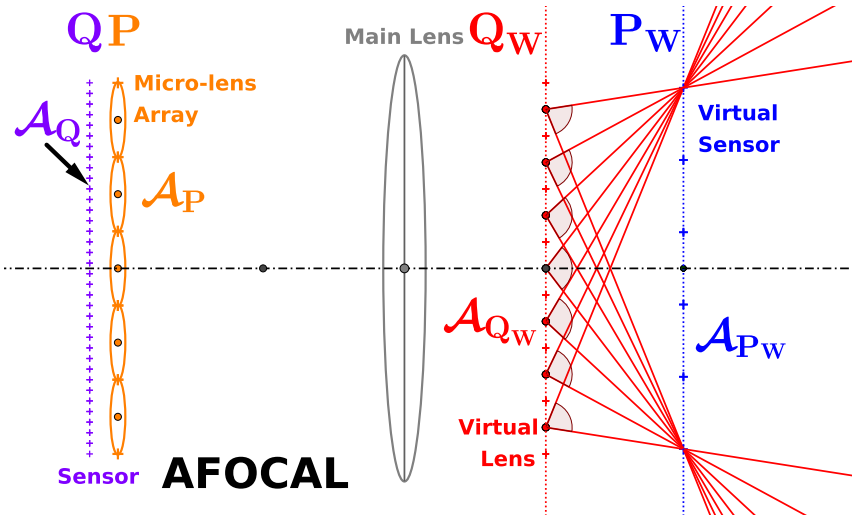
where Φ_H is the power of the lens defined as $\Phi_H = \frac{1}{f_H}$.

For the microlens array, each microlens is given an index t between 1 and n_p. Even though the microlens array is at the origin, it is not necessarily centered with

[1] Note that the imaging matrices are only 3×3 in this case.



(a) Focused configuration



(b) Afocal configuration

Figure 5.2: ECA of the focused and afocal light field camera. The fields of view of the individual cameras of the ECA are shown in red. The layout of this figure is the same for the following figures plotting the different optical properties. The parameters of the different components have been chosen to have a compact figure and do not represent a realistic imaging system. There would usually be an overlap of the virtual sensors for neighboring virtual cameras in the focused configuration.

the optical axis so we introduce a shift t_p to account for this possible misalignment. The matrix of a microlens of power $\Phi_p = \frac{1}{f_p}$ is thus:

$$\mathcal{L}_{DMU} = T_t \, L_{\Phi_p} \, T_t^{-1} \tag{5.5}$$

with L_{Φ_p} the lens matrix and T_t the position matrix for the microlens t:

$$T_t = \begin{pmatrix} 1 & 0 & t_p - d_p \left(\frac{n_p}{2} + \frac{1}{2}\right) + d_p \, t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{5.6}$$

Position of the Microlens in Object Space

We also introduce the matrix K_p that converts the index t of a microlens to the homogeneous coordinates of the microlens center:

$$K_p = \begin{pmatrix} d_p & t_p - d_p \left(\frac{n_p}{2} + \frac{1}{2}\right) \\ 0 & 0 \\ 0 & 1 \end{pmatrix} . \tag{5.7}$$

The position in homogeneous coordinates for the image of the center of a microlens t in object space is:

$$\widehat{\mathbf{C}}_{P_W} = \left(T_H L_{\frac{1}{f_H}} T_{(0,L_H)} T_H^{-1} \right) K_p \begin{pmatrix} t \\ 1 \end{pmatrix} . \tag{5.8}$$

The explicit form of the microlens image in world coordinates is:

$$\mathbf{C}_{P_W} = \begin{pmatrix} -\dfrac{f_H \, (2\, t_p - d_p \, (n_p - 2\, t + 1))}{2\, z_H - 2\, f_H} \\ \dfrac{{z_H}^2 + L_H \, z_H - L_H \, f_H}{z_H - f_H} \end{pmatrix} . \tag{5.9}$$

In order to simplify and shorten the equations, we introduce the following adimensional parameters:

$$\alpha = \frac{L_H}{f_H}, \; \gamma = \frac{z_H}{f_H}, \; \beta = \frac{b}{f_p}, \; \delta = \frac{f_p}{f_H}. \tag{5.10}$$

The variable α tells about the thickness of the main lens. The parameter γ tells about the imaging mode of the main lens for the MLA plane, that is, the image of the microlens array plane is real when $\gamma > 1$, located at infinity when $\gamma = 1$ and virtual when $\gamma < 1$. The variable β tells about the imaging mode of the microlens array for the sensor. It is Galilean for $\beta > 1$, afocal for $\beta = 1$ and Keplerian for $\beta < 1$. Finally, δ is simply the ratio of the focal lengths of the microlens and the main lens.

We can substitute these parameters into Equation 5.9 to obtain:

$$\mathbf{C}_{P_W} = \begin{pmatrix} -\dfrac{t_p - d_p\,\left(\frac{n_p}{2} - t + \frac{1}{2}\right)}{\gamma - 1} \\ f_H\,\left(\alpha + \dfrac{\gamma^2}{\gamma - 1}\right) \end{pmatrix}. \tag{5.11}$$

Position of the Pixels in Object Space

The sensor is centered on the optical axis and the pixels are indexed by v from 1 to n_s, so the matrix K_s transforming the pixel index into the coordinate of its center is:

$$K_s = \begin{pmatrix} d_s & -d_s\,\left(\frac{n_s}{2} + \frac{1}{2}\right) \\ 0 & -b \\ 0 & 1 \end{pmatrix}. \tag{5.12}$$

with $b = f_p$ in the afocal case.

The position in homogeneous coordinates for the image of the center of a pixel v through the microlens t and main lens in object space is:

$$\widehat{\mathbf{C}}_{Q_W} = \left(T_H L_{\frac{1}{f_H}} T_{(0,L_H)} T_H^{-1}\right)\left(T_t\, L_{\frac{1}{f_p}}\, T_t^{-1}\right) K_s \begin{pmatrix} v \\ 1 \end{pmatrix}. \tag{5.13}$$

The explicit form of the pixel image in world coordinates is then given by:

$$\mathbf{C}_{Q_W} = \begin{pmatrix} \dfrac{\beta\,\left(t_p - d_p\,\left(\frac{n_p}{2} - t + \frac{1}{2}\right)\right) + d_s\,\left(\frac{n_s}{2} - v + \frac{1}{2}\right)}{\beta\,\delta - (\beta - 1)\,(\gamma - 1)} \\ f_H\,\left(\alpha + \gamma + 1 - \dfrac{\beta - 1}{\beta\,\delta - (\beta - 1)\,(\gamma - 1)}\right) \end{pmatrix}. \tag{5.14}$$

In the **afocal case**, $\beta = 1$, so the expression for $\mathbf{C}_{Q_W}$ in the previous equation can be simplified further:

$$\mathbf{C}_{Q_W} = \begin{pmatrix} \dfrac{t_p + d_s\,\left(\frac{n_s}{2} - v + \frac{1}{2}\right) - d_p\,\left(\frac{n_p}{2} - t + \frac{1}{2}\right)}{\delta} \\ f_H\,(\alpha + \gamma + 1) \end{pmatrix}. \tag{5.15}$$

5.1.2 Afocal Case

There exists a specific case for which the role of the apertures as the virtual sensor or the virtual camera can be switched. This occurs when the distance between Q and P is equal to the focal length of the microlens array. In this case, the image of the pixels are sent to infinity by the microlenses and then the main lens images the pixels in its front focal plane as can be seen in Figure 5.2b. The pixels with the same relative position to the center of their assigned microlens integrate light rays of the in-camera light field with the same direction.

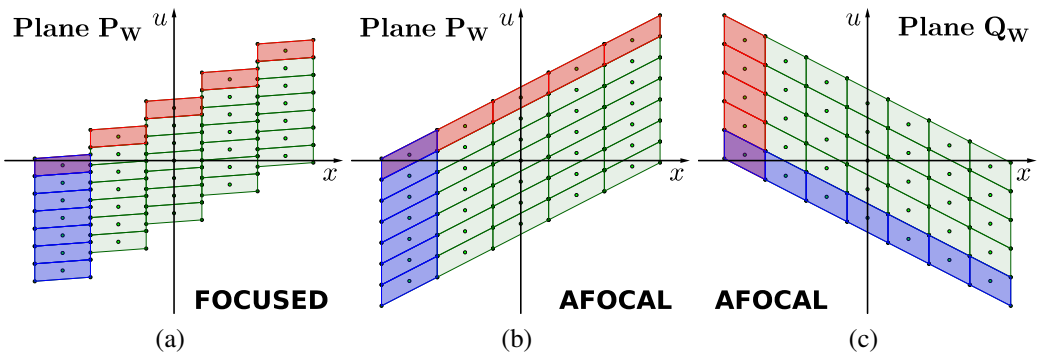


Figure 5.3: Phase space of the ECA. A camera view is obtained by summing all cells from one column. In the focused case in (a), the evaluation plane is the P_W plane, a column of cells in blue corresponds to different pixels from the same DMU element. For the afocal configuration, there are two ways to obtain a column of cells: when the evaluation plane is P_W as for the focused case in (b) or when the evaluation plane is Q_W in (c). In the latter case, a column of cells in red corresponds to the same relative pixel position from different DMU elements.

The difference with the focused configuration is better explained in a phase space diagram as shown in Figure 5.3. In this space, a virtual camera is represented by a vertical column of contiguous pixels. The only position for which this condition is satisfied in the focused case is at the P_W plane, see Figure 5.3a.

In the afocal configuration, though, there are two configurations that yield virtual cameras, 1. for the evaluation plane positioned at the P_W plane and 2. for the evaluation plane positioned at Q_W, see Figure 5.3b and Figure 5.3c. Choosing one plane or the other allows for creating two different ECAs, that, however, describe the same set of rays. It may be noted that the two apertures of the ECA model determine the boundaries of the phase space parallelograms.

5.1.3 Simulation

The microlens array plane is fixed at the origin, so, the effect of moving the object plane or (physically) refocusing with the main lens is simulated by only changing the position of the main lens along the optical axis and computing the camera array properties. The value of the refocus is the distance between the microlens array and the back principal plane of the lens. We also investigate the effect of a varying distance between the sensor and the microlens array.

The main lens is approximated by a thin lens. Vignetting effects (see Section 4.3) between the microlenses and the main lens are not taken into account as it would change the pairing between pixels and microlenses and make the following discussions more difficult. The additional complexity is studied with the light field microscope setup in Section 6.2.2. Here, we prefer to focus on the system

Table 5.1: Values for the parameters for the setup described in [Bishop and Favaro, 2012] used in our comparison study.

Parameter	Pixel	Microlens	Main Lens
Pitch/Diameter (mm)	0.009	0.135	40.0
Focal length (mm)	-	0.35	80.0
Number	4095	273	1

parameters.

In the following study, we present the results for the light field camera from [Bishop and Favaro, 2012] as an illustrative example as its properties show the effects of microlens-based light field camera designs most clearly. This camera has been designed to be used in the focused configuration and as such it is representative of similar existing systems. However, it was not particularly designed to be used in the afocal configuration. In order to compare the focused and afocal configurations on a common basis, we created an afocal version of [Bishop and Favaro, 2012].

The parameters of the components are summarized in Table 5.1. The only difference between the afocal and focused setups is the distance between the sensor and the microlens array. It is equal to the microlens focal length for the afocal setup and it is 1.2 times this value for the focused setup.

The Appendix presents the array properties computed for other values of the distance between the sensor and the microlens array as well as for two other light field cameras: the Lytro and the Raytrix from their respective patents.

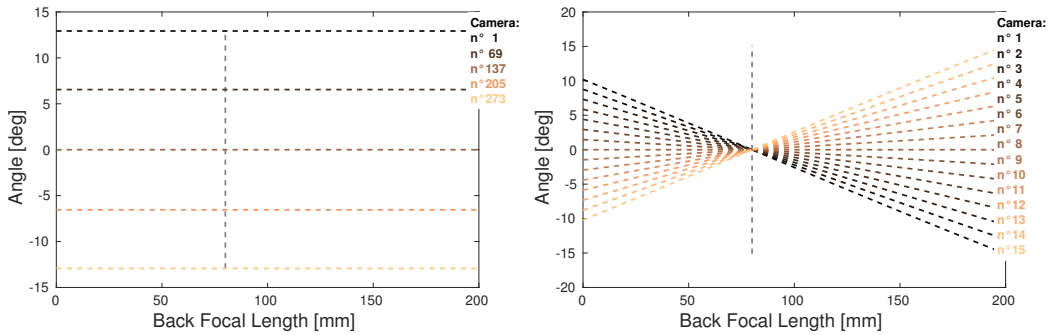
5.1.4 Properties

The several optical properties of the ECA of the focused and afocal light field camera are derived from the position of the two aperture elements in object space, $\mathbf{C}_{P_W}$ and $\mathbf{C}_{Q_W}$ from Equations 5.11, 5.15 and 5.14. The analytical expressions of the properties are regrouped in Table 5.3 and are discussed one by one in the following section. The ECA of the afocal and the focused configurations are made differently, hence the properties show specific behaviors. Analyzing and understanding these behaviors as well as their dependences can help design an optical system for a specific application.

Plane positions The first interesting property is the position of the planes Q_W and P_W along the optical axis. Table 5.2 shows a summary of the possible positions of these planes depending on the position of the main lens (either at one focal length away from Q' or P or neither) and the configuration of the sensor and microlens array. The distance separating the sensor and the microlens array can either be lower, equal, or greater than the focal length of the microlenses and corresponds to configurations called respectively galilean, afocal, and keplerian. The virtual

Table 5.2: Position of the two aperture planes P_W and Q_W in object space. The letters R, F, V respectively indicate that the plane is real, in the front focal plane of the main lens or virtual.

Lens Pos	Img P		Img Q'		Neither	
Plane Pos	P_W	Q_W	P_W	Q_W	P_W	Q_W
Galilean	∞	R	V	∞	R/V	R/V
Afocal	∞	F	-	F	R/V	F
Keplerian	∞	V	R	∞	R/V	R/V



(a) Focused case (b) Afocal case

Figure 5.4: **Viewing direction** of each of the virtual cameras of the ECA. There are as many virtual cameras as microlenses in the focused case (a), so 273. For the afocal case (b), the number of virtual cameras is the number of pixels behind a microlens or equivalently the ratio of the number of pixels to the number of microlenses, so 15. The vertical gray dashed line indicates the focal length of the main lens.

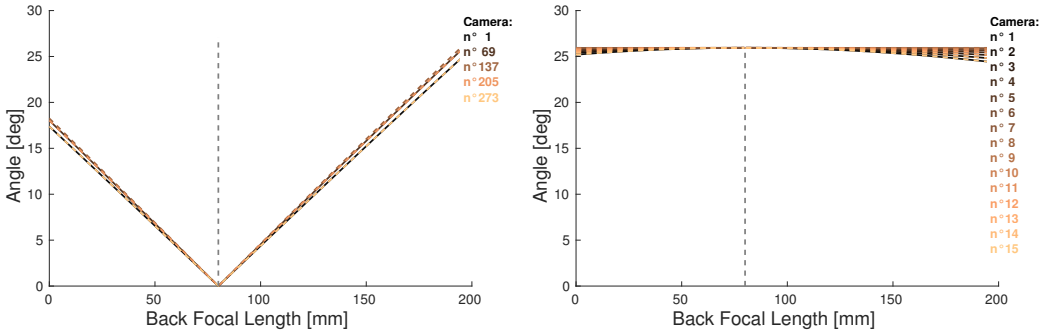
camera array can be made of perspective or directional cameras, looking at a real or virtual plane, at a finite or infinite distance. In the afocal setup, Q_W is always located in the front focal plane of the main lens whatever its position is. This also means that, for this case, the pitch and height of the pixel aperture in object space do not depend on the position of the main lens, they are independent of γ.

View direction The view direction is the angle between the line connecting $\mathbf{C}_{P_W}$ and the center of $\mathcal{A}_{Q_W}$, and the optical axis. In the focused case, both the center and aperture planes move and the view direction is constant, Figure 5.4a. However, in the afocal case, the plane P_W is static but not Q_W, so the view direction changes, see Figure 5.4b.

Angle of view The field of view is delimited by the cone of rays centered in $\mathbf{C}_{P_W}$ and bounded by $\mathcal{A}_{Q_W}$. Generally, in the same ECA, the magnification for planes P_W and Q_W is different causing the relative position of $\mathbf{C}_{P_W}$ and $\mathcal{A}_{Q_W}$ of two neigh-

Table 5.3: Analytical expressions for the optical properties of the ECA model of the focused and afocal light field cameras.

Focused	Afocal
Virtual DMU pitch D_p	
$-\dfrac{d_p}{\gamma-1}$	$-\dfrac{d_p}{\gamma-1}$
Virtual pixel pitch D_s	
$-\dfrac{d_s}{\beta\,\delta-(\beta-1)\,(\gamma-1)}$	$-\dfrac{d_s}{\delta}$
View direction	
$-\arctan\left(\dfrac{t_p-d_p\,\left(\frac{n_p}{2}-t+\frac{1}{2}\right)}{f_H}\right)$	$-\arctan\left(\dfrac{\delta\,t_p+(\gamma-1)\;D}{\delta\,f_H}\right)$
Angle of view	
$\arctan(A+B)$ $-\arctan(A-B)$	$\arctan\left(E-\dfrac{n_p\,d_p}{2\,f_H}\right)$ $-\arctan\left(E+\dfrac{n_p\,d_p}{2\,f_H}\right)$
Depth of field	
$-\dfrac{2\,\beta\,\delta\,d_p\,f_H\,d_s}{{d_s}^2(\gamma-1)^2-{d_p}^2(\beta\delta-(\beta-1)\,(\gamma-1))^2}$	$\dfrac{2\,\delta\,d_p\,f_H\,d_s}{{d_s}^2\,(\gamma-1)^2-\delta^2\,{d_p}^2}$
Baseline	
$-\dfrac{d_p\,\left((\gamma-1)\,\left(\alpha+\gamma-\frac{z}{f_H}+1\right)+1\right)}{\beta\,\delta\,(\gamma-1)\,\left(\alpha+\gamma-\frac{z}{f_H}+1\right)}$	$-\dfrac{n_p\,d_p\,\left(\alpha+\gamma-\frac{z}{f_H}+1\right)}{\left(\alpha-\frac{z}{f_H}\right)\,(\gamma-1)+\gamma^2}$
Transversal Accuracy	
$-\dfrac{d_s\,\left(\left(\alpha-\frac{z}{f_H}\right)\,(\gamma-1)+\gamma^2\right)}{\beta\,\delta}$	$d_p\,\left(\alpha+\gamma-\dfrac{z}{f_H}+1\right)$
Longitudinal Accuracy	
$\dfrac{2\,f_H\,d_s\,((\gamma-1)\;C+1)\;C}{{d_p}^2-{d_s}^2\,(\gamma-1)^2\,C^2}$	$\dfrac{2\,n_p\,f_H\,\left(\left(\alpha-\frac{z}{f_H}\right)\,(\gamma-1)+\gamma^2\right)\,C}{\left(\left(\alpha-\frac{z}{f_H}\right)\,(\gamma-1)+\gamma^2\right)^2-{n_p}^2}$
with $A=\dfrac{t_p-d_p\,\left(\frac{n_p}{2}-t+\frac{1}{2}\right)}{f_H}$, $B=\dfrac{d_p\,(\gamma-1)}{2\,\beta\,\delta\,f_H}$, $C=\left(\alpha+\gamma-\dfrac{z}{f_H}+1\right)$, $D=t_p+d_s\,\left(\dfrac{n_s}{2}-v+\dfrac{1}{2}\right)-d_p\,\left(\dfrac{n_p}{2}-t+\dfrac{1}{2}\right)$, $E=\dfrac{\delta\,t_p+(\gamma-1)\;D}{\delta\,f_H}$.	



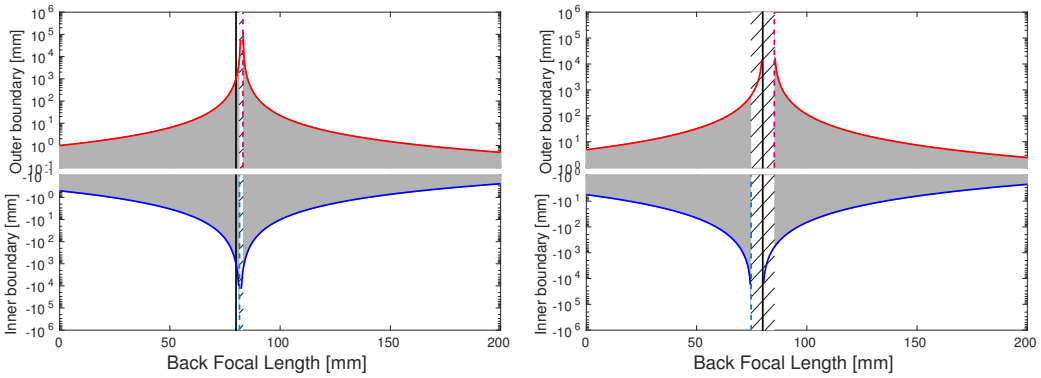
(a) Focused case (b) Afocal case

Figure 5.5: **Angle of view** of each of the virtual cameras of the ECA.

boring cameras to be different. Consequently, the angle of view of two neighboring cameras is slightly different. In the focused case, see Figure 5.5a, when the plane of the virtual cameras of the ECA, P_W, is imaged to infinity, the angle of view becomes zero when the back focal length is equal to the focal length of the main lens. In a classic configuration where the distance between the main lens and the MLA a bit larger than the focal length of the main lens, the angle of view remains low because the number of pixels of the virtual sensor is small. In the afocal case, see Figure 5.5b, the plane of the virtual cameras Q_W never goes to infinity. So, the angle of view hits the maximum value when the plane P_W goes to infinity. Moreover, the number of virtual pixels per camera is high and so is the angle of view.

Depth of Field In a classical camera, the depth of field is by definition located around the Q_W plane, which is also the plane of best focus. The same is true for the virtual cameras of the ECA. In addition, for light field cameras, the depth of field of the ECA cameras determines the range of synthetic refocusing. The ECA cameras all have the same depth of field as it depends solely on the pitch of the virtual apertures and the distance between their planes. As the virtual focus plane moves further away from the virtual lens plane, the depth of field grows larger till becoming infinite. This effect can be observed in Figure 5.6 when the back focal length approaches the main lens focal length. The asymptote position determines the hyperfocal distance of the system where the sharpness range in the image is the largest.

Baseline The baseline is a step function that only takes values that are integer multiples of the distance between two neighboring cameras. It is a positive function that is bounded by the maximum distance between the cameras of the ECA. The results are shown in Figure 5.7. For both the focused and afocal cases, the baseline is minimum at the virtual camera center plane position. However, the baseline is maximum at the position of the front focal plane of the main lens in the focused case and at the virtual sensor plane position in the afocal case. The important region is the



(a) Focused case (b) Afocal case

Figure 5.6: **Depth of field**. The blue and red curve represent the distance of the boundaries of the depth of field from the virtual sensor plane. The depth of field is common for all the virtual cameras. The dashed blue and red vertical lines indicate the hyperfocal distances of the camera. The hashed area between these positions is the area where the depth of field is infinite.

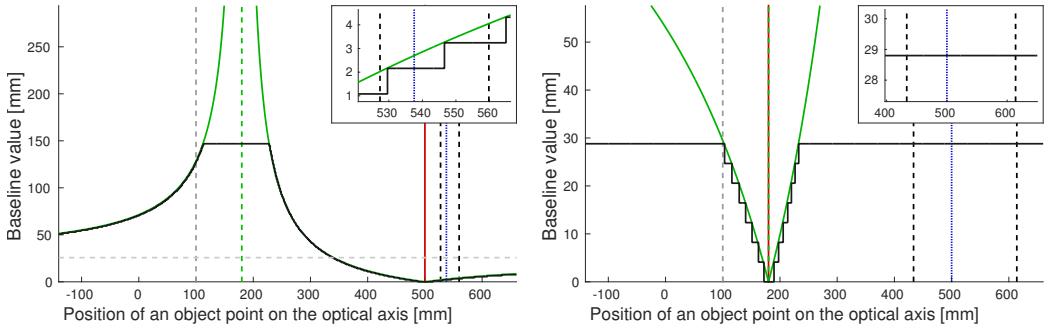
one situated between the depth of field limits. In the focused case, see Figure 5.7a, the baseline per camera is low because the overlap between neighboring cameras is reduced as the zero-disparity plane is behind the cameras. In the afocal case, see Figure 5.7b, the baseline is at its largest on the full depth of field range as a result of the zero-disparity plane being at the virtual sensor plane.

Accuracy The results are shown in Figure 5.8. The transversal measure of the accuracy is linear with the absolute distance of the evaluation point to the plane of the virtual camera center. The longitudinal measure of the accuracy is a more complex curve.

Figure 5.9 regroups the plots of the baseline and accuracies for the field around the optical system.

5.1.5 Notes on the Focused/Afocal Comparison

The two configurations provide two distinct solutions for the spatio-angular resolution trade-off linked to the arrangement of the virtual cameras. From the baseline and angle of view properties, the ECA of the focused configuration is made of many cameras with a small angle of view, each looking at a different location of the focus plane. On the contrary, the ECA in the afocal configuration is very similar to a physical array of cameras. Each camera has a large angle of view and they all observe the same part of the focus plane. The focused configuration has a lower number of cameras seeing a common region of space and so its angular resolution is less than



(a) Focused case (b) Afocal case

Figure 5.7: **Baseline** for an evaluation point on the optical axis for a Back Focal Length of $100mm$. The theoretical baseline (green curve) assumes a continuity of infinitely many cameras in the ECA. The real baseline (black curve) is a step curve computed from the actual position of the ECA cameras. The possible values for the baseline (the distances between the centers of the virtual cameras) are discrete, so this curve is a step curve that has a maximum equaling the distance between the two furthest cameras of the ECA. The horizontal gray dashed line indicates the theoretical value of the baseline for an evaluation point at infinity. The vertical lines indicate positions of interest such as the main lens plane (in gray), its front focal plane (in green), virtual camera plane (in red), the virtual sensor plane (in blue) and the depth of field boundaries (in black).

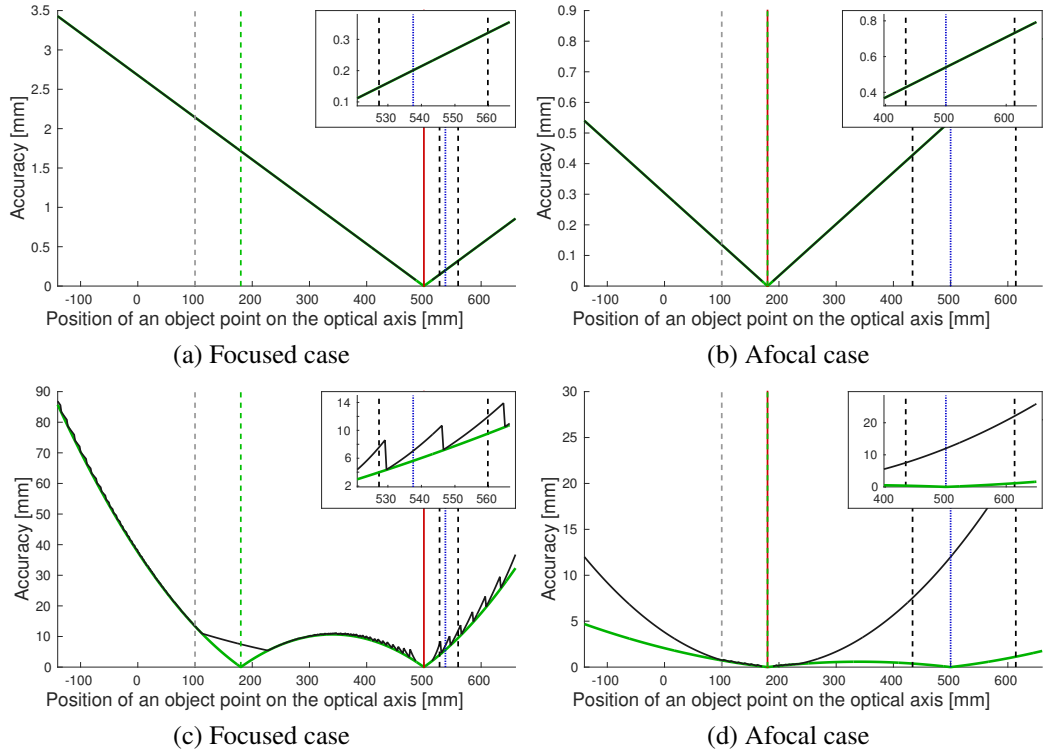
for the afocal configuration. This distinction was described as an improvement to the afocal case to retrieve lost spatial resolution [Lumsdaine and Georgiev, 2009].

The previous simulation and analysis did not take into account the influence of the vignetting effect, where some finite apertures in the system are blocking the theoretical path of light rays, resulting in cropped two-aperture elements. Vignetting reduces the spatial and angular resolution of the system so the depth of field and the angle of view of outer cameras in the ECA are most affected.

5.2 Validation

We validate our model through experimentation by obtaining the properties of the ECA for a real light field camera. As a result, we can estimate some unknown physical specifications of the system. In contrast to our first-order model, we observe effects of non-linearities in the real data which we point out in the discussion of our results.

Our selected light field camera is the first generation Lytro [Lytro, 2017], which can be categorized under the afocal type systems similar to Figure 4.3c. We use this light field camera to create a data set of correspondences between world rays and sensor pixels that are later used to construct a generalized imaging model as



(a) Focused case

(b) Afocal case

(c) Focused case

(d) Afocal case

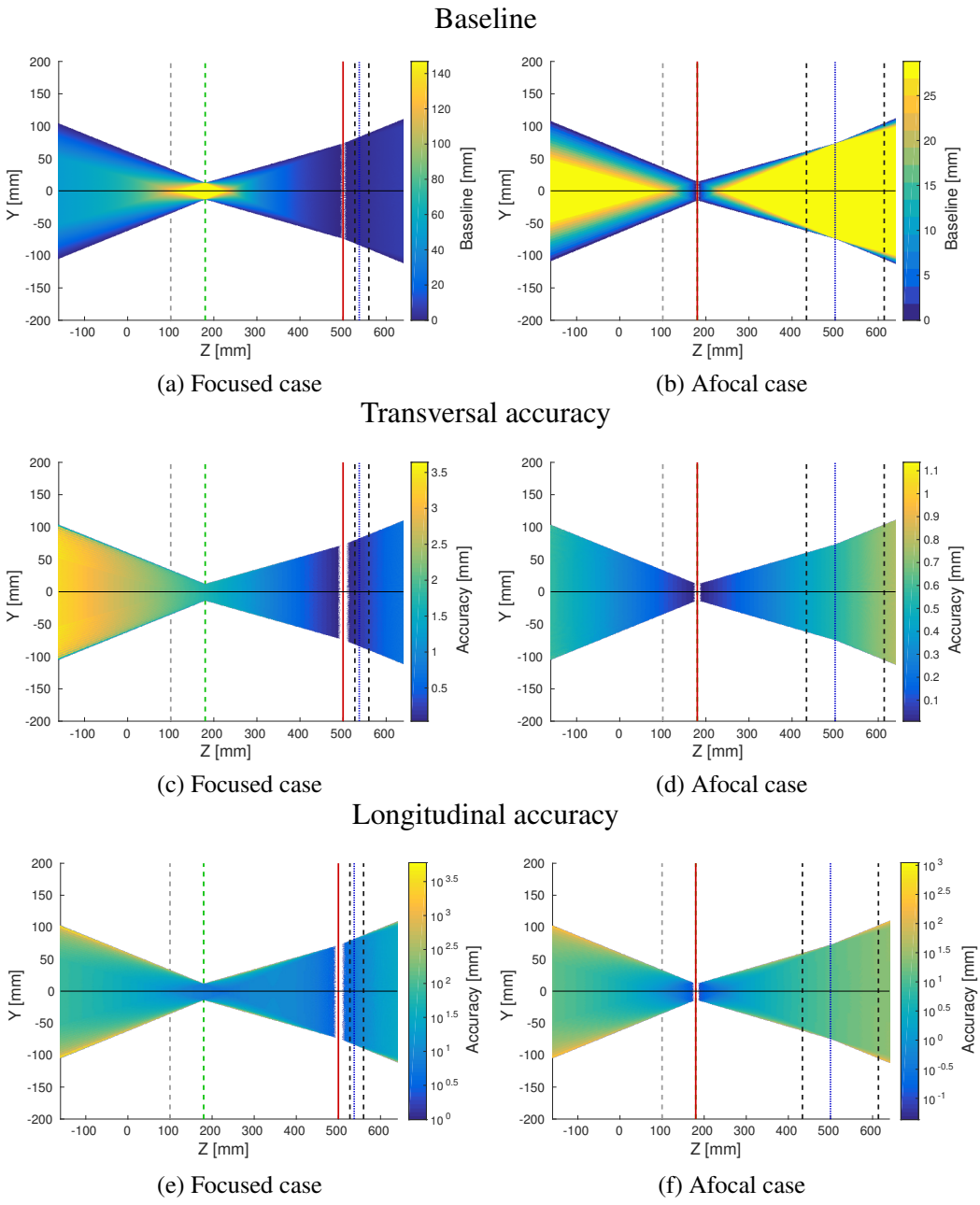
Figure 5.8: **Transversal** (top row) **and longitudinal** (bottom row) **accuracies** for an evaluation point on the optical axis for a Back Focal Length of $100mm$. The real curves (in black) have undefined values for positions where the baseline is either zero or undefined (close to the virtual camera center plane in red). The theoretical longitudinal accuracy (in green) and the real accuracy (in black) differ for positions where the baseline is clamped. Overall, the discontinuous behavior of the real curve is due to the discrete change of the baseline value. The vertical lines have the same definitions as in Figure 5.7.

described in [Grossberg and Nayar, 2005]. Given the interpretation of a light field as an array of subview images of the scene, we obtain a generalized imaging model per subview that is directly analogous to our proposed ECA.

The experiment was realized with the help of John Restrepo. The estimation of the properties from the measured data is the work of John Restrepo only, and the fit of the properties with the ECA model is the work of the author of this dissertation.

5.2.1 Experimental Setup

Our experimental setup is displayed in Figure 5.10a. We modify the camera by removing the main lens from its encasing. This way, we can control its distance to the sensor and microlens array. The Lytro camera provides optical refocus and zoom. For simplification purposes, we fix the main lens settings, keeping its optical

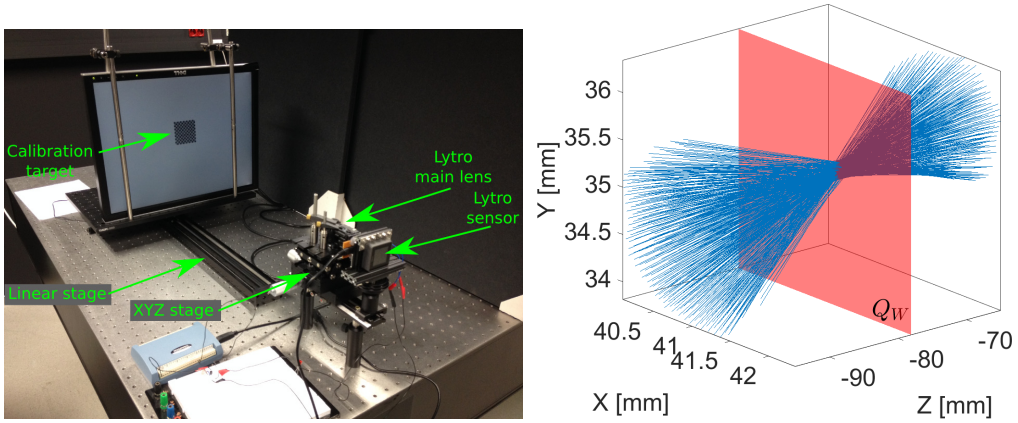


(a) Focused case (b) Afocal case

(c) Focused case (d) Afocal case

(e) Focused case (f) Afocal case

Figure 5.9: Results for the baseline and accuracies on the field. The BFL is $100mm$.



(a) (b)

Figure 5.10: (a) Photograph of the experimental setup. (b) Ray bundle for a single subview with a reduced number of rays for visualization. The intersecting plane identifies the location of the center of perspective.

properties constant.

The separated lens and light field sensor (microlens array-sensor couple) are set up independently. Given our mechanical conditions, we cannot guarantee an optimal alignment as compared to the original camera. As a consequence, our data shows some irregularities that we will discuss in the following sub-section. We measure world space rays by recording multiple positions of a calibration target [Atcheson *et al.*, 2010] displayed on a computer screen for a series of known distances.

5.2.2 Estimation of Properties

The pre-processing of the acquired data involves decoding the light fields [Dansereau *et al.*, 2013b], detecting the corners of the calibration target [Atcheson *et al.*, 2010], and upsampling the resulting ray data set to obtain a corresponding ray for every pixel in each subview. In order to construct property plots such as Figure 5.4, we repeat the experiment for multiple displacements of the main lens ΔZ, which produces property measurements for different back focal length values.

An example of the ray data is shown in Figure 5.10b (right), for a single subview and main lens position. From the ray distribution, we can conclude that the imaging is effectively a perspective transformation.

We equate the locus point of each ray bundle with the center of perspective for each subview; these points should agree with our ECA positions in order to validate our model. The obtained centers for all subviews are plotted in Figure 5.11a, the plane containing the centers is the aperture plane Q_W of our model. We ignore the

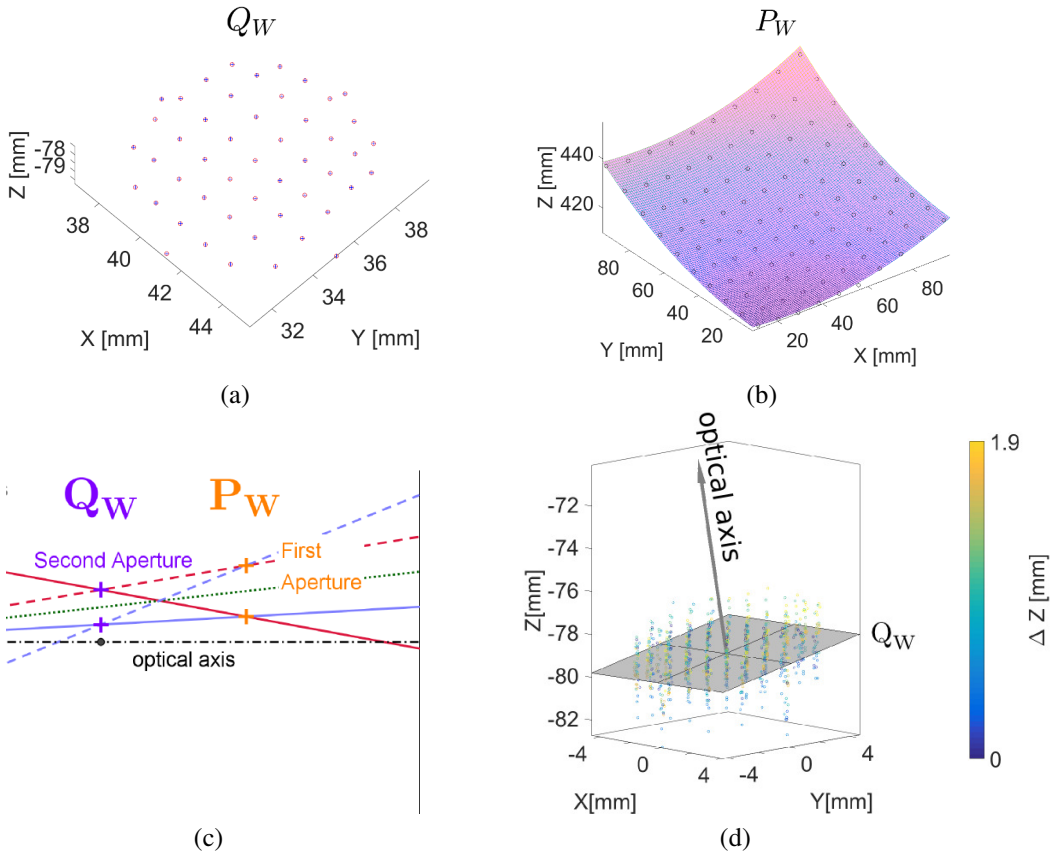


Figure 5.11: Planes of the ECA model. (a) Centers of perspective for a single main lens position. (b) Best focus plane, with the markers as the intersections of rays and a surface as spherical fit. (c) Sketch for an afocal setting. (d) Plot of all centers for all lens positions and computed optical axis.

outer-most centers because of low contrast. The array of centers indicates regularity with the exception of the extreme subviews. This validates the ECA model for the central paraxial region but also indicates the presence of aberrations for the outer subviews.

Our ECA model predicts that the array of centers is located on a plane perpendicular to the optical axis as presented in the sketch in Figure 5.11c. However, our modification of the Lytro camera creates a misalignment between the axis of the lens and the microlens array, which effectively produces a Scheimpflug effect, tilting the imaging planes of the main lens. We do not account for these effects in our model and instead use our experimental data to obtain a new optical axis. We compute a unique optical axis for all lens positions ΔZ. This new optical axis is defined as the normal of the least squares fitted plane to all estimated centers of perspective. The result is shown in Figure 5.11d. The optical axis obtained with this procedure is used to obtain the experimental properties of the subviews.

Plane Positions

We compute the perspective centers, again, for multiple positions of the main lens represented by the distance ΔZ of the main lens with respect to its initial position. Furthermore, to cancel variations between different experiments for different main lens positions, that are due to mechanical influences, we re-center all cameras to the central one. This implies that we cannot measure the central camera, only the separation inside the array. The results are displayed in Figures 5.13b, 5.13b, and 5.13c. In order to match the format of plots employed in the previous sections, we restrict ourselves to the central column of the array of subviews.

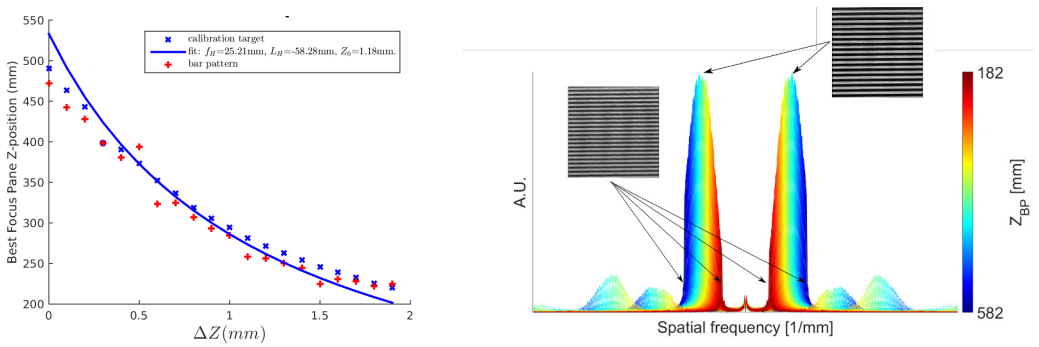
View Direction

Similar to its definition for our ECA model, the view direction is measured as the angle between the central ray of each subview and the optical axis of the ECA. For our experiment, we use the central ray of the bundle as shown in Figure 5.10b and the normal from the fitted plane to all centers as in Figure 5.11d. The view direction is the angle between these two vectors. For the same column of cameras as before, the experimental view direction is displayed in Figure 5.13d.

Best Focus Plane Distance

This plane corresponds to the plane P_W in our ECA model. It is simultaneously the plane of zero disparity between subviews. Therefore, each corresponding ray from the same pixel in all subviews must intersect in this plane. A computed example, for a smaller sample of rays, is displayed in Figure 5.11b. We observe that, in fact, we do not obtain a plane but a curved surface. This clearly indicates the presence of optical aberrations that are global to all subviews, in particular, a field curvature of the main lens. We compute the axial location of this surface as the average of the Z coordinates for all points in it. The resulting plot is displayed in Figure 5.12a.

We further support our experimental findings by measuring the best focus plane location with an alternative method. We use the same experimental setup, now displaying a binary bar target for multiple positions of the computer screen Z_{BP}. Treating each subview independently, we use a metric for contrast to establish the best focus plane location and, subsequently, averaging for all subviews. Our metric for contrast is based on the Fourier spectrum of the bar target. We take an average of several line profiles of each image and compute the $2D$ spectrum. In it, we measure the height of the secondary spectral peaks as a function of the bar target position. The highest contrast corresponds to the maximum peak. The spectra for multiple bar pattern profiles at different distances are displayed in Figure 5.12b. There are two subviews with high and low contrast to exemplify the peak heights to which they correspond. The frequency location of the peaks varies with the position of the target Z_{BP} due to the magnification effect of changing the object position.



(a) (b)

Figure 5.12: Estimation of the best focus plane location. (a) Best focus plane location for the calibration target (blue markers) corresponding fit (solid line) and for comparison, the bar pattern results (red markers). (b) Superposition of the spectra at multiple target positions (with DC term removed). There are two subviews of the bar pattern indicating the positions for highest and lowest contrast. Colors indicate different target positions.

We compare the results of these two strategies for the axial position of the best focus plane in Figure 5.12a. The agreement between the data points is a good indicator of the accuracy of our experiments.

5.2.3 Properties Fit

We use all the accumulated experimental data to validate our ECA model. As it is physically challenging to measure the position of the main lens with respect to the sensor accurately, we do not have information to forward-simulate our ECA model generating identical plots to superimpose with the experimental results. Instead, we perform a fit of the experimentally attained properties to their corresponding analytical equations. We first show what the unknown system parameters are and what the formulas of the camera properties becomes for the fit. Then, we perform a global fit of all the properties to get an averaged result for the system parameters.

Fit Model for the Properties

In our experiment, the absolute position of the lens z_H is unknown and for each data point, the lens is translated along the z-axis by the amount ΔZ from its original position where $\Delta Z = 0$. So we have, $z_H = \Delta Z + f_H + Z_0$, where f_H, the focal length, and Z_0, the constant separation between the sensor and the lens, are both unknown. Moreover, we use a different indexing for the cameras:

$$d_s v' = d_s(v - \frac{n_s + 1}{2}) - d_p(t - \frac{n_p + 1}{2}) \tag{5.16}$$

where v' is an integer varying between -3 and 3 indicating the camera index. The ratio $\frac{d_p}{d_s}$ is equal to 10 for the Lytro sensor.

Consequently, the unknown parameters of the system are:

- the focal length of the main lens: f_H,
- the first-order thickness of the main lens: L_H,
- the constant separation between the sensor and the main lens: Z_0, and
- the shift between the sensor center and the optical axis: t_p.

The properties of the ECA obtained from the experiment are the position of the virtual cameras along the x-, y-, and z-axis ($CamXYZ$), their view direction (VD), and the axial location of the best focus plane (BF) .

The equations listed in Table 5.3 are valid for a 2D case so we consider fitting only the properties of a line of cameras in the array.

Cameras x and y positions: The upper term of the vector in Equation 5.15 gives the x or y position of the ECA cameras depending on the chosen line (along the x- or y-axis) of cameras:

$$CamXY = f_H \frac{t_p + d_s\left(\frac{n_s}{2} - v + \frac{1}{2}\right) - d_p\left(\frac{n_p}{2} - t + \frac{1}{2}\right)}{f_p} \tag{5.17a}$$

With the previous reindexing from Equation 5.16, the previous equation becomes:

$$CamXY = \frac{f_H(t_p - d_s v')}{f_p} \tag{5.17b}$$

Cameras z position: The lower term of the vector in Equation 5.15 gives the z position of the cameras:

$$CamZ = L_H + z_H + f_H = L_H + \Delta Z + Z_0 + 2f_H \tag{5.18}$$

but we measure $CamZ$ from the first surface of the lens, so it becomes independent of z_H and consequently independent of ΔZ.

Best focus plane location: This property is obtained from the second element of the vector in Equation 5.11, it is rewritten as:

$$BF = f_H\left(\alpha + \frac{\gamma^2}{\gamma - 1}\right) = \frac{z_H^2 + L_H(z_H - f_H)}{z_H - f_H} \tag{5.19a}$$

Replacing z_H by its relation to the experimental measurement, we obtain:

$$BF = \frac{(\Delta Z + f_H + Z_0)^2 + (\Delta Z + Z_0)L_H}{(\Delta Z + Z_0)} \tag{5.19b}$$

View direction The view direction formula is rewritten from its equation in Table 5.3 as:

$$VD = -\arctan\left(\frac{t_p}{f_H} + \frac{(z_H - f_H)\left(t_p + d_s\left(\frac{n_s}{2} - v + \frac{1}{2}\right) - d_p\left(\frac{n_p}{2} - t + \frac{1}{2}\right)\right)}{f_H f_p}\right) \tag{5.20a}$$

Replacing the numerator with more appropriate variables gives:

$$VD = -\arctan\left(\frac{t_p f_p + (\Delta Z + Z_0)(t_p - d_s v')}{f_H f_p}\right) \tag{5.20b}$$

Independently fitting each equation to the experimental data yields the results indicated in Figure 5.13.

Global Fit

The system parameters are estimated from the individually fitted properties. However, the fits are independent so the values of the parameters from the different properties are not in exact accordance. Since the properties are derived from the same system, a simultaneous fit would output a single value for each parameter and distribute the error among all the properties.

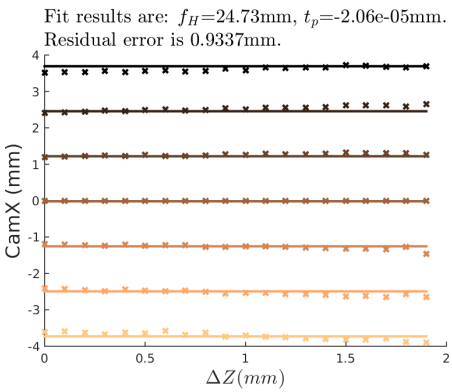
The global fit is a least square minimization problem on the system parameters $\mathbf{p}$:

$$\min_{\mathbf{p}}\left(\sum_i \alpha_i g_i(\mathbf{p})^2\right), \tag{5.21}$$

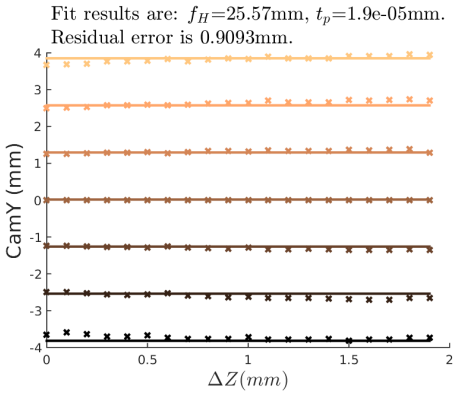
with $g_i(\mathbf{p}) = \mathit{Data}_i - \mathit{Model}_i(\mathbf{p})$ the error function for the property i, and α_i being a weighting coefficient. The properties $\mathit{Model}_i(\mathbf{p})$ used in this global fit are the ECA camera position in X, Y and Z, the view direction and the Z position of the best focus plane. The results for the fit are shown together with the experimental data in Figures 5.14 and 5.12a.

The output coefficients $\mathbf{p}$ of the global fit are the parameters of the main lens of the system: the focal length f_H (25.21mm), the thickness L_H (-58.3mm), its absolute distance to the sensor $f_H + Z_0$ (26.4mm from the back principal plane for $\Delta Z = 0$) and the alignment shift between the sensor and the microlens array t_p. However, this last parameter is lost in the fit due to the re-centering of the data for the view direction and the ECA camera position in X and Y. We use three independent pseudo parameters, one for each property to represent it instead.

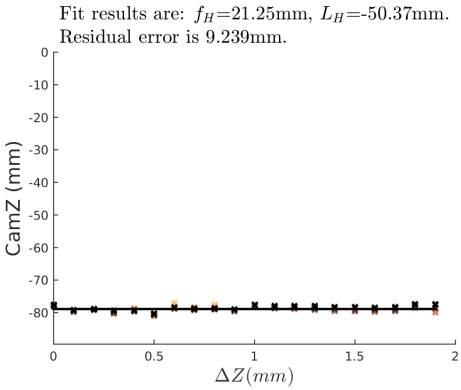
To culminate our validation, we perform a characterization of the main lens, employing a Shack-Hartmann sensor [Platt and Shack, 2001]. This experiment delivers a focal length f_H for the main lens of $23 \pm 2mm$, which provides a reasonable error with respect to our estimation from the fit. Measuring the absolute smallest distance of the main lens to the sensor $f_H + Z_0$ (*i.e.,* at $\Delta Z = 0$) is a challenging



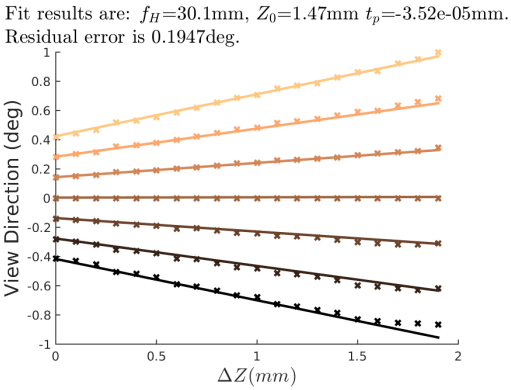
(a) Horizontal position of the virtual cameras



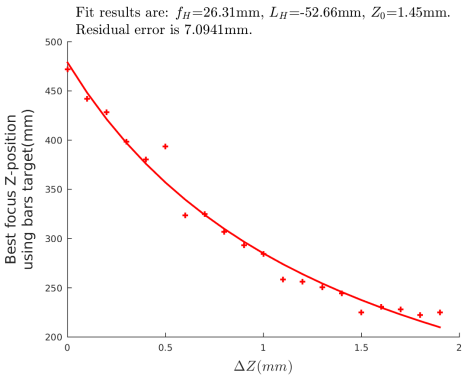
(b) Vertical position of the virtual cameras



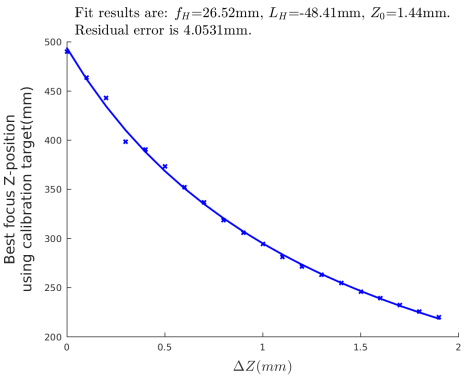
(c) Position of the virtual cameras on the z-axis



(d) View direction of the virtual cameras

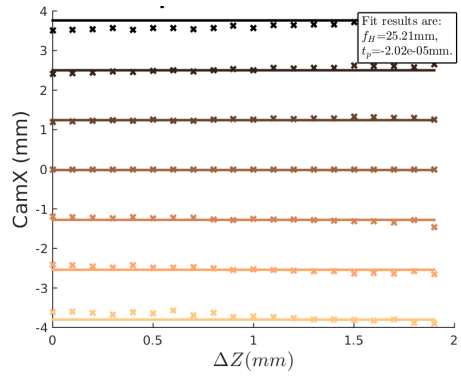


(e) Best focus using bar target

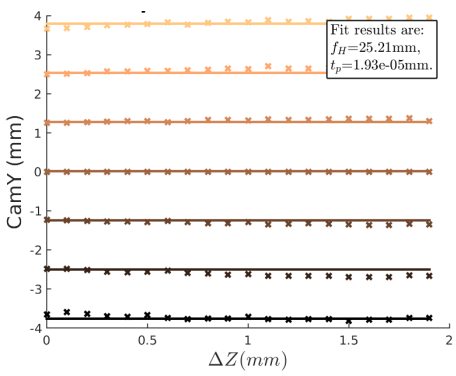


(f) Best focus using calibration target

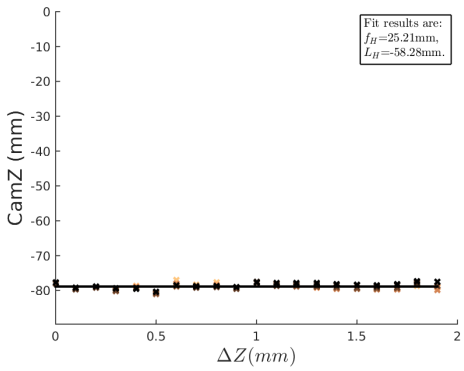
Figure 5.13: Results of the independent fit of the properties. Markers indicates the measured data points and the solid lines indicate the fit curve of the property. The central camera is excluded from the fit because it was used to re-center the other ones. ΔZ indicates the main lens displacement with respect to the microlens array.



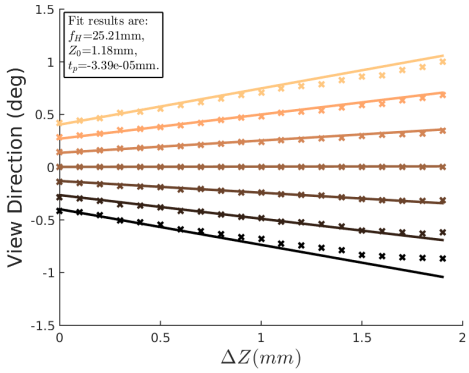
(a) Horizontal position of the virtual cameras



(b) Vertical position of the virtual cameras



(c) Position of the virtual cameras on the z-axis



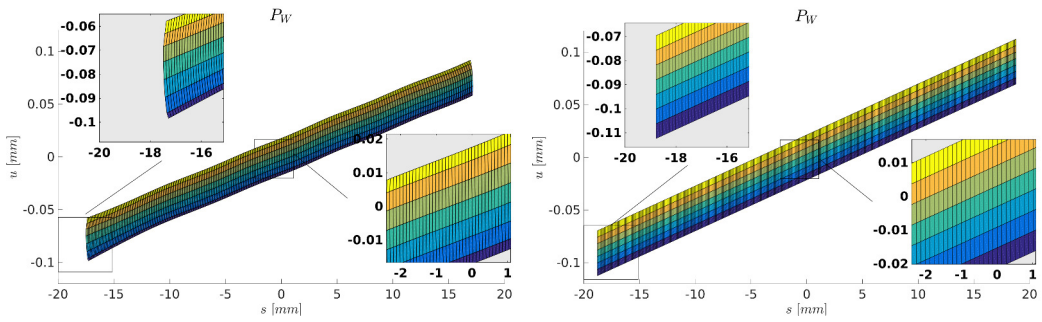
(d) Vertical position of the virtual cameras

Figure 5.14: Results of the global fit of the properties. The results of the global fit on the best focus plane location are shown in Figure 5.12a.

experimental endeavor. From our experiment, we can *qualitatively* confirm that the distance obtained from the fit corresponds to the experimental setup because we purposely place the main lens at the physically closest distance to the sensor. The last pseudo parameter t_p has a different value for each property fit and this value is close to zero because of the re-centering of the data.

5.2.4 Discussion

Aberrated rays correspond to a deformation of the phase space as displayed in the phase space from the experiment plotted in Figure 5.15a. In contrast to the simulated phase space, see 5.15b, the cells are no longer aligned in the plane P_W, as the aberrations redirect the ray bundle, distorting the phase space. This fact is more prominent on the edges of the subviews, corresponding to the left and right sides of Figure 5.15a. The simulated phase space is only provided by the parameters obtained from the fit, as again, we lack some physical specifications to accurately model the real light field camera. A future higher order model may be able to ade-



(a) (b)

Figure 5.15: (a) Experimental phase space at the plane P_W . Inaccuracies of the ray estimation, stemming from corner detection and ray fitting, provoke the undulations observed in the light field cells. (b) Corresponding approximate simulated phase space.

quately describe the deformations of the phase space and correlate them with optical aberrations.

5.3 Conclusion

We have applied the ECA model to microlens-based light field cameras. In particular, for a 2D setup made of a sensor, a microlens array and a single main thick lens, we have derived analytic expressions for each property of its equivalent camera array. These expressions are parameterized by the focal length of the main lens. Camera properties can be computed for the distances comprised by the scene volume, which makes them a useful tool for optical design. Moreover, we have discussed the similarities and differences between the focused and afocal configurations of the system.

We have also validated our model through experimentation with a real light field camera. In the absence of a genuine ground truth, we fitted our model to the experimental data, extracting system parameters that were checked by alternative methods. This validates our model for the selected light field camera. A similar procedure can be used without any difficulty for other camera architectures.

Chapter 6

Application to Light Field Microscopy

The light field microscope has been introduced and improved by [Levoy *et al.*, 2006, 2009; Broxton *et al.*, 2013]. Its practical implementation relies on the fabrication of a custom microlens array, which presents a hurdle for experimenting with the technology. However, the applications are numerous and propose a new paradigm in the field of microscopy.

This chapter explores the design constraints in the construction of light field microscopes. It is divided into two parts. The first part considers the making of inexpensive and accessible means of exploring light field microscopy with good quality, albeit at a reduced optical magnification. The design constraints are governed by the vignetting effects that rise through the association of different optical components. In the second part, a light field microscope, the design of which is close to the original work from [Levoy *et al.*, 2006], is built in the laboratory. The important aspects of its design are analyzed and discussed through the scope of the ECA model.

As stated in the introduction, the work and results of the first part of this chapter are prior to the ECA model. Actually, this work revealed the need for a more abstract tool for light field camera design and motivated the creation of the ECA model as it is presented in the previous chapters. Consequently, the ECA model is not restricted to light field cameras only and can also be used in the context of light field microscopy as this was its original purpose. After the model was created and sufficiently developed, it was decided not to apply the ECA model to the optical systems presented in the first part of this chapter as they had already shown their limited range of applications. Instead, the ECA model was applied to the light field microscope of the second part.

6.1 Light Field Microscopy with the Lytro Camera

This section explores the use of commercial light field cameras for microscopic imaging applications. For this purpose, we associate two different imaging systems: the Lytro camera, an inexpensive consumer-grade light field sensor, and a microscope objective. Each part is an optical system with its own specificities so we study the limits in which they can be used together.

As we show, the major problem in combining the Lytro and additional magnification optics (in addition to f-number matching), is the loss of information due to spatial vignetting. Our main finding is the possibility of using the Lytro in what we term an *inverse regime*: in this setting the camera picks up a virtual object that is located far behind its imaging optics. To our knowledge, this is the first time that such a light field imaging mode is described.

We investigate two different setups based on this inverse regime and we achieve an inexpensive and accessible means of exploring light field microscopy with good quality, albeit at a reduced optical magnification.

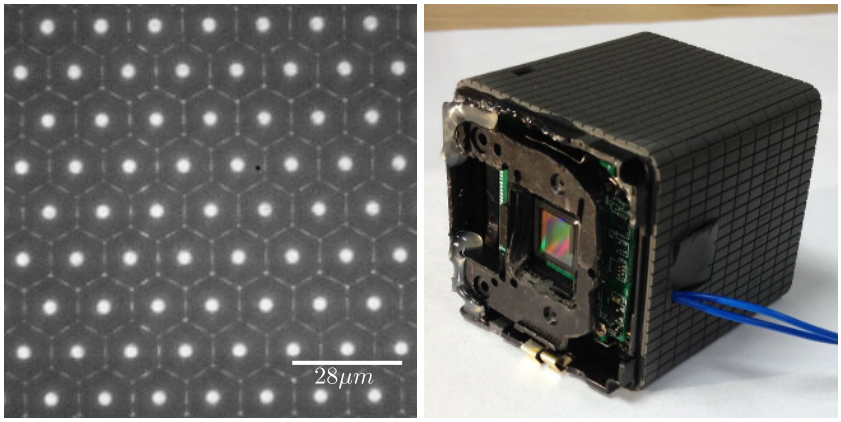
6.1.1 Lytro Features

The Lytro camera is made of an optical system that is forming an image in the plane of a microlens array that is, in turn, redirecting the light rays to a sensor. It has a 3280×3280 pixel CMOS sensor with 12-bit A/D and $1.4\mu m \times 1.4\mu m$ pixels [Kučera, 2014], see Figure 6.1b. Each microlens has a diameter of $14\mu m$ which is equivalent to 10 pixels. The microlenses are packed on a hexagonal lattice, see Figure 6.1a. The effective spatial resolution is, therefore, 328×328 pixels whereas the angular sampling rate is 10×10 values. The Lytro main objective lens has a fixed f-number of $F = 2$ and features an $8\times$ optical zoom. We explore its potential as an imaging parameter in Section 6.1.4. Another important feature of the objective lens is that it can focus from $0mm$ to infinity.

6.1.2 F-number Mismatch

A prerequisite for non-vignetted imaging (see also Section 6.1.4) is that the f-number of the microlens array and that of the microscope objective match. This is the solution employed in conventional light field microscopy [Levoy *et al.*, 2006; Oldenbourg, 2008; Levoy *et al.*, 2009]. A custom $F = 20$ microlens array (MLA) is typically employed as it is compatible with a large number of existing microscope objectives. However, this $F = 20$ MLA is not readily available and has to be custom manufactured.

The Lytro is designed for macroscopic imaging, hence, its f-number ($F = 2$) is not adapted to the microscopic situation. If we were to use the Lytro microlens array as is, the angular sampling in one direction would be divided by the ratio of the f-number of the two systems and only one pixel would be lit under each



(a) (b)

Figure 6.1: Lytro light field sensor. (a) Picture of the microlens array taken with a microscope with top illumination. We can see the hexagonal structure of the microlens array. The bright dot in the middle of each hexagon is an image of the light source reflected by the surface of the microlens. (b) Picture of the Lytro camera without its optics.

microlens (instead of approximately one hundred), see Figure 6.3 (bottom). This would remove any interest for light field purposes as only a single view would be recorded and 99% of the sensor would remain unused, see Section 6.1.4 for examples.

Therefore, in order to use the Lytro camera for microscopic imaging successfully, the f-numbers of the two optical systems need to be adapted. There is, however, a trade-off due to the Helmholtz-Lagrange invariant, see Equation 1.12. The object-side and image-side numerical apertures are linked to the magnification.

Consequently, we opt for an optical demagnification scheme to increase the numerical aperture of the cone of light rays that is incident on the Lytro's light field sensor. Theoretically, we need to divide the microscope objective's f-number by 10 to reach the same f-number as the Lytro camera. An immediate consequence from Equation 1.27 is that the combination of all optical elements must, therefore, have a magnification divided by 10, *i.e.,* we are aiming to convert the system to unit-magnification. Due to the small size of the Lytro's microlenses, the optical resolution of the system is still satisfactory, even at this low magnification (see Section 6.1.6).

The magnification of the combined system M_{final} can be written as the product of the magnification of each individual system:

$$M_{final} = M_{micro} M_{lens_1} ... \, M_{lens_N} M_{Lytro}, \tag{6.1}$$

where $M_{lens_i,\, i=1..N}$ indicates the magnification of N to-be-designed intermediate lens systems. The microscope objective has a fixed magnification of 10, whereas the

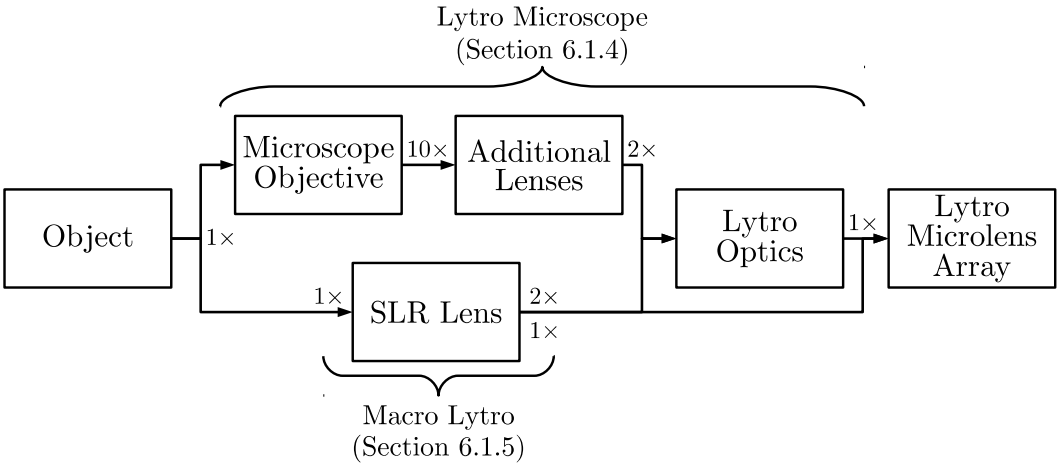


Figure 6.2: The two proposed solutions to achieve low magnification. The first option (first row) keeps the camera and the microscope intact. The second option (second row) replaces the microscope with an SLR objective.

lowest magnification setting of the Lytro has a value of $M_{Lytro} = 0.5$. The resulting magnification is $M_{final} = 5$ without additional optical components ($N = 0$). This is too large to prevent angular information loss.

We explore two different options to implement the adapted system, see Figure 6.2. The first option (see Section 6.1.4) is to demagnify the image of the microscope with additional lenses ($N = 2$). This solution lets us use the microscope and the Lytro camera unmodified. The second option is to remove the microscope, replacing it with a single-lens reflex (SLR) objective in macro-imaging mode. Here, we compare a setup with and without the Lytro optics (see Section 6.1.5).

6.1.3 Vignetting

Before getting interested in the actual setups, we need to investigate the vignetting effects present in the system and study the settings of the Lytro camera to overcome the vignetting issues.

Sensor Coverage

Generally, there are two types of mechanical vignetting: spatial and angular vignetting. Spatial vignetting directly translates into a loss of field of view, which may reduce the image size at the sensor, see Figure 6.3 (top). Angular vignetting occurs when the size of the pupil in the system is reduced voluntarily (with a controlled diaphragm) or involuntarily, see Figure 6.3 (bottom). Angular vignetting is not an issue in a standard camera: it only affects the exposure and is directly linked to the depth of field of the camera. In a light field camera, however, it is crucial

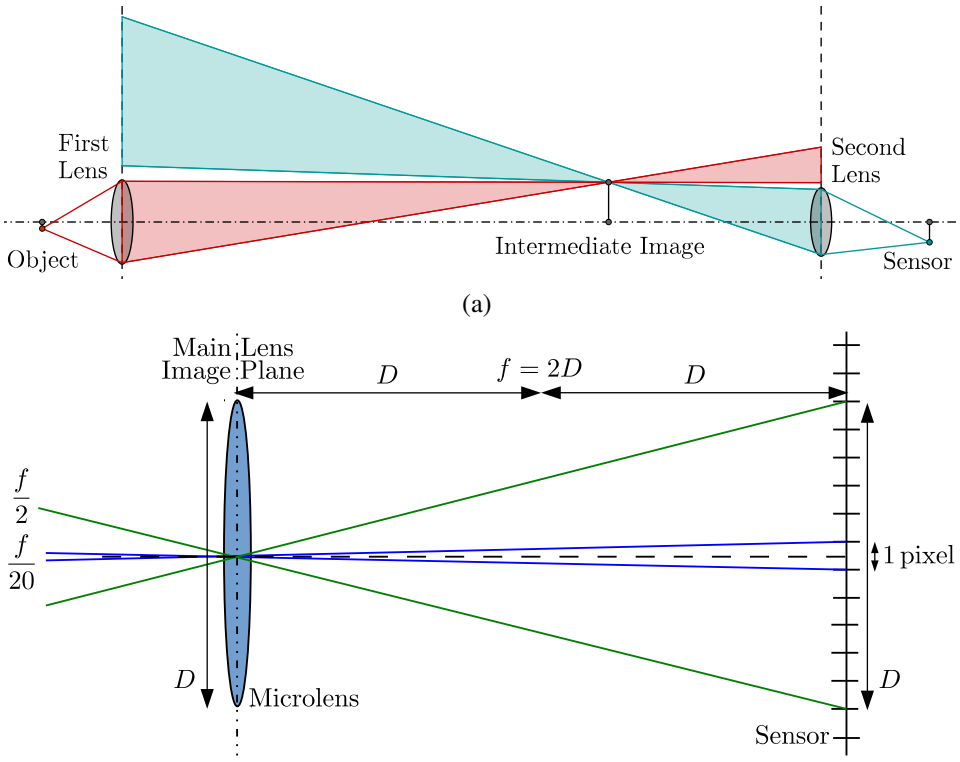


Figure 6.3: Spatial and angular vignetting. (a) Spatial vignetting occurs when light rays from the object (in red) do not pass through the second lens. For non-vignetted imaging, light rays (in blue) converging to a point at the sensor edge should include all the red light rays emerging from the object. (b) Representation of the sampling of a cone of light rays by a $F = 2$ microlens. Green rays symbolize the angular cone that can be acquired by the microlens, while blue rays emerge from an $F = 20$ optical system. Angular vignetting prevents an effective sensor utilization.

to minimize angular vignetting in order to prevent the loss of directional light field information.

We propose to measure the vignetting in terms of its adaptation to the recording light field sensor. An ideal optical system that is adapted to a particular sensor would fully cover all its sensor elements. The raw pixel resolution is divided into spatial and angular parts, consequently, the sensor coverage c_{sensor} can be approximately expressed as

$$c_{sensor}[\%] = c_{spatial}[\%] \times c_{angular}[\%], \tag{6.2}$$

where $c_{spatial}$ is the spatial coverage of the sensor in percent, and $c_{angular}$ is the angular coverage of one microlens micro-image, also in percent. We define the spatial coverage as the ratio of the number of the microlenses that receives any light to the total number of microlenses. And, we define the angular coverage as the ratio of the number of pixels of the vignetted micro-image to the number of pixels

of the micro-image obtained when the f-number matching rule is satisfied. In our experimental validation, we measure the spatial coverage in the center view and the angular coverage in the center lenslet. This choice is motivated by the simpler estimation of the relevant coverage areas as compared to using the side views/edges of the field. The measure can be considered to be an approximation of the upper bound of the system space-bandwidth product, *i.e.,* the optical information capacity of the system [Lohmann *et al.*, 1996].

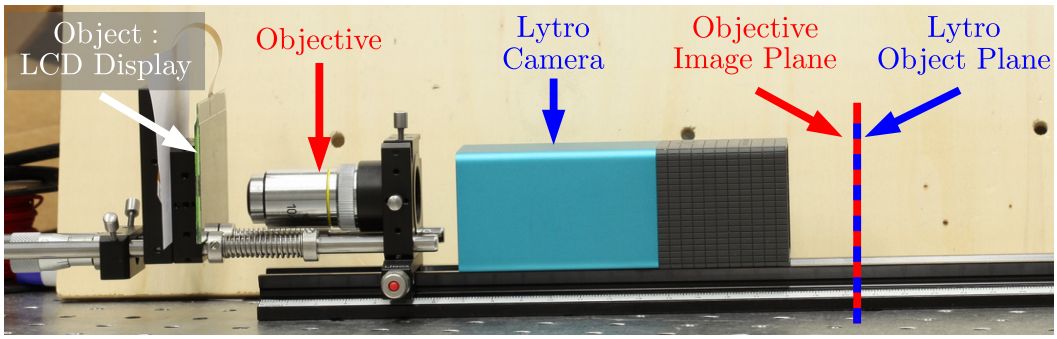
Unmodified Use of the Lytro's Main Optics

The main optics of the Lytro camera, *i.e.,* the optics without the microlens array, is designed to avoid angular vignetting when imaging onto the microlens array, *i.e.,* the microlens array and the main optics have been designed with the same f-number of $F = 2$. We have observed that the main optical system can be used in two different ways with a microscope. These two imaging regimes can be used differently in designing an optical matching system.

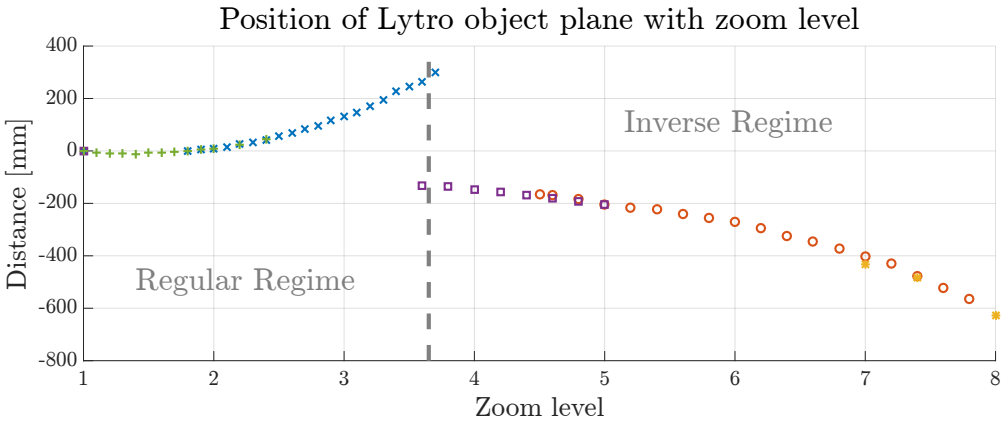
Regular regime The Lytro camera can image a plane as close as the first surface of its optics for a zoom level of $1\times$. This minimal focus distance increases with the zoom level. In order to use the Lytro with the microscope, it has to be positioned such that the near focus of the camera is placed at the image plane of the microscope objective. The image size of the microscope objective is rather large (typically $50mm \times 50mm$) whereas the Lytro's entrance pupil is only $20mm$ in diameter, approximately. Therefore, spatial vignetting incurs a loss of sensor coverage of up to 94% as shown in Figure 6.5 (left). The angular vignetting is stronger with only 16% angular coverage.

Inverse regime We discovered that, in a specific configuration where the camera is set to focus to the closest possible plane for a zoom level of $1\times$, the camera can enter into a virtual object regime. The camera is then able to image an object plane that is located *behind* its first lens, *i.e.,* in the direction of the sensor, see Figure 6.4. This configuration enables the positioning of the camera close to the microscope objective and therefore reduces the spatial vignetting as a larger number of rays can be captured by the lens surface, see Figure 6.5 (right). This mode of operation inverts the image.

For both imaging regimes described above, the magnification M_{Lytro} (≈ 0.5) is not low enough for achieving a good angular coverage. While the spatial vignetting problem can be successfully addressed with the inverse regime, the angular vignetting can only be dealt with by using a low magnification optical system. We, therefore, investigate the different options.



(a)



(b)

Figure 6.4: Inverse regime configuration. (a) Picture of the setup. (b) Plot of the distance of the Lytro focus plane to its front lens with the variation of zoom. An abrupt change of position of the zoom lens group occurs for a zoom level of $3.7\times$, enabling to switch between the regular and inverse regime.

6.1.4 The Lytro Microscope

Our first option to achieve the matching of the f-numbers discussed in Section 6.1.2, consists of designing an optical demagnification system (placed between the microscope objective and the Lytro) that increases the angular extent of the light (see Equation 1.12). This solution keeps desirable properties like the large numerical aperture of the microscope and its fixed working distance, while at the same time, the Lytro camera can remain unmodified. The major task is to find a good trade-off between the vignetting and the magnification of the resulting light field microscope.

Our best solution along this direction employs two lenses. This configuration serves two goals: 1) to decrease the magnification successively, simplifying the task of each individual lens, and 2) to move the image behind the Lytro camera so that it can be used in the inverse regime which offers a better spatial coverage $c_{spatial}$. The ray-diagram in Figure 6.6 illustrates the two-lens setting: an intermediate image that is slightly demagnified is created in front of the microscope. Then, the second lens creates a further demagnified image behind the Lytro camera. The Lytro camera is

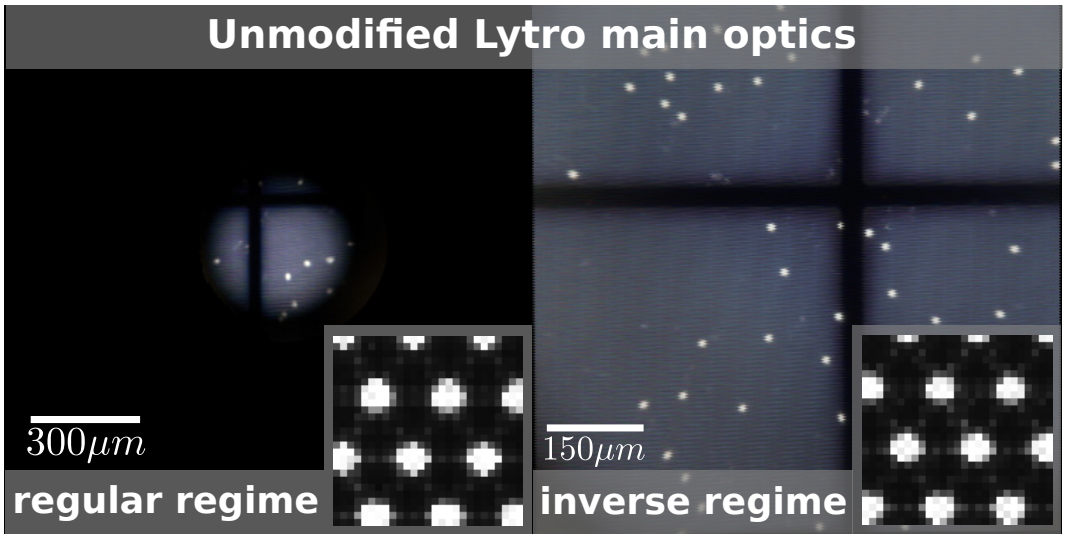


Figure 6.5: Direct imaging through a $\times 10$ microscope objective with an unmodified Lytro. The object is a blue LCD panel. It consists of a square black grid that is separating the different pixels of $0.5mm \times 0.5mm$ size. We hypothesize the white dots to be bubbles inside the liquid crystal. The images show the equivalent camera image computed from the light field, while their insets show a close-up on the microlens images of the raw sensor data for different zoom-levels. Spatial and angular vignetting are easily observed in the large and small images, respectively. The zoom level setting is $1\times$ (regular regime) on the left and $5\times$ (inverse regime) on the right. The spatial vignetting is strong in the regular regime (6% of spatial coverage), while it is greatly improved in the inverse regime (100% of spatial coverage). Angular coverage is similar in both cases ($c_{angular} \approx 25\%$).

operated in its inverse imaging regime in order to pick up this virtual image. We determine the positions and focal lengths of the additional lenses using the following equation derived from the thin lens formula:

$$z = f\frac{M-1}{M}\,. \tag{6.3}$$

The focal length f and the position of the lens z are chosen according to the desired magnification M (negative because the image is inverted). In our implementation, the first additional lens has a focal length of $50mm$ and is put close to the microscope objective. The second additional lens has a focal length of $85mm$ and is put close to the Lytro. The effect of using two lenses is that the individual focal lengths are larger and the aberrations are reduced.

6.1.5 Macro Lytro

Our second option (see Figure 6.7) is to use a single SLR camera lens in front of the Lytro to achieve unit magnification. ($M_{final} = M_{SLR}M_{Lytro} = 1$). As for

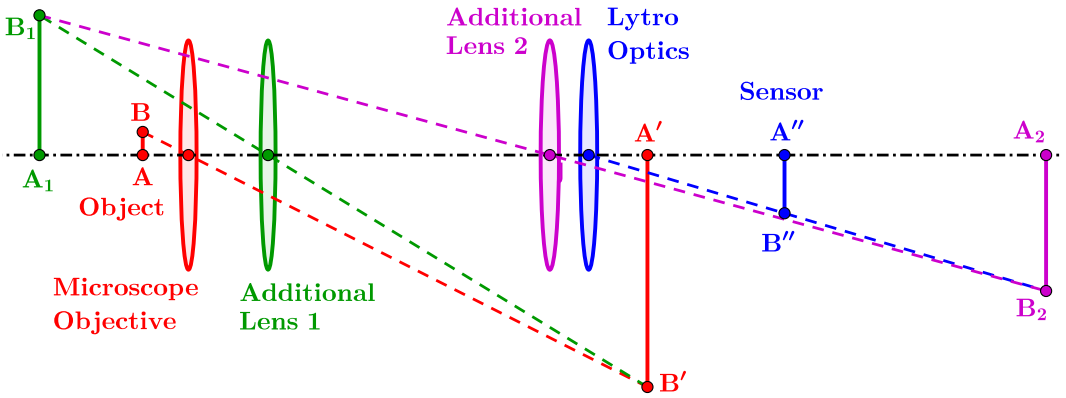


Figure 6.6: Diagram of the system using two additional lenses. Each sub-system is indicated in a different color and has to be interpreted independently from the other sub-systems. Their operation can be understood in sequence: the objective images object AB to $A'B'$, the first additional lens images $A'B'$ to A_1B_1 so that the second lens images A_1B_1 to A_2B_2 behind the Lytro camera. Finally, the Lytro's main optics in the inverse regime images this virtual object A_2B_2 to its sensor in the plane $A''B''$.

the microscope, and because the focal length of the SLR lens is large ($50mm$ and $100mm$ in our experiments), we want to use the Lytro in the inverse regime to keep the spatial coverage as high as possible. In practice, the Lytro is set as close as possible to the SLR lens. A variation of this setup is to remove the Lytro optics, and only use the SLR lens so that $M_{final} = M_{SLR} = 1$.

These designs have only one or two optical components and relieve the hurdle of undoing the work of the microscope objective with many lenses. However, the SLR lens is not specifically designed for the magnification of close objects and its aperture is not meant to be maintained at a constant value for the microlens array. The relations used to establish Equation 1.27 are not valid in this macro-configuration. Instead, the relevant quantity is the effective f-number, see Equation 1.21. The minimal value of effective f-number that can be achieved with a camera lens is close to $F_w = 2$. It is reached for a limit f-number of $F = 1$ and a magnification of $M = -1$. This condition would be optimal for the Lytro microlens array. However, commercial lenses usually have a limit f-number between 1.4 and 3.5 increasing the effective f-number to between 2.8 and 7.0.

Compared to the light field microscope, on the one hand, this setup is more versatile. The magnification can be set to the desired value by simply moving the object and the microlens array. It does not require the difficult alignment of several optics. On the other hand, this setup does not benefit from the structure of the microscope that already includes lighting and moving the sample through micrometer stages in three dimensions. In addition, the working distance is not fixed which changes the magnification as well as the object-side numerical aperture when moving the sample. However, the strong point of this design is its accessibility. Building a light field

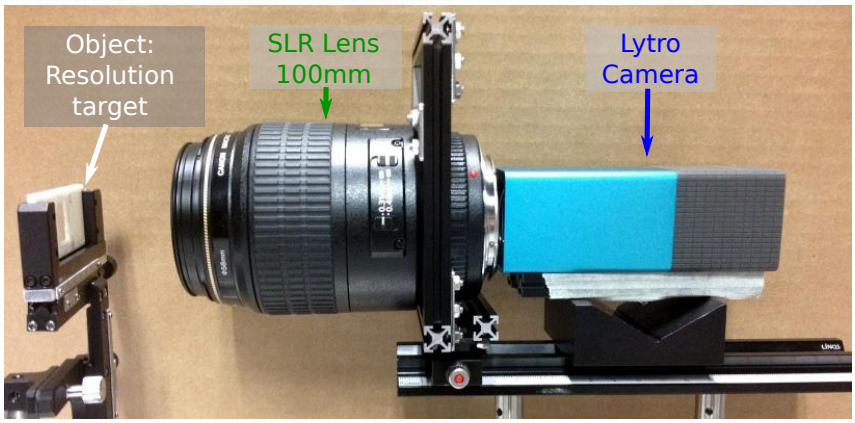


Figure 6.7: Setup from the SLR + Lytro experiment using a $100mm$ SLR lens ($100mm$ SLR + Lytro).

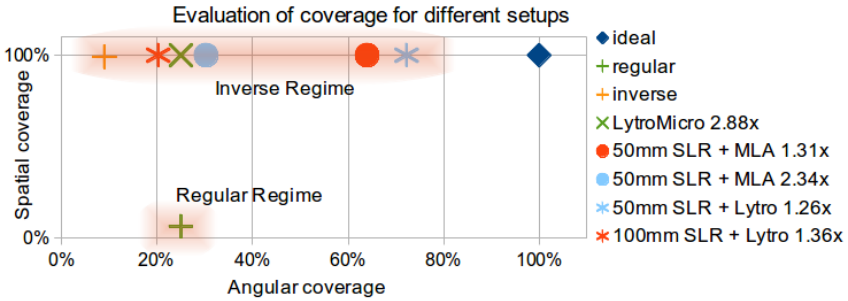


Figure 6.8: Combined results of the sensor coverage in the experiments from Section 6.1.4 and Section 6.1.5.

macrography setup is done quickly and without the need for a deep understanding of the operating principles of a microscope.

6.1.6 Results

Before showing results, we describe and compare the different implementations. The different sensor coverages, as well as their spatio-angular coverage values, can be found in Figure 6.8. It is clearly visible that directly using the Lytro camera in its regular imaging regime is unsuitable for microscopic light field imaging. The inverse imaging regime improves the spatial coverage, but the angular coverage is limited. The best combination is achieved with a $50mm$ SLR lens in front of the Lytro which yields the best overall sensor coverage c_{sensor}.

Resolution Test Chart

In order to compare the resolution of the two different techniques, we use a 1951 USAF resolution test chart. The results of the experiments are summarized in Ta-

Table 6.1: Performance results from the experiments of Section 6.1.4 and Section 6.1.5. Theoretical resolution is derived from the microlens pitch. The microlens footprint is the size of a microlens in object space.

Setup	Magnification	Resolution $[lp/mm]$		Microlens footprint $[\mu m]$
		Measured	Theoretical	
Lytro Microscope	2.88	80	103	4.87
$50mm$ SLR + MLA	1.31	32	47	10.7
$50mm$ SLR + MLA	2.34	51	84	5.99
$50mm$ SLR + Lytro	1.26	32	45	11.1
$100mm$ SLR + Lytro	1.36	36	49	10.3

ble 6.1.

The first option (Section 6.1.4) "Lytro Microscope" was implemented using a Canon $50mm$ SLR lens and a Nikon $85mm$ SLR lens as additional lenses. Those lenses were put on top of a Leitz Ergolux microscope using an objective of magnification 10 with an object-side numerical aperture $NA_o = 0.2$. This microscope has a lens tube with a magnification of 0.8 so the f-number is $F = 20$. The images have been taken with a magnification of 2.88, *i.e.,* a microlens covers $4.87\mu m$ in object space, see Figure 6.9 (top). The spatial coverage is above 99% but due to the large magnification the angular coverage is low (between 9% and 25%). The resolution is between 80 and 90 line pairs per millimeter $[lp/mm]$.

The resolution indicated above is computed for the center view. It decreases with further distance from the center. A loss of image quality due to aberrations can be observed. They are introduced because the observed area is larger than usual for the microscope objective. Microscope objectives are typically designed so that only a reduced inner portion of the full field is very well corrected. In addition, our matching lenses introduce further aberrations. Due to the angular vignetting being strong, the contrast of viewpoints far from the center is low. It should be noted that even viewpoints inside the vignetted area can be computed, albeit at a poor signal-to-noise ratio, see Figure 6.9.

The second option (Section 6.1.5) was implemented in three ways: two times with the Lytro placed behind two different lenses, a $50mm$ and a $100mm$ Canon SLR lens (referred as SLR + Lytro), and with the Lytro microlens array without the Lytro main optics, see Figure 6.1 (right), placed behind the $50mm$ lens (referred as SLR + MLA), see Figure 6.7. Magnifications from 1.26 to 2.34 were achieved. The spatial coverage is always 100% and angular coverage is good (up to 70%). In this case, chromatic aberrations are present which degrades the image. The aberrations are reduced in the SLR + Lytro case as compared to the SLR + MLA case because the magnification of the SLR lens is lower in this setting. It is most noticeable in the side views as, for these views, imaging is performed through the outer pupil regions of the SLR lens. We suspect that the chromatic aberration is introduced by the SLR

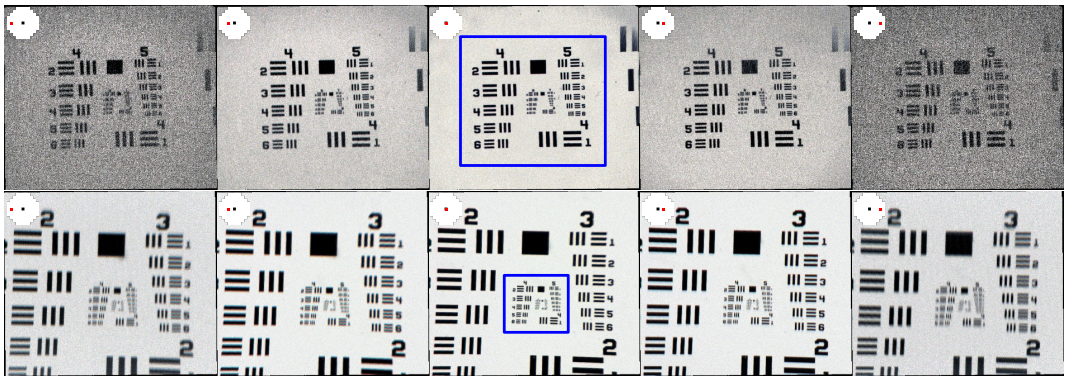


Figure 6.9: Set of different viewpoints of the resolution target with the center view in the middle (the red dot in the top left inset indicates the position of the view). *Top:* Images taken with the "Lytro Microscope" (Section 6.1.4). The magnification is 2.88. *Bottom:* Images taken with the "$50mm$ SLR + Lytro" (Section 6.1.5). A red-green color shift due to strong chromatic aberrations can be observed in the side views. Note that the top row has a higher resolution: it shows the pattern that is visible in the center of the bottom row (level 4 and 5). The contrast of the images of the same row has been set to a similar level for comparison.

lens because it is not intended for macro-imaging. Using a dedicated macro-lens instead would likely remove this effect.

Microscopic Sample

The most direct application of the light field microscope is the study of microscopic samples. The low magnification and the large field of view allow us to see in detail an object area that is between $1.5mm \times 1.5mm$ and $3.5mm \times 3.5mm$ with a magnification of 2.88 and 1.3 respectively. Cell tissues or rough surfaces of different materials have a structure close to the millimeter so high magnification is not always necessary to analyze them.

We illustrate this technique in Figure 6.10 (right). Several images of microscopic specimen were taken with the same settings as in the previous section. The magnification is 2.88 and we can clearly see the structure of different kind of surfaces that are invisible to the naked eye.

The light field data allows for the reconstruction of the depth of the sample when the number of views is sufficient. We took a picture of a ground truth aluminium staircase with stairs of $1.00mm$ width and $0.50mm$ height with an accuracy of $\pm 5\mu m$ with the $50mm$ SLR + Lytro setup. We obtained the depth map in Figure 6.11 (bottom view) using a modified variational multi-scale optical flow algorithm for light field depth estimation [Manakov *et al.*, 2013]. Although only a small slice of the staircase is in focus (the depth of field is $1mm$), the depth can be computed outside of this area. Essentially, out-of-focus regions are naturally con-

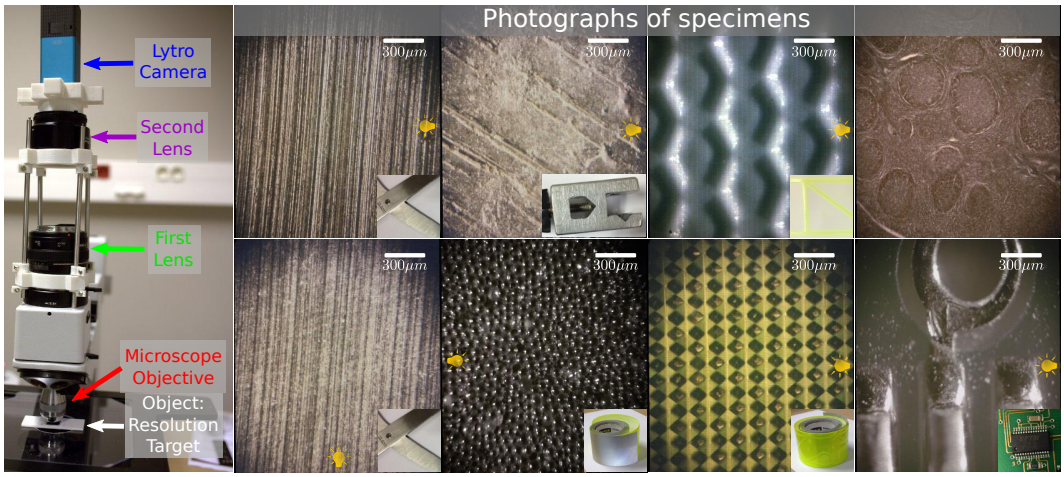


Figure 6.10: *Left:* Microscope setup with the Lytro camera on top of two additional SLR lenses. Images of several samples have been taken under different illumination conditions. Magnification is 3.0. *Top and bottom left:* brushed steel from scissors blade with lighting from the right and bottom. *Top middle left:* Scratched surface of a piece of metal. *Bottom middle left:* Plastic surface with highly retroreflective properties. The material is made of plastic micro bubbles that are invisible to the naked eye. *Top middle right:* Fabric with a hexagonal structure. The lighting is coming from the side and casts shadows and strong highlights on the three-dimensional structure of the fabric. *Bottom middle right:* Highly retro-reflective material from security reflective tape. *Top right:* Tonsil tissue with bright field illumination. *Bottom right:* Pins of an electronic component on a circuit board.

sidered as a different scale by the algorithm, so, the estimation of the depth of the closest and furthest steps is correct. This behavior nicely interacts with the scene properties, the parallax being larger in out-of-focus regions. The optical system can be seen as supporting the part of the algorithm that is handling large displacements. The detailed properties of the Lytro main optics are unknown, therefore, it is necessary to adjust the scale of the computed depth. We use the aluminium staircase as the reference. We find the affine transformation between the depth map and the staircase model by fitting planes to match the stairs. After the transformation, we measure an RMS error of $75\mu m$ for vertical planes and $17\mu m$ for horizontal planes. The difference is due to the different slopes of the horizontal and vertical steps with relation to the camera view direction. The magnification is equal to 1.32. We also applied this depth reconstruction to a daisy head, see Figure 6.11 (top row). In practice, the depth inside a cube of about $3.5 \times 3.5 \times 3.5mm^3$ can be estimated, which is rather large for a microscopic setting.

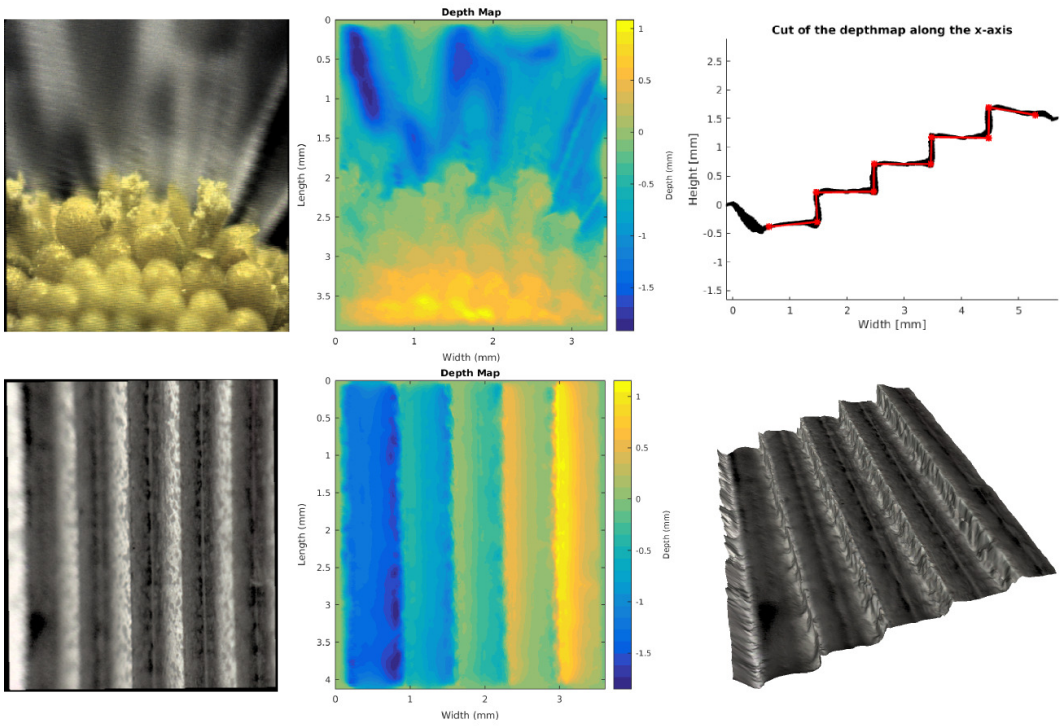


Figure 6.11: Subview (top left) and computed depth map (top middle) of a daisy flower. Subview (bottom left) and computed depth map (bottom middle) of the aluminium staircase. The bottom right picture is a 3D visualization of the depth map after the calibrating and the top right picture is a projection of a region in the middle of the calibrated depth map onto the xz plane showing the plane fits in red.

6.2 ECA Simulation for the Light Field Microscope

This section investigates the design of a light field microscope according to the configuration introduced by [Levoy *et al.*, 2006] with the help of the ECA model. The setup serves as an additional application of the ECA model in a different optical context than the light field cameras in Chapter 5. In general, the model allows to characterize the microscope properties which can indicate the practical aspects of the setup for a particular application like the 3D reconstruction of a certain microscopic volume. The following study includes a practical realization of the setup, simulation results with vignetting considerations, and experimental results to confront the simulation with the actual setup.

6.2.1 Setup

The implementation of the light field microscope, see Figure 6.12, is separated into two main parts: the magnifying part (in green) and the relay part (in red), that are separated by the DMU (in blue), in this case, a microlens array.

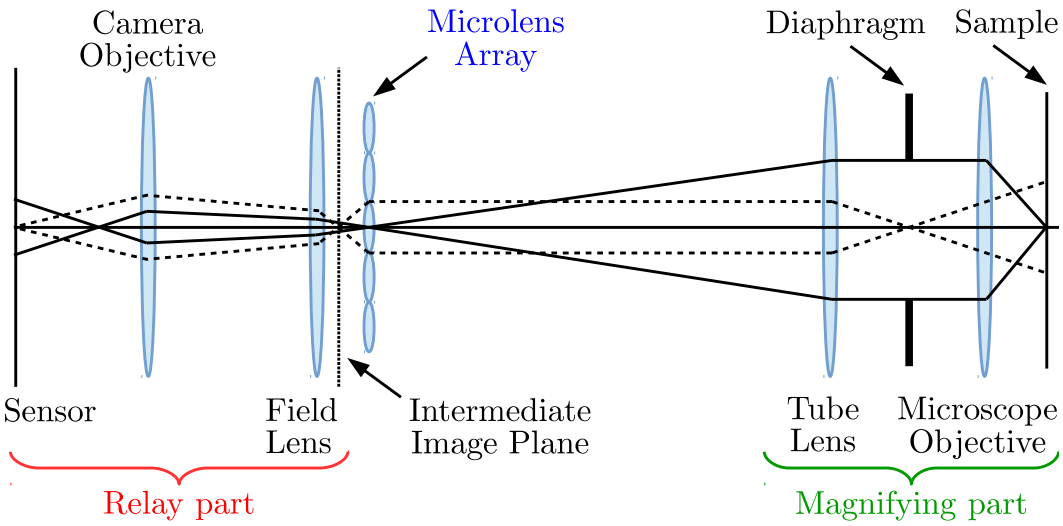


Figure 6.12: Sketch of our light field microscope.

Magnifying part The magnification is performed by an infinity-corrected objective with a factor of 20 and a numerical aperture of 0.5 together with a tube lens with a focal length of $200mm$ and a diameter of $30mm$. The image-side f-number of this optical system is 20. Both, the objective and the tube lens function together to magnify an object placed at the working distance of the objective. The DMU is placed at the image plane of the tube lens. The DMU is a microlens array from RPC Photonics. Each microlens is plano-convex and has a focal length of $2.1mm$. Its aperture is a square with a side length of $100\mu m$, hence, an f-number of 21. The whole array is arranged in a square grid and its size is $50.8mm \times 50.8mm$ or $2inch \times 2inch$ for a total number of 508×508 microlenses. The array is placed with the curved side facing the tube lens.

Relay part The relay system consists of a doublet field lens with a focal length of $75mm$ and a diameter of $48mm$ placed close to the back (flat side) of the microlens array and a DSLR Canon camera. The camera objective is placed behind the field lens and serves as a relay lens for its sensor. The Canon EF macro objective has a focal length of $100mm$ and operates in a 1:1 macro configuration at $f/2.8$. The sensor of the Canon 5D Mark II camera has 5616×3744 pixels and a total size of $36mm \times 24mm$. Each pixel has the shape of a square with a $6.41\mu m$ side length and has a bit depth of 14 bits. The sensor also comprises an RGB Bayer filter for color imaging.

Design considerations After the MLA is chosen, the magnifying and light field parts of the system are fixed. The relay part must be designed such as to minimize vignetting and to ensure a target spatial vs directional resolution trade-off. The following considerations influence the design.

1. Without the field lens, the camera objective images directly the back focal plane of the microlens array onto its sensor. It also sets an additional magnification but it is constrained by an important rule for realizing the ECA in the afocal configuration. Virtual cameras can only be constructed from object space pixel apertures if they share the same center of perspective. Consequently, in the image plane of the microlenses, the ratio of the microlens separation to the pixel size must be an integer. The camera objective magnification must be chosen to ensure this condition as well as possible. The simulation actually needs a perfect integer ratio to work but, in the real setup, it is possible to compensate for a non integer ratio by re-sampling the microlens image accordingly. Interpolation may introduce additional blur though.

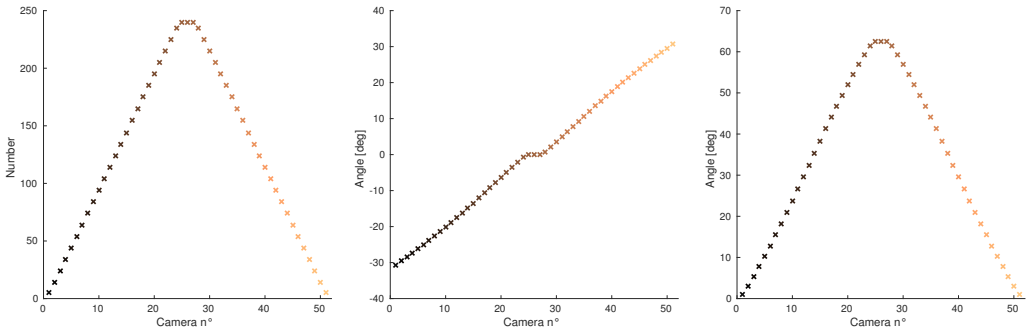
2. The role of the field lens is to reduce the amount of vignetting in the relay part of the system due to the divergence of the light beam after passing through the microlens array. In theory, the best is to place it directly in the microlens image plane. In practice, it is placed as close as possible to the lens array and it influences the magnification condition of the camera objective.

3. The free parameters for the relay part are the position of the field lens and of the camera as well as the focal length of the field lens and the magnification of the camera. The goal is first, to satisfy the integer condition and second, to adapt the magnification factor of the relay part to the desired spatio-angular trade-off. The magnifying part which is fixed assures the spatial resolution of the system, so the real trade-off does not influence this aspect but it influences the field width that can be imaged onto the sensor. The real trade-off is the number of microlenses that can be observed versus the angular sampling, or number of pixel in a micro-image. However, the width of the field itself is limited by the vignetting of the magnifying part. In the end, the desired trade-off is to match the field width to the sensor width and accept the resulting angular sampling or to control the angular sampling at the sacrifice of the covered field. With a magnification of the relay part larger than 1 in absolute value, the micro-images cover more pixels which increases the angular sampling and inversely, a magnification smaller than 1 implies less angular sampling.

4. The main sources of vignetting are the microscope objective that limits initially the field but also the relay part. In order to minimize an additional decrease of the field, the apertures of the relay components should be as large as possible and positioned as close as possible to each other.

6.2.2 Simulation

For the simulation of the ECA of the microscope, we consider the vignetting effects due to the apertures of the different optical components. Its main effect is to change



(a) Number of pixels per camera (b) View direction (c) Angle of view

Figure 6.13: Virtual camera properties of the ECA of the microscope setup.

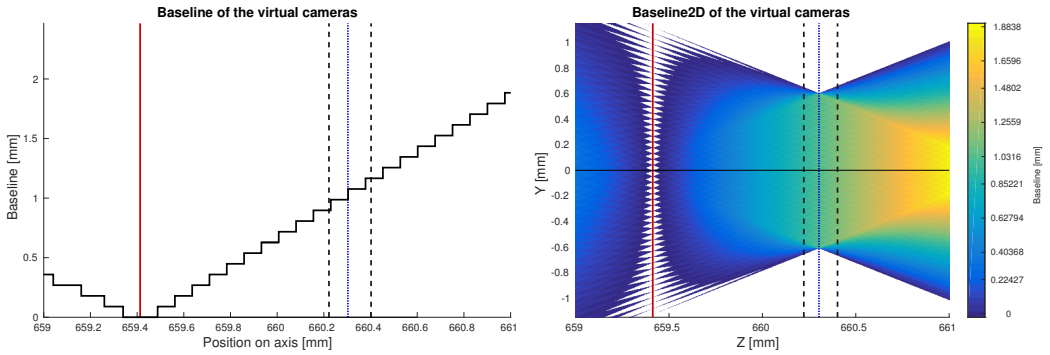
the number of cameras and the number of pixels of each camera, see Figure 6.13a, which affect all the other properties, see 5.1.3.

Properties

The virtual camera array is made of 51 cameras instead of the 21 cameras the naive approach would predict. This difference comes from the fact that the system is in the afocal configuration where virtual camera apertures are the image of the pixel apertures in object space. This aspect was not discussed in Section 4.3.5 and completes the vignetting effects for the afocal configuration. The position of the virtual camera centers depends on the relative position of the pixel center to the orthogonal projection of the microlens center on the sensor plane, the blue cross in Figure 4.9. In the naive approach, see Section 4.3.1, the micro-images are centered on this microlens center orthogonal projection making the number of virtual cameras equals to the number of pixels in one dimension in the micro-image: 21. With vignetting considerations, the microlens center orthogonal projection is not centered on the center of the micro-image anymore as illustrated in Figure 4.9. So, the pixels of the micro-image are more distant to this reference point and the number of virtual cameras increases but they have less virtuals pixels, see Figure 6.13a.

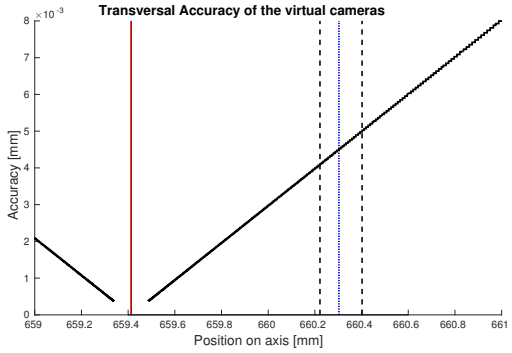
The virtual camera array is located $890\mu m$ in front of the plane containing the virtual pixels, see the distance between the red and blue dotted line in Figure 6.14b. The virtual camera apertures have a width of $45\mu m$, a step height of the black curve in Figure 6.14a is twice this value. The separation between the two most distant virtual cameras is $2.28mm$, see the extent along the red line in Figure 6.14b. The virtual pixels are $4.5\mu m$ in width, see the black curve value as it intersects the blue dotted line in Figure 6.14c and the size of the area covered by at least one camera is $1.22mm$, see the extent along the blue dotted line in Figure 6.14b. The magnification of the magnifying part is the ratio of the physical size of a microlens to its width in object space: $\frac{100}{4.5} = 22.2$. The depth of field is $180\mu m$, see the distance between the black dashed lines in Figure 6.14a.

Virtual camera properties The number of pixels per camera varies from 1 for the outermost ones to 240 for the center ones. It never equals 270, the number of simulated microlenses because of vignetting. The view direction of the cameras spans a wide range of angles between $-30°$ and $30°$, see Figure 6.13b. The angle of view is restricted for the side cameras as they have fewer pixels and reaches $62.5°$ for the center ones, see Figure 6.13c.

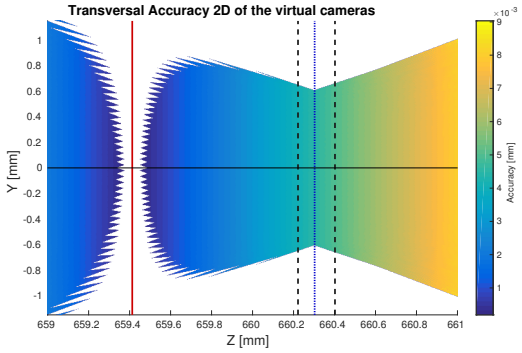


(a) Baseline on axis

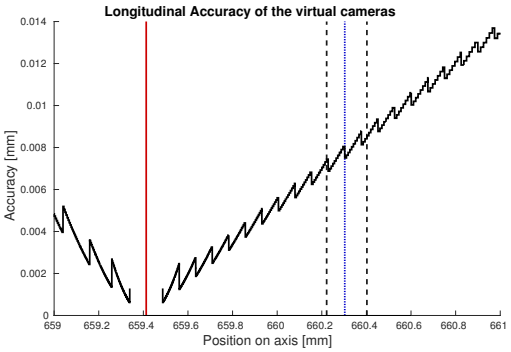
(b) Baseline on field



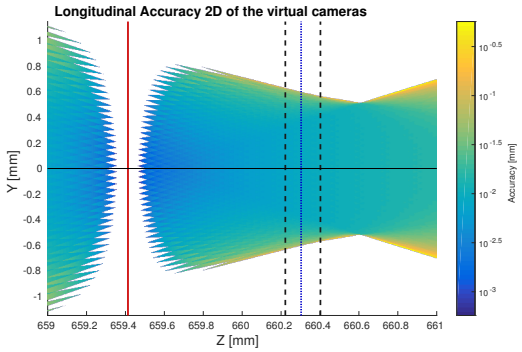
(c) Transversal Accuracy on axis



(d) Transversal Accuracy on field



(e) Longitudinal Accuracy on axis



(f) Longitudinal Accuracy on field

Figure 6.14: Field properties of the ECA of the microscope setup.

Field properties The effect of vignetting on the field properties is mainly to change the baseline around the virtual pixel plane. With the naive vignetting approach of Section 4.3.1, each camera observes the same part of this plane, so the baseline equals the separation between the most distant cameras. With the improved vignetting approach of Section 4.3.2, cameras see different parts of the focus plane, see the prolongation of the dark blue cone that originates from the center of the virtual cameras on the red line in Figure 6.14(a) and (b). Still, on axis, the baseline is approximately $1mm$. The transversal accuracy is between $4\mu m$ and $5\mu m$ in the depth of field range, see Figure 6.14(c) and (d), and the longitudinal accuracy is between $7\mu m$ and $9\mu m$, see Figure 6.14(e) and (f).

6.2.3 Experimental Results

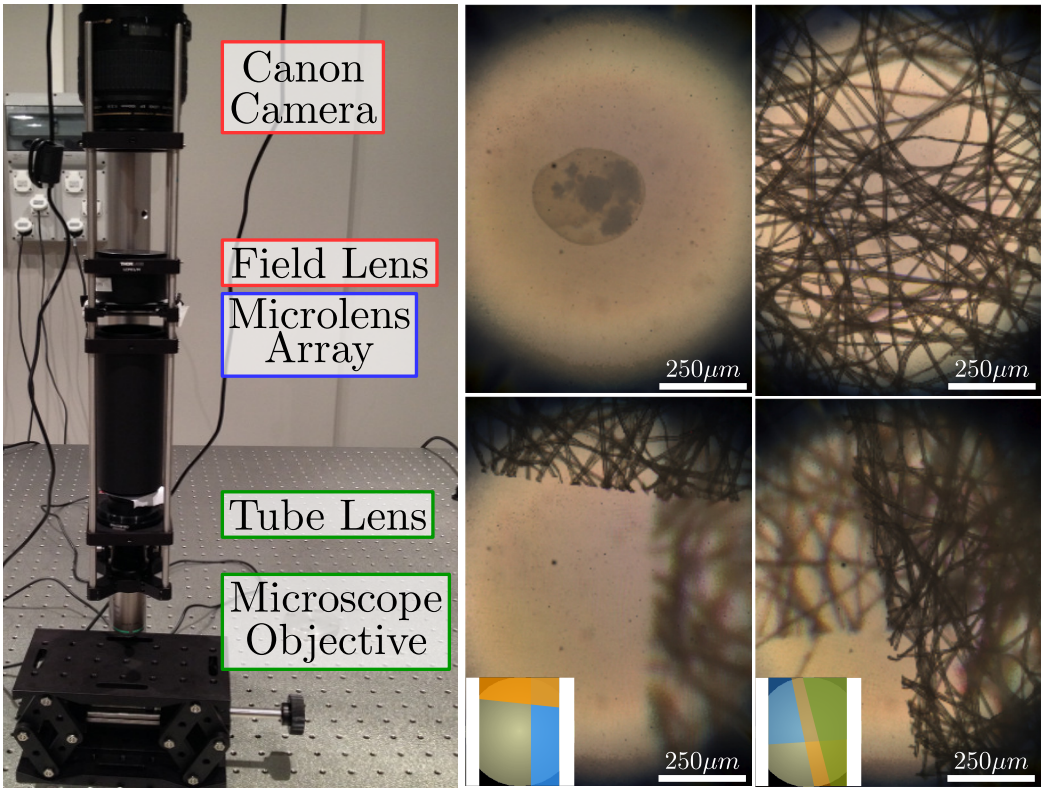
Construction and Alignment We use a cage system to hold the different components, see Figure 6.15a. It is made for optics no larger than $50.8mm$ which perfectly fits the size of our microlens array.

In order to facilitate the alignment and positioning of some components, a controllable diaphragm is positioned between the microscope objective and the tube lens. As a result of the microscope objective being telescopic on the image side, closing down the diaphragm aperture is equivalent to reducing the cone of light emitted by the object points. It also helps to locate the image plane of the microlenses and in adjusting the position of the camera to image this plane sharply.

Sensor raw image On the raw image, the micro-images are disks with a diameter of 21 pixels arranged in a rectangular grid. It means that the magnification of the relay part is 1.35. The number of microlenses covering the sensor is 270×180. The f-number is matched, so, the angular coverage $c_{angular}$ should be 100% but because of the packing of the microlenses in a square grid, there is an additional gap in the diagonal between microlenses and approximately 21% of pixels are lost. The spatial coverage $c_{spatial}$ is not 100% either because the field is a circle centered on the sensor with a diameter equal to 88% of the sensor width. The spatial coverage is actually 79% and the total sensor coverage c_{sensor} is 62%.

Performance We use a 1951 USAF test chart to evaluate the performance of the system. The system reaches a resolution of $114lp/mm$ for a magnification of the first part of the system of 22.2. The size of a microlens in object space being $4.5\mu m$, this is at the upper limit of the expected performance. The overall aberrations are low and constant even on the side views, see Figure 6.16.

Samples Several samples were observed with the light field microscope. They are presented in Figure 6.15b. The cell clusters are not well resolved with the naive approach used to synthesize the picture and would certainly benefit from state of



(a) Setup (b) Samples

Figure 6.15: (a) Setup of the light field microscope. (b) Series of samples acquired with the microscope. On the top left, the sample is a plastic shell containing groups of cells. The rest are pictures of optical paper. The top right picture has only one layer. The bottom left and right have two and three layers respectively. Their arrangement is indicated in the inset. The color scheme indicates the superposition ordering: blue is below orange which is below green. The distance between consecutive layers is approximately $350\mu m$.

the art super-resolution techniques, see Section 2.1.4. On the opposite, the fiber structures of the optical paper are large enough to be resolved correctly. Figure 6.17 presents the reconstructed depth map for the optical paper sample from Figure 6.15b (top right) rotated clockwise. In this single layer, the depth difference between the closest and furthest fiber is approximately $100\mu m$. The depth reconstruction algorithm is the same one used in Section 6.1.6. Strong vignetting on the side of the image and a lack of fibers are the causes for the deep hole on the top right corner.

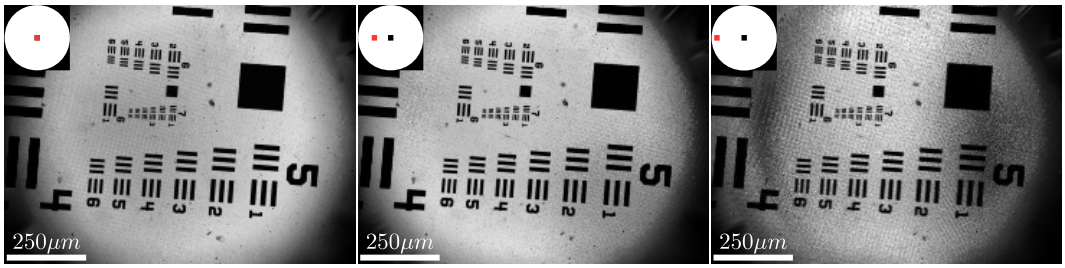


Figure 6.16: Set of different viewpoints of the calibration target. The left picture is from the center viewpoint, the middle picture is from the fifth side view and the right one is from the ninth side view. The optical quality is constant but the contrast and signal to noise ratio decreases dramatically on the side view. The total field is as large as the sensor width. The intensity of the pictures of the side views has been leveled to the intensity of the center view.

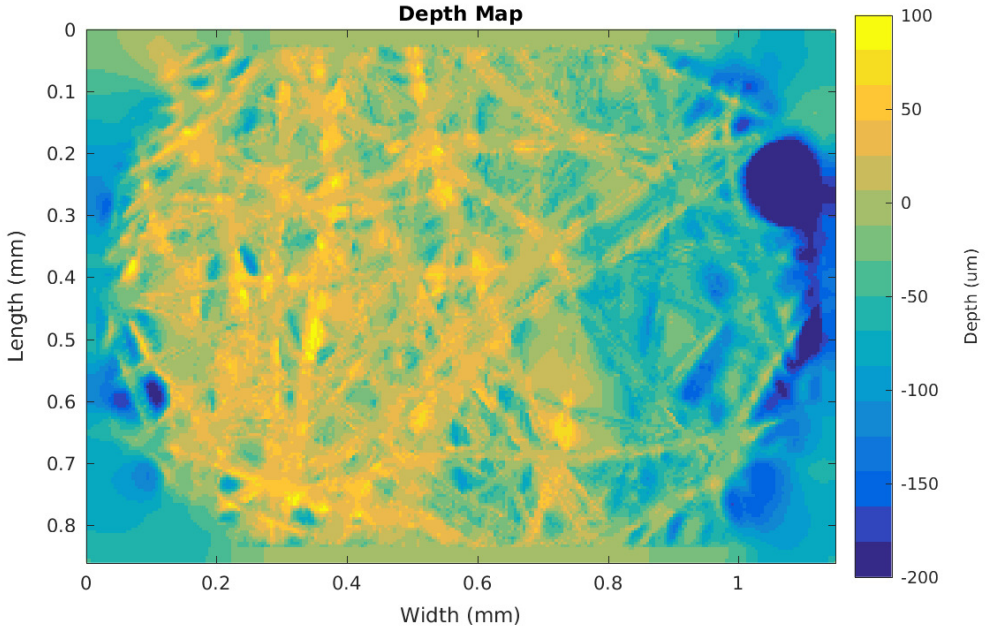


Figure 6.17: Reconstructed depth map for the top right sample of Figure 6.15b.

6.3 Conclusion

In this chapter, we have developed and tested several adaptations of the Lytro consumer light field camera to enable an entry-level experimentation with light field microscopy. While the fixed f-number of the Lytro's microlens array prevents its

direct use with a standard microscope (regular regime), it is possible to trade the overall system magnification for light field features and to avoid spatial vignetting with the inverse imaging regime. Lytro microscopy is therefore an option for low-magnification work as is common in industrial settings, or for investigations into the meso- and large-scale micro-structure of materials. Even though the achieved optical magnification between 1 and 3 appears to be low, the small size of the microlenses still yields a decent optical resolution of up to $6.25\mu m$ in object space which already shows interesting optical structures that are imperceivable by the naked eye.

In addition, we have implemented our own light field microscope using the predictions of the ECA model as guidelines in its design. We have shown that the vignetting caused by the apertures of the various components, mainly those of the microscope objective and of the relay part, reduces the number of pixels of the virtual cameras on the sides, thereby affecting the ECA properties considerably. Nonetheless, the microscope achieves a magnification of 22.2 and an overall good optical performance with an optical resolution of $4.5\mu m$ and few aberrations.

The ECA of the light field microscope has 51 virtual cameras which view directions range from $-30°$ to $30°$ and so, this microscope seems a good candidate for an application in tomography. Tomography requires a large number of cameras with the largest range of view directions around the object to be reconstructed accurately. The first requirement can be addressed with a higher number of pixels per microlens, for instance, by reducing the pixel width. The second can be addressed using a microscope objective with a higher NA.

Part IV

Implementation

Chapter 7

Software Implementation

In Chapter 5, we have derived the analytical formulation of the properties of the ECA virtual cameras in the case where the optical system was simple: it was constituted of a sensor, a microlens array and a single main lens. However, deriving the equations for the ECA properties may not be so simple in systems with a more complex arrangement of optical components or when vignetting strongly matters. To handle more general systems, we implemented the Geometrical Optical Design (GeODe for short) MATLAB Toolbox. In practice, we used it to compute the simulation results of the light field microscope in Chapter 6 and the field properties in Chapter 5.

In this chapter, we discuss the design choices that lead to the structure of the toolbox. These choices were guided by the intent to include the different mathematical and algorithmic processes introduced in the previous chapters such as the imaging matrices of Chapter 3 and the ECA construction of Chapter 4. The goal is to provide enough important high level aspects so that the ECA model can be implemented by a third party. We also propose a solution to the computation of the vignetting function in phase space introduced in Section 4.3.3 as we consider this part to be essential to the ECA model.

7.1 Design Choices for the GeODe Library

The design of optical systems is usually performed using some specific commercial software such as CODE V, ZEMAX or OSLO. These pieces of software are based on propagating rays through a sequence of optical surfaces, one after another. The sequence of surfaces is represented as a table where each element is a row and the element properties (radius of curvature, thickness, aperture, etc.) are indicated in columns. The row ordering specifies the ordering of the sequence from top to bottom. This sequential imaging removes the need to find the order in which the rays intersect the consecutive surfaces and makes the computation faster. This system is inherited from the historical *y-nu* trace described in [O'Shea, 1985] that is well suited for rotationally symmetric systems. It is also possible to run a non-sequential

ray trace for more complex systems where the path of the ray and the order of the surface intersections needs to be determined as the ray propagates. Optical design software is known for their often difficult description of the decentering of optical components. Since each element is positioned and oriented relative to the previous element in the sequence, it sometimes makes it hard for the user to enter the desired decentering and it requires a deep understanding of the program to master. Once the system is set, the computation of the system properties is obtained from the propagation of rays.

The goal of the GeODe library is to provide the tools to properly model first-order light field acquisition devices. Unlike other optical design libraries, it is not based on ray tracing. Moreover, the design of this library incorporates the previous observations of the optical design software operating principles. The main features that such a library should implement are the following:

- to provide an easy positioning and orienting of optical components and other objects in 3D space,
- to image objects of interest (like the pixels of a sensor) through the optical components,
- to build a camera (light field or not) and evaluate its optical properties.

This section presents and discusses the design choices of the structure of the library. This structure is organized to address the three tasks in parts that are independent from each other. The first part is a tree class structure to represent the data of the system. The second part is a class that performs the sequential imaging operations. The first and second part both implement the matrix tools defined in Chapter 3. The third part is a class that implements the ECA construction and evaluation processes described in Chapter 4.

7.1.1 Data Structure

The data of the optical components and objects to be imaged through the system is organized in a hierarchical tree structure. This data structure is similar to the scene graph structure used in Computer Graphics. Its advantages are numerous. In the tree, each node may have many child nodes but it has only a single parent node. An inherent property of the tree is that the operation on a parent node propagates to its child nodes. As we have seen in Chapter 3, the imaging matrix of the system is obtained by concatenating the optical matrix and the transformation matrix of the optical components in the correct order. The tree structure provides a particularly efficient and natural way to concatenate these matrices in the correct order. Moreover, a common feature of the optical design software is to group several optical elements to form an optical group like a zoom group or a focus group. This is directly implemented when all the components share the same parent. Consequently,

the decentering of this parent affects all its elements and facilitates better control for the user.

The role of this data structure is not to organize the objects in the order of imaging but only to store and organize the geometrical and optical data of the components.

7.1.2 Sequential Imaging

The ECA model and the matrix system are both based on sequential imaging. However, our design choice is to decorrelate the imaging operations from the data structure. We propose a class that stores an ordered list of references to the optical components to represent an imaging sequence and performs forward or backward imaging of objects through the sequence. Defining an imaging sequence as a list of references has two main advantages. The first one is to give the hand to the user so that he can quickly explore several sequences by adding or removing components from the sequence without having to remove them from the data structure. The second advantage is the possibility to add the same element multiple times. A good example for this feature is the Kaleido Camera with its parallel mirrors facing each other. The light rays, depending on their angle of incidence, can bounce off the mirrors zero times, once, twice or multiple times. With our ordered list, this configuration is easily implemented.

However, the order of the components in the sequence is not forced by their position in the 3D space so it is possible to represent a sequence that is not physically possible. For instance, in a system made of three lenses placed at an increasing position on an axis, the sequence defined by the first lens, then the last lens and finally the middle lens is impossible.

7.1.3 ECA implementation

The class defined previously handles sequential imaging but does not make a camera. For this specific purpose, we define a new class which implements the ECA model. As described in Chapter 4, the imaging by a light field system is separated into four parts: the sensor, a first component sequence, the directional multiplexing unit, and a second component sequence. The role of this class is to implement the ECA building process and to compute its properties. The building process consists of imaging pixels and DMU elements to object space, evaluating the vignetting function of the pixel/DMU element pairs according to the specified vignetting method, and creating virtual cameras. The properties of the ECA are the object space properties of the individual cameras, the field properties of the entire array, and the white sensor image. The evaluation of the vignetting function is essential to the construction of the ECA. We address its computation in the following section.

7.2 Evaluation of the Vignetting Function

In Section 4.3.3, we explained how to compute the vignetting function of every pixel/DMU pair of a given optical system with several system apertures. For a pixel/DMU pair, its phase space cell is to be intersected with the phase space cells of the other two-element apertures, the ones made of the pixel aperture and each of the system apertures. The volume of the remaining region of intersection is the value needed for computing the vignetting coefficient, see Equation 4.1. Moreover, the boundaries of this region are also essential for computing of the optical properties of the virtual cameras of the ECA. We also demonstrated that the initial phase space cells and their intersections are convex because the involved apertures are convex.

Below, we present methods to compute the exact vignetting function in 2D and 3D space as well as optimization schemes and an alternative estimation based on Monte Carlo estimation.

7.2.1 2D System

For a 2D optical system, an aperture is the space between its boundaries which are two vertices. The bundle of rays defined by two apertures is comprised between four rays, the rays that pass through the two boundary vertices of the two apertures. In the phase space, these four rays map to 2D points and the ray bundle is the convex region delimited by these points forming a quadrilateral. Consequently, the problem of computing the phase space volume and shape of the intersected region corresponding to a vignetted ray bundle is relatively simple. It consists in intersecting convex polygons in 2D which can be solved with polygon clipper algorithms. The volume of the phase space cell is actually the area of the final polygon.

7.2.2 3D System

In Section 4.3, we described the ray bundles that traverse an optical system with set theory. However, in order to compute the etendue of the these ray bundles numerically, we need a geometric description of their corresponding phase space regions. For a 3D optical system, the phase space regions are four-dimensional, so the problem to solve is to compute the intersection of 4D shapes. These shapes are convex, and consequently, we first need to find which phase space points or their equivalent rays in 3D space make up their convex hull. As stated before, we consider that apertures are convex polygons defined in a plane. Curvy apertures like circular lens rims or some diaphragm blades can be well approximated by polygons with a sufficiently large number of vertices. Therefore, the convex hull is a polytope (the 4D equivalent of a polyhedron in 3D space) and is fully defined by a finite number of points. For a system of two apertures, the phase space points corresponding to rays that pass through the contour of both apertures definitely belong to the convex hull of the ray bundle delimited by these two apertures. As the contours are defined by

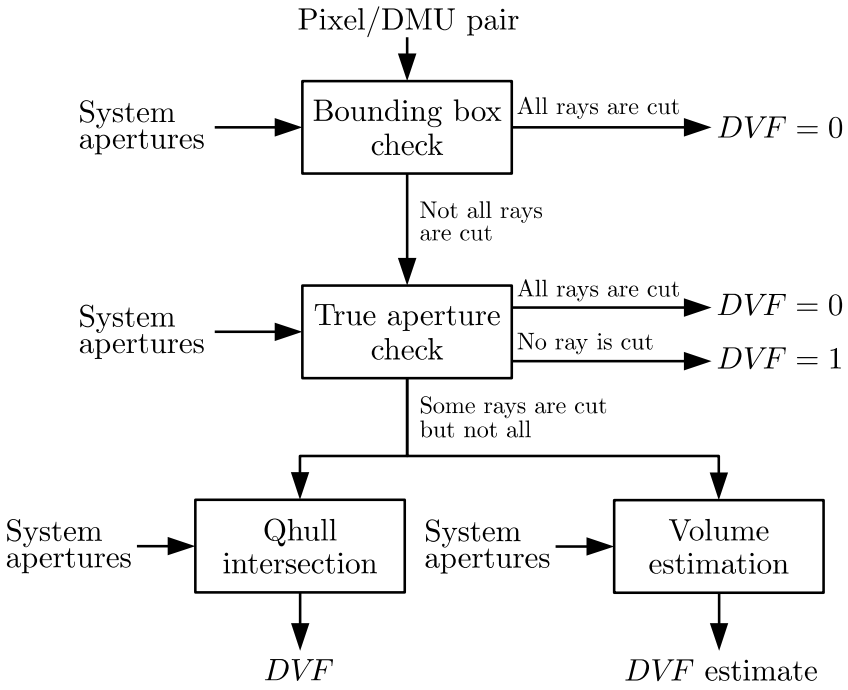


Figure 7.1: Steps to compute the vignetting function of a pixel/DMU pair.

the straight edges that connects two consecutive vertices, the rays that pass through every combination of the first aperture vertices and the second aperture vertices are the sought points defining the convex hull.

We use the Qhull library [Barber *et al.*, 1996] to compute the intersection between two convex hulls. It outputs the vertices of the convex hull of the intersection and it also computes the volume of the hull.

Computing the intersection with Qhull is computationally expensive because the library uses general algorithms for any dimensions which are not optimized for our four-dimensional problem. We describe optimization schemes as well as an alternative evaluation method in the following. The steps of the process are illustrated in Figure 7.1.

Optimization

The computational cost of evaluating the vignetting function is high for one main reason: a large number of pixel/DMU pairs. In the case of the light field microscope of Chapter 6, the sensor has 21.4 million pixels and the DMU has 48.6 thousand elements for a total of approximately 1 billion pairs. Most of the time though, only a small part of the sensor receives light from a single DMU, for instance, in the light field microscope, the micro-images are 21×21 pixels large which represents only 0.002% of the total number of pixels. So, it is useless to perform the computation of the intersection with Qhull if a pixel/DMU pair is completely vignetted but it is

also unnecessary if the pair is not vignetted at all.

The phase space region of the ray bundle of a pixel/DMU pair is defined only by a small set of rays, hence, we can determine if it is necessary to use the full intersection algorithm by checking first whether the rays pass through the system apertures. If all rays or none are cut, the value of the vignetting function is respectively 0 and 1. If some rays pass but not the others, the Qhull library is called to compute the intersection volume.

An additional optimization is to conservatively use bounding boxes that enclose the apertures instead of the true aperture polygonal shape. The bounding box of an aperture is simply a rectangle in the plane of the aperture that includes the true polygonal aperture. Using a bounding box for the two apertures reduces the number of rays to check for down to 16. The bounding box is also used for the system apertures. It allows to discard most of the completely vignetted pairs more efficiently, but the remaining pairs have to be checked again with their true apertures to be correctly classified as completely vignetted, partially vignetted or unvignetted.

The whole optimization process can also accelerate the 2D case even though the bounding box part is unnecessary.

Alternative Method: Monte Carlo Estimation

In the case where the Qhull library cannot be used or fails to deliver correct results, the intersection volume can be estimated using the Monte Carlo technique. With this technique, the volume is estimated by randomly sampling the phase space cell of the pixel/DMU pair and counting the number of points that belong to the intersection region. In practice, it is similar to the optimization scheme but with more rays. It requires to sample the pixel aperture and the DMU aperture uniformly, trace rays through these points, and check whether they are cut by the system apertures. The estimated volume $\widetilde{V}$ is obtained with the following formula:

$$\widetilde{V} = V \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{p}_i) \tag{7.1}$$

where V is the sampled volume of the phase space cell of the pixel/DMU pair, N is the number of samples, $\{\mathbf{p}_i\}_{i=1..N}$ is the set of samples and f is the indicator function of the intersection. This estimate is associated with a variance value the square root of which indicates the error of the measure.

$$Var(\widetilde{V}) = \frac{V^2}{N} \frac{1}{N-1} \sum_{i=1}^{N} (f(\mathbf{p}_i) - \langle f \rangle)^2 \tag{7.2}$$

where $\langle f \rangle$ is the average of f over the samples $\{\mathbf{p}_i\}_{i=1..N}$.

The drawback with Monte Carlo estimation is the slow convergence of the estimate with respect to the number of samples. Moreover, the estimation also suffers

from the curse of dimensions. For a small intersection volume, the number of samples to get an acceptable variance can be more than what is reasonable. However, the variance is known, so, it is possible to stop the computation earlier if the resulting error is smaller than a pixel quantification step or its noise level.

In addition, the shape of the phase space intersection region is estimated as the convex hull of the valid phase space samples.

7.3 Conclusion

We have shown the underlying concepts behind the GeODe toolbox: the arrangements of the optical components and other objects of the systems in a hierarchical data structure, the sequential imaging, and the construction of the ECA. The goal behind this toolbox is to provide sufficient features for the investigation of light field camera design. However, it is still lacking many aspects of a true optical design software. Ideally, the user would specify the conditions defined by a set of properties in object space and the system would optimize the component properties and positions to satisfy these conditions. Moreover, when a proper theory of aberrations for light field cameras is achieved, it should integrate ray tracing with real surfaces.

We also proposed a method to compute the vignetting function that is essential to the construction of the ECA, and we implemented it in the GeODe toolbox. Even though we proposed some optimization schemes, overall this evaluation is the most computationally expensive part in the library. In order to improve the performance, a parallelization of this task may be considered, or a reimplementation in a more efficient language, effectively merging the Qhull library calls in the code.

The GeODe source code is available on InriaForge[1].

[1]The Git repository is available at https://scm.gforge.inria.fr/anonscm/git/geodem/geodem.git

Conclusion

Light field imaging proposes a new paradigm to traditional imaging. Consequently, a large number of light field acquisition devices exist, with as many different designs to capture the spatial and angular components of light rays. Even so, the literature implicitly suggests that they all share in common a similar aspect. In a sense, the light field sampling they perform is equivalent to that of a camera array. This dissertation proved this assertion, assuming first-order imaging, and built on it by demonstrating how it can be used to reveal guidelines to the optical design of light field devices. This model opens the path to better optical systems optimized in a broad range of applications from the consumer market to the industry and fields from robotics to biology.

Summary

In this dissertation, we provided a solution for the analysis of the optical design of light field acquisition systems such as light field cameras and light field microscopes. This solution, the equivalent camera array model, considers imaging through first-order optics by breaking the system into a set of two-aperture elements. These elements are pairs of pixels and directional multiplexing units that are imaged to the object space and reorganized into an array of virtual cameras. The main advantage of this model over previous ones is to keep the true pixel and DMU apertures that define the elementary light field sampling kernel.

As a result of the need to image a large number of these primitives (the two-aperture elements) through the system components, we developed a mathematical tool based on point imaging matrices. This matrix scheme exploits the fact that imaging through an optical component like a lens connects the object points and their conjugate image points through a linear projection of the 3D space. This system benefits from the concept of homogeneous matrices from Computer Graphics to represent the imaging projection with a 4×4 matrix and decentering with a straightforward composition of geometrical transformation matrices. Even though the tool neglects ray propagation, its results are identical to those of standard ray tracing techniques such as ABCD ray matrices for first-order conditions and conversion between rays and points is easily done.

The ECA model also considers the vignetting effects caused by other apertures in the system that influence the number and position of the virtual cameras as well as the number and position of their virtual pixels. We presented a way to evaluate the vignetting of all two-aperture elements using the properties of optical phase space. The model consequently establishes a set of optical properties in object space for each of the virtual cameras as well as for the full array. We described these properties as they serve as indicators of the performance of the optical system and guide the designer to reach the targeted application.

We applied the model to microlens-based light field cameras in both the afocal and focused configuration. These configurations differ only by the distance between their sensor and their microlens array. We derived analytical formulas for their optical properties to compare them and discussed the consequences of the slight change in their design on their optical performances. We also validated the model with a commercial light field camera in the afocal configuration. We realized this validation by retrieving the lens parameters of the main objective of the camera through the fitting of several properties measured experimentally. A good estimation of the lens parameters was achieved even though the system is prone to many aberrations that deviate from our first-order model.

This dissertation also investigated light field imaging for microscopy applications. We first studied the association of a commercial light field camera and a microscope objective. We showed that the mismatch between their f-number caused large spatial and angular vignetting. Even though we managed to use the light field camera in an unexpected imaging mode that reduced spatial vignetting, angular vignetting prevented the utilization of the system with a high magnification factor. Consequently, we realized and tested several setups using a macro SLR objective. These systems proved that macro light field imaging was possible by reconstructing the depth in a cubic volume of a few millimeters in length. However, our setup certainly performs worse than one designed specifically for this macrography task. Then, we built and analyzed our own light field microscope. We discussed the guidelines of its design and used the setup to additionally validate the ECA model.

Lastly, we proposed a software implementation of the ECA model by the GeODe toolbox. This toolbox regroups aspects from standard optical design softwares such as sequential imaging and from Computer Graphics such as a scene graph to handle the positioning of optical components. These aspects are combined together to represent the imaging inside a light field optical system and extract its equivalent camera array and its properties. We also described in detail how to compute the vignetting function associated to each two-aperture element in the system and provided some optimization schemes as well as an alternative Monte Carlo estimation of this function.

Future work

The equivalent camera array model we presented is a first-order optical model and, as such, it completely neglects two main aspects of imaging: aberrations and diffraction. In true optical design, the first-order analysis serves as a layout to define the main functions of the optical system. Then, the system is refined with the true surfaces, analyzed with ray tracing and optimized to reduce the aberrations. Aberrations and diffraction are very common in all imaging devices. The first goal would be to develop an aberration model for light field systems that keeps the virtual camera structure of the ECA, each embedded with its own aberration function. The second goal would be to additionally include a diffraction model also in relation to the ECA model. With these two expansions, the ECA model would describe accurately the imaging of light field acquisition devices.

We saw how the vignetting inside the light field camera affects the number of cameras and their field of view as well as the calibration image. An interesting development would be to use this information for better extraction and reconstruction of the light field captured by the camera.

In this thesis, we have studied the use of the ECA model for light field cameras in depth. However, we believe that the model can also beneficially be employed for the study of light field display systems such as auto-stereoscopic screens. Future work would focus on defining which systems can be modeled as an array of virtual projectors and replicate our findings with similar validation methods and analysis. Actually, we wanted to test this idea with the light field microscope by adding an illumination path using the same microlens array as for the acquisition. Unfortunately, a shortage of time prevented us to get adequate results, so we leave this work unfinished.

In terms of applications, the measurement of micro- and meso-bidirectional reflectance distribution functions (BRDF) and their relation to macroscopic BRDF models also appears to be an interesting development. This type of measurement proposes new challenges as the whole hemisphere of directions from the surface needs to be sampled for both the illumination and acquisition. These constraints could definitely be simulated and tested with our model. The obtained data would serve to establish a bridge between different scales of observation and a better understanding of the connection between them.

Appendix A

Simulation Study

The next pages present the ECA properties for different systems from the literature. For each system, there are four configurations. The differences between the configurations are indicated in the title of the section. This title is divided into three parts separated by a dash. First is the name of the system. Second is the value of β (0.8/1.0/1.2) and the last is whether the system is in the focused or afocal configuration. Apart from the Lytro system which is in the afocal configuration, the original configuration (the one from the corresponding paper or patent) of each system is best represented by the $\beta = 1.2$ configuration.

A.1 System from the work of Bishop and Colleagues

From the work of [Bishop *et al.*, 2009].

A.1.1 Bishop-Beta0.80-focused

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	4095
Width (mm)	36.8550
Dist to MLA (µm)	280.0000

Lens Array Parameters	Value
Diameter (mm)	0.1350
Focal length (mm)	0.3500
F-number	2.5926
Pitch (mm)	0.1350
Number	273
Width (mm)	36.8550

Lens Parameters	Value
Diameter (mm)	40
Focal length (mm)	80
F-number	2
Thickness (mm)	0

Figure A.1: System parameters

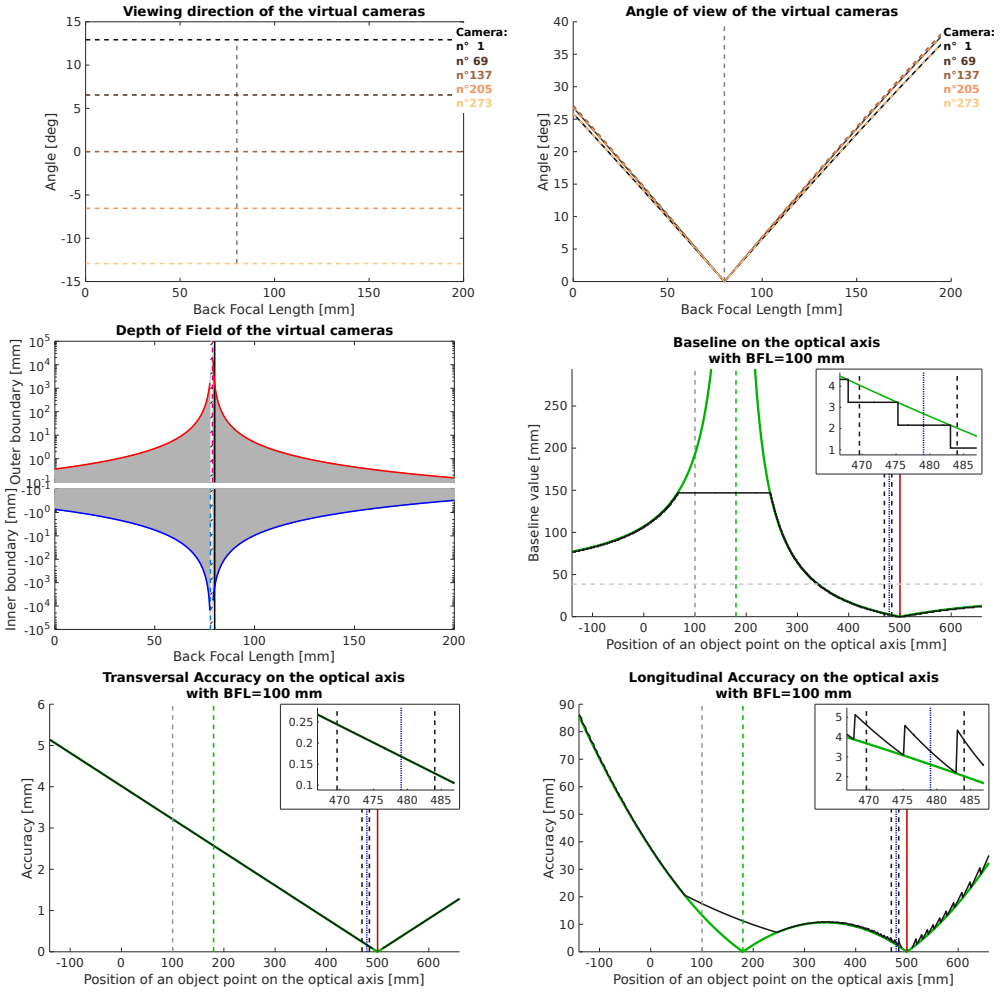


Figure A.2: ECA properties

A.1.2 Bishop-Beta1.00-afocal

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	4095
Width (mm)	36.8550
Dist to MLA (µm)	350

Lens Array Parameters	Value
Diameter (mm)	0.1350
Focal length (mm)	0.3500
F-number	2.5926
Pitch (mm)	0.1350
Number	273
Width (mm)	36.8550

Lens Parameters	Value
Diameter (mm)	40
Focal length (mm)	80
F-number	2
Thickness (mm)	0

Figure A.3: System parameters

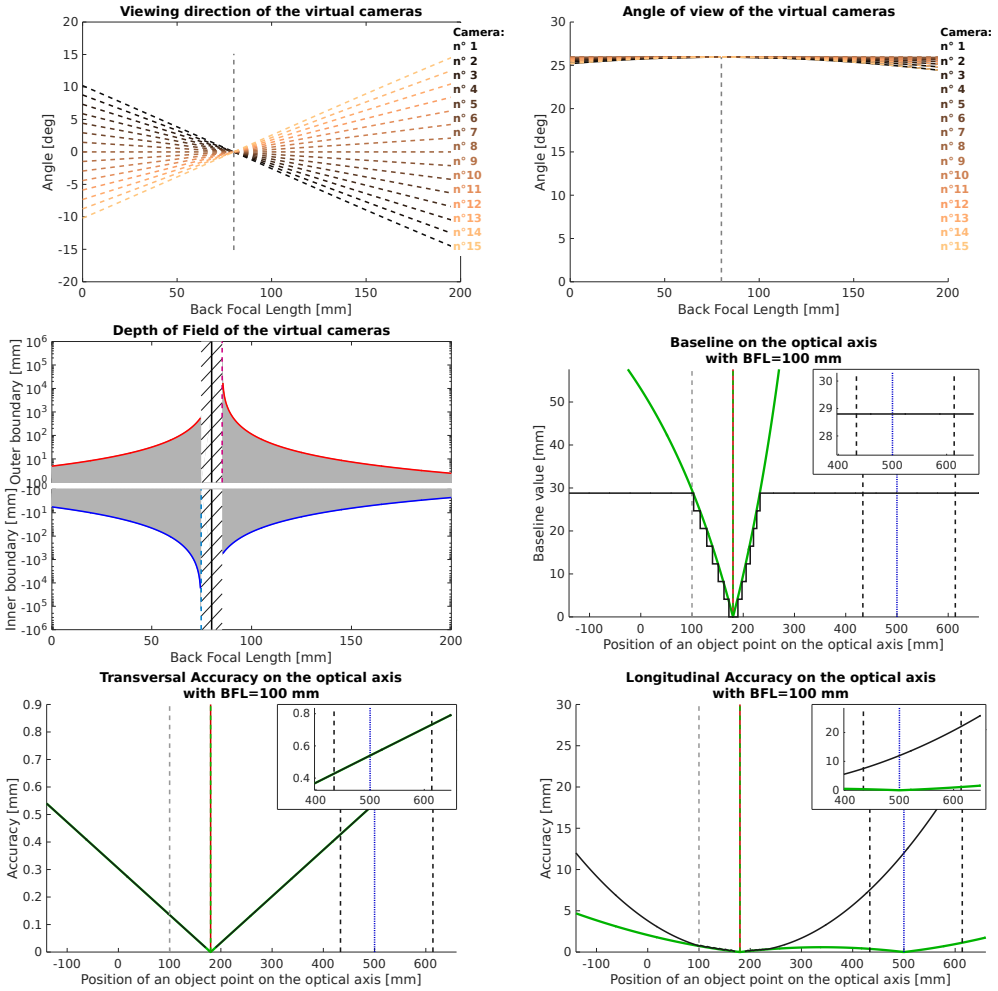


Figure A.4: ECA properties

A.1.3 Bishop-Beta1.00-focused

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	4095
Width (mm)	36.8550
Dist to MLA (µm)	350

Lens Array Parameters	Value
Diameter (mm)	0.1350
Focal length (mm)	0.3500
F-number	2.5926
Pitch (mm)	0.1350
Number	273
Width (mm)	36.8550

Lens Parameters	Value
Diameter (mm)	40
Focal length (mm)	80
F-number	2
Thickness (mm)	0

Figure A.5: System parameters

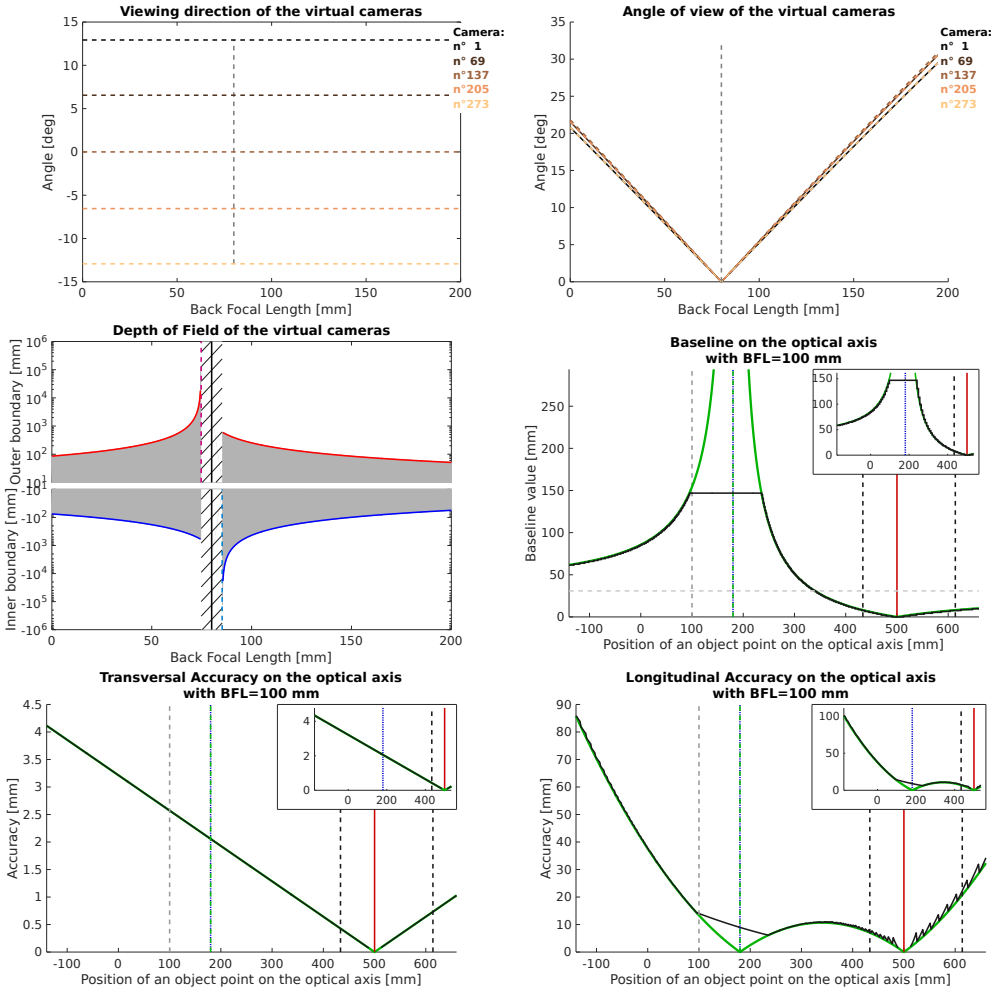


Figure A.6: ECA properties

A.1.4 Bishop-Beta1.20-focused

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	4095
Width (mm)	36.8550
Dist to MLA (µm)	420

Lens Array Parameters	Value
Diameter (mm)	0.1350
Focal length (mm)	0.3500
F-number	2.5926
Pitch (mm)	0.1350
Number	273
Width (mm)	36.8550

Lens Parameters	Value
Diameter (mm)	40
Focal length (mm)	80
F-number	2
Thickness (mm)	0

Figure A.7: System parameters

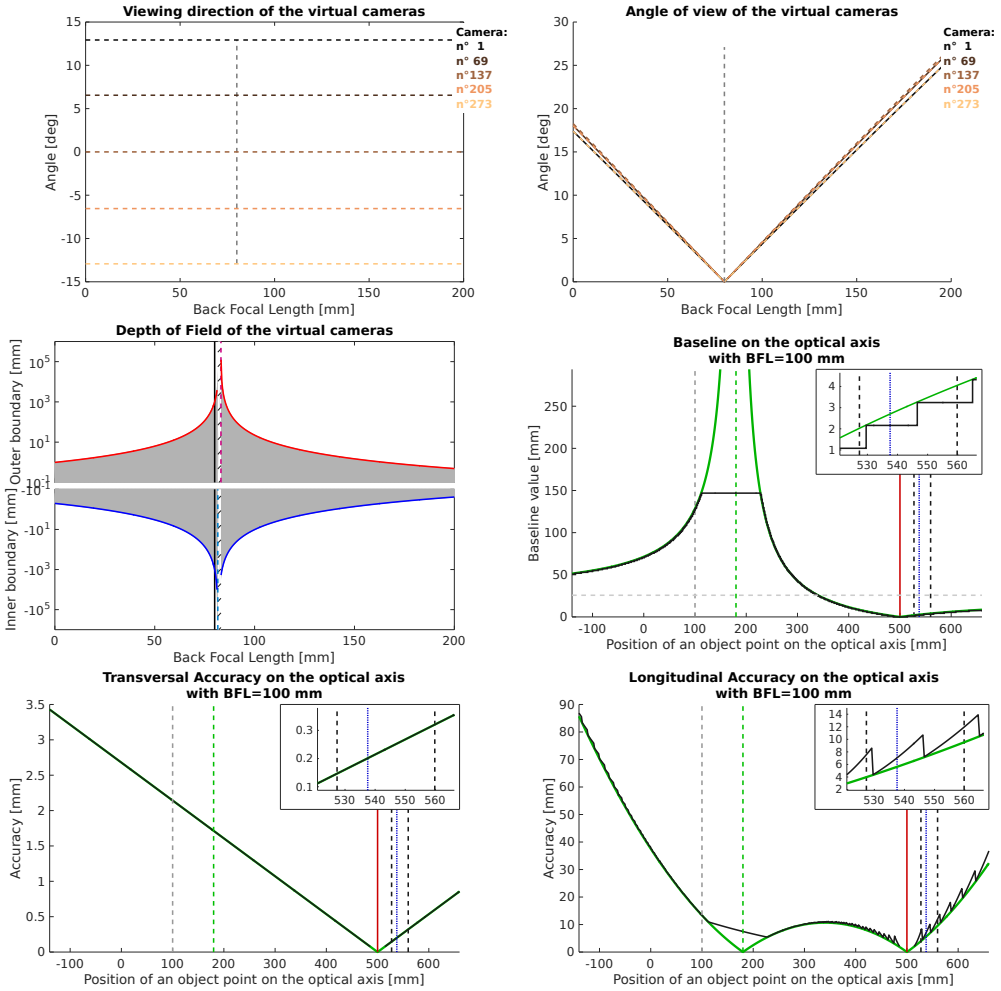


Figure A.8: ECA properties

A.2 System from the Lytro Camera Patent

From patent US8289440.

A.2.1 Lytro-Beta0.80-focused

Sensor Parameters	Value
Diameter (µm)	1.4000
Pitch (µm)	1.4000
Number	3280
Width (mm)	4.5920
Dist to MLA (µm)	22.4000

Lens Array Parameters	Value
Diameter (mm)	0.0140
Focal length (mm)	0.0280
F-number	2
Pitch (mm)	0.0140
Number	328
Width (mm)	4.5920

Lens Parameters	Value
Diameter (mm)	3.5000
Focal length (mm)	7
F-number	2
Thickness (mm)	0

Figure A.9: System parameters

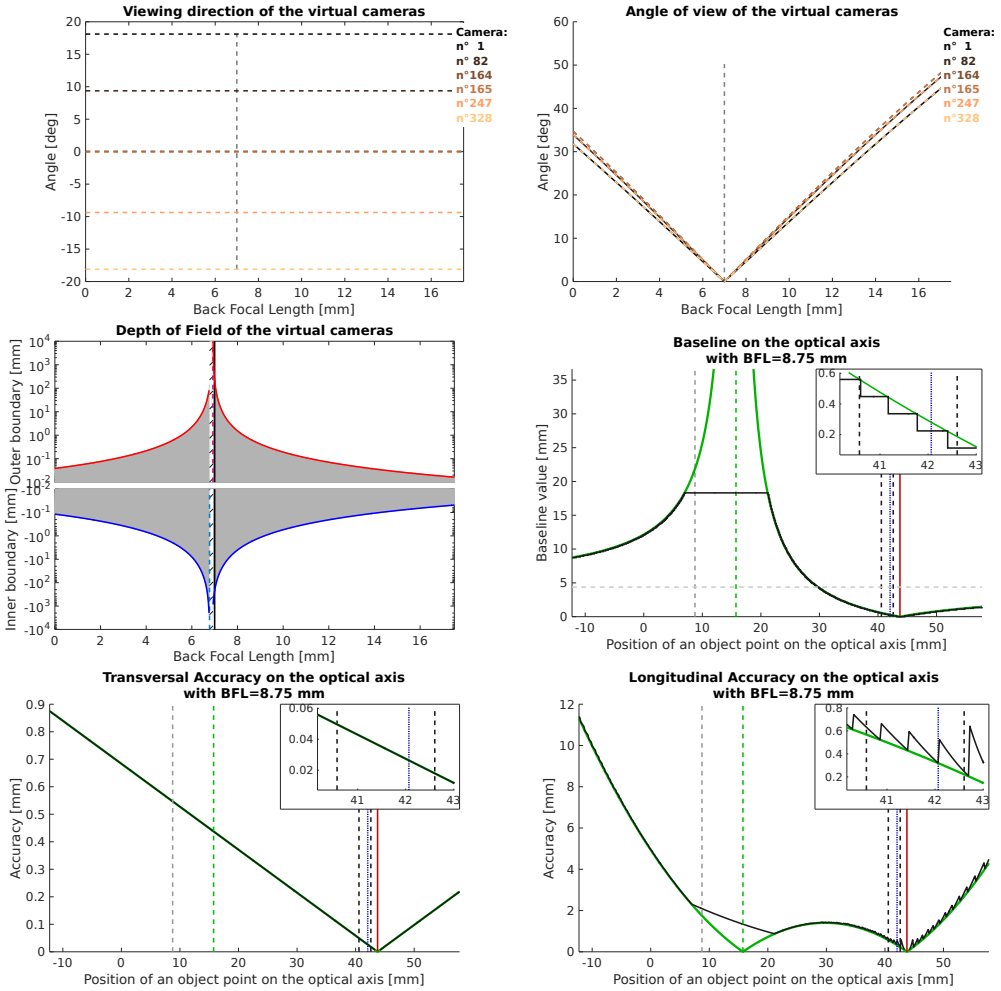


Figure A.10: ECA properties

A.2.2 Lytro-Beta1.00-afocal

Sensor Parameters	Value
Diameter (µm)	1.4000
Pitch (µm)	1.4000
Number	3280
Width (mm)	4.5920
Dist to MLA (µm)	28

Lens Array Parameters	Value
Diameter (mm)	0.0140
Focal length (mm)	0.0280
F-number	2
Pitch (mm)	0.0140
Number	328
Width (mm)	4.5920

Lens Parameters	Value
Diameter (mm)	3.5000
Focal length (mm)	7
F-number	2
Thickness (mm)	0

Figure A.11: System parameters

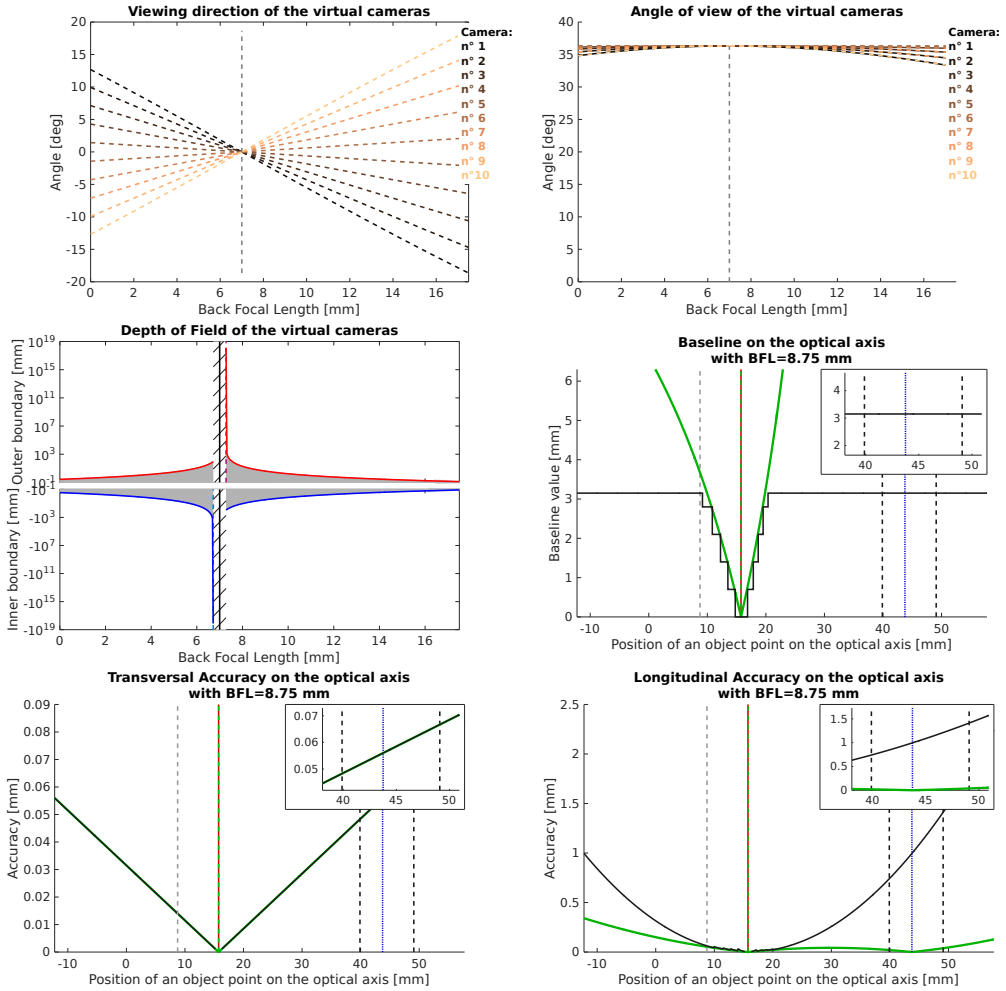


Figure A.12: ECA properties

A.2.3 Lytro-Beta1.00-focused

Sensor Parameters	Value
Diameter (µm)	1.4000
Pitch (µm)	1.4000
Number	3280
Width (mm)	4.5920
Dist to MLA (µm)	28

Lens Array Parameters	Value
Diameter (mm)	0.0140
Focal length (mm)	0.0280
F-number	2
Pitch (mm)	0.0140
Number	328
Width (mm)	4.5920

Lens Parameters	Value
Diameter (mm)	3.5000
Focal length (mm)	7
F-number	2
Thickness (mm)	0

Figure A.13: System parameters

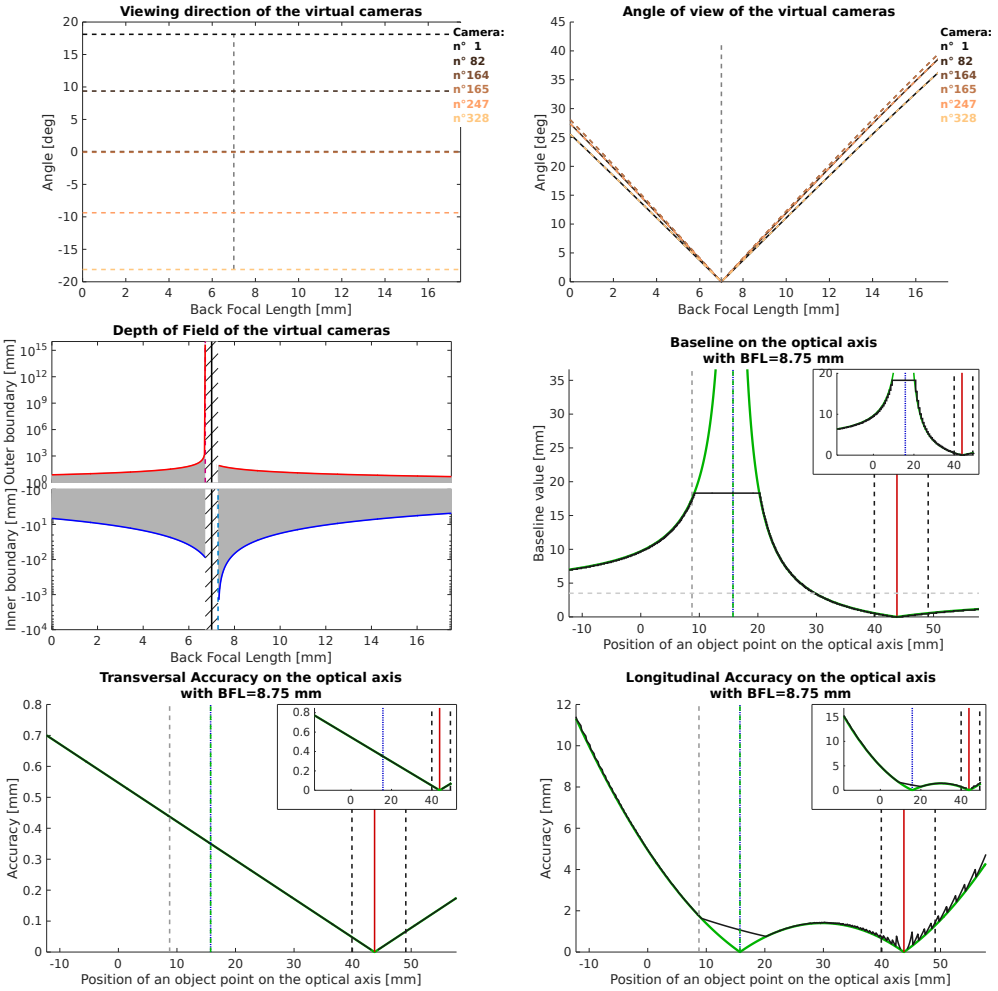


Figure A.14: ECA properties

A.2.4 Lytro-Beta1.20-focused

Sensor Parameters	Value
Diameter (µm)	1.4000
Pitch (µm)	1.4000
Number	3280
Width (mm)	4.5920
Dist to MLA (µm)	33.6000

Lens Array Parameters	Value
Diameter (mm)	0.0140
Focal length (mm)	0.0280
F-number	2
Pitch (mm)	0.0140
Number	328
Width (mm)	4.5920

Lens Parameters	Value
Diameter (mm)	3.5000
Focal length (mm)	7
F-number	2
Thickness (mm)	0

Figure A.15: System parameters

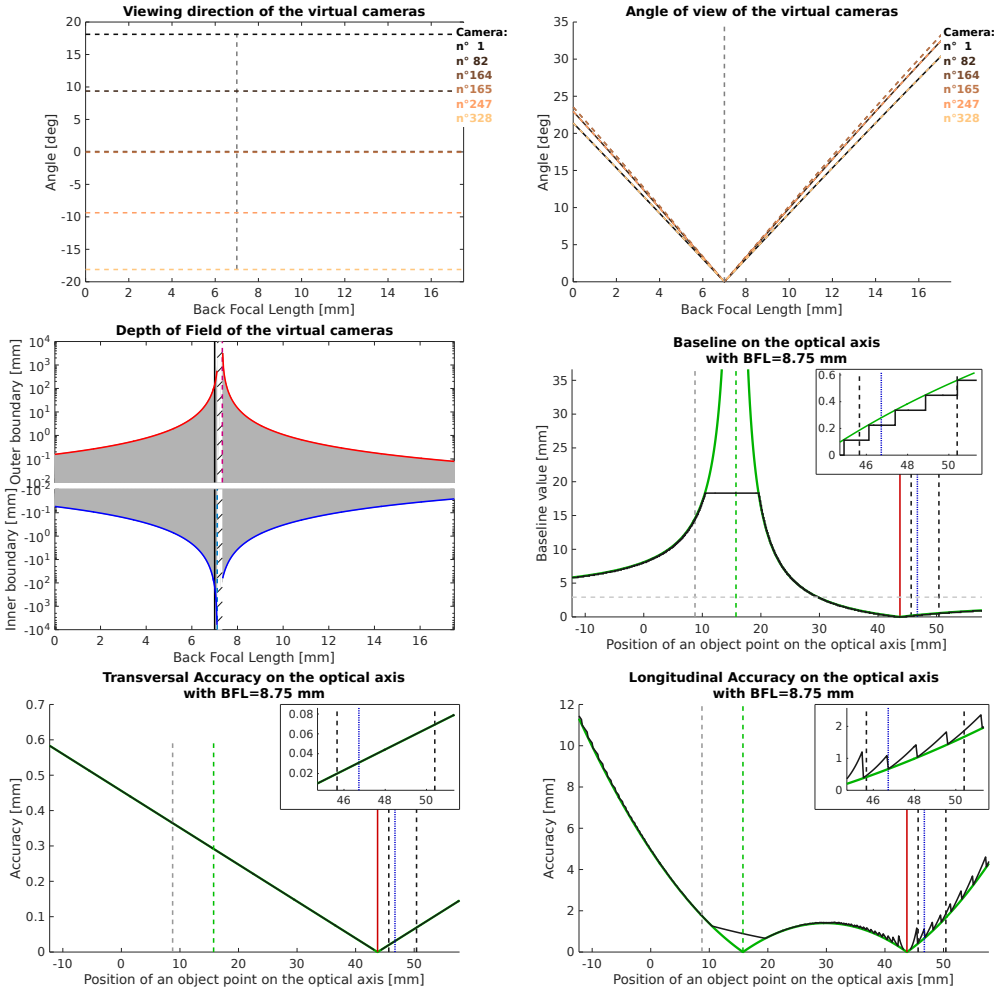


Figure A.16: ECA properties

A.3 System from the Raytrix Camera Patent

From patent US8619177.

A.3.1 Raytrix-Beta0.80-focused

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	3706
Width (mm)	33.3540
Dist to MLA (µm)	420.8000

Lens Array Parameters	Value
Diameter (mm)	0.1530
Focal length (mm)	0.5260
F-number	3.4379
Pitch (mm)	0.1530
Number	218
Width (mm)	33.3540

Lens Parameters	Value
Diameter (mm)	17.5000
Focal length (mm)	35
F-number	2
Thickness (mm)	0

Figure A.17: System parameters

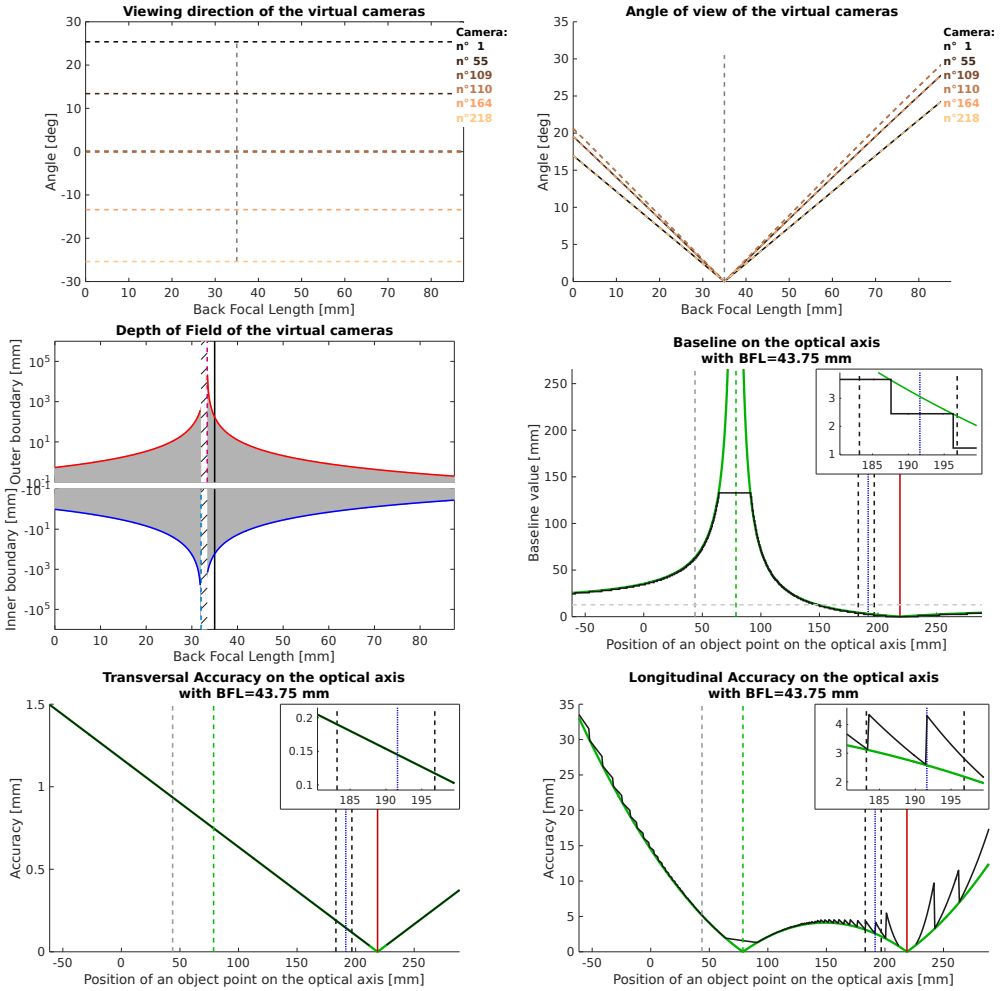


Figure A.18: ECA properties

A.3.2 Raytrix-Beta1.00-afocal

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	3706
Width (mm)	33.3540
Dist to MLA (µm)	526

Lens Array Parameters	Value
Diameter (mm)	0.1530
Focal length (mm)	0.5260
F-number	3.4379
Pitch (mm)	0.1530
Number	218
Width (mm)	33.3540

Lens Parameters	Value
Diameter (mm)	17.5000
Focal length (mm)	35
F-number	2
Thickness (mm)	0

Figure A.19: System parameters

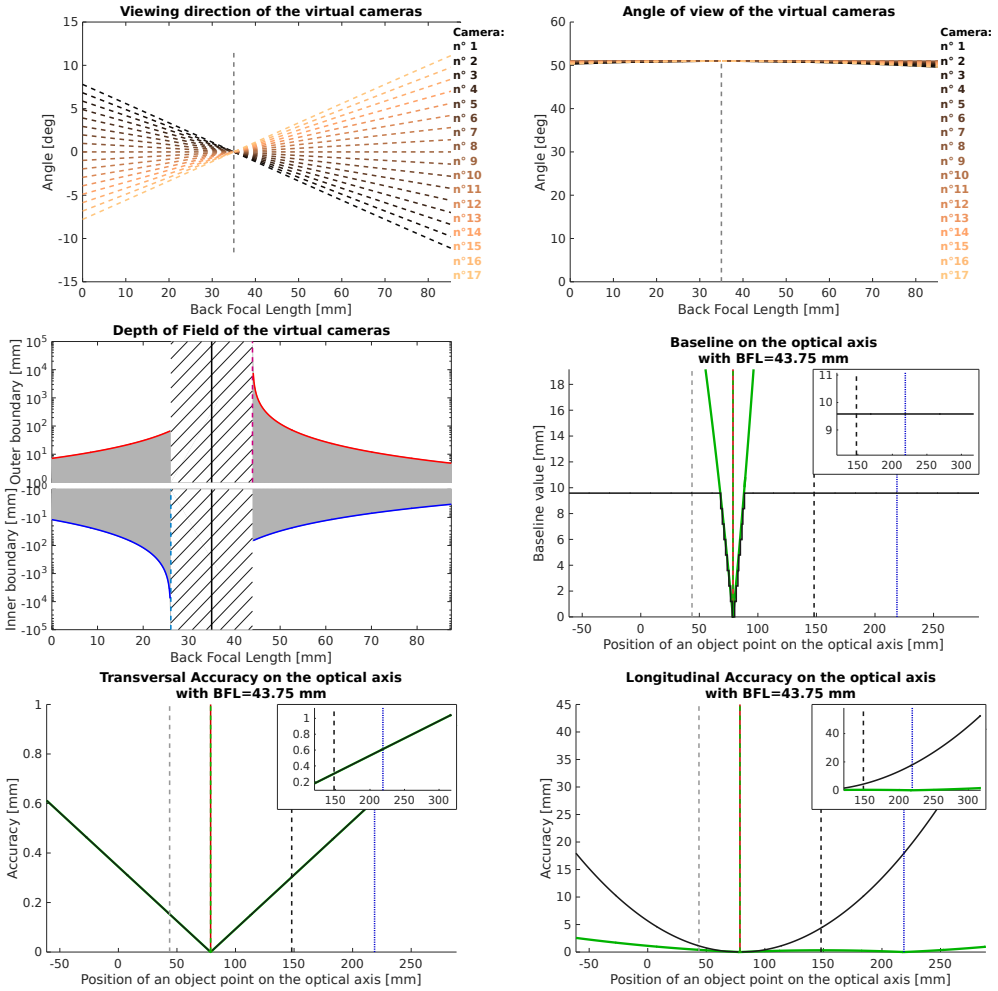


Figure A.20: ECA properties

A.3.3 Raytrix-Beta1.00-focused

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	3706
Width (mm)	33.3540
Dist to MLA (µm)	526

Lens Array Parameters	Value
Diameter (mm)	0.1530
Focal length (mm)	0.5260
F-number	3.4379
Pitch (mm)	0.1530
Number	218
Width (mm)	33.3540

Lens Parameters	Value
Diameter (mm)	17.5000
Focal length (mm)	35
F-number	2
Thickness (mm)	0

Figure A.21: System parameters

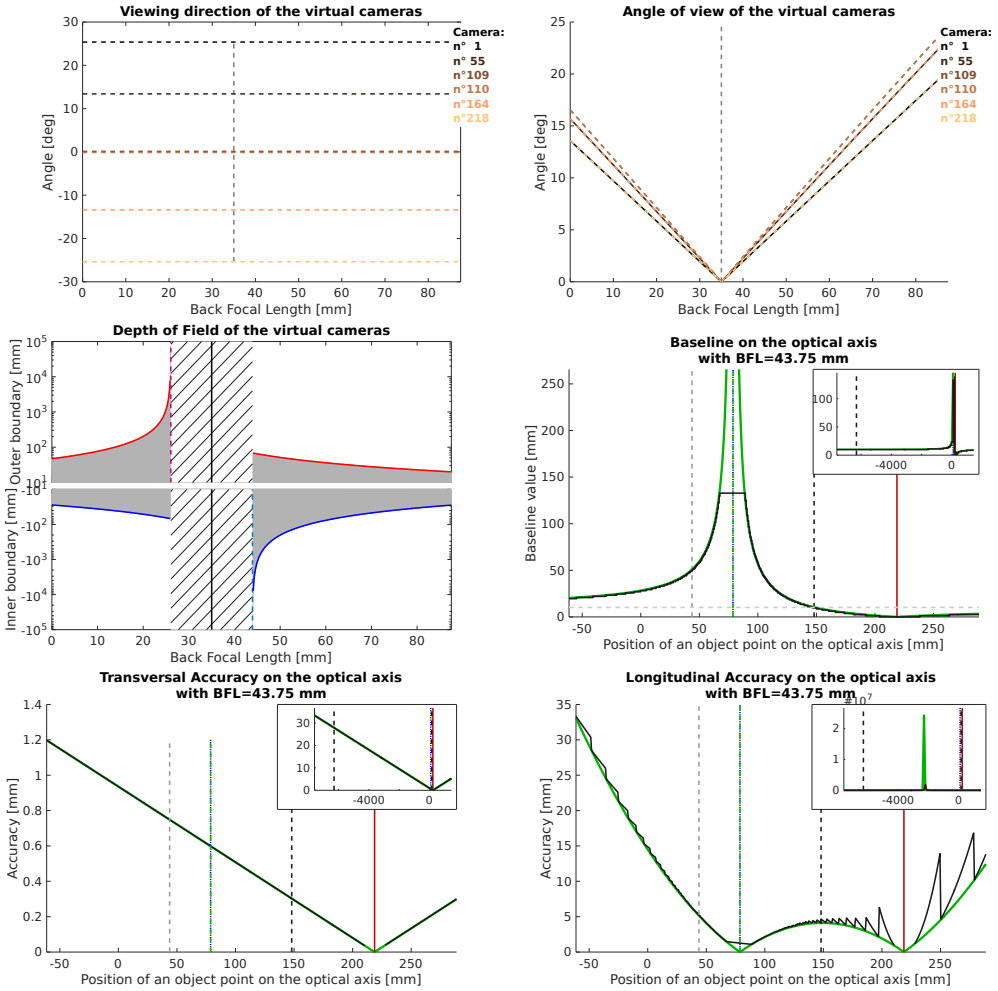


Figure A.22: ECA properties

A.3.4 Raytrix-Beta1.20-focused

Sensor Parameters	Value
Diameter (µm)	9
Pitch (µm)	9
Number	3706
Width (mm)	33.3540
Dist to MLA (µm)	631.2000

Lens Array Parameters	Value
Diameter (mm)	0.1530
Focal length (mm)	0.5260
F-number	3.4379
Pitch (mm)	0.1530
Number	218
Width (mm)	33.3540

Lens Parameters	Value
Diameter (mm)	17.5000
Focal length (mm)	35
F-number	2
Thickness (mm)	0

Figure A.23: System parameters

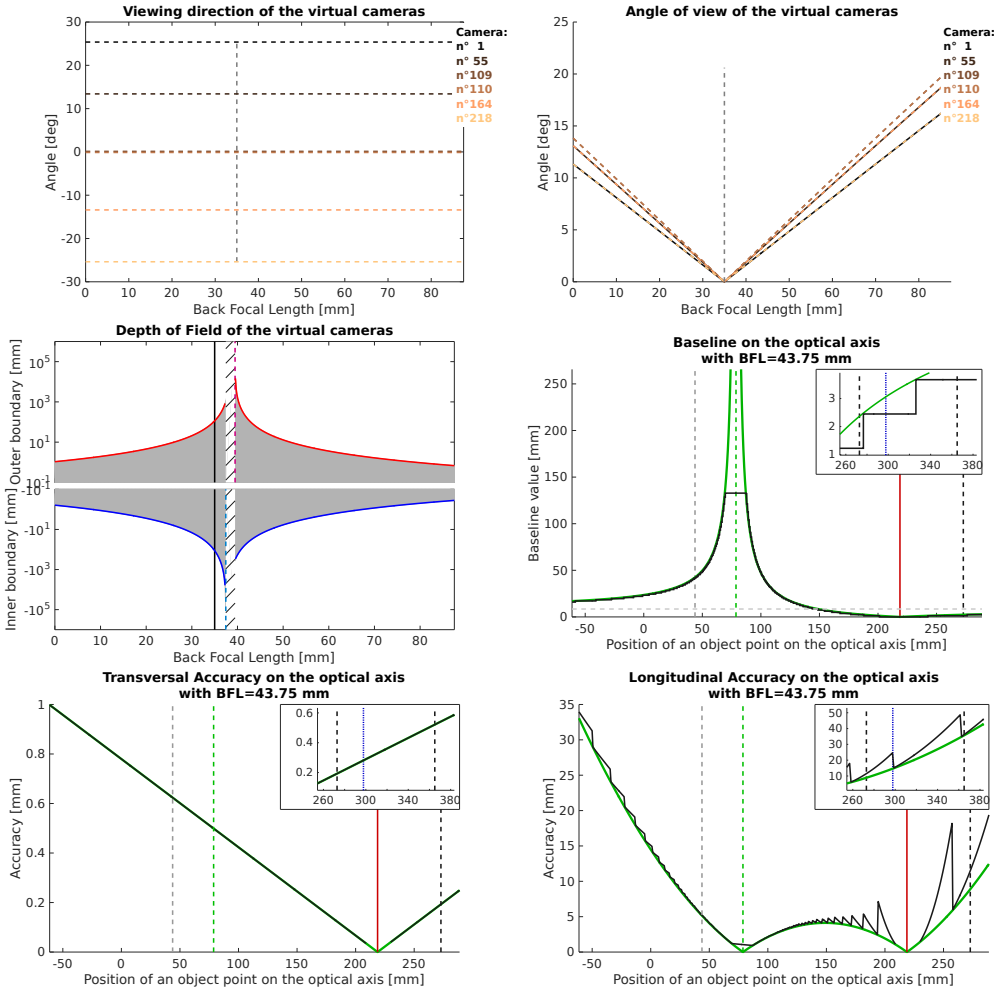


Figure A.24: ECA properties

Bibliography

ADELSON, Edward H and BERGEN, James R, 1991. *The plenoptic function and the elements of early vision*. Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology.

ADELSON, Edward H and WANG, John Y. A., 1992. Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2):99–106.

ATCHESON, Bradley, HEIDE, Felix and HEIDRICH, Wolfgang, 2010. Caltag: High precision fiducial markers for camera calibration. In *VMV*, volume 10, pages 41–48. Citeseer.

BAEK, Byung Joon, LEE, Hyeong Koo, KIM, YoungJin and KIM, TaeChan, 2013. Mirrorless interchangeable-lens light field digital photography camera system. In *Consumer Electronics (ICCE), 2013 IEEE International Conference on*, pages 226–227. IEEE.

BARBER, C Bradford, DOBKIN, David P and HUHDANPAA, Hannu, 1996. The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4):469–483.
URL http://www.qhull.org

BISHOP, Tom E and FAVARO, Paolo, 2012. The light field camera: Extended depth of field, aliasing, and superresolution. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(5):972–986.

BISHOP, Tom E, ZANETTI, Sara and FAVARO, Paolo, 2009. Light field superresolution. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–9. IEEE.

BORN, Max and WOLF, Emil, 1980. *Principles of optics*. Pergamon Press.

BROXTON, Michael, GROSENICK, Logan, YANG, Samuel, COHEN, Noy, ANDALMAN, Aaron, DEISSEROTH, Karl and LEVOY, Marc, 2013. Wave Optics Theory and 3-D Deconvolution for the Light Field Microscope. *Optics Express*, 21(21):25418–25439.

CHAI, Jin-Xiang, TONG, Xin, CHAN, Shing-Chow and SHUM, Heung-Yeung, 2000. Plenoptic sampling. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 307–318. ACM Press/Addison-Wesley Publishing Co.

CHEN, Chih-Chieh, CHIANG, Shih-Chieh Fan, HUANG, Xiao-Xuan, SU, Ming-Shing and LU, Yi-Chang, 2010. Depth estimation of light field data from pinhole-masked dslr cameras. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 1769–1772. IEEE.

CHO, Donghyeon, LEE, Minhaeng, KIM, Sunyeong and TAI, Yu-Wing, 2013. Modeling the calibration pipeline of the lytro camera for high quality light-field image reconstruction. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 3280–3287. IEEE.

COHEN, Noy, YANG, Samuel, ANDALMAN, Aaron, BROXTON, Michael, GROSENICK, Logan, DEISSEROTH, Karl, HOROWITZ, Mark and LEVOY, Marc, 2014. Enhancing the performance of the light field microscope using wavefront coding. *Optics express*, 22(20):24817–24839.

DAMGHANIAN, Mitra, 2013. *The Sampling Pattern Cube: A Framework for Representation and Evaluation of Plenoptic Capturing Systems*. Ph.D. thesis, Mid Sweden University.

DAMGHANIAN, Mitra, OLSSON, Roger and SJOSTROM, Marten, 2012. Extraction of the lateral resolution in a plenoptic camera using the spc model. In *3D Imaging (IC3D), 2012 International Conference on*, pages 1–5. IEEE.

DAMGHANIAN, Mitra, OLSSON, Roger and SJÖSTRÖM, Marten, 2014. Performance analysis in lytro camera: Empirical and model based approaches to assess refocusing quality. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 559–563. IEEE.

DANSEREAU, Donald G, BONGIORNO, Daniel L, PIZARRO, Oscar and WILLIAMS, Stefan B, 2013a. Light field image denoising using a linear 4d frequency-hyperfan all-in-focus filter. In *Computational Imaging*, page 86570P.

DANSEREAU, Donald G, PIZARRO, Oscar and WILLIAMS, Stefan B, 2013b. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1027–1034. IEEE.

DRAZIC, Valter, 2010. Optimal depth resolution in plenoptic imaging. In *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, pages 1588–1593. IEEE.

FENG, Tao and SHUM, Heung-Yeung, 2000. An optical analysis of light field rendering. In *Proceedings of Fifth Asian Conference on Computer Vision*, pages 394–399.

FUCHS, Martin, KÄCHELE, Markus and RUSINKIEWICZ, Szymon, 2013. Design and fabrication of faceted mirror arrays for light field capture. In *Computer Graphics Forum*, volume 32, pages 246–257. Wiley Online Library.

GEORGIEV, Todor and INTWALA, Chintan, 2003. Light field camera design for integral view photography. Technical report, Adobe System, Inc.

GEORGIEV, Todor and INTWALA, Chintan, 2006. Light field camera design for integral view photography. Technical report, Adobe System, Inc.

GEORGIEV, Todor, INTWALA, Chintan, BABAKAN, Sevkit and LUMSDAINE, Andrew, 2008. Unified frequency domain analysis of lightfield cameras. *Computer Vision–ECCV 2008*, pages 224–237.

GEORGIEV, Todor, YU, Zhan, LUMSDAINE, Andrew and GOMA, Sergio, 2013. Lytro camera technology: theory, algorithms, performance analysis. In *IS&T/SPIE Electronic Imaging*, pages 86671J–86671J. International Society for Optics and Photonics.

GEORGIEV, Todor, ZHENG, Ke Colin, CURLESS, Brian, SALESIN, David, NAYAR, Shree and INTWALA, Chintan, 2006. Spatio-angular resolution tradeoffs in integral photography. *Rendering Techniques*, 2006:263–272.

GORTLER, Steven J, GRZESZCZUK, Radek, SZELISKI, Richard and COHEN, Michael F, 1996. The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 43–54. ACM.

GROSSBERG, Michael D. and NAYAR, Shree K., 2005. The raxel imaging model and ray-based calibration. *International Journal of Computer Vision*, 61(2):119–137. doi:10.1023/B:VISI.0000043754.56350.10.
URL http://dx.doi.org/10.1023/B:VISI.0000043754.56350.10

HAHNE, Christopher, 2016. *The standard plenoptic camera: applications of a geometrical light field model*. Ph.D. thesis.

HAHNE, Christopher, AGGOUN, Amar, HAXHA, Shyqyri, VELISAVLJEVIC, Vladan and FERNÁNDEZ, Juan Carlos Jácome, 2014a. Light field geometry of a standard plenoptic camera. *Optics Express*, 22(22):26659–26673.

HAHNE, Christopher, AGGOUN, Amar, HAXHA, Shyqyri, VELISAVLJEVIC, Vladan and FERNANDEZ, Juan CJ, 2014b. Baseline of virtual cameras acquired

by a standard plenoptic camera setup. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2014*, pages 1–3. IEEE.

HEINZE, Christian, SPYROPOULOS, Stefano, HUSSMANN, Stephan and PERWASS, Christian, 2015. Automated robust metric calibration of multi-focus plenoptic cameras. In *Instrumentation and Measurement Technology Conference (I2MTC), 2015 IEEE International*, pages 2038–2043. IEEE.

IHRKE, Ivo, RESHETOUSKI, Ilya, MANAKOV, Alkhazur, TEVS, Art, WAND, Michael and SEIDEL, Hans-Peter, 2012. A kaleidoscopic approach to surround geometry and reflectance acquisition. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pages 29–36. IEEE.

IHRKE, Ivo, RESTREPO, John and MIGNARD-DEBISE, Lois, 2016. Principles of light field imaging: Briefly revisiting 25 years of research. *IEEE Signal Processing Magazine*, 33(5):59–69. doi:10.1109/MSP.2016.2582220.

IHRKE, Ivo, STICH, Timo, GOTTSCHLICH, Heiko, MAGNOR, Marcus and SEIDEL, Hans-Peter, 2008. Fast incident light field acquisition and rendering. *Journal of WSCG (WSCG'08)*, 16(1-3):25–32.

ISERINGHAUSEN, Julian, GOLDLÜCKE, Bastian, PESHEVA, Nina, ILIEV, Stanimir, WENDER, Alexander, FUCHS, Martin and HULLIN, Matthias B, 2017. 4d imaging through spray-on optics. *ACM Transactions on Graphics (TOG)*, 36(4):35.

IVES, Frederic E, 1903. Parallax stereogram and process of making same. US Patent 725,567.

JOHANNSEN, Ole, HEINZE, Christian, GOLDLUECKE, Bastian and PERWASS, Christian, 2013. On the calibration of focused plenoptic cameras. In *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, pages 302–317. Springer.

KRISHNAN, Sreenath, 2016. Capturing wide field of view light fields using spherical mirrors.

KUČERA, Jan, 2014. Lytro meltdown. http://optics.miloush.net/lytro/Default.aspx. Accessed: 2017-06-01.

LAM, Edmund Y, 2015. Computational photography with plenoptic camera and light field capture: tutorial. *JOSA A*, 32(11):2021–2032.

LANMAN, Douglas, CRISPELL, Daniel, WACHS, Megan and TAUBIN, Gabriel, 2006. Spherical catadioptric arrays: Construction, multi-view geometry, and calibration. In *3D Data Processing, Visualization, and Transmission, Third International Symposium on*, pages 81–88. IEEE.

LEE, Joon-Jae, SHIN, Donghak, LEE, Byung-Gook and YOO, Hoon, 2012. 3D Optical Microscopy Method based on Synthetic Aperture Integral Imaging. *3D Research*, 3(4):1–6.

LEE, Sungkil, EISEMANN, Elmar and SEIDEL, Hans-Peter, 2010. Real-time lens blur effects and focus control. *ACM Transactions on Graphics (TOG)*, 29(4):65.

LEVIN, Anat, FREEMAN, William T and DURAND, Frédo, 2008. Understanding camera trade-offs through a bayesian analysis of light field projections. In *European Conference on Computer Vision*, pages 88–101. Springer.

LEVOY, Marc, 2006. Light fields and computational imaging. *Computer*, (8):46–55.

LEVOY, Marc, CHEN, Billy, VAISH, Vaibhav, HOROWITZ, Mark, MCDOWALL, Ian and BOLAS, Mark, 2004. Synthetic aperture confocal imaging. In *ACM Transactions on Graphics (ToG)*, volume 23, pages 825–834. ACM.

LEVOY, Marc and HANRAHAN, Pat, 1996. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42. ACM.

LEVOY, Marc, NG, Ren, ADAMS, Andrew, FOOTER, Matthew and HOROWITZ, Mark, 2006. Light field microscopy. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 924–934. ACM.

LEVOY, Marc, ZHANG, Zhengyun and MCDOWALL, Ian, 2009. Recording and Controlling the 4D Light Field in a Microscope using Microlens Arrays. *Journal of Microscopy*, 235:144–162.

LIANG, Chia-Kai, LIN, Tai-Hsu, WONG, Bing-Yi, LIU, Chi and CHEN, Homer H, 2008. Programmable aperture photography: multiplexed light field acquisition. *ACM Transactions on Graphics (TOG)*, 27(3):55.

LIANG, Chia-Kai and RAMAMOORTHI, Ravi, 2015. A light transport framework for lenslet light field cameras. *ACM Transactions on Graphics (TOG)*, 34(2):16.

LIGHT, Inc., 2017. Light website. https://light.co/. Accessed: 2017-08-21.

LIN, Xing, WU, Jiamin and DAI, Qionghai, 2015. Camera array based light field microscopy. In *Optical Molecular Probes, Imaging and Drug Delivery*, pages JT3A–48. Optical Society of America.

LIPPMANN, Gabriel, 1908. Epreuves reversibles donnant la sensation du relief. *J. Phys. Theor. Appl.*, 7(1):821–825.

LOHMANN, Adolf W., DORSCH, Rainer G., MENDLOVIC, David, ZALEVSKY, Zeev and FERREIRA, Carlos, 1996. Space-Bandwidth Product of Optical Signals and Systems. *Journal of the Optical Society of America A*, 13(3):470–473.

LU, Chien-Hung, MUENZEL, Stefan and FLEISCHER, Jason, 2013. High-resolution Light-Field Microscopy. In *Computational Optical Sensing and Imaging*, pages CTh3B–2. Optical Society of America.

LUMSDAINE, Andrew and GEORGIEV, Todor, 2008. Full resolution lightfield rendering. *Indiana University and Adobe Systems, Tech. Rep*.

LUMSDAINE, Andrew and GEORGIEV, Todor, 2009. The focused plenoptic camera. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–8. IEEE.

LYTRO, Inc., 2017. Lytro website. https://www.lytro.com/. Accessed: 2017-08-21.

MAGNOR, Marcus and GIROD, Bernd, 2000. Data compression for light-field rendering. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(3):338–343.

MANAKOV, Alkhazur, RESTREPO, John, KLEHM, Oliver, HEGEDÜS, Ramon, SEIDEL, Hans-Peter, EISEMANN, Elmar and IHRKE, Ivo, 2013. A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging. *ACM Trans. on Graphics (SIGGRAPH'13)*, 32(4):47:1 – 47:14.

MIGNARD-DEBISE, Lois and IHRKE, Ivo, 2015. Light-field microscopy with a consumer light-field camera. In *2015 International Conference on 3D Vision*, pages 335–343. IEEE. doi:10.1109/3DV.2015.45.

MIGNARD-DEBISE, Lois, RESTREPO, John and IHRKE, Ivo, 2017. A unifying first-order model for light-field cameras: The equivalent camera array. *IEEE Transactions on Computational Imaging*, 3(4):798–810. doi:10.1109/TCI.2017.2699427.

MURPHY, Douglas B and DAVIDSON, Michael W, 2012. *Fundamentals of Light Microscopy and Electronic Imaging*. John Wiley & Sons.

NG, Ren, 2005. Fourier slice photography. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 735–744. ACM.

NG, Ren, 2006. *Digital Light Field Photography*. Ph.D. thesis, Stanford University, Department of Computer Science.

NG, Ren and HANRAHAN, Pat, 2007. Digital correction of lens aberrations in light field photography. volume 6342, pages 63421E–63421E–14. doi:10.1117/12.692290.
URL http://dx.doi.org/10.1117/12.692290

NG, Ren, LEVOY, Marc, BRÉDIF, Mathieu, DUVAL, Gene, HOROWITZ, Mark and HANRAHAN, Pat, 2005a. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11).

NG, Ren, LEVOY, Marc, BRÉDIF, Mathieu, DUVAL, Gene, HOROWITZ, Mark and HANRAHAN, Pat, 2005b. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11).

NOVIKOVA, Tatiana, PIERANGELO, Angelo, MARTINO, Antonello De, BENALI, Abdelali and VALIDIRE, Pierre, 2012. Polarimetric Imaging for Cancer Diagnosis and Staging. *Opt. Photon. News*, 23(10):26–33.

OLDENBOURG, Rudolf, 2008. Polarized Light Field Microscopy: An Analytical Method using a Microlens Array to Simultaneously Capture both Conoscopic and Orthoscopic Views of Birefringent Objects. *Journal of Microscopy*, 231(3):419–432.

O'SHEA, Donald C, 1985. *Elements of modern optical design*. Wiley.

PACHECO, Shaun and LIANG, Rongguang, 2014. Snapshot, reconfigurable multispectral and multi-polarization telecentric imaging system. *Optics express*, 22(13):16377–16385.

PÉREZ, Fernando, PÉREZ, Alejandro, RODRÍGUEZ, Manuel and MAGDALENO, Eduardo, 2012. Fourier slice super-resolution in plenoptic cameras. In *Computational Photography (ICCP), 2012 IEEE International Conference on*, pages 1–11. IEEE.

PERWASS, Christian and WIETZKE, Lennart, 2012. Single lens 3d-camera with extended depth-of-field. In *IS&T/SPIE Electronic Imaging*, pages 829108–829108. International Society for Optics and Photonics.

PLATT, B. C. and SHACK, R., 2001. History and Principles of Shack-Hartmann Wavefront Sensing. *Journal of Refractive Surgery*, 17(Sept/Oct):573–577. doi:10.3928/1081-597X-20010901-13.

RANDER, Peter, NARAYANAN, PJ and KANADE, Takeo, 1997. Virtualized reality: Constructing time-varying virtual worlds from real world events. In *Visualization'97., Proceedings*, pages 277–283. IEEE.

RAYTRIX, Inc., 2017. Raytrix website. http://raytrix.de/. Accessed: 2017-08-21.

REDDY, Dikpal, BAI, Jiamin and RAMAMOORTHI, Ravi, 2013. External mask based depth and light field camera. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 37–44.

RESHETOUSKI, Ilya, MANAKOV, Alkhazur, SEIDEL, Hans-Peter and IHRKE, Ivo, 2011. Three-dimensional kaleidoscopic imaging. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 353–360. IEEE.

SINGER, Wolfgang, TOTZECK, Michael and GROSS, Herbert, 2006. *Handbook of optical systems, physical image formation*, volume 2. John Wiley & Sons.

TAGUCHI, Yuichi, AGRAWAL, Amit, RAMALINGAM, Srikumar and VEERARAGHAVAN, Ashok, 2010a. Axial light field for curved mirrors: Reflect your perspective, widen your view. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 499–506. IEEE.

TAGUCHI, Yuichi, AGRAWAL, Amit, VEERARAGHAVAN, Ashok, RAMALINGAM, Srikumar and RASKAR, Ramesh, 2010b. Axial-cones: Modeling spherical catadioptric cameras for wide-angle light field rendering. *ACM Trans. Graph.*, 29(6):172–1.

TAGUCHI, Yuichi, AGRAWAL, Amit, VEERARAGHAVAN, Ashok, RAMALINGAM, Srikumar and RASKAR, Ramesh, 2010c. Axial-cones: Modeling spherical catadioptric cameras for wide-angle light field rendering. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2010)*, 29(6):172:1–172:8.

TAKAHASHI, Keita, NAEMURA, Takeshi and HAMSHIMA, H, 2003. Depth of field in light field rendering. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 1, pages I–409. IEEE.

THOMAS, MG, MONTILLA, I, MARICHAL-HERNANDEZ, JG, FERNANDEZ-VALDIVIA, JJ, TRUJILLO-SEVILLA, JM and RODRIGUEZ-RAMOS, JM, 2013. Depth Map Extraction from Light Field Microscopes. In *12th Workshop on Information Optics (WIO)*, pages 1–3. IEEE.

THOMASON, Chelsea, THUROW, B and FAHRINGER, T, 2014. Calibration of a microlens array for a plenoptic camera. In *52nd Aerospace Sciences Meeting: American Institute of Aeronautics and Astronautics*, page 0396.

TORRE, Amalia, 2005. *Linear ray and wave optics in phase space: Bridging ray and wave optics via the Wigner phase-space picture*. Elsevier.

TUROLA, Massimo, 2016. *Investigation of plenoptic imaging systems: a wave optics approach*. Ph.D. thesis, City University London.

TUROLA, Massimo and GRUPPETTA, Steve, 2014. Wave optics simulations of a focused plenoptic system. In *Frontiers in Optics*, pages JTu3A–24. Optical Society of America.

UEDA, Kensuke, KOIKE, Takafumi, TAKAHASHI, Keita and NAEMURA, Takeshi, 2008a. Adaptive integral photography imaging with variable-focus lens array. *Proc SPIE: Stereoscopic Displays and Applications XIX*, pages 68031A–9.

UEDA, Kensuke, LEE, Dongha, KOIKE, Takafumi, TAKAHASHI, Keita and NAEMURA, Takeshi, 2008b. Multi-focal compound eye: liquid lens array for computational photography. In *ACM SIGGRAPH 2008 new tech demos*, page 28. ACM.

UNGER, Jonas, WENGER, Andreas, HAWKINS, Tim, GARDNER, Andrew and DEBEVEC, Paul, 2003. Capturing and rendering with incident light fields. Technical report, DTIC Document.

VEERARAGHAVAN, Ashok, RASKAR, Ramesh, AGRAWAL, Amit, CHELLAPPA, Rama, MOHAN, Ankit and TUMBLIN, Jack, 2008. Non-Refractive Modulators for Encoding and Capturing Scene Appearance and Depth. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8.

VEERARAGHAVAN, Ashok, RASKAR, Ramesh, AGRAWAL, Amit, MOHAN, Ankit and TUMBLIN, Jack, 2007. Dappled Photography: Mask Enhanced Cameras For Heterodyned Light Fields and Coded Aperture Refocussing. *ACM Trans. on Graphics (TOG)*, 26:69.

VENKATARAMAN, Kartik, LELESCU, Dan, DUPARRÉ, Jacques, MCMAHON, Andrew, MOLINA, Gabriel, CHATTERJEE, Priyam, MULLIS, Robert and NAYAR, Shree, 2013. Picam: An ultra-thin high performance monolithic camera array. *ACM Transactions on Graphics (TOG)*, 32(6):166.

WANNER, Sven and GOLDLUECKE, Bastian, 2012. Spatial and angular variational super-resolution of 4d light fields. In *European Conference on Computer Vision*, pages 608–621. Springer.

WEI, Li-Yi, LIANG, Chia-Kai, MYHRE, Graham, PITTS, Colvin and AKELEY, Kurt, 2015. Improving light field camera sample design with irregularity and aberration. *ACM Transactions on Graphics (TOG)*, 34(4):152.

WENDER, Alexander, ISERINGHAUSEN, Julian, GOLDLUECKE, Bastian, FUCHS, Martin and HULLIN, Matthias B., 2015. Light field imaging through household optics. In *Vision, Modeling & Visualization*, pages 159–166. The Eurographics Association. ISBN 978-3-905674-95-8. doi:10.2312/vmv.20151271.

WETZSTEIN, Gordon, IHRKE, Ivo, LANMAN, Douglas and HEIDRICH, Wolfgang, 2011. Computational plenoptic imaging. In *Computer Graphics Forum*, volume 30, pages 2397–2426. Wiley Online Library.

WILBURN, Bennett, JOSHI, Neel, VAISH, Vaibhav, LEVOY, Marc and HOROWITZ, Mark, 2004. High-speed videography using a dense camera array. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–294. IEEE.

WILBURN, Bennett, JOSHI, Neel, VAISH, Vaibhav, TALVALA, Eino-Ville, ANTUNEZ, Emilio, BARTH, Adam, ADAMS, Andrew, HOROWITZ, Mark and LEVOY, Marc, 2005. High performance imaging using large camera arrays. *ACM Transactions on Graphics (TOG)*, 24(3):765–776.

WILT, Brian A, BURNS, Laurie D, HO, Eric Tatt Wei, GHOSH, Kunal K, MUKAMEL, Eran A and SCHNITZER, Mark J, 2009. Advances in Light Microscopy for Neuroscience. *Annual Review of Neuroscience*, 32:435.

XU, Yichao, MAENO, Kazuki, NAGAHARA, Hajime and TANIGUCHI, Rin-ichiro, 2014. Mobile camera array calibration for light field acquisition. *CoRR*, abs/1407.4206.
URL http://arxiv.org/abs/1407.4206

XU, Yichao, MAENO, Kazuki, NAGAHARA, Hajime and TANIGUCHI, Rin-ichiro, 2015. Camera array calibration for light field acquisition. *Frontiers of Computer Science*, 9(5):691–702. doi:10.1007/s11704-015-4237-4.
URL http://dx.doi.org/10.1007/s11704-015-4237-4

YAMAMOTO, Tomoyuki, KOJIMA, Masaru and NAEMURA, Takeshi, 2004. Liflet: Light field live with thousands of lenslets. In *ACM SIGGRAPH 2004 Emerging technologies*, page 16. ACM.

YAMAMOTO, Tomoyuki and NAEMURA, Takeshi, 2004. Real-time capturing and interactive synthesis of 3d scenes using integral photography. In *Proceedings of SPIE*, volume 5291, pages 155–166.

YANG, Jason C, 2000. *A light field camera for image based rendering*. Ph.D. thesis, Massachusetts Institute of Technology.

YANG, Jason C, EVERETT, Matthew, BUEHLER, Chris and MCMILLAN, Leonard, 2002. A real-time distributed light field camera. In *Proceedings of the 13th Eurographics workshop on Rendering*, pages 77–86. Eurographics Association.

ZHENG, Guoan, HORSTMEYER, Roarke and YANG, Changhuei, 2013. Wide-field, High-resolution Fourier Ptychographic Microscopy. *Nature Photonics*, 7(9):739–745.

ZHOU, Changyin and NAYAR, Shree K, 2011. Computational cameras: convergence of optics and processing. *IEEE Transactions on Image Processing*, 20(12):3322–3340.