Displaying Multiple 3D Scenes with a Single Layered Display

Chisaki Sato, Chihiro Tsutake, Keita Takahashi, Toshiaki Fujii

Corresponding Author: Chisaki Sato (c.sato@fujii.nuee.nagoya-u.ac.jp)
Department of Information and Communication Engineering, Graduate School of Engineering, Nagoya University, Japan
Keywords: 3D display, layered display, light field

ABSTRACT

We propose a method of displaying two different 3D scenes on a single layered light-field display, where the layer patterns are optimized for the two scenes simultaneously. We demonstrate that both scenes can be displayed in high quality when the viewing zones for them are separated sufficiently.

1 Introduction

It is convenient for us if a single display device can show different 3D images towards different viewing zones. For example, in the case of an in-vehicle display shown in Fig. 1, 3D drive navigation should be shown to the driver while another 3D entertainment content could be presented for the other passengers. To this end, we investigate a method of displaying two different light fields (two sets of multi-view images) towards two different viewing zones using a single layered light-field display.

A layered light-field display [1, 2] consists of several LCD panels stacked in front of a backlight. The patterns (layer patterns) displayed on the LCD panels are computed from a target 3D scene. The display has the capability to display dozens of different viewpoint images simultaneously towards different viewing directions. Therefore, not only binocular parallax but also motion parallax caused by the head movement can be perceived by the observers.

The layered displays have an advantage over several other displays with respect to the viewpoint-resolution trade-off. Glasses-free displays with lenticular lens [3] and parallax barriers [4] suffer from the obvious trade-off between "number of viewpoints to be displayed" and the "number of pixels per viewpoints", because a single display panel should be divided among the viewpoints. Meanwhile, in the case of a layered display, an increase in the number of views does not immediately decrease the resolution for each viewpoint image, but it only causes moderate quality degradation. Although the viewpoint-resolution trade-off can be overcome by the displays with rear projections [5], the projection optics involves a significantly large form factor in contrast to the thin structure of a layered display.

We use a layered light field display to show not a single but two distinct scenes simultaneously, which is a challenging task. Conventionally, a single light field, a set of multi-view images of a single target scene, is given as the input,

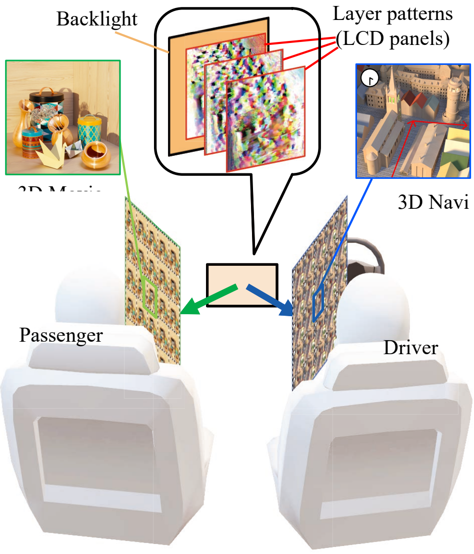

Fig. 1: Application scenario of our method.

and the layer patterns are optimized so as to reproduce the target scene accurately. Only a few layers are sufficient to reproduce the light field, because the images constituting the light field are redundant; these images are very similar with each other because they capture the same target scene from slightly different viewpoints. In other words, the layer patterns can be interpreted as a compressed representation for these images, where the compression capability depends on the redundancy in the target light field. However, in our case, the layer patterns are used to show two distinct light fields, which has no redundancy or coherence between them. To enhance the representation capability of the layers, we adopt up to four-fold time-division multiplexing. Moreover, we found that the display quality was significantly affected by the distance between the two viewing zones for the two light fields. We experimentally show that under appropriate design conditions, the layer patterns can accommodate two distinct light fields simultaneously, and can display both of them with high quality.

2 Layered Light-Field Display

2.1 Display Principle

A layered light-field display consists of several LCD panels stacked in front of a backlight and can display different

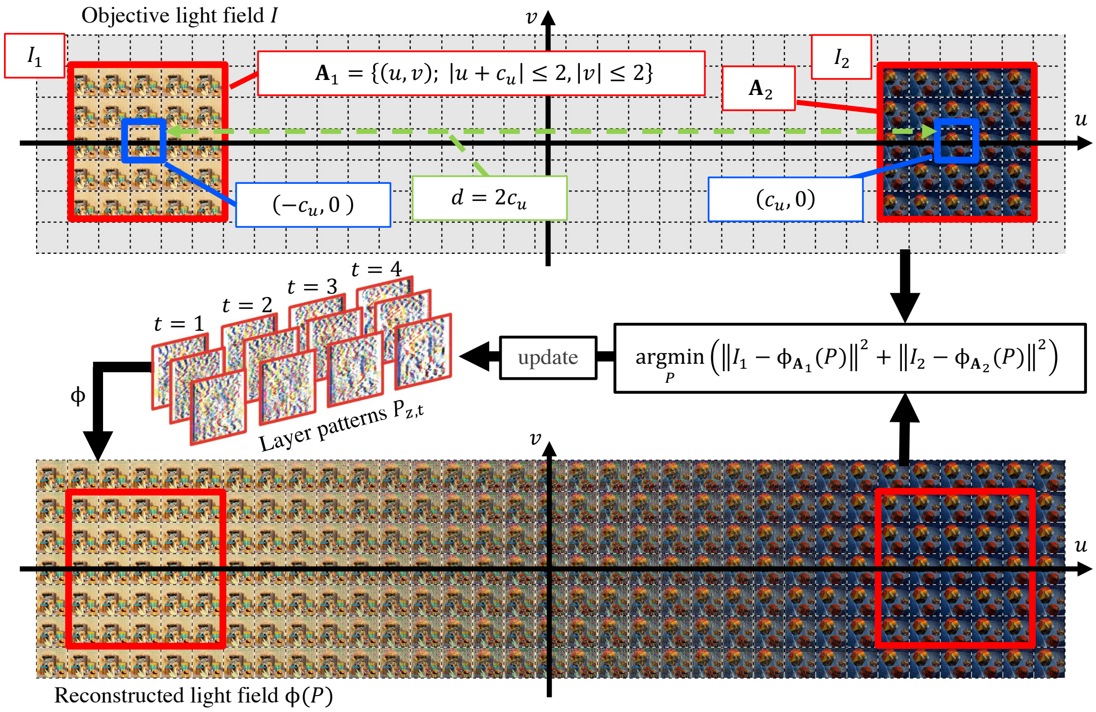

Fig. 2: Overview of our method.

images towards different viewing directions. A light field L emitted from the display is described as the multiplication of the layer patterns:

$$L_{u,v}(x,y) = \prod_{z \in Z} P_z(x + uz, y + vz) \tag{1}$$

where (x, y) and (u, v) denote the pixel and viewpoint coordinates in the light field. $P_z(x, y)$ denotes the layer pattern (transmittance) placed at depth z. In this paper, the number of layers is assumed to be three and $Z = \{-1, 0, 1\}$. In some cases, time-division multiplexing is used to extend the layers' representation capability. When T sets of layer patterns $P_{z,t} (t \in [1, T])$ are displayed repeatedly at a high speed, the light field generated by the display is written as

$$L_{u,v}(x,y) = \frac{1}{T} \sum_{t=1}^{T} \prod_{z \in Z} P_{z,t}(x + uz, y + vz). \tag{2}$$

We adopt up to four-fold ($T = 4$) time multiplexing.

2.2 Optimizing Layer Patterns

The layer patterns should be optimized for a target scene so as to reproduce the target light field as accurately as possible. Let $I = \{L_{u,v}(x, y)\}_{u,v,x,y}$ denote the target light field. The operation of the right-hand side of Eqs. (1) or (2) is denoted as Φ. The optimization of the layer patterns is formulated as

$$P^* = \arg\min_{P} ||I - \Phi(P)||^2. \tag{3}$$

This minimization can be solved by two methods: an iterative method derived from non-negative tensor factorization (NTF) [1] and a learning-based method constructed on deep neural networks [2]. We adopt the former method in this paper, but extend it to accommodate two distinct light fields as the target simultaneously.

3 Displaying Two Distinct Scenes

We use a single layered display to show two distinct 3D scenes. As shown in Fig. 1, two sets of different light fields are displayed towards the respective viewing zones from the same layer patterns. An overview of our method is illustrated in Fig. 2.

For the target light field I, the viewpoint coordinate (u, v) is defined over a wide range of viewpoints, but only the viewpoints in the two viewing zones, surrounded by the red boxes in the figure, are treated as the target viewpoints. We denote these zones as $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively. These zones cover the areas of 5×5 viewpoints and are located symmetrically with respect to $(u, v) = (0, 0)$, centered at $(-c_u, 0)$ and $(c_u, 0)$, respectively.

$$\mathbf{A}_1 = \{(u,v);\ |u + c_u| \leq 2, |v| \leq 2\}, \tag{4}$$

$$\mathbf{A}_2 = \{(u,v);\ |u - c_u| \leq 2, |v| \leq 2\} \tag{5}$$

The distance between two viewing zones is defined as $d = 2c_u$. The target light fields for $\mathbf{A}_1$ and $\mathbf{A}_2$ are denoted as I_1 and I_2, respectively, each of which includes 5×5 viewpoint images. Note again that I_1 and I_2 are unrelated to each other (e.g., 3-D navigation and 3-D movie content).

The objective function for layer pattern optimization is written as

$$P^* = \arg\min_P (||I_1 - \Phi_{\mathbf{A}_1}(P)||^2 + ||I_2 - \Phi_{\mathbf{A}_2}(P)||^2). \quad (6)$$

Here, $\Phi_{\mathbf{A}_1}$ and $\Phi_{\mathbf{A}_2}$ denotes the process of Eq. (2), where the light fields are generated for the viewing zones, $\Phi_{\mathbf{A}_1}$ and $\Phi_{\mathbf{A}_2}$, respectively, from the same set of layer patterns P. We use the NTF-based iterative method to solve this optimization.

It should be noted the display is not designed to "black-out" the viewpoints outside the two viewing zones; we can see something from the viewpoints not included in $\mathbf{A}_1$ and $\mathbf{A}_2$, as shown in the bottom of Fig. 2. Interestingly, we can observe a cross dissolve effect over the viewpoints between $\mathbf{A}_1$ and $\mathbf{A}_2$; as the viewpoint moves from $\mathbf{A}_1$ to $\mathbf{A}_2$, the image generated from the display gradually changes from the first scene to the second.

4 Experiments

We evaluated the effect of the distance between two viewing zones on the quality of the displayed images. The locations of the viewing zones were varied with $3 \leq c_u \leq 20$, resulting in the distance between them $6 \leq d \leq 40$. Two light fields (Origami: scene 1, Platonic: scene 2) in the HCI dataset [6] were used as the targets. The degree of time-division multiplex was set to $T = 1, 2, 4$, and the number of iterations for layer pattern optimization was set to 500. The displayed images were computational generated using Eq. (2) from the optimized layer patterns, and compared against the target light fields for quantitative evaluation.

Figure 3 depicts the relation between the distance between the viewing zones, d, and the quality (PSNR) of the displayed images. As expected, a larger T led to better quality of the displayed images. Moreover, the quality significantly improved as the distance d increased, which can be explained as below. When the distance is sufficiently large, the displayed image can change gradually from the first scene to the second as the viewpoint moves from $\mathbf{A}_1$ to $\mathbf{A}_2$. Meanwhile, if the distance is small, this change should happen more abruptly. Such an abrupt change is difficult for the layer patterns to reproduce, which leads to insufficient quality of displayed images. In other words, rapid changes along the viewpoints correspond to the high-frequency components over the viewpoints, which are difficult to be represented compressively by the layer patterns. To conclude, a layered display with time-multiplexing can accommodate two distinct light fields with high quality, if the two viewing zones are separated sufficiently.

We also present several visual results. Figure 4 shows the layer patterns obtained with $c_u = 20$ and $T = 4$. They seems to be meaningless patterns not resembling the scenes to be displayed, but the two scenes were successfully generated from them. Figure 5 visualizes several displayed images on the same condition, accompanied by the viewpoint-wise

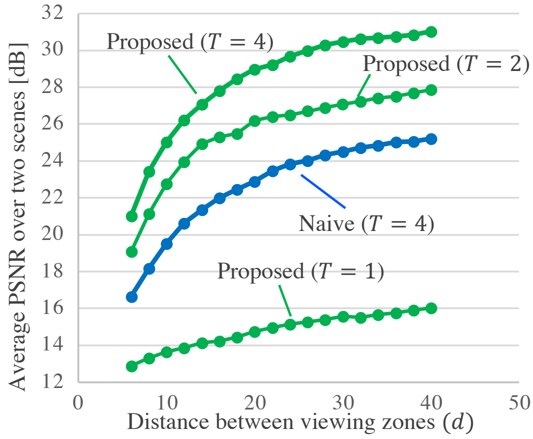

Fig. 3: Display quality (PSNR) against the distance between viewing zones (d).

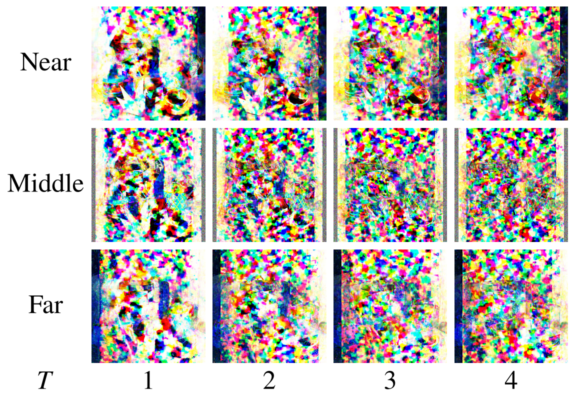

Fig. 4: Layer patterns with $c_u = 20$ and $T = 4$.

and mean PSNR values. In both of the viewing zones, we achieved sufficient quality for the displayed images.

For further validation, we finally mention a naive method with $T = 4$; the four temporal frames were divided evenly for the two distinct scenes. More specifically, the first two frames are optimized so as to display I_1 for $\mathbf{A}_1$ and the black images for $\mathbf{A}_2$, while the latter two frames are to display I_2 for $\mathbf{A}_2$ and the black images for $\mathbf{A}_1$. As shown in Figs. 3 and 5, the performance of the naive method was quite limited. Our method achieved much better quality because the four temporal frames were used in not a divisive but an united manner to display the two distinct light fields.

5 Conclusion

We propose a method of displaying two different light fields towards two different viewing zones using a single layered light-field display. We experimentally show that the layer patterns can accommodate two distinct light fields simultaneously and can display both of them with high quality when the distance between the viewing zones is sufficiently large. Our future work will include hardware implementation of our method.

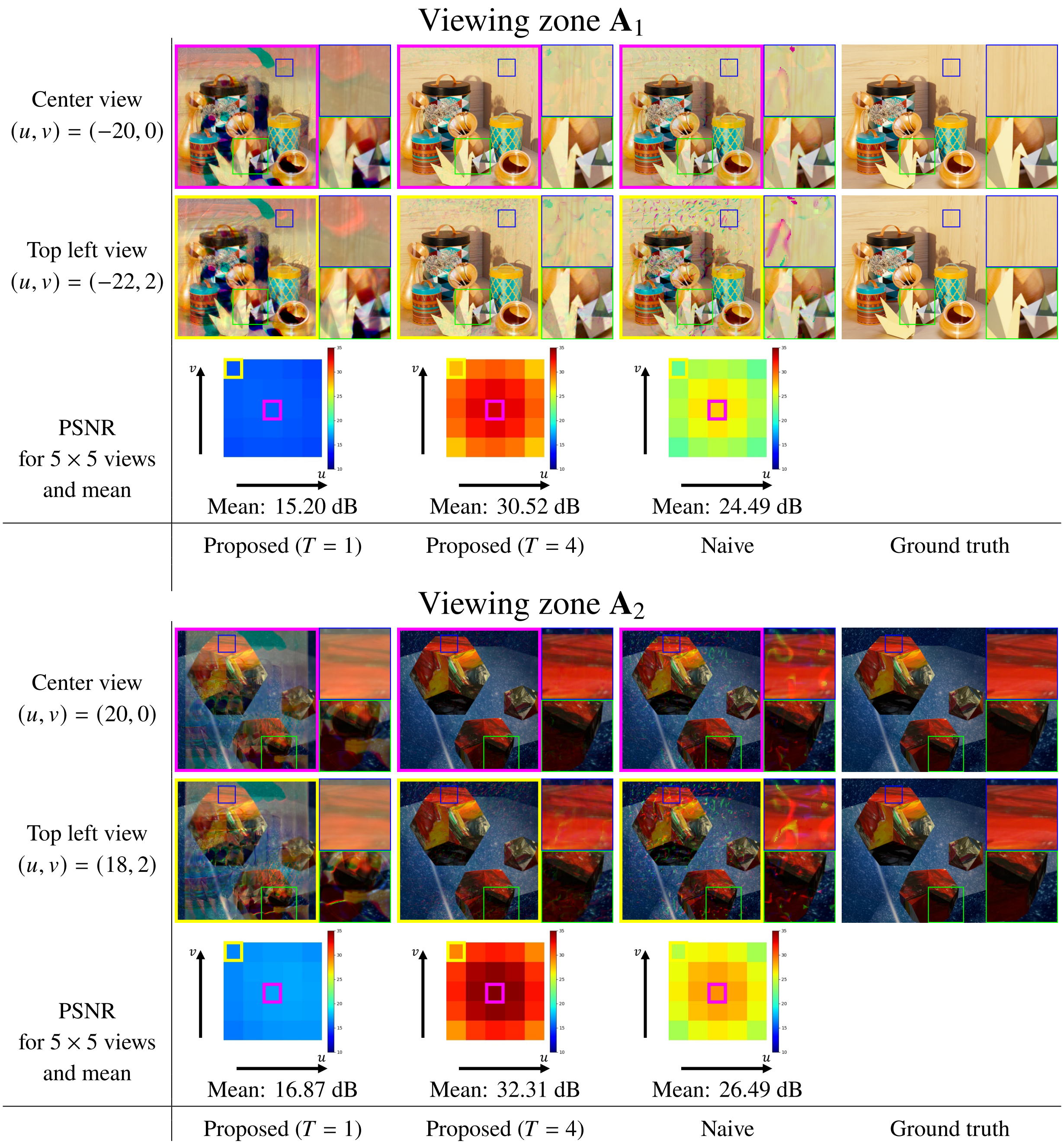

Fig. 5: Displayed results with $c_u = 20$.

References

[1] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar, "Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting," *ACM Trans. Graph. (Proc. SIGGRAPH)*, vol. 31, no. 4, pp. 1–11, 2012.

[2] K. Maruyama, K. Takahashi, and T. Fujii, "Comparison of layer operations and optimization methods for light field display," *IEEE Access*, vol. 8, pp. 38 767–38 775, 2020.

[3] 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," *Journal of Display Technology*, vol. 6, no. 10, pp. 422–430, 2010.

[4] 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 TVCG*, vol. 14, no. 3, pp. 487–499, 2008.

[5] S. Iwasawa and M. Kawakita, "Behind the screen of NICT's 200-in. automultiscopic display," *International Display Workshop*, pp. 1903–1096, 2012.

[6] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, "A dataset and evaluation methodology for depth estimation on 4D light fields," *ACCV*, 2016.