

REVIEW

Open Access

Review of light field technologies



Shuyao Zhou^{1,2}, Tianqian Zhu¹, Kanle Shi¹, Yazi Li¹, Wen Zheng¹ and Junhai Yong^{3*}

Abstract

Light fields are vector functions that map the geometry of light rays to the corresponding plenoptic attributes. They describe the holographic information of scenes by representing the amount of light flowing in every direction through every point in space. The physical concept of light fields was first proposed in 1936, and light fields are becoming increasingly important in the field of computer graphics, especially with the fast growth of computing capacity as well as network bandwidth. In this article, light field imaging is reviewed from the following aspects with an emphasis on the achievements of the past five years: (1) depth estimation, (2) content editing, (3) image quality, (4) scene reconstruction and view synthesis, and (5) industrial products because the technologies of light fields also intersect with industrial applications. State-of-the-art research has focused on light field acquisition, manipulation, and display. In addition, the research has extended from the laboratory to industry. According to these achievements and challenges, in the near future, the applications of light fields could offer more portability, accessibility, compatibility, and ability to visualize the world.

Keywords: Light field imaging, Holographics, Human-machine graphic interaction

Introduction

A light field is the totality of light rays or radiance in three-dimensional (3D) space through any position and in any direction, as defined by Gershun [1] in 1936. Formally,

$$L : g \rightarrow c$$

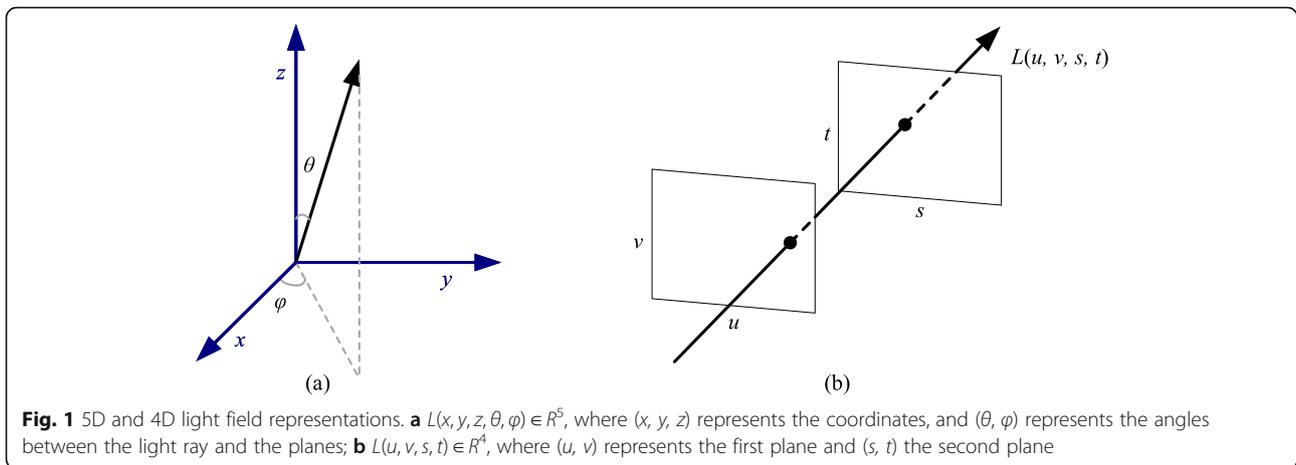
where a light field L is defined by mapping the geometry of a light ray g to the attributes of the corresponding light c . Here, c is a vector that describes the intensity of every component of the light such as red, green, and blue (RGB). Geometrically, g has various definitions in different light field models. A plenoptic function describes all visual information [2]. Gershun [1] defined a five-dimensional (5D) plenoptic function $L(x, y, z, \theta, \phi) \in R^5$ for the light field because each ray can be parameterized by three coordinates (x, y, z) and two angles (θ, ϕ) . Compared with the previous 5D representation, Levoy and Hanrahan [3] assumed in their four-dimensional (4D) representation $L(u, v, s, t) \in R^4$ that the light field is

composed of oriented lines in free space, successfully reducing the redundancy of the total dataset and simplifying the reconstruction of the plenoptic function. $L(u, v, s, t)$ parameterizes lines by their intersections with two planes in an arbitrary position, where (u, v) represents the first plane and (s, t) represents the second plane (see Fig. 1 for the 5D and 4D light field representations and Fig. 2 for two different visualizations of the light field). Meanwhile, Levoy and Hanrahan introduced light fields to the computer graphics field. In addition, if one describes a light field captured by a camera moving on a sphere centered on the target object, then geometry g can be defined as $(\theta, \phi, s, t) \in R^4$, where $(\theta, \phi) \in R^2$ is a spherical surface and $(s, t) \in R^2$ is a plane surface with light projecting to it. "Ray space" is a synonym of "light field" [4, 5] to describe rays in a 3D space. The light field is the same as the orthogonal ray space. In the field of free-viewpoint televisions [6, 7], the term "ray space" is often used to describe ray-based 3D information display systems.

In 2005, the light field made the transition from mainly pure research to large-scale industrial applications. For example, Ng et al. [8] developed the first handheld plenoptic camera. It was not until 2010 that

* Correspondence: yongjh@tsinghua.edu.cn

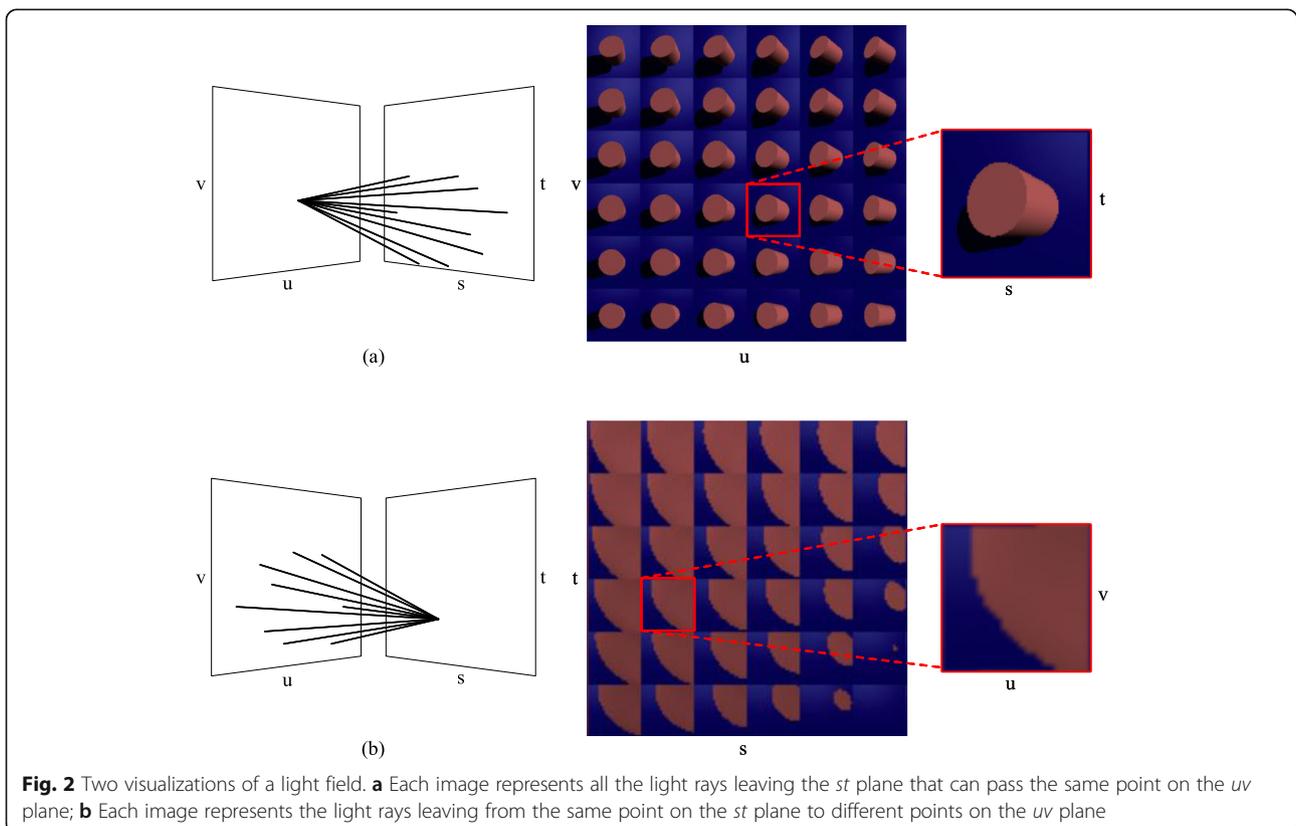
³School of Software, BNRist, Tsinghua University, Beijing, China
Full list of author information is available at the end of the article



light field technology was commercialized to capture a light field. With the development of commercial light field cameras [9], plenoptic cameras that make it possible to refocus provide many benefits, and they have been widely used in light field applications. Subsequently, the commercial potential of the light field has been greatly illustrated in image editing, holographically perceived light fields, augmented reality, and physical entities such as plenoptic cameras. The number of

publications has increased geometrically as light fields have gained increasing attention from researchers (see Fig. 3 for the timeline of light field imaging).

Light field acquisition is the preliminary light field imaging process. Wu et al. [10] comprehensively highlighted methods and devices for light field acquisition in their survey, including (1) multisensor capture (using multiple cameras to capture a light field at one time, with most of them being camera arrays [11–17]),



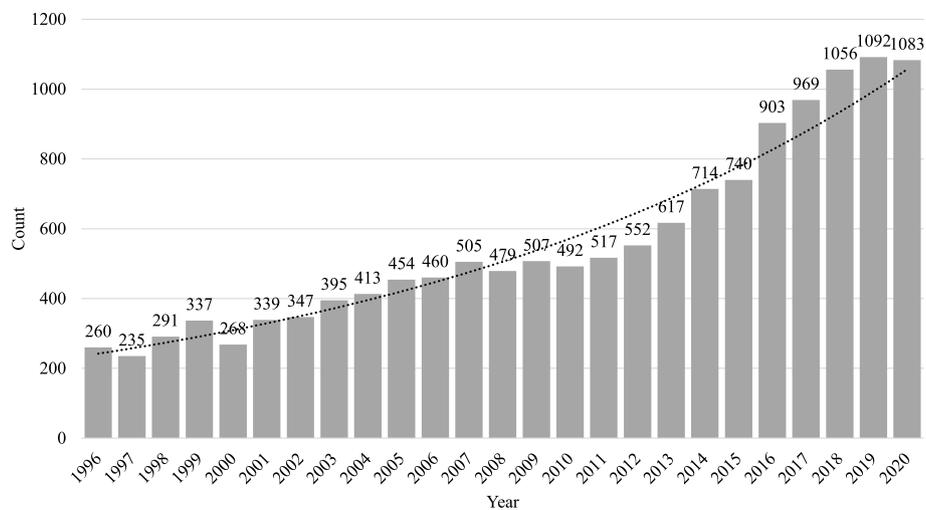


Fig. 3 Light field imaging timeline: numbers of publications whose titles contain keywords “light field” or “light fields” or “light-field” or “ray space” from 1996 to 2020 (the dotted line is the geometric approximation)

(2) time-sequential capture (using one camera to capture a light field with multiple exposures, which is time consuming [18–23]), and (3) multiplexed imaging [encoding high-dimensional data into a simpler two-dimensional (2D) image, which is the most popular method [8, 24–37].

Herein, this paper aims to review light field imaging, revealing the current deficiencies and exploring the future possibilities (see Fig. 4 for an overview). Five aspects have been reviewed: current depth estimation methods (Depth estimation Section), light field editing techniques (Editing Section), light field enhancements with an emphasis on increasing the quality of images (Enhancement Section), 3D reconstruction and view synthesis (Reconstruction and view synthesis Section), and the light field industry, which is categorized into light field acquisition and light field displays (Section 6).

Depth estimation

Depth estimation involves inferring 3D information from 2D images, which is a foundation for light field editing and rendering. Light field data record the spatio-angular information of light rays; thus, a light field image contains many depth cues to make depth estimation possible. Conventionally, depth cues include, but are not limited to, correspondence cues, defocus cues, binocular disparity, aerial perspective, and motion parallax. Occlusion often occurs when two or more objects come too close and, therefore, hide some information from each other. Specifically, when people want to see an occluded object, they usually move slightly to avoid the occluder. This commonsensical solution explains why light fields have special benefits in solving the depth map with occlusion. Therefore, researchers have mainly focused on traditional approaches and convolutional neural network (CNN) approaches for depth

estimation with examinations of occlusion handling: (1) constraint-based estimation, (2) epipolar plane image (EPI)-based estimation, and (3) CNN-based estimation.

Constraint-based estimation utilizes different constraints of the light field structure to estimate the depth. Bishop and Favaro [38] estimated the depth from multiple aliased views and demonstrated that this could be done at each pixel of a single light field image. Williem and Lee [39] utilized the correspondence cue and defocus cue, which were robust against both occlusion and noise, and they introduced two data costs: the constrained angular entropy cost and constrained adaptive defocus cost. Zhu et al. [40] addressed a multiocluder occlusion by regularizing the depth map with an antiocclusion energy function. Some researchers have considered the relationship between occluders and natural light reflections. For example, Baradad et al. [41] estimated the 4D light field of a hidden scene from 2D shadows cast by a known occluder on a diffuse wall by determining how light, which naturally reflected off surfaces in the hidden scene, interacted with the occluder. Chen et al. [42] detected partially occluded boundary regions (POBRs) by using superpixel-based regularization. After a series of shrinkage and reinforcement operations on the labeled confidence map and edge strength weights over the POBR, they produced a depth estimate with a low average disparity error rate and high occlusion boundary precision-recall rate. To proceed with occlusion handling from image to video, Lueangwattana et al. [43] examined the structure from motion to improve light field rendering and, hence, addressed fence occlusion in videos while preserving background details.

However, the EPI, proposed by Bolles et al. [44] in 1987, simplifies depth measurement by restricting

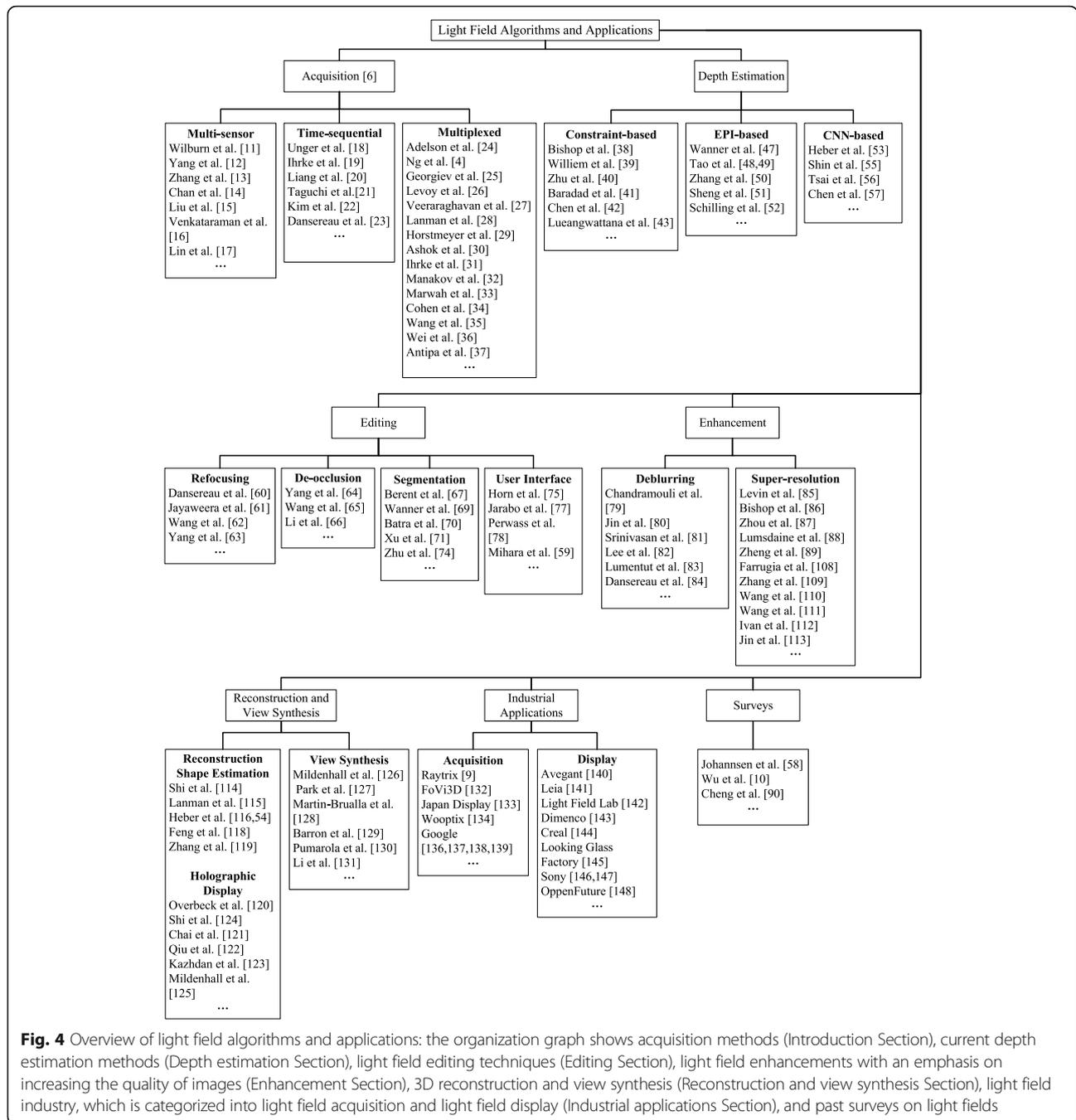


Fig. 4 Overview of light field algorithms and applications: the organization graph shows acquisition methods (Introduction Section), current depth estimation methods (Depth estimation Section), light field editing techniques (Editing Section), light field enhancements with an emphasis on increasing the quality of images (Enhancement Section), 3D reconstruction and view synthesis (Reconstruction and view synthesis Section), light field industry, which is categorized into light field acquisition and light field display (Industrial applications Section), and past surveys on light fields

motion to straight lines and working with a series of closely spaced images, thereby reducing the 3D problem into a set of 2D problems. Some studies on EPI representations are highlighted here. Matoušek et al. [45] suggested a dynamic-programming-based algorithm find correspondences in EPIs by extracting lines with similar intensities in an EPI separately for each row. In other research, Criminisi et al. [46] worked with EPI volume, a dense horizontally rectified spatio-temporal volume that results from a linearly translating camera, for automated

layer extraction. They relied on an EPI tube, which is a collection of EPI lines of the same depth.

Unlike the above works that refine EPI representations, the following works established how to apply EPIs in depth estimation. Wanner and Goldluecke [47] used the dominant directions of EPIs from the structure tensor method to estimate depth. However, this method is sensitive to noise and occlusion. In addition, estimation based on a 2D EPI is vulnerable to noise and sometimes fails because of very dark and bright image features.

Multiorientation EPIs are epipolar plane images in all available directions and provide rich light field angular information. To achieve a better depth map from light field images, Tao et al. [48] computed dense depth estimation by combining defocus and correspondence depth cues based on full 4D EPI. Moreover, defocus cues perform better in repeating textures and noise, and correspondence cues are robust in terms of bright points and features. Tao et al. [49] obtained defocus cues by computing the spatial variance after angular integration and correspondence depth cues by computing the angular variance. Furthermore, they performed depth estimation on glossy objects with both diffuse and specular reflections and one or more light sources by exploiting the full EPIs. Similarly, Zhang et al. [50] proposed a spinning parallelogram operator (SPO) to locate lines and calculate their orientations in an EPI for local depth estimation, which further handled occlusions and was more robust to noise. In addition, Sheng et al. [51] combined a multiorientation SPO with edge orientation to improve depth estimation around occlusion boundaries. They proved that the direction of the optimal EPI was parallel to the boundary of the occlusion. In contrast to the work of Sheng et al. [51], Schilling et al. [52] incorporated both depth and occlusion using an inline occlusion-handling scheme, OBER-cross+ANP, to improve object boundaries and smooth surface reconstruction.

Deep CNNs have been extensively applied to depth estimation because they have a better balance between accuracy and computational cost. In 2017, Heber et al. [53] extended the previous work [54] in which the network operated on EPIs, and they replaced all 2D operators with 3D counterparts. Then, they used a CNN that predicted disparity based on RGB EPI volumes, and the proposed network learned to recover depth information for shape estimation. For supervised training, researchers require large labeled datasets. Shin et al. [55] solved the data insufficiency problem by incorporating a multi-stream network, which encoded each EPI separately for depth estimation, into their CNN model EPINET. Tsai et al. [56] proposed an attention-based view selection network that exploited the priorities of light field images and the correlations between them to reduce redundancy and computation time. In 2021, Chen et al. [57] applied an attention-based multilevel fusion network. They grouped four directions (0°, 45°, 90°, and 135°) of light fields into four branches. Then, they combined the branches with two feature fusion methods to generate depth maps: intrabranch feature fusion based on channel attention and interbranch feature fusion based on branch attention. Researchers usually employ deep CNNs for accurate depth estimation, combining them with traditional approaches to produce better results.

Depth estimation has been a focus of much research. Researchers have worked on constraint-based methods and have explored different depth cues and their combinations. They also simplified the estimation by using EPIs and applying learning-based methods. There have been other studies that evaluate depth estimation methods. The work of Johannsen et al. [58] covers more depth estimation algorithms before and including 2017. The key to enhancing other light field-related applications, such as refocusing or rendering, is to develop more-precise and more robust depth estimation methods.

Editing

Because most light field datasets contain redundancy, researchers are interested in fully using redundancy and manipulating the light field images. Editing light fields is challenging [59] because (1) the light fields are 4D, whereas most tools on the market are for 2D, (2) local edits need to preserve the redundancy of the 4D light field, and (3) the depth information of the 4D light field is implicit. Light field image editing can be divided into (1) refocusing, (2) removing the occlusion, (3) segmenting the light fields to make the editing experience as smooth as editing a 2D image (e.g., removing the scene objects or changing their color), and (4) improving the user interface of the light field editing.

Because light field images contain not only textural information but also geometrical information, researchers can explore refocusing after capturing that cannot be accomplished with 2D images. In 2015, Dansereau et al. [60] demonstrated that a hyperfan-shaped passband can achieve refocusing over a wide range of depths, which they called “volumetric refocusing.” However, the approach only worked for a single volumetric region. To overcome this, Jayaweera et al. [61] proposed a simultaneous refocusing approach for multiple volumetric regions in light fields. They employed a 4D sparse finite-extent impulse response filter, which is a series of two 2D filters composed of multiple hyperfan-shaped passbands. Noncrucial parts of images produced by digital single-lens camera arrays often experience blurs (bokeh). Wang et al. [62] proposed a light field refocusing method to improve bokeh rendering and image quality. They first estimated the disparity map and rendered the bokeh on the center-view sub-image. The rendered bokeh image was then used as a regularization term to generate refocused images. Moreover, Yang et al. [63] proposed a refocusing framework that produced coordinates for interpolation, and they aligned the images onto the focal plane.

Occlusion removal is another typical task in light field editing. The nature of light field sub-aperture images (SAIs) provides complementary information so that

hidden scenes can be seen from other views. Yang et al. [64] partitioned an image into multiple visibility layers and propagated the visibility information through layers. The visibility layer is defined as all the occlusion-free rays in any camera, computed by energy minimization. For CNN methods, Wang et al. [65] suggested a deep encoder-decoder network for automatically extracting foreground occlusions by analyzing scene structures. The SAIs were first encoded with spatial and angular information and then decoded for center-view reconstruction. However, they only considered the one-dimensional (1D) connections among the SAIs. To improve this, Li et al. [66] proposed another CNN-based encoder-decoder method (Mask4D) to learn the occlusion mask with center-view reconstruction. They applied a 5D tensor to explore spatial connections among SAIs. Occlusion removal is also useful for reconstruction.

Segmentation is a specific research focus in light field editing. Berent and Dragotti [67] proposed an algorithm to extract coherent regions based on a level set method [68]. Wanner et al. [69] carried out globally consistent multilabel assignment for light field segmentation. It used appearance and disparity cues, similar to the multiple-view object segmentation method developed by Batra et al. [70], a method that could automatically segment calibrated images from multiple viewpoints with an energy minimization framework that combined stereo and appearance cues. Xu et al. [71] proposed an approach for localizing transparent objects in a light field image. They used light field linearity, the linearity of the light field distortion feature, which modeled refraction in objects between views captured by a light field camera, to separate Lambertian objects (good light field linearity) and transparent objects (poor light field linearity), and they found the occlusion area by using the occlusion detector, which detected occlusion points by checking the consistency of the forward and backward matches between a pair of viewpoints. As a result, the method could finish the transparent object segmentation automatically without any human interaction.

Superpixel algorithms [72] group pixels into perceptually meaningful atomic regions, which can be used to replace the rigid structure of the pixel grid. Previous methods of segmenting 2D images, such as the simple linear iterative clustering superpixels [73], adopted k-means for superpixel generation. As mentioned in the first paragraph of this section, there are three difficulties with light field editing. In 2017, Zhu et al. [74] defined a light field superpixel as a light ray set that contains all rays emitted from a proximate, similar, and continuous surface in the 3D space, and they essentially eliminated the defocus and occlusion ambiguities in traditional 2D superpixels. Unlike in previous works, they focused on a smaller unit — the superpixel — illustrating that

superpixel segmentation on a 4D light field performed better in representing the proximity regions.

For existing user interfaces, users can employ tools to edit the light field, even though the depth map is imperfect. Horn and Chen [75] designed LightShop to manipulate light fields by operating on view rays. Image warping is a process of distorting an image [76]. When a user defines how the view rays warp, the LightShop renderer composites and renders multiple light fields by executing the user-defined ray-shading program. However, the system has limitations when compositing different light fields because of the fixed illumination of a light field. In 2014, Jarabo et al. [77] provided an overview of different light field editing interfaces, tools, and workflows from a user perspective. Some products, such as the single-lens 3D-camera with extended depth of field presented by Perwaß and Wietzke [78], can refocus or apply predefined filters to light field images. Building on previous work [75], Mihara et al. [59] were the first to use a graph-cut approach for 4D light field segmentation based on a learning-based multilabel segmentation scheme. The user needs to specify a target region so that the algorithm can identify the appropriate regions and evaluate whether each ray is included in the selected region. Moreover, they defined appropriate neighboring relationships to preserve redundancies.

Overall, light field editing methods remain largely unexplored compared with 2D image editing. Researchers have experimented further on light field images, such as adding mosaic or special effects or filters, editing body features, combining several images into a short movie, and changing background, which are promising.

Enhancement

Light field enhancement optimizes the quality of light field images. Modern light field research mainly focuses on deblurring and super-resolution (SR).

Motion deblurring has two approaches: blind motion deblurring and nonblind deblurring. Blind motion deblurring has been extensively studied using 2D images. In 2014, Chandramouli et al. [79] first investigated motion deblurring for light field images and assumed constant depth and uniform motion for the simplicity of the model. Jin et al. [80] explored bilayer blind deconvolution that removed the motion blur in each layer to recover 2D textures from a motion-blurred light field image. Moreover, Srinivasan et al. [81] introduced a light field deburring algorithm by analyzing the motion-blurred light field in the primal and Fourier domains. Unlike in previous studies, they recovered a full 4D light field image. Lee et al. [82] addressed six-degree-of-freedom (6-DOF) blind deblurring by considering the 3D orientation change of the camera. Lumentut et al. [83] proposed a deblurring deep neural network with 16,000

times faster speed than in prior work [81] and deblurred a full-resolution light field in less than 2 s. Dansereau et al. [84], unlike in the above approaches, adopted a nonblind algorithm for 6-DOF motion blur, assuming that the ground truth camera motion was known.

Levin et al. [85] illustrated that there is a trade-off between spatial and angular resolutions. Therefore, many researchers have endeavored to improve the spatial resolution of the captured light field using an SR algorithm by exploiting additional information from the available data. For instance, in 2009, Bishop et al. [86] studied SR algorithms using a depth map. They characterized the point-spread function of a plenoptic camera under Gaussian optics assumptions for a depth-varying scene and formulated the reconstruction of the light field in a Bayesian framework. Therefore, they restored the images at a resolution higher than the number of microlenses. Zhou et al. [87] applied the ray-tracing method to analyze the subpixel shifts between the angular images extracted from the defocused light field data and the blur in the angular images, and they obtained an SR result with a magnification ratio of 8. In contrast to restoring high-resolution images from low-resolution images, Lumsdaine and Georgiev et al. [88] rendered high-resolution images by adopting positional and angular information in captured radiance data. They rendered images from a 542-megapixel light field to produce a 106-megapixel final image. Zheng et al. [89] presented a convolutional deep neural network using cross-scale warping to the reference-based SR, which involves applying an extra high-resolution image as a reference to help super-resolve a low-resolution image that shares a similar viewpoint. In 2019, Cheng et al. [90] categorized the existing SR methods into projection-based [91–94], optimization-based [95–101], and learning-based [102–107]. Moreover, Farrugia and Guillemot [108] reduced the light field angular dimension using low-rank approximation and then applied CNNs to achieve peak signal-to-noise ratio (PSNR) gains of 0.23 dB over the second-best-performing method. Zhang et al. [109] proposed a residual convolutional network for higher spatial resolution. They first improved the spatial resolution of the central view image because it contained more subpixel information. Then, they trained the network to improve the spatial resolution of the entire light field image. Wang et al. [110] suggested a spatial-angular interactive network. They began by extracting the spatial and angular features independently from the input light field images, and then the information was processed by many interaction groups to achieve spatial-angular interaction features. Finally, the interacted features were fused to achieve high-resolution SAIs. Similar to previous researchers [110], who used feature collection and distribution, Wang et al. [111] proposed a deformable

convolution network where all side-view features were aligned with the center-view feature and then aligned with the original features. Consequently, angular information is encoded into views for SR performance. Ivan and Williem [112] investigated an end-to-end encoder-decoder style for a joint spatial and angular light field SR model from only a single image without relying on physical-based rendering or secondary networks so that end users can experience the advantages of light field imaging. Jin et al. [113] proposed another learning-based light field spatial SR framework that uses deep combinatorial geometry embedding and structural consistency regularization. Their method improved the average PSNR by more than 1.0 dB and preserved more-accurate parallax details at a lower computational cost.

Researchers have devoted great effort to light field image deblurring and resolution trade-offs. For individual users, the capturing process involves more randomness, which requires stabilization, and they expect to achieve a high-resolution display similar to what they can obtain from a normal camera.

Reconstruction and view synthesis

Geometric reconstruction involves reconstructing an object from its geometric information. Sparsity in the Fourier domain is an important property that makes 4D light field reconstruction possible from a small set of samples. Shi et al. [114] proposed a method for recovering non-Lambertian light fields from a small number of 1D viewpoint trajectories optimized for sparsity in the continuous Fourier domain. In general, 3D reconstruction includes shape estimation, and some research has extended the results to holography. In addition, view synthesis creates new views from a given set of views.

For shape estimation, Lanman et al. [115] proposed surround structured lighting to achieve full 360° reconstructions using a single camera position, but it could only scan relatively small volumes. Subsequently, multi-view reconstruction methods have become increasingly popular. Heber et al. [116] proposed a variational multi-view stereo method, in which they used a circular sampling scheme inspired by a technique called “active wavefront sampling” (AWS) [117], where the AWS module is an off-axis aperture that moves along a circular path around the optical axis. In 2016, Heber and Pock [54] trained a CNN to predict 3D scene points from the corresponding 2D hyperplane orientation in the light field domain by using horizontal and vertical EPIs and dividing each EPI into patches. Similarly, Feng et al. [118] focused on 3D face reconstruction from 4D light field images using a CNN. They constructed 3D facial curves, rather than a complete face, to make up a 3D face at once by combining all the horizontal and vertical curves of a face to form horizontal and vertical depth

maps separately. However, unlike Heber and Pock [54], they exploited a complete EPI for depth prediction. Zhang et al. [119] applied a light field camera as a virtual 3D scanner to scan and reconstruct 3D objects, which enabled dense surfaces to be reconstructed in real time. They illustrated that, with five light fields, the reconstructed 3D models were satisfactory.

Another application of light field 3D reconstruction is the holographically perceived light field because the light field can provide continuous focus cues. Overbeck et al. [120] developed “welcome to light fields” to enhance the virtual reality experience by setting up a 16-GoPro rotating rig, processing multiview depth maps, adopting disk-based light field rendering to make seamless connections among pictures, and varying compression levels with the movement of the eyeballs. Many studies have demonstrated high-quality scene rendering [121–123]. In 2017, Shi et al. [124] introduced near-eye light field computer-generated rendering with spherical waves for wide-field-of-view interactive 3D computer graphics. Nonetheless, Mildenhall et al. [125] proposed a deep-learning method for view synthesis from an irregular grid of sampled views that first expands each sampled view into a local light field by means of a multiplane image scene representation and then blends adjacent local light fields. They used up to 4000 times fewer views and presented a plenoptic sampling framework by clearly specifying how users should sample input images for view synthesis with portable devices.

Recent view synthesis methods have approached new view generation from only a few input images. The neural radiance field (NeRF) [126] is a nonconvolutional deep network representation that characterizes the volume space with a multilayer perceptron. It takes a continuous function $(x, y, z, \theta, \phi) \in R^5$ as input, and it outputs the volume density and view-dependent RGB color. Overall, it reconstructs the surface geometry and appearance from a small set of images. Subsequent studies have extended the NeRF in a variety of ways. Park et al. [127] modeled shape deformations by augmenting the NeRF. They can reconstruct free-viewpoint selfies from photographs. Nonetheless, the NeRF in the wild [128] considers the photometric and environmental variations between images to reconstruct real-world scenes. Compared with NeRF, Mip-NeRF [129] takes in a 3D Gaussian that represents the region over which the radiance field should be integrated. As a result, it can show the same scene at multiple levels of sharp detail. For video synthesis, Pumarola et al. [130] and Li et al. [131] extended the NeRF to dynamic objects with an additional parameter, time t .

The reconstruction and view synthesis processes involve constructing, rendering, and displaying. Obtaining 3D reconstruction through light fields provides

researchers with a bright outlook for real-time augmented reality and virtual reality. In addition, light fields provide insights into various view synthesis methods.

Industrial applications

As mentioned earlier, Lytro and Raytrix have industrialized light field acquisition devices. Other companies, such as FoVi3D [132] and Japan Display [133], have also produced light field projection solutions. Unlike for the Lytro camera, in which a microlens array is placed in front of an image sensor, Woptix [134] reduced the resolution trade-off by using a liquid lens [135] of the optical chain in front of the sensor to make it possible to change the focal planes quickly, providing full resolution of the sensor in real time. Furthermore, Google published many patents related to light field capturing [136–138]. It also published “Capturing Light Field Images with Uneven and/or Incomplete Angular Sampling” [139] in 2018. It designed a camera to capture light field images with uneven and incomplete angular sampling. The results showed an improvement in not only the spatial resolution but also the quality of the depth data.

For light field display, Avegant [140], Leia [141], Light Field Lab [142], Dimenco [143], and Creal [144] facilitated realistic digital photographs with display screens. Likewise, Looking Glass Factory [145] created a light field image display providing 45 different viewpoints as long as the viewer is within a 58° viewing cone, and it is far less expensive than previous products and, therefore, more attainable. In 2020, Sony published two light field-related technologies: 3D Spatial Reality Display Technology [146] and Atom View [147]. The former tracks the position of the eyes of the user, enabling the user to see real-world images or creations in 3D. In addition, it achieves a relatively high resolution and glasses-free 3D using real-time light field rendering technology. Atom View is an application in volumetric virtual production through point-cloud rendering, editing, and coloring. It can digitize space instead of a physical set and, therefore, can reproduce locations and sets. These light field display solutions are physical products, and light fields can also be applied in immersive online experiences. For instance, OppenFuture [148] provides Hololux Light Field Solution, which focuses on 3D reconstruction. It can reconstruct complex materials at full angle, and it is closely working with e-commerce companies to enhance the shopping experience. Furthermore, Google has produced its glasses-free light field display technology, Project Starline [149], which is used for real-time communication. People can communicate with each other as if they are sitting across from each other. However, Project Starline relies on custom-built hardware and highly specialized equipment.

The above products solve problems in two important parts of the light field imaging pipeline: acquisition and display. In the near future, one can expect more-portable devices for capturing light fields to emerge. In addition, the use of light field displays may extend from fixed screens to display extremely small or extremely large pictures and can further benefit medical microscopy or cinematic displays. Finally, the light field can contribute to closer-to-truth communication, which should make the “smart life” more attainable.

Conclusions

Depth estimation, which is essential for light field applications, was introduced. Then, the trend of light field applications was evaluated in terms of editing, enhancement, reconstruction, and current industrial products. The light field has been a research focus in computer graphics since 1996 and has progressed into the commercial market since 2010. Starting in 2010, the number of publications has increased rapidly, showing that many researchers are exploring the potential applications of light fields. These studies emphasize the critical role of the light field in enhancing visual experience. However, they require considerable expertise in utilizing light field technology. Therefore, there are still many human-light field interaction challenges to address not only in holography and augmented reality, but also in free image editing and interactive 3D reconstruction. Overall, light field imaging is commercially practical for businesses and individual users.

Abbreviations

1D: One-dimensional; 2D: Two-dimensional; 3D: Three-dimensional; 4D: Four-dimensional; 5D: Five-dimensional; AWS: Active wavefront sampling; CNN: Convolutional neural network; 6-DOF: Six-degree-of-freedom; EPI: Epipolar plane image; NeRF: Neural radiance field; POBR: Partially occluded boundary region; PSNR: Peak signal-to-noise ratio; RGB: Red, green, and blue; SAI: Sub-aperture images; SPO: Spinning parallelogram operator; SR: Super-resolution

Acknowledgements

The authors are grateful to Kuaishou Technology, University of California, Berkeley, and Tsinghua University.

Authors' contributions

All authors read and approved the final manuscript.

Funding

The last author was supported by the National Key R&D Program of China, No. 2019YFB1405703.

Availability of data and materials

All data analysed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

All authors give ethics approval and consent to participate.

Consent for publication

All authors give consent for publication.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Y-tech, Kuaishou Technology, Beijing 100085, China. ²EECS Department, University of California, Berkeley, CA 94720, USA. ³School of Software, BNRIST, Tsinghua University, Beijing, China.

Received: 6 July 2021 Accepted: 29 October 2021

Published online: 03 December 2021

References

1. Gershun A (1939) The light field. *J Math Phys* 18(1–4):51–151. <https://doi.org/10.1002/sapm193918151>
2. Adelson EH, Bergen JR (1991) The plenoptic function and the elements of early vision. In: Landy M, Movshon JA (eds) *Computational models of visual processing*. MIT Press, Cambridge, pp 3–20
3. Levoy M, Hanrahan P (1996) Light field rendering. In: *Abstracts of the 23rd annual conference on computer graphics and interactive techniques*. Association for Computing Machinery, New York. <https://doi.org/10.1145/237170.237199>
4. Fujii T (1994) A basic study on the integrated 3-D visual communication. Dissertation, The University of Tokyo
5. Fujii T, Kimoto T, Tanimoto M (1996) Ray space coding for 3D visual communication. Paper presented at the international picture coding symposium, Electronic Imaging, Melbourne, 13–15 March 1996
6. Tanimoto M (2006) Overview of free viewpoint television. *Signal Process Image Commun* 21(6):454–461. <https://doi.org/10.1016/j.image.2006.03.009>
7. Kubota A, Smolic A, Magnor M, Tanimoto M, Chen T, Zhang C (2007) Multiview imaging and 3DTV. *IEEE Signal Process Mag* 24(6):10–21. <https://doi.org/10.1109/MSP.2007.905873>
8. Ng R, Levoy M, Brédif M, Duval G, Horowitz M, Hanrahan P (2005) Light field photography with a hand-held plenoptic camera. Stanford University Computer Science Tech Report
9. Raytrix. 3D light-field camera technology. <https://raytrix.de/>. Accessed 31 May 2021
10. Wu GC, Masia B, Jarabo A, Zhang YC, Wang LY, Dai QH et al (2017) Light field image processing: an overview. *IEEE J Sel Top Signal Process* 11(7):926–954. <https://doi.org/10.1109/JSTSP.2017.2747126>
11. Wilburn BS, Smulski M, Lee H-HK, Horowitz MA (2002) Light field video camera. In: *Media Processors 2002*, vol 4674, pp 29–36. <https://doi.org/10.1117/12451074>
12. Yang JC, Everett M, Buehler C, McMillan L (2002) A real-time distributed light field camera. In: *Abstracts of the 13th eurographics workshop on rendering*. Eurographics Association, Pisa
13. Zhang C, Chen T (2004) A self-reconfigurable camera array. In: *Abstracts of the ACM SIGGRAPH 2004 sketches*. Association for Computing Machinery, New York. <https://doi.org/10.1145/1186223.1186412>
14. Chan SC, Ng KT, Gan ZF, Chan KL, Shum HY (2005) The plenoptic video. *IEEE Trans Circuits Syst Video Technol* 15(12):1650–1659. <https://doi.org/10.1109/TCSVT.2005.858616>
15. Liu YB, Dai QH, Xu WL (2006) A real time interactive dynamic light field transmission system. In: *Abstracts of the 2006 IEEE international conference on multimedia and expo*. IEEE, Toronto. <https://doi.org/10.1109/CME.2006.262686>
16. Venkataraman K, Lelescu D, Duparré J, McMahon A, Molina G, Chatterjee P et al (2013) PiCam: an ultra-thin high performance monolithic camera array. *ACM Trans Graph* 32(6):166. <https://doi.org/10.1145/2508363.2508390>
17. Lin X, Wu J, Zheng G, Dai Q (2015) Camera array based light field microscopy. *Biomed Opt Express* 6(9):3179–3189. <https://doi.org/10.1364/BOE.6.003179>
18. Unger J, Wenger A, Hawkins T, Gardner A, Debevec P (2003) Capturing and rendering with incident light fields. In: *Abstracts of the 14th eurographics workshop on rendering*. Eurographics Association, Leuven
19. Ihrke I, Stich T, Gottschlich H, Magnor M, Seidel HP (2008) Fast incident light field acquisition and rendering. *J WSCG* 16(1–3):25–32
20. Liang CK, Lin TH, Wong BY, Liu C, Chen HH (2008) Programmable aperture photography: multiplexed light field acquisition. *ACM Trans Graph* 27(8):1–10. <https://doi.org/10.1145/1399504.1360654>
21. Taguchi Y, Agrawal A, Ramalingam S, Veeraraghavan A (2010) Axial light field for curved mirrors: reflect your perspective, widen your view. In: *Abstracts of the 2010 IEEE computer society conference on computer vision and pattern recognition*. IEEE, San Francisco. <https://doi.org/10.1109/CVPR.2010.5540172>

22. Kim C, Zimmer H, Pritch Y, Sorkine-Hornung A, Gross M (2013) Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans Graph* 32(4):73. <https://doi.org/10.1145/2461912.2461926>
23. Dansereau DG, Schuster G, Ford J, Wetzstein G (2017) A wide-field-of-view monocentric light field camera. In: Abstracts of the IEEE conference on computer vision and pattern recognition. IEEE, Honolulu. <https://doi.org/10.1109/CVPR.2017.400>
24. Adelson EH, Wang JYA (1992) Single lens stereo with a plenoptic camera. *IEEE Trans Pattern Anal Mach Intell* 14(2):99–106. <https://doi.org/10.1109/34.121783>
25. Georgev T, Zheng KC, Curless B, Salesin D, Nayar S, Intwala C (2006) Spatio-angular resolution tradeoffs in integral photography. In: Abstracts of the 17th eurographics conference on rendering techniques. Eurographics Association, Aire-la-Ville
26. Levoy M, Ng R, Adams A, Footer M, Horowitz M (2006) Light field microscopy. *ACM Trans Graph* 25(3):924–934. <https://doi.org/10.1145/1141911.1141976>
27. Veeraraghavan A, Raskar R, Agrawal A, Mohan A, Tumblin J (2007) Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans Graph* 26(3):69–es. <https://doi.org/10.1145/1276377.1276463>
28. Lanman D, Raskar R, Agrawal A, Taubin G (2008) Shield fields: modeling and capturing 3D occluders. *ACM Trans Graph* 27(5):131. <https://doi.org/10.1145/1457515.1409084>
29. Horstmeyer R, Euliss G, Athale R, Levoy M (2009) Flexible multimodal camera using a light field architecture. In: Abstracts of the 2009 IEEE international conference on computational photography. IEEE, San Francisco. <https://doi.org/10.1109/ICCPHOT.2009.5559016>
30. Ashok A, Neifeld MA (2010) Compressive light field imaging. In: Abstracts of the three-dimensional imaging, visualization, and display 2010 and display technologies and applications for defense, security, and avionics IV. SPIE, Orlando. <https://doi.org/10.1117/12.852738>
31. Ihrke I, Wetzstein G, Heidrich W (2010) A theory of plenoptic multiplexing. In: Abstracts of the 2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, San Francisco. <https://doi.org/10.1109/CVPR.2010.5540174>
32. Manakov A, Restrepo JF, Klehm O, Hegedüs R, Eisemann E, Seidel HP et al (2013) A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging. *ACM Trans Graph* 32(4):47. <https://doi.org/10.1145/2461912.2461937>
33. Marwah K, Wetzstein G, Bando Y, Raskar R (2013) Compressive light field photography using overcomplete dictionaries and optimized projections. *ACM Trans Graph* 32(4):46. <https://doi.org/10.1145/2461912.2461914>
34. Cohen N, Yang S, Andalman A, Broxton M, Grosenick L, Deisseroth K et al (2014) Enhancing the performance of the light field microscope using wavefront coding. *Opt Express* 22(20):24817–24839. <https://doi.org/10.1364/OE.22.024817>
35. Wang YP, Wang LC, Kong DH, Yin BC (2015) High-resolution light field capture with coded aperture. *IEEE Trans Image Process* 24(12):5609–5618. <https://doi.org/10.1109/TIP.2015.2468179>
36. Wei LY, Liang CK, Myhre G, Pitts C, Akeley K (2015) Improving light field camera sample design with irregularity and aberration. *ACM Trans Graph* 34(4):152. <https://doi.org/10.1145/2766885>
37. Antipa N, Necula S, Ng R, Waller L (2016) Single-shot diffuser-encoded light field imaging. In: Abstracts of the 2016 IEEE international conference on computational photography. IEEE, Evanston. <https://doi.org/10.1109/ICCPHOT.2016.7492880>
38. Bishop TE, Favaro P (2011) Full-resolution depth map estimation from an aliased plenoptic light field. In: Kimmel R, Klette R, Sugimoto A (eds) *ACCV 2010: computer vision - ACCV 2010, 10th Asian conference on computer vision*, Queenstown, New Zealand, November 2010, Lecture notes in computer science, vol 6493. Springer, Berlin, pp 186–200. https://doi.org/10.1007/978-3-642-19309-5_15
39. Williem PIK, Lee KM (2018) Robust light field depth estimation using occlusion-noise aware data costs. *IEEE Trans Pattern Anal Mach Intell* 40(10):2484–2497. <https://doi.org/10.1109/TPAMI.2017.2746858>
40. Zhu H, Wang Q, Yu JY (2017) Occlusion-model guided antiocclusion depth estimation in light field. *IEEE J Sel Top Signal Process* 11(7):965–978. <https://doi.org/10.1109/JSTSP.2017.2730818>
41. Baradad M, Ye V, Yedidia AB, Durand F, Freeman WT, Wornell GW et al (2018) Inferring light fields from shadows. In: Abstracts of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Salt Lake City. <https://doi.org/10.1109/CVPR.2018.00656>
42. Chen J, Hou JH, Ni Y, Chau LP (2018) Accurate light field depth estimation with superpixel regularization over partially occluded regions. *IEEE Trans Image Process* 27(10):4889–4900. <https://doi.org/10.1109/TIP.2018.2839524>
43. Lueangwattana C, Mori S, Saito H (2019) Removing fences from sweep motion videos using global 3D reconstruction and fence-aware light field rendering. *Comput Vis Media* 5(1):21–32. <https://doi.org/10.1007/s41095-018-0126-8>
44. Bolles RC, Baker HH, Marimont DH (1987) Epipolar-plane image analysis: an approach to determining structure from motion. *Int J Comput Vis* 1(1):7–55. <https://doi.org/10.1007/BF00128525>
45. Matoššek M, Werner T, Hlavac V (2002) Accurate correspondences from epipolar plane images. In: Likar B (ed) *Proc. Computer Vision Winter Workshop*, pp 181–189
46. Criminisi A, Kang SB, Swaminathan R, Szeliski R, Anandan P (2005) Extracting layers and analyzing their specular properties using epipolar-plane-image analysis. *Comput Vis Image Underst* 97(1):51–85. <https://doi.org/10.1016/j.cviu.2004.06.001>
47. Wanner S, Goldluecke B (2012) Globally consistent depth labeling of 4D light fields. In: Abstracts of the 2012 IEEE conference on computer vision and pattern recognition. IEEE, Providence. <https://doi.org/10.1109/CVPR.2012.6247656>
48. Tao MW, Hadap S, Malik J, Ramamoorthi R (2013) Depth from combining defocus and correspondence using light-field cameras. In: Abstracts of the IEEE international conference on computer vision. IEEE, Sydney. <https://doi.org/10.1109/ICCV.2013.89>
49. Tao MW, Wang TC, Malik J, Ramamoorthi R (2014) Depth estimation for glossy surfaces with light-field cameras. In: Agapito L, Bronstein MM, Rother C (eds) *ECCV 2014: computer vision - ECCV 2014 workshops*, European conference on computer vision, Zurich, Switzerland, September 2014, Lecture notes in computer science, vol 8926. Springer, Cham, pp 533–547. https://doi.org/10.1007/978-3-319-16181-5_41
50. Zhang S, Sheng H, Li C, Zhang J, Xiong Z (2016) Robust depth estimation for light field via spinning parallelogram operator. *Comput Vis Image Underst* 145:148–159. <https://doi.org/10.1016/j.cviu.2015.12.007>
51. Sheng H, Zhao P, Zhang S, Zhang J, Yang D (2018) Occlusion-aware depth estimation for light field using multi-orientation EPIs. *Pattern Recogn* 74:587–599. <https://doi.org/10.1016/j.patcog.2017.09.010>
52. Schilling H, Diebold M, Rother C, Jähne B (2018) Trust your model: light field depth estimation with inline occlusion handling. In: Abstracts of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Salt Lake City. <https://doi.org/10.1109/CVPR.2018.00476>
53. Heber S, Yu W, Pock T (2017) Neural EPI-volume networks for shape from light field. In: Abstracts of the IEEE international conference on computer vision. IEEE, Venice. <https://doi.org/10.1109/ICCV.2017.247>
54. Heber S, Pock T (2016) Convolutional networks for shape from light field. In: Abstracts of the IEEE conference on computer vision and pattern recognition. IEEE, Las Vegas. <https://doi.org/10.1109/CVPR.2016.407>
55. Shin C, Jeon HG, Yoon Y, Kweon IS, Kim SJ (2018) EPINET: a fully-convolutional neural network using epipolar geometry for depth from light field images. In: Abstracts of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Salt Lake City. <https://doi.org/10.1109/CVPR.2018.00499>
56. Tsai YJ, Liu YL, Ouhyoung M, Chuang YY (2020) Attention-based view selection networks for light-field disparity estimation. *Proc AAAI Conf Artif Intell* 34(7):12095–12103. <https://doi.org/10.1609/aaai.v34i07.6888>
57. Chen JX, Zhang S, Lin YF (2021) Attention-based multi-level fusion network for light field depth estimation. *Proc AAAI Conf Artif Intell* 35(2):1009–1017
58. Johannsen O, Honauer K, Goldluecke B, Alperovich A, Battisti F, Bok Y et al (2017) A taxonomy and evaluation of dense light field depth estimation algorithms. In: Abstracts of the IEEE conference on computer vision and pattern recognition workshops. IEEE, Honolulu. <https://doi.org/10.1109/CVPRW.2017.226>
59. Mihara H, Funatomi T, Tanaka K, Kubo H, Mukaigawa Y, Nagahara H (2016) 4D light field segmentation with spatial and angular consistencies. In: Abstracts of the 2016 IEEE international conference on computational photography. IEEE, Evanston. <https://doi.org/10.1109/ICCPHOT.2016.7492872>
60. Dansereau DG, Pizarro O, Williams SB (2015) Linear volumetric focus for light field cameras. *ACM Trans Graph* 34(2):15–20. <https://doi.org/10.1145/2665074>

61. Jayaweera SS, Edussooriya CUS, Wijenayake C, Agathoklis P, Bruton LT (2020) Multi-volumetric refocusing of light fields. *IEEE Signal Process Lett* 28:31–35. <https://doi.org/10.1109/LSP.2020.3043990>
62. Wang YQ, Yang JG, Guo YL, Xiao C, An W (2019) Selective light field refocusing for camera arrays using bokeh rendering and superresolution. *IEEE Signal Process Lett* 26(1):204–208. <https://doi.org/10.1109/LSP.2018.2885213>
63. Yang JG, Xiao C, Wang YQ, An CJ, An W (2020) High-precision refocusing method with one interpolation for camera array images. *IET Image Process* 14(15):3899–3908. <https://doi.org/10.1049/iet-ipr.2019.0081>
64. Yang T, Zhang YN, Yu JY, Li J, Ma WG, Tong XM et al (2014) All-in-focus synthetic aperture imaging. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds) *ECCV 2014: computer vision - ECCV 2014, 13th European conference on computer vision, Zurich, Switzerland, September 2014, Lecture notes in computer science, vol 8694*. Springer, Cham, pp 1–15. https://doi.org/10.1007/978-3-319-10599-4_1
65. Wang YQ, Wu TH, Yang JG, Wang LG, An W, Guo YL (2020) DeOccNet: learning to see through foreground occlusions in light fields. In: Abstracts of the IEEE winter conference on applications of computer vision. IEEE, Snowmass. <https://doi.org/10.1109/WACV45572.2020.9093448>
66. Li YJ, Yang W, Xu ZB, Chen Z, Shi ZB, Zhang Y et al (2021) Mask4D: 4D convolution network for light field occlusion removal. In: Abstracts of the IEEE international conference on acoustics, speech and signal processing. IEEE, Toronto. <https://doi.org/10.1109/ICASSP39728.2021.9413449>
67. Berent J, Dragotti PL (2007) Unsupervised extraction of coherent regions for image based rendering. In: Abstracts of the British machine vision conference. BMVA Press, Coventry. <https://doi.org/10.5244/C.21.28>
68. Sethian JA (1996) Theory, algorithms, and applications of level set methods for propagating interfaces. *Acta Numer* 5:309–395. <https://doi.org/10.1017/S0962492900002671>
69. Wanner S, Straehle C, Goldluecke B (2013) Globally consistent multi-label assignment on the ray space of 4D light fields. In: Abstracts of the IEEE conference on computer vision and pattern recognition. IEEE, Portland. <https://doi.org/10.1109/CVPR.2013.135>
70. Batra D, Kowdle A, Parikh D, Luo JB, Chen T (2010) iCoseg: interactive co-segmentation with intelligent scribble guidance. In: Abstracts of the 2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, San Francisco. <https://doi.org/10.1109/CVPR.2010.5540080>
71. Xu YC, Nagahara H, Shimada A, Taniguchi RI (2015) TransCut: transparent object segmentation from a light-field image. In: Abstracts of the IEEE international conference on computer vision. IEEE, Santiago. <https://doi.org/10.1109/ICCV.2015.393>
72. Ren XF, Malik J (2003) Learning a classification model for segmentation. In: Abstracts of the 9th IEEE international conference on computer vision. IEEE, Nice
73. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S (2012) SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal Mach Intell* 34(11):2274–2282. <https://doi.org/10.1109/TPAMI.2012.120>
74. Zhu H, Zhang Q, Wang Q (2017) 4D light field superpixel and segmentation. In: Abstracts of the IEEE conference on computer vision and pattern recognition. IEEE, Honolulu. <https://doi.org/10.1109/CVPR.2017.710>
75. Horn DR, Chen B (2007) LightShop: interactive light field manipulation and rendering. In: Abstracts of the 2007 symposium on interactive 3D graphics and games. Association for Computing Machinery, Washington. <https://doi.org/10.1145/1230100.1230121>
76. Wolberg G (1990) *Digital image warping*. IEEE Computer Society Press, Los Alamitos
77. Jarabo A, Masia B, Bousseau A, Pellacini F, Gutierrez D (2014) How do people edit light fields? *ACM Trans Graph* 33(4):146. <https://doi.org/10.1145/2601097.2601125>
78. Perwaß C, Wietzke L (2012) Single lens 3D-camera with extended depth-of-field. In: Abstracts of the human vision and electronic imaging XVII. SPIE, Burlingame. <https://doi.org/10.1117/12.909882>
79. Chandramouli P, Favaro P, Perrone D (2014) Motion deblurring for plenoptic images. arXiv preprint arXiv: 1408.3686
80. Jin MG, Chandramouli P, Favaro P (2015) Bilayer blind deconvolution with the light field camera. In: Abstracts of the IEEE international conference on computer vision workshops. IEEE, Santiago. <https://doi.org/10.1109/ICCVW.2015.36>
81. Srinivasan PP, Ng R, Ramamoorthi R (2017) Light field blind motion deblurring. In: Abstracts of the IEEE conference on computer vision and pattern recognition. IEEE, Honolulu. <https://doi.org/10.1109/CVPR.2017.253>
82. Lee D, Park H, Park IK, Lee KM (2018) Joint blind motion deblurring and depth estimation of light field. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) *ECCV 2018: computer vision - ECCV 2018, 15th European conference on computer vision, Munich, Germany, September 2018, Lecture notes in computer science, vol 11220*. Springer, Cham, pp 300–316. https://doi.org/10.1007/978-3-030-01270-0_18
83. Lumentut JS, Kim TH, Ramamoorthi R, Park IK (2019) Fast and full-resolution light field deblurring using a deep neural network. arXiv preprint arXiv: 1904.00352
84. Dansereau DG, Eriksson A, Leitner J (2017) Richardson-Lucy deblurring for moving light field cameras. In: Abstracts of the IEEE conference on computer vision and pattern recognition workshops. IEEE, Honolulu. <https://doi.org/10.1109/CVPRW.2017.225>
85. Levin A, Freeman WT, Durand F (2008) Understanding camera trade-offs through a bayesian analysis of light field projections. In: Forsyth D, Torr P, Zisserman A (eds) *ECCV 2008: computer vision - ECCV 2008, 10th European conference on computer vision, Marseille, France, October 2008, Lecture notes in computer science, vol 5305*. Springer, Berlin, pp 88–101. https://doi.org/10.1007/978-3-540-88693-8_7
86. Bishop TE, Zanetti S, Favaro P (2009) Light field superresolution. In: Abstracts of the 2009 IEEE international conference on computational photography. IEEE, San Francisco. <https://doi.org/10.1109/ICCPHOT.2009.5559010>
87. Zhou SB, Yuan Y, Su LJ, Ding XM, Wang JC (2017) Multiframe super resolution reconstruction method based on light field angular images. *Opt Commun* 404:189–195. <https://doi.org/10.1016/j.optcom.2017.03.019>
88. Lumsdaine A, Georgiev T (2008) Full resolution lightfield rendering. *Indiana Univ Adobe Syst, Tech Rep* 91:92
89. Zheng HT, Ji MQ, Wang HQ, Liu YB, Fang L (2018) CrossNet: an end-to-end reference-based super resolution network using cross-scale warping. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) *ECCV 2008: computer vision - ECCV 2008, 15th European conference on computer vision, Munich, Germany, September 2018, Lecture notes in computer science, vol 11210*. Springer, Cham, pp 88–104. https://doi.org/10.1007/978-3-030-01231-1_6
90. Cheng Z, Xiong ZW, Chen C, Liu D (2019) Light field super-resolution: a benchmark. In: Abstracts of the IEEE/CVF conference on computer vision and pattern recognition workshops. IEEE, Long Beach. <https://doi.org/10.1109/CVPRW.2019.00231>
91. Georgiev T, Chunev G, Lumsdaine A (2011) Superresolution with the focused plenoptic camera. In: Abstracts of the computational imaging IX. SPIE, San Francisco Airport. <https://doi.org/10.1117/12.872666>
92. Liang CK, Ramamoorthi R (2015) A light transport framework for lenslet light field cameras. *ACM Trans Graph* 34(2):16–19. <https://doi.org/10.1145/2665075>
93. Lim J, Ok H, Park B, Kang J, Lee S (2009) Improving the spatio resolution based on 4D light field data. In: Abstracts of the 16th IEEE international conference on image processing. IEEE, Cairo
94. Nava FP, Luke JP (2009) Simultaneous estimation of super-resolved depth and all-in-focus images from a plenoptic camera. In: Abstracts of the 2009 3DTV conference: the true vision-capture, transmission and display of 3D video. IEEE, Potsdam
95. Alain M, Smolic A (2018) Light field super-resolution via LFBM5D sparse coding. In: Abstracts of the 25th IEEE international conference on image processing. IEEE, Athens. <https://doi.org/10.1109/ICIP.2018.8451162>
96. Bishop TE, Favaro P (2012) The light field camera: extended depth of field, aliasing, and superresolution. *IEEE Trans Pattern Anal Mach Intell* 34(5):972–986. <https://doi.org/10.1109/TPAMI.2011.168>
97. Farag S, Velisavljevic V (2018) A novel disparity-assisted block matching-based approach for super-resolution of light field images. In: Abstracts of the 2018-3DTV-conference: the true vision-capture, transmission and display of 3D video. IEEE, Helsinki. <https://doi.org/10.1109/3DTV.2018.8478627>
98. Mitra K, Veeraraghavan A (2012) Light field denoising, light field superresolution and stereo camera based refocusing using a GMM light field patch prior. In: Abstracts of the 2012 IEEE computer society conference on computer vision and pattern recognition workshops. IEEE, Providence. <https://doi.org/10.1109/CVPRW.2012.6239346>
99. Rossi M, El Gheche M, Frossard P (2018) A nonsmooth graph-based approach to light field super-resolution. In: Abstracts of the 25th IEEE international conference on image processing. IEEE, Athens. <https://doi.org/10.1109/ICIP.2018.8451127>
100. Rossi M, Frossard P (2017) Graph-based light field super-resolution. In: Abstracts of the 19th international workshop on multimedia signal processing. IEEE, Luton. <https://doi.org/10.1109/MMSP.2017.8122224>

101. Wanner S, Goldluecke B (2014) Variational light field analysis for disparity estimation and super-resolution. *IEEE Trans Pattern Anal Mach Intell* 36(3): 606–619. <https://doi.org/10.1109/TPAMI.2013.147>
102. Fan HZ, Liu D, Xiong ZW, Wu F (2017) Two-stage convolutional neural network for light field super-resolution. In: Abstracts of the 2017 IEEE international conference on image processing. IEEE, Beijing. <https://doi.org/10.1109/ICIP.2017.8296465>
103. Farrugia RA, Galea C, Guillemot C (2017) Super resolution of light field images using linear subspace projection of patch-volumes. *IEEE J Sel Top Signal Process* 11(7):1058–1071. <https://doi.org/10.1109/JSTSP.2017.2747127>
104. Gul MSK, Gunturk BK (2018) Spatial and angular resolution enhancement of light fields using convolutional neural networks. *IEEE Trans Image Process* 27(5):2146–2159. <https://doi.org/10.1109/TIP.2018.2794181>
105. Wang YL, Liu F, Zhang KB, Hou GQ, Sun ZN, Tan TN (2018) LFNNet: a novel bidirectional recurrent convolutional neural network for light-field image super-resolution. *IEEE Trans Image Process* 27(9):4274–4286. <https://doi.org/10.1109/TIP.2018.2834819>
106. Yoon Y, Jeon HG, Yoo D, Lee JY, Kweon IS (2015) Learning a deep convolutional network for light-field image super-resolution. In: Abstracts of the IEEE international conference on computer vision workshops. IEEE, Santiago. <https://doi.org/10.1109/ICCVW.2015.17>
107. Yuan Y, Cao ZQ, Su LJ (2018) Light-field image superresolution using a combined deep CNN based on EPI. *IEEE Signal Process Lett* 25(9):1359–1363. <https://doi.org/10.1109/LSP.2018.2856619>
108. Farrugia RA, Guillemot C (2020) Light field super-resolution using a low-rank prior and deep convolutional neural networks. *IEEE Trans Pattern Anal Mach Intell* 42(5):1162–1175. <https://doi.org/10.1109/TPAMI.2019.2893666>
109. Zhang S, Lin YF, Sheng H (2019) Residual networks for light field image super-resolution. In: Abstracts of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Long Beach. <https://doi.org/10.1109/CVPR.2019.01130>
110. Wang YQ, Wang LG, Yang JG, An W, Yu JY, Guo YL (2020) Spatial-angular interaction for light field image super-resolution. In: Vedaldi A, Bischof H, Brox T, Frahm JM (eds) ECCV 2020: computer vision - ECCV 2020, 16th European conference on computer vision, Glasgow, United Kingdom, August 2020, Lecture notes in computer science, vol 12368. Springer, Cham, pp 290–308. https://doi.org/10.1007/978-3-030-58592-1_18
111. Wang YQ, Yang JG, Wang LG, Ying XY, Wu TH, An W et al (2021) Light field image super-resolution using deformable convolution. *IEEE Trans Image Process* 30:1057–1071. <https://doi.org/10.1109/TIP.2020.3042059>
112. Ivan A, Williem PIK (2020) Joint light field spatial and angular super-resolution from a single image. *IEEE Access* 8:112562–112573. <https://doi.org/10.1109/ACCESS.2020.3002921>
113. Jin J, Hou JH, Chen J, Kwong S (2020) Light field spatial super-resolution via deep combinatorial geometry embedding and structural consistency regularization. In: Abstracts of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Seattle. <https://doi.org/10.1109/CVPR42600.2020.00233>
114. Shi LX, Hassaneih H, Davis A, Katabi D, Durand F (2014) Light field reconstruction using sparsity in the continuous Fourier domain. *ACM Trans Graph* 34(1):12–13. <https://doi.org/10.1145/2682631>
115. Lanman D, Crispell D, Taubin G (2009) Surround structured lighting: 3-D scanning with orthographic illumination. *Comput Vis Image Underst* 113(11):1107–1117. <https://doi.org/10.1016/j.cviu.2009.03.016>
116. Heber S, Ranftl R, Pock T (2013) Variational shape from light field. In: Heyden A, Kahl F, Olsson C, Oskarsson M, Tai XC (eds) EMMCVPR 2013: energy minimization methods in computer vision and pattern recognition, 9th international conference on energy minimization methods in computer vision and pattern recognition, Lund, Sweden, August 2013, Lecture notes in computer sciencel, vol 8081. Springer, Berlin, pp 66–79. https://doi.org/10.1007/978-3-642-40395-8_6
117. Frigerio F (2006) 3-dimensional surface imaging using active wavefront sampling. Dissertation, Massachusetts Institute of Technology
118. Feng MT, Gilani SZ, Wang YN, Mian A (2018) 3D face reconstruction from light field images: a model-free approach. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) ECCV 2018: computer vision - ECCV 2018, 15th European conference on computer vision (ECCV), Munich, Germany, September 2018, Lecture notes in computer science, vol 11214. Springer, Cham, pp 508–526. https://doi.org/10.1007/978-3-030-01249-6_31
119. Zhang YL, Li Z, Yang W, Yu PH, Lin HT, Yu JY (2017) The light field 3D scanner. In: Abstracts of the 2017 IEEE international conference on computational photography. IEEE, Stanford. <https://doi.org/10.1109/ICCPHOT.2017.7951484>
120. Overbeck RS, Erickson D, Evangelakos D, Pharr M, Debevec P (2018) A system for acquiring, processing, and rendering panoramic light field stills for virtual reality. *ACM Trans Graph* 37(6):197. <https://doi.org/10.1145/3272127.3275031>
121. Chai JX, Tong X, Chan SC, Shum HY (2000) Plenoptic sampling. In: Abstracts of the 27th annual conference on computer graphics and interactive techniques. ACM Press, New Orleans. <https://doi.org/10.1145/344779.344932>
122. Qiu WC, Zhong FW, Zhang Y, Qiao SY, Xiao ZH, Kim TS et al (2017) UnreaL.V: virtual worlds for computer vision. In: Abstracts of the 25th ACM international conference on multimedia. Association for Computing Machinery, Mountain View. <https://doi.org/10.1145/3123266.3129396>
123. Kazhdan M, Hoppe H (2013) Screened poisson surface reconstruction. *ACM Trans Graph* 32(3):29. <https://doi.org/10.1145/2487228.2487237>
124. Shi L, Huang FC, Lopes W, Matusik W, Luebke D (2017) Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3D computer graphics. *ACM Trans Graph* 36(6):236. <https://doi.org/10.1145/3130800.3130832>
125. Mildenhall B, Srinivasan PP, Ortiz-Cayon R, Kalantari NK, Ramamoorthi R, Ng R et al (2019) Local light field fusion: practical view synthesis with prescriptive sampling guidelines. *ACM Trans Graph* 38(4):29. <https://doi.org/10.1145/3306346.3322980>
126. Mildenhall B, Srinivasan PP, Tancik M, Barron JT, Ramamoorthi R, Ng R (2020) NeRF: representing scenes as neural radiance fields for view synthesis. In: Vedaldi A, Bischof H, Brox T, Frahm JM (eds) ECCV 2020: computer vision - ECCV 2020, 16th European conference on computer vision, Glasgow, United Kingdom, August 2020, Lecture notes in computer science, vol 12346. Springer, Cham, pp 405–421. https://doi.org/10.1007/978-3-030-58452-8_24
127. Park K, Sinha U, Barron JT, Bouaziz S, Goldman DB, Seitz SM et al (2020) Nerfies: deformable neural radiance fields. arXiv preprint arXiv: 2011.12948
128. Martin-Brualla R, Radwan N, Sajjadi MSM, Barron JT, Dosovitskiy A, Duckworth D (2021) Nerf in the wild: neural radiance fields for unconstrained photo collections, Paper presented at the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Nashville
129. Barron JT, Mildenhall B, Tancik M, Hedman P, Martin-Brualla R, Srinivasan PP (2021) Mip-nerf: a multiscale representation for anti-aliasing neural radiance fields. arXiv preprint arXiv: 2103.13415
130. Pumarola A, Corona E, Pons-Moll G, Moreno-Noguer F (2021) D-nerf: neural radiance fields for dynamic scenes, Paper presented at the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Nashville
131. Li ZQ, Niklaus S, Snavely N, Wang O (2021) Neural scene flow fields for space-time view synthesis of dynamic scenes, Paper presented at the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Nashville
132. FoV3D - products. <https://www.fovi3d.com/ffd>. Accessed 31 May 2021
133. JDI's latest display technology. <https://www.j-display.com>. Accessed 31 May 2021
134. Wootpix - the ultimate image solutions. <https://wootpix.com/>. Accessed 31 May 2021
135. Kuiper S, Hendriks BHW (2004) Variable-focus liquid lens for miniature cameras. *Appl Phys Lett* 85(7):1128–1130. <https://doi.org/10.1063/1.1779954>
136. Ng YR, Cheng E, Liang CK, Fatahalian K, Evans DJ, Wampller K et al (2018) Depth-assigned content for depth-enhanced virtual reality images. US Patent 10,129,524, 13 Nov 2018
137. Kuang JT, Liang CK (2018) Automatic lens flare detection and correction for light-field images. US Patent 9,979,909, 22 May 2018
138. Knight TJ, Pitts C, Ng YR, Fishman A, Romanenko Y, Kalt J et al (2015) Light-field processing and analysis, camera control, and user interfaces and interaction on light-field capture devices. US Patent 8,995,785, 31 Mar 2015
139. Pitts C, Liang CK, Akeley K (2018) Capturing light-field images with uneven and/or incomplete angular sampling. US Patent 10,033,986, 24 Jul 2018
140. Avegant - engineering. <https://www.avegant.com>. Accessed 31 May 2021
141. LeiaPix. <https://leiainc.com.cn/platform/leia-loft/>. Accessed 31 May 2021
142. Light field lab. <https://www.lightfieldlab.com/>. Accessed 31 May 2021
143. Simulated reality - 3D display technology. <https://www.dimenso.eu/>. Accessed 31 May 2021
144. Creal. <https://www.creal.com/>. Accessed 31 May 2021
145. Looking glass factory. <https://lookingglassfactory.com/>. Accessed 31 May 2021
146. Sony corporation - spatial reality display - about spatial reality display. <https://www.sony.net/Products/Developer-Spatial-Reality-display/en/develop/AboutSRDisplay.html>. Accessed 31 May 2021

147. Sony innovation studios - Sony pictures. <https://www.sonyinnovationstudios.com/>. Accessed 31 May 2021
148. Opendtech - Hololux™ light field reconstruction solutions. <https://www.opendtech.com/en>. Accessed 31 May 2021
149. Project starline: feel like you're there, together. <https://blog.google/technology/research/project-starline/>. Accessed 31 May 2021

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
