This is the accepted version of the paper available at https://doi.org/10.1117/12.2589128

Realistic physical camera motion for light field visualization

Mary Guindy[ab], Attila Barsi[b], Peter A. Kara[c], Tibor Balogh[a], and Aniko Simon[d]

[a]Holografika, Budapest, Hungary
[b]Pazmany Peter Catholic University, Budapest, Hungary
[c]Budapest University of Technology and Economics, Budapest, Hungary
[d]Sigma Technology, Budapest, Hungary

ABSTRACT

Light field displays depict the real world by emitting light rays corresponding to the 3D space of the scene or object that is to be represented. Therefore, such displays are suitable for utilization in a multitude of use cases, including the cinema. Cinematography is the art encompassing motion picture photography. It describes a great number of methods for working with physical cameras. For example, a camera can be mounted on a car in a typical chase scene, attached to a spring, suspended or mounted on a colliding object and many more. Although these camera rigs may produce exceptional visuals, achieving similar results by using light field cameras will be extremely challenging. In order to capture light fields, two sets of possible solutions can occur. These include capturing light fields by means of an actual light field camera or via an array of cameras, where the latter may be planar or curved. Using one way or the other is rather demanding, as each method has its own limitations and difficulties. When using a light field camera, the captured light field should map to that of the light field display, on which it shall be visualized. Whereas capturing a scene with a camera array, can unfortunately lead to self-capture. Moreover, the portability of the camera arrays is evidently problematic, due to the sheer size and weight. In addition to these challenges, light field rendering on its own is far from being trivial. While rendering to conventional 2D displays from camera arrays may use image-based rendering, where many views of the scene can be set up from pre-captured images, light fields are represented by 5D plenoptic functions that are not easy to capture with conventional camera arrays. Moreover, image-based rendering techniques often fail to produce convincing results for light field displays. For some use cases, they can be reduced into 4D for horizontal-parallax-only light fields, since our eyes are horizontally separated and horizontal motion is more frequent than vertical. In practice, the creation of a light field scene from a set of images requires injecting each 2D image into a 4D light field representation. In this paper, we visualize different simulations for realistic physical camera motions on a real light field display. In order to overcome the aforementioned problems, virtual cameras were used to simulate a set of different physical camera motions used in cinematography for light field displays. In applications, physics simulation libraries include algorithms for the dynamics of soft as well as rigid bodies. Moreover, collision detection is also accounted for. Many tools have been devised to simulate physics, among which is the Bullet Physics library. In our work, we used the Bullet Physics library in order to generate realistic physical camera motions as well as physical environment simulations for light field displays. The limitations and challenges imposed by light field displays when simulating physical camera motions are discussed, along with the results and the produced outputs.

Keywords: Light field, virtual cameras, physical camera simulation, camera motion, cinematography.

1. INTRODUCTION

Among the recent technological advancements in 3D visualization, the immersive, glasses-free technology of light field has emerged in the past decades. The original concept of light field was specified a century ago[1] and continued evolving throughout the years. The technical term "light field" was first defined by Gershun in 1936[2] as "the amount of light traveling in every direction through every point in space".[3] By the beginning of the 21th century, real light field displays have been developed as a means for light field visualization. Such systems evolved over time to include volumetric displays, multi-focal displays, super multi-view displays, etc.[4]

Due to the fact that light field displays do not require additional viewing devices for the natural 3D experience, they provide a greater sense of immersion compared to conventional 3D displays. Thus, they are the most suitable for cinematographic purposes as conventional 3D solutions evoke a sense of isolation for viewers.

Over the years, cinematography has evolved from 2D to 3D, where viewers can watch 3D movies in cinemas by means of 3D glasses. With the aforementioned technological evolution, light field displays offer a huge leap in the field of cinematography by providing viewers with 3D glasses-free experience, adding more sense of immersion, as such displays act as a 3D window to the real world. One of the major aspects of cinematography is the camera control which takes into account the camera motion and path planning to produce realistic motion paths.

Camera motion and path planning has long been investigated for conventional 2D cameras. Applying the same techniques to light field cameras can be challenging, especially for wide-baseline devices. In order to overcome the problems imposed by wide-baseline light field cameras, the usage of virtual cameras to simulate a set of different physical camera motions is suggested.

In this paper, we introduce a novel tool to simulate realistic physical camera motions, as well as physical environments for light field displays. Physics is generated by means of the "Bullet Physics library".[5] Challenges imposed when testing the simulated realistic physical camera motions on light field displays are discussed, along with the results and outputs.

The remainder of this paper is structured as follows: Section 2 discusses the related scientific efforts. Light field camera technology is elaborated in Section 3. The simulation itself, which is the primary contribution of this paper, is introduced and discussed in Section 4. The paper is concluded in Section 5, also addressing the potential future continuations of this work.

2. RELATED WORK

Light fields were introduced to describe the surrounding real world by means of light rays that occupy the 3D space under representation.[3] In 1991, Adelson and Bergen introduced a 7D plenoptic function describing light fields.[6] The dimensionality of this function was further reduced into 5D, and finally into 4D for free spaces, where the 4D light field is composed of 2D image slices.[7]

The state-of-the-art light field displays already aim to enable cinematographic use cases. For example, the HoloVizio C80[8] is a large-scale light field cinema, with an exceptionally high angular resolution (0.5 degrees) and a 3-meter-wide screen. Such displays can be scaled up to serve more conventional numbers of viewers.[9]

Cinematography is defined as the process of rendering all the collected ideas, tones, actions, words and other phenomena in a visual or audiovisual context. In his book "The five C's of Cinematography", Maschelli defined the main elements of cinematography to be camera angles, continuity, cutting, close-ups and composition.[10] Cinematography encompasses a set of rules and conventions in order to efficiently communicate actions, among which is the camera placement, the actor or object of interest, and the Region of Interest (ROI).[11]

Considering the camera placement, path planning for cameras takes into account a lot of aspects. While adhering to the cinematographic rules, path planning aims at displaying the actor or object of interest,[12] as well as the details of significant motions while providing realistic motion paths.[13] Various parameters are crucial in achieving the required camera movements in cinematography, among which are the orientation of the camera, focal length, frame rate, etc. In addition to the camera parameters, a camera may be mounted on a moving platform. Hence, the camera platform's acceleration / deceleration and turn speed are taken into account when designing the movement of the camera. Finally, the physical / visual obstacles obstructing the camera's path must be considered as well. Limitations imposed by these parameters must be taken into account by directors

when planning the camera's path for cinematographic capture. Realistic camera motions, as well as the engaging experience for viewers are the most essential objectives in this aspect of film making. Many tricks are used by directors to achieve these objectives, for example the usage of multiple cameras along the required path where switching between cameras is carried out. While this is possible using conventional cameras, applying the same techniques to light field cameras can be extremely expensive and difficult, especially for wide-baseline cameras, as they are composed of a 1D or 2D array of cameras. Therefore, the usage of virtual cameras is recommended.

With the evolution of 3D graphics and the technological advancements in film making, the use of virtual cameras became essential. Applying the conventions of cinematography to virtual cameras in order to achieve realistic, plausible videos is now considered crucial. Thus, the evolved practical concept of combining cinematography with computer graphics and artificial intelligence as a means of achieving virtual camera path planning has become common practice.[14]

In addition to using virtual cameras, physics simulations need to be considered. Many real-time dynamic engines have been developed to simulate realistic physics. Among the most promising dynamic engines is the Havok physics, NVidia PhysX, Bullet Physics, Open Tissue, etc.[15] Dynamics engines offer a multitude of physical properties, including simulation of rigid bodies, constraint implementations such as suspensions and springs, collision detection, material properties, object representations, integrator and simulator paradigms.[16]

As mentioned earlier, applying cinematography techniques on light field displays can be challenging and it imposes certain difficulties. In this paper, we investigate the simulation of realistic physical cameras on light field displays.

3. LIGHT FIELD CAMERAS

As stated in the previous section, light fields can be described minimally as a 4D representation, where they are are visualized on light field displays. With the complexity of light fields, challenges are being imposed on their capture. Hence, the implementation for light field cameras becomes challenging. Numerous cameras were devised to capture light fields, among which are the plenoptic cameras, camera arrays, time-of-flight (TOF) cameras and numerous others.

3.1 Light field acquisition

Light field cameras are used to capture 4D light field via capturing the 2D position on the image plane, as well as the 2D incident direction. Wetzstein classified light field acquisition into three main categories:[17]

- Multiple sensors: As the name indicates, a 1D or 2D array of cameras is used in this method to capture multiple photographs. Each photograph is a 2D slice, constituting the 4D light field.

- Temporal multiplexing: Instead of using multiple cameras as in the previous method, a single camera is used. In order to obtain multiple images via the single camera, various solutions have been proposed. One of them is to place the object of interest on a turntable. Another one is to move the camera while reorienting it towards the object of interest over a spherical or planar path. Other methods have also been proposed to capture light fields by a single camera, such as programmable aperture photography.[17] Although temporal multiplexing has a great number of advantages over using the multiple sensors technique, including the reduction of complexity and costs, it cannot capture dynamic scenes appropriately.

- Spatial and frequency multiplexing: This category for light field acquisition aims at solving the problems encountered in the two other methods. Many solutions have been proposed, such as integral photography, parallax barriers, etc. High-speed cameras became essential to capture light fields dynamically. An example for high-speed cameras is the plenoptic camera, where an array of micro lenses is positioned in front of the sensor of camera.[18] These types of cameras evidently fall under the category of narrow-baseline, and are thus not suitable to create wide-baseline contents.

3.2 Challenges and obstacles

Light fields are characterized by the baseline: they are either narrow-baseline or wide-baseline light fields. For each category, there are corresponding light field cameras and displays. For light field cameras, the captured light fields should map to those of the light field displays. Unlike narrow-baseline devices, wide-baseline light field cameras (i.e., camera arrays) impose many challenges. Using camera arrays is challenging due to its weight, physical size, possibility of self-capture and high price.[19] Furthermore, it is hard to generate uniform lighting in a scene captured similarly by all cameras in the region. In addition to the aforementioned issues, using dynamic camera arrays introduce additional challenges, among which is the camera tracking and calibration. Many solutions were proposed for the calibration of camera arrays, such as simply calibrating each view point by means of a single camera calibration. An additional solution is the usage of plane plus parallax for calibrating the camera array used for light field acquisition.[20] The portability of camera arrays can be difficult, especially if they are used outdoors due to their reliance on power supply as well as the usage of complicated transmission lines.[21]

4. REALISTIC PHYSICAL SIMULATION FOR LIGHT FIELD CAMERAS

Due to the aforementioned issues elaborated in the previous section, it is advised to use virtual cameras in order to simulate realistic physical environments. Nonetheless, the physical properties for realistic light field cameras should be taken into consideration when using virtual cameras in order to accurately simulate a realistic camera path.

Our work aims at fusing path planning for wide-baseline light field cameras and the simulation of physical cameras for film making. The goal of this paper is to introduce an application used to build physical environments in which we can evaluate camera motions that typically come up during film production. Physical parameters used in path planning can be entered in the application, such as the speed (acceleration / deceleration), mass, etc. of the camera rig. However, the remaining properties (optical and sensor properties) are determined by the light field display, thus ensuring the consistency between the properties of both the light field cameras and displays, and hence, avoiding any unnecessary conversions. The physical environment consists of simple shapes, including a ground, boxes, cylinders, cars, suspensions and planes.

## 4.1 Physical properties for the simulation

Light field cameras are characterized by their sensor properties (frame rate, focal range, aperture size, intrinsic camera parameters), optical properties (baseline, arrangement of the cameras, resolution), physical dimensions (shape of the cameras and the mount) and weight (weight of the camera and the mount). Whereas the optical and sensor properties for light field cameras cannot be changed, physical properties can be easily altered. In our application, we allow users to change the weight of the camera within given limits, as well as the size, speed and suspension properties of the camera mount platform. For the objects in the scene, change is possible for their sizes, positions, orientations and weights. Although the baseline for the light field display (and thus for the light field camera simulation) cannot be changed, changing the ROI achieves the desired effect. ROI defines a region of space that is box-shaped in the virtual scene, within which everything is visible on the light field display.[22] Changing the ROI is the same as changing the baseline and the arrangement of the cameras. ROI is used to transform between the display's physical coordinate system and the world coordinate system. In other words, changing the object of interest is applicable by means of changing the ROI.

4.2 Metrics

Testing the different realistic camera motions in the physical environments require some metrics to define the plausibility of the achieved results. In our work, we specify some of the metrics to be measured, including the collisions, objects entering the blurry region of light field display, occlusions, and hence, image stability is based on these aforementioned metrics. It is important to take into account that these measurements are performed for horizontal-only-parallax (HOP) light field displays. Therefore, the observer line is considered. Measuring camera collisions is achieved by means of axis-aligned bounding boxes (AABBs). The number of objects colliding with the camera is done by counting the number of intersections between the AABBs of the objects in the scene and the AABB of camera.
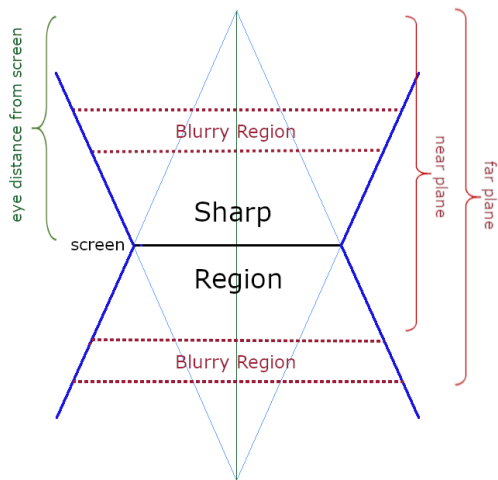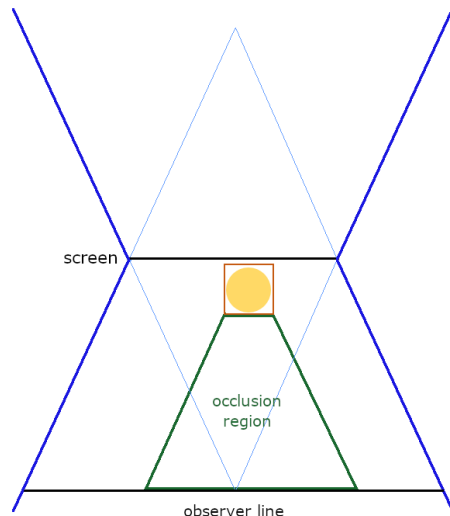
(a) Blurry regions   (b) Occlusion region

Figure 1: Metrics for light field visualization.

For light field displays, double viewing frustums (in front of the screen and behind the screen) are constructed instead of one as in case of conventional displays. Figure 1a shows a top view for the setup of light field displays. The black line depicts the screen, whereas the blue lines show the viewing angles. The dotted lines encompass the areas where objects are rendered blurry. The number of objects in the blurry region is calculated by counting the number of intersections between the AABBs of the objects and the frustum encompassing the blurry region.

For third person cameras, camera tracks and follows an object of interest. Occlusions indicate the existence of other objects in the path between the camera and the object of interest. In order to measure occlusions on light field displays, a frustum in front of the object of interest in the direction of camera is considered. Figure 1b shows the top view for the frustum with respect to the object of interest, where the front plane of the frustum is similar to that of the display. The back plane of frustum is the same as the front plane of the AABB of the object. The right and left planes are parallel to those of the display, but bounding the object of interest instead. As for the top and bottom plane constituting the frustum, they are constructed from the top and bottom lines on the AABB of the object passing through the observer line. The number of objects existing in the occluded region is measured by calculating the number of intersections between the AABBs of the objects and the constructed frustum.

4.3 Evaluation and Results

In this paper, we discuss the simulation of realistic physical camera motions on light field displays. The main goal is to provide a tool for film makers to build an environment in which realistic physical light field camera motions that typically come up during film production can be tested and evaluated. In order to test those camera motions, several scenes were built and tested on the HoloVizio C80 light field display. Bullet Physics library was used for simulating the physical properties of the system. A physical scene can be built by means of some shapes, including ground, planes, boxes, cylinders, cars and suspensions. Our applications allow users to enter some physical parameters used in path planning, such as the mass of camera and the speed of the camera platform. Accordingly, path planning for light field displays is tackled by defining the scene, path and physical properties. Many scenes were built and tested for the different camera motions.

Figure 2 shows three of the scenes. The first scene includes a car moving towards a set of columns. The car accelerates to collide with one of the columns, resulting in its fall. The camera is mounted once on the car and the other time on the column to detect the effect of collisions on the light field camera. The second scene contains a suspension element on which the camera is mounted. It depicts the effect of a camera mounted on a suspended platform. The third scene depicts a camera falling towards the ground.
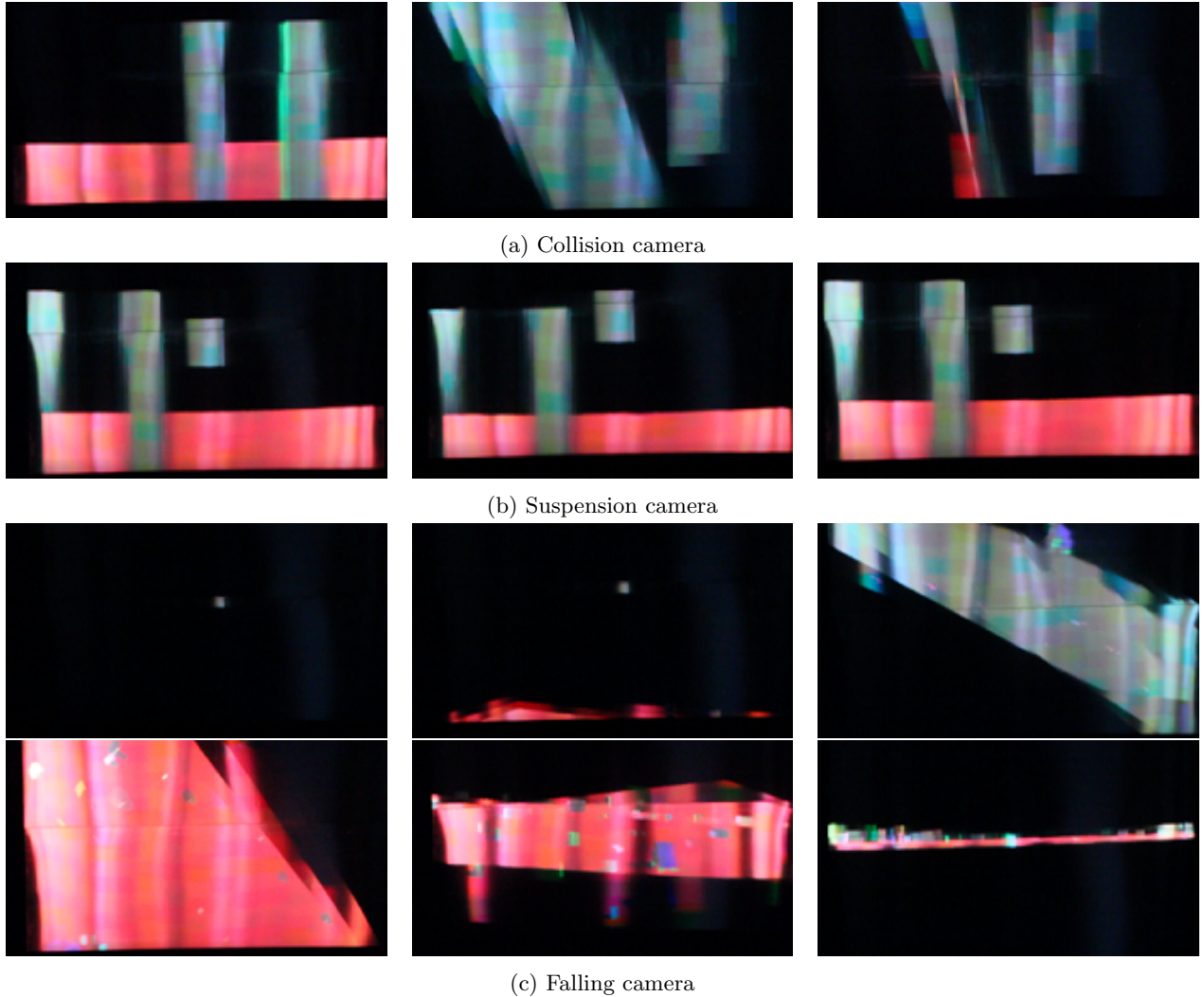
(a) Collision camera

(b) Suspension camera

(c) Falling camera

Figure 2: Physical simulation of cameras on light field display.

5. CONCLUSION

Light field displays present the viewers with an immersive, glasses-free 3D environment, and according to the state-of-the-art technologies, at the time of this paper, it can be stated that they are the most suitable for cinematographic purposes. Light fields are captured by means of light field cameras (narrow-baseline and wide-baseline). Applying the conventional rules of cinematography and techniques for light field cameras can be challenging, especially for wide-baseline systems. Hence, we present a tool for those involved in cinematographic content creation to virtually simulate the different realistic physical light field camera motions.

As for potential future continuations of our scientific contribution, additional physical properties could be further added to the simulation, including physical stress on the frame of the mount, oscillations when accelerating / decelerating, camera arrays with re-configurable frames and numerous others. Moreover, the planning of lighting from different angles should also be considered. In other words, making sure that the necessary lighting for simulation should be exactly the same from all angles.

Acknowledgments

REFERENCES

[1] Lippmann, G., "Épreuves réversibles donnant la sensation du relief," *J. Phys. Theor. Appl.* **7**(1), 821–825 (1908).

[2] Gershun, A., "The light field," *Journal of Mathematics and Physics* **18**(1-4), 51–151 (1939).

[3] Balram, N., "Fundamentals of light field imaging and display systems," (2016).

[4] Huang, F.-C., Wetzstein, G., Barsky, B. A., and Raskar, R., "Eyeglasses-free display: Towards correcting visual aberrations with computational light field displays," *ACM Trans. Graph.* **33**(4) (2014).

[5] Coumans, E. et al., "Bullet physics library," *Open source: bulletphysics. org* **15**(49), 5 (2013).

[6] Adelson, E. H. and Bergen, J. R., [*The plenoptic function and the elements of early vision*], vol. 2, Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology (1991).

[7] Levoy, M. and Hanrahan, P., "Light field rendering," in [*Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*], *SIGGRAPH '96*, 31–42, Association for Computing Machinery, New York, NY, USA (1996).

[8] Holografika, "HoloVizio C80 glasses-free 3D cinema system." https://holografika.com/c80-glasses-free-3d-cinema/.

[9] Kara, P. A., Martini, M. G., Nagy, Z., and Barsi, A., "Cinema as large as life: Large-scale light field cinema system," in [*2017 International Conference on 3D Immersion (IC3D)*], 1–8 (2017).

[10] Mascelli, J. V., [*The five C's of cinematography*], Grafic Publications (1965).

[11] He, L.-w., Cohen, M. F., and Salesin, D. H., "The virtual cinematographer: A paradigm for automatic real-time camera control and directing," in [*Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*], *SIGGRAPH '96*, 217–224, Association for Computing Machinery, New York, NY, USA (1996).

[12] Yeh, I.-C., Lin, C.-H., Chien, H.-J., and Lee, T.-Y., "Efficient camera path planning algorithm for human motion overview," *Comput. Animat. Virtual Worlds* **22**(2–3), 239–250 (2011).

[13] Davis., K., Viswanathan., V. K., Clark., C. M., Gambin., T., and Wood., Z. J., "Cinematographic and geometric criteria for virtual camera path generation for the visualization of shipwreck data," in [*Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 1: GRAPP,*], 191–198, INSTICC, SciTePress (2018).

[14] Lino, C., Christie, M., Ranon, R., and Bares, W., "The Director's Lens: An Intelligent Assistant for Virtual Cinematography," in [*ACM Multimedia*], (2011).

[15] Glondu, L., Marchal, M., and Dumont, G., "Evaluation of physical simulation libraries for haptic rendering of contacts between rigid bodies," in [*ASME 2010 World Conference on Innovative Virtual Reality*], 41–49 (2010).

[16] Boeing, A. and Bräunl, T., "Evaluation of real-time physics simulation systems," in [*Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia*], *GRAPHITE '07*, 281–288, Association for Computing Machinery, New York, NY, USA (2007).

[17] Wetzstein, G., *Computational plenoptic image acquisition and display*, PhD thesis, University of British Columbia (2011).

[18] Rogge, S., Schiopu, I., and Munteanu, A., "Depth estimation for light-field images using stereo matching and convolutional neural networks," *Sensors* **20**(21) (2020).

[19] Cserkaszky, A., Kara, P. A., Tamboli, R. R., Barsi, A., Martini, M. G., and Balogh, T., "Light-field capture and display systems: limitations, challenges, and potentials," in [*Novel Optical Systems Design and Optimization XXI*], International Society for Optics and Photonics (2018).

[20] Xu, Y., Maeno, K., Nagahara, H., and Taniguchi, R.-i., "Camera array calibration for light field acquisition," *Frontiers of Computer Science* **9**(5), 691–702 (2015).

[21] Pei, Z., Li, Y., Ma, M., Li, J., Leng, C., Zhang, X., and Zhang, Y., "Occluded-object 3D reconstruction using camera array synthetic aperture imaging," *Sensors* **19**(3), 607 (2019).

[22] Darukumalli, S., Kara, P. A., Barsi, A., Martini, M. G., and Balogh, T., "Subjective quality assessment of zooming levels and image reconstructions based on region of interest for light field displays," in [*2016 International Conference on 3D Imaging (IC3D)*], 1–6 (2016).