Comparative Study on Layered Light-Field Displays and Optimization Methods

Keita Maruyama [1], Keita Takahashi [1], Toshiaki Fujii [1],
Munekazu Date [2], Hideaki Kimata [2]

[1] Graduate School of Engineering, Nagoya University, Japan
[2] NTT Media Intelligence Laboratories, NTT Corporation, Japan

Keywords: layered light-field display, light field, CNN

ABSTRACT

We focus on two factors that affect the performance of layered light-field displays: the layer device and optimization method. We quantitatively compared the performances of different architecture of layered light-field displays (LCD, HOE, and S-IPS LCD) and their optimization methods (analytical method and CNN-based method).

1. INTRODUCTION

Light-field displays have attracted much attention from researchers because of their ability to provide not only binocular depth sensation but also natural motion parallax in relation with head motion, which invokes strong feeling of immersion. To develop such a display, several approaches using parallax barriers [1], lenticular lenses [2], and rear projections [3, 4] have been proposed. We focus on a layered light-field display [5–9], which is composed of a stack of layer devices that are capable of controlling the pixel value in each pixel unit as shown in Fig. 1. With this structure, the appearance of the display changes in accordance with the observed direction, because the light rays emitted from a single point of the backlight pass through different combinations of pixels depending on the directions. Therefore, this display can express many viewpoints with only a few layer patterns. To calculate layer patterns, we prepare a light field (i.e., a set of multi-view images) as input. Next, layer patterns are optimized so that the light rays emitted from the display reproduce input light field.

In this paper, we focus on two factors of layered light-field displays: the layer device and optimization method. These factors affect the quality of the reproduced light field and computation time. Therefore, we quantitatively compared the performances of different architecture of layered light-field displays (LCD, HOE, and S-IPS LCD) and their optimization methods (analytical method and CNN-based method).

2. LAYERED LIGHT FIELF DISPLAY

In this chapter, we first describe the operation of three layer devices. Next, we introduce two layer pattern optimization methods. A light field is defined in a 4-D space, but in this chapter, we represent them in a 2-D subspace for simplicity. Furthermore, in this chapter, we set the number of layers to 2, but we can easily extend it to more than 2.

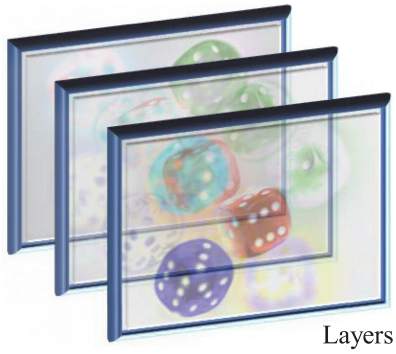

Fig. 1: Structure of layered light-field display.

2.1 Layer Devices

We have three candidates for the layer devices, which carry out different operations for the light rays as shown in Fig. 2.

The light rays modulated by the stacked liquid crystal panels (LCDs) are described as

$$l_{mul}(s,u) = a(u+s)b(u) \tag{1}$$

where, $a(u)$ and $b(u)$ denote the transmittance of the front and rear layers respectively. The outgoing direction is represented as s. In the case of LCD, the output light rays are attenuated significantly as the number of stacked layers increases, because two polarizers are attached to the both surfaces of each panel.

Next, in the case of holographic optical element (HOE), the light rays generated by the stacked layers are described as

$$l_{add}(s,u) = a(u+s)+b(u). \tag{2}$$

In this configuration, each layer requires a projector that provides the corresponding pattern.

Another candidate for the layer implementation is super-in plane switching (S-IPS) LCD. The operation of light rays carried out by them has been clarified [10] and is represented as

$$l_{S-IPS}(s,u) = a(u+s)+b(u)-2a(u+s)b(u). \tag{3}$$

In the case of S-IPS LCD, the polarizers are attached only to the outermost surfaces regardless of the number of stacked layers. Therefore, we can reduce the attenu-

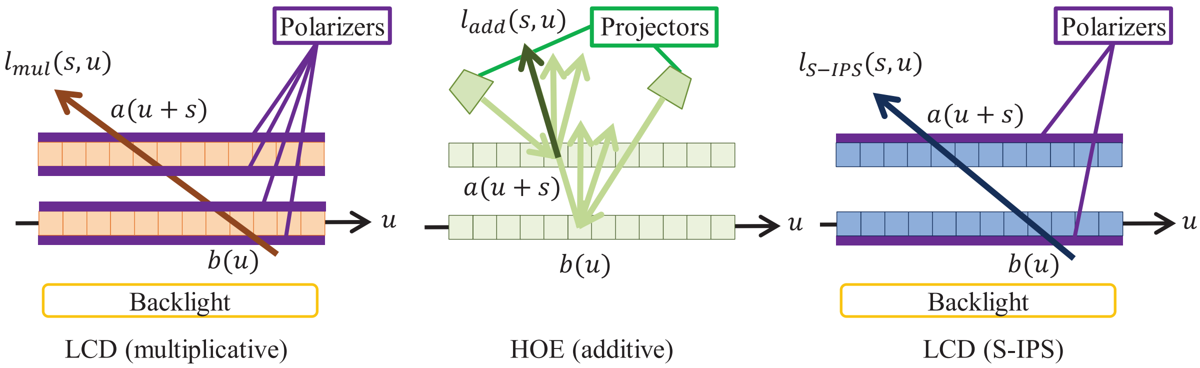

Fig. 2: Light rays output from each of the stacked layers.

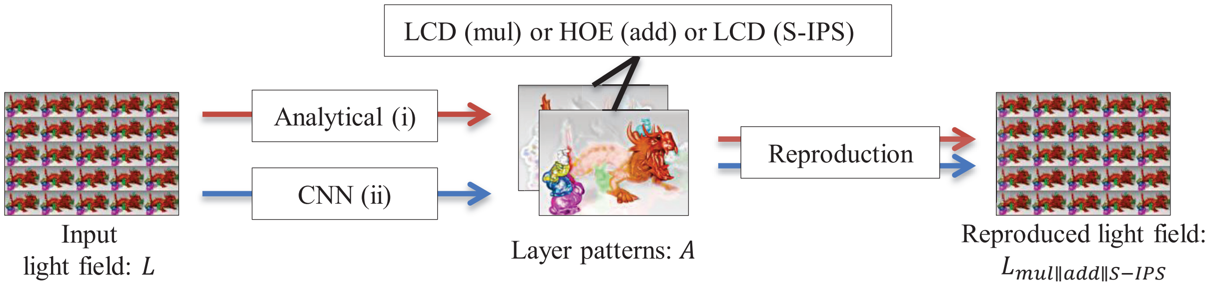

Fig. 3: Diagram of different layer devices and optimizations.

ation of output light rays even if the number of stacked layers increases. In previous work, the display quality with LCDs [5–8] or HOEs [9] has already been investigated. Meanwhile, the display quality with S-IPS LCDs has not been evaluated.

2.2 Optimization Methods

As for the optimization method, we have two candidates as shown in Fig. 3. Formally, the optimizations are given as:

$$\arg\min_{a,b} \sum_{s,u} ||l(s,u) - l_{mul\|add\|S-IPS}(s,u)||^2 \tag{4}$$

where $l(s,u)$ denotes the input light field.

Conventionally, these optimizations have been solved using analytical methods (Fig. 3 (i)) [5–7,9]. However, in this approach, the computation is slow because each layer should be alternatively updated with many iterations.

Another method is to use a convolutional neural network (CNN) to optimize layer patterns (Fig. 3 (ii)). The optimization process can be written in a form of mapping as

$$f : L \rightarrow A \tag{5}$$

where L represent a tensor that contains all the pixels of $l(s,u)$ for all s. Similarly, A represents a tensor that contains all the pixels of $a(u)$ and $b(u)$. The mapping from the layer patterns to the light field (Eqs. (1), (2), and (3)) is written as

$$g : A \rightarrow L_{mul\|add\|S-IPS} \tag{6}$$

where L_{mul}, L_{add}, and L_{S-IPS} are tensors that contain all the pixels of $l_{mul}(s,u)$, $l_{add}(s,u)$, and $l_{S-IPS}(s,u)$ for all s respectively. We constructed a CNN that corresponds to a composite mapping $g \circ f$ and minimized the squared error loss given as

$$\arg\min_{f} ||L - L_{mul\|add\|S-IPS}||^2 \tag{7}$$

over a massive amount of training samples. In our previous work [11], we applied a CNN-based optimization to obtain the layer patterns from a compressively sampled light field. However, we have not compared the display quality obtained with the three layer devices mentioned in 2.1.

3. EXPERIMENTS

In this experiment, the number of viewspoints in the target light field and the number of layers were set to 25 and 3, respectively. First, we optimized layer patterns using analytical methods for three layer devices. In the case of LCD and HOE, we optimized layer patterns based on [5–7, 9]. In the case of S-IPS LCD, we optimized them by extending the multiplicative update rule used in [5–7]. We implemented these optimizations with an open-source matrix library CuPy to make a fair comparison with the CNN-based optimization. Each layer was updated for 50 times for all layer devices. Next, we describe our CNN-based optimization. We implemented this optimization by modifying the framework developed in [11], as illustrated in Fig. 4. Thoroughout the network, the size of images is unchanged, but the number of channels is changed. Tensors L and $L_{mul\|add\|S-IPS}$ have 25

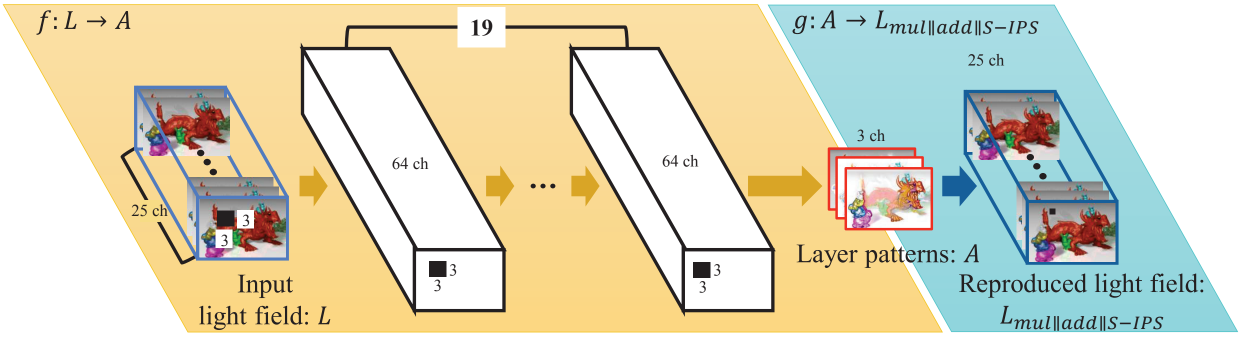

Fig. 4: Network architecture for optimizing layer patterns

channels, each of which corresponds to a viewpoint. Tensor A has 3 channels corresponding to the 3 layer patterns of the display. We trained this network with 295,200 samples according to our previous work [11].

We compared three layer devices (LCD, HOE, and S-IPS LCD) and two optimization methods in terms of the quality of the reproduced light field and computation time. Shown in Fig. 5 is the reproduction quality for the three datasets, Dragon and Bunnies (a), Dino (b), and Medieval2 (c), which were not included in the training data for the CNN-based method. In the case of LCD, the CNN-based method (ii) achieved the same quality as the analytical method (i). Meanwhile, the CNN-based method outperformed the analytical method in HOE, but the results was reversed in S-IPS LCD. As for the layer devices, the quality was better in the order of LCD, HOE, and S-IPS LCD. Next, the computation time are shown in Fig. 6. The CNN-based method (ii) was much faster than the analytical method (i) regardless of the layer device. As for the layer devices, S-IPS LCD was slower than the other two due to the complexity of the operation. Figure 7 shows simulated displayed images with different layer devices and optimizations. LCD had good performance regardless of the optimization method. In the case of HOE, the CNN-based method caused degradation of the displayed image. For S-IPS LCD, the object behind was blurred compared to the other two layer devices.

4. CONCLUSION

In this paper, we quantitatively compared the performances of different architecture of layered light-field displays (LCD, HOE, and S-IPS LCD) and their optimization methods (analytical method and CNN-based method). In the case of LCD and HOE, the CNN-based optimization method was superior to the analytical method in terms of the quality of reproduced light field and computation time. However, the analytical method achieved better reproduction quality in S-IPS LCD. As for the layer devices, LCD achieved better performance than the other two layer devices. For future work, we will further improve the reproduction quality of the light field and speed up the computation time by devising the network structure in the CNN-based method.

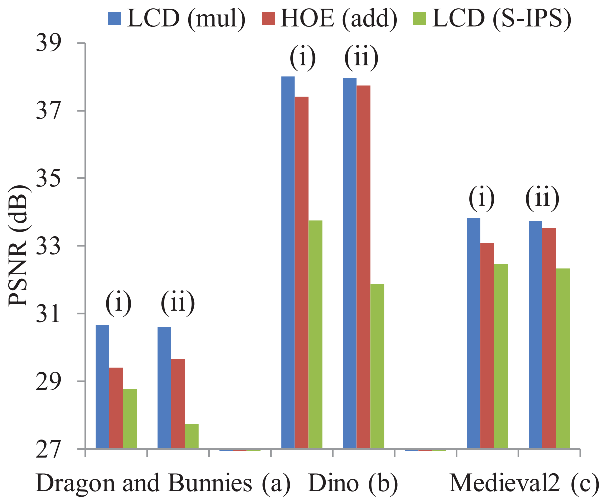

Fig. 5: Quantitative quality of light fields reproduced with analytical (i) and CNN-based (ii) methods.

Method	Layer device	(a)	(b)	(c)
Analytical (i)	LCD (mul)	3.538	1.892	1.905
	HOE (add)	0.857	0.524	0.476
	LCD (S-IPS)	9.146	5.056	5.136
CNN (ii)	LCD (mul)	0.281	0.138	0.137
	HOE (add)	0.278	0.136	0.137
	LCD (S-IPS)	0.308	0.151	0.151

Computation time (s)

Fig. 6: Computation time.

References

[1] F. E. Ives, "Parallax stereogram and process of making same.," *U.S. Patent US725567A*, 1903.

[2] G. Lippmann, "Epreuves reversibles donnant la sensation du relief," *Journal of Physics*, vol. 7, no. 4, pp. 821–825, 1908.

[3] W. Matusik, H. Pfister, "3D TV: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes," *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 814–824, 2004.

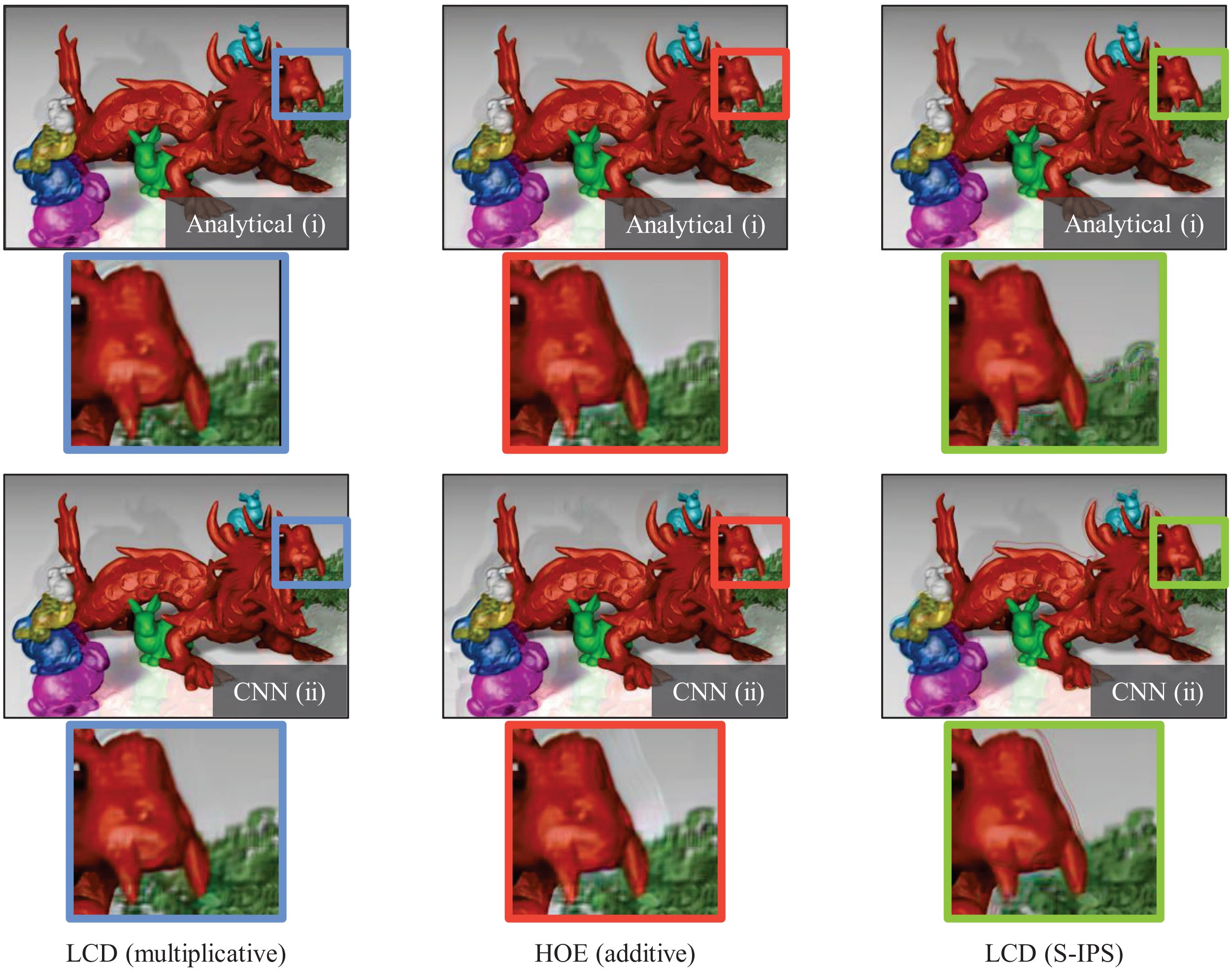

Fig. 7: Display simulations with different layer devices and optimizations.

[4] Y. Takaki, N. Nago, "Multi-projection of lenticular displays to construct a 256-view super multi-view display," *Optics express*, vol. 18, no. 9, pp. 8824–8835, 2010.

[5] G. Wetzstein *et al.*, "Layered 3D: tomographic image synthesis for attenuation-based light field and high dynamic range displays," *ACM Transactions on Graphics (ToG)*, vol. 30, no. 95, 2011.

[6] D. Lanman *et al.*, "Polarization fields: dynamic light field display using multi-layer LCDs," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 186, 2011.

[7] G. Wetzstein *et al.*, "Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting," *ACM Transactions on Graphics (TOG)*, vol. 31, no. 80, 2012.

[8] K. Takahashi, Y. Kobayashi, T. Fujii, "From focal stack to tensor light-field display," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4571–4584, 2018.

[9] S. Lee *et al.*, "Additive light field displays: realization of augmented reality with holographic optical elements," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 60, 2016.

[10] M. Date *et al.*, "Luminance addition of a stack of multidomain liquid-crystal displays and capability for depth-fused three-dimensional display application," *Applied Optics*, vol. 44, no. 6, pp. 898–905, 2005.

[11] K. Maruyama *et al.*, "A 3-D display pipeline from coded-aperture camera to tensor light-field display through CNN," *IEEE International Conference on Image Processing (ICIP)*, 2019.