

Supplementary Materials:

Saliency Guided Depth Calibration for Perceptually Optimized Compressive Light Field 3D Display

Shizheng Wang¹, Wenjuan Liao¹, Phil Surman², Zhigang Tu², Yuanjin Zheng², and Junsong Yuan³

¹Institute of Microelectronics of Chinese Academy of Sciences, China

²School of EEE, Nanyang Technological University, Singapore

³CSE department, University at Buffalo, USA

In this document, we provide additional introduction and results in support of the primary manuscript. Part A presents an extended introduction of the duality properties of light field capture and display on the Lytro camera and multi-layer display for highlighting the proposed motivation. Part B includes additional descriptions and results on the proposed contrast-enhanced light field saliency detection and on the saliency-guided light field display in order to assess the advances. The results and demo EXE of the proposed light field saliency detection and display are introduced in Part C.

A. Duality of Light Field Capture and Display

This section provides a brief introduction to light field capture and display using the Lytro camera and multi-layer LCDs for further clarifying the special properties of light field images and the motivation of the proposed display.

A.1. Introduction of Lytro Camera and Saliency Detection Dataset

Here, Figure S.1 illustrates the light field imaging structure of a light field camera [1-2]. This type of camera has two sets of lenses; the main lens and the microlens array. Rays can be seen as intersecting behind and in front of the microlens array as a' and b' . The range a to b over which rays can intersect is regarded as the re-focusable range. This re-focusable depth information can be outputted from the Lytro software which has two components; the front re-focusable position a and the rear re-focusable position b .

Based on this type of Lytro camera, the public dataset [3] used in this paper provides the following materials: (1) a set of re-focus images, named 'focus stack', (2) the all-focus image which can be generated based on the focus stack with official Lytro software [4], (3) the depth map related to the all-focus image, (4) the raw light field data which can be mapped to the original lenslet image and related multi-view images based on the public MATLAB toolbox [5] and (5) the manual ground truth of the saliency detection result.

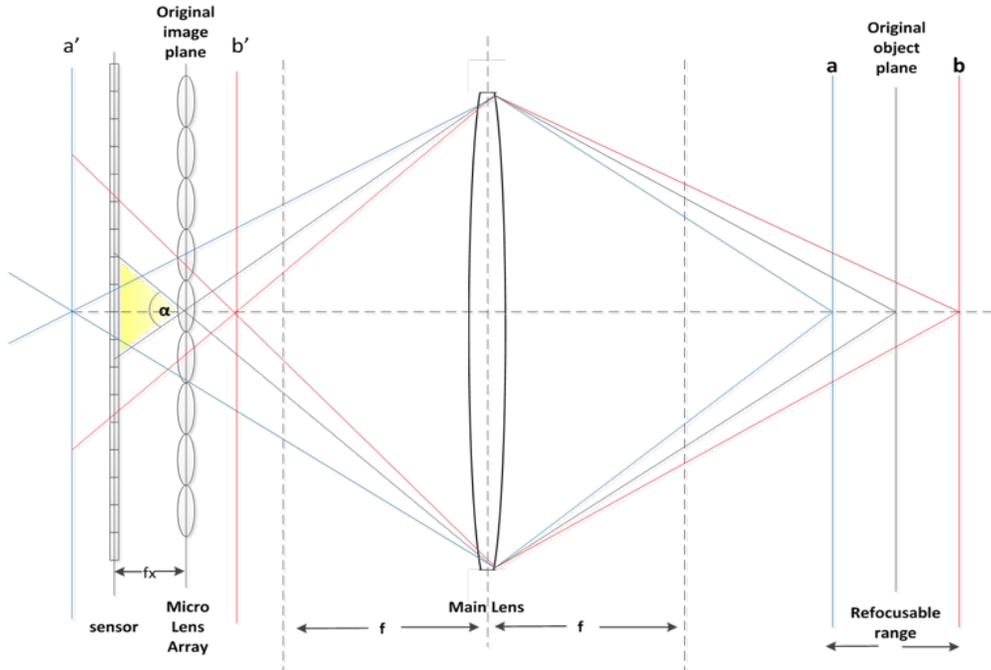


Figure S.1. The structure of Lytro light field camera. The black rays denote the trajectory of original light rays. In the image domain, virtual planes a' and b' correspond to object plane a and b , respectively.

A.2. Properties Demonstration of Light Field Images

Figure S.2 illustrates the all-focus image, depth image and lenslet image based on the Lytro camera. The yellow and blue lenslet images are related to the yellow and blue disks on the all-focus image. Figure S.3 demonstrates the multi-view performance and refocus ability of the light field based on the Lytro camera [6].

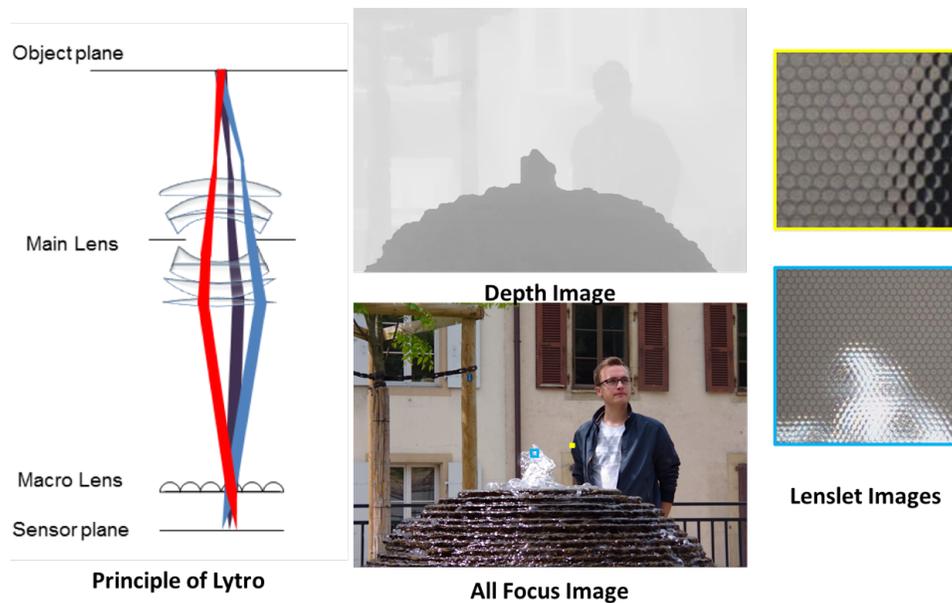
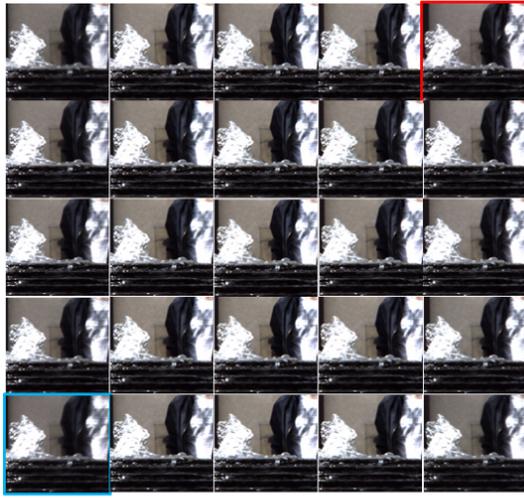
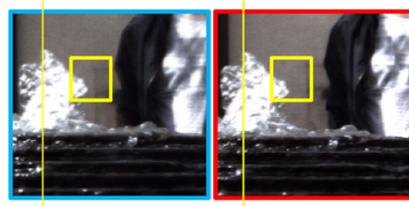


Figure S.2. The illustrations of all-focus image, depth image and lenslet image based on Lytro camera.



5x5 Multi-view Images



Differences between Multi-view Images



Refocus Images Stacks

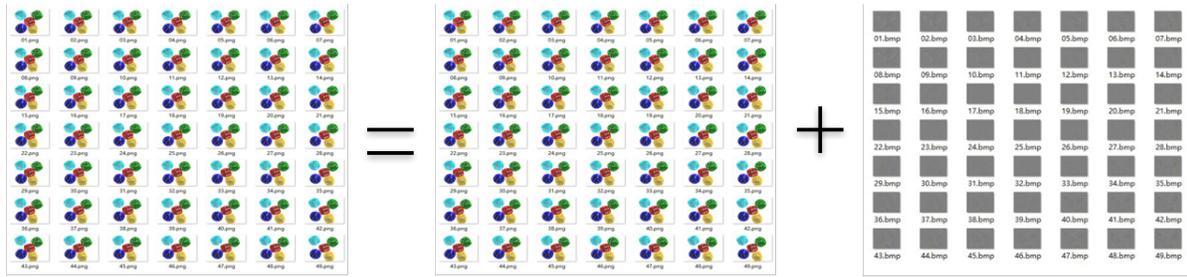
Figure S.3. The illustrations of light field properties of multi-view display and refocus display with the same raw data.

A.3. Multi-layer Light Field Display

In this section, the multi-layer light field display is explained from the viewpoint of image coding [7]. Here we take two-layer optimization for example. L is the original light field matrix. A and B are the vectors represented the transmittance of the front layer and rear layer, respectively. W is a 0/1 value matrix which controls the optimized region for target light field. The reconstructed light field can be generated by the Hadamard product of W and AB^T , as shown in Equation (1):

$$\arg \min_{\{0 \leq A \leq 1, 0 \leq B \leq 1\}} \frac{1}{2} \|L - W \circ (AB^T)\|_2^2 = \mathbf{Min}\{Residual\} = \mathbf{Min}\{Original - Reconstruction\}. \quad (1)$$

As illustrated in Figure S.4, inspired by the idea of 3D image coding, taking the light field image “dice” [8] for instance, the layered optimization for light field display could be regarded as encoding 7x7 multi-view images into two-layer multi-layer images. The whole optimization of Equation (1) could also be seen as the minimization of residual between original light field and reconstructed light field.



Original Light Field **Reconstructed Light Field** **Residual**
 Predicted Light Field **Reconstruction Error**

Prediction Using Layered Images :



Figure S.4. An example of layered optimization in a view of light field 3D image coding.

A.4. Proposed Motivation

Figure S.5 further demonstrates the proposed motivation using an example from the public Lytro dataset [3]. Based on the saliency detection result the initialization method for light field reconstruction with multi-layer LCDs could be different. Since the depth range for this type of display is only comparable with 2D spatial resolution in a small range from the middle of LCDs, the proposed optimization could cover more saliency objects in the displayed depth range compared to the conventional initialization way which displays light field just from the center of depth range.

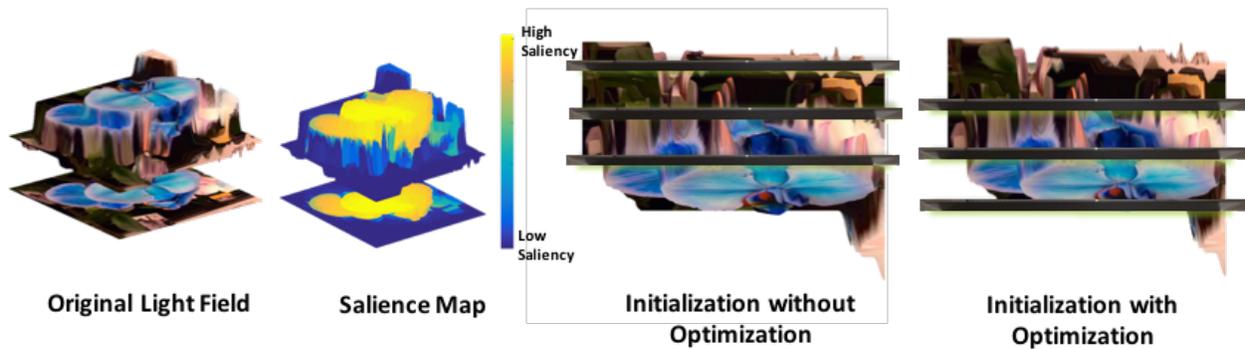


Figure S.5. An example of saliency guided depth calibration for perceptually optimized compressive light field 3D display.

B. Additional Experimental Results of Light Field Saliency Detection and Layered Display

In the primary manuscript, we mainly compare the proposed saliency detection approach with the state-of-the-art light field saliency detection algorithms which also have publications based on the public light field saliency dataset. Here, following the survey of 2D saliency detection models in [9], we will discuss more comparative tests with more state-of-the-art saliency detection methods on 2D inputs. The subjective performance test and related supporting materials are also discussed and released in this section.

B.1. Light Field Saliency Detection

We use standard precision recall (PR) and receiver operating characteristic (ROC) curves for evaluations. When computing the overall quality on the whole dataset, we consider three metrics for determining the accuracy of saliency detection: F-measure, area under curve (AUC) and mean absolute error (MAE), whose accurate definition and computation equations can be found in references [9, 10].

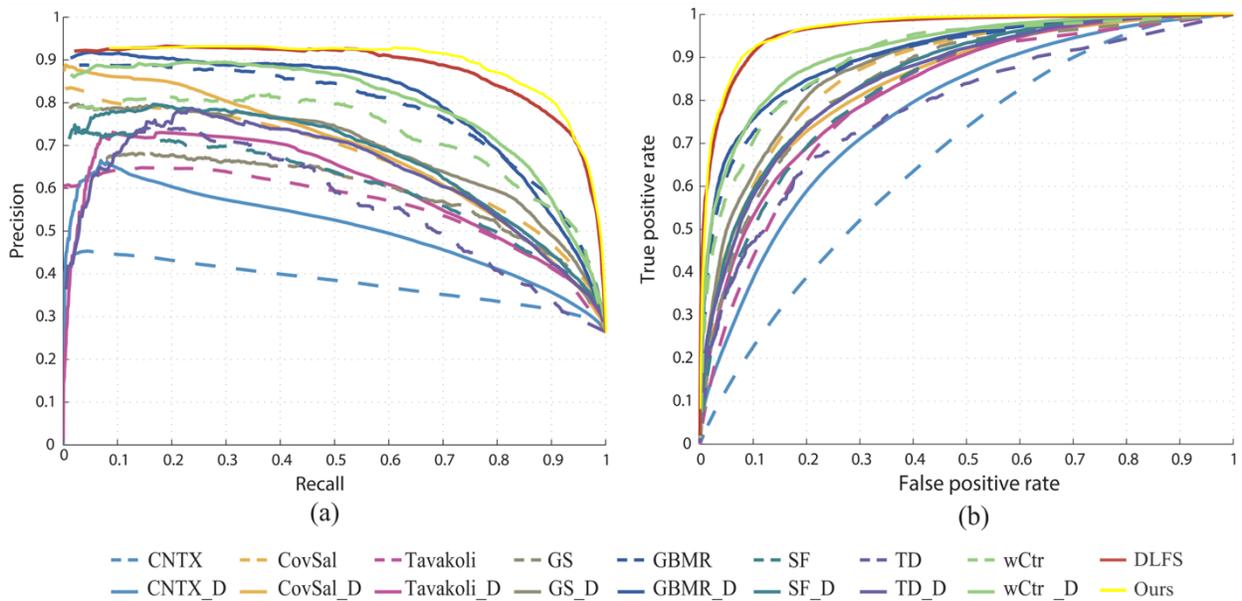


Figure S.6. Quantitative results of our method, state-of-the-art 2D methods and depth-extended versions. (a) PR curves; (b) ROC curves.

State-of-the-art 2D saliency approaches have previously not been tailored for light field data, but only for 2D image inputs. To validate the benefit of light field depth, we extend the experiments with 8 state-of-the-art 2D saliency approaches. These methods include Tavakoli [11], CNTX [12], GS [13], SF [14], TD [15], CovSal [16], GBMR [17] and wCtr [18]. We set the models all with default parameters in their original implementations.

Figure S.6 presents the PR and ROC curves of our results. The comparisons of F-measure, ROC and MAE are given in Table S.1. Here the suffix ‘_D’ denotes further fusing 2D saliency maps with light field depth contrast saliency maps into final ones through the standard pixel-wise summation. Note that our approach is superior to all the state-of-the-art 2D models, even combined with the light field depth-induced saliency and the latest tailored light field model [9]. Although the accuracy from all the 2D saliency methods improve about 1~5% and 3~6% on F-measure and MAE, respectively, by incorporating the light field depth saliency, they still cannot exceed the light field models. This is because this type of comparison is unfair from the input of raw data, light field saliency models can make use of more cues like focus cue for salience detection.

The experimental results of the post-optimization are also highlighted in Table S.2, which shows that the proposed algorithm is better than the state-of-the-art [9] by either using or not using post-processing. Here the suffix ‘_w/oP’ denotes the results without post-optimization.

MODEL	F-MEASURE	AUC	MAE
CNTX	0.3643	0.6700	0.3574
CNTX_D	0.4123	0.7718	0.3514
COVSAL	0.6335	0.8599	0.2417
COVSAL_D	0.6373	0.8466	0.2850
TAVAKOLI	0.5498	0.8078	0.2551
TAVAKOLI_D	0.5711	0.8276	0.2903
GS	0.5944	0.8443	0.2395
GS_D	0.6217	0.8792	0.2843
GBMR	0.7461	0.8965	0.1822
GMRR_D	0.7536	0.9072	0.2415
SF	0.4678	0.8301	0.2468
SF_D	0.4704	0.8552	0.2903
TD	0.5766	0.7775	0.2623
TD_D	0.5999	0.849	0.2951
WCTR	0.6996	0.8991	0.1878
WCTR_D	0.7382	0.9156	0.2475
DLFS	0.8186	0.9641	0.1363
OURS	0.8424	0.9640	0.1040

Table S.1. Comparison of F-measure, AUC and MAE from our method, state-of-the-art 2D methods and their light field extended method (red: the best; blue: the second best).

MODEL	F-MEASURE	AUC	MAE
DLFS	0.8186	0.9641	0.1363
DLFS_w/oP	0.7749	0.8982	0.1605
OURS	0.8424	0.9640	0.1040
OURS_w/oP	0.8276	0.9716	0.1573

Table S.2. Comparison of F-measure, AUC and MAE from our method, state-of-the-art 3D methods and their light field extended method (red: the best; blue: the second best).

B.2. Light Field Layered Display

SCORE	EVALUATION
0	Bad (salient objects too blurred to see)
1	Middle(salient objects are blurred but can see)
2	Good(salient objects are not blur)
3	Best(salient objects are clear in 3D)

Table S.3. Evaluation scores.

In this paper, 12 individuals whose ages from 24 to 70 also evaluate the subjective performances of variable prototype configurations on the prototype shown in Figure 7 of the primary manuscript. The ranks of the display performance are recorded to determine the subjective score that subjects could assess according to the related weightings shown in TABLE 3. The configuration which display light field test images with clear salient object as well as real 3D performance will be recoded as best. Different configurations could have the same score, even two coordinate best for one test image if their subjective performances seem the same.

This evaluation is implemented on images one by one. The performance of optimized depth initialization was viewed as the best with 75 cases in the 100 test images, and its average score is 264. The configuration A has 53 best with 238 score, while configuration B and configuration C only have 5 best and 3 best in the subjective experiment.

C. Experimental Results and Demo EXEs

More results and demo EXEs of proposed light field salience detection and display are released and introduced in this section, including:

Material 1: Demo for a simulated three-layer light field display based on the test image “Messerschmitt” from [8].

Material 2: A light field media player for multi-layer LCDs. It is based on the public media player [19] and public light field dataset [8].

Material 3: The GPU speed-up layered decomposition software tailored from the simultaneous algebraic reconstruction technique (SART) algorithm and its cg open source code [20] is also released and demonstrated on the public Lytro camera dataset [3].

Material 4: The original salience detection results of the proposed method both with and without post-optimization are released. The reported numbers are generated on the platform of public salience test [10].

All these Materials can be downloaded from <http://shizhengwang.info/projects/LFSDisplay.html>.

References

- [1] Maeno K, Nagahara H, Shimada A, Taniguchi RI. Light field distortion feature for transparent object recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2013 (pp. 2786-2793).
- [2] Xu Y, Maeno K, Nagahara H, Taniguchi RI. Mobile camera array calibration for light field acquisition. arXiv preprint arXiv:1407.4206. 2014 Jul 16.
- [3] Li N, Ye J, Ji Y, Ling H, Yu J. Saliency detection on light field. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2014 (pp. 2806-2813).
- [4] Official lytro software. <https://www.lytro.com>
- [5] Light field toolbox 0.4. <http://lightfield-forum.com/tag/lightfield-toolbox/>
- [6] Rerabek M, Ebrahimi T. New light field image dataset. In 8th International Conference on Quality of Multimedia Experience (QoMEX) 2016 (No. EPFL-CONF-218363).
- [7] Wang Q, Lin H, Ma Y, Kang SB, Yu J. Automatic Layer Separation using Light Field Imaging. arXiv preprint arXiv:1506.04721. 2015 Jun 15.
- [8] Wetzstein G, Lanman D, Hirsch M, Raskar R. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting.
- [9] Zhang, J., Wang, M., Gao, J., Wang, Y., Zhang, X., & Wu, X. (2015, July). Saliency Detection with a Deeper Investigation of Light Field. In IJCAI (pp. 2212-2218).
- [10] Li, X., Li, Y., Shen, C., Dick, A., & Van Den Hengel, A. (2013). Contextual hypergraph modeling for salient object detection. In Proceedings of the IEEE international conference on computer vision (pp. 3328-3335).
- [11] Tavakoli, H. R., Rahtu, E., & Heikkilä, J. (2011, May). Fast and efficient saliency detection using sparse sampling and kernel density estimation. In Scandinavian Conference on Image Analysis (pp. 666-675). Springer Berlin Heidelberg.
- [12] Stas Goferman, Lihi Zelnik-Manor, and Ayellet Tal. Context-aware saliency detection. TPAMI, 34(10):1915–1926, 2012.
- [13] Wei, Y., Wen, F., Zhu, W., & Sun, J. (2012, October). Geodesic saliency using background priors. In European conference on computer vision (pp. 29-42). Springer Berlin Heidelberg.
- [14] Perazzi, F., Krähenbühl, P., Pritch, Y., & Hornung, A. (2012, June). Saliency filters: Contrast based filtering for salient region detection. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 733-740). IEEE.
- [15] Scharfenberger, C., Wong, A., Fergani, K., Zelek, J. S., & Clausi, D. A. (2013). Statistical textural distinctiveness for salient region detection in natural images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 979-986).
- [16] Erdem, E., & Erdem, A. (2013). Visual saliency estimation by nonlinearly integrating features using region covariances. Journal of vision, 13(4), 11-11.
- [17] Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang. Saliency detection via graph-based manifold ranking. In CVPR, pages 3166–3173, 2013.
- [18] Wangjiang Zhu, Shuang Liang, Yichen Wei, and Jian Sun. Saliency optimization from robust background detection. In CVPR, pages 2814–2821, 2014.
- [19] Wang S, Ong KS, Surman P, Yuan J, Zheng Y, Sun XW. Quality of experience measurement for light field 3D displays on multilayer LCDs. Journal of the Society for Information Display. 2016 Dec 1;24(12):726-40.
- [20] Lanman, D., Wetzstein, G., Hirsch, M., Heidrich, W., Raskar, R. "Polarization Fields: Dynamic Light Field Display using Multi-Layer LCDs". ACM Trans. Graph. (Siggraph Asia) 2011.