Potential of Plenoptic Cameras in the Field of Automotive Safety

Sinan Hasirlioglu[1,2(✉)], Mahesh Karthik[1], Andreas Riener[1,2], and Igor Doric[1]

[1] Technische Hochschule Ingolstadt, CARISSMA, 85049 Ingolstadt, Germany
Sinan.Hasirlioglu@thi.de
[2] Johannes Kepler University Linz, 4040 Linz, Austria
http://www.carissma.eu

Abstract. Cameras in general are exteroceptive sensors and provide visual information about the local environment. For automotive safety applications, image data is processed to identify critical situations and avoid traffic accidents. Plenoptic cameras which are capable of recording both, angular and spatial information of light, have found increasing interest in optical 3D measurement techniques in recent years. Major advantages are multiple perspective views of captured scene and post focus capability. Therefore, the depth of the scene can be estimated similar to stereo camera.

In this paper, the potential of plenoptic cameras in the field of automotive safety is investigated in three aspects (1) depth estimation performance, (2) feature detection quality, and (3) weather robustness. For comparative results, adverse conditions such as low light or rain are replicated in a test hall. Initial results show that the additional microlens array of plenoptic cameras enable to achieve significant improvements over conventional cameras.

Keywords: Plenoptic camera · Lightfield camera · Computer vision
Object detection · Depth estimation · Automotive safety
Vehicle safety

1 Introduction

In the field of automotive safety, camera sensors play a major role in terms of object classification, feature detection and depth estimation. These aspects are even more important when it comes to adverse weather condition such as rain. Lightfields in general are referred to as the 4D spatio-angular light ray distribution incident on a 2D light sensor [1]. A plenoptic camera unlike a conventional digital camera, captures not only a 2D image information, but also the directions of the incoming light rays [2]. Thus, the lightfield implicitly captures 3D scene geometry and reflectance properties. It is also important to note that images are formed by extracting same pixels under each microlens which is placed in between main lens and the image sensor [3]. In adverse weather conditions, feature or object detection and distance estimation becomes a challenging task which is crucial in the field of automotive safety to avoid major

T. Kováčiková et al. (Eds.): INTSYS 2017, LNICST 222, pp. 164–173, 2018.
https://doi.org/10.1007/978-3-319-93710-6_18

accidents [4]. There are several sensors available for object classification and distance estimation such as radar, lidar and camera sensors. These sensors have its restrictions in adverse weather conditions [4]. Plenoptic cameras render many perspective views of the captured scene, from which a depth map of the scene can be estimated [7–9] and with additional light direction information it is possible to refocus on objects of interest after the image has been captured [5,6]. Investigation was also made whether the perspective view changes or post focus capability can be used for increasing the quality of environmental information.

Outline

This paper is organized as follows. Section 2 gives an overview on related work done in the field of depth estimation, depth selective focus methods, object detection, and influence of rain on plenoptic imagery. Section 3 describes aspects such as depth estimation, depth selective focus on objects, and feature detection along with rain simulator test setup. Section 4 discusses the experimental results in terms of adverse weather conditions and Sect. 5 summarizes the results and the scientific contribution of this paper.

2 Related Work

Georgiev and Lumsdaine [11] proposed a normalized cross correlation technique between microlens images for estimating the disparity map. Bishop and Favaro [12] introduced a method for recovering 3D depth of scene by matching anti-aliased lightfield views and then by performing deconvolution. In Wanner and Goldluecke [8], disparity map is estimated locally using dominant directions on epipolar plane images, which are computed with the structure tensor. In Yu et al. [9], 3D geometry of lines in a lightfield image was analyzed and the disparity maps are calculated through line matching between the sub-aperture images. In our research, the algorithm proposed by Jeon et al. [13] was used for estimating the disparity map using a lenslet lightfield camera. Proposed algorithm in [13] was used for solving the problem of narrow baseline between the sub-aperture images by using phase shift theorem.

Focus is an important aspect of photography which allows us to gather more light information and it also increases the signal-to-noise ratio (SNR). Lightfield cameras gather more light information than conventional cameras [7] and provide sufficient information for post focus capability. It blurs out scene elements which falls outside the plane of focus [5,6,14]. First frequency planar filter [14], 4D planar filter [15], shift and sum filter [16] which are capable of focusing on object at particular depth and volumetric filtering which is capable of selecting objects over range of depths is discussed in [10]. Regarding influence of environmental factors on plenoptic camera images, [10] has discussed depth selective focus on objects and noise reduction under murky water and particulate matter scenarios. To the best of our knowledge no one has investigated the influence of rain on plenoptic imagery and feature detection in adverse weather condition.

3 Materials and Methods

The proposed test methodology for investigating the potential of plenoptic cameras is mainly focused on depth estimation and feature detection quality under adverse weather conditions. These aspects are discussed below.

3.1 Depth Estimation

As a first step, the disparity map was estimated with an assumption regarding the baseline b in triangulation equation for finding depth Z of the object. In given method the main lens aperture diameter is considered as the baseline of the camera. It is clear from [7], that the perspective views are equivalent to viewing through different sub-apertures on the main lens. Therefore, top and bottom perspective views can be considered as viewing through top and bottom sub-apertures of the main lens as illustrated by Fig. 1. Main lens aperture diameter b is given by

$$b = \frac{f_L}{F} \tag{1}$$

where f_L is the focal length of the main lens and F the constant F-number of the camera. The F-number for Lytro Illum is a constant value of 2. For all test runs a focal length of 30 mm was set, which is estimated in pixels by using method proposed by Dansereau et al. [17]. Pixel disparity d of scene is determined for object at known depth values Z, with the help of basic triangulation equation gior ven below

$$d = \frac{f_L \cdot b}{Z} \tag{2}$$

where b is the baseline of the camera and Z the known depth value of object. In commercially available lightfield cameras such as Lytro, the pixel disparity range is very small in between adjacent sub-aperture images (1 pixel) as reported by

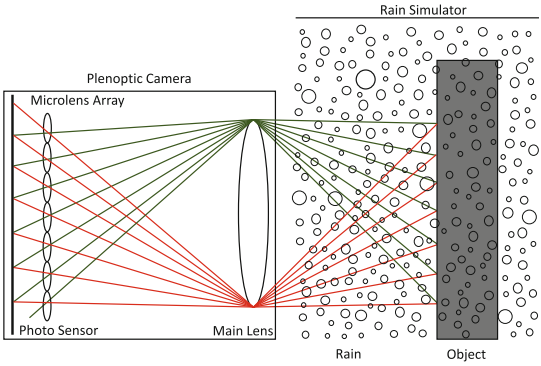

Fig. 1. Pictorial representation of plenoptic camera capturing 3D scene geometry using microlens array.

[9]. For investigating this, the disparity map was determined by extracting the left most and right most sub-aperture images. Triangulation equation (see Eq. 2) was used to find the approximate depth of the object. The depth map estimated by this method resulted in larger deviations in depth values up to 2600 mm for object placed at 500 mm, which is due to the narrow baseline between the sub-aperture images (see Fig. 3). Object placed at 500 mm and 1000 mm from the camera showed only 4 pixel and 2 pixel disparity respectively.

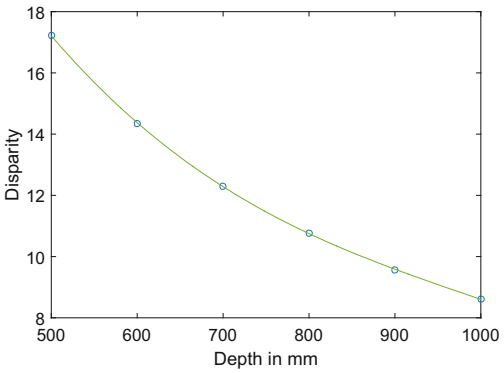

Fig. 2. Required disparity values vs. known depth values in mm.

Therefore, the algorithm proposed in [13] is used for solving the problem of narrow baseline between the sub-aperture images by phase shifting of the sub-aperture images using phase shift theorem. For the estimation of sub-pixel shifts of sub-aperture image, phase shift theorem is used in the fourier domain and thus by enabling the estimation of the stereo correspondences at sub-pixel accuracy. The subpixel shifted sub-aperture image I_r can be obtained as

$$I_r = I(x + \delta x) = F^{-1}[F\{I(x)\}exp^{2\pi i \delta x}]. \tag{3}$$

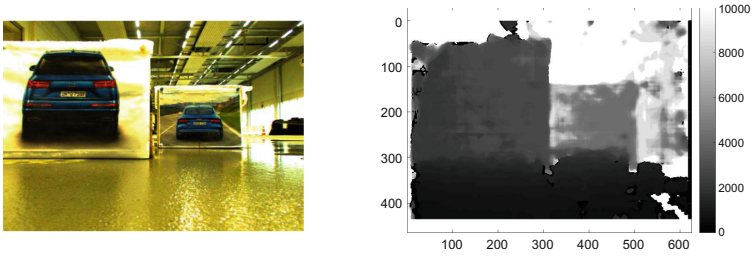

Fig. 3. Estimated depth map by using two sub-aperture images in normal light scenario. For object placed at 500 mm the estimated depth output showed larger deviation up to 2600 mm.

Similarity measurement is performed in between the sub-aperture images shifted at different sub-pixel locations and the center view sub-aperture image as reference. Cost volume based stereo [18] was used in order to construct cost volume and it gives a clear picture of how well a pixel x in image I matches with same pixel in second image I_r. Per pixel cost volume for different disparity labels are calculated. Two cost functions namely sum of absolute differences and sum of gradient differences were used in order to match sub-aperture images and to construct the cost volume. The cost functions are defined as below.

$$C_{SGD} = \sum_{s \epsilon V} \beta min[|\nabla_x I(s_c, x) - \nabla_x I(s, x + \delta x)|, \tau_2] + (1 - \beta) min[|\nabla_y I(s_c, x) - \nabla_y I(s, x + \delta x)|, \tau_2] \quad (4)$$

$$C_{SAD} = \sum_{s \epsilon V} min[|I(s_c, x) - I(s, x + \delta x)|, \tau_1] \quad (5)$$

$$C_{Total} = \alpha C_{SAD} + (1 - \alpha) C_{SGD} \quad (6)$$

Parameters α and β can vary between 0 and 1. Parameter α is used for balancing color and gradient term. As proposed in [13], β is independent of the scene and is purely determined by using the relative position between phase shifted view s and center view s_c. τ_1 and τ_2 are the truncation values for suppressing the outliers. V contains all perspective views, except the center view s_c. The shift vector δx varies linearly with respect to increase in angular deviations from center view point as following,

$$\delta x = lk(s - s_c) \quad (7)$$

where k represents the pixel shift unit and l the cost label. Median filtering was performed in order to preserve the edges. And based on winner takes all strategy, label l will be assigned to pixel x. As a final step, graph cut technique [19] was utilized to correct the disparities using neighbour estimation. Followed by estimation of object depths using triangulation equation (see Eq. 2).

3.2 Depth Selective Focus

Depth selective focus using the additional light direction information allows to post focus on objects at different depths [5]. For testing purpose, shift and sum depth selective focus method which is similar to planar focus method [15] was used in order to detect the object features under adverse weather conditions such as rain. In plenoptic camera each lenslet in microlens array is considered as sub-image, which has a corresponding shift vector up to which it must be shifted in order to focus on objects at certain depth. The basic principle is that the filter shifts all the sub-images to a common depth and then adds them together to yield a single 2D output. As mentioned in [10,15], the parameter slope gives the amount by which the lightfield slices would be shifted in order to focus on object at particular depth. In general, slope value of 0 lies near the center of

captured depth of field, such that slope will be having negative values for objects in foreground and positive values for object in background. Parameter slope is given by

$$slope = \frac{D}{P_Z} \tag{8}$$

where D represents separation in between two planes described by [10,15] and P_Z the depth of point P in scene. The relationship between slope and depth depends on lightfield parameterization.

3.3 Feature Detection Based on Normalized Cross-Correlation

Feature detection using normalized cross correlation technique [20] was performed on images captured under challenging conditions including rain scenario. Normalized cross correlation is a template based matching technique, whose output results in correlation scores which can be used as criteria for feature detection. Therefore, the cross correlation technique was used to investigate the influence of rain as well as low lighting condition on plenoptic images. Verification was also made whether depth selective focus and field of view changes offered by plenoptic cameras can be utilized in order to achieve an improvement in correlation scores. For testing purpose the head light part of the car which was captured under normal light condition was considered as template image. As a next step, normalized cross correlation was performed in between the template image and test images taken under various challenging scenarios. The normalized cross correlation term [20] is given by

$$c(u,v) = \frac{\sum_{x,y}[I(x,y) - \bar{I}_{u,v}][t(x-u,y-v) - \bar{t}]}{\left\{\sum_{x,y}[I(x,y) - \bar{I}_{u,v}]^2 \sum_{x,y}[t(x-u,y-v) - \bar{t}]^2\right\}^{0.5}} \tag{9}$$

where I is the test image, t is the template image, $\bar{t}$ is the mean of the feature, and $\bar{I}_{u,v}$ is the mean of $I(x;y)$, which is the region under the feature. The denominator term consisting of magnitude of I and magnitude of t is mainly used for normalizing the cross correlation process. Equation 9 results in correlation scores $c(u,v)$, which ranges from the value -1 to $+1$, with $+1$ being the best match. The condition for detecting the head light part of the car is minimum of five valid matches which are very close to each other having correlation scores greater than the threshold value of 0.85. Test results are presented in Sect. 4.

3.4 Test Setup

For simulating the rain scenario, a setup similar to the one suggested by [4] was used in indoor conditions and a rain intensity of about 120 mmh was used for testing purpose. Rain simulators are reproducible, controllable and save test effort. The Lytro lightfield camera was placed in the rain area and car image objects were placed at 500 mm and 1000 mm. Measurements were also done in normal and low light scenarios by switching on and switching off the lighting setup inside the indoor test facility.

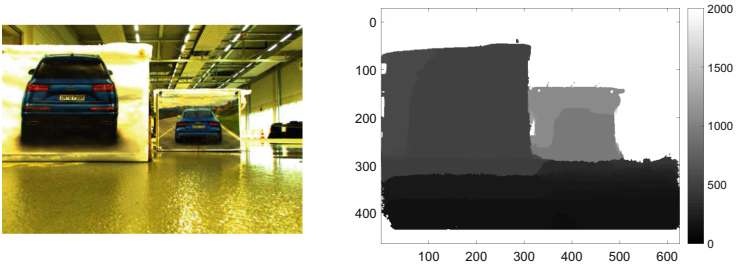

Fig. 4. Estimated depth map using phase shift theorem in normal light scenario showing estimated depth values of 508 mm and 1020 mm.

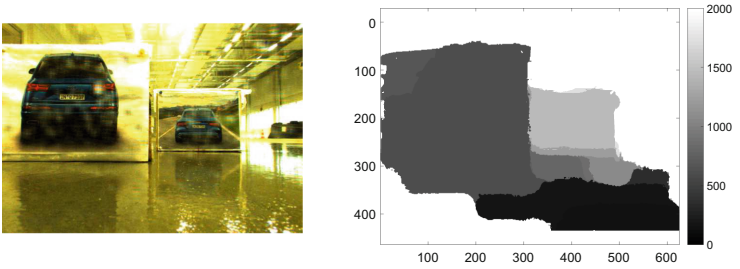

Fig. 5. Estimated depth map after using phase shift in rain scenario showing estimated depth value of 617 mm and 1234 mm for object placed at 500 mm and 1000 mm.

4 Experimental Results

4.1 Depth Estimation Performance

Depth estimation performance of plenoptic camera was investigated in normal light (see Fig. 4) and rain scenarios (see Fig. 5). For a pixel shift unit k within the range of 0.015 to 0.017, lightfield camera gave good depth estimation results for the object placed at 500 mm and 1000 mm. It can be seen that there are variations compared to normal light scenario. This is mainly due to the presence of atmospheric disturbances like rain in the image. It will also be difficult to compute the pixel correspondence in between the sub-aperture images, which will result in deviations in the disparity and depth maps. Value of 0 will be assigned to pixels for which correspondence matching cannot be found in the disparity map. There are higher possibilities of false correspondence matching which will also result in deviations in disparity and depth maps.

4.2 Feature Detection Performance Based on Normalized Cross Correlation

Feature detection based on normalized cross correlation was performed in between head light images captured under normal lighting condition (template

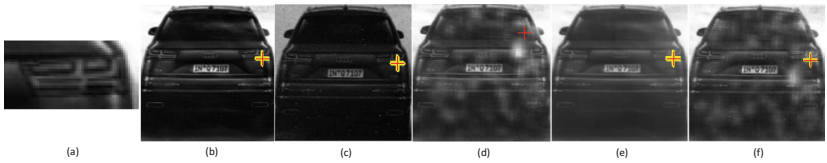

Fig. 6. (a) Template image, (b) Feature detection using normalized cross correlation in normal light, (c) Low light scenario showing good detection result, (d) No detection in rain scenario (center sub-aperture image), (e) Feature detection in rain scenario after depth selective focus, (f) Feature detection using field of view change. (Color figure online)

image) and images captured under adverse conditions (see Fig. 6). The best matching correlation score is indicated in red color mark. All other valid matches are indicated in yellow color marks.

Figure 6(b) shows the best matching correlation score +1 and 11 valid matches which are very close to each other. All resulting correlation scores were greater than the threshold value of 0.85.

In low lighting condition (see Fig. 6(c)), the correlation process resulted in best matching score of 0.9755 and a total of 16 valid matches. Feature detection in low light conditions show good results, which clearly illustrates the higher light gathering ability of plenoptic cameras.

In adverse weather conditions such as rain scenario (see Fig. 6(d)), the correlation process resulted in best matching score of 0.7809 which is already below the valid threshold value. The resultant score and detected region in main image is not the match of the template head light image. Invalid detection occurs due to rain droplets in scene, which act as disturbances by blurring some regions of the head light part of the car. Due to invalid detection and zero near valid matches, head light feature in car could not be detected in rain scenario.

For reducing the effects and disturbances, depth selective focusing provided by plenoptic cameras was utilized for focusing on the car object. It can be clearly seen in Fig. 6(e) that effects of rain droplets are completely removed. It resulted in best matching score of 0.9780. Total of 17 valid matches which are very close to each other were also obtained as final result. These results clearly illustrates that post focus capability offered by plenoptic cameras is helpful in removing the disturbances caused by rain.

Verification was also made whether by utilizing the change in perspective view can be used for achieving an improvement in feature detection performance. It can be clearly seen from center sub-aperture image shown in Fig. 6(d) and top most sub-aperture image represented by Fig. 6(f) that by changing the field of view, the rain droplet which blocks and blurs the head light part of car is shifted in top most perspective view such that the head light part of the car is made visible. The best matching score is increased to 0.9298 and the number of near valid matches from 0 to 6. It is clear, that changing the perspective view can make a significant contribution towards improving the feature detection (see Table 1).

Table 1. Normalized cross-correlation scores for different test scenarios.

Test scenarios	Cross-correlation scores
Normal light	1
Low light	0.9755
Rain	0.7809
Depth selective focus in rain	0.9780
Perspective view change in rain	0.9298

5 Conclusion and Future Work

This paper presents the potential of plenoptic cameras in the field of automotive safety, especially in adverse weather conditions. Initial results show, that rain drops cause deviations in depth values, but can be corrected by the plenoptic camera. Object feature detection by utilizing the field of view changes and post focus capability provided by the plenoptic camera also proves the robustness of plenoptic cameras in adverse weather conditions. Higher light gathering ability will be helpful in low light scenarios to gather sufficient information about the scene. These aspects are possible mainly due to the additional hardware inside the camera which is nothing but the microlens array.

Although the plenoptic cameras offer several advantages, it has its own disadvantages. The spatial resolution of images rendered by plenoptic cameras are always less when compared to conventional cameras which is due to the fact that it produces more perspective views of the captured scene. This results in higher angular resolution and lower spatial resolution. Since many pairs of images are compared with each other for finding disparity, the computational time regarding depth estimation and pre-processing of raw images is quite high.

Future work includes the use of focused plenoptic cameras, which have the ability to render final images with higher resolution. Furthermore, tests can be conducted in dynamic scenarios with both moving objects and camera.

Acknowledgment. This work is supported under the FH-Impuls program of the German Federal Ministry of Education and Research (BMBF) under Grant No. 13FH7I01IA.

References

1. Levoy, M., Hanrahan, P.: Light field rendering. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, pp. 31–42. ACM (1996)
2. Wetzstein, G., Ihrke, I., Lanman, D., Heidrich, W.: State of the art in computational plenoptic imaging. In: EUROGRAPHICS. Citeseer (2010)
3. Adelson, E.H., Wang, J.Y.: Single lens stereo with a plenoptic camera. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 99–106 (1992)

4. Hasirlioglu, S., Kamann, A., Doric, I., Brandmeier, T.: Test methodology for rain influence on automotive surround sensors. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 2242–2247. IEEE (2016)
5. Ng, R.: Fourier slice photography. ACM Trans. Graph. (TOG) **24**(3), 735–744 (2005)
6. Schedl, D.C., Birklbauer, C., Gschnaller, J., Bimber, O.: Generalized depth-of-field light-field rendering. In: Chmielewski, L.J., Datta, A., Kozera, R., Wojciechowski, K. (eds.) ICCVG 2016. LNCS, vol. 9972, pp. 95–105. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46418-3_9
7. Ng, R., Levoy, M., Bredif, M., Duval, G., Horowitz, M., Hanrahan, P.: Light field photography with a hand-held plenoptic camera. Comput. Sci. Tech. Rep. CSTR **2**(11), 1–11 (2005)
8. Wanner, S., Goldluecke, B.: Variational light field analysis for disparity estimation and super-resolution. IEEE Trans. Pattern Anal. Mach. Intell. **36**(3), 606–619 (2014)
9. Yu, Z., Guo, X., Lin, H., Lumsdaine, A., Yu, J.: Line assisted light field triangulation and stereo matching. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2792–2799 (2013)
10. Dansereau, D.G., Pizarro, O., Williams, S.B.: Linear volumetric focus for light field cameras. ACM Trans. Graph. **34**(2), 15–1 (2015)
11. Georgiev, T., Lumsdaine, A.: Reducing plenoptic camera artifacts. Comput. Graph. Forum **29**(6), 1955–1968 (2010). Wiley Online Library
12. Bishop, T.E., Favaro, P.: The light field camera: extended depth of field, aliasing, and superresolution. IEEE Trans. Pattern Anal. Mach. Intell. **34**(5), 972–986 (2012)
13. Jeon, H.-G., Park, J., Choe, G., Park, J., Bok, Y., Tai, Y.-W., So Kweon, I.: Accurate depth map estimation from a lenslet light field camera. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1547–1555 (2015)
14. Isaksen, A., McMillan, L., Gortler, S.J.: Dynamically reparameterized light fields. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, pp. 297–306. ACM Press/Addison-Wesley Publishing Co. (2000)
15. Dansereau, D., Bruton, L.: A 4D frequency-planar IIR filter and its application to light field processing. In: Proceedings of the 2003 International Symposium on Circuits and Systems, ISCAS 2003, vol. 4, pp. IV. IEEE (2003)
16. Georgiev, T., Intwala, C.: Light field camera design for integral view photography. Adobe Technical report (2006)
17. Dansereau, D.G., Pizarro, O., Williams, S.B.: Decoding, calibration and rectification for lenselet-based plenoptic cameras. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1027–1034 (2013)
18. Hosni, A., Rhemann, C., Bleyer, M., Rother, C., Gelautz, M.: Fast cost-volume filtering for visual correspondence and beyond. IEEE Trans. Pattern Anal. Mach. Intell. **35**(2), 504–511 (2013)
19. Kolmogorov, V., Zabih, R.: Multi-camera scene reconstruction via graph cuts. Comput. Vis. ECCV **2002**, 8–40 (2002)
20. Lewis, J.P.: Fast normalized cross-correlation. In: Vision Interface, vol. 10, no. 1, pp. 120–123 (1995)