A Pipeline from Raytrix to Tensor Display

Wenzhe Ouyang*, Shu Fujita*, Keita Takahashi*, and Toshiaki Fujii*
*Department of Information and Communication Engineering, Nagoya University, Japan

***Abstract*—A tensor display, which consists of a few light attenuating layers stacked in front of a backlight, has potential to visualize a light field with high quality. A Raytrix camera is a focused plenoptic camera and has many advantages, such as higher spatial resolution and ability to capture light field videos, compared to traditional plenoptic cameras. In this paper, we established a process pipeline from capturing data by a Raytrix cameara to displaying it on a tensor display. To demonstrate the effectiveness of our pipeline, we conducted both computer simulations and experiments on a prototype display.**

I. INTRODUCTION

Light fields (LF), which contains both spatial and angular information, provide a rich representation of real-world scenes. Technically, LFs are often interpreted as a set of dense multi-view images, and a light field display should be capable of emitting these multi-views simultaneously. Thus, a light field display can bring better stereoscopic experience compared to traditional binocular displays which only display two views separately to the left and right eyes. To develop such a display, several methods have been proposed, such as the ones using parallax barrier [1], lenticular lenses [2] and rear projection [3][4]. Among them, a tensor display, as an emerging and promising technology, was introduced by [5]. This display is composed of a few liquid crystal displays (LCD) stacked in front of a backlight. By controlling the transmittance of each layer, this configuration can display a view-dependent images that correspond to the target 3-D object, as shown in Fig 1.

In order to display real-world scenes using a tensor display, a process pipeline from acquisition to display was proposed in [6], in which a Lytro Illum camera [7] and a multi-view camera were used to capture light field data. Meanwhile, in this paper, we presented a process pipeline from a Raytrix camera [8]. A Raytrix camera, is also a plenoptic camera, but has a little structural difference from the previous Lytro Illum cameras. Thanks to this difference, a Raytrix camera can capture light fields with higher spatial resolution. In addition, it can also capture light field videos. Based on several previous works [9][10], we completed the pipeline that can convert input from a Raytrix camera into layer patterns for a tensor display. To demonstrate the effectiveness of our pipeline, we conducted both computer simulations and experiments on a prototype display.

II. HARDWARE AND PROCESS PIPELINE

In this section, we introduce two most important hardware used in our experiments, input and output devices. Then, we introduce the algorithm used in our pipeline.

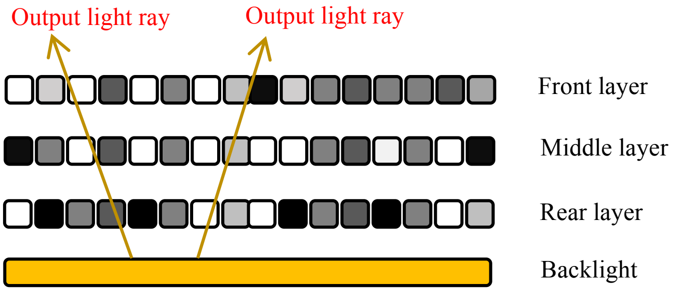
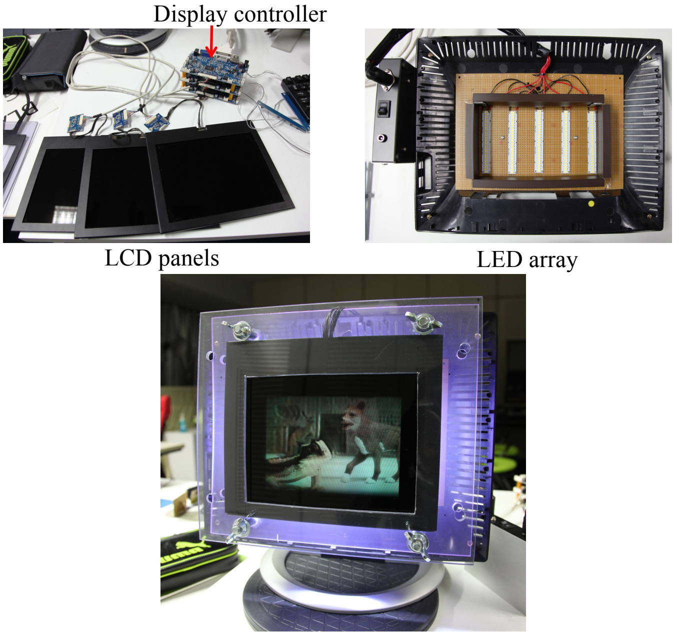


Fig. 1: Tensor display with 3 layers



LCD panels
LED array

Fig. 2: Tensor display prototype [6]

A. Hardware

As for the input device, we used a Raytrix R5 camera to capture light field video contents at up to 180 fps. Specifically, the Raytrix R5 has 2048×2048 pixels; the number of microlens is 8,787 and each covers a hexagonal region whose diameter is 23 pixels. As for the output device, we used our own tensor display prototype, which has been presented in [6]. Figure 2 shows the hardware components and appearance of our display prototype. Our prototype consists of three semi-transparent LCD panels and a hand crafted backlight. Each layer of LCD has a resolution of 1024×768 pixels on a 9.7 inch panel and a refresh rate of 60 or 75 Hz.

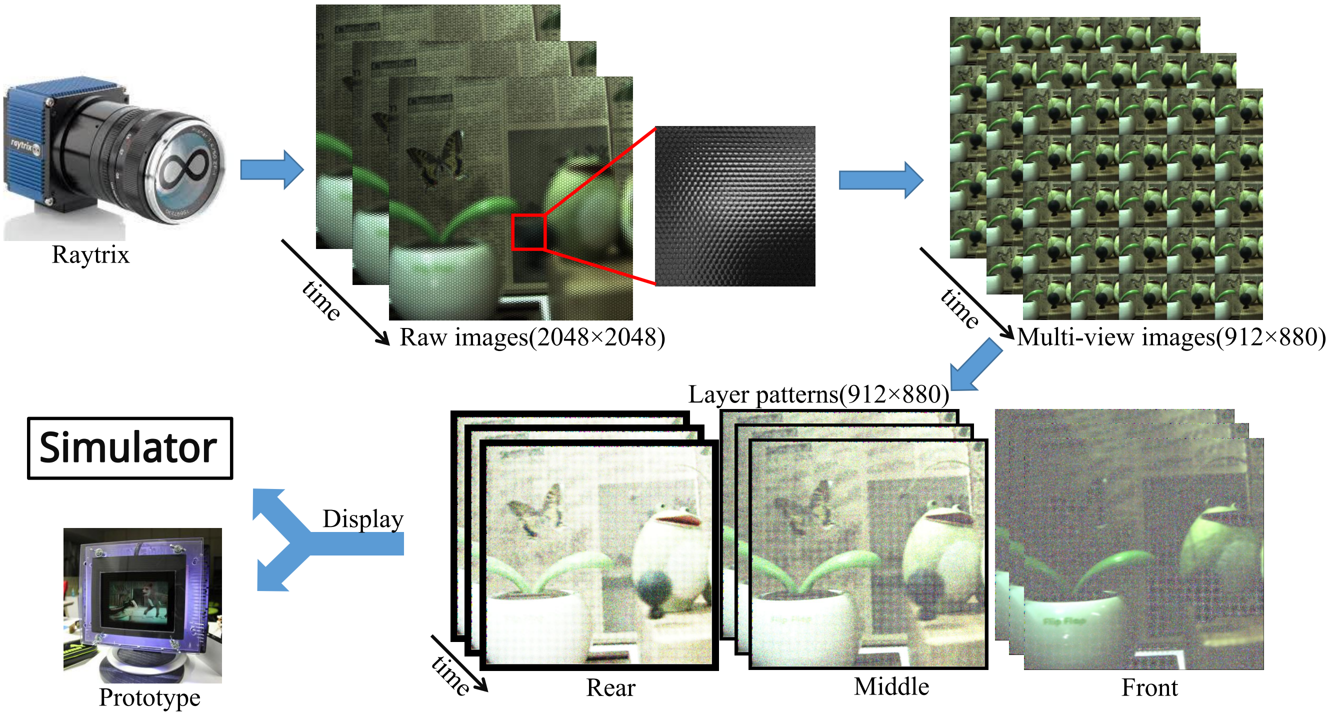


Fig. 3: Processing pipeline

B. Process Pipeline

Our pipeline can be divided into three parts: renderinging multi-view images from raw data taken by a Raytrix camera, generating layer patterns from multi-view images, and displaying the layer patterns in the tensor display, as shown in Fig. 3.

A Raytrix cameara is a multi-focused plenoptic camera, and it usually needs a sophisticated method for rendering multi-view images from raw data. In this study, we utilized a similar method with [9] to render multi-view images, in which patch size estimation (equivalent to depth estimation for each microlens) is conducted. In addition, we modify the convergence plane of the scene by adding and deleting boundaries to/from multi-view images.

As for generating layer patterns, we utilized the method proposed by [10], which introduced a generalized framework for light field factorization that can handle both the orthographic and perspective models. In this paper, our target is to display real-world 3-D contents, which is captured by a perspective camera, therefore we adopted the perspective model only.

After extracting multi-view images and calculating layer patterns, we have got a set of layer patterns of every frame. We can use those layer patterns as the input to computer simulation and our prototype display.

III. SIMULATION AND EXPERIMENTS

To demonstrate the effectiveness of our pipeline, we first introduce datasets captured by a Raytrix R5 camera. These datasets offer three different configurations, as shown in Tab. I. The target scene with Configuration 3 is shown in Fig. 4. Please notice that although these three configurations look similar, the contents are not synchronized. Every dataset has 100 frames with 30 fps. We extracted 5×5 multi-view images from raw data of every frame, and used them to calculate layer patterns using the method in [10].

We used a simulator developed in [11], which can change the viewpoint from the user interface directly. Furthermore, the simulator in [11] can change several parameters that influence how the 3-D image looks, such as the distance between the observers viewpoint and the display, the intervals between the layers, and the distance among the virtual viewpoints to which input images are assigned.

We modified this simulator to display several frames along time in sequence, which can display light field videos at up to 60 fps with a Intel Quad-core processor and 16GB RAM.

The simulation results with Configurations 1, 2 at the 10th frame are shown in Figs. 6, 7, where the left, middle, and right columns correspond to the View A, View B and View C marked in Fig. 5. Comparing the results of View A and View B, we can observe obvious stereoscopic effects, especially in

TABLE I: Parameter setting of Dataset

Object	Dataset		
	Configuration1	Configuration2	Configuration3
Tadpole	150mm	225mm	300mm
Leaves	200mm	300mm	400mm
Butterfly	300mm	450mm	600mm

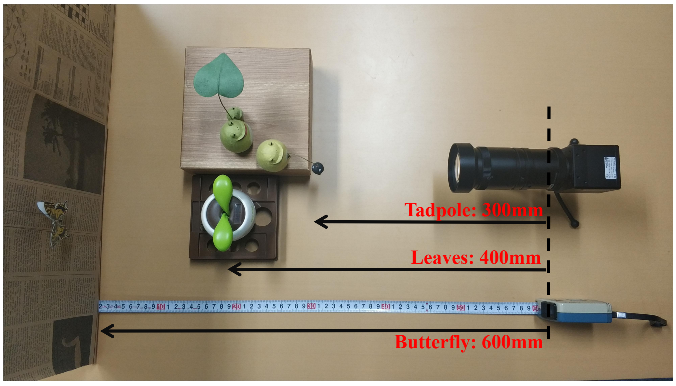


Fig. 4: Target scene from above

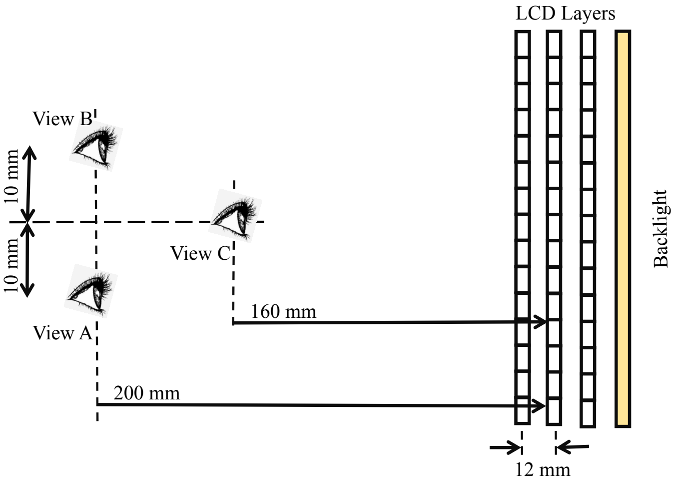


Fig. 5: Simulation setting with configuration 1

Configuration 1, which has the maximum range of disparities among the three configurations (please observe the frog's eyes, where the details are enlarged).

We conducted experiments on our prototype of tensor display [6]. The experimental results with Configuration 1 at the 10th frame are shown in Fig. 8. We perceived 3D sensation when we observed the display from 50 cm away.

IV. CONCLUSIONS

In this paper, we presented a pipeline from a Raytrix camera to a tensor display. In order to show the effectiveness of our pipeline, we first presented a computer simulation of a tensor display, and then we conducted experiments on our display prototype. We successfully visualized real 3D scenes in videos using our pipeline.

REFERENCES

[1] F. E. Ives, "Parallax stereogram and process of making same," *U.S. Patent US72556A*, 1903.

[2] G. Lippmann, "Épreuves réversibles donnant la sensation du relief," *Journal de Physique Théorique et Appliquée*, vol. 7, no. 1, pp. 821–825, 1908.

[3] W. Matusik and H. Pfister, "3D TV: A Scalable System for Real-Time Acquisition, Transmission and Autostereoscopic Display of Dynamic Scenes," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 814–824, 2004.

[4] Y. Takaki and 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, D. Lanman, M. Hirsch, and R. Raskar, "Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 1–11, 2012.

[6] Y. Kobayashi, S. Kondo, K. Takahashi, and T. Fujii, "A 3-D Display Pipeline: Capture, Factorize, and Display the Light Field of a Real 3-D Scene," *ITE Transactions on Media Technology and Applications*, vol. 5, no. 3, pp. 88–95, 2017.

[7] "Lytro," https://www.lytro.jp/.

[8] "Raytrix," https://www.raytrix.de/.

[9] S. Wanner, J. Fehr, and B. Jhne, "Generating EPI Representations of 4D Light Fields with a Single Lens Focused Plenoptic Camera," *Advances in Visual Computing: 7th International Symposium*, ISVC 2011, Las Vegas, NV, USA, September 26–28, 2011.

[10] S. Kondo, Y. Kobayashi, K.Takahashi, and T.Fujii, "Physically-Correct Light-Field Factorization for Perspective Images," *IEICE Transactions on Information and Systems*, vol. E100.D, no. 9, pp. 2052–2055, 2017.

[11] S. Mikawa, K. Takahashi, and T. Fujii, "Simulating a Stacked-Layer Light-Field Display Using Orthographic/Perspective View Models," *International Workshop on Advanced Image Technology*, 2018.

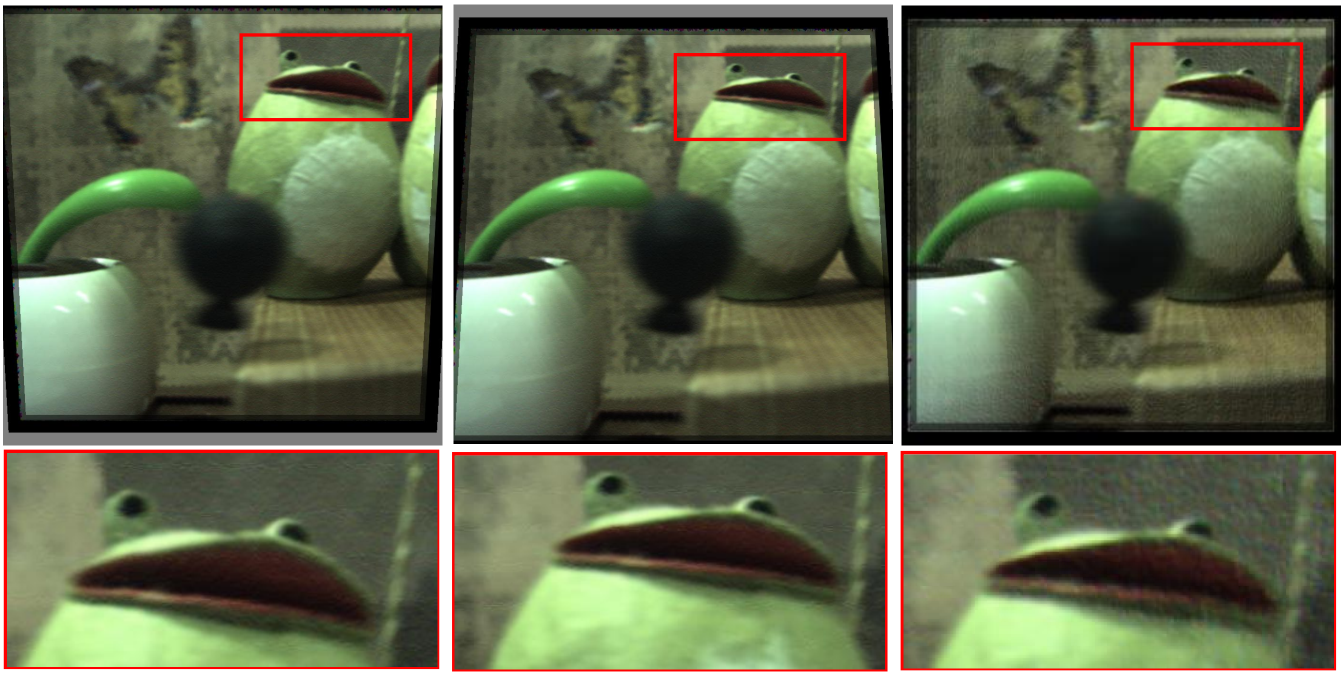


Fig. 6: Simulation results with Conf. 1 (the left, middle, and right are corresponding to View A, B, and C in Fig. 5, respectively)

Fig. 7: Simulation results with Conf. 2 (the left, middle, and right are corresponding to View A, B, and C in Fig. 5, respectively)

Fig. 8: Prototype experiment results with Conf. 1 (observe from the left, center, and right direction, respectively)