https://doi.org/10.2352/lim.2022.1.1.11

Towards non-Lambertian scenes for tensor displays

Eline Soetens*[1]*, Armand Losfeld*[1]*, Daniele Bonatto, Sarah Fachada, Laurie Van Bogaert, Gauthier Lafruit, Mehrdad Teratani
Laboratories of Image Synthesis and Analysis, Université Libre de Bruxelles, Brussels, Belgium

Abstract

Tensor displays are screens able to render a light field with correct depth perception without wearing glasses. Such devices have already been shown to be able to accurately render a scene composed of Lambertian objects. This paper presents the model and prototyping of a tensor display with three layers, using repurposed computer monitors, and extends the light field factorization method to non-Lambertian objects. Furthermore, we examine the relation and limitations between the depth-of-field and the depth range with Lambertian and non-Lambertian scenes. Non-Lambertian scenes contain out-of-range disparities that can not be properly rendered with the usual optimization method. We propose to artificially compress the disparity range of the scene by using two light fields focused on different depths, effectively solving the problem and allowing to render the scene clearly on both simulated and prototyped tensor display.

Introduction

A light field [1] encodes the distribution of light in any direction, keeping the parallax information intact. Rendering a scene using a light field is more immersive than using traditional 2D media and more realistic than 3D models. However, while light field capturing devices [2, 3, 4] evolved overtime, only limited displays are able to render them accurately. Notably, head-mounted displays [5] with view synthesis [6] allow a 6 degrees-of-freedom experience, but require the use of special glasses. Considering glass-free technologies, holographic stereograms [7] are printed to display only one static scene, while Super-MultiView light field displays such as the Holografika [8] are able to render dynamic light fields, but both solutions require very expensive materials for manufacturing. Finally, tensor displays [9, 10, 11] promise results close to the Holografika while being cheaper.

In this paper, we extend the tensor display model [9], a screen composed of n semi-transparent Liquid-crystal display (LCD) panels and usually a back-light, to render light fields of non-Lambertian content. Tensor displays encode the rays of a light field as combinations of the color each ray traverses over the n layers, allowing to view the scene from any viewpoint. Nagoya's university [10, 11] implemented a version of the original tensor display [9]. However, they focused their work on analyzing the required number of layers and a novel method to reconstruct the light field. Both papers show that the tensor display is limited by the number of layers and their gaps due to light intensity loss.

All previous works related to tensor display did not explicitly analyze the limitation of such displays for non-Lambertian content, neither on a simulated nor a prototyped display. Our main contributions are (1) Prototyping a tensor display with 3 layers using old LCD monitors, (2) Understanding the limitations for displaying non-Lambertian objects on tensor displays, and (3) Proposing a method to factorize non-Lambertian objects.

[1]The first two authors contributed equally.

Tensor Display Model

A tensor display allows one to render a light field without wearing glasses. It is defined as n aligned layers (LCD panels) and a homogeneous back-light [9]. This configuration allows to display oriented light rays as a color combination of the n layers. The whole light field can be encoded this way allowing a glass-free experience. By using the formalism of [9], a light ray $\tilde{l}(x,y,t)$ is expressed as

$$\tilde{l}(x+u,y+v,t) = \prod_{i=1}^{j} \beta_i \left(x + \frac{d_i}{t}u, y + \frac{d_i}{t}v \right) I(x,y) \tag{1}$$

with (u,v,t) describing the light rays in the light field according to where the light field is seen, β_i, the transmittance at a certain position in layer i and constrained to $[0,1]$, d_i, the z-position of layer i and $I(x,y)$ the initial intensity of the emitted light. Without loss of generality, $I(x,y)$ is set to 1 by considering the back-light emitting at maximal intensity. By considering $t \geq d_n$, the product ranges over all the layers, thus, $j = n$.

The tensor values are computed by a least square error minimization process with the target light field $l(x,y,z)$.

$$\beta_i^* = \arg\min_{\beta_i} \int\int (l(x',y',t) - \tilde{l}(x',y',t))^2 dx'dy' \tag{2}$$

Optimal values are obtained by considering each layer as an array of M transmittance adjustable pixels, then, equation (1) is expressed as an M elements order n tensor.

$$T = B_1 \otimes B_2 \otimes \ldots \otimes B_n \tag{3}$$

All combinations of the n layers are contained in this tensor, each representing trajectory of a light ray. However, most are physically wrong and do not represent a possible light ray. A correction tensor W is introduced to keep only valid configurations.

$$\tilde{L} = W \odot T \quad w = \begin{cases} 1 & \text{if possible configuration} \\ 0 & \text{otherwise} \end{cases} \quad w \in W \tag{4}$$

With $\odot$, the Hadamard product.

As LCDs are not able to display negative color values, the minimization is expressed as a Nonnegative Tensor Factorization (NTF) problem, which is linked to Nonnegative Matrix Factorization [12].

$$B_i^* = \underset{B_i}{\text{argmin}} \quad ||L - W \odot T||^2 \tag{5}$$

With an iterative update step:

$$B_i \leftarrow B_i \odot \left(\frac{(W_{(n)} \odot L_{(n)}) B_*^{(i)}}{(W_{(n)} \odot (B_i (B_*^{(i)})^T)) B_*^{(i)}} \right) \tag{6}$$

where $A_{(n)}$ the tensor matricization, $*$ is the Khatri-Rao product [13] and $B_*^{(i)}$ is defined by

$$B_*^{(i)} = B_n * \ldots * B_{i+1} * B_{i-1} * \ldots * B_1 \tag{7}$$

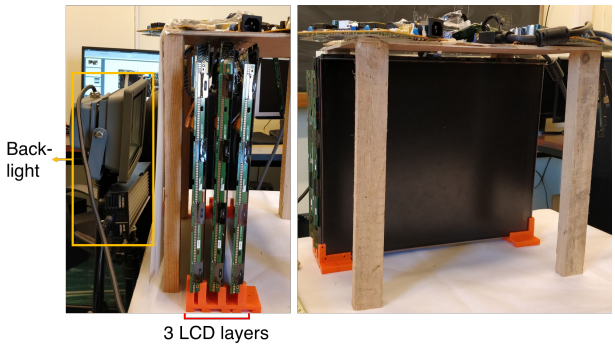

Figure 1. *Side view and 3/4 view of our tensor display. The layers are aligned using a 3D-printed piece (in red) and the back-light is centered behind the screen (in yellow).*

Nagoya University's implementation [10] allows to output each layer from a light field using the extracted multi-views, which are an array of slightly different views that can be acquired by an array of cameras or the use of a plenoptic camera [14]. This method differs from the model as: (1) each view corresponds to a fixed direction, and (2) the gap between layers is constant. As the field-of-view (FoV) is determined by the angle of the farthest view, it is linked to the first hypothesis. Moreover, the number and space of the layers also limits the quality of the device since the light intensity is impacted. In the model, the transmittance of a pixel is considered as perfect, while in practice, the number of layers increases the attenuation factor. A previous work has shown that a tensor display can be made up to five layers [11], but this number depends on the screens specifications and has to be adjusted with the quality of the screens.

The depth of field capacity of the tensor display has been studied for strictly Lambertian scenes. Wetzstein *et al.* [9] shows that augmenting the number of layers or using time multiplexing can increase the display depth range. In addition, the Nagoya implementation [10] focused on the property of a display made with 3 layers without time multiplexing, and shows a *depth-selectivity* as well as the expected depth range. However, these analyses were narrowed to Lambertian content.

Prototyped Tensor Display

We prototyped a 3-layers tensor display using three *Samsung SyncMaster 710v* (1280×1024 pixels with 75Hz as fresh rate) monitors (figure 1). The back-light is made of two lights fixed with a tripod, and thus is not entirely homogeneous, but still permits to use a 3 layers tensor display.

The gap between the layers is important as a narrow gap limits the depth of the scene while a large gap reduces the field-of-view. Furthermore, with large gaps, objects tend to be sharp in one layer corresponding to their physical depth and blurred in all others, converging to a multi-plane image. Non-uniform gaps are allowed by the model, however, we used a fixed gap for simplicity with the layers positions at 120, 145 and 170 mm from the back-light.

As each layer attenuates the traversing light, the maximal number of layers is usually set to 3. To improve the field-of-view or the available disparities, polarizing films are removed from all the layers except the first (for light polarization) and the last (orientation blocking) polarizing films.

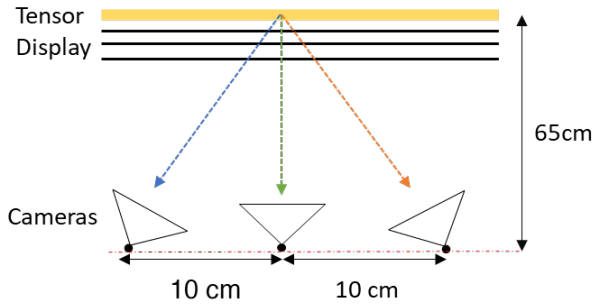

Figure 2. *Position of the camera for capturing displayed light field*

Analysis and approaches

Tensor display for Lambertian content

First, we prove that our prototype is able to display a light field with only Lambertian objects and visually compare three simulated and acquired views with what it is possible to see on the tensor display. Thus, it is important to capture only the rays associated with one of the three views. The capturing cameras was placed as in figure 2. To compare the original views with the displayed ones, two datasets are used: (1) A light field simulated 3D scene with simple objects at various depth values and rendered as 5×5 views, it is an open-source simulated light field dataset [15] used in [9] (cf. figure 3); (2) A cameras acquired scene captured as 17×17 views, coming from an open-source light field dataset published by the Stanford University [16] (cf. figure 4).

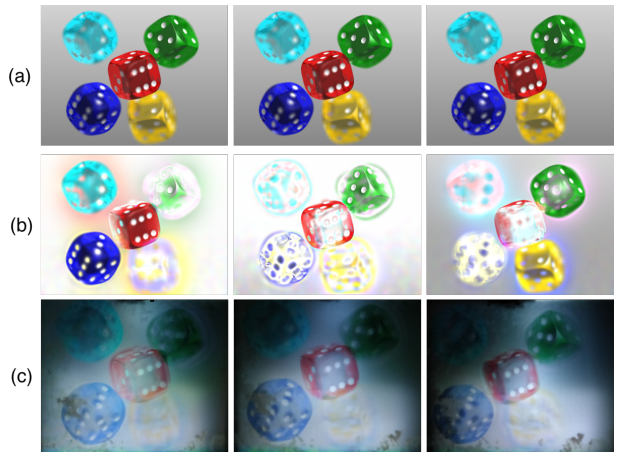

Figure 3. *Simulated light field (a) Original Views. (b) Layered images computed with multi-view. (c) Photographed views on the tensor display. Note that, on the borders of the captured views, artifacts are displayed. It comes from the glue that was fixing, before, the polarizing films.*

Due to low-quality screens and the non-homogeneous back-light, the quality of the displayed images decreases while the parallax is preserved. The non-uniform back-light gives also black regions on the border of the screens, as clearly visible in figure 3(c) and figure 4(c). The captured images are more blurred due to a loss of sharpness in the light field acquisition process. In the simulation, this problem does not exist since we consider perfect cameras.

Additionally to these results, a video demonstrating the capability of the tensor display with simulated and cameras acquired light field can be found on Youtube [2].

[2] `https://youtu.be/yy1b5XoHFGo`

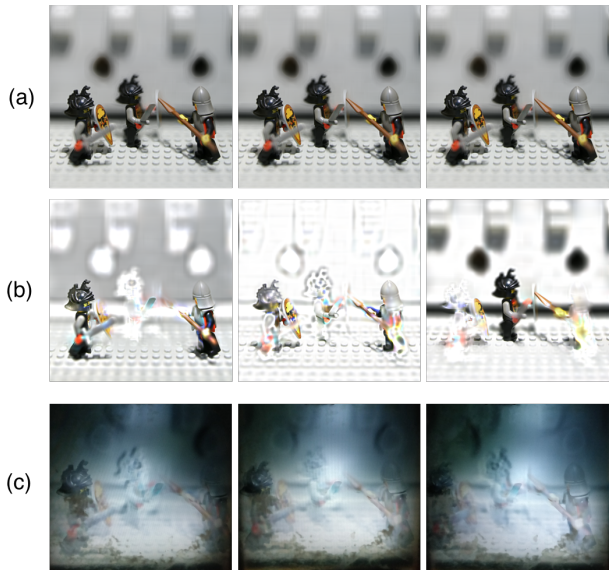

Figure 4. *(a) Original Views. (b) Layered images computed with multi-view. (c) Views displayed by tensor display. Note that, on the borders of the captured views, weird artifacts are displayed. It comes from the glue that was fixing, before, the polarizing films.*

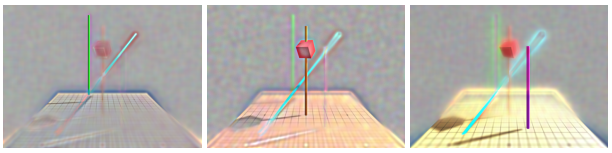

Figure 5. *Layers of the tensor display. Farthest, middle, frontest*

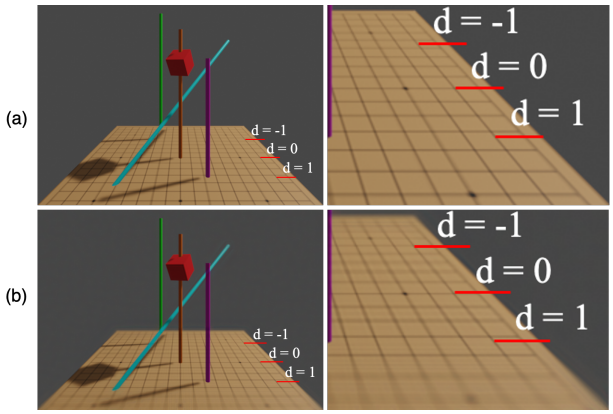

Figure 6. *(a) Center view and zoom-in of the light field captured and (b) the light field produced by the tensor display*

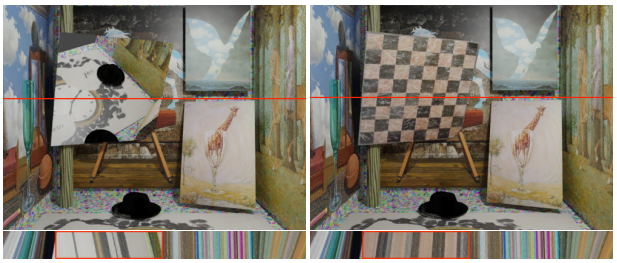

Figure 7. *Center view of a light field capturing a scene with a mirror plane (upper left) and with a simple Lambertian plane (upper right). The corresponding EPIs are below. EPI taken for each light field with rows at the height of the red line. Disparity for the section that corresponds to the plane are clearly different for the mirror plane*

Depth Range

In this section we analyze the characteristics of our tensor display and its limitations. In particular, the constraints placed on the depth of the scene represented and their implications for the rendering of non-Lambertian objects.

We observe the *depth-selectivity*, Takehashi *et al.* [10] highlighted in addition to the limitation on the depth range [17] of the information that can be rendered by the display. We note that Lambertian objects appearing in the scene are separated according to their depth and are distributed across the layers of the tensor display. As shown in figure 5, the closest objects are mainly displayed by the first layer, objects at a halfway depth are represented by the second layer and objects more in the background appear mostly on the third layer.

It follows that objects are sharper when their depth corresponds to the layer depth. More precisely, it is the disparity of an object that affects the quality of its rendering in the tensor display.

Indeed, the first, second and third layers render objects with a disparity of respectively 1, 0 and -1. Thus the baseline between the cameras used to capture the light field must be carefully chosen in order to obtain exploitable data. The depth of the objects and the focal length of the cameras must be taken into account to make sure the objects we wish to render present a disparity included in $\{-1,0,1\}$. This is done by tuning the frustum to have a near value of disparity of -1 and the far value set to 1. The light field shown in figure 6 and used to compute the layers in figure 5 was obtained using a grid of converging cameras [3]. Those cameras are placed in such a way that the red cylinder has a disparity of 0, the purple one has a disparity of 1 and the green one has a disparity of -1.

We can observe in figure 6 that the lines of the go board at the disparity of 1, 0 or -1 appear cleanly, while the lines in between appear more blurry. In addition, the lines that are at the front or at the back of the board also appear blurry since they are out of the depth range of the tensor display.

Tensor display for non-Lambertian content

The previous examples of the tensor display all showed scenes that only contain Lambertian objects. Such objects are defined as projecting the same color and light intensity independently of the viewing direction. In an epipolar plane image (EPI) [18] of a light field (i.e. image obtained by stacking on each other rows taken at the same height in lateral views, which amounts to fixing the two vertical coordinates in the light field [19]), Lambertian content appears as straight lines. The slope of these lines corresponds to the object's disparity, and hence is inversely proportional to its depth. Since tensor displays describe linear relationships, they are well suited to represent Lambertian objects. However, non-Lambertian materials present a unique challenge since they show non-linear movement in the EPI [20, 21].

In the previous section, we analyzed the impact of the depth on the quality. For non-Lambertian content, the disparity becomes even more relevant. Indeed, due to the nature of such objects, the disparity of the pixels that represent them does not correspond to their depth, so that the minimum and maximum disparities of the light field can exceed the disparities corresponding to the minimal and maximal depths of the scene [22].

For example, as we can see in figure 7, the disparity of a pixel that appears in a mirror gives information about the depth

[3] `https://github.com/dbonattoj/blender-addon`

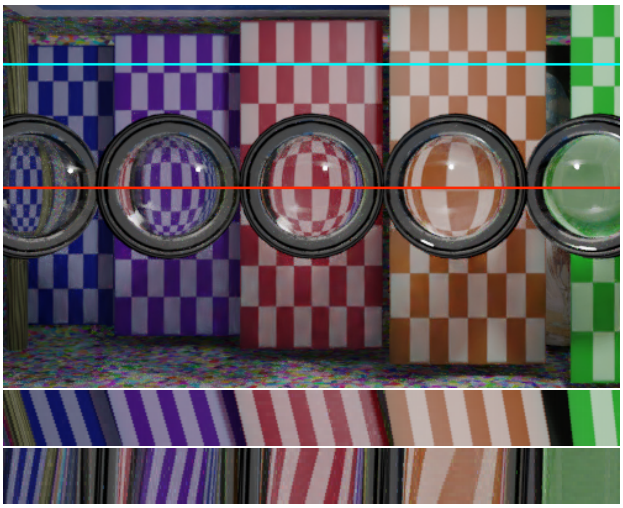

Figure 8. *Comparison of disparity of the object and its texture. First EPI is taken at height of blue line, second EPI is taken at height of red line*

of the virtual scene reflected in the object. This perceived depth does not match the physical depth of the mirror.

For a transparent sphere or a half-sphere, the opposite effect appears. The texture that appears on the surface of the non-Lambertian object has a disparity that places it in front of the object itself. We can notice the phenomena in figure 8, where the light field was captured with an array of cameras converging on the position of the half-sphere. Thus the half-spheres themselves have a disparity close to 0, as we can see in the second EPI.

This discrepancy between the depth of an object and the disparity of its pixel will have an impact on our ability to render non-Lambertian objects in a tensor display. In the scene represented in figure 8 the disparity range of the scene is actually greater than its depth range. The texture that appears on the non-Lambertian object behaves as if it belongs to an object in front of the actual non-Lambertian object, which means that from a disparity point of view the range is actually bigger.

In order to accurately represent non-Lambertian objects, the sampling of the light field must be carefully chosen. The disparity of the objects and the perceived disparity of their texture must correspond to disparities one of the layers is able to render. If this is not the case, either the border of the object or its texture or both will appear blurry in the display. However, this may be difficult to achieve if the scene includes a background and several objects with disparate depths. This situation is similar to a scene with only Lambertian objects, with a too-large depth range.

Our proposed solution is to use several light fields of the scene, taken with different baselines, in order to compute the image for each layers using equation (2). To obtain the final result (figure 9) two light fields were used. In the first one, the refracted texture of the non-Lambertian objects has a disparity of 1 but the disparity of the background is below -1. In the second one, the objects themselves have a disparity of 0 and the plane in the background has a disparity around -1, however the refracted texture has a disparity above 1. Values of the layers are computed by minimizing the error between the light field produced by the tensor display and a target light field (cf. equation (2)). Our proposal is to use the first light field as a target when computing the first layer then the second light field when computing the second and third layers. Thus, the first layer of the device displays the refracted texture while the second and third layers represent the objects and the background. The result is visible in figure 9.

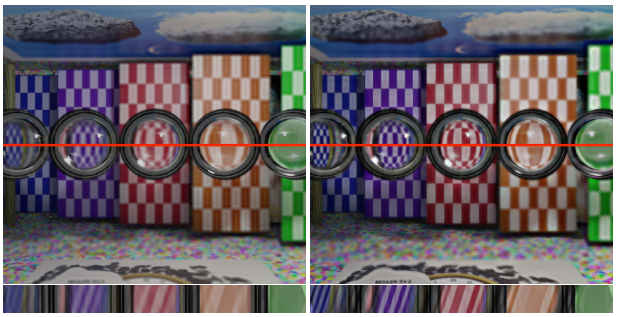

Figure 9. *Simulation of the result produced by the display using a single light field (left) or two different light fields (right). The EPI of the simulated result were taken at the height of the red line. The refracted texture for the three leftmost half-spheres appears significantly sharper, as does the texture for the red checkerboard. This is visible in the view of the simulated light field as well as in the EPI*

This manipulation changes the disparity range of the scene. Thus the rendered scene is physically biased, however it is a way to circumvent the physical limitation of the tensor display. Indeed, the user is less sensitive to the exact disparity in non-Lambertian objects than in diffuse objects, where he relies on the perceived disparity to interpret the depth information. Such technique may also be used to artificially compress the depth range of a scene containing only Lambertian objects, but in that case may distort the perceived depth.

Conclusion and Discussion

We demonstrated the possibility to build a functional tensor display using easily available material. Such devices abide by the same constraints as other displays similarly built. Since the display available disparities are limited to $\{-1, 0, 1\}$, thus, the depth range of a Lambertian scene is limited to the corresponding depths. However, for non-Lambertian scenes, the limitation also depends on the disparities of the non-Lambertian surfaces even if the depth range corresponds to the bounds of $[-1, 1]$. We overcome those limitations by artificially compressing the disparity of the scene using different light fields. Even if the rendered scene did not accurately represent the original scene, it gave a sharper result and preserved the depth perception. While our method to render non-Lambertian objects showed promising results, to obtain the results in figure 9, we used two different synthetic light fields. For real scenes, this approach is cumbersome. Changing the software to resample existing light fields is a possible software solution.

Currently, our prototype is being improved by replacing the back-light with a strong uniform or directional light, and the layer quality will be enhanced by increasing the resolution and the color contrast. Finally, we will extend our datasets to perform subjective user studies to analyze the limits of the proposed light field distortion method.

Acknowledgments

Sarah Fachada is a Research Fellow of the Fonds de la Recherche Scientifique - FNRS, Belgium. This work was supported in part by the HoviTron project that received funding from the European Union's Horizon 2020 research and innovation program under grant agreement N° 951989; in part by the FER 2021 project, Grant N° 1060H000066-FAISAN, and in part by the Emile DEFAY 2021 project, Grant N° 4R00H000236-EmileDEFAY, Belgium. Authors appreciate technical support by Mr. Ir. Rudy ERCEK during the development of the project.

References

[1] Marc Levoy and Pat Hanrahan. Light field rendering. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, page 31–42, New York, NY, USA, 1996. Association for Computing Machinery.

[2] Ren Ng, Marc Levoy, Mathieu Bredif, Gene Duval, Mark Horowitz, Pat Hanrahan, and Duval Design. Light Field Photography with a Hand-held Plenoptic Camera. *Stanford Technical Report CTSR*, page 11, April 2005.

[3] Andrew Lumsdaine and Todor Georgiev. The Focused Plenoptic Camera. *2009 IEEE International Conference on Computational Photography, ICCP 09*, April 2009.

[4] Christian Perwass and Lennart Wietzke. Single Lens 3D-Camera with Extended Depth-of-Field. *Proc. SPIE*, 8291:4, February 2012.

[5] George Koulieris, Kaan Akşit, Michael Stengel, R. Mantiuk, Katerina Mania, and C. Richardt. Near-eye display and tracking technologies for virtual and augmented reality. *Computer Graphics Forum*, 38:493–519, 05 2019.

[6] Daniele Bonatto, Sarah Fachada, Segolene Rogge, Adrian Munteanu, and Gauthier Lafruit. Real-Time Depth Video Based Rendering for 6-DoF HMD Navigation and Light Field Displays. *IEEE Access*, pages 1–20, October 2021.

[7] Sarah Fachada, Daniele Bonatto, and Gauthier Lafruit. High Quality Holographic Stereograms Generation using four RGBD Images. In *Applied Optics*. OSA Publishing, October 2020. ISSN: 1559-128X, 2155-3165.

[8] Tibor Balogh, Tamás Forgács, Tibor Agács, Olivier Balet, Eric Bouvier, Fabio Bettio, Enrico Gobbetti, and Gianluigi Zanetti. A scalable hardware and software system for the holographic display of interactive graphics applications. *Eurographics (Short Presentations)*, pages 109–112, 2005.

[9] Gordon Wetzstein, Douglas Lanman, Matthew Hirsch, and Ramesh Raskar. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):1–11, 2012.

[10] Keita Takahashi, Yuto Kobayashi, and Toshiaki Fujii. From focal stack to tensor light-field display. *IEEE Transactions on Image Processing*, 27(9):4571–4584, 2018.

[11] Keita Maruyama, Keita Takahashi, and Toshiaki Fujii. Comparison of layer operations and optimization methods for light field display. *IEEE Access*, 8:38767–38775, 2020.

[12] Daniel Lee and H. Sebastian Seung. Algorithms for non-negative matrix factorization. In T. Leen, T. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems*, volume 13. MIT Press, 2000.

[13] Tamara G. Kolda and Brett W. Bader. Tensor Decompositions and Applications. *SIAM Review*, 51(3):455–500, August 2009.

[14] Daniele Bonatto, Sarah Fachada, Takanori Senoh, Jiang Guotai, Xin Jin, Gauthier Lafruit, and Mehrdad Teratani. Multiview from micro-lens image of multi-focused plenoptic camera. In *2021 International Conference on 3D Immersion (IC3D)*, pages 1–8, 2021.

[15] Gordon Wetzstein. Synthetic light field archive. `https://web.media.mit.edu/~gordonw/SyntheticLightFields/`, Feb 2003.

[16] Stanford University. The (new) stanford light field archive. `http://lightfield.stanford.edu/lfs.html`.

[17] D Lanman, G Wetzstein, M Hirsch, and R Raskar. Depth of field analysis for multilayer automultiscopic displays. *Journal of Physics: Conference Series*, 415:012036, feb 2013.

[18] Robert C. Bolles, H. Harlyn Baker, and David H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International journal of computer vision*, 1(1):7–55, 1987.

[19] Gauthier Lafruit and Mehrdad Teratani. *Virtual Reality and Light Field Immersive Video technologies for Real-World applications*. IET, the Institute of Engineering and Technology, 2022. p.232-249.

[20] Kazuki Maeno, Hajime Nagahara, Atsushi Shimada, and Rin-Ichiro Taniguchi. Light field distortion feature for transparent object recognition. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2786–2793, 2013.

[21] Sarah Fachada, Daniele Bonatto, Mehrdad Teratani, and Gauthier Lafruit. Light Field Rendering for non-Lambertian Objects. In *Electronic Imaging*, February 2021.

[22] Sergio Moreschini, Robert Bregovic, and Atanas Gotchev. Shearlet-based light field reconstruction of scenes with non-lambertian properties. In *2019 8th European Workshop on Visual Information Processing (EUVIP)*, European Workshop on Visual Information Processing, pages 140–145. IEEE, October 2019. European Workshop on Visual Information Processing ; Conference date: 01-01-1900.

Author Biography

Eline Soetens is currently pursuing a master's degree in Computer Science and Engineering at the Université Libre de Bruxelles, Brussel, Belgium, where she intends to pursue a Ph.D degree. Her research interests lie in light field and non-Lambertian analysis.

Armand Losfeld is currently pursuing a master's degree in Computer Science and Engineering at the Université Libre de Bruxelles, Brussels, Belgium, and intends to become a Ph.D candidate at the same University. His research interests are plenoptic imaging and tensor display technologies.

Daniele Bonatto received the master's degree in computational intelligence (applied sciences) from the Université Libre de Bruxelles (ULB), Belgium in 2016. He is currently pursuing the Ph.D. degree in real-time 3D computing, jointly between the ULB and the Vrije Universiteit Brussel. He works on real-time free-viewpoint rendering of natural scenery with sparse multi-camera acquisition setups. Jointly with the Moving Picture Experts Group (MPEG) standardization, he developed the reference view synthesis software.

Sarah Fachada (Student Member, IEEE) graduated from the Ecole polytechnique, France, and the Trinity College of Dublin, Ireland in 2017, majoring in computer science. She is currently pursuing the Ph.D. degree with the Université Libre de Bruxelles, Belgium, working on acquisition and rendering in light fields and DIBR. She is a Research Fellow of the Fonds de la Recherche Scientifique—FNRS, Belgium. Jointly with MPEG, she developed the reference view synthesis software.

Laurie Van Bogaert is pursuing her Ph.D degree at the Université Libre de Bruxelles, Belgium. She received her master's degree of science in Computer Science and Engineering in 2021.

Gauthier Lafruit (Member, IEEE) received the M.Sc. and Ph.D. degrees from Vrje Universiteit Brussel (VUB), in 1989 and 1995, respectively. He is currently a Professor of virtual reality and light field technologies at the l'Université Libre de Bruxelles (ULB), Brussels, Belgium. He works in the domain of visual data analysis and compression, participating to compression standardization of CCSDS, JPEG and MPEG. His research interests include depth image-based rendering, immersive video, point-clouds, and digital holography technologies.

Mehrdad Teratani (Senior member, IEEE) received his PhD degree in Information Electronics from Nagoya University, Japan, in 2004. Since 2020 he is a professor of Light Field Video Engineering at the Université Libre de Bruxelles (ULB), Belgium. He has been active in MPEG standardization since 2009, and chairing Lenslet Video Coding AhG. He holds 16 granted patents. His research are 3-D Imaging Systems, Light Field Video Processing and Compression and Intelligent Video System.