ROBUST DEPTH ESTIMATION FROM MULTI-FOCUS PLENOPTIC IMAGES

Francisco Cunha[1], Lucas A. Thomaz[1,2], Luis M. N. Tavora[2], Pedro A. A. Assunção[1,2], Rui Fonseca-Pinto[1,2], and Sérgio M. M. Faria[1,2]

[1]Instituto de Telecomunicações, Portugal
[2]ESTG, Polytechnic of Leiria, Portugal,
e-mails: {francisco.cunha,lucas.thomaz, amado, sergio.faria}@co.it.pt,
{luis.tavora, rui.pinto}@ipleiria.pt

ABSTRACT

This paper describes a robust depth estimation algorithm for multi-focus plenoptic images. The main feature of the proposed method consists of a hybrid template matching scheme built-upon intensity and local phase information, which adapts to the blurriness of neighbouring lenslet microimages. By reducing the impact of defocus-blur on the template matching accuracy, the proposed method efficiently handles the varying triangulation baseline over the depth-of-field, thus discarding the need for scene-related information such as the expected range of disparities. Experimental results demonstrate the robustness of the proposed method over the most used commercially available depth estimation algorithm, achieving a reduction of 73% on the depth estimation error.

***Index Terms*—** Light Field, Depth Estimation, Multi-focus, Lenslet

1. INTRODUCTION

Light-field (LF) cameras consist of an optical system composed by a principal lens and a micro-lens array (MLA) mounted in front of an image sensor. Its main characteristic is the ability to capture both the intensity and angular information of the light rays reaching the main lens, i.e., dense 3D radiance data, from a single-shot capture. Given its inherent 3D capabilities, plenoptic technology has been driving research efforts in different application fields, such as computer vision, computational photography, among others.

Multi-focus plenoptic images are acquired by the so-called plenoptic cameras 2.0 using specific arrangements of micro-lenses with different focal types [1]. One of such cameras is the Raytrix [2], comprising interlaced arrangement of micro lenses with 3 different focal lengths according to the layout depicted in Figure 1. Due to the inclusion of micro-lenses with variable focal length, multi-focus plenoptic images exhibit different blurring for a given working distance because adjacent micro-images represent the same image texture with different focus levels. This effect constitutes an issue for depth estimation (DE), as state-of-the-art lenslet-based techniques rely on the definition of point-to-point correspondences between neighbouring micro-images. As such methods are based on conventional template matching approaches, whose performance is highly sensitive to image degradation such as out-of-focus blur, DE from multi-focus plenoptic images is not consistently accurate for any type of scene.

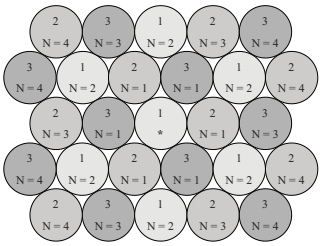


Fig. 1: Multi-focus MLA arrangement. The categorical index encodes the f-type and N denotes the neighbourhood order with respect to the central micro-lens marked with *.

This paper addresses the problem of DE from multi-focus plenoptic images through the adoption of a defocus-robust local texture descriptor. This allows the definition of multi-view point correspondences, while mitigating the impact of degradation caused by micro-images with different levels of blur. An adaptive approach based on the concept of defocus coherence of neighbouring micro-images is proposed, thus enabling the definition of point correspondences in complementary domains (intensity and frequency-based), according to the blurriness of the matched micro-images. The overall results show that the proposed algorithm preserves the accuracy of the DE throughout the depth-of-field of the camera, which is important in applications where the precision is critical *e.g.* medical and industrial ones.

The remainder of the paper is organized as follows: Section 2 presents the related work and describes Local Phase Quantization (LPQ) as a defocus-robust local texture descriptor. Section 3 introduces the proposed method and Section 4 discusses the experimental results. Finally, Section 5 draws the conclusions and final remarks.

2. BACKGROUND

The dense 4D data obtained by LF cameras offers multiple possibilities for DE. One of the most common approaches involves sampling 2D slices of fixed angular and spatial direction, usually referred to as Epipolar Plane Images (EPIs) [3], which allows one to compute the scene depth by correlating the respective edges slopes to disparity measurements [4, 5]. Since for 2.0 LF cameras the process of view demultiplexing on the basis of EPI formation depends on a preliminary DE [6], the following review is restricted to DE methods that

This work was supported by Programa Operacional Regional do Centro, project PLenoISLA POCI-01-0145-FEDER-028325 and by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/EEA/50008/2020, Portugal.

exclusively operate on raw lenslet images, which are available without any further processing or additional information.

The authors of [7] propose a dense DE scheme relying on the semi-global matching paradigm [8]. The method involves a dense disparity estimation based on photometric cost volume aggregation computed over an expected range of disparities. It adaptively adjusts the set of neighboring micro images that are considered for point matching, according to a spatial coverage criterion, and further includes a semi-global regularization step, as well as a supplementary regularization for insufficiently textured regions.

A different approach is presented in [9], where a feature-based scheme aims to reduce the complexity of the previous semi-global matching. The complexity reduction results from sparse detection and geometrically-constrained matching of points of interest. The obtained point correspondences are triangulated using a method based on a RANSAC. Nonetheless, the main contribution of such method consists of a depth refinement procedure based on quartile sectioning of the depth range, aiming to weight the contribution of each micro lens type, depending on the axial position of the triangulated correspondences. Compared to [7], this method presents lower computational complexity at the cost of a reduced accuracy.

The discussed methods [7, 9] are not adaptable, as they both heavily depend on a set of operating parameters that must be specifically tuned beforehand. For instance, in [7], the range of expected disparities, the maximum triangulation baseline, as well as smoothness factors related to the regularization steps must be set according to focus and magnification settings. Analogously, in [9], both the predetermined partitioning scheme of the depth range, as well as the definition of linear combinators to the proposed weighting scheme for depth refinement constitute an operating downside.

To overcome these limitations, an alternative approach for the definition of correspondences based on LPQ [10] is proposed. As further detailed below, this method does not require scene-related information (*e.g* maximum allowed triangulation baseline / range of expected disparities), though it still enables handling the varying triangulation baseline of 2.0 LF cameras.

2.1. Local phase quantization and spatial defocus invariance

The discrete model that relates an isotropically blurred image *I'* and its original version *I* is expressed in the frequency domain as

$$\mathscr{F}\{I'\} = \mathscr{F}\{I\} \cdot \mathscr{F}\{h\}, \tag{1}$$

where $\mathscr{F}$ denotes the discrete Fourier transform and h the blur response function. By decomposing Eq. (1) into its magnitude and phase components, the latter is given by

$$\angle I'(u,v) = \angle I(u,v) + \angle h(u,v), \tag{2}$$

in which $\angle$ represents the phase angle of the transformed elements and u, v denote a pair of arbitrary frequency components. Assuming the blur response function h to be centrally symmetric, as in the case of defocus blur kernels approximated by Gaussian-like shapes, it can be demonstrated that the corresponding Fourier transform only admits real values, thus forcing the $\angle h(u,v)$ to assume the form of a binary-valued function given by

$$\angle h(u,v) = \begin{cases} 0, & \text{if } h(u,v) \geqslant 0 \\ \pi, & \text{else} \end{cases}. \tag{3}$$

Hence, by applying Eq. (3) into Eq. (2), blur invariance can be achieved with respect to the Fourier phase according to

$$\angle I'(u,v) = \angle I(u,v),\ \forall\,(u,v) : h(u,v) \geqslant 0. \tag{4}$$

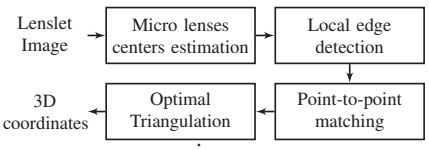


Fig. 2: Workflow of the proposed depth estimation framework.

Therefore, a defocus robust local description can be obtained by properly characterizing local phase information. Since defocus blur has a low-pass essence, LPQ allows for such pixel-based description through a compact representation of the short-term Fourier transform (STFT) coefficients of the four lowest frequency components *i.e.* $(u,v) \in \{(1/s,0),(0,1/s),(1/s,1/s),(1/s,-1/s)\}$, with s standing for the size of the STFT support window. Then, the concatenation of the resulting real and imaginary parts of the computed Fourier responses of each frequency yields a vector G, composed by 8 coefficients that undergo scalar quantization according to

$$q_i = \begin{cases} 0, & \text{if } G_i \leqslant 0 \\ 1, & \text{else} \end{cases}. \tag{5}$$

Lastly, the resulting quantized coefficients are transformed through binary encoding, enabling the assignment of a blur insensitive 8-bit integer to each image pixel. Different results of local phase description can be obtained by changing the transform scheme and/or by introducing variants to the quantizer, as proposed in [11, 12]. For the reported system, STFT was used to compute the local frequency coefficients given its efficiency and robustness when compared, for instance, with Gabor filter banks [13]. In this work, the originally proposed LPQ quantization scheme [10] has been adopted due to its simplicity.

3. PROPOSED METHOD

The proposed method consists on four sequential steps, as represented in Figure 2. Besides the improved awareness to defocus-related artifacts offered by LPQ, its adoption in the DE framework allows this to run without any prior guess of the maximum triangulation baseline. In fact, as small triangulation baselines allow DE for both near and far objects [14], one can refrain from assuming a maximum allowed triangulation baseline if correspondences are iteratively established between consecutive first-order neighbouring micro images. However, in the case of multi-focus LF cameras, such process involves the definition of correspondences based on radiance information acquired by micro lenses with different focal length, as depicted in Figure 1. A possible solution to enable template matching in this case employs image restoration algorithms. However, such methods require the knowledge of the point-spread functions of both micro lenses, which typically involves a complex and ill-posed estimation process [15]. Alternatively, the adoption of LPQ into the DE framework is exploited in this work to bolster the definition of point correspondences between different micro lenses, even when defocus-related degradation occur.

3.1. Micro lenses centers estimation

Since the MLA can be modeled as a camera array, the first step of the proposed method involves the estimation of the center of the micro lenses, defined in image coordinates given that the MLA is assumed

fronto-parallel relatively to the image sensor. This is done according to the methodology proposed in [16].

3.2. Local edge detection

Given the distribution of micro lenses centers, depending on the texture of the scene, a suitable edge detection algorithm is locally executed at micro images of interest, in order to define a sparse support domain for subsequent matching. The adaptation of edge detection to operate *per* micro image is quite relevant for DE in 2.0 plenoptic cameras, since point correspondences can only be unambiguously defined in regions with sufficient detail [2]. The method proposed in this paper employs threshold-based binarization followed by morphological opening and dilation operations for automatic detection of salient micro images, *i.e.* micro images that present object edges, combined with local Otsu-based thresholding and binary border following [17] to detect edges directly at micro images of the lenslet (see Figure 3c).

3.3. Point-to-point matching

In the context of the proposed DE algorithm, the previously obtained local edge maps have a two-fold relevance. First, they constitute the sparse domain Ω of reference blocks on a given micro image m_{Ref}, to be matched against their homologous at a target micro image, m_{Tgt}. Furthermore, local edge maps allow to estimate the blurriness b_σ of each micro image of interest.

Using the defocus blur as a depth cue, it is proposed to build a decision criterion around b_σ to adapt the domain on which the point correspondences between m_{Ref} and m_{Tgt} are defined. Since for specific narrow depth bands, neighbouring micro lenses with different focal length still retain similar texture detail, the desired criterion should allow to eliminate the need of LPQ-based template matching on such cases. This condition is desirable as intensity-based template matching boosts the density of depth estimates. This notion is clearer in light of the formulation of LPQ for local texture description (see section 2.1). Briefly, the size s of the discrete support window used to compute the local frequency responses dictates the trade-off between blur robustness and local discriminability [10], which consequently determines the minimum patch size to define templates for matching on the LPQ-based image. The blurriness b_σ for each micro image of interest m_* was estimated, according to the principle of gradient ratio [18], as

$$b_\sigma^{m_*} = \sum_{i,j} \frac{\nabla m_*[i,j]}{\nabla m'_*[i,j]}, i, j \in \Omega, \tag{6}$$

where m'_* stands for a Gaussian filtered version of m_* using sigma = 1.5. Given the circular shape of the microimages, a normalized convolution was used since the binary mask of each microimage domain was available. Then, the quantity

$$\Delta_{b_\sigma} = ||b_\sigma^{m_{Ref}} - b_\sigma^{m_{Tgt}}|| \tag{7}$$

was computed and aggregated for the set of first-order neighbouring micro images, leading to a list of Δ_{b_σ} values used for K-Means clustering with $K = 2$ as an approach to categorize the difference between the gradient ratios of two micro images in terms of a defocus coherence index K_c. Accordingly, depending on the value of K_c returned by the clustering algorithm for a given pair of micro images, the template matching domain can adaptively change. In this paper two operating domains for template matching are considered, either based on intensity or on LPQ-based image.

Algorithm 1: Finding point-to-point correspondences between neighbouring micro images (m_{Ref}, m_{Tgt})

input : $\{p_I, p_{LPQ}\}$ (Intensity and LPQ-based template matching patch sizes)
output: $\Omega_{m_{Ref}} \longleftrightarrow \Omega_{m_{Tgt}}$
Compute Δ_{b_σ} (m_{Ref}, m_{Tgt});
Find nearest K-Means centroid, K_c;
if $K_c = 0$ **then**
  $operatingDomain \leftarrow$ intensity template matching;
  $metric \leftarrow$ normalized cross correlation;
else
  $operatingDomain \leftarrow$ LPQ-based template matching;
  $metric \leftarrow$ sum of absolute differences;
end
Get sparse ref. domain, $\Omega_{m_{Ref}}$ and target domain, $\Omega_{m_{Tgt}}$;
for $x_R \in \Omega_{m_{Ref}}$ **do**
  for $x_T \in \Omega_{m_{Tgt}}$ **do**
    Test for epipolar consistency $Ec(x_R, x_T)$;
  end
  for *all* $x_T \iff Ec(x_R, x_T) = 1$ **do**
    Compute similarity $s(x_R, x_T)$;
  end
  $x_T^* \leftarrow \underset{x_T}{\operatorname{argmax}}\, s(x_R, x_T)$;
  Test for reciprocal consistency, $Rc(x_R, x_T^*)$;
  if $Rc(x_R, x_T^*) = 1$ **then**
    if $x_T^* \rightarrowtail x_R$ **then**
      Assign correspondence;
    else
      Discard correspondence (x_R, x_T^*);
    end
  else
    Discard correspondence (x_R, x_T^*);
  end
end

Having set the decision criterion in the defocus coherence space, the next step involves the definition of point-to-point correspondences between neighbouring micro images. The process is generic and independent for all the remaining pairs of micro images, thus being suitable for parallel processing.

Briefly, the patch centered at each edge coordinate of m_{Ref} serves as reference template for a given iteration. Then, by taking the edge coordinates of m_{Tgt}, as well as the coordinates of both m_{Ref} and m_{Tgt} centers, all the target candidate points located farther than 1 pixel from the epipolar line starting at x_R are discarded $(E_c(x_R, x_T))$. Next, the patch similarity of each geometrically-consistent candidate is computed, thus allowing to retain the best matching candidate and to define a temporary correspondence. Next, such process is performed in the reversed order of the elements that compose the temporary correspondence in order to ascertain its bidirectionality $(R_c(x_R, x_T^*))$. The last validation focuses on the injectivity of the final correspondences. To this end, if more than one point of the domain of m_{Tgt} has been matched against x_R, only the one with higher similarity score is kept.

The previous operations lead to a set of point correspondences between m_{Ref} and m_{Tgt}. A single correspondence between points from the two mentioned micro images is given by $x_R \longleftrightarrow x_T^*$ If m_{Tgt} can be considered as source of reference points matched against candidates at a micro image different than m_{Ref} a new correspondence is defined by $x_T^* \longleftrightarrow x_T^{**}$. By deduction, a simple aggregation scheme can be employed to group correspondences that refer to the same feature, which are replicated in a variable number of micro images. This simple mechanism precludes the need for any prior information about the varying triangulation baseline over the depth of field.

3.4. Optimal triangulation

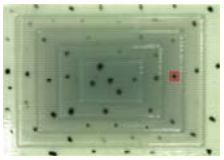
This step involves estimating a 3D point for each group of correspondences. In 2.0 LF cameras, metric depth estimates can be obtained if the distance between the MLA and the image sensor is known. This happens as depth estimates are proportional of such distance, thus resulting in the so-called virtual depth [2]. In practice, the virtual depth can be obtained if the position of an image formed behind the sensor is known. From a geometrical point of view, the distance between the MLA and the image sensor equals the projection distance that a ray passing through the center of a micro lens traverses until it hits the image sensor. Assuming the MLA as a grid of pinhole cameras at the plane $Z = 0$, each group of correspondences can be represented by a group of rays (*i.e.* lines) with origin at the plane $Z = 0$ that hit the sensor plane at $Z = -b$. Thus, the point correspondences and the respective micro lenses centers (in image coordinates), can be expanded into three-dimensional space, allowing one to formulate the estimation of an image point behind the sensor as a line intersection problem; in this work, a least-squares approach [19], which optimizes the 3D estimation by minimizing the re-projection error, was used.

4. EXPERIMENTAL ASSESSMENT

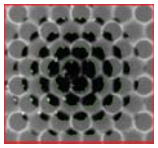
To evaluate the performance of the proposed DE method, a practical experiment was designed using a 3D object, consisting of five step planes equally spaced (0.250 mm) in the axial dimension, as represented in Figure 3a. A commercially available 2.0 plenoptic camera Raytrix R42 [20] was used to obtain a raw lenslet image of the object. For the purpose of converting virtual depth estimates into metric depth the plenoptic camera was calibrated, as described in [21]. Furthermore, to allow image-based DE, texture was applied (painted black dots) over the originally uniform 3D object (Figure 3b). The raw lenslet image was processed to detect groups of micro images of interest, as reported in Section 3, and an example of such detection is shown (Figure 3c). The proposed point-to-point matching and triangulation were applied in each detected region, resulting in multiple 3D coordinates. For this experiment, an optimal set of execution parameters has been heuristically determined, namely: a size of 7 for both the intensity and LPQ-based templates, normalized cross-correlation as similarity function for intensity-based template matching and sum of absolute differences, as the similarity function for LPQ-based template matching. Since the micro images were found to have a valid inner domain (free of vignetting) of approximate 31 pixels, and the STFT support window size used to generate the LPQ image representation was 5 pixels, thus the employed patch sizes was 7 pixels.

The limits of the five object planes were known in image coordinates, allowing to group the depth estimations *per* plane using the central coordinate of each salient region. This methodology resulted in five distributions of depth estimates. The metric depth values were, then, obtained according to the camera calibration. For comparison purposes, the raw depth map provided by the proprietary Raytrix software was also obtained. As such map is defined on lenslet image coordinates, the same experimental methodology was applied to it. Also, a minimum confidence threshold of normalized cross correlation of 0.90 was set to preserve depth estimates at the most salient micro images. Using both methods, the average of the depth estimations in each object level was assessed, and the step size between neighbour planes was estimated by subtracting their depths of consecutive planes, as indexed in Figure 3a. The results, as well as the error between the estimation and the experimental value of 0.250 mm are presented in Table 1.

(a) Side view of the 3D model, with plane levels of distances of 0.25 mm.



(b) Texturized 3D piece.



(c) Micro images and local edges.

Fig. 3: Depth estimation experimental setup.

Table 1: Metric distance in mm (and estimation difference for the experimental value of 0.25mm) for the DE by the proposed and Raytrix algorithms.

Step Size Estimation	Proposed	Raytrix [2]
Δ_{AB}	**0.274 (0.024)**	0.325 (0.075)
Δ_{BC}	**0.248 (0.002)**	0.242 (0.008)
Δ_{CD}	**0.249 (0.001)**	0.191 (0.059)
Δ_{DE}	**0.284 (0.034)**	0.168 (0.082)
Average	**0.264 (0.015)**	0.232 (0.056)

Table 1 evidences the improved depth estimations of the proposed method relatively to the Raytrix software. On average, the obtained error is around 0.015 mm, compared to 0.056 mm of the Raytrix method, thus a reduction of 73% in the absolute error. Furthermore, the proposed method presents higher invariance throughout the depth range, which suggests the validity of LPQ to defocus insensitive DE. It should be noted, however, that a detailed error analysis was not performed as the uncertainty of depth estimations varies with respect to the local image contrast and this relation was not possible to derive. For the best of the authors knowledge, no publicly available datasets of both ground truth depth maps and metric calibration parameters (including the distance between the MLA and sensor) exist. Hence, the presented experiment focused on the comparison with the Raytrix proprietary-software, since all the necessary information for the geometric triangulation is available.

5. CONCLUSIONS

This paper proposes a DE algorithm for multi-focus plenoptic images that simultaneously addresses the problems of sensitivity to defocus-blur and variable triangulation baseline over the depth of field. By adaptively switching between intensity and LPQ domains, the proposed matching framework offers robustness to blur. Furthermore, by embedding LPQ in the matching framework, it finds correspondences between neighbouring micro images of varying focal length, with high impact on the generalization of the algorithm, regardless the range of disparities of the input lenslet image. The proposed method proves to be superior to the commercially available Raytrix algorithm, reducing the DE absolute error from 0.056 mm to 0.015 mm and reveals improved invariance across the depth range. In the future it is expected to further associate an uncertainty map to support subsequent post-processing steps such as filtering and depth inpainting.

6. REFERENCES

[1] T. Georgiev and A. Lumsdaine, "The multi-focus plenoptic camera," 1 2012, vol. 8299, SPIE.

[2] C. Perwaß and L. Wietzke, "Single lens 3D-camera with extended depth-of-field," in *Human Vision and Electronic Imaging XVII*, Burlingame, CA, USA, January 2012, vol. 8291, pp. 45 – 59, SPIE.

[3] S. Wanner, J. Fehr, and B. Jähne, "Generating epi representations of 4D light fields with a single lens focused plenoptic camera," in *Advances in Visual Computing"*, Berlin, Heidelberg, 2011, pp. 90–101, Springer.

[4] S. Zhang, H. Sheng, C. Li, J. Zhang, and Z. Xiong, "Robust depth estimation for light field via spinning parallelogram operator," *Computer Vision and Image Understanding*, vol. 145, pp. 148–159, April 2016.

[5] S. Wanner and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. March, pp. 606–619, March 2014.

[6] T. Georgiev and A. Lumsdaine, "Focused plenoptic camera and rendering," *Journal of Electronic Imaging*, vol. 19, pp. 1–11, 2010.

[7] O. Fleischmann and R. Koch, "Lens-based depth estimation for multi-focus plenoptic cameras," in *Pattern Recognition*, Münster, Germany, September 2014, pp. 410–420, Springer.

[8] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 328–341, February 2008.

[9] R. Ferreira and N. Gonçalves, "Fast and accurate micro lenses depth maps for multi-focus light field cameras," in *Pattern Recognition*, Hannover, Germany, September 2016, pp. 309–319.

[10] E. Rahtu, J. Heikkilä, V. Ojansivu, and T. Ahonen, "Local phase quantization for blur-insensitive image analysis," *Image and Vision Computing*, vol. 30, pp. 501—-512, August 2012.

[11] J. Päivärinta, E. Rahtu, and J. Heikkilä, "Volume local phase quantization for blur-insensitive dynamic texture classification," in *Image Analysis*, A. Heyden and F. Kahl, Eds., Berlin, Heidelberg, 2011, pp. 360–369, Springer.

[12] Y. Xiao, Z.-G. Cao, L. Wang, and T. Li, "Local phase quantization plus: A principled method for embedding local phase quantization into fisher vector for blurred image recognition," *Information Sciences*, vol. 420, pp. 77–95, December 2017.

[13] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *Image and Signal Processing*, Octeville, France, July 2008, vol. 5099, pp. 236–243.

[14] L. Palmieri and R. Koch, "Optimizing the lens selection process for multi-focus plenoptic cameras and numerical evaluation," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, July 2017, pp. 1763–1774.

[15] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Magazine*, vol. 13, no. 3, pp. 43–64, May 1996.

[16] C. Zhang, Z. Ji, and Q. Wang, "Decoding and calibration method on focused plenoptic camera," *Computational Visual Media*, vol. 2, pp. 57–69, March 2016.

[17] S. Suzuki and K. Abe, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, pp. 32–46, December 1985.

[18] C. Tang, C. Hou, and Z. Song, "Defocus map estimation from a single image via spectrum contrast," *Optics letters*, vol. 38, pp. 1706–8, May 2013.

[19] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, USA, 2 edition, 2003.

[20] "Raytrix GmbH," `https://raytrix.de/`, Accessed: 2018-09-24.

[21] O. Johannsen, C. Heinze, B. Goldluecke, and C. Perwass, *On the Calibration of Focused Plenoptic Cameras*, vol. 8200, pp. 302–317, Springer, Berlin, Heidelberg, January 2013.