

This is the author's copy of the publication as archived with the DLR's electronic library at <http://elib.dlr.de>. Please consult the original publication for citation.

Benefits of plenoptic cameras for robot vision during close range on-orbit servicing maneuvers

Lingenauber, Martin; Strobl, Klaus H.; Oumer, Nassir W.; Kriegel, Simon

Copyright Notice

©2017 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Citation Notice

```
@INPROCEEDINGS{lingenauber2017benefits,
author    = {{L}ingenauber, {M}artin and {S}trobl, {K}laus {H.} and {O}umer, {N}assir {W.} and {K}riegel, {S}imon},
title     = {{B}enefits of {P}lenoptic {C}ameras for {R}obot {V}ision {D}uring {C}lose {R}ange {O}n-{O}rbit {S}ervicing {M}aneuvers},
booktitle = {{2017 IEEE Aerospace Conference},
title     = {Benefits of plenoptic cameras for robot vision during close range on-orbit servicing maneuvers},
year      = {2017},
pages     = {1-18},
doi       = {10.1109/AERO.2017.7943666},
month     = {March},
}
```

Benefits of Plenoptic Cameras for Robot Vision During Close Range On-Orbit Servicing Maneuvers

Martin Lingenauber
 German Aerospace Center (DLR)
 Institute of Robotics and Mechatronics
 Department of Perception and Cognition
 Münchener Str. 20
 82234 Wessling, Germany
 +49-(0)-8153-28-1595
 martin.lingenauber@dlr.de

Nassir W. Oumer
 German Aerospace Center (DLR)
 Institute of Robotics and Mechatronics
 Department of Perception and Cognition
 Münchener Str. 20
 82234 Wessling, Germany
 +49-(0)-8153-28-1478
 nassir.oumer@dlr.de

Klaus H. Strobl
 German Aerospace Center (DLR)
 Institute of Robotics and Mechatronics
 Department of Perception and Cognition
 Münchener Str. 20
 82234 Wessling, Germany
 +49-(0)-8153-28-2482
 klaus.strobl@dlr.de

Simon Kriegel
 German Aerospace Center (DLR)
 Institute of Robotics and Mechatronics
 Department of Perception and Cognition
 Münchener Str. 20
 82234 Wessling, Germany
 +49-(0)-8153-28-1045
 simon.kriegel@dlr.de

Abstract—This paper discusses the potential benefits of plenoptic cameras for robot vision during on-orbit servicing missions. Robot vision is essential for the accurate and reliable positioning of a robotic arm with millimeter accuracy during tasks such as grasping, inspection or repair that are performed in close range to a client satellite. Our discussion of the plenoptic camera technology provides an overview of the conceptual advantages for robot vision with regard to the conditions during an on-orbit servicing mission. A plenoptic camera, also known as light field camera, is basically a conventional camera system equipped with an additional array of lenslets, the micro lens array, at a distance of a few micrometers in front of the camera sensor. Due to the micro lens array it is possible to record not only the incidence location of a light ray but also its incidence direction on the sensor, resulting in a 4-D data set known as a light field. The 4-D light field allows to derive regular 2-D intensity images with a significantly extended depth of field compared to a conventional camera. This results in a set of advantages, such as software based refocusing or increased image quality in low light conditions due to recording with an optimal aperture while maintaining an extended depth of field. Additionally, the parallax between corresponding lenslets allows to derive 3-D depth images from the same light field and therefore to substitute a stereo vision system with a single camera. Given the conceptual advantages, we investigate what can be expected from plenoptic cameras during close range robotic operations in the course of an on-orbit servicing mission. This includes topics such as image quality, extension of the depth of field, 3-D depth map generation and low light capabilities. Our discussion is backed by image sequences for an on-orbit servicing scenario that were recorded in a representative laboratory environment with simulated in-orbit illumination conditions. We mounted a plenoptic camera on a robot arm and performed an approach trajectory from up to 2 m towards a full-scale satellite mockup. Using these images, we investigated how the light field processing performs, e.g. in terms of depth of field extension, image quality and depth estimation. We were also able to show the applicability of images derived from light fields for the purpose of the visual based pose estimation of a target point.

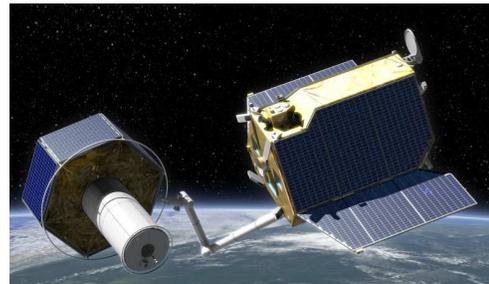


Figure 1: Rendering of an example configuration from the German DEOS on-orbit servicing mission study. The servicer on the right performs a close range maneuver as it reaches out to the client satellite in order to grasp it. (Image credit: DLR)

TABLE OF CONTENTS	
1. INTRODUCTION	1

2. RELATED WORK	2
3. LIGHT FIELD THEORY	3
4. PLENOPTIC CAMERAS	3
5. PLENOPTIC CAMERAS FOR SPACE	7
6. EXPERIMENTS	8
7. RESULTS	9
8. DISCUSSION	11
9. CONCLUSION	14
ACKNOWLEDGMENTS	15
REFERENCES	15
BIOGRAPHY	18

1. INTRODUCTION

This paper introduces the use of plenoptic cameras for robot vision during On-Orbit Servicing (OOS) operations with a robotic arm in close range to a client satellite, i.e. from a few meters to contact as shown exemplarily in Fig. 1. The tasks during close range operations are amongst others grasping, docking, inspection, repair, or maintenance. Although it is desired to achieve an autonomous operation of the robot arm, depending on the specific mission, teleoperation can also be

applied. These tasks require movements of the arm in close range to the satellite with positioning accuracy of sometimes a few millimeters in order to guarantee a collision free and reliable operation of the robot arm. The goal is to avoid any generation of new space debris, including abrasion of small chunks of material as it could happen due to undesired contact with the satellite's surface with the robotic arm or its tool.

Due to the necessary resolution when targeting small structures or attachments on the client, passive vision sensors, i.e. cameras, are the preferred choice to support robot arm operations. The images should be well in focus over the complete working range of the robot arm in order to allow for reliable tracking during the whole operation, i.e. the Depth of Field (DoF) is required to be relatively large. This can require to sacrifice light throughput, in terms of the lens aperture opening, for image sharpness, thus leading to a low image brightness. The latter might be alleviated by extending the exposure time or by increasing the sensor sensitivity or both, which might either lead to motion blur in the images or to an increase of the noise level, respectively. Hence, a decoupling of the DoF from the aperture size is desired.

In order to relax those limitations, the authors propose the use of plenoptic cameras. They provide an extended DoF with a more open aperture, thus allowing sharp images over a larger range with improved light gathering capabilities. With a single exposure of a plenoptic camera, it is possible to create 2-D images and 3-D range images at the same time [1]. Therefore, a plenoptic camera is a passive, monocular 3-D sensor that can potentially replace a stereo camera system. A plenoptic camera can be thought of as a conventional camera with an additional matrix of lenslets, the so called Micro Lens Array (MLA). Each lenslet has a diameter in the range of tens of micrometers and the MLA is mounted within a small distance in front of the camera sensor. An off-the-shelf main lens is used for the plenoptic camera without any further modification [1–4]. The camera records a light field [5, 6], a 4-D data structure that contains not only the location of arrival of each recorded light ray on the sensor but also the direction of arrival within the camera, i.e. after the main lens. The parallax between the lenslets of the MLA in combination with the 4-D light field data enable several advantages over conventional cameras, whose usability is worth to be investigated. First, a plenoptic camera can achieve images with an extended DoF while maintaining the optimum aperture. This results in better quality images and the increased gathering of light [7]. Second, due to the parallax between the lenslets it is possible to record 3-D range images with a single exposure. As 2-D images and 3-D images are derived from the same light field, a direct mapping of texture and geometry is possible. Third, the extended DoF and the light field data structure allow to shift certain hardware operations to software. One example is the purely digital refocusing to different focal planes [3] which might be beneficial during teleoperation scenarios to give the operator a better impression.

In general, plenoptic cameras can be used as conventional cameras, but the mentioned advantages come into play most effectively in the close range domain. The DoF extension is more significant as well as the 3-D capabilities [1]. By concept, plenoptic cameras offer valuable improvements for camera based vision in orbit, but they come with the costs of higher resolution sensors and higher computational demands. In the authors' opinion, plenoptic cameras can be made space ready with a relatively low amount of additional effort given that it is a new technology and also given the presented advan-

tages. Hence, for the given use case of close range operations and given the possible improvements to conventional cameras a critical analysis of the plenoptic technology regarding the benefits for OOS operations is advisable.

As plenoptic cameras are a new technology in the area of OOS, the contribution of this paper is the introduction of the underlying concepts, the current state of the art with respect to robotic OOS operations and an analysis of the potential benefits. By this we intent to support the discussion of the question to which extent plenoptic cameras can improve camera based robot vision in orbit. In detail, we investigate the general image quality, the DoF extension, the 3-D capabilities, the improved low light capabilities, and the usefulness as input for an OOS robot vision algorithm. We base this on experimental data from image sequences of an OOS scenario that were recorded with a laboratory environment which provides controlled illumination conditions that simulate the ones in orbit [8].

2. RELATED WORK

Robotics for OOS

The major part of the publicly known OOS missions dealt with approaching a client and performing maneuvers with different levels of autonomy. Some of them made use of optical sensors, but only the DARPA Orbital express mission involved the use of a robotic arm. Examples are the DART mission by NASA and the 18-month mission of XSS-11 by the U.S. Airforce, both launched in 2005 [9]. In 2010 the Swedish PRISMA mission performed experiments regarding autonomous formation flying and rendezvous maneuvers with a servicer-client configuration [10]. It contained a set of maneuver experiments at different distances between servicer and client using GPS navigation as well as optical sensors, including a vision based final approach maneuver in the range from 3 m to 0 m [11].

In 2007, the DARPA mission Orbital Express successfully completed what can be regarded to be the first automatic OOS mission with a robotic arm [9]. It consisted of a small client satellite NEXTSat and the larger servicer ASTRO. Both satellites were specifically designed for the OOS experiments, which included a docking interface and a robotic arm on ASTRO. ASTRO was equipped with optical and infrared sensors for the approach as well as with a camera system on the robotic arm. The mission successfully demonstrated the fully autonomous capture and servicing of a satellite using the docking interface for the first time, as well as the first automatic capture of a client with a robotic arm [12]. In detail, this included approaching the client, circumnavigating it for inspection, grasping it with the robot arm prior to docking, transferring propellant and the exchange of a battery and a computer unit.

The teleoperation of the robotic arms on board the ISS can also be regarded as OOS operations in close range. As a specific example, the Robotic Refueling Mission (RRM), which started in 2011 on board the ISS, is a multi-phase experiment about tasks and operations close to or at a client satellite. A set of different tools is used to test robotic servicing technologies by accomplishing tasks such as removal of screws, caps or even thermal blanket on different activity boards attached to the outside of the ISS. These demonstrations were performed with the Canadian "Dextre" robot during phase one [13]. Phase two contains amongst others testing visual aids for machine vision under in-orbit illumination conditions

as well as the tests with the Visual Inspection Poseable Invertebrate Robot (VIPIR), a close range inspection tool that also allows to look inside cavities with an endoscopic camera [14]. The Raven technology demonstrator [15, 16], to be launched in 2017, will be attached outside the ISS in order to test and demonstrate real-time, relative navigation techniques for future autonomous rendezvous and capture missions. Raven combines a flash LIDAR, a camera in the visible range, an IR-camera, and an Inertial Measurement Unit (IMU) in one package [15].

Plenoptic cameras and their application

The plenoptic function [17] is a multi-dimensional function that describes the complete set of information that is visible at a certain point in space (cf. Sec. 3). Plenoptic computational imaging is a set of techniques to capture the plenoptic function and to computationally reconstruct images or other data formats from it [18]. The survey in [18] provides an overview of the possible techniques to record and to use plenoptic data and its derivatives.

This work focuses on the usage of plenoptic cameras, which record a light field with a single exposure and not by combining multiple camera images. The light field is a 4-D subset of the plenoptic function introduced by [5,6]. Plenoptic cameras incorporate an MLA in front of the image sensor in order to achieve plenoptic imaging capabilities while avoiding to block light as it would be the case with a pinhole mask or a coded aperture [19].

Plenoptic cameras allow the recording of directional and lateral samples [2, 20], i.e. the resolution of a derived 2-D image is lower than the raw sensor resolution for the cost of sampling additional information (e.g. depth) about the scene. Although the concept of plenoptic cameras was first described in 1908 by Lippmann [21], the first actual plenoptic camera was presented in 1992 by Adelson and Wang [2]. Different concepts of plenoptic cameras with increased performance were developed over the years for example by Ng et al. [3], by Georgiev and Lumsdaine [22] or by Perwass and Wietzke [1]. These developments helped to increase the resolution of the derived 2-D images [4] as well as to extend the DoF significantly [1]. Additionally, a number of approaches for the super-resolution of the 2-D images has been proposed by e.g. [23–26].

Reports on the application of plenoptic cameras in robotics are sparse and for space applications no previous work is known to the authors at the time of writing. Dansereau discusses methods for the calibration of plenoptic cameras as well as the processing of light fields with respect to the usage for field robotics [7]; the application of plenoptic cameras for underwater robotics with respect to challenging illumination conditions is discussed in [27]. Dong et al. explore the potential of real-time vision-based navigation with plenoptic cameras for robotics [28]. Visual odometry is important for autonomous robot navigation and can be improved with plenoptic cameras as shown by Dansereau et al. for underwater robotics [29] and by Zeller et al. for indoor navigation [30] [31]. Due to their recording concept and because of the subsequent processing of the light field, plenoptic cameras are beneficial in the case of demanding low-light conditions. Processing noisy light fields as well as removing distracting particles in the imaged scene was shown in [32–34].

3. LIGHT FIELD THEORY

The plenoptic function

The plenoptic function introduced by Adelson and Bergen [17] is a theoretical concept which defines the radiance of light arriving at a point in space and time. It parametrizes each light ray incoming to a point V from each direction as a 7-D function $P = f(\theta, \phi, \lambda, \tau, V_x, V_y, V_z)$, where θ and ϕ are the angular coordinates of the direction of a light ray and λ is its wavelength. The point itself is parametrized by time τ and its coordinates in space V_x, V_y, V_z . For recordings of static scenes, the plenoptic function can be reduced to five dimensions $P = f(\theta, \phi, V_x, V_y, V_z)$ as single images including all wavelengths are recorded. A recorded plenoptic function would for example allow to generate novel views of a scene.

The 4-D light field

In order to record a scene, it is sufficient to sample light rays emanating from the outside of the convex hull of objects. In other words, as long as a light ray is not blocked on its way to the sensor, i.e. as it travels through free space, its radiance does not change and therefore the plenoptic function for different V_z remains the same. In other words, the z -axis is thought of to represent the optical axis of a potential sensor and therefore the position V_z of the sensor along the z -axis would be redundant. This concept, introduced by Levoy and Hanrahan [5] and Gortler et al. [6], allows to remove the redundancy in the z -direction of the plenoptic function which results in the 4-D light field $P_L = f(\theta, \phi, V_x, V_y)$. Hence each ray is now described by the location of the view point and the direction from which it arrives at that point.

The two-plane parametrization

The two-plane parametrization that was proposed in conjunction with the light field [5, 6] describes the position and direction of a ray by its intersections with two parallel planes. The u, v -plane and the s, t -plane enclose an object. A ray emanating from the object is now described by the Cartesian coordinates of its intersections with the two planes. According to [5], the sampling of a single pair of u, v - and s, t -planes is termed a light slab. In order to record the complete light field $L = P_L = f(u, v, s, t)$ of an object, [5] and [6] propose to sample four to six light slabs at different sides of an object.

4. PLENOPTIC CAMERAS

When light rays enter a camera through a focusing main lens, they converge at the image side focal distance b , which is related to the object distance a via the focal length f as

$$1/f = 1/a + 1/b \quad (1)$$

which is the imaging or thin lens equation [35]. The remainder of the paper assumes thin lens behavior of all lenses and paraxial rays only.

A plenoptic camera consists of a refractive main lens with focal length f_M , an array of refractive elements each with a focal length f_L and an image sensor located at distance b_L behind the array as shown in Fig. 2. Although the array can be composed of pinholes, specifically shaped apertures or other types of refractive elements [18], we focus on cameras with an array of lenslets, the MLA, as it allows an increased throughput of light. Hence, the index M represents parameters of the main lens and the index L represents parameters of

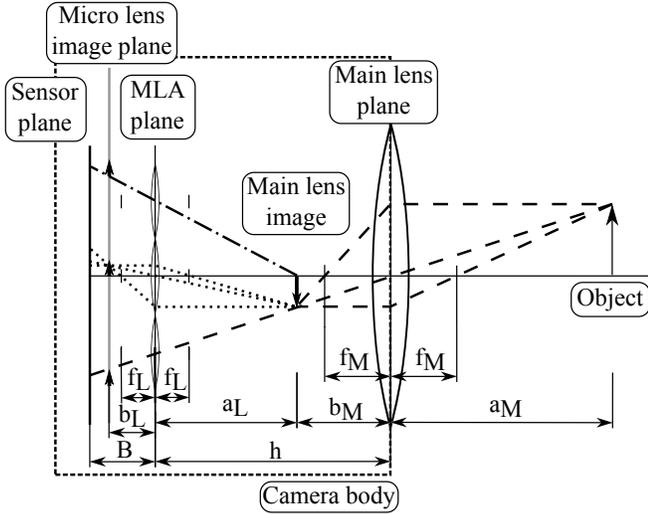


Figure 2: The concept and imaging process of a focused plenoptic camera. The index M is for values related to the main lens; the index L is for values related to the lenslets of the MLA.

the lenslets of an MLA.

In order to realize the two-plane parametrization of [5] in a plenoptic camera, i.e. the object being between two planes (cf. section 3), the main lens plane represents the (u, v) -plane and the MLA plane represents the (s, t) -plane [3]. As a plenoptic camera is a system of two consecutive lenses, the main lens image is considered to be the object for the lenslets. As they are either focused on the main lens plane or at a point within the camera, the complete theory for plenoptic sampling can be discussed within the camera body space. In the end, the resulting distances can be projected back to object space via the thin lens equation in order to maintain the connection with the real world. This requires a precise calibration of the plenoptic camera system (cf. Sec. 4 - Calibration of MFPCs).

Sampling a light field

The MLA splits the pencil of incoming rays such that rays from different directions hit the sensor at different positions as shown in Fig. 3. This preserves the angular information contained in the pencil of rays, as rays from different directions arrive at different sensor pixels. This is in contrast to a conventional camera, where the sensor is positioned at a distance b_M behind the main lens principle plane such that all rays coming from an object point at distance a_M will converge in one point on the sensor plane (cf. Fig. 4). In other words, a conventional camera integrates all rays arriving at a point x, y on the sensor from all directions s, t , which corresponds to a projection of the 3-D world to a 2-D plane, and it results in a loss of depth information.

Plenoptic camera concepts

Two different plenoptic camera concepts are possible depending on the distance B between the MLA plane and the sensor plane [36]. For the first plenoptic camera concept, shown in Fig. 5a, it holds $B = f_L$. It was introduced by Adelson and Wang [2] and improved to a hand-held version by Ng et al. [3]. Here, the main lens is focused on the MLA plane and the lenslets split the incoming pencil of rays in a similar manner as the pinholes in Fig. 3. As the

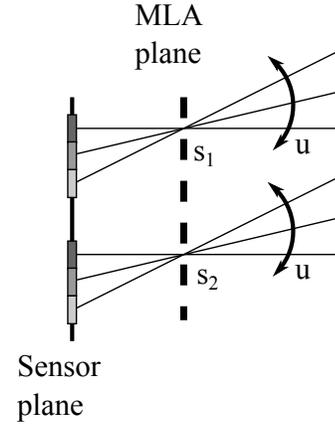


Figure 3: The plenoptic sampling principle with an ideal pinhole array shows how the converged light ray coming from the right is split up at position s on the pinhole plane into single rays that arrive from different directions u and hit the sensor plane at different pixels. By this, the pixel index within the field of view of a pinhole represents the direction u .

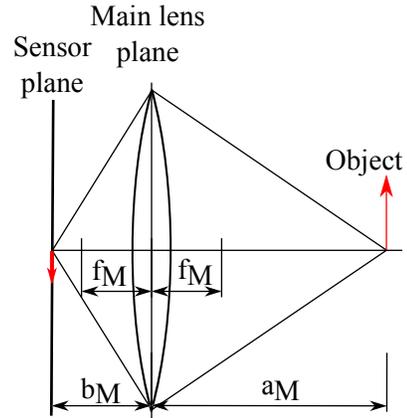


Figure 4: A conventional camera focuses all rays from a single object point to a single image point. In case the object is in focus, the rays are focused to a single point on the sensor plane.

lenslets are considered to focus at infinity on their object side, i.e. the main lens image side, this concept is known as the Unfocused Plenoptic Camera (UPC) or plenoptic camera 1.0 [4] and it is commercially available from Lytro, Inc. [37]. The second plenoptic concept, shown in Fig. 5b, is known as the Focused Plenoptic Camera (FPC) or plenoptic camera 2.0 [4]. It is achieved if $B \neq f$ and by focusing the lenslets to a distance a_L on the main lens image [22], thus the name. This concept allows a trade-off between the angular and the spatial resolution by adapting B to the requirements of the use case [36].

The condition that defines either one of the two concepts is solely based on the optical geometry of the system [36]

$$B_L > f_L \left(1 + \frac{\delta}{d_L}\right) \quad (2)$$

where δ is the sensor pixel size and d_L is the lenslet's aperture diameter. If the condition holds, a FPC (2.0) is achieved. Else the UPC (1.0) is achieved [36].

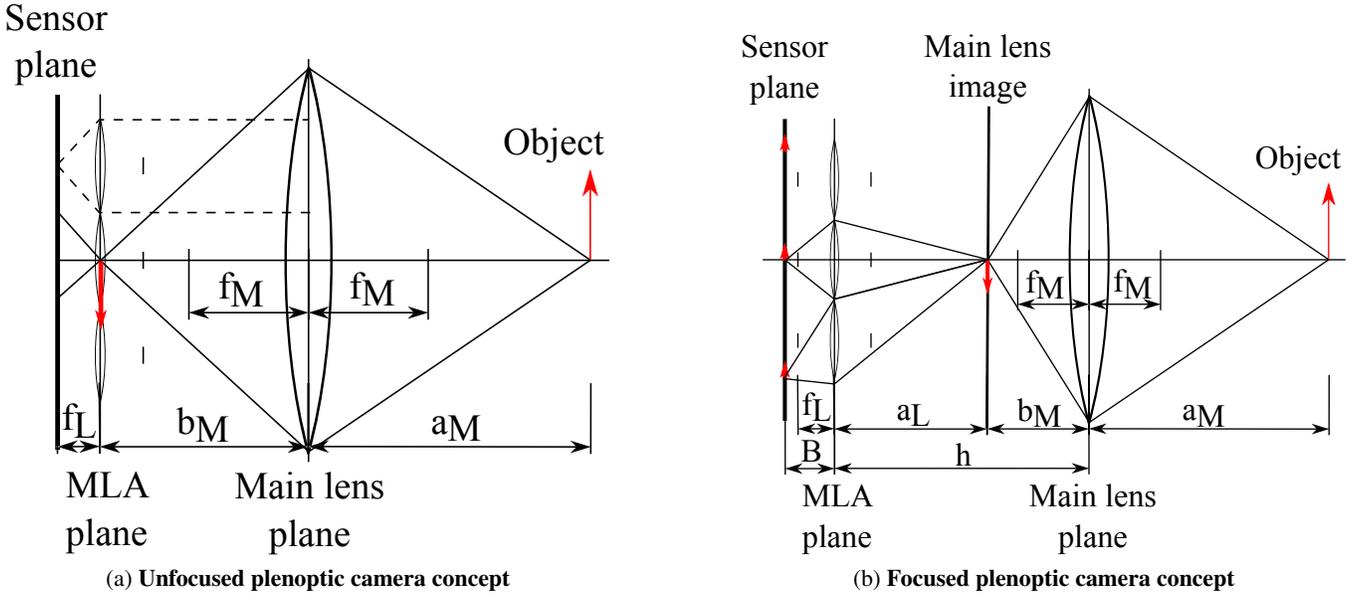


Figure 5: Plenoptic camera concepts and their focusing behavior. The lenslets of the unfocused plenoptic camera are focused to optical infinity (i.e. unfocused), whereas the main lens is focused to the MLA. This results in the split of the pencil of rays to different pixels. The lenslets of the focused plenoptic camera are focused on the main lens image, hence they work as cameras at different vantage points.

The Multi-Focus Plenoptic Camera (MFPC)

An enhancement of the FPC concept is the so called Multi-Focus Plenoptic Camera (MFPC) introduced by [1] and [38]. It applies an MLA with multiple different types of lenslets, where for example a type 1 lenslets is adjacent to a type 2 and a type 3 lenslet forming a lenslet group which is repeated throughout the MLA. Each lenslet type is focused to a different object distance a_L such that ideally the DoF of the different lenslet types touch and the overall DoF of the plenoptic camera is increased. MFPCs are commercially available from Raytrix GmbH [39].

Resolution of plenoptic cameras

The potential spatial resolution of the final image depends on the applied plenoptic camera concept, i.e. the distance B_L [36], the depth of the scene [4] as well as on the distance a_L at which the image is formed with respect to the MLA [1, 4, 36].

For the UPC concept, the spatial and the directional sampling resolution is fixed due to the design condition of $B_L = f_L$. The maximum spatial resolution is achieved for images that are formed at the MLA, and it is equal to the amount of lenslets N , i.e. $R_{max} = R(a_L = 0) = N$. In other words, the resolution of a rendered 2-D image equals N . The directional resolution of the sampled light field is equal to the number of pixels in the micro image below each lenslet [4].

The FPC concept allows a trade-off between the spatial and directional resolution, which depends on the chosen magnification factor a_L/b_L of each lenslet. In other words, the spatial resolution is not governed by the amount of lenslets alone but by the optical geometry through a_L and b_L [4]. The spatial resolution is at its maximum at the distance a_L to which the lenslets are focused, i.e.

$$R_{max} = R(a_L) = NK \frac{b_L}{a_L} \quad (3)$$

where K is the amount of pixels per lenslet [36]. In other

words, the achievable resolution is up to $\frac{b_L}{a_L}$ of the sensor resolution. These theoretically achievable resolution levels are decreased by shadowing effects and by effects of the lenslets on the boundary pixels of a micro image. The affected pixels are usually left out during the light field processing in order to avoid artifacts in the final images [4].

The MFPC, as a sub-concept of the FPC, follows the aforementioned conditions. The MFPC is reported to achieve an effective spatial resolution of up to a fourth of the sensor resolution [1].

Reasons for MFPCs

Based on the so far specified features of the different plenoptic camera concepts, we argue to use FPCs for robot vision in OOS due to their increased spatial resolution and because of the possible spatio-directional trade-off, which yields room for the adaption to specific requirements. Additionally, the MFPC variant provides an increased lateral resolution as well as an extended DoF when compared to a single focus plenoptic camera [1]. Furthermore, according to [4] the FPC concept allows to increase the space between the MLA and the sensor, which provides more freedom during hardware implementation. Additionally, this allows lenslets with a larger diameter, which in the end can improve the overall image quality. In our opinion, lateral resolution in 2-D as well as in 2.5-D depth maps is essential in order to distinguish objects, area boundaries, and edges. Thus, we rate spatial resolution higher than directional resolution. Of course, on the one hand, an increased directional resolution can be important for finer depth increments. On the other hand, higher resolved micro images can lead to better results in the currently applied correspondence-based depth estimation for light fields, resulting in finer depth steps. A final argument for the use of MFPCs over other plenoptic camera concepts is that resolution of space-qualified sensors is valuable and thus it should be used as effectively as possible. Therefore, the remainder of this paper discusses the usage of MFPCs and

their comparison with conventional cameras.

Calibration of MFPCs

Using cameras as measurement devices requires the accurate calibration of the intrinsic and the extrinsic parameters. A MFPC measures the virtual depth v of a main lens object [1] as

$$v = a_L/B \quad (4)$$

The metric depth in the MLA's object space z_L is related to the virtual depth v via

$$z_L = v \cdot B + h. \quad (5)$$

where h is the distance between the main lens principle plane and the MLA's principle plane [40] (cf. Fig. 2). Therefore, in order to achieve *metric* depth measurements a determination of the plenoptic parameters B and h is required in addition to the conventional intrinsic parameters.

Several calibration approaches for MFPCs have been proposed by [41–43] for which images of a calibration plate are recorded at multiple known distances and an optimization procedure tries to fit the estimated depth to the ground truth distances in order to determine the required parameters.

A more practical approach was introduced by Strobl and Lingenauber [40], which only requires the camera's pixel size as prior knowledge. It is operated similar to established conventional camera calibration methods [44, 45] but extended with an additional step for the calibration of the plenoptic parameters. The camera records a calibration pattern from different viewpoints and view angles. With the help of the known geometry of the pattern and due to perspective deformations, the intrinsics of the camera are determined based on 2-D images rendered with an extended DoF (cf. Sec. 4 - DoF of MFPCs). In the second step, estimated depth maps (cf. Sec. 4 - Depth estimation) are used to determine the plenoptic parameters B and h [40]. In combination with a robot arm, the extrinsic or hand-eye calibration is provided as well [45].

Besides its increase in accuracy and computational speed if compared to other available methods, the stepwise metric calibration is interesting from a satellite operations point of view. During an OOS mission, the cameras can be re-calibrated, e.g. in the course of the commissioning phase, as the method only requires a known geometry within the work range of the robot arm. As long as the calibration geometry provides enough well-known feature points and good contrast, any structure might be sufficient. The procedure would only require a set of at least three images recorded from different view positions and view angles in order to deliver the intrinsic, plenoptic and extrinsic parameters automatically. For an intrinsic calibration alone, a sole image ought to be sufficient [45]. However, for an increased statistical stability of the solution and for a better feature density, we regularly use eight images. In addition, the requirement of a known structure could be potentially lifted [46].

DoF of MFPCs

In general, the Depth of Field (DoF) is defined as the range an object point can be moved back and forth on the optical axis until the diameter c of the circle of confusion becomes larger than the pixel pitch of the camera sensor. In other words, it is the range of in-focus distances a of a camera system within which an extended object appears to be imaged sharply when the focal length f and the image distance b

remain fixed. The DoF depends on the sensor pixel size, the focal length, the aperture, and the focal distance of a lens. In case of a plenoptic camera, a two-lens system must be considered as well as the decrease in image resolution due to the combination of different sensor pixels for the 2-D image generation (cf. Sec. 4 - 2-D image rendering).

The DoF of a conventional camera reduces quickly when the focus distance is reduced to only a couple of millimeters at close distances. Given the case of close range operations in OOS, these are the situations where high accuracy and thus sharp images are required in order to avoid collisions. In-focus images are important for detecting edges and boundaries of objects but also for the reliable and accurate depth estimation (cf. Sec. 4).

Regarding FPCs, [1] shows that a plenoptic camera's DoF can be approximately 1.9 times as large as the one of a conventional camera at a focus distance of 5 m. In case of the MFPC, they achieve a DoF extension of approximately 1.8 for the 5 m distance but at 0.5 m, they achieve a DoF 5.4 times larger than the conventional camera DoF. Hence, a MFPC can provide a larger DoF than a conventional camera, but in close range the DoF becomes significantly larger.

2-D image rendering

The rendering of 2-D images from the raw light field equals to compute the intensity value at the final image position. A possible procedure for an MFPC is described in [1] and requires an already known depth estimate for each image position (cf. Sec. 4 - Depth estimation). The estimation of depth prior to the 2-D image generation is necessary in order to avoid rendering artifacts [1, 4]. Depending on the depth, the lenslets that image the object point at (x, y, v) best are chosen and for each lenslet a projection of this point onto the image plane, based on the known parameters of the lenslets, is performed. Due to the depth dependent choice of lenslets for an object point, the procedure results in an image with extended DoF. The data resulting from this process is shown in Fig. 6 for an MFPC, i.e. the raw image, the depth map, and the final 2-D image.

Further approaches for FPCs are reported in the literature which also perform a depth estimation prior to the 2-D image rendering. The approach by [4] is similar to the described one. Bishop and Favaro [24] show how the antialiasing of the light field prior to the depth estimation and the 2-D rendering can provide super-resolution images. Georgiev et al. show super-resolution rendering for FPCs with a color pattern [47]. Wanner et al. [48] discuss the generation of Epipolar Plane Images (EPI) [49] from light fields recorded with a MFPC in order to enable the advantageous features of the EPI data format regarding image analysis or 3-D reconstruction.

Depth estimation

The depth estimation of FPCs can be regarded as a multi-view stereo vision problem as each lenslet acts as a single camera that views an object point from a different vantage point. The lenslets are positioned on a regular grid, thus their relative position is known. Due to the parallax between lenslets, a triangulation of the depth is possible. As the main lens image's size is approximately of the same magnitude as the lenslet dimensions, the available parallax is assumed to be sufficient. Therefore, depth estimation experiences the same decrease in accuracy due to the object distance as in stereo vision.

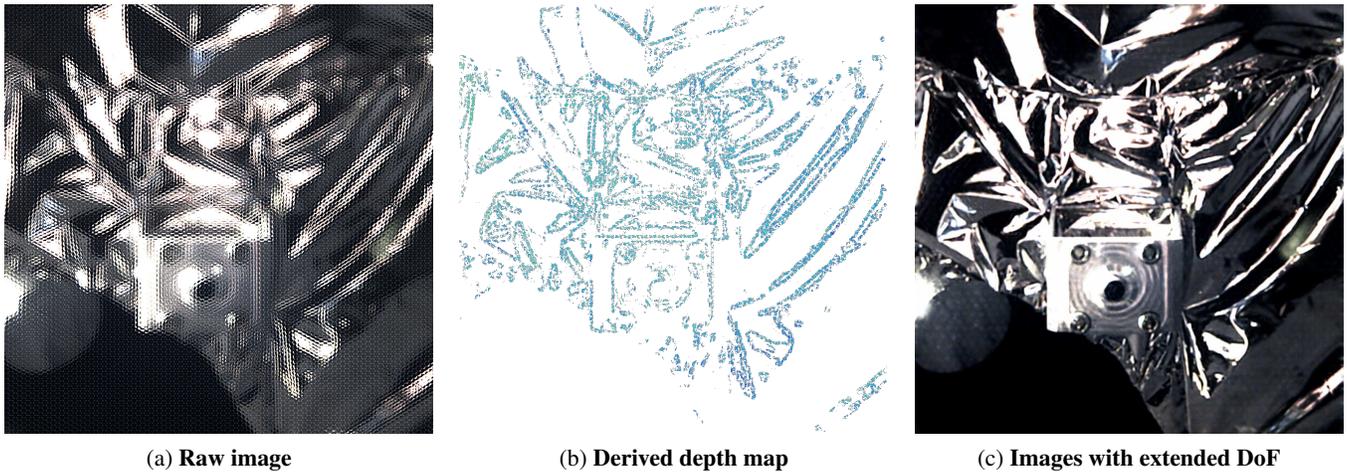


Figure 6: Example images for the processing pipeline for images from an MFPC. The 4 MPx raw image (left) allows a derivation of a depth map (middle) whose depth values are used for the generation of a 1 MPx 2-D image with an extended DoF (right). The distance between the camera and the tip of the object was approximately 40 cm.

Depth is estimated in terms of the virtual depth in Eq. 4, which is also required for correct 2-D image rendering. The calibration discussed in Sec. 4 - Calibration of MFPCs relates the virtual depth to the real depth in object space. For FPCs, Georgiev and Lumsdaine [4] propose cross-correlation with image patches. For MFPCs, Perwass and Wietzke [1] propose the use of Sum of Absolute Differences (SAD) to establish the correspondence between image patches of different micro images. Both variants require sufficient image contrast and thus limit the depth estimation primarily to edges and alike structures. An example for the depth estimates of the latter algorithm is shown in Fig. 6b. More enhanced depth estimation approaches take the aliasing of real plenoptic cameras into account [50, 51] or make use of a probabilistic, pixel-based depth estimation [31]. The approach in [34] works on EPIs instead of combining micro images. A discussion of the different approaches is beyond the scope of this paper but can be found for example in [31, 34, 51].

5. PLENOPTIC CAMERAS FOR SPACE

Hardware

Given that plenoptic cameras are a new technology, the effort to make the hardware space-ready is relatively straightforward. As the example of the Raytrix camera and multiple examples in literature show [3, 4, 26], it is possible to achieve a plenoptic camera by carefully modifying an existing camera with a MLA. Hence, achieving a space ready plenoptic camera becomes a question of material choice for the MLA, its manufacturing and the integration of the MLA into a space qualified conventional camera.

MLAs for terrestrial applications are for example made with polymer on glass [39]. Due to the ionizing radiation in space, which can alter the MLA material and results in changes of its optical properties, radiation resistant optical materials must be used. Depending on the chosen material, this might induce additional technical challenges for the manufacturing due to the small structures of the MLA. As the discussion of the optimal material choice for a MLA and of the appropriate manufacturing process is beyond the scope of this paper it is left as future work.

Regarding the integration of the MLA into an existing space qualified camera, first the sensor and the MLA have to be mounted closer to the main lens if compared to a conventional camera in order to achieve the maximum lateral resolution as well as the full amount of DoF extension in an MFPC [1]. This allows to use a main lenses with a smaller exit pupil and therefore a smaller diameter, thus saving mass and volume. Second, a rigid lens and sensor mount is important as the distances h and B (cf. Fig. 5b) need to be well calibrated in order to achieve accurate depth estimations. Hence, the MLA will require a mount added to the camera body which can ensure the required level of rigidity. In case a camera calibration can be performed during the commissioning phase and not only prior to launch, the requirements on rigidity might be lowered to the extend that movements of the robotic arm do not alter the optical system. Although the MLA can be adjusted to the available f-number and focal length of an available space qualified main lens, it seems to be advisable to adapt both, the main lens and the MLA, to the specific requirements of a mission.

The required sensor resolution of a MFPC of approximately four times the desired 2-D image resolution [1] seems to be feasible for close range application. We assume a 2-D image resolution of 1 MPx to be sufficient, i.e. a sensor of 4 MPx is required. To our knowledge 4 MPx CMOS sensors are available for space usage, e.g. the CMOSIS CMV4000 sensor [52]. Based on our current experience from experiments, it is also advisable to use a stray light baffle in order to avoid image artifacts as the ones shown in Fig. 9. A monochromatic sensor is expected to increase the image quality as well.

Software

Due to their higher dimensionality, processing light fields requires more computational resources in terms of processing time and memory as 2-D images. Given the limited computational resources on board a spacecraft, this makes the software side of plenoptic imaging far more challenging. In contrast to the relatively low amount of required hardware modifications, the space-ready, on-board processing of light field data requires a complete new development including tests and verifications.

In order to decrease the computational load, it is possi-

ble to perform light field processing in the Fourier domain [33, 53, 54] instead of the 4-D spatial domain. For example Ng [53] achieves a reduction of algorithm complexity from $O(n^4)$ to $O(n^2 \log n)$ for refocusing. As shown by Dansereau et al. [32], processing in the Fourier domain significantly increases the speed but it requires more memory than in the spatial domain as the complete light field has to be loaded into the memory at once. Which domain is more beneficial for processing depends on the size of the light field and the available resources. As processing power is more difficult to achieve in space than memory size, it seems to be advisable to perform the processing in the Fourier domain. Additionally, hybrid processing, i.e. some parts are computed in the spatial domain and other parts in the Fourier domain, is also possible as a compromise and to balance the limitations of either processing type [32].

Due to their data structure, rendering from light fields can be parallelized, e.g. for processing on a GPU [55], and an implementation in hardware, e.g. with FPGAs, seems to be feasible [56]. Hence, it is not only possible to achieve a trade-off between computational load and memory usage but also to leverage parallelized processing in order to increase the computational speed of light field processing. As with related image processing tasks, we expect a significant decrease of the required computational power once the appropriate algorithms are implemented closer to hardware, e.g. in FPGAs.

6. EXPERIMENTS

We performed experiments based on two data sets recorded with a close range OOS test setup (cf. Sec. 6 - Experimental setup). This gave us first insight regarding the currently possible DoF extension and the quality of 2-D images, 2.5-D depth maps and 3-D views of a possible target point. Furthermore, we were able to investigate the performance of a vision based pose estimation when used with extended DoF 2-D images.

Example mission scenario

The experiments and the setup are based on a potential scenario of an OOS mission to repair or de-orbit a client satellite. The client satellite is assumed to be at least attitude controlled from ground and its geometry is known ahead of the mission. It is considered to be a non-cooperative target because no markers or aids to support the vision based OOS operations are available, in contrast to the Orbital Express mission.

For the client satellite, we assume a mid-size earth observation satellite in the Low Earth Orbit (LEO) in a sun-synchronous dawn-dusk orbit. As a result, the sun light always shines from the side onto the satellite. Although the operations are assumed to be performed outside of the eclipse phase, this results in challenging low light conditions as visible in Fig. 8. The mission scenario and the servicer satellite are assumed to be comparable to the German DEOS mission study [57]. Thus, a light weight robot arm is attached to the servicer satellite in order to perform different types of operations close or at the client satellite. The robotic arm is equipped with a tool and with a camera system mounted close to the Tool Center Point (TCP) in order to support the operations with live image data fed in a robot vision processing pipeline. This allows to operate the robot arm either autonomously or from ground via teleoperation.

Our experiments cover an approach of the servicer’s robotic

arm towards a dedicated target point or grip, which is the final step of a rendezvous and grasping maneuver in order to establish a firm connection between the two satellites. As the client is attitude controlled, we assume the relative motion to be close to zero in all six degrees of freedom between the two satellites. Due to its mechanical stability, the Launch Interface (LIF) is chosen as the target point for a grasp (cf. Fig. 8). Because of its limited size and in order to avoid damage of the surrounding Multi Layer Insulation (MLI) foil, a pose estimation accuracy of 10 mm is required for the grasping operation.

The current experimental setup represents the worst case scenario in terms of illumination conditions as we assume to have only sun illumination available. In the nominal case an Artificial Illumination Unit (AIU), which is attached close to the cameras, can provide illumination of the scene.

Experimental setup

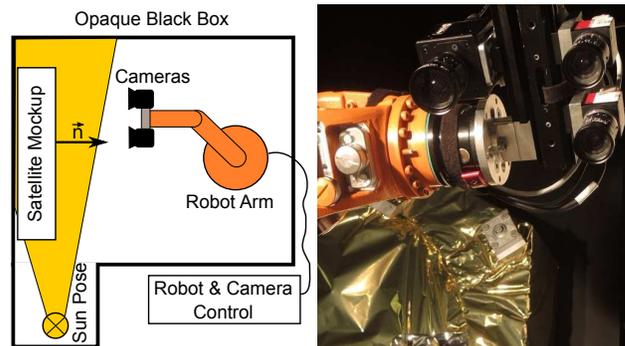


Figure 7: Overview of the experimental setup. Left: an industrial robot, a mockup of a client satellite and a simulated sun are surrounded by an opaque black box for space like conditions. Right: A single plenoptic camera (black) is attached to the robot’s TCP for obtaining images while approaching the satellite mockup. The mockup and a LIF is visible in the background.

In order to collect representative data for the experiments, we used the test setup as described in [8] and shown in Fig. 7 but with a plenoptic camera instead of conventional cameras. The setup allows to simulate the approach of a robotic arm towards a Launch Interface Bracket (LIF), which is meant to be used as a point for grasping, under controlled illumination conditions. It consists of a full-scale mockup of a LEO earth observation satellite, a strong light source to simulate sun illumination and an industrial KUKA KR-16 robot with cameras mounted to its TCP. The setup is surrounded by a black box to block out any external light. We used a Raytrix R5-C-K plenoptic color camera with a 4 MPx CMOSIS CMV4000 sensor together with a Pentax 12.5 mm main lens. The optics were set to an f-number of approximately 3.8, as for MFPCs the f-number of the main lens should match the one of the micro lenses [1]. The optics were focused to the far end of the assumed work range of 220 cm and 90 cm, respectively.

Following the naming convention of [8], we recorded data for two trajectories named LIF0_0 and LIF3_0 with the sun at position 0, i.e. at 90° to the satellite surface normal direction \vec{n} (cf. Fig. 7). The LIF3_0 trajectory has a range from approximately 110 cm to 20 cm with respect to the center bolt surface of the LIF (cf. Fig. 8c) and the LIF0_0 trajectory is from 220 cm to 20 cm. The ground truth of the trajectory is provided by the inverse kinematics of the robot. LIF3 is a

milled aluminum part and shows typical milling marks (cf. Fig. 9c) whereas LIF0 is a 3-D printed part. Hence, they show differences in surface structure and reflective behavior (cf. Fig 8c and Fig. 9c). In contrast to the trajectories in [8], the position of the light field camera was slightly shifted and the LIF was not always in the center of the Field Of View (FOV). As in [8], the position of the satellite mockup was calibrated within the robot reference frame and a CAD model of it is available as ground truth for geometrical comparison.

Previous to the dataset recording, images of a checkerboard pattern were recorded with the plenoptic camera and used to perform the intrinsic, plenoptic, and extrinsic calibration with the procedure in [40]. If not stated otherwise, the RxLive software version 2.10 was used to process the recorded light fields and the derived data.

Experiment 1: Extended DoF and image quality

In order to evaluate the quality of the extended DoF 2-D images we investigated the recorded image data sets in terms of image sharpness at different distances. Additionally, the overall image appearance in terms of brightness, contrast and rendering artifacts was examined.

The results are based on 2-D images rendered with the RxLive software without any further enhancements but with the extended DoF option enabled.² The image quality was only controlled by setting the exposure time that appealed to provide well lit images at the time of recording.

Experiment 2: Depth maps and 3-D capabilities

The second experiment was about investigating the depth estimation potential of the currently available plenoptic camera technology. From the recorded light field data we derived 2.5-D depth maps and 3-D renderings by using the RxLive software and thus the depth estimation procedure of [1]. The applied camera settings might not have been optimal for depth estimation at each distance, because they remained fixed throughout a trajectory. In our opinion this might also be the case during an operation in orbit.

We experienced that the depth estimation algorithm is very sensitive to changes in brightness, contrast and noise level in the raw images. In order to reveal the full currently available potential for the depth estimation we manually improved the aforementioned parameters as to increase the quality of the raw images prior to the depth estimation. Additionally, the depth estimation parameters were set such that a balance between the possible accuracy, a low level of depth noise and a certain density of the depth maps was achieved. The correlation based depth estimation only works in regions with sufficient texture and therefore best at edges and other regions with high contrast [1].

The RxLive software allows to create different types of depth maps as shown in Fig. 10. First, the original depth map which only contains actual depth estimates. Second, an enhanced depth map where values are added around the depth estimates. Third, a complete fill of the remaining holes in the depth map by area-wise interpolation. We derived all three types of depth map from the recorded light field data as well as the 3-D renderings shown in Fig. 12. Thanks to the plenoptic camera calibration we were able to derive 3-D point clouds with metric values in the object space from all three types of depth maps. For the conversion from depth maps to

point clouds, the same procedure as in [40] was applied.

Experiment 3: Vision based pose estimation

The accurate tracking and pose estimation of a target point for a grasping maneuver or related tasks can be achieved with algorithms as the one presented in [8]. We employed this vision based pose estimation algorithm with model support to the sequence of 2-D images with extended DoF from the LIF3.0 trajectory and validated the pose estimation against the ground truth data. From the CAD model of the satellite mockup of [8] we derived a simplified version containing the main edges as seen in Fig. 13.

The edge-based tracking method, introduced in [8], is based on the work of [58]. Similar variants [59, 60] of edge-based tracking methods that do not extract image edges explicitly exist. However, they require line parametrization of edges. The applied algorithm searches correspondences between model and image edges, to minimize the re-projection error in pose-space, and updates the predicted pose with respect to rotation and translation parameters. The algorithm relies on a local, non-linear Least-Squares Estimation (LSE) that is fast, accurate and provides a drift-free pose estimation. However, the presence of spurious undesired edges limits the range of convergence and leads to a local minimum close to the global optimum in six degrees of freedom. This is particularly critical with the challenging conditions posed by the MLI specular reflections [61, 62], the metal parts, and the harsh illumination in space. Regarding our image data, the challenges are the poorly illuminated object and a significant change of illumination over the course of an image sequence. This can result in the failure of the state of the art edge-based tracking methods. To address this problem, we employ a template matching technique to refine the last pose, in order to improve the prediction capability for the current pose. An alternative to this prediction approach is introducing a motion model and the Kalman filter, which requires a precise motion model and state statistics. Therefore, we improve the robustness by first applying a 2-D template matching procedure [63] to obtain an approximate planar transformation between previous and current camera images. This transformation is used to refine the prediction available from the last estimate that will be closer to the true pose, hence reducing the risk of failure for the subsequent LSE.

As a pre-processing stage, we removed non-linear distortion using the available camera calibration data in order to use a linear projection model during pose estimation. Also we computed the edge maps using the Canny edge detector [64] and their normal directions. The template matching based on image edges improves the performance of a standard model-based tracking, by providing a priori information to the LSE optimization. However, this is at a cost of slightly increased computation, which is generally acceptable for the requirements of an OOS mission.

7. RESULTS

Results 1: Extended DoF and image quality

Examples of the image sequence for the LIF0.0 trajectory are shown in Fig. 8. The images show the beginning, the middle and the end of the image sequence. For the qualitative investigation of image sharpness, a zoom on the targeted LIF is provided for the first two images. All images are extended DoF images but without any post-processing in terms of adjusting the brightness, contrast or noise reduction.

²In Raytrix terms these are called total focus images.

The achievable difference with post-processing is visible in comparison with Fig. 10.

As the image sequence in Fig. 8 contains images without the raw image enhancement used for experiment 2 (cf. Sec. 6 - Experiment 2: Depth maps and 3-D capabilities) it shows the actual brightness and contrast level due to the low light conditions. The zoomed detail views in Fig. 8 show the achievable sharpness caused by the extended DoF of the MFPC. The LIF remains distinguishable throughout the trajectory, given the image resolution of 1 MPx. Despite the short exposure time of 14 ms $\approx 1/71$ s and thanks to the low f-number of approximately 3.8, the image is bright enough to contain sufficient content for the vision based pose estimation (cf. Sec. 7 - Results 3: Vision based pose estimation).

In general, all images from the sequence possess good quality and allow the usage in robot vision algorithms as shown exemplarily in Sec. 7 - Results 3: Vision based pose estimation. We found two sources of artifacts specific to plenoptic cameras shown in Fig. 9. At certain constellations of camera and illumination source a ghost image of the MLA, as the one in Fig. 9b on the right side, is present in the images. We assume it is due to stray light and that it can be reduced or avoided by using a baffle on the optics. Since the 2-D image rendering depends on depth information, the ghost images lead to small but visible rendering artifacts as under these conditions the determination of correspondences for the depth estimation is error prone. A second source for rendering artifacts are saturated or highlighted regions, in which the correspondence search is also not possible (cf. Fig. 9b on the left). This leads to jumps of the rendered highlight regions between consecutive images, which can potentially avoid a successful feature detection and tracking in these areas. These two sources for rendering artifacts also lead to noise in the depth images. While the ghost images often contribute to only a single lenslet type's depth estimation, the highlights can affect all types.

The effect that a single object point is sampled in multiple micro images is visible in Fig. 9c. Here a cut-out of the raw image is compared to its counterpart from the rendered image, which is shown for reference below. The red and green colored reflections are visible as dots in several micro images and the 2-D rendering blends them to form a single spot of reflection. In a gray scale version of this image, this would result in an equally shaped bright spot. So far we could not determine a negative effect on the further processing due to the color images.

Results 2: Depth maps and 3-D capabilities

The results in Fig. 10 show extended DoF images with increased brightness and contrast in comparison to the ones in Fig. 8. The adjustments were necessary, in order to achieve depth maps with a sufficient density (cf. second row of Fig. 10). As the depth estimation depends on texture, most measurements are located at object edges that provide sufficient contrast for the correspondence search. Hence, the depth map density also depends on the distance to the object as more edges and details become visible in close range. In general the actual depth maps appear to be sparse including a significant amount of depth noise. The results of the artificial increase of the point density by the RxLive software are shown in the third row of Fig. 10. It appears, that the software uses a sort of growing operator to increase the areas around depth estimates. Please note, the depth maps in row two and three of Fig. 10 show only virtual depth values (cf. Eq. 4).

By extracting the depth values from the depth maps and by combining them with the calibration we achieved metric depth measurements in object space in form of 3-D point clouds. Due to the sparse actual depth maps, we used the enhanced depth maps for the point clouds shown in Fig. 11. Please note, that in order to be able to visualize the difference in point density at different distances, we had to thin out the point clouds. Only every 20th lateral position contains a point, if one was present in the raw data, resulting in a lateral resolution of 50×50 in the point clouds shown in Fig. 11. In the top row of Fig. 11, the point clouds show the back of the recorded scene. In the bottom row they show a side view, where y_C indicates the vertical direction of the images in Fig. 10 and z_C indicates the metric distance from the camera to the object point in real space. As visible for the distance of approximately 19 cm in Fig. 11c, the plenoptic depth estimation is coherent in general but with a high amount of depth noise. However, it shows the accuracy of the stepwise metric calibration, as the noise is a result of the virtual depth estimation (cf. Fig. 12). Furthermore, the diagonal pattern in all point clouds show that the actual metric distance to the camera vantage point of the perspective projection is provided due to the successful rectification of the calibration. The LIF, whose surface was approximately 60 mm closer to the camera than the rest of the satellite, is clearly distinguishable on the right side of Fig. 11c. As expected, the noise is more present in areas with the reflective MLI. Interestingly, the depth noise is mostly directed away from the camera, resulting in a relatively well perceptible front surface of the object. Fig. 11b and Fig. 11a clearly show the decrease in depth accuracy and point density at larger distances. It also shows the decrease of depth increments as the steps between the LIF surface and the rest of the satellite become larger. For the point cloud of the beginning of the trajectory in Fig. 11a, the reflections of the large wrinkles of the MLI in the upper part of the FOV seem to dominate the depth estimate and result in false depth values (deep blue dots). Nevertheless, the distance to the LIF at $y_C \approx 50$ cm is estimated correctly.

Only after the increase in point density, 3-D renderings as shown in the first row of Fig. 12 became possible. By filling the remaining holes in the depth map with interpolated values, the renderings in the second row of Fig. 12 were achieved. Please note for both cases, these are interpolated depth values which should not be regarded as real measurements. The renderings in Fig. 12 show virtual depths (cf. Eq. 4) which were scaled with the same factor for all three renderings in order to illustrate different depth levels. The extended DoF images were used as texture. As visible, the larger the distance to the object gets the more the difference in depth vanishes, resulting in a single plane over the complete image at the maximum distance. Additionally, as visible in Fig. 12a, the complete fill of the depth maps leads to renderings that include the background of the scene, i.e. the black box wall, in the same depth plane although it was located approximately 90 cm behind the satellite mockup.

Results 3: Vision based pose estimation

The estimated pose of the simplified model used for the vision based pose estimation algorithm (cf. Sec. 6 - Experiment 3: Vision based pose estimation) is shown in the top row of Fig. 13 for the trajectory LIF3_0. The model only contains the most prominent edge of the targeted LIF and the surrounding area in order to minimize the required computation. As visible, the model was fitted well on the corresponding edges in the image, which indicates a good pose estimation. In order to quantify this impression, the plots in the middle row of Fig. 13 show the absolute deviation of the pose estimate T_k

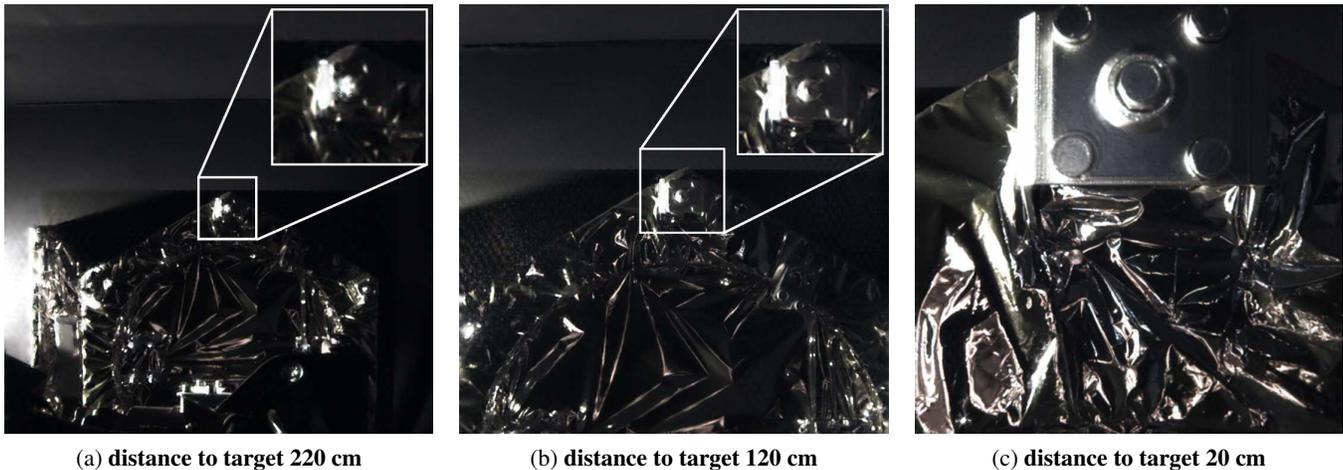


Figure 8: Example images from the sequence for trajectory LIF0.0 at different distances. Zooms on the targeted LIF are provided to show the effect of the extended DoF of the MFPC.

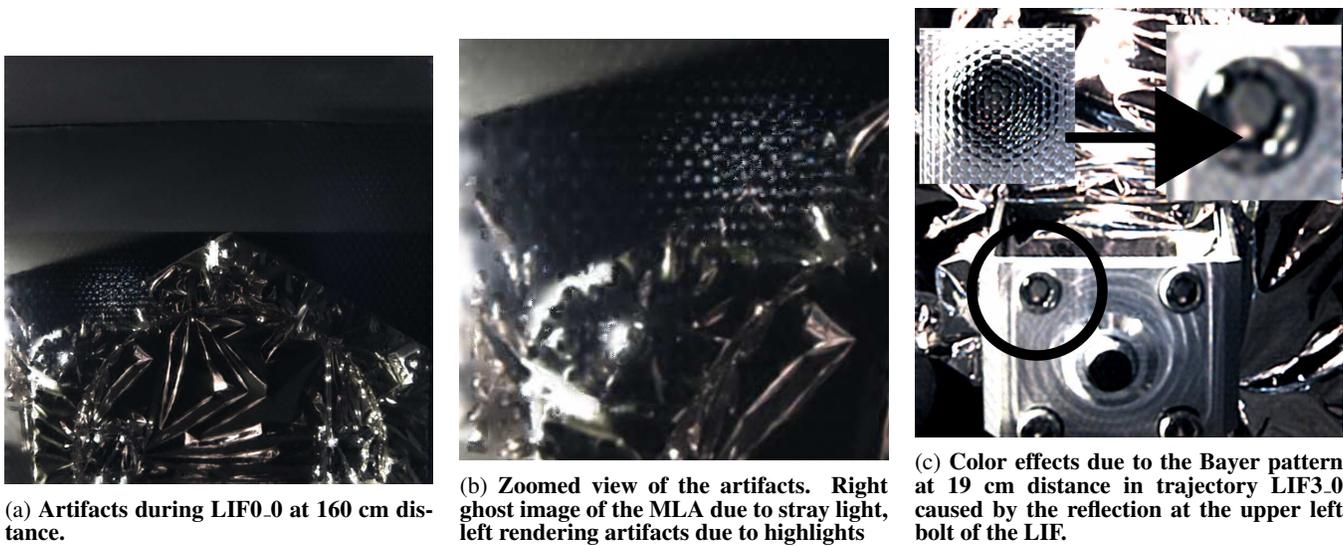


Figure 9: Example of artifacts and effects visible in the rendered 2-D images.

provided by the vision based pose estimation relative to the ground truth pose \bar{T}_k as measured by the robot’s kinematics for the LIF3.0 trajectory in all three dimensions. We remark here that the orientation is constant, hence the pose estimate is for a pure translation and that the algorithm is initialized with a ground truth pose in the beginning. Please note, that the results were achieved without pre-processing for brightness or contrast enhancement.

The vision based pose estimation with extended DoF images provides an accurate position estimate, with the desired accuracy less or close to 10 mm at the end of the trajectory. In the lateral direction (Fig. 13d and Fig. 13e) the maximum error is approximately ± 3 mm. In the direction of the optical axis, the initial error grows after initialization but reduces to approximately 12 mm at the end of the trajectory. In order to cross check the absolute deviation plots, the plots in the bottom row of Fig. 13 show the estimated relative translation between consecutive images. In the ideal case they would show no deviation in the horizontal and vertical direction and a constant offset of 10 mm in the direction of the optical axis, which was the distance between consecutive images.

In the lateral direction, the maximum relative translation estimate is at first unsteady with values between -1 mm and $+5$ mm but during the final 20 cm it is approximately zero. In the direction of the optical axis, the consecutive relative translation shows a more steady behavior with only a few peaks, one of them as large as approximately 37 mm, but overall it remains close or at the desired 10 mm.

8. DISCUSSION

The light field concept, i.e. the reduction of the plenoptic function to 4-D, requires a static scene. By using a single shot plenoptic camera instead of a moving camera as in [5,6], the redundancy in the z-direction is still present even with motion in the scene. The raw image recorded by a plenoptic camera is in fact a 2-D image with redundant information (cf. Fig. 6a). Hence, each image recorded by a plenoptic camera is a single 4-D light field of a static scene. This is valid as long as the motion in the scene is less than the exposure time of the camera, i.e. as long as no motion blur is present in the images. Once this is violated the reduction to a 4-D

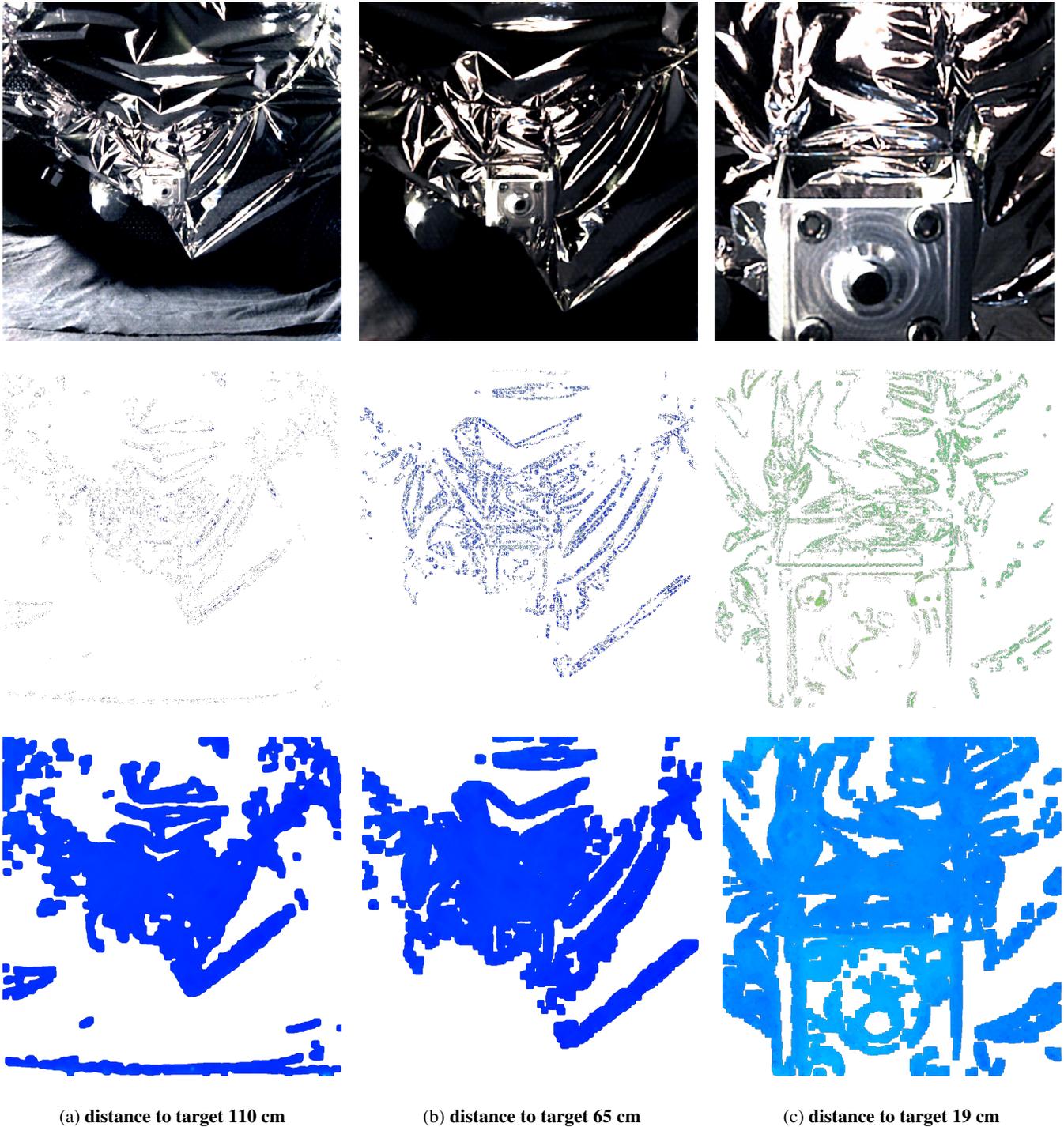


Figure 10: 2-D renderings and depth maps of LIF3.0 at different distances. First row: 2-D images with an extended DoF and after pre-processing of the raw file. Second row: Map of the actual computed virtual depth values. Third row: Maps of the enhanced virtual depth values.

light field is not valid anymore and artifacts in the processed data can appear. For the given OOS scenario, as long as the movement of the robot arm towards the target point is less than the exposure time, it still allows to rely on the 4-D light field concept. Although, the data for this work was recorded in stop motion [8], the motion during an actual close range operation is assumed to be small enough to avoid any motion related effects.

The results in Sec. 7 - Results 1: Extended DoF and image quality show, that MFPCs and the standard light field processing provide 2-D images with a good level of brightness and contrast in an OOS scenario, despite the low-light conditions in this specific data set. Due to the plenoptic sampling concept it was possible to record images with a low f-number of approximately 3.8 and still provide a DoF of 2 m as visible in Fig. 8. Thanks to the low f-number, the exposure time as

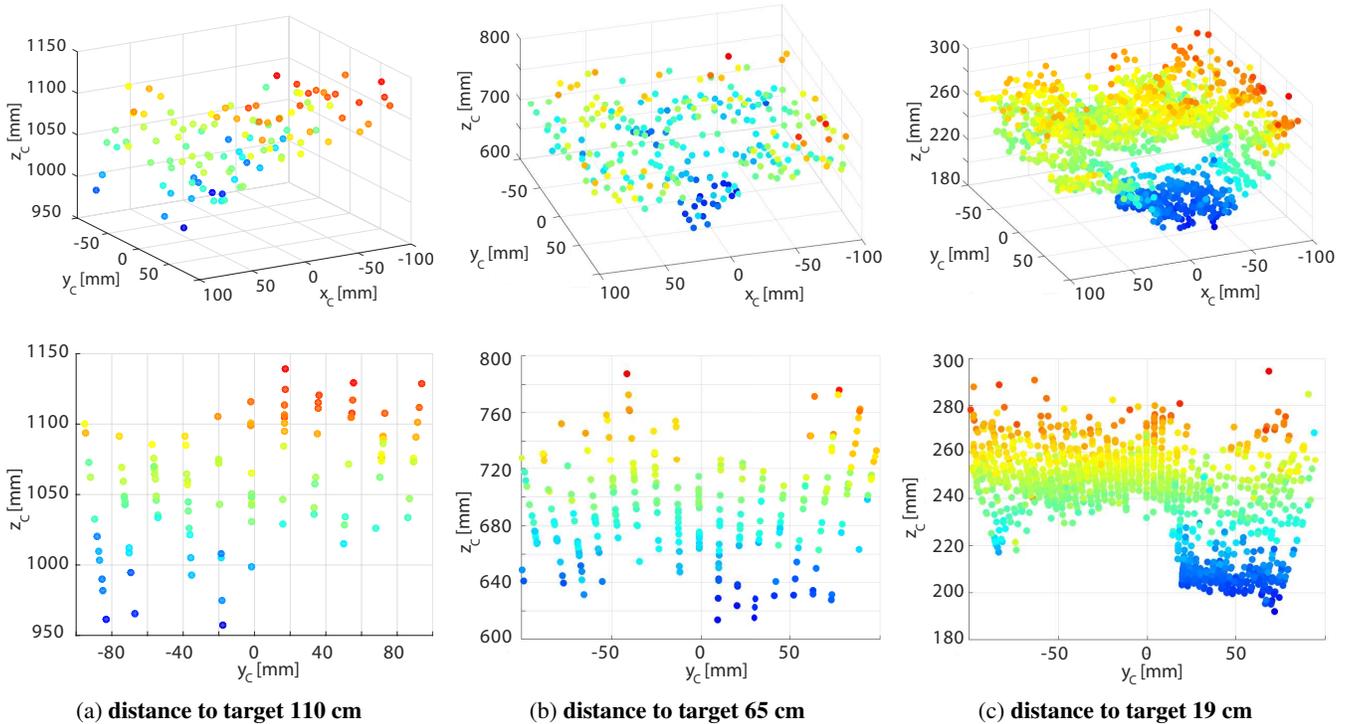


Figure 11: Calibrated point clouds of trajectory LIF3_0 derived from the depth maps in Fig. 10. Top row: Tilted view on the object’s back side showing the extend of the scene. Bottom row: View on the side showing the depth increments. Depth values are in millimeters and color coded from blue (close) to red (far) with respect to the distance of the camera to the object point. y_C is in the vertical direction of the images in Fig. 10, z_C denotes the depth in the direction of the optical axis.

well as the sensor gain could be left low, which assures blur-free images and potentially low noise levels, respectively. Hence, already at recording time a good image quality can be achieved as it was expected from the conceptual discussion in Sec. 4. This reduces the effort and the side effects of noise reduction and similar post-processing steps. Due to the additional rendering step that is required to generate 2-D images from the raw plenoptic recordings, artifacts in the images are possible and were also present in ours. As said in Sec. 7 - Results 1: Extended DoF and image quality, we expect to be able to minimize the ghost images of the MLA by using a baffle in front of the main lens. However, the artifacts that were caused by highlights on the reflective surfaces might require dedicated processing in order to avoid or at least minimize the effects they have on image quality. As visible in the metric point clouds in Fig. 11, they can also cause false depth estimates. As shown in [65], it is possible to exploit the light field data format in order to decrease the effects of reflective surfaces. In case it also applies to images in an OOS scenario, this would further increase the usefulness of plenoptic cameras.

Regarding the 3-D capabilities, the results show that the standard depth estimation algorithm provides sparse depth maps with a significant level of noise in the case of our recordings. We would expect similar behavior from a stereo vision system, as the depth from triangulation concept requires texture and gradients for the correspondence search. Given the challenging low-light conditions as well as the reflective and also low-textured surfaces, the results are as expected. The metrically calibrated point clouds in Fig. 11 show, that the depth accuracy decreases with an increase in camera to object distance, as known from stereo vision, which

is also according to our expectations. Interestingly, the depth noise is mostly directed away from the camera, resulting in a predictable front surface of the object. Therefore, the MFPCs might at least qualify as a passive obstacle detection device for hazard avoidance purposes. Nevertheless, given the ongoing work in the field of robust depth estimation from plenoptic data [31, 34] and also in the presence of reflective surfaces [65], which aim at exploiting the 4-D light field data, we see the potential of the plenoptic camera as a passive, single camera 3-D vision sensor. This is already visible in the close-range results of Fig. 11c, where the LIF is clearly distinguishable from the background, both in size and depth, and the depth values are within the expected range.

By using extended DoF images for an actual OOS vision-based pose estimation (cf. Sec. 7 - Results 3: Vision based pose estimation) it is possible to achieve an accurate solution. In the lateral direction, the accuracy was well below the required 10 mm whereas in the direction of the optical axis it slightly exceeded it. An additional pre-processing in order to increase the brightness and contrast in the 2-D images might have helped to detect the required edges more reliable. But given the low brightness level in the images in Fig. 13, the results are already promising. For a comprehensive analysis, additional rotation in the trajectory and more variety of the illumination conditions is required, but this first proof of concept shows that images from a plenoptic camera work well as input to a robot vision algorithm.

Overall, the experiments provided a first insight on what can be expected from plenoptic cameras during close range OOS operations. The 2-D images already provide a good level of quality, whereas the 3-D capabilities are limited by the same

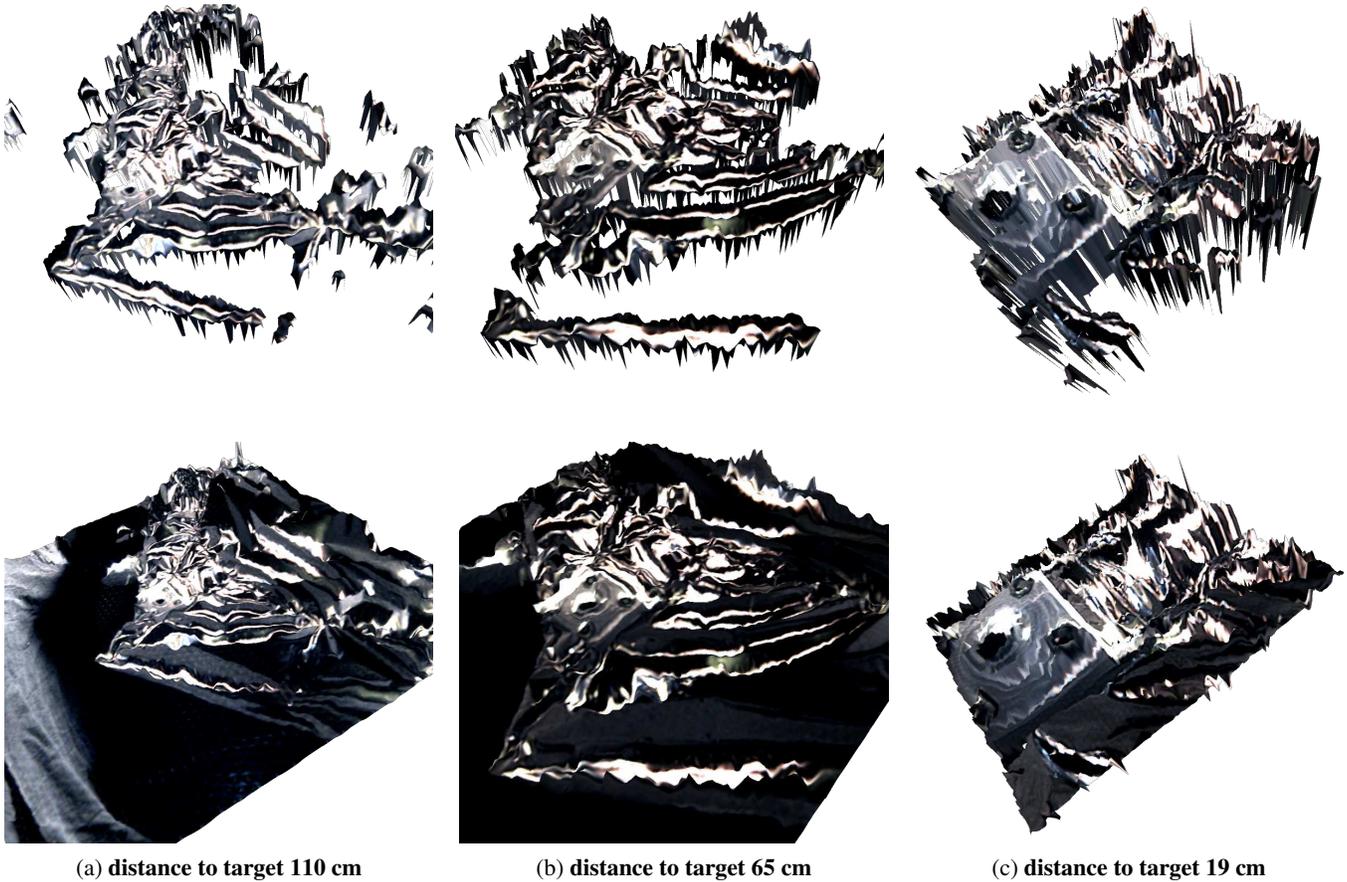


Figure 12: Tilted views of 3-D renderings for trajectory LIF3.0 at different distances. They are based on depth maps with either the enhanced depth map (top row) or with the complete fill option (bottom row). The extended DoF 2-D images are used as the texture in both images.

challenges as other passive vision sensors.

9. CONCLUSION

In this work, we introduced the plenoptic camera technology to the field of on-orbit servicing and discuss its potential as a passive vision sensor for robotic operations in orbit with an extended depth of field, single view 3-D capabilities and improved low-light performance. We covered the basic light field theory, the concepts of plenoptic vision, the state of the art in calibration and the processing of light field data to gain 2-D images with extended depth of field as well as depth maps and 3-D data. We reason that multi-focus plenoptic cameras are the most suitable concept for on-orbit servicing due to their higher lateral resolution and the larger depth of field in comparison to other plenoptic concepts. In our opinion, space-ready plenoptic cameras can be achieved with minor hardware modifications of existing conventional cameras but the development of the processing software requires the same effort as any other next generation image processing technology. We presented the results of our experiments regarding plenoptic vision for on-orbit servicing with image sequences that were recorded under representative low-light illumination conditions. This included the first application of the stepwise metric calibration of focused plenoptic cameras of [40]. We found that the extended depth of field provides sufficient image sharpness over an approach trajectory of up to 2 m distance, although at certain illumination situations

the current rendering of the 2-D images can lead to artifacts. Regardless, our on-orbit servicing vision based pose estimation of a target provided an accuracy of at least 12 mm at 20 cm distance from the camera to the target point in the direction of the optical axis. The lateral accuracy was in the range of ± 3 mm. The 3-D capabilities were investigated using the standard depth estimation for multi-focus plenoptic cameras [1] which showed to be sensitive to differences in brightness and contrast as well as to noise in the plenoptic raw images. Due to the limitation of the depth estimation algorithm to strong texture, the depth maps appeared sparse and with significant depth noise. Nevertheless, the metric depth values show the general coherence of the measurements and, even at the current quality level, the potential usability, e.g. as a hazard avoidance sensor.

In order to provide robust and reliable plenoptic vision sensors, future work requires the development of a more detailed imaging model of plenoptic camera systems than it is currently available. Furthermore, the recent developments in depth estimation and 2-D image generation require an in-depth investigation in terms of accuracy, robustness and processing speed with relevant data. Therefore, we plan to record more extensive data sets with more complex trajectories that also include rotational motion. Further investigations of light fields as a data structure, regarding their usefulness e.g. to reduce the effects of reflective surfaces [65] that are omnipresent on satellites, are of interest. Last but not least, increasing the effectiveness and speed of the light field

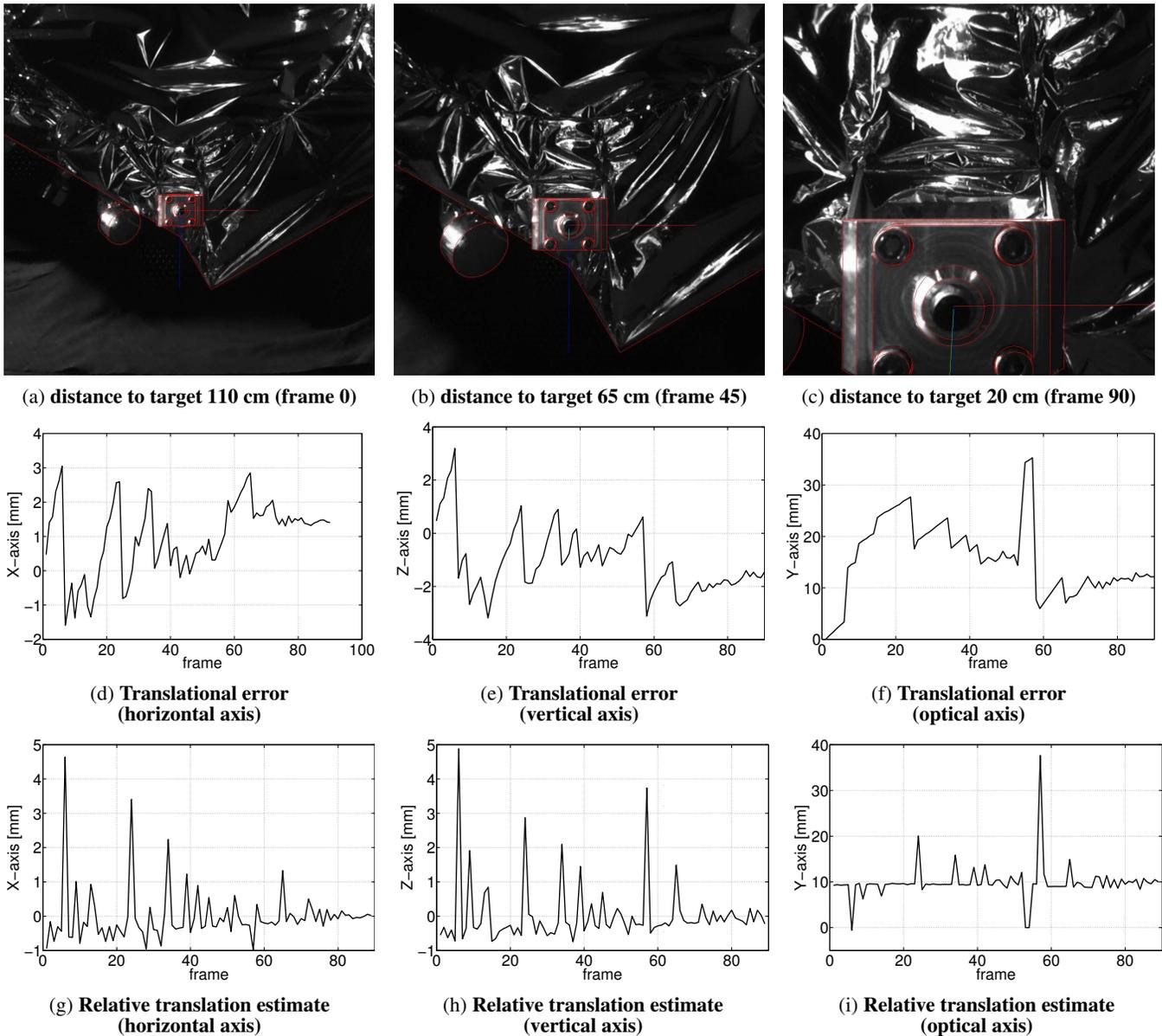


Figure 13: Performance of the vision-based pose estimation with extended DoF images. The top row shows the input images with the simplified model that was used for the algorithm. The model is positioned according to the the estimated pose. The middle row shows the absolute deviation of the estimate T_k from the ground truth \bar{T}_k whereas the bottom row shows the relative translation estimate between consecutive images

processing is an important step towards space ready plenoptic vision systems.

In the field of computer vision the topic of plenoptic vision is already well represented and with the availability of commercial plenoptic cameras as well as the required computational power, it is also gaining speed in the field of robotics. Given the long-lasting development cycles for space technology, it seems to be advisable to now start working on space-ready plenoptic systems, as we see high potential in plenoptic cameras as a future passive vision system for on-orbit servicing.

ACKNOWLEDGMENTS

We would like to thank Giorgio Panin for the technical support with the visual tracking algorithm and Michael Kaßecker for his help in setting up and calibrating the experimental setup that was used for this work. Thanks to Florian A. Fröhlich for the discussions about light field theory and plenoptic cameras and to Sabine Weiß for the support during the processing of the light field data.

REFERENCES

- [1] C. Perwass and L. Wietzke, “Single lens 3D-camera with extended depth-of-field,” in *Proc. SPIE 8291 Human Vision and Electronic Imaging XVII*, vol. 8291,

- Feb. 2012, pp. 829 108–829 108–15.
- [2] E. H. Adelson and J. Y. A. Wang, “Single lens stereo with a plenoptic camera,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 99–106, feb 1992.
 - [3] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, “Light field photography with a hand-held plenoptic camera,” Stanford University, Tech. Rep. 11, 2005.
 - [4] T. Georgiev and A. Lumsdaine, “Focused plenoptic camera and rendering,” *Journal of Electronic Imaging*, vol. 19, no. 2, pp. 021 106–021 106, 2010.
 - [5] M. Levoy and P. Hanrahan, “Light Field Rendering,” in *SIGGRAPH ’96 Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH ’96. New York, NY, USA: ACM, 1996, pp. 31–42.
 - [6] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen, “The Lumigraph,” in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH ’96*, - 1996, p. nil.
 - [7] D. G. Dansereau, “Plenoptic signal processing for robust vision in field robotics,” Ph.D. dissertation, University of Sydney, 2014.
 - [8] M. Lingenauber, S. Kriegel, M. Kaßecker, and G. Panin, “A dataset to support and benchmark computer vision development for close range on-orbit servicing,” in *Proceedings of ASTRA 2015 - 13th Symposium on Advanced Space Technologies in Robotics and Automation. ESA (European Space Agency)*, 2015.
 - [9] NASA Goddard Space Flight Center, “On-Orbit Satellite Servicing Study Project Report,” NASA, Tech. Rep., 2010.
 - [10] S. Persson, S. D’Amico, and J. Harr, “Flight Results from PRISMA Formation Flying and Rendezvous Demonstration Mission,” in *61st International Astronautical Congress*, 2010.
 - [11] S. D’Amico, J.-S. Ardaens, and R. Larsson, “Spaceborne Autonomous Formation-Flying Experiment on the PRISMA Mission,” *Journal of Guidance, Control, and Dynamics*, vol. 35, no. 3, pp. 834–850, 2012.
 - [12] Stephen Clark, “In-space satellite servicing tests come to an end,” web, Spaceflight Now, 2007. [Online]. Available: <http://spaceflightnow.com/news/n0707/04orbitalexpress/>
 - [13] NASA, “RRM Task Overview,” 2016. [Online]. Available: https://sspd.gsfc.nasa.gov/rrm_tasks.html
 - [14] —, “Robotic Refueling Mission — Phase 2,” 2016. [Online]. Available: https://sspd.gsfc.nasa.gov/rrm_phase2.html
 - [15] M. Strube, R. Henry, E. Skeleton, J. V. Eepoel, N. Gill, and R. McKenna, “Raven : An On-Orbit Relative Navigation Demonstration Using International Space Station Visiting Vehicles,” in *2015 AAS GN&C Conference*, 2015.
 - [16] NASA, “Creating Autopilot for Spacecraft And How Raven Helps,” 2016. [Online]. Available: <https://sspd.gsfc.nasa.gov/raven.html>
 - [17] E. H. Adelson and J. R. Bergen, “The Plenoptic Function and the Elements of Early Vision,” *Computational Models of Visual Processing*, pp. 3–20, 1991. [Online]. Available: http://web.mit.edu/persci/people/adelson/pub_abstracts/elements91_abs.html
 - [18] G. Wetzstein, I. Ihrke, D. Lanman, and W. Heidrich, “Computational Plenoptic Imaging,” *Computer Graphics Forum*, vol. 30, no. 8, pp. 2397–2426, 2011.
 - [19] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, “Image and depth from a conventional camera with a coded aperture,” *ACM transactions on graphics (TOG)*, vol. 26, no. 3, p. 70, 2007.
 - [20] T. G. Georgiev and A. Lumsdaine, “Resolution in plenoptic cameras,” in *Computational Optical Sensing and Imaging*. Optical Society of America, 2009, p. CTuB3.
 - [21] G. Lippmann, “Épreuves réversibles. Photographies intégrales,” *Comptes Rendus de l’Académie des Sciences*, vol. 146, no. 9, pp. 446–451, 1908.
 - [22] A. Lumsdaine and T. Georgiev, “The focused plenoptic camera,” in *Computational Photography (ICCP), 2009 IEEE International Conference on*. IEEE, 2009, pp. 1–8.
 - [23] T. G. Georgiev and A. Lumsdaine, “Superresolution with plenoptic 2.0 cameras,” in *Signal recovery and synthesis*. Optical Society of America, 2009, p. STuA6.
 - [24] T. E. Bishop and P. Favaro, “The Light Field Camera: Extended Depth of Field, Aliasing, and Superresolution,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 972–986, May 2012.
 - [25] F. Pérez, A. Pérez, M. Rodríguez, and E. Magdaleno, “Super-Resolved Fourier-Slice Refocusing in Plenoptic Cameras,” *Journal of Mathematical Imaging and Vision*, vol. 52, no. 2, pp. 200–217, 6 2015.
 - [26] T. E. Bishop, S. Zanetti, and P. Favaro, “Light field superresolution,” in *Computational Photography (ICCP), 2009 IEEE International Conference on*, Apr. 2009, pp. 1–9.
 - [27] O. Pizarro, S. B. Williams, M. V. Jakuba, M. Johnson Roberson, I. Mahon, M. Bryson, D. Steinberg, A. Friedman, D. Dansereau, N. Nourani Vatani, D. Bongiorno, M. Bewley, A. Bender, N. Ashan, and B. Douillard, “Benthic monitoring with robotic platforms - The experience of Australia,” in *Underwater Technology Symposium (UT), 2013 IEEE International*, March 2013, pp. 1–10.
 - [28] F. Dong, S.-H. Ieng, X. Savatier, R. Etienne Cummings, and R. Benosman, “Plenoptic cameras in real-time robotics,” *The International Journal of Robotics Research*, vol. 32, no. 2, pp. 206–217, 2013. [Online]. Available: <http://ijr.sagepub.com/content/32/2/206.abstract>
 - [29] D. G. Dansereau, I. Mahon, O. Pizarro, and S. B. Williams, “Plenoptic flow: Closed-form visual odometry for light field cameras,” in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2011, pp. 4455–4462.
 - [30] N. Zeller, F. Quint, and U. Stilla, “Narrow Field-Of Visual Odometry Based on a Focused Plenoptic Camera,” *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 1, pp. 285–292, 2015.
 - [31] Niclas Zeller, Franz Quint, and Uwe Stilla, “Depth estimation and camera calibration of a

- focused plenoptic camera for visual odometry,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 118, pp. 83 – 100, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924271616300508>
- [32] D. G. Dansereau, O. Pizarro, and S. B. Williams, “Linear volumetric focus for light field cameras,” *ACM Transactions on Graphics (TOG)*, vol. 34, no. 2, p. 15, 2015.
- [33] D. G. Dansereau, D. L. Bongiorno, O. Pizarro, and S. B. Williams, “Light field image denoising using a linear 4D frequency-hyperfan all-in-focus filter,” in *Proceedings of the SPIE Conference on Computational Imaging (SPIE’13)*, vol. 8657, 2013.
- [34] Shuo Zhang, Hao Sheng, Chao Li, Jun Zhang, and Zhang Xiong, “Robust depth estimation for light field via spinning parallelogram operator,” *Computer Vision and Image Understanding*, vol. 145, pp. 148–159, 2016, light Field for Computer Vision. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1077314215002714>
- [35] Optical Society of America, *Handbook of Optics*, 3rd ed., M. Bass, C. DeCusatis, J. Enoch, V. Lakshminarayanan, G. Li, C. Macdonald, V. Mahajan, and E. Van Stryland, Eds. New York, NY, USA: McGraw-Hill Education, 2010, vol. 1.
- [36] A. Lumsdaine, T. G. Georgiev, and G. Chunev, “Spatial analysis of discrete plenoptic sampling,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 829 909–829 909.
- [37] Lytro Inc., “Lytro - Home,” web, 2016. [Online]. Available: <https://www.lytro.com/>
- [38] T. Georgiev and A. Lumsdaine, “The multifocus plenoptic camera,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 829 908–829 908.
- [39] Raytrix GmbH, “3D light field camera technology,” web, 2016. [Online]. Available: <https://www.raytrix.de/>
- [40] Klaus H. Strobl and Martin Lingenauber, “Stepwise calibration of focused plenoptic cameras,” *Computer Vision and Image Understanding*, vol. 145, pp. 140 – 147, 2016, light Field for Computer Vision. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S107731421500274X>
- [41] N. Zeller, F. Quint, and U. Stilla, “Calibration and accuracy analysis of a focused plenoptic camera,” *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 1, pp. 205–212, 2014.
- [42] N. Zeller, C. Noury, F. Quint, C. Teulière, U. Stilla, and M. Dhome, “Metric Calibration of a Focused Plenoptic Camera Based on a 3d Calibration Target,” *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 449–456, 2016.
- [43] O. Johannsen, C. Heinze, B. Goldluecke, and C. Perwaß, “On the Calibration of Focused Plenoptic Cameras,” in *GCPR Workshop on Imaging New Modalities*, 2013.
- [44] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov 2000.
- [45] Klaus H. Strobl and Gerd Hirzinger, “Optimal Hand-Eye Calibration,” in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2006, pp. 4647–4653.
- [46] K. H. Strobl and G. Hirzinger, “More accurate pinhole camera calibration with imperfect planar target,” in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1068–1075.
- [47] T. Georgiev, G. Chunev, and A. Lumsdaine, “Superresolution with the focused plenoptic camera,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2011, pp. 78 730X–78 730X.
- [48] S. Wanner, J. Fehr, and B. Jähne, “Generating EPI representations of 4D light fields with a single lens focused plenoptic camera,” in *International Symposium on Visual Computing*. Springer, 2011, pp. 90–101.
- [49] H. H. Baker and R. C. Bolles, “Generalizing epipolar-plane image analysis on the spatiotemporal surface,” *International Journal of Computer Vision*, vol. 3, no. 1, pp. 33–49, 1989.
- [50] T. E. Bishop and P. Favaro, “Full-resolution depth map estimation from an aliased plenoptic light field,” in *Asian Conference on Computer Vision*. Springer, 2010, pp. 186–200.
- [51] H.-G. Jeon, J. Park, G. Choe, J. Park, Y. Bok, Y.-W. Tai, and I. S. Kweon, “Accurate depth map estimation from a lenslet light field camera,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2015, pp. 1547–1555.
- [52] C. BVBA, “CMV4000,” 2016. [Online]. Available: http://www.cmosis.com/products/product_detail/cmv4000
- [53] R. Ng, “Fourier slice photography,” in *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3. ACM, 2005, pp. 735–744.
- [54] D. G. Dansereau, “A 4-D dual-fan filter bank for depth filtering in light fields,” *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 542–549, 2007.
- [55] A. Lumsdaine, G. Chunev, and T. Georgiev, “Plenoptic rendering with interactive performance using GPUs,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 829 513–829 513.
- [56] A. Madanayake, R. Wimalagunaratne, D. G. Dansereau, and L. T. Bruton, “Design and FPGA-implementation of 1 st-order 4D IIR frequency-hyperplanar digital filters,” in *2011 IEEE 54th international midwest symposium on circuits and systems (MWSCAS)*. IEEE, 2011, pp. 1–4.
- [57] D. Reintsema, J. Thaeter, A. Rathke, W. Naumann, P. Rank, and J. Sommer, “DEOS—the German robotics approach to secure and de-orbit malfunctioned satellites from low earth orbits,” in *Proceedings of the i-SAIRAS*, 2010, pp. 244–251.
- [58] T. Drummond and R. Cipolla, “Real-time Tracking of Complex Structures with On-line Camera Calibration,” in *BMVC*, 1999, pp. 57.1–57.10.
- [59] E. Marchand, P. Bouthemy, and F. Chaumette, “A 2D-3D model-based approach to real-time visual tracking,” *IVC*, vol. 19, no. 13, pp. 941–955, 2001.
- [60] A. I. Comport, D. Kragic, É. Marchand, and F. Chaumette, “Robust Real-Time Visual Tracking: Comparison, Theoretical Analysis and Performance Evaluation,” in *IEEE ICRA*, Barcelona, Spain, April 2005, pp. 2841–2846.

- [61] N. W. Oumer, G. Panin, Q. Mühlbauer, and A. Tseneklidou, "Vision-based Localization for On-Orbit Servicing of a partially Cooperative Satellite," *Acta Astronautica*, vol. 117, pp. 19–37, 2015.
- [62] A. Petit, E. Marchand, and K. Kanani, "Tracking complex targets for space rendezvous and debris removal applications," in *IEEE/RSJ IROS*, 2012, pp. 4483–4488.
- [63] A. Hofhauser, C. Steger, and N. Navab, "Edge-Based Template Matching and Tracking for Perspectively Distorted Planar Objects," in *Proceedings of the 4th International Symposium on Advances in Visual Computing (ISVC)*, vol. 5358, 2008, pp. 35–44.
- [64] J. Canny, "A Computational Approach to Edge Detection," *IEEE PAMI*, vol. 8, no. 6, pp. 679–698, Jun. 1986.
- [65] S. Wanner and B. Goldluecke, *Reconstructing Reflective and Transparent Surfaces from Epipolar Plane Images*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 1–10. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40602-7_1

BIOGRAPHY



Martin Lingenauber is a research scientist at the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR) in Oberpfaffenhofen, Germany. In 2010 he graduated with distinction as a Dipl.-Ing. in aerospace engineering from the Technical University of Berlin. From 2010 until 2011 he was with the control systems department of the European Space Agency (ESA) in Noordwijk, Netherlands where he worked on testing components for attitude control systems and on new vision sensor concepts. Since 2011 he is with the Institute of Robotics and Mechatronics and works on robot vision for space applications and plenoptic cameras for robotics.



Klaus H. Strobl is a research scientist at the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR) in Oberpfaffenhofen, Germany. His research interests focus on computer vision, 3-D graphics, camera calibration, mobile robotics, and deep learning. Klaus studied electrical engineering (automatic control) at the Universidad de Navarra (Spain), the Vienna University of Technology (Austria), the Technische Universität München (Germany), and the Norwegian University of Science and Technology (Norway). He held a visiting researcher position at the Department of Computing, Imperial College London in 2009 and earned his Ph.D. *summa cum laude* in electrical engineering in 2014 at Technische Universität München.



Nassir W. Oumer received his M.Tech degree in Electrical Engineering from the Indian Institute of Technology Madras in 2006, M.Sc. degree in Space Robotics and Automation from the Helsinki University of Technology, with distinction in 2009, and a Ph.D. in Computer Science from the University of Osnabrück in 2016. Since 2010, Nassir has been working at the German Aerospace Center (DLR), and he has been involved in various On-orbit Servicing projects, researching and developing vision-based methods for rendezvous and capture of a satellite. His responsibilities include research and development of vision-based 3-D tracking methods for rendezvous, capture and vision-based control of a satellite.



Simon Kriegel is a senior researcher at the Department of Perception and Cognition of the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR). From the Technical University of Munich, he received his Dipl.-Ing. degree in Information Technology in 2008 and his Ph.D. in computer science in 2015. He started his career at KUKA Roboter GmbH, where he worked on different research projects in the field of object detection and manipulation for unloading tasks with industrial robots. Since 2009 he is with the DLR where his research focuses on 3-D modeling, Next-Best-View planning and exploration.