

Received January 14, 2020, accepted February 9, 2020, date of publication February 20, 2020, date of current version March 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2975209

Comparison of Layer Operations and Optimization Methods for Light Field Display

KEITA MARUYAMA¹, KEITA TAKAHASHI¹, (Member, IEEE),
AND TOSHIKI FUJII¹, (Member, IEEE)

Graduate School of Engineering, Nagoya University, Nagoya 464-8603, Japan

Corresponding author: Keita Maruyama (maruyama@fujii.nuce.nagoya-u.ac.jp)

ABSTRACT A light-field display provides not only binocular depth sensation but also natural motion parallax with respect to head motion, which invokes a strong feeling of immersion. Such a display can be implemented with a set of stacked layers, each of which has pixels that can carry out light-ray operations (multiplication and addition). With this structure, the appearance of the display varies over the observed directions (i.e., a light field is produced) because the light rays pass through different combinations of pixels depending on both the originating points and outgoing directions. To display a specific 3-D scene, these layer patterns should be optimized to produce a light field that is as close as possible to that produced by the target three-dimensional scene. To deepen the understanding for this type of light field display, we focused on two important factors: light-ray operations carried out using layers and optimization methods for the layer patterns. Specifically, we compared multiplicative and additive layers, which are optimized using analytical methods derived from mathematical optimization or faster data-driven methods implemented as convolutional neural networks (CNNs). We compared combinations within these two factors in terms of the accuracy of light-field reproduction and computation time. Our results indicate that multiplicative layers achieve better accuracy than additive ones, and CNN-based methods perform faster than the analytical ones. We suggest that the best choice in terms of the balance between accuracy and computation speed is using multiplicative layers optimized using a CNN-based method.

INDEX TERMS Light field, 3D display, CNN.

I. INTRODUCTION

Three-dimensional (3-D) displays have been the subject of study for many years [1]–[5]. These displays can be categorized on the basis of several criteria such as the necessity of wearing glasses and the number of supported viewing directions. Glasses-free (naked-eye) displays have attracted attention because they enable a more natural viewing experience than glasses-based ones. Multi-view displays have more potential than conventional stereo-only displays because they not only provide depth perception by showing different images to the left and right eyes but also present natural motion parallax in accordance with the movement of observers. In particular, multi-view displays that can support many and dense viewing directions are often referred to as light-field displays.

The associate editor coordinating the review of this manuscript and approving it for publication was Ganesh Naik¹.

To develop glasses-free multi-view/light-field displays, researchers have devised several methods, including those that use parallax barriers [1], [6]–[8], specially designed lenses (lenticular screens or integral photography lenses) [2], [3], [9]–[11], and stacked layers [12]–[17]. In this paper, we focus on the third method. This type of display, called a “layered display,” can be viewed from many directions (angles) simultaneously without the resolution of each viewing direction being sacrificed (but the quality of each view degrades), which is deemed as one of the desirable properties for glasses-free multi-view/light-field displays.

The structure of a layered display is illustrated in Fig. 1. A few layers, each of which has pixels that can carry out light-ray operations (multiplication and addition), are stacked with small intervals. With this structure, the appearance of the display varies over the observed directions, because the light rays pass through different combinations of pixels depending on both the originating points and outgoing directions. To display a specific 3-D scene with this structure, the layer patterns

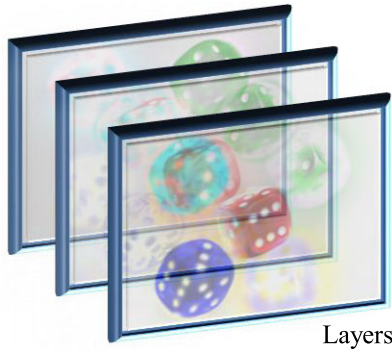


FIGURE 1. Structure of layered light-field display.

should be designed to make the direction-dependent views consistent with the appearance of the target 3-D scene. More precisely, a light field [18], [19] (i.e., tens of images), which is expected to be observed from different viewing directions, is given as the input, then the layer patterns are optimized to reproduce the light field as accurately as possible. This design can also be applied to light-field projections [20], [21], head-mounted displays [22], [23], and table-top displays [24].

To deepen the understanding for this type of light field display, we conducted comparisons based on two important factors: light-ray operations carried out using layers and optimization methods for the layer patterns. We compared multiplicative and additive layers. Multiplicative layers are implemented using liquid crystal display (LCD) panels and backlight [16]. Additive layers are constructed with holographic optical elements (HOEs) and projectors [21]. Regarding the optimization of the layer patterns, we compared analytical methods used in previous studies [16], [21], which are slow due to heavy computation, with faster data-driven methods implemented as convolutional neural networks (CNNs) [25]. Such a CNN-based method requires significant time for training, but inference using a trained network is very fast, which paves the way for light-field displays running at video-rate speed.

We compared combinations within these factors (layers operations and optimization methods) in terms of the accuracy of light-field reproduction and computation time. To the best of our knowledge, we are the first to present such comparisons for layered light-field displays.¹ Our results indicate that multiplicative layers achieve better accuracy than additive ones, and CNN-based methods are faster than the analytical methods, with comparable accuracy to the best achievable accuracy of analytical methods. We recommend that in terms of the balance between accuracy and computation speed, one should adopt multiplicative layers optimized using a CNN-based method.

¹A preliminary version of this paper was presented at a conference [26]. A more complete description, thorough discussions, and additional experimental results are included in the present paper.

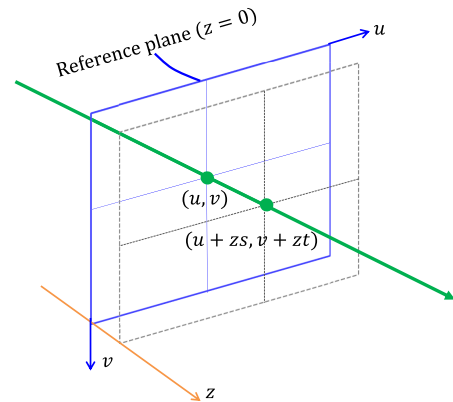


FIGURE 2. Definition of light field. Each light ray is parameterized by the point of intersection with reference plane (u, v) and outgoing direction (s, t) . It intersects with plane located at depth z (shown with dotted lines) on $(u + zs, v + zt)$.

II. LAYERED LIGHT-FIELD DISPLAY

We first introduce a parameterization for a light field. We then mention how a light field is produced using stacked layers with which multiplication or addition of light rays are carried out, as illustrated in Fig. 3. Finally, we describe the coordinate system we used in this study.

A. LIGHT-FIELD PARAMETERIZATION

A light field is defined as a 4-D function describing all the light rays that travel straight in free space [18], [19]. In this paper, we adopt a plane + angle parameterization, as shown in Fig. 2. A reference plane ($z = 0$) is defined, and a light ray is parameterized by the point of intersection with the reference plane $[(u, v)]$ and the outgoing direction with respect to the z axis $[(\theta, \phi)]$. The intensity of each light ray is described as $L(s, t, u, v)$ with $s = \tan(\theta)$ and $t = \tan(\phi)$. We assume that all the elements of $L(s, t, u, v)$ take non-negative values because the light intensity is non-negative.

B. MULTIPLICATIVE LAYERS

Regarding multiplicative layers [16], we assume that a few light-attenuating panels (e.g. LCD panels) are stacked with evenly spaced intervals in front of a backlight. Let us consider a light ray passing through point (u, v) on the reference plane and going in the direction of (s, t) . We can see that the intersection of this light ray with a layer located at depth z is $(u + zs, v + zt)$. Therefore, the intensity of a light ray (normalized by the intensity of the backlight) emitted from this display can be described as

$$L_{mul}(s, t, u, v) = \prod_{z \in \mathcal{Z}} P_z(u + zs, v + zt), \quad (1)$$

where $P_z(u, v)$ denotes the transmittance of a layer located at z and \mathcal{Z} denotes a set of depths where the layers are located.

C. ADDITIVE LAYERS

Regarding additive layers [21], we assume that they are evenly spaced and the luminance of layer pixels are summed

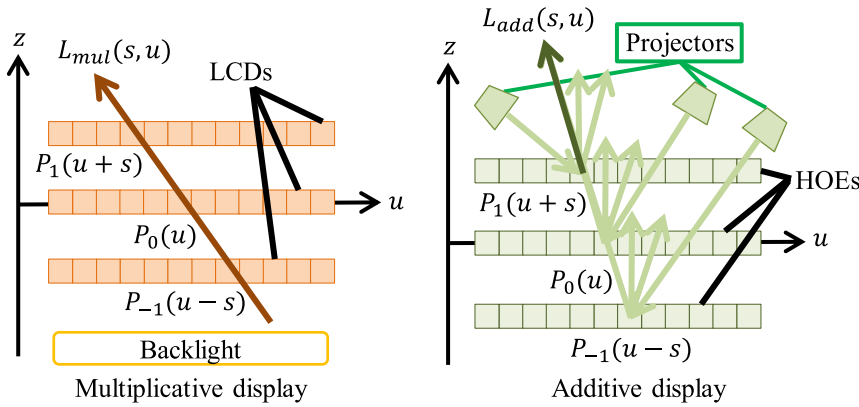


FIGURE 3. Configurations of layered light-field displays, in which layers are evenly spaced in parallel to reference plane. Light field is defined in 4-D space, but we represent it in 2-D subspace for simplicity.

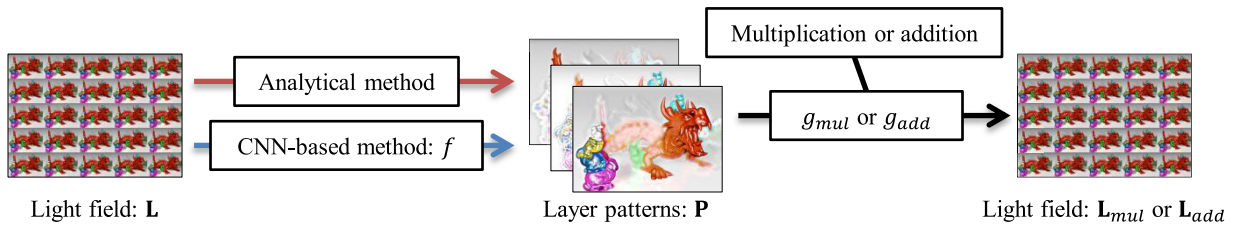


FIGURE 4. Diagram of different layer operations (additive and multiplicative) and optimization methods.

along the path of a light ray. Such operations can be implemented using HOEs and projectors. A light ray emitting from this display can be given as

$$L_{add}(s, t, u, v) = \sum_{z \in \mathcal{Z}} P_z(u + zs, v + zt), \quad (2)$$

where $P_z(u, v)$ denotes the luminance of a pixel (u, v) on the layer located at z .

D. COORDINATE SYSTEM

Throughout the paper, we assume that all four variables (s, t, u, v) in a light field are integers. With this assumption, a light field can be regarded as a set of directional views: $L_{s,t}(u, v) = L(s, t, u, v)$, where (s, t) corresponds to an index of a viewpoint (viewing direction) and (u, v) indicates a discrete pixel position. We assume that a light field consists of 5×5 views; thus, s and t are limited within the range of $[-2, 2]$. We also assume that a light-field display is composed of three layers located at $\mathcal{Z} = \{-1, 0, 1\}$. Note that z corresponds to the disparity among the directional views rather than the physical length.

III. OPTIMIZATION METHODS

We now describe how the multiplicative or additive layer patterns are optimized to display a target 3-D scene. A light field that should be emitted from the display is described as $L(s, t, u, v)$. The optimization goals for the layer patterns are

expressed as

$$\arg \min_{P_z | z \in \mathcal{Z}} \sum_{s,t,u,v} \|L(s, t, u, v) - L_{mul}(s, t, u, v)\|^2, \quad (3)$$

$$\arg \min_{P_z | z \in \mathcal{Z}} \sum_{s,t,u,v} \|L(s, t, u, v) - L_{add}(s, t, u, v)\|^2. \quad (4)$$

To achieve these goals, we use the two types of optimization methods shown in Fig. 4. We first describe the analytical methods for optimizing the layer patterns both for multiplicative and additive layers then data-driven CNN-based methods as effective alternatives.

A. ANALYTICAL METHODS

We first describe an analytical optimization method for multiplicative layers, which is equivalent to the multiplicative update rule mentioned in a previous study [16]. Since (3) is a non-convex problem, we resort to an alternative optimization. First, all the layer patterns are initialized with random values in the range of $[\epsilon, 1]$, where ϵ is a sufficiently small positive number. Next, we carry out optimization for one layer at a time and circulate the optimization for all the layers until convergence.

When optimizing a specific layer $P_z(u, v)$, we assume that the other layers $P_{z'}(u, v)$ ($z' \in \mathcal{Z} \setminus \{z\}$) are fixed. We define a column vector \mathbf{l} that includes all the elements of the light field $L(s, t, u, v)$. Similarly, we also define a column vector \mathbf{p}_z including all the elements of $P_z(u, v)$. We can also introduce

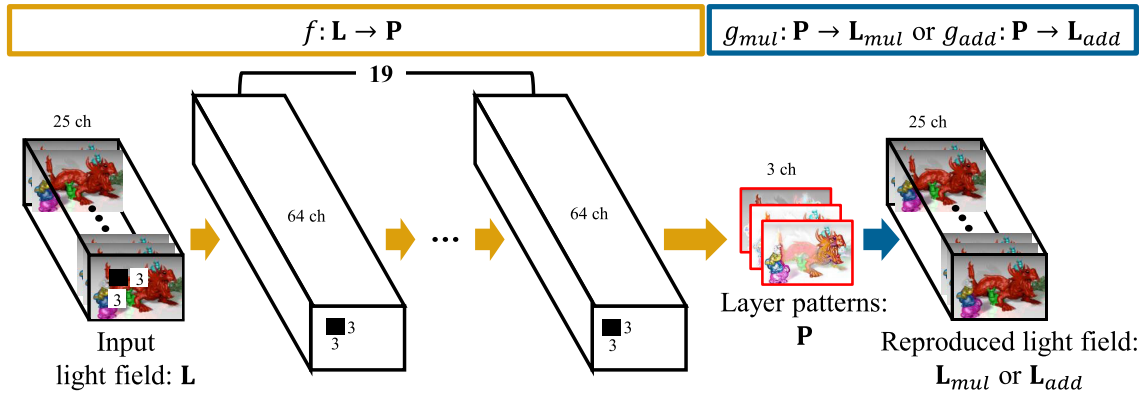


FIGURE 5. Network architecture for optimizing layer patterns.

a matrix $\mathbf{A}_{z'}$, with which $\mathbf{A}_{z'}\mathbf{p}_z$ corresponds to the light-field vector generated by the layer patterns. Note that $\mathbf{A}_{z'}$ is determined by the fixed layers $z' \in \mathcal{Z} \setminus \{z\}$, and all the elements are non-negative. The optimization of the target layer $P_z(u, v)$ is formulated in a non-negative least square problem as

$$\arg \min_{\mathbf{p}_z} \|\mathbf{l} - \mathbf{A}_{z'}\mathbf{p}_z\|^2. \quad (5)$$

The squared error can be reduced using the multiplicative update rule as follows:

$$\mathbf{p}_z \leftarrow \mathbf{p}_z \odot (\mathbf{A}_{z'}^T \mathbf{l}) / (\mathbf{A}_{z'}^T \mathbf{A}_{z'} \mathbf{p}_z) \quad (6)$$

where \odot and $//$ represent element-wise product and division, respectively. Since all the elements in the right-hand side are non-negative, the left hand-side is ensured to be always non-negative. After (6) is applied, all the elements of \mathbf{p}_z are clipped to $[\epsilon, 1]$.

Next, we describe a similar analytical method for the additive layers. In contrast to the multiplicative case, (4) is a convex problem; thus, the layer patterns for all depths $z \in \mathcal{Z}$ can be optimized simultaneously. We define a column vector \mathbf{p} that contains all the elements of all the layer patterns $P_z(u, v)$ ($z \in \mathcal{Z}$). We can also define a matrix \mathbf{A} with which $\mathbf{A}\mathbf{p}$ corresponds to the light field produced by the layer patterns \mathbf{p} . All the elements of \mathbf{A} are also non-negative (more specifically, 0 or 1). Using these variables, the optimization is formulated also in a non-negative least square problem [21] as

$$\arg \min_{\mathbf{p}} \|\mathbf{l} - \mathbf{A}\mathbf{p}\|^2. \quad (7)$$

Similarly to the case with multiplicative layers, we can apply the multiplicative update rule described as

$$\mathbf{p} \leftarrow \mathbf{p} \odot (\mathbf{A}^T \mathbf{l}) / (\mathbf{A}^T \mathbf{A} \mathbf{p}) \quad (8)$$

After (8) is applied, all the elements of \mathbf{p} are clipped to $[\epsilon, 1]$. This process is repeated until convergence after initializing all layer patterns with random values in $[\epsilon, 1]$.

B. CNN-BASED METHODS

We also evaluated CNN-based methods for optimizing multiplicative and additive layer patterns, respectively. These methods were constructed based on our previous study [25], but with one significant difference. In that study [25], we obtained the layer patterns from a compressively sampled light field. In this study, however, the layer patterns were obtained from a full light field.

The optimization process for the layer patterns can be written in a form of mapping as

$$f: \mathbf{L} \rightarrow \mathbf{P}, \quad (9)$$

where \mathbf{L} represents a tensor that contains all the pixels of $L(s, t, u, v)$ for all (s, t) . Similarly, \mathbf{P} represents a tensor that contains all the pixels of $P_z(u, v)$ for all $z \in \mathcal{Z}$. To make the notations consistent, the mappings from the layer patterns to the light field ((1) and (2)) can be rewritten as

$$g_{mul}: \mathbf{P} \rightarrow \mathbf{L}_{mul}, \quad (10)$$

$$g_{add}: \mathbf{P} \rightarrow \mathbf{L}_{add}, \quad (11)$$

where \mathbf{L}_{mul} and \mathbf{L}_{add} represent all the light rays in $L_{mul}(s, t, u, v)$ and $L_{add}(s, t, u, v)$, respectively. We constructed two CNNs that correspond to the composite mappings $g_{mul} \circ f$ and $g_{add} \circ f$, respectively, and minimized the squared error loss given as

$$\arg \min_f \|\mathbf{L} - \mathbf{L}_{mul}\|^2, \quad (12)$$

$$\arg \min_f \|\mathbf{L} - \mathbf{L}_{add}\|^2 \quad (13)$$

over a massive amount of training samples.

The network architecture is rather straight-forward, as illustrated in Fig. 5. The network consisted of 20 2-D convolutional layers stacked in a sequence. Throughout the networks, the spatial size of the tensors was constant, but only the number of channels was changed. Tensors \mathbf{L} , \mathbf{L}_{mul} , and \mathbf{L}_{add} had 25 channels, each of which corresponds to a viewpoint. Tensors \mathbf{P} had 3 channels, each of which corresponds to the 3 layer patterns of the display. The other intermediate feature maps had 64 channels. During the training stage,

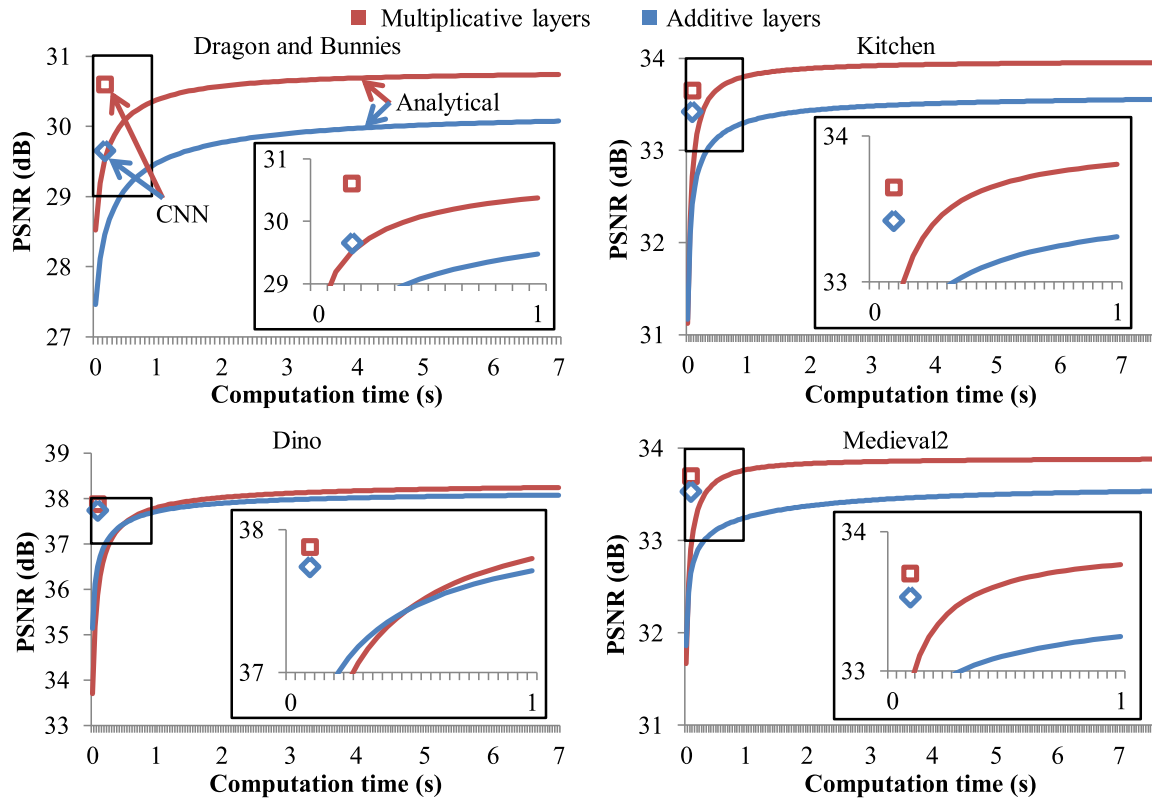


FIGURE 6. Computation time and accuracy of reproduced light fields.

training samples passed through the entire network. However, in a real application, only the mapping f is conducted on a computer, but the mapping g_{mul} or g_{add} is conducted using the physical display hardware.

IV. EXPERIMENTS

A. IMPLEMENTATION

We used a Linux-based PC equipped with a NVIDIA Geforce GTX 1080 Ti. We executed all the methods on the GPU for fair comparison. To implement the analytical methods for multiplicative and additive displays, we used the open-source matrix library CuPy, which enables the methods to be executed on a GPU.² Note that the analytical methods update the solution in an iterative manner; therefore, we need to investigate the trade-off between the computation time (the number of iterations) and accuracy of the solution. Regarding the CNN-based methods, we constructed two networks for multiplicative and additive displays. Following a previous study [25], we gathered training samples from several light-field datasets [27]–[30]. Three color channels of each dataset were used as three individual datasets. Each training sample was a set of 25 (corresponding to 5×5 views) 2-D image blocks with 64×64 pixels that were extracted from the same spatial positions in a light-field dataset. With data augmentation in

²These GPU implementations run much faster than the CPU counterpart written in C++ provided in a previous study [17].

the intensity levels, we finally collected 295,200 samples. The networks were implemented using Chainer version 3.2.0, a Python-based framework for neural networks. The batch size for training was set to 15. We used a built-in Adam optimizer. The number of epochs was 20 for both networks. Once the training finished, inference was conducted almost in a constant time. We made our software available from our website [31].

B. RESULTS

We evaluated four combinations (a multiplicative or additive display optimized with an analytical or CNN-based method) in terms of the computation time and the accuracy of the reproduced light fields. The accuracy was measured using the peak signal-to noise ratio (PSNR) against the original light field, which was obtained from the mean squared error over all 5×5 multi-view images and three color channels. The quantitative results with four light fields (which were not included in the training data) are summarized in Fig. 6. When the analytical methods were used, accuracy gradually increased along the computation time, and it took several seconds to converge. However, the CNN-based methods were very fast and achieved good accuracy, which is comparable to that analytical methods can reach after many iterations. We can conclude that CNN-based methods are better than analytical ones in terms of the balance between computation time and accuracy. Regarding the light-ray operations carried

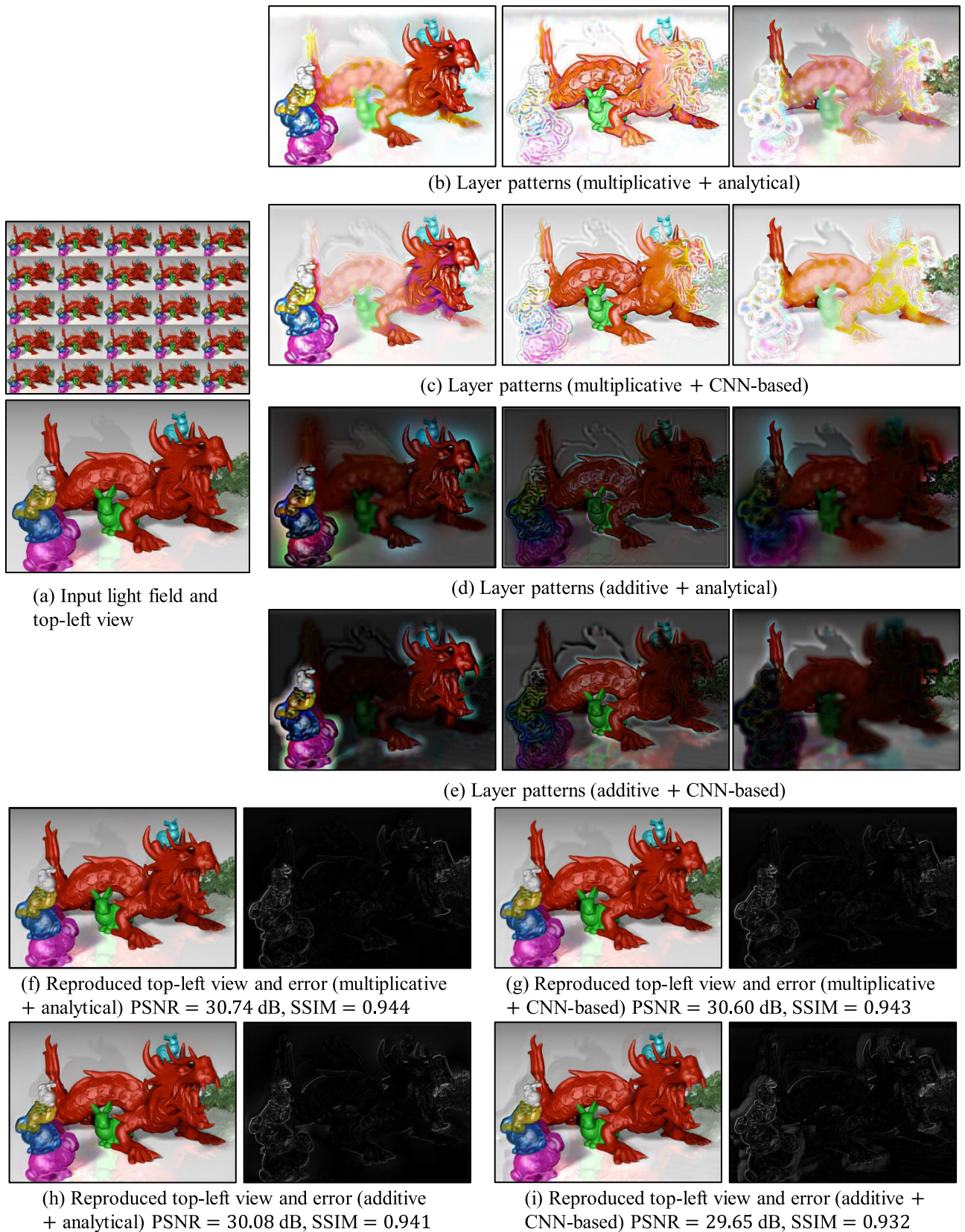


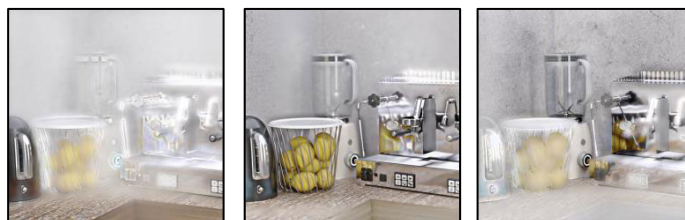
FIGURE 7. Visual results with *Dragon and Bunnies*: (a) input light field and top-left view, (b)–(e) layer patterns for multiplicative or additive display optimized with analytical or CNN-based method, (f)–(i) reproduced top-left views and errors.



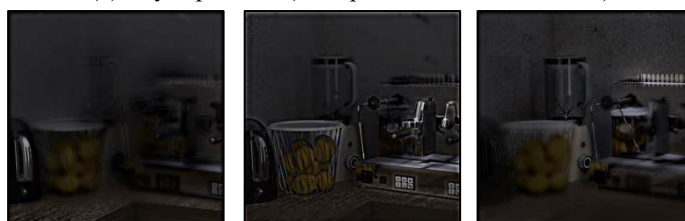
(a) Input light field and top-left view



(b) Layer patterns (multiplicative + analytical)



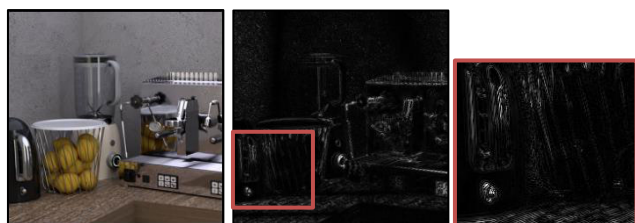
(c) Layer patterns (multiplicative + CNN-based)



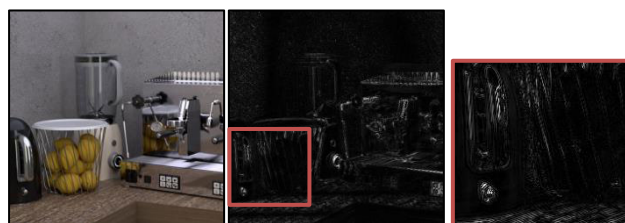
(d) Layer patterns (additive + analytical)



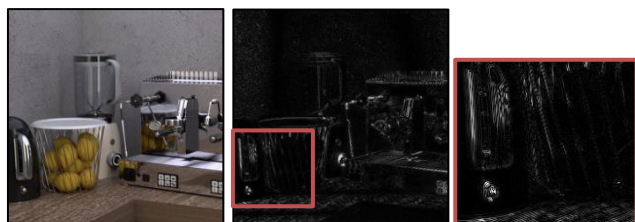
(e) Layer patterns (additive + CNN-based)



(f) Reproduced top-left view and error (multiplicative + analytical) PSNR = 33.95 dB, SSIM = 0.940



(g) Reproduced top-left view and error (multiplicative + CNN-based) PSNR= 33.65 dB, SSIM = 0.938



(h) Reproduced top-left view and error (additive + analytical) PSNR = 33.55 dB, SSIM = 0.938



(i) Reproduced top-left view and error (additive + CNN-based) PSNR = 33.42 dB, SSIM = 0.937

FIGURE 8. Visual results with *Kitchen*: (a) input light field and top-left view, (b)–(e) layer patterns for multiplicative or additive display optimized with analytical or CNN-based method, (f)–(i) reproduced top-left views and errors.

out using layers, multiplicative layers yielded better results than additive ones.

We also present visual results obtained with two datasets in Figs. 7 and 8. In each figure, we show input light fields and top-left views in (a), layer patterns obtained with the four combinations mentioned above in (b)–(e), reproduced top-left views and errors from the ground truth (magnified by 2 for better visualization) in (f)–(i). The number of iterations for the analytical methods were fixed to 200, which produced sufficiently converged results. Each of the reproduced views was obtained from the corresponding layer patterns by calculating each of the light rays produced through the stack of those layers. We also reported the PSNR and structural similarity (SSIM) values. The SSIM values reported here are the averages over all 5×5 multi-view images. The errors are perceivable mainly around the object edges. We can see that an additive display causes larger errors than a multiplicative one regardless of the optimization method. Please refer to the supplementary video for more details.

V. CONCLUSION

We compared the performance of layered light-field displays with different light-ray operations (multiplication and addition) conducted using layers and optimization methods (analytical and CNN-based methods) for the layer patterns. Our results indicate that multiplicative layers achieve better accuracy than additive ones, and CNN-based methods are faster than analytical ones, with comparable accuracy to the best achievable accuracy of analytical methods. We recommend that in terms of the balance between accuracy and computation speed, one should adopt multiplicative layers optimized using a CNN-based method. Our future work includes further improvement of the CNN structure for better accuracy with less computation time. We also need to increase the number of images in light fields to support wider viewing directions. Development of better display hardware³ is also necessary. These efforts will pave the way for high-quality light-field displays running at video-rate speeds.

REFERENCES

- [1] F. E. Ives, "Parallax stereogram and process of making same," U.S. Patent 725 567 A, Apr. 14, 1903.
- [2] G. Lippmann, "Épreuves réversibles donnant la sensation du relief," *J. Phys.*, vol. 7, no. 1, pp. 821–825, 1908.
- [3] T. Okoshi, "Three-dimensional imaging techniques," *Leonardo*, vol. 22, no. 3/4, p. 444, 1989.
- [4] S. Pastoor and M. Wöpking, "3-D displays: A review of current technologies," *Displays*, vol. 17, no. 2, pp. 100–110, Apr. 1997.
- [5] B. Javidi and F. Okano, *Three-Dimensional Television, Video, and Display Technologies*. Springer, 2002.
- [6] H. Isono, M. Yasuda, and H. Sasazawa, "Autostereoscopic 3-D display using LCD-generated parallax barrier," *Electron. Commun. Jpn.*, vol. 76, no. 7, pp. 77–84, 1993.
- [7] K. Sakamoto and T. Morii, "Multiview 3D display using parallax barrier combined with polarizer," *Proc. SPIE*, vol. 6399, Oct. 2006, Art. no. 63990R.
- [8] T. Peterka, R. L. Kooima, D. J. Sandin, A. Johnson, J. Leigh, and T. A. DeFanti, "Advances in the dynallax solid-state dynamic parallax barrier autostereoscopic visualization display system," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 3, pp. 487–499, May 2008.
- [9] R. Börner, "Autostereoscopic 3D-imaging by front and rear projection and on flat panel displays," *Displays*, vol. 14, no. 1, pp. 39–46, Jan. 1993.
- [10] M. McCormick, "Integral 3D imaging for broadcast," in *Proc. 2nd Int. Display Workshops*, vol. 3, 1995, pp. 77–80. [Online]. Available: <https://ci.nii.ac.jp/naid/10029343968/>
- [11] J. Arai, F. Okano, M. Kawakita, M. Okui, Y. Haino, M. Yoshimura, M. Furuya, and M. Sato, "Integral three-dimensional television using a 33-megapixel imaging system," *J. Display Technol.*, vol. 6, no. 10, pp. 422–430, Oct. 2010.
- [12] S. Suyama, H. Takada, and S. Ohtsuka, "A direct-vision 3-D display using a new depth-fusing perceptual phenomenon in 2-D displays with different depths," *IEICE Trans. Electron.*, vol. 85, no. 11, pp. 1911–1915, 2002.
- [13] D. Lanman, M. Hirsch, Y. Kim, and R. Raskar, "Content-adaptive parallax barriers: Optimizing dual-layer 3D displays using low-rank light field factorization," in *Proc. ACM SIGGRAPH Asia papers*, 2010, pp. 1–10.
- [14] D. Lanman, G. Wetzstein, M. Hirsch, W. Heidrich, and R. Raskar, "Polarization fields: Dynamic light field display using multi-layer LCDs," *ACM Trans. Graph.*, vol. 30, no. 186, pp. 1–10, 2011.
- [15] D. Lanman, G. Wetzstein, M. Hirsch, W. Heidrich, and R. Raskar, "Beyond parallax barriers: Applying formal optimization methods to multilayer automultiscopic displays," *Proc. SPIE*, vol. 8288, Feb. 2012, Art. no. 82880A.
- [16] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar, "Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting," *ACM Trans. Graph.*, vol. 31, no. 80, pp. 1–11, 2012.
- [17] K. Takahashi, Y. Kobayashi, and T. Fujii, "From focal stack to tensor light-field display," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4571–4584, Sep. 2018.
- [18] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. ACM SIGGRAPH*, 1996, pp. 31–42.
- [19] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The Luminograph," *Proc. ACM SIGGRAPH*, vol. 96, no. 30, pp. 43–54, 1996.
- [20] M. Hirsch, G. Wetzstein, and R. Raskar, "A compressive light field projection system," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1–12, Jul. 2014.
- [21] S. Lee, C. Jang, S. Moon, J. Cho, and B. Lee, "Additive light field displays: Realization of augmented reality with holographic optical elements," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–13, Jul. 2016.
- [22] D. Lanman, F. Heide, D. Reddy, J. Kautz, K. Pulli, and D. Luebke, "Cascaded displays: Spatiotemporal superresolution using offset pixel layers," *ACM Trans. Graph.*, vol. 33, no. 4, p. 60, 2014.
- [23] F.-C. Huang, K. Chen, and G. Wetzstein, "The light field stereoscope: Immersive computer graphics via factored near-eye light field displays with focus cues," *ACM Trans. Graph. (TOG)*, vol. 34, no. 60, pp. 1–12, 2015.
- [24] K. Maruyama, H. Kojima, K. Takahashi, and T. Fujii, "Implementation of table-top light-field display," in *Proc. Int. Display Workshops*, 2018, p. 1–6.
- [25] K. Maruyama, Y. Inagaki, K. Takahashi, T. Fujii, and H. Nagahara, "A 3-D display pipeline from coded-aperture camera to tensor light-field display through CNN," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 1064–1068.
- [26] K. Maruyama, K. Takahashi, T. Fujii, M. Date, and H. Kimata, "Comparative study on layered light-field displays and optimization methods," in *Proc. Int. Display Workshops*, 2019, p. 1–20.
- [27] MIT Media Lab's Camera Culture Group. (2018). *Compressive Light Field Camera*. [Online]. Available: <http://cameraculture.media.mit.edu/projects/compressive-light-fieldcamera/>
- [28] Computer Graphics Laboratory, Stanford University. (2018). *The (New) Stanford Light Field Archive*. [Online]. Available: <http://lightfield.stanford.edu>
- [29] Heidelberg Collaboratory for Image Processing. (2016). *Datasets and Benchmarks for Densely Sampled 4D Light Fields*. [Online]. Available: http://lightfieldgroup.iwr.uniheidelberg.de/?page_id=713
- [30] Heidelberg Collaboratory for Image Processing. (2018). *4D Light Field Dataset*. [Online]. Available: <http://hci-lightfield.iwr.uni-heidelberg.de/>
- [31] K. Takahashi. (2019). *Light Field Display Project*. [Online]. Available: <http://www.fujii.nuee.nagoyau.ac.jp/Research/LFDisplay/index.html>
- [32] Y. Kobayashi, S. Kondo, K. Takahashi, and T. Fujii, "[Paper] a 3-D display pipeline: Capture, factorize, and display the light field of a real 3-D scene," *ITE Trans. Media Technol. Appl.*, vol. 5, no. 3, pp. 88–95, 2017.

³We already developed a prototype with limited performance [32].



KEITA MARUYAMA received the B.E. degree in electrical engineering from Nagoya University, Japan, in 2018, where he is currently pursuing the degree with the Graduate School of Engineering. His research interests are in light-field acquisition and rendering for 3D displays.



KEITA TAKAHASHI (Member, IEEE) received the B.E., M.S., and Ph.D. degrees in information and communication engineering from The University of Tokyo, in 2001, 2003, and 2006, respectively. From 2006 to 2011, he was a Project Assistant Professor with The University of Tokyo. From 2011 to 2013, he was an Assistant Professor with The University of Electro-Communications. He is currently an Associate Professor with the Graduate School of Engineering, Nagoya University, Japan.

His research interests include computational photography, image-based rendering, and 3D displays.



TOSHIAKI FUJII (Member, IEEE) received the B.E., M.E., and Dr.E. degrees in electrical engineering from The University of Tokyo, in 1990, 1992, and 1995, respectively. From 2008 to 2010, he was with the Graduate School of Science and Engineering, Tokyo Institute of Technology. Since 1995, he has been with the Graduate School of Engineering, Nagoya University, where he is currently a Professor. His current research interests include multidimensional signal

processing, multicamera systems, multiview video coding and transmission, free-viewpoint television, and their applications.

• • •