INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG2017/ M 41648
October 2017, Macau, CN

Source Gwangju Institute of Science and Technology (GIST)
Status Report
Title [MPEG-I-Visual]: Depth Estimation using Light Field EPI for EE1
Author Ji-Hun Mun (GIST, jhm@gist.ac.kr)
Yo-Sung Ho (GIST, hoyo@gist.ac.kr)

Abstract

The light field epipolar plane image(EPI) is efficiently used for depth estimation. In this document we show the estimated depth map by considering occlusion area via light field EPI depth estimation method. We perform the depth estimation experiments using ULB data set.

1 Introduction

Last 119th Torino MPEG meeting, we show the depth estimation method using a light field EPI data in M41017[1]. That experiment is conducted through the CG dataset to evaluate the performance of proposed method using CG ground truth depth data. In order to more precisely convince our depth estimation technique, we conduct an experiment through a dense light field data set. The estimated depth map quality is evaluated by measuring the bad pixel rate (BPR) and root mean square (RMSE) value throughout the ground truth data.

2 EPI Depth Estimation Experiment for EE1

The experiment conditions and evaluation method for EE1 is explained in N17070[3]. Especially, the objective of EE1 is preparing an anchor depth data for many test sequences (e.g., ULB [2], Painter, Fraunhofer, Toy Train). Except for the ULB test sequences, the other sequences are captured with a different size of camera array (Painter: 4x4, Toy Train: 5x5, Fraunhofer: 3x5). Since the efficiency of our light field EPI based depth estimation technique is determined by distance of image viewpoints, we choose the dense camera array data set such as ULB data set. Fig. 1 shows the EPI of ULB data set along with the horizontal direction of band 1.

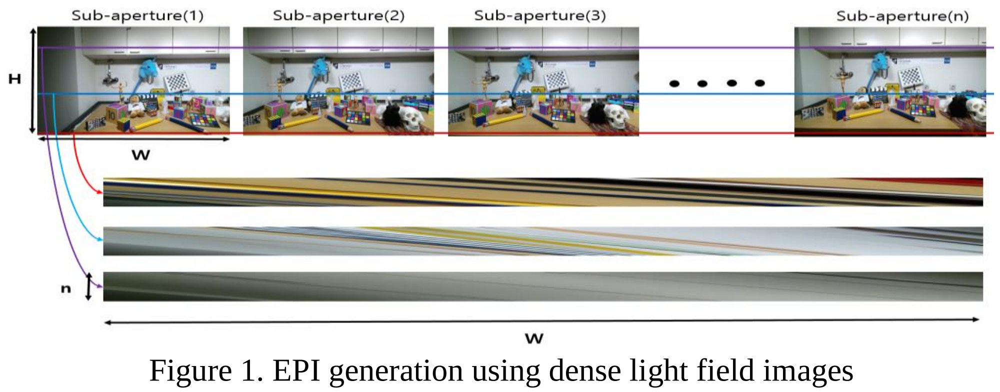


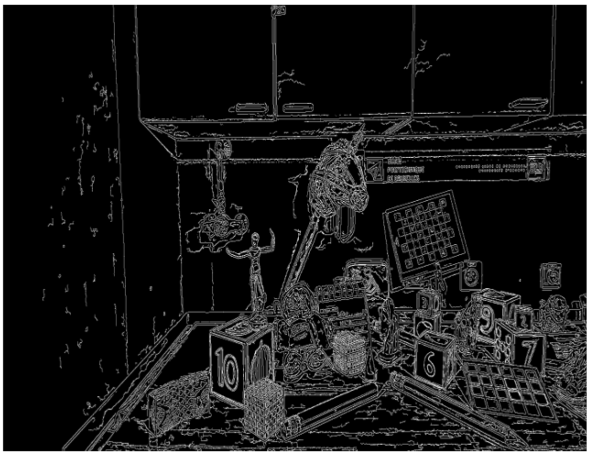
Figure 1. EPI generation using dense light field images

3 Occlusion Aware Depth Estimation

Conventionally proposed depth estimation method [3] from the light field EPI does not consider the occlusion area. The occlusion area affects to the estimated depth map quality. Not only the EPI based depth estimation method but also a local window based light field depth estimation method is affected by occlusion area. In order to handle this problem occlusion aware depth estimation is newly added to our previous work [3].

When we compute a cost value to find an optimal angular direction, occlusion causes an improper cost value computation. Since the occlusion area appear near the object boundary, firstly we extract edge map from light field images. Then, we compute an edge orientation from the edge map as indicated in figure 2 and apply the n x n size window patch for edge orientation comparison. If the difference of orientation pixel value is bigger than pre-defined threshold value, we can define that patch includes occlusion area using (1).

$$Occ = \begin{cases} 1, & if |L_{n_{RGB}}(x', y') - |L_{n+1_{RGB}}(x', y')| \geq threshold \\ 0, & else \end{cases} \tag{1}$$



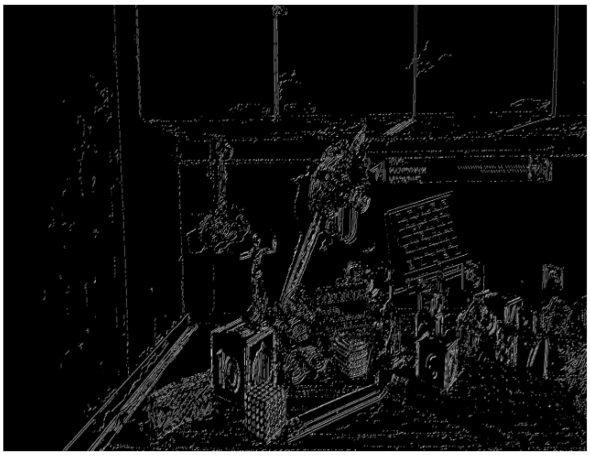


Figure 2. Edge map(left) and edge orientation(right)

In order to compute a cost for an optimal angular value selection from light field EPI, the occlusion area patch is replaced with similar patch from neighbor light field image. Even though the occlusion area is partially included in window patch, the whole patch region is replaced from other images to improve the cost value accuracy. The patch is replaced after computing variance and mean value which has smallest difference with current and neighbor window patch. The angular coordinate of edge orientation between two different light field image are (s_1, s_2) and (u_1, u_2). We compute the mean (2) and variance (3) value within two window patches.

$$\overline{LF_k}(x, y) = \frac{1}{N_k} \sum_{s,u} LF(x, y, s_k, u_k), \;\; k = 1, \; 2 \tag{2}$$

$$V_k(x, y) = \frac{1}{N_k - 1} \sum_{s,u} (LF(x, y, s_k, u_k) - \overline{LF_k}(x, y))^2, \qquad k = 1, 2 \tag{3}$$

where N_k is the number of pixels in window patch k, and (x, y) represents the spatial coordinates. The new window patch area is detected by selecting a minimum variance value within the specific window patch as indicated in (4).

$$L_{rep}(x,y) = arg \min_{k=1,2}\{V_k(x,y)\} \quad (4)$$

The $L_{rep}(x,y)$ indicates a center coordinate of new window patch for occlusion area replacement. Then the new cost computation term (5) is added to the originally invented cost computation equation (6) for optimal angular value selection.

$$C_{Occ} = V_{L_{rep}}(x,y) \quad (5)$$

$$C_{total}(E_l,P) = \frac{1}{N(\theta)} \sum_{P_n \in N(\theta)} \begin{matrix} (1-\alpha)|E(P_n) - E(P)| \\ +\alpha\left[|E_x(P_n) - E_x(P)| + \left|E_y(P_n) - E_y(P)\right|\right] \end{matrix} + C_{Occ} \quad (6)$$

The optimal angular value is selected from the angular candidates as explained in [3]. Since the occlusion area is considered when computing a cost value, the accuracy of optimal angular value selection is improved.

4 Experiment Results

In order to estimate the depth map via our algorithm we use ULB test sequences. Especially, the band 1 sequences are used, and that have been captured with 2mm physical camera distance. The band 1 data is composed with 51 lines of 425pictures. We only use first line image sets for experiment. The figure 3 and 4 indicate the estimated depth map result with different view point image along the band-1 and line-1.

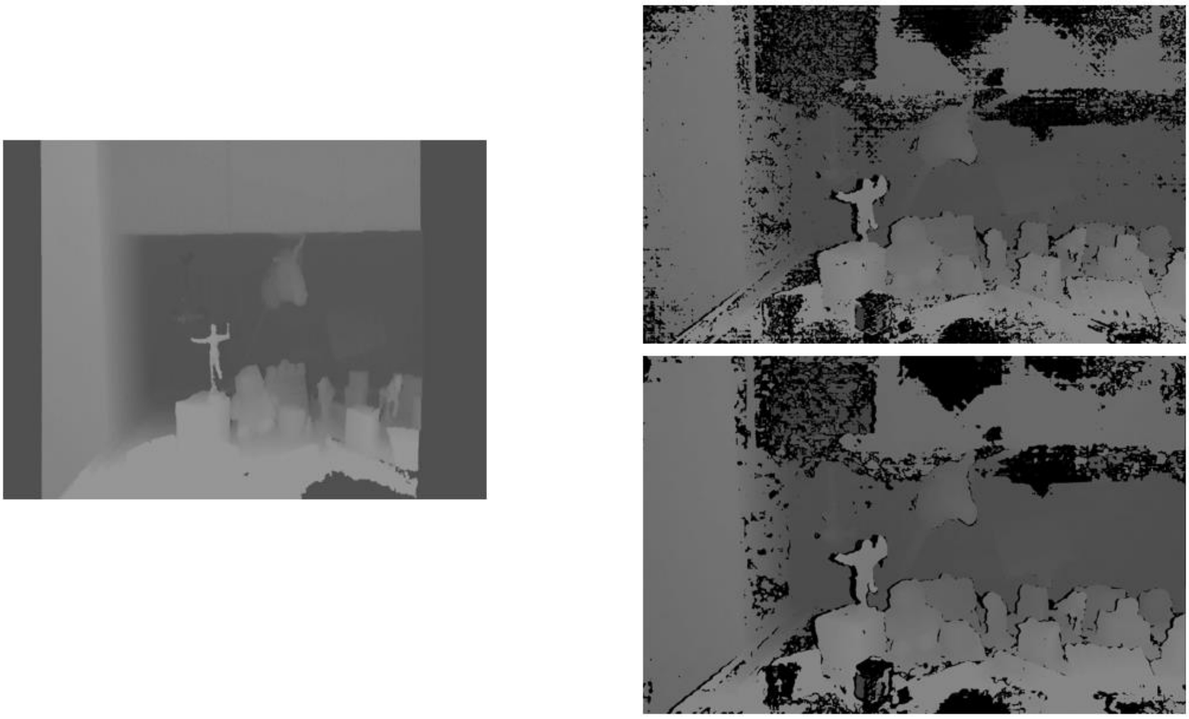


Figure 3. View-1(band-1, texture-1)
Ground truth (left)
Estimated disparity (right/row1: method-1)
Estimated disparity (right/row2: method-2)

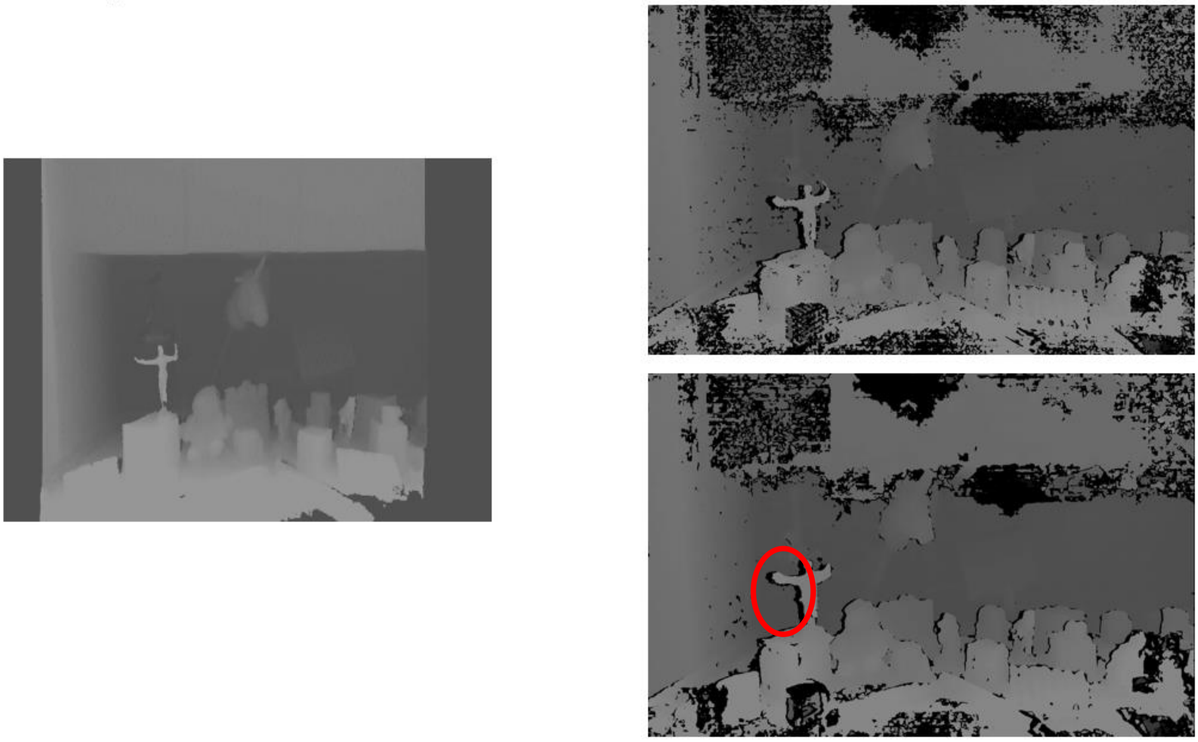


Figure 4. View-195(band-1, texture-1)
Ground truth (left)
Estimated disparity (right/row1: method-1)
Estimated disparity (right/row2: method-2)

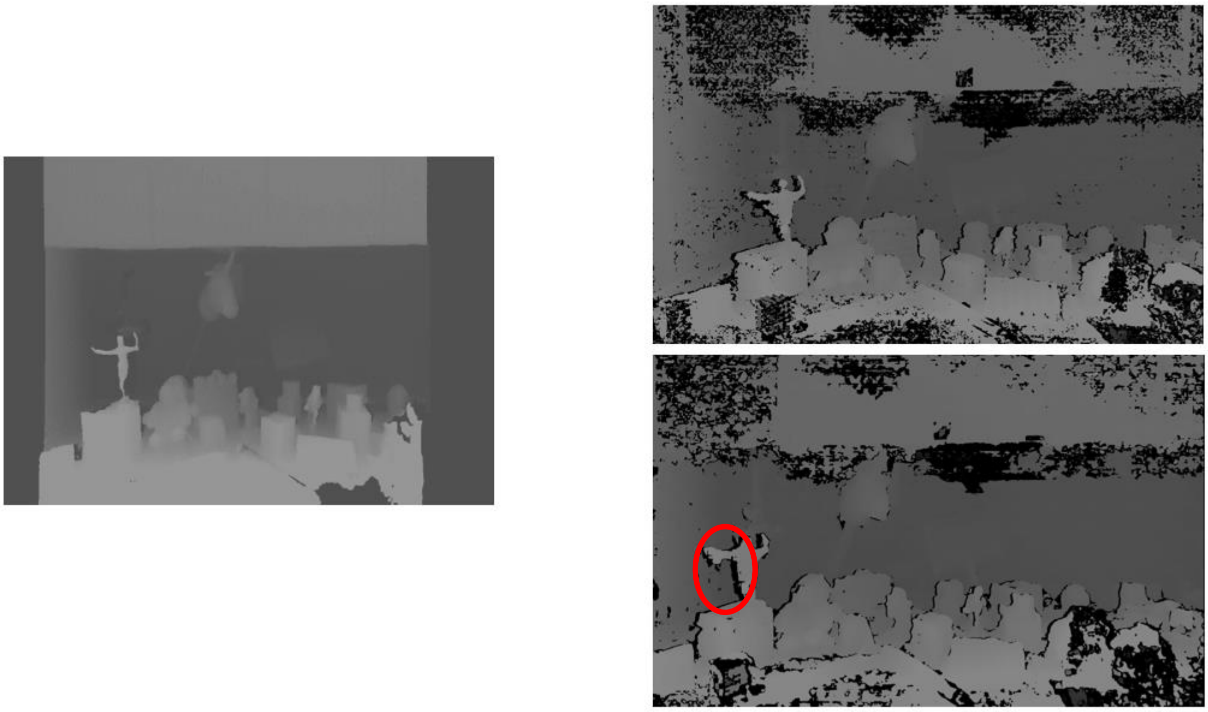


Figure 5. View-341(band-1, texture-1)
Ground truth (left)
Estimated disparity (right/row1: method-1)
Estimated disparity (right/row2: method-2)

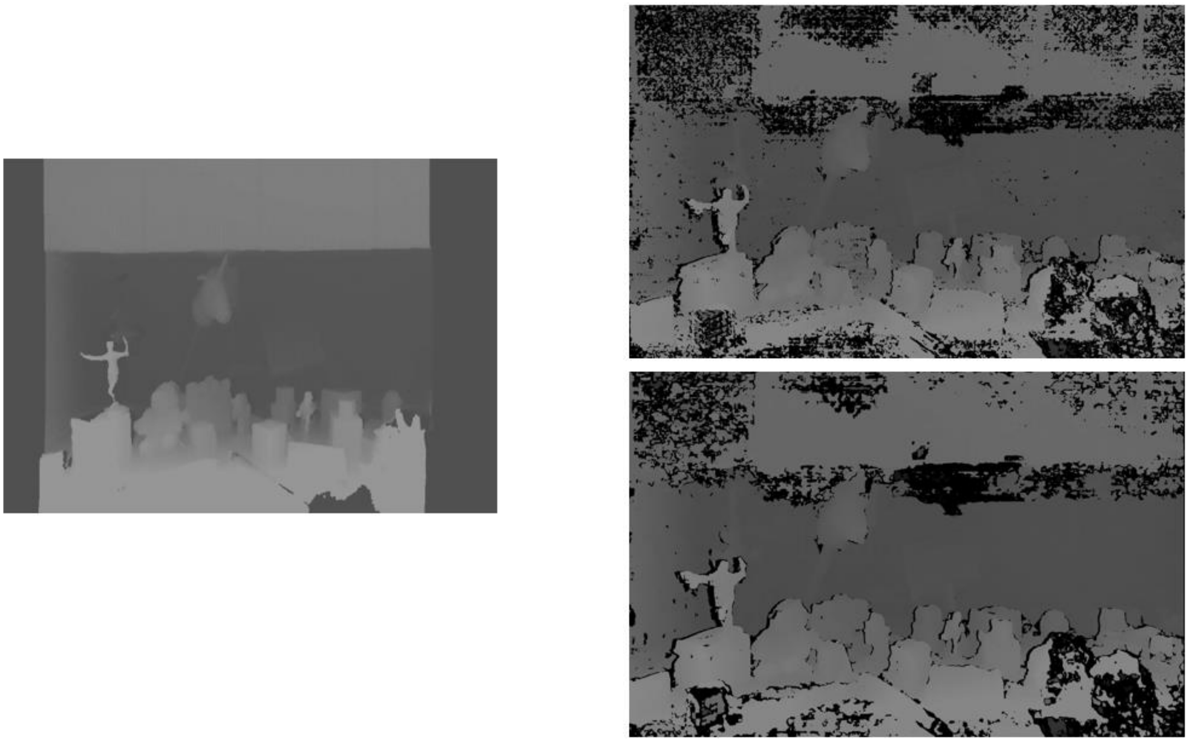


Figure 6. View-379(band-1, texture-1)
Ground truth (left)
Estimated disparity (right/row1: method-1)
Estimated disparity (right/row2: method-2)

The estimated depth map quality is objectively evaluated by computing BPR and RMSE value. In order to evaluate the depth map quality without the image padding area (horizontally appeared in both side), we crop that area and also same region is applied to ground truth depth map for evaluation.

	BPR(%)		RMSE	
	Method 1	Method 2	Method 1	Method 2
View-1	33.17	31.81	15.44	14.71
View-195	33.45	32.19	15.78	15.10
View-341	34.66	32.83	15.91	15.36
View-379	35.21	33.61	16.27	15.82

Table 1. BPR and RMSE comparison results

The performance of the proposed depth estimation method is compared with conventionally invented depth estimation method [3]. In table 1, method 1 and method 2 indicate conventionally invented method and occlusion aware depth estimation method respectively. Even though the occlusion aware based depth estimation method is applied, the results contain many error regions. Especially, the occlusion affects a depth quality where the view point is abruptly varied region such as red circle in figure 4 and 5. Since the used number of image is diminished when estimating the depth map from the horizontal light field image, the occlusion problem is gradually increased.

Conclusion

In this document we have explained occlusion aware depth estimation method with light field EPI data. Since the EPI based depth map estimation method is affected by the density of test sequences, we use the ULB data set for experiment. For the objective quality evaluation, we compute the BPR and RMSE value. From the experiment results, we notice that the proposed method improves the depth map quality rather than conventionally invented depth map estimation method. However, we have to improve the depth map quality on the occlusion area and other error issues in future work.

Acknowledgement

This work was supported by the 'Civil-Military Technology Cooperation Program' grant funded by the Korea government.

References

[1] J. H. Mun, and Y. S. Ho, "Depth Map Generation from Light Field Epipolar Plane Image" ISO/IEC JTC1/SC29/WG11 MPEG2017/M41017, July 2017, Torino, Italy.

[2] D. Bonattto, A. Schenkel, T. Lenertz, Y. Li, and G. Lafruit, "ULB High Density 2D/3D Camera Array Data Set, version2," ISO/IEC JTC1/SC29/WG11 MPEG2017/M14083, July 2017, Torino, Italy.

[3] G. Lafruit, J. Jung, K. Wegner, and D. Doyen, "Exploration Experiments for MPEG-I: Windowed-6DoF," ISO/IEC JTC1/SC29/WG11 MPEG2017/N17070, July 2017, Torino, Italy.