https://doi.org/10.2352/ISSN.2470-1173.2019.11.IPAS-270

Depth from Stacked Light Field Images using Generative Adversarial Network

Ji-Hun Mun and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
123 Cheomdangwagi-ro, Buk-gu, Gwangju 61005, South Korea

Abstract

The estimated depth map provides valuable information in many computer vision applications such as autonomous driving, semantic segmentation and 3D object reconstruction. Since the light field camera capture both the spatial and angular light ray, we can estimate a depth map throughout that properties of light field image. However, estimating a depth map from the light field image has a limitation in term of short baseline and low resolution issues. Even though many approach have been developed, but they still have a clear flaw in computation cost and depth value accuracy. In this paper, we propose a network-based and epipolar plane image (EPI) light field depth estimation technique. Since the light field image consists of many sub-aperture images in a 2D spatial plane, we can stack the sub-aperture images in different directions to handle occlusion problem. However, usually used many light field sub-aperture images are not enough to construct huge datasets. To increase the number of sub-aperture images for stacking, we train the network with augmented light field datasets. In order to illustrate the effectiveness of our approach, we perform the extensive experimental evaluation through the synthetic and real light field scene. The experimental result outperforms the other depth estimation techniques.

1. Introduction

Light fields have a complex structure due to that collects the light ray information in any direction to hold the various direction of light ray. Due to that properties of light field, the light filed image includes spatial and angular directional spaces. The depth map estimation from light filed image is not simple problem, because of the low resolution and the distance between neighbor view point images which called sub-aperture images. Conventional depth estimation techniques have relied on various methods such as 2D stereo matching [1, 2, 3], geometric priors [4, 5] and depth from different light conditions [6, 7]. However, those methods are targeted on estimating a depth from a single-view camera or multi-view camera image, so depth blurring and hole regions often appear in estimated depth map. In addition, those methods need a precisely calibrated image pairs such as rectification, color correction and light condition pre-processing [8, 9].

In order to overcome that kind of limitations in conventional depth estimation, a lot of attention is given to generate accurate and robust algorithm for light field depth estimation. Since the light field image which captured from the lenslet array is convertible to multi-view image, it can be changed with slightly different view point image using geometric converting method. From the converted light field image, the depth map is estimated thanks to the similar camera structure with multi-view camera system [10].

(a) (b) (c)

Figure 1. Light field depth estimation comparison results. (a) Input light field image. (b) Tao et. Al method result [11]. (c) Our learning based depth estimation method result

Except for the light field image has narrow baseline and row resolution properties, depth estimation approach is very similar with the stereoscopic and monocular image based method. Widely used stereoscopic image depth estimation method mimic the binocular human visual mechanism. In addition, many of the dense baseline depth estimation methods have been developed on stereo vision and it also properly working in real time. However, those method is not optimized to apply in light field image depth estimation, due to the geometric condition of light field image is different from the general stereo and multi-view image.

Various algorithms for robust light field depth estimation method are developed depending on different kind of light field images, such as epipolar plane images (EPI), light field to multi-view, and defocus and correspondence. Existing light field EPI depth estimation method [12] suffer from the occlusion issue when measuring the slope of EPI. Thanks to the continuity sampling of light field sub-aperture image, we can estimate a depth map by exploring the slope of EPI. However, the EPI which stacked by angular directional sub-aperture image contains the cross edge region, that causes difficulty to estimate a correct depth label.

Among various deep learning methods, the *unsupervised* learning technique which do not require a pair of input and ground truth depth data for network training. Since obtaining ground truth depth maps from real and synthetic is very expensive and extremely hard work. Even though the data sets are abundant to training the network for depth estimation, we can easily contact the undesirable artifacts, such as blurring, inaccurate depth, and hole region in estimated depth map.

In this paper, we propose the accurate depth estimation throughout the various light field EPI with different angular sub-aperture images. We explore the possibility of network training for depth estimation on synthetic and real light field image dataset. To secure the abundant dataset for training we propose various angular sub-aperture based EPI data generation and data augmentation techniques. In addition, previously existing light filed depth estimation problem is handled via domain style transfer network which is very well known generative adversarial network (GAN) with newly proposed loss function for training. By combining those two contribution, we estimate an accurate depth map from light field input image.

2. Related Works

Over the several years, various depth estimation methods using light field images are developed [12, 13, 14]. Also, various domain translation deep learning techniques [15, 16] are intensively activated to utilize for light field depth estimation.

2.1 Depth from Light Field EPI

The light field EPI is composed of directional sub-aperture images. Many of the EPI-based depth estimation method focus on the horizontal and vertical direction for generating EPIs, since that is easy to extract from a set of light field sub-aperture images. But, we can consider more available orientation for multi-directional EPIs as indicated in Fig. 2.

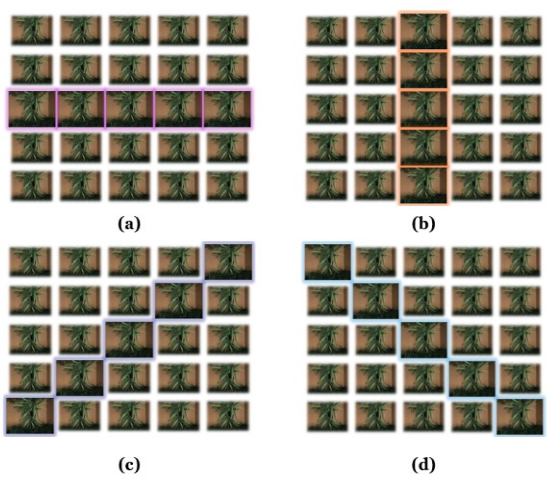


Figure 2. Light field multi-directional sub-aperture images for EPI. (a) horizontal. (b) vertical. (c) 135° -directional sub-aperture. (d) 45°-directional sub-aperture

The EPI contains linear combination of sub-aperture lines that is projected on the image and camera plane. In order to extract rich sub-aperture image for more EPIs, we need to consider the light filed ray structure in general case as shown in figure 3. From the image and camera plane, to express the EPI image plane y is fixed to y^* and camera plane t is fixed to t^*, and simultaneously changing the light field coordinate $L(x, s)$ as represented in (1).

$$I_{y^*,t^*}(x,s) = L(x, y^*, s, t^*) \tag{1}$$

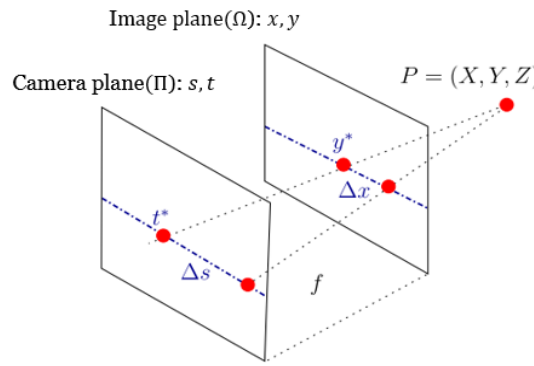


Figure 3. 4D light filed light field system with Image plane(Ω) and camera plane(π).

As the EPI consist of line with various directional sub-aperture images, the depth value can be derived by measuring the slope of EPI. Wanner *et al*. [14] apply the structure tensor to measure the slope in EPI. But, the tensor structure highly relied on an angular resolution so that have affected by occlusion when orient estimation. Zhang *et al*. [17] use a spinning parallelogram operator (SOP) as an approach to compute the slop of each EPI for local depth value estimation. They measure the EPI slope from maximized distribution distance in two part of parallelogram window. Tosic *et al*. [18] invented ray detection for depth estimation throughout the normalized section derivation in Gaussian kernel. This method considers occlusion region by analyzing overlapped rays to determine the object order in image. If some region has small variance value, then that is determined as foreground area. However, this method assumes that the initially estimated depth map is accurate and simultaneously variance condition have to be satisfied.

2.2 Learning based Depth Estimation

Thanks to the improvement of machine learning technique and hardware improvement, variety of computer vision applications are developed such as 3D reconstruction, multi view synthesis [19], and semantic segmentation [20].

Even though various patch based depth estimation for light field images, they still contain inaccuracy and blurring artifacts in estimated depth map. To conquer those problems, learning based depth estimation techniques are highly recommended. For light field depth estimation Heber et al. [21] propose a CNN for EPI slope measurement and global optimization technique. They train the CNN to find out accurate orientation of EPIs. Their previous work also estimating a depth map via encoding and decoding end-to-end deep neural network. Anna et al. [22] propose the encoder-decoder network to evade a specular while estimating a depth map. The network effort to distinguish the specular and diffuse for accurate depth estimation without considering given circumstances.

Unlike aforementioned learning techniques for light field depth estimation, we propose to train a network which fully adopt the unsupervised learning technique which called GAN. However, our network is trained using a pair of light filed EPI and depth EPI, due to the duality of EPI in training procedure. At the same time, we train the network with multi-directional EPI, since each epipolar image has intrinsic characteristics which can treat the occluded region while estimating a depth map from EPIs.

2.3 Domain Transfer

Image domain transfer methods are recently invented based on the GAN. After then, various astonishing approach and improvement have been suggested to change the domain from one to another. Some approach utilizes a MRF concept to control the image patch in order to transfer the image to aimed domain. Other method uses pixel updating approach that means they directly change the pixel value in the output image to translate an image domain [24]. In order to evade directly using pixel values to change the domain, some method pre-train the network through amount of training dataset [25].

Changing an image domain is similar concept with distribution optimization at target domain. Particularly, it is the same condition with minimize the distance between source image domain and target image domain to mimic the different domain style each other. Depending on those properties, we utilize a domain transfer concept to adapt our source image distribution (augmented light field images) to target domain distribution (ground truth depth dataset)

3. Proposed Approach

Our proposed method composed of two main components. Firstly, to generate an abundant EPI training dataset, multi directional sub-aperture images are exploited and augmented to extremely use the sub-aperture images for training. Then, the network is trained using GAN with newly proposed loss function which is devised to overcome the previously existing depth inaccuracy problem.

3.1 Data Preparation

There are many available light field datasets at the internet, also they provide various type of scene with ground truth data especially in case of synthetic datasets. However, those datasets are not enough to train the network in terms of variety of situation in capturing light field scene not only for the synthetic or rendered light field scene. In this paper, we exploit the 8 synthetic light field scene and real light field datasets that provides pair of texture and ground truth depth data. In this paper, we augment the input dataset through rotation, flipping, scaling and color range variation. The objective of this data augmentation is increasing the network training efficiency and yielding a high quality depth map.

The maximum angular resolution of light filed image which used in this paper is restricted to 7×7, due to the accuracy of estimated depth map does not show a critical differences when it compared with 9×9 and 11×11 as indicated in Table 1. Even though the mean square error (MSE) and bad pixel ratio (BPR) which determine the error when the pixel difference is larger than 0.7, those value decreasing ratio is not show the enough performance with respect to the training cost.

Table1. Depth map accuracy with different light field angular resolutions

	Light field angular resolution				
	3 x 3	**5 x 5**	**7 x 7**	**9 x 9**	**11 x 11**
MSE	2.762	2.172	1.834	1.779	1.714
BPR	7.47	6.87	3.79	3.27	3.04

While training the network through the input light field image, the resolution of image is gradually decreased. The low resolution image affects to the feature extraction on backpropagation procedure. In order to prevent the low resolution problem, we up-scaled the input light field image for training by multiple of 2.

The light field EPI includes characteristics of multi-directional sub-aperture linearity. By measuring the EPI orientation, the depth values are easily estimated, due to we perform the rotation augmentation on the light field sub-aperture images. The translation augmentation technique already has been widely used in depth estimation, optical flow and scene flow. The conventionally used rotation augmentation method only focus on the captured 2-dimensional domain. But, due to the light ray recording property of light field, it composed of 4-dimensional (x, y, u, v) domain. While rotating the image, the properties of light field EPI have to be preserved. The horizontal directional based EPI properties is different with the vertical directional based EPI properties. As indicated in Fig. 3, if the horizontal directional sub-aperture images are rotated in 90° then horizontal EPI property is changed to the vertical EPI property.

However, when we apply not an addition of 90 degree for image rotation such as 30 degree and 45 degree, the EPI properties is not well preserved. This limitation is handled by applying a learning technique, and that will be explained in section 3.2. Considering various angular directions for light field EPI shows robust depth map result especially on the occluded region. Since Hao *et al.* [23] assume that the occluded region in camera plane is very close to the slope of occluder, they adopt the detected occluded region during optimize the cost volume. With the similar perspective of that, we use the different orientation based generated EPIs for network training.

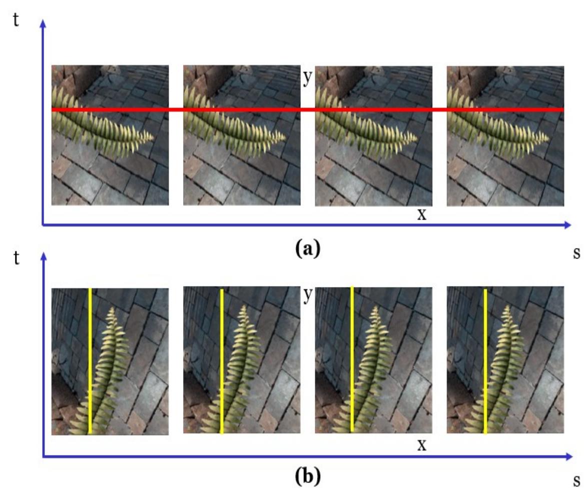


Figure 3. EPI property transformation via sub-aperture rotation augmentation. (a) horizontal EPI property. (b) vertical EPI property.

Lastly, generally used flipping method is applied for data augmentation. We flip the light field image through up/down and left/right. If the image was symmetrically flipped through left and right, then slope of the EPI also flipped. Due to the flipping, the estimated disparity value also has reversed, so we have to consider the flipped image training for the network.

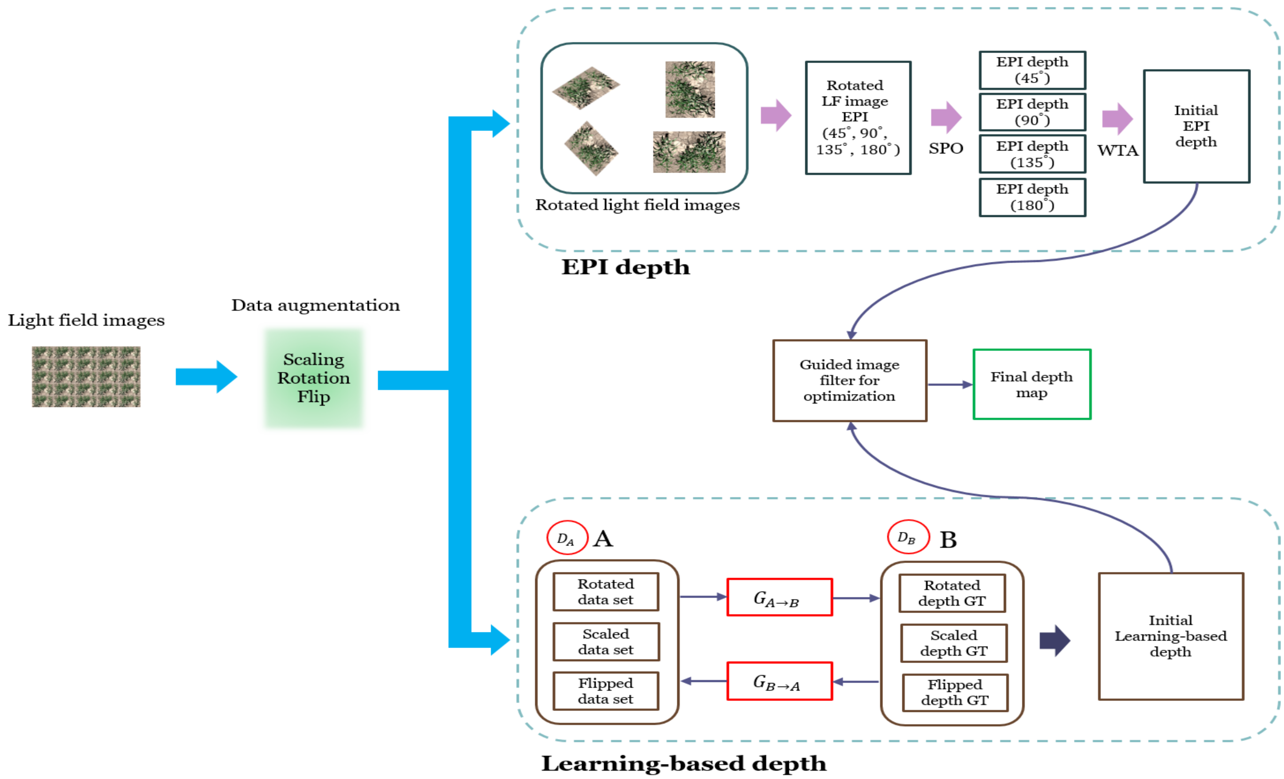


*Figure 4. Proposed light field depth estimation pipeline. The **EPI depth** estimation block compute the slope of EPI which generated through the rotated light field image via SPO. **Learning-based depth** generates a depth map by using GAN with domain transfer concept.*

3.2 Proposed Architecture

Our proposed light field depth estimation pipeline is represented in Fig. 4. The proposed method composed with two main part. Firstly, we generate EPIs through the augmented light field image especially using rotation augmentation for depth map creation. Secondly, neural network is adopted to estimates another depth map. Those estimated depth map is optimized via guided image filtering for final depth map.

EPI depth estimation: The objective of light field image augmentation is increasing a dataset whilst yield wide orientation-based light field EPI generation. In EPI depth estimation operation, we adapt rotated light filed image through 4-angle $\theta = \{45^{\circ}, 90^{\circ}, 135^{\circ}, 180^{\circ}\}$. The EPIs which structured with different rotation angle sub-aperture image still preserve the property between the angular direction and viewpoint. In addition, from the various angular EPIs, we can estimate more robust depth map than horizontal or vertical directional EPIs.

Multi-orientation EPIs also can be applicable to conventional light field EPI-based depth estimation methods [14, 17]. In this paper, to compute a direction of EPI, we adopt the spinning parallelogram operator (SPO) which proposed by Zhang *et al* [13]. Since SPO measure the slope of EPI lines by maximizing distribution distance in parallelogram window, it shows accurate EPI slope measurement performance than Zhang *et al*. [26] method which measure the slope within angular candidates. Instead of commonly used distance measure L_2 norm, SPO use the Earth Mover's distance (EMD) or χ^2 to measure the difference between the distance of pixel colors in window. The difference of χ^2 is utilized in color histogram as defined in (2)

$$\chi^2(g_\theta, h_\theta) = \sum_i \frac{(g_\theta(i) - h_\theta(i))^2}{g_\theta(i) + h_\theta(i)} \tag{2}$$

where $g_\theta(i)$ and $h_\theta(i)$ are the histogram of separated parts. If the χ^2 value is large, that means the two parallelogram are different. As a result, we can assume that there exists straight edge in hypothetical matching line. Throughout the (2), we construct a cost volume along the axis (Ω, Π, θ) for disparity space image construction. From the measured EPI slope via SOP on EPI, the EPI-based depth on each rotation angle is estimated using (3).

$$\theta_\lambda(\Omega, \Pi) = \arg\max_\theta d_\lambda(\Omega, \Pi, \theta) \tag{3}$$

where λ is a set of angular EPI, θ represents corresponding maximum response of (3), and d is indicates the histogram distance measured by the SPO. This value indicates the depth of the point in the center view. So, we can get the depth value for the each angular EPI image $d_\lambda(\Omega, \Pi, \theta)$. However, we fix the λ when finding a maximally matched cost value in (3), due to the proposed algorithm find an optimal EPI depth value by applying winner takes all (WTA) method.

Learning-based depth estimation: Assuming the initially estimated depth map from light field EPI is accurate, but it still includes blurring artifact near the object edge region. The goal of our proposed network is providing a supplementary depth data for optimization in final step.

As represented in Fig. 4, augmented light field images are used for input data in domain A. Simultaneously, the ground truth depth data is charged in domain B. The network learn a mapping function $G : A \longrightarrow B$ from the source domain A(augmented light field image) to target domain B(depth image). So, the network effort to make a same result with $G(A)$ and B. From augmented light field source domain data, we train the network to estimate a depth map in target domain.

Changing an image domain from one to other domain is already have been invented, such as Pix2Pix, Cycle-GAN[15], and Disco-GAN. Among them, we utilize the style transfer approach in generative adversarial network which demonstrated in [15]. They suggest the cycle-consistency to minimize the discordance between source and target domain. That help to recover the specific domain image their own original domain with constraint condition.

In order to define an adversarial network with cycle-consistency, we need generator and discriminator which changing the domain and measure the faithful respectively. As represented in Fig. 4, $G_{A\rightarrow B}$ and $G_{B\rightarrow A}$ match the departure domain to destination domain. D_A discriminating between A and $G_{B\rightarrow A}(B)$, D_B discriminating between B and $G_{A\rightarrow B}(A)$ in each domain. By combining those parameters, we can define the adversarial loss function for network training as shown in (4).

$$\begin{aligned}\mathcal{L}_{GAN}(G_{A\rightarrow B}, D_B) &= \min_{G_{A\rightarrow B}} \max_{D_B} \mathbb{E}_{b\sim p_{data(b)}}[logD_B(b)] \\ &\quad + \mathbb{E}_{a\sim p_{data(a)}}[\log(1 - D_B(G(a))] \\ \mathcal{L}_{GAN}(G_{B\rightarrow A}, D_A) &= \min_{G_{B\rightarrow A}} \max_{D_A} \mathbb{E}_{a\sim p_{data(a)}}[logD_A(a)] \\ &\quad + \mathbb{E}_{b\sim p_{data(b)}}[\log(1 - D_A(G(b))]\end{aligned} \tag{4}$$

where $p_{data(a)}$ and $p_{data(b)}$ indicate data distribution in A and B domain with sample a and sample b respectively. The style transfer loss function can be defined as consequence of min-max problem.

In theoretically, adversarial network can train the mapping function G to precisely coincidence with each other domain. However, among the amount of training dataset the network does not guarantee a source image correctly mapping on the target image vice versa. For example, we randomly select augmented light filed image from domain A to change the domain to B for depth inference. But, if we do not provide any constraint conditions, the back-propagated image will not correctly match with that own domain.

As a result of that, the adversarial loss function defined in (4) cannot be used alone. In order to constraint the adversarial loss, a cycle-consistency loss is added which help the source image A transform the domain to target domain via $G_{A\rightarrow B}(a)$, then it correctly back into the source domain via $G_{B\rightarrow A}(G_{A\rightarrow B}(a))$ that is same with the A. The cycle consistency loss is defined in (5).

$$\begin{aligned}\mathcal{L}_{Cycle} = \mathbb{E}_{a\sim p_{data(a)}}[\|G_{B\rightarrow A}(G_{A\rightarrow B}(a) - a\|_1] + \\ \mathbb{E}_{b\sim p_{data(b)}}[\|G_{A\rightarrow B}(G_{B\rightarrow A}(b) - b\|_1]\end{aligned} \tag{5}$$

For network training our final objective loss function is defined by combining (4) and (5) as follows:

$$\mathcal{L}_{CG} = \mathcal{L}_{GAN}(G_{A\rightarrow B}, D_B) + \mathcal{L}_{GAN}(G_{B\rightarrow A}, D_A) + \mathcal{L}_{Cycle} \cdot \sigma \tag{6}$$

At last step, we optimize the estimated depth maps which are obtained from EPI-based and learning-based depth estimation via the guided image filter.

Implementation and training: All implementation of our network is coded by *PyTorch*. The learning based depth estimation network is built upon 3-convolutional layers with batch normalization. In addition, 6-residual block is added to preserve the image quality during training. In order to directly update parameters, the batch size is fixed to 1. For the back-propagation optimization approach, we use the ADAM with momentum $\beta_1 = 0.5$, $\beta_2 = 0.999$. The multiplied weight value in cycle consistency σ is set to 10. Overall training procedure, we set the learning rate to 0.0002.

For the network training, we use the light field scene which captured by Lytro Illum from Wang *et al*. [27] and synthetic dataset. The training time for our network with those datasets takes 4 and half days through a NVIDIA GTX TITAN X.

4. Experimental Results

We evaluate our proposed method using both qualitative and quantitative comparison with other light field depth estimation methods. We use synthetic and captured light field images for evaluation. The estimated depth map accuracies of the proposed method and other method. Fig. 5 and Fig. 6 show the experimental result with proposed method and other methods on Illum dataset synthetic datasets.

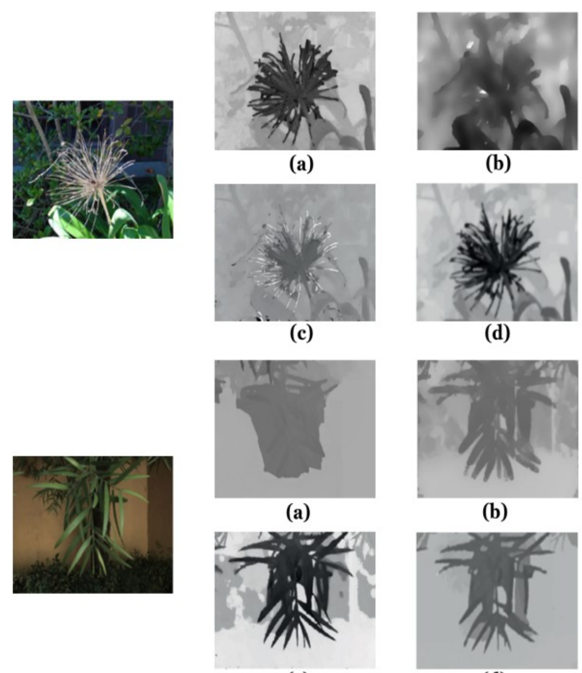


Figure 5. Experimental results of depth estimation with other methods in Illum datasets. (a) Lytro Illum, (b) Zhang et al. (c) Chen et al. (d) proposed result

Our method shows better handle noise by combining the EPI depth and learning based depth. Especially, EPI depth consider various angular direction during computing a slope via SPO method. Due to that reason, EPI provide more robust depth value near the object boundary region compare to the other light field depth estimation methods.

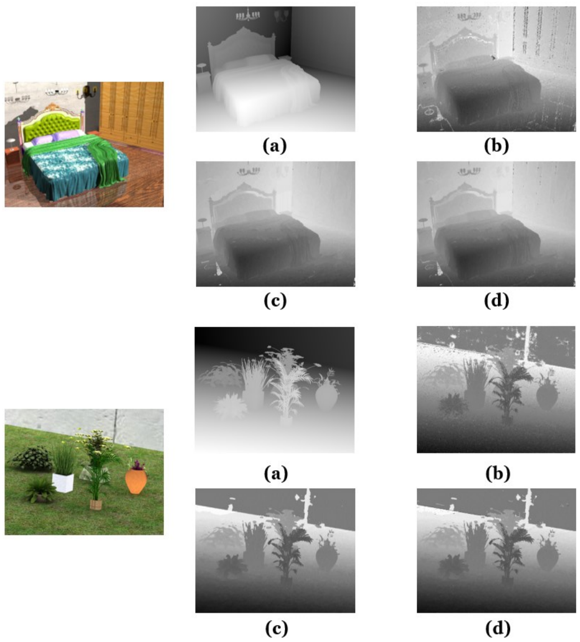


Figure 6. Experimental results of depth estimation with other methods in synthetic dataset. (a) ground truth depth, (b) Zhang et al, (c) Chen et al, (d) proposed result

In order to numerically evaluate our proposed method in synthetic dataset, a bad pixel error ratio is computed in Table 2. If the pixel value differences between estimated depth map is larger than 0.7 then that pixel is determined as bad pixel.

Table1. BPR comparison results in synthetic datasets (%)

Sequences		Zhang et al	Chen *et al*	Proposed
Livingroom	Non-occ	7.4	6.7	5.8
	All	12.3	11.9	11.2
Sculptures	Non-occ	7.7	7.3	6.9
	All	12.7	12.2	11.3
Bedroom	Non-occ	7.1	6.5	5.1
	All	12.1	11.8	10.9
Plant	Non-occ	7.3	6.8	5.5
	All	11.8	11.1	10.3

5. Conclusion

In this paper we propose depth estimation approach by combining the EPI-based depth result and learning-based depth result via guided image filter. In order to handle the object boundary depth in accuracy problem, we augmented the light field dataset. EPI depth estimation operation only use the rotation augmented data and learning operation use all augmented datasets. Depth estimation network through augmented light field images is trained using adversarial and cycle-consistency loss function. By combining two loss function we can generate depth map which robust to the noise in an image. From the experimental result, we demonstrate that our proposed method generate more accurate depth map than other methods.

Acknowledgements

This work was supported 'Civil-Military Technology Cooperation Program' grant funded by the Korea govetnment and 'Brain Korea 21 Plus Project' of the Ministry of Education & Human Resources Development, Republic of Korea (ROK) [F16SN26T2205].

References

[1] D. Scharstein and R. Szelisk, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. Journal Computer Vision*, vol. 47 no. 3, pp. 7-42, Aug. 2002.

[2] O. Hamilton and T. Breckon, "Generalized dynamic object removal for dense stereo vision based scene mapping using synthesised optical flow," *In Proc. Int. Conf. Image Processing*, pp. 3439-3443, Aug. 2016.

[3] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. on TPAMI*, vol. 30, no. 2, pp.328-341, Dec. 2007.

[4] V. Hedau, D. Hoiem, and D. Forsyth, "Thinking inside the box: Using appearance models and context based on room geometry," *in Proc. ECCV.*, pp. 224–237, Nov. 2010.

[5] A. Gupta, M. Hebert, T. Kanade, and D. M. Blei, "Estimating spatial layout of rooms using volumetric reasoning about objects and surfaces," *in Proc. NIPS*, pp. 1288–1296, Dec. 2010.

[6] M. W. Tao, P. P. Srinivasan, J. Malik, S. Rusinkiewicz, and R. Ramamoorthi, "Depth from shading, defocus, and correspondence using light-field angular coherence," *in Proc. CVPR*, pp. 1940–1948, Oct. 2015.

[7] A. Abrams, C. Hawley, and R. Pless. Heliometric, "stereo: Shape from sun position," *in Proc. ECCV*, pp. 357–370, Nov. 2012.

[8] J. Liu, X. Gong, and J. Liu, "Guided inpainting and filtering for kinect depth maps," *in Proc. Int. Conf. Pattern Recognition*, pp. 2055–2058. Feb. 2013.

[9] F. Qi, J. Han, P. Wang, G. Shi, and F. Li, "Structure guided fusion for depth map Inpainting," *Pattern Recognition Letters*, vol. 34, no. 1 pp.70–76, June 2012.

[10] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," *in Proc. CVPR*, pp. 1027-1034, June 2013.

[11] M.W. Tao, S. Hadap, J. Malik, and R. Ramamoorthi, "Depth from combining defocus and correspondence using light-field cameras," *in Proc. ICCV*, pp. 673–680, Dec. 2013.

[12] S. Wanner, and B. Goldluecke, "Globally consistent depth labeling of 4D light fields," *in Proc. of CVPR*, pp. 41–48 July 2012.

[13] S. Zhang, H. Sheng, C. Li, J. Zhang, and Z. Xiong, "Robust depth estimation for light field via spinning parallelogram operator," *Comput. Vision Image Understanding,* vol. 145, no. C, pp.148-159, April 2016.

[14] S. Wanner, and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," *IEEE Trans. on TPAMI*, vol. 36, no. 3, pp. 606-619, Aug. 2013.

[15] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image- to-image translation using cycle-consistent adversarial networks," *in Proc. ICCV*, pp. 2242 – 2251, Nov. 2017.

[16] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image translation with conditional adversarial networks," *in Proc. CVPR*, arXiv:1611.07004v2, Nov. 2016.

[17] 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 (CVIU)*, vol. 145, pp.148–159, April 2016.

[18] I. Tosic, and K. Berkner, "Light field scale-depth space transform for dense depth estimation," *in Proc. CVPRW*, pp. 441–448, Sept. 2014.

[19] N. K. Kalantari, T. C. Wang, and R. Ramamoorthi, "Learning-based view synthesis for light field cameras," *in Proc. SIGGRAPH Asia*, vol. 35, no. 6, DOI:10.1145/2980179.2980251, Sept. 2016.

[20] E. Shelhamer, J. Long, and T. Darrell, "Fully convolution networks for semnatic segmentation," *IEEE Trans. on TPAMI,* vol. 39, no. 4, pp. 640-651, April 2017.

[21] S. Heber and T. Pock, "Convolutional networks for shape from light field," *in Proc. CVPR*, DOI:10.1109, June 2016.

[22] A. Alperovich, O. Johannsen, M. Strecke, and B. Goldluecke, "Light field intrinsic with a deep encoder-decoder network," *in Proc. CVPR*. pp. 9145-9154, June 2018.

[23] H. Sheng, P. Zhao, S. Zhang, J. Zhang, and D. Yang, "Occlusion-aware depth estimation for light field using multi-orientation EPIs," *Pattern Recognition*. vol. 74, pp. 587-599, Feb. 2018.

[24] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," *in Proc. CVPR*, pp. 2414–2423, June 2016.

[25] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution" *in Proc. ECCV*, pp. 694–711, Mar. 2016.

[26] Y. Zhang, H. Lv, Y. Liu, H. Wang, X. Wang, Q. Huang, X. Xiang, and Q. Dai, "Light field depth estimation via epipolar plane image and locally linear embedding" *TCSVT*, vol. 27, no. 4, pp. 739–747, April. 2017.

[27] T. C. Wang, A. A. Efros, and R. Ramamoorthi, "Occlusion-aware depth estimation using light field cameras," *in Proc. ICCV*, pp. 3487-3495. Dec. 2015.

Author Biography

Ji-Hun Mun received his B.S in electronic engineering from the Chonbuk National University, Chonbuk, Korea (2013) and his M.S. in Science of Information and Communication from the Gwangju Institute of Science and Technology, Gwangju, Korea (2015). Now he takes Ph.D. course in Science of Information and Communication from the Gwangju Institute of Science and Technology, Gwangju, Korea. He works has focused on the video codec H.265/HEVC, 3D video technology and stereo matching.

Yo-Sung Ho received his B.S. and M.S in electronic engineering from the Seoul National University, Seoul, Korea (1981) and his Ph.D. in electrical and computer engineering from the University of California, Santa Barbara (1990). He worked in Philips Laboratories from 1990 to 1993. Since 1995, he has been with the Gwangju Institute of Science and Technology, Gwangju, Korea, where he is currently a professor. His research interests include image analysis, 3D television, and digital video broadcasting.

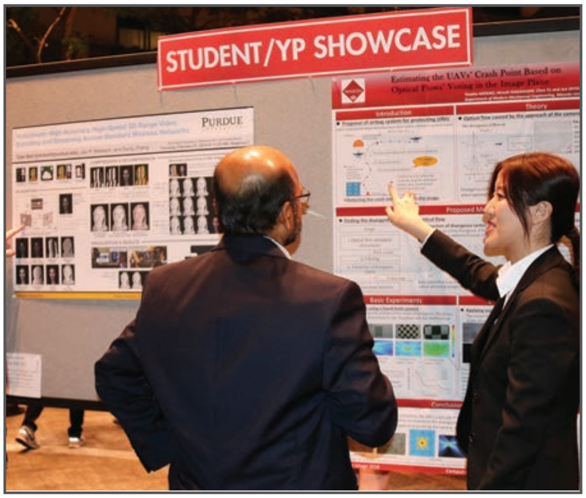
JOIN US AT THE NEXT EI!

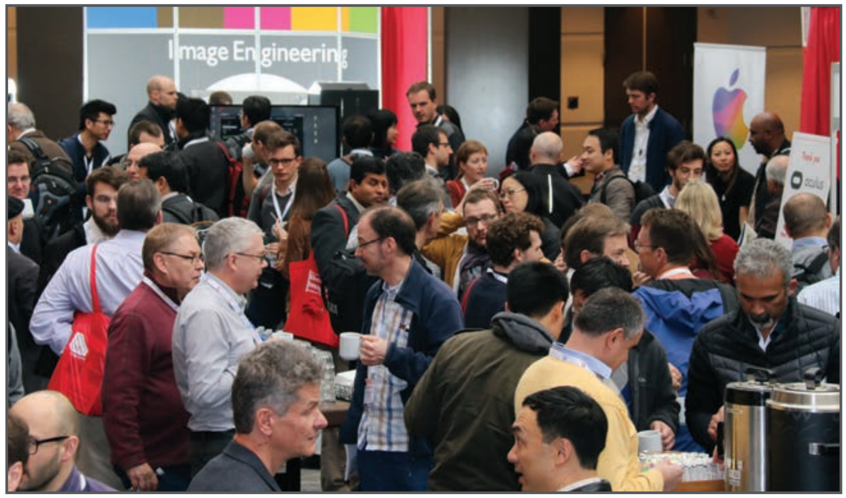
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org