



Review

Lessons Learned from Implementing Light Field Camera Animation: Implications, Limitations, Potentials, and Future Research Efforts

Mary Guindy ^{1,*}  and Peter A. Kara ² 

¹ Faculty of Information Technology and Bionics, Pazmany Peter Catholic University, Práter Str. 50/A, 1083 Budapest, Hungary

² Department of Networked Systems and Services, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Műegyetem rkp. 3, 1111 Budapest, Hungary; kara@hit.bme.hu

* Correspondence: guindy.mary.mohsen.messak@itk.ppke.hu

Abstract: Among the novel 3D visualization technologies of our era, light field displays provide the complete 3D visual experience without the need for any personal viewing device. Due to the lack of such constraint, these displays may be viewed by any number of observers simultaneously, and the corresponding use case contexts may also involve a virtually unlimited numbers of users; any number that the valid viewing area of the display may accommodate. While many instances of the utilization of this technology operate with static contents, camera animation may also be relevant. While the topic of light field camera animation has already been addressed on an initial level, there are still numerous research efforts to be carried out. In this paper, we elaborate on the lessons learned from implementing light field camera animation. The paper discusses the associated implications, limitations, potentials, and future research efforts. Each of these areas are approached from the perspectives of use cases, visual content, and quality assessment, as well as capture and display hardware. Our work highlights the existing research gaps in the investigated topic, the severe issues related to visualization sharpness, and the lack of appropriate datasets, as well as the constraints due to which novel contents may be captured by virtual cameras instead of real capture systems.

Keywords: light field; camera animation; visual quality; use case; hardware



Citation: Guindy, M.; Kara, P.A.

Lessons Learned from Implementing Light Field Camera Animation: Implications, Limitations, Potentials, and Future Research Efforts.

Multimodal Technol. Interact. **2024**, *8*, 68.
<https://doi.org/10.3390/mti8080068>

Academic Editor: Cristina Portales

Received: 28 June 2024

Revised: 29 July 2024

Accepted: 30 July 2024

Published: 1 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The evolution of 3D visualization technologies in the last century continuously and steadily strengthened the need to represent 3D information. Similar to how the eyes perceive the real world via two distinct perspectives (i.e., stereopsis), stereoscopic devices capture two views of the scene or object of interest. These views are then visualized on stereoscopic displays with the aid of viewing gears to recreate 3D perception. However, not every single 3D visualization technology relies on the classic concept of stereoscopic content provision (i.e., one perspective per eye). In the scientific literature, the visualization of stereoscopic contents that do not require viewing devices is called autostereoscopy. However, this phenomenon does not cover solutions that are not based on stereoscopic pairs.

Unlike the case of 2D displays—where the visual content is perceived from the same perspective regardless of the viewing angle—light field (LF) visualization enables the 3D perception of the content with the correct corresponding perspective based on the viewing angle, without the need of any personal viewing device (e.g., special glasses or headgears). This statement is somewhat applicable to 3D multiview displays as well, yet they only provide such visualization over a very limited viewing angle—also referred to as “sweet spots”—and the same interval of perspectives repeat over these viewing positions. In contrast, light field displays (LFDs) utilize the entire field of view (FOV) of the system to provide a continuous change in perspective and a smooth parallax effect.

At the time of writing this paper, LFDs have not yet emerged on the consumer market, although certain implementations are commercially available. However, the scientific community develops novel solutions (i.e., different implementations of LF visualization) and uses the currently available state-of-the-art displays to carry out numerous research efforts in a myriad of topics that need to be addressed in order to pave the way towards future use cases. A high portion of such use cases primarily rely on static contents. The term “static” in this context refers to both the LF content itself (i.e., the content does not change over time) and its presentation (i.e., the parameters of the camera do not change over time).

There are many other use cases—including interactive scenarios—that rely on or benefit from camera animation, which is the gradual change in camera parameters. Typical examples of this include the change in zoom level, camera position, and camera orientation. However, implementing LF camera animation is not necessarily a straightforward task, as many factors—such as capture configuration or the human implications of visualization (i.e., impacts on the visual experience)—may complicate its success and efficiency.

Recently, we implemented camera animation on a large-scale LFD [1–3] and assessed the achieved results. Beyond the initial output of such a research effort, there are numerous aspects that need to be considered in order to successfully integrate camera animation into LF use case contexts; in essence, the road ahead is still long.

In this paper, the lessons learned from implementing LF camera animations are discussed in detail with regards to different aspects. First of all, we elaborate the implications of LF camera animation: what the currently available technology and the state-of-the-art findings imply. This is followed by the associated limitations: constraints that may either make progress and practical deployment challenging or even prohibitive. Then, we discuss the potentials of the investigated field: advantages and benefits enabled by employing LF camera animation. Finally, we specify future research efforts: the upcoming steps that need to be taken to make LF camera animation not only a feasible but an efficient method for the future deployment of LF visualization technology.

Instead of a generic approach to the four aforementioned topics, this paper discusses each of them through the perspective of four relevant contexts. First and foremost, the use cases that are meaningful and feasible for LF camera animation are analyzed. In essence, the potential utilization contexts determine the success of the emergence of LF visualization in the future. Another important—and closely related—area of LF technology is the LF content itself that is captured and visualized. We also address LF quality assessment, which is a significant portion of the published and ongoing research efforts. Additionally, the paper covers hardware-related considerations for both capture and display. Again, the primary aspect of each section is investigated through these four contexts. The structure of our contribution is summarized in Figure 1.

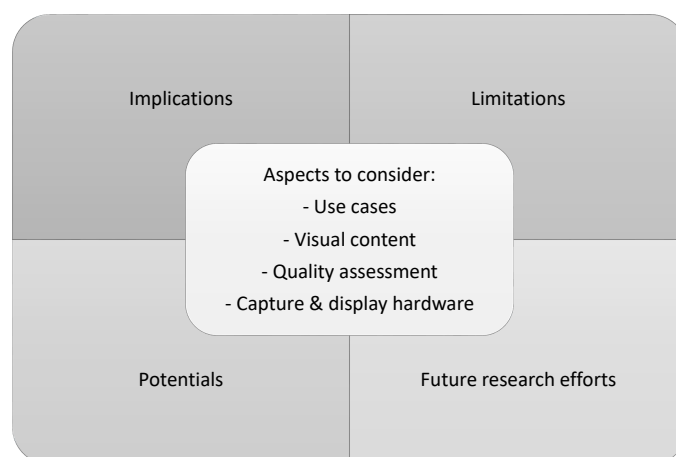


Figure 1. Structure of contribution.

The main contribution of this review paper is to provide a thorough understanding regarding the implications, limitations, potentials, and future research efforts related to LF camera animation. Its goal is to support the future research and development in LF technology and, ultimately, to aid the emergence of LF camera animation. For example, the future research efforts related to use cases, visual content, quality assessment, and hardware specify gaps in the scientific literature that should be bridged in order to support the successful deployment of LF camera animation. A novelty of the work is that it touches upon topics with respect to LF camera animation that are important but still in early phases of implementation, such as split-domain LF visualization, angular super resolution, and the visual Turing test.

The remainder of this paper is structured as follows. The methodology employed by our work is briefly overviewed in Section 2. The trends and topics that have characterized LF research since the Millennium (i.e., the year 2000) are analyzed in Section 3. The state-of-the-art scientific literature is reviewed in Section 4. Our previous research effort on implementing LF camera animation is summarized in Section 5. The implications, limitations, potentials, and future research efforts are discussed in Sections 6, 7, 8 and 9, respectively. The limitations of the work itself are stated in Section 10, and the paper is concluded in Section 11.

2. Materials and Methods

Our work is comprised of two main approaches, each with their own appropriate methodologies. The first one is the analysis of research trends and topics within the domain of LF technology. We investigated published works—scientific articles in particular—from the year 2000 to 2023. The analysis was performed on the well-known Scopus database. A total of 50 queries was performed, ranging from rather generic (e.g., LF imaging) to very specific (e.g., LF interpolation) research topics. The queries considered the fact that in articles, the scientific term “light field” may also appear as “light-field”, “lightfield”, or simply as “LF” (such as in this very article). This variation extends to terms such as “dataset” as well, which may be spelled as “data-set” or phrased as “database”. Therefore, as an example, the query for LF datasets was the following: “light field dataset” OR “lightfield dataset” OR “LF dataset” OR “light field data set” OR “lightfield data set” OR “LF data set” OR “light field database” OR “lightfield database” OR “LF database”. Note that writing keywords with or without hyphens yields the same results, as Scopus does not differentiate them. The results of the analysis of research trends and topics are introduced in Section 3.

The second one is a conventional in-depth analysis of the related scientific literature. For this purpose, beyond relying on Scopus, we searched relevant databases, such as IEEE. In this analytical context, the yields of the various research topics are less important. In fact, as stated before, while the low yields do serve as a notable limitation, they also emphasize the under-investigated nature of LF camera animation. Additionally, in contrast to the analysis of research trends and topics that relies only on published articles, the detailed review of the relevant scientific literature includes other forms of publication as well, such as conference papers.

3. Analysis of Research Trends and Topics

As already mentioned in the previous section, we performed a total of 50 queries within the scope of this analysis. The first ones investigated the current state of terminologies—i.e., how the term “light field” appears in the scientific literature. We found that 98.94% of the published articles use either “light field” or “light-field”. However, this percentage may be somewhat misleading, as this very article would fall into this category as well, even though we consistently use “LF” throughout the paper. The reason for this is that the term appears as “light field” in the title, and also appears within the text—such as in this very sentence. Therefore, it is a more accurate statement that 1.06% of all the LF-related articles write “LF” exclusively.

Regarding the topics covered by the scientific literature and their frequencies, we investigated the following LF-related terms (and all the possible variations, as stated ear-

lier): image; imaging; camera; display; model; super resolution; reconstruction; dataset; application; rendering; quality assessment; system; analysis; manipulation; compression, acquisition; depth estimation; video; photography; feature; capture; algorithm; enhancement; processing; geometry; synthesis; refocusing; generation; use case; visualization; estimation; editing; calibration; optimization; denoising; resolution; interpolation; segmentation; method; registration; subjective test; depth map; integration; computation; workflow; and camera animation. Additionally, we performed queries for camera animation in general as well.

The analysis of the LF-related research topics that yield at least 100 articles in total for the investigated period is presented in Figure 2. First of all, it is applicable to the vast majority of the research topics that there has been a notable increase since 2013. This may be primarily due to the fact that it was at that time the first commercial LFDs emerged, which enabled numerous research efforts, such as quality assessment. However, quality assessment is not represented in the figure, as there were only 64 articles in total published on the topic by the end of 2023. The eight research topics that are represented in the figure are (in descending order of total articles) image, imaging, camera, display, model, super resolution, reconstruction, and dataset. The first two are relatively generic research terms, which explains their frequent occurrences. While “image” simply refers to static LF contents, “imaging” represents the process of creating LF contents suitable for visualization (i.e., visual reproduction). Regarding more specific topics, it can be stated that since 2015, in every single year, more articles were published on LF cameras than on LFDs. Super resolution is also among the most frequent research topics, yet it should be noted that this term covers both high angular density and resolution enhancement. The works listed by the queries dominantly refer to the latter. In Section 4, the different research efforts assigned to the term “super resolution” are thoroughly explained. It is also important to highlight that there is a noteworthy increase in novel LF datasets, which are crucial to a great number of research activities, such as developing new metrics of objective quality assessment, performing subjective studies, validating use cases, and many more.

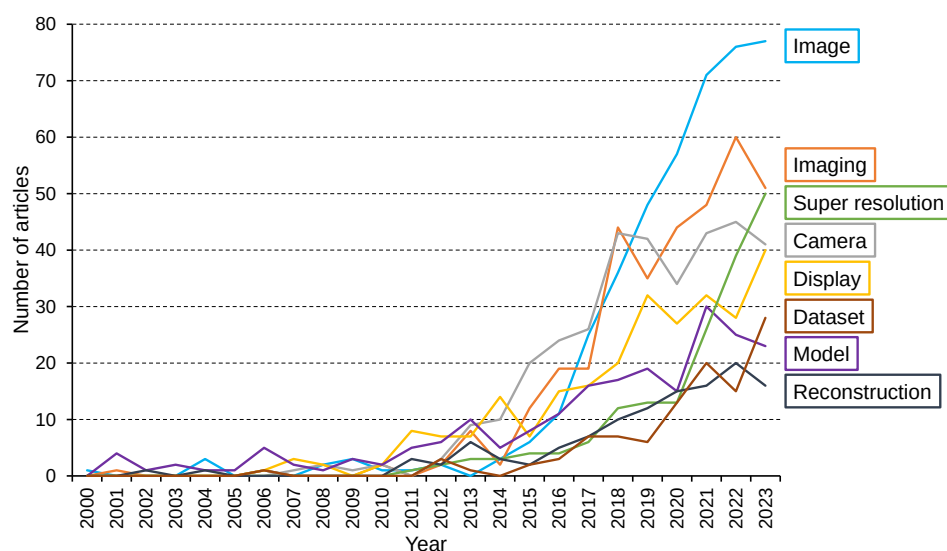


Figure 2. Scopus analysis of LF-related research topics from 2000 to 2024 with at least 100 articles in total. The topics in descending order of the total number of articles are the following: image (i.e., static LF content), imaging (i.e., visual reproduction), camera (i.e., LF acquisition), display (i.e., LF visualization), model (i.e., LF object), super resolution (i.e., resolution enhancement), reconstruction (i.e., generating continuous LF from samples), and dataset (i.e., novel LF contents).

Generally, LF-related research is very steadily increasing. While between 2000 and 2010, the annual number of new articles was between 100 and 200, it reached 300 by 2016, 500 by 2019, and 700 by 2022. However, these trends are not applicable to LF camera animation. According to the obtained results, only a single article was published in the investigated period—which is our own earlier work from 2022. However, camera animation in general is also an under-investigated research area, yielding only eight works. Naturally, camera animation as a generic topic is not of much research interest in this day and age. However, applying it to such novel technology holds potential benefits, which are later elaborated in this paper, along with the implications, limitations, and future research efforts.

4. Related Work

In this section, we revise the efforts and published works of the scientific community that are related to our contribution. First, we provide a general overview of the theoretical framework behind LF representation. This is followed by the classification of the relevant hardware, and the explanation of the associated concepts, including rendering. Since compression is one of the most important research areas of LF technology—and thus, relevant to every instance of utilization—the section then elaborates the contemporary solutions of LF compression in a chronological order. The focus is then narrowed down to user-centric considerations, due to the fact that user experience ultimately determines the success of the deployment of LF use cases—including utilization contexts based on LF camera animation. As such subjective studies often rely on LF datasets, the section also briefly reviews the databases available at the time of writing this paper. Finally, the section addresses super resolution and distinguishes it from angular super resolution, both of which are crucial research topics for the future emergence of LF technology. The related research efforts on LF camera animation are discussed in a subsequent separate section, namely, Section 5.

The technical term “light field” was first introduced by Gershun in 1936 [4]. Nonetheless, the original concept was specified more than a century ago as a means to encapsulate the visual information of the physical world, while evolving throughout the years along with the advancements of digital and optical technologies [5–8]. LF describes the radiance at a point in a certain direction [9]. In other words, it is “the amount of light traveling in every direction through every point in space” [10]. One of the major advantages of LFs is their ability to improve the comprehension of how the human visual system (HVS) interprets the surrounding world. Accordingly, LF images provide tremendous amounts of visual information regarding the represented scenes, as they describe the light traversing in all directions for all the points of 3D space. Therefore, unlike conventional photography—which only captures a 2D image—LF imaging demands the acquisition of multi-dimensional data (i.e., spatial and angular information) [11,12]. LF in general is meant to represent a portion of 3D space. In terms of visualization, this means that we can imagine a plane in front of and behind the visualized content. In order to characterize the light rays within such a constrained portion of 3D space, one may take the two coordinate pairs on these parallel planes where the line representing the light ray intersects them [9]—which, as shown later, is already a simplification of LF representation [13,14]. This, however, poses a notable limitation on the scope of LF visualization, since the constrained portion of 3D space is indeed finite, which means that it is impossible to visualize portions of a content that are virtually infinitely far away.

As a means to represent LF, the plenoptic function was proposed by Adelson and Bergen [15], describing everything that can be perceived within a given segment of space, hence the name “plenoptic” (made up from “plenus” meaning complete, and “optic”). The plenoptic function is a 7D function formulated as $P = P(\theta, \phi, \lambda, t, V_x, V_y, V_z)$, depicting light rays with wavelength λ passing by an angle (θ, ϕ) through the pupil of the eye placed at location (V_x, V_y, V_z) at time t . Due to the high dimensionality of the plenoptic function, calculations can be extremely complex and hard to process. In an effort to reduce the complexity, McMillan and Bishop [16] reduced the dimensionality of LF representation to

5D by registering a set of panoramic images captured at various 3D positions. Later in 1996, Levoy and Hanrahan [9] further reduced the LF representation to 4D in the case of free space (i.e., no occluders). Their idea was based on the parametrization of light rays using their intersections with two planes while traveling in straight lines. In other words, the two intersection point pairs (u, v) and (s, t) on the planes are used to represent LFs. Figure 3 depicts the progression of the LF representation over the years.

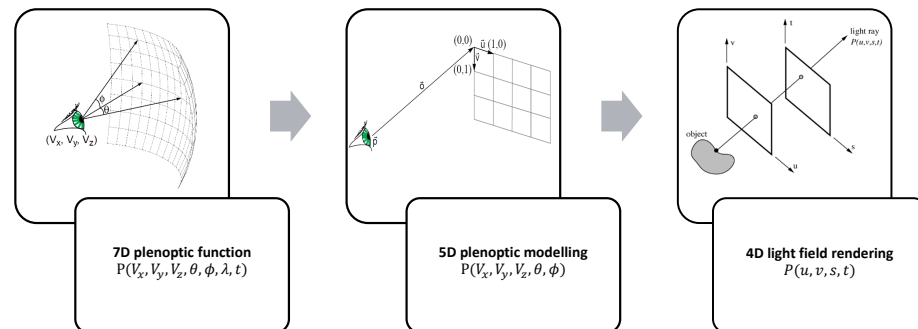


Figure 3. Simplification of LF representation [13,14].

Capture LF hardware allows the storage of almost all the information of the viewed scene from the camera's point of view, where the amount of light in each light ray arriving at the camera sensor is recorded [17]. This information can be useful in further applications requiring additional knowledge about the visualized scene. For additional information about the scene, dynamic cameras can be used for navigation.

Nowadays, LF capture systems can be either horizontal-only parallax (HOP) or full-parallax (FP) systems, with the latter capturing the parallax in both directions. Considering the baseline length of LF capture and display devices, they can be categorized as narrow- and wide-baseline systems. Furthermore, in the case of projection-based LFDs, the location of the projector array with respect to the screen and the observers creates two categories as well. If the projectors are on the same side of the screen as the observers, then it is a front-projection LFD, and if they are on the other side, then it is a back-projection LFD. The two types of projections are illustrated in Figure 4. Note that in the case of front-projection solutions, the projectors are typically above the viewers, and they may also be located behind the viewers.

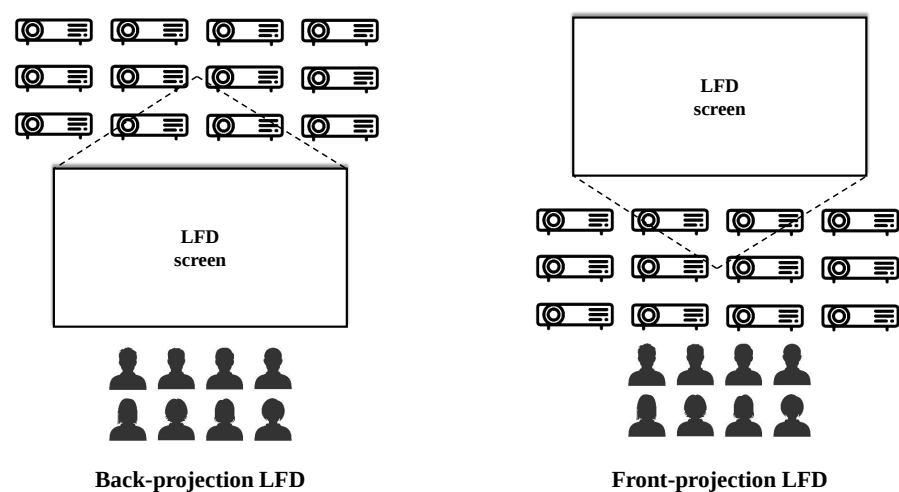


Figure 4. Back-projection and front-projection LFDs.

For LFDs, the number of views for a given content is directly proportional to the achieved motion parallax within the FOV. Hence, the aim of LFDs is to provide as many views as possible for contents in order to have a continuous motion parallax, which in turn requires an adequate degree of angular resolution [18]. Similarly to LF capture systems, LFDs can be either HOP or FP. Since the eyes are horizontally separated, HOP LFDs are more practical than the implementation of a vertical parallax, in addition to being less complex than FP solutions. For HOP LFDs, farther viewing distances are only possible with sufficient corresponding angular resolution [19]—as insufficient light ray density makes the visualized content look flat 2D, as no two or more distinct rays with respect to a given point on the screen can reach the two pupils—which also extends the valid viewing area (VVA), the angle of which is originally determined by the FOV. The FOV itself is determined by the baseline of the system. Theoretically, an LFD is considered to be a narrow-baseline system when the FOV ranges between 10° and 15° , whereas for FOV values greater than 30° , the LFD is counted as a wide-baseline system [20]. However, at the time of writing this paper, there is no scientific-community-wide consensus regarding this classification. Regarding the relation of the content and the display, as a general rule, the more the LFs match between both the capture and visualization systems, the higher the quality of the visualized content, as fewer content adjustments are required.

Similarly to the idea of viewing frustums for general 3D rendering [21], double viewing frustums—inside which the elements within the region of interest (ROI) are visualized—are used for LFDs. The ROI defines a box-shaped volume in the virtual scene, within which all contents are visualized on the LFD [22,23]. For projection-based LFDs, the characteristics of the ROI can have considerable effects on the perceived visualization quality, where poorly chosen ROI values can greatly deteriorate the quality of the displayed content [22]. The ROI plays a major role in generating LF camera animations virtually, the topic of which is discussed in detail in Section 5. In the concept of a double viewing frustum, the screen is placed between the frustums. Contents rendered around the screen are considered to be in the sharp region of the LFD and, hence, are rendered sharply. However, contents further away from the screen enter the blurry region of the display and, thus, suffer blurriness, resulting in a lower perceived quality.

Since LF rendering requires the storage of almost all the visual information related to the captured scene by means of storing multiple views of the given scene, compression techniques are needed to accommodate the huge amount of information. In order to achieve a smooth continuous motion parallax effect, the number of captured views needs to be sufficiently high. In efforts to solve the data storage and bandwidth problem, various compression techniques—taking into account the similarities between the LF images representing the scene—were suggested for LFs. This includes disparity compensation for compressing synthetic 4D LFs [24–26], which was already highly investigated prior to the emergence of the first modern projection-based LFDs. Approximation through factorization [27] and geometry estimation using Wyner–Ziv coding [28] were also relevant approaches from that scientific era. From the beginning of the 2010s, various compression methods for LF images captured by hand-held devices were proposed [29–35]. The subsequent efforts relied on homography-based low-rank approximation (HLRA) [36], disparity-guided sparse coding [37], deep-learning-based assessment of the intrinsic similarities between LF images [38], and Fourier disparity layer representation—where the Fourier domain can effectively construct a set of layers for LF representation given very few views [39]. In recent years, the contemporary solutions included low-bitrate LF compression based on structural consistency [40], disparity-based global representation prediction [41], compression by means of a generative adversarial network (GAN) [42], a spatial-angular-decorrelated network (SADN) [43], bit allocation based on a coding tree unit (CTU) (which takes into account the HVS to remove perceptual redundancy) [44], compressed representation via multiplane images (MPIs) comprised of semi-transparent stacked images [45], and neural-network-based compression by using the visual aspects of sub-aperture images (SAIs), incorporating descriptive and modulatory kernels [46]. Further lossy compression methods

for LFs include transform coding [47–51], predictive coding [35,52–54], pseudo-sequence coding methods [55–57], and utilizing a two-dimensional prediction coding structure [58]. Table 1 summarizes the different LF compression techniques.

As previously stated, the storage and transmission of LF images present significant challenges due to their large data volumes. Consequently, various compression techniques have been employed to address this issue. As illustrated in Table 1, it is evident that the majority of these techniques are of the lossy type. Many of these techniques handle LF images from a single scene similarly to how video sequences are managed. In other words, these methods apply video coding strategies to LF images that represent a single scene. Whereas these methods produce plausible results, they mostly consider the case of static scenes where a static camera is deployed. However, in the case of LF dynamic camera animations, a significantly larger number of images must be stored, particularly due to the substantial scene changes caused by extensive camera movements. This presents a major challenge—and in some instances, limitations—in the context of LF dynamic camera animations.

Table 1. LF compression techniques.

LF Compression Technique	Citations	Type of Compression
Disparity compensation for compressing synthetic 4D LFs	[24–26]	lossy
Approximation through factorization	[27]	lossy
Geometry estimation using Wyner–Ziv coding	[28]	lossy
Compression methods for LF images captured by hand-held devices	[29,30,32–35] [31]	lossy lossless
Homography-based low-rank approximation	[36]	lossy
Disparity-guided sparse coding	[37]	lossy
Deep-learning-based assessment of the intrinsic similarities between LF images	[38]	lossy
Fourier disparity layer representation	[39]	lossy
Low-bitrate LF compression based on structural consistency	[40]	lossy
Disparity-based global representation prediction	[41]	lossy
Compression by means of a generative adversarial network	[42]	lossy
Spatial-angular-decorrelated network	[43]	lossy
Bit allocation based on a coding tree unit	[44]	lossy
Compressed representation via multiplane images comprised of semi-transparent stacked images	[45]	lossy
Neural-network-based compression by using the visual aspects of sub-aperture images, incorporating descriptive and modulatory kernels	[46]	lossless
Transform coding	[47–51]	lossy
Predictive coding	[35,52–54]	lossy
Pseudo-sequence coding methods	[55–57]	lossy
2D prediction coding framework	[58]	lossy

While certain LFD implementations have already emerged and are now commercially available, a vast amount of research effort is still required to enhance and optimize the visualization quality of LFDs. Such efforts are centered around human-centric considerations and quality of experience (QoE). These primarily include perceptual thresholds and personal preference, and there are many open research questions that are yet to be investigated [59]. The empirical data for QoE research are collected via subjective quality assessment, the methodology of which has been recently standardized [60]. At the time of writing this paper, various subjective tests were performed on HoloVizio LFDs of Holografika [61–68]—particularly the 721RC, the 722RC (also known as 640RC), the C80 cinema system, and the 80WLT television-like display. The test variables of the state-of-the-art scientific literature include spatial and angular resolution [69,70], spatial and angular distortion [71–74], compression [75], interpolation [76], viewing distance [19], zoom level [23], content size [77,78], content characteristics (e.g., complexity [79], alignment [22] and orientation [80]), human–computer interaction (HCI) [81,82], and more. Additionally, novel dedicated prototypes (i.e., designed for specific use cases) were introduced and tested, such as for the future use case of LF telepresence [83,84].

As the vast majority of the aforementioned efforts of subjective quality assessment fundamentally rely on LF contents, there is an evident need for LF datasets. Such datasets vary a lot in their characteristics and provide information. They can contain real-world captured contents, synthetic ones (i.e., rendered), or a combination of both. Datasets targeted for evaluating the quality of LFs usually incorporate high-quality contents, along with their impaired counterparts. According to Shafiee and Martini [85], LF datasets can be categorized into three groups: (i) content-only datasets, (ii) task-based datasets, and (iii) QoE datasets. As the name implies, content-only datasets contain the LF contents only, and nothing more. Real-world captured contents may be acquired by a lenslet camera [86–91], a single-lens camera [90,92–96], or an array of cameras [90,97]. In the case of rendered content-only datasets [94,96,98], the camera is virtual, which is, of course, applicable to the other dataset types as well, along with the classification of real cameras. Task-based datasets include additional information on the task for which the dataset was created. Similarly to content-only datasets, task-based datasets can also be captured by a lenslet camera [99–103], a single-lens camera [104,105], an array of cameras [106], or by a virtual camera [103,104,107,108]. Finally, QoE datasets contain subjective ratings that were acquired through extensive testing with numerous test participants. The currently available QoE datasets were captured by either a lenslet camera [109–113], a single-lens camera [72,114], or a virtual camera [113,115]. The dominant portion of the datasets covered in this section contain LF images. LF video datasets—such as the work of Guillo et al. [92]—are exceptionally rare at the time of writing this paper. The different types of LF datasets [85]—along with relevant examples for each data capture method—are summarized in Table 2.

Table 2. LF dataset types [85].

LF Dataset Type	Definition	Data Capture Methods	Examples
Content-only	Contains the LF contents only	- Lenslet camera - Single-lens camera - Array of cameras - Virtual camera	[86–91] [90,92–96] [90,97] [94,96,98]
Task-based	Includes additional information on the task for which the dataset was created	- Lenslet camera - Single-lens camera - Array of cameras - Virtual camera	[99–103] [104,105] [106] [103,104,107,108]
QoE	Contains subjective ratings that were acquired through extensive testing with numerous test participants	- Lenslet camera - Single-lens camera - Virtual camera	[109–113] [72,114] [113,115]

A common research topic within the scientific community of LF technology is super resolution. However, as there are two distinct interpretations for the same terminology, it requires clarification via a prefix. Among the published works, the most frequent interpretation of super resolution is image resolution enhancement. As this may be considered the default interpretation, the terminology is often used without a prefix. An accurate prefix for such may be spatial super resolution, but image super resolution also describes the notion faithfully. Different methods were devised to achieve LF image super resolution, including projection-based methods [116–119] and optimization-based methods [120–126]. A great number of novel attempts employ convolutional neural networks (CNNs) and deep networks, specifically targeted for data captured by LF cameras, since such devices have limited spatial and angular resolutions. These networks aiming to achieve spatial super resolution for LF images include a two-stage CNN, exploiting the correlations among the LF images both internally and externally [127]; a bidirectional recurrent network [128]; a deformable convolution network, taking into account the angular information among images while handling disparities [129]; residual networks, where the LF images are first grouped and then fed into different network branches from which the residual information along different directions is calculated [130]; and an algorithm applying optical flow to align LFs, after which the angular dimension is reduced by means of low-rank approximation, and then, a deep CNN is used for spatial super resolution [131]. Additionally, among the other methods of spatial super resolution are the LF-DFnet (deformable convolution network) [129]; the LF-IINet (intra-inter view interaction network), preserving the system parallax while exploiting the correlations among images [132]; dense dual-attention networks [133]; and end-to-end networks using epipolar geometry, in order to learn the details of sub-pixels per view image [134]. In efforts to reduce the dimensionality, the complexity, and the cost of 4D LF data, the work of Van Duong et al. [135] proposes a network that decomposes LF data into a lower data subspace while exploiting the information resulting from the possible 4D LF representations, including epipolar plane image (EPI), as well as spatial and angular information. Regarding networks enhancing both spatial and angular resolutions, the work of Yoon et al. [136] proposes the LF convolution neural network (LFCNN), composed of spatial and angular super resolution networks. Furthermore, LF-InterNet [137] enhances both the spatial and angular resolutions by extracting their features from LFs separately, with interactions occurring between them later, ending up by fusing the interacted features. Another method uses two super resolution networks, targeted for spatial and angular super resolutions separately, generating multiview features that are later remixed by an adaptive feature remixing (AFR) module [138].

The other interpretation of super resolution is angular super resolution. It refers to an angular density so high that not only two distinct light rays with respect to a point on the LFD screen address the two pupils of the observer—which is essential to the 3D visual experience [19]—but also a single pupil. Based on the state-of-the-art LF visualization technology and its current usage, angular super resolution has not yet been achieved. The reason why the word “usage” is involved in this statement is that angular super resolution evidently depends on the viewing distance as well. After all, the farther the observer, the lower the perceivable light ray density—making LF visualization appear flat 2D beyond certain distances. The most significant benefit of reaching angular super resolution is that it allows observers to change their focal distance. While with lower angular resolution, one may only focus on the plane of the screen of the LFD, with angular super resolution, one may focus on closer and farther portions of the visualized content. Although achieving such a goal may greatly benefit LF use cases, it poses great challenges on multiple fronts. As angular super resolution is a future research topic that is applicable to and relevant for both use cases, visual content, and quality assessment, as well as hardware, we address it in Section 9. In the next section, we briefly review the currently available research efforts on LF camera animation.

5. Implementing Light Field Camera Animation

Camera animations comprise changing the focal length, animating camera movements, rotations, and switching between different cameras. Considering general camera animations, we can categorize them into cinematography [139] and simulation camera animations. Cinematography camera animations include pan, tilt, zoom, dolly, truck, and pedestal movements. As for simulation camera animations, gaming and other interactive environments are the most common use cases, in which observers and players may perceive the surrounding virtual environment. Thus, the main camera attributes—such as position and orientation—should be determined [140]. Simulation camera animations include first-person cameras—also known as point-of-view (POV) shots [141]—second-person cameras, and third-person cameras.

Considering LF capture systems, specific customized hardware should be used, whether it is plenoptic cameras, 2D camera arrays, or conventional cameras customized for LF capture. This further complicates the capture process for LFs, in addition to the issues of cost and portability. To overcome the aforementioned problems, virtual LF cameras are adopted. Similarly to the idea of LF cameras capturing the light distribution about real scenes, virtual LF cameras achieve the same for virtual scenes with the added advantage of being free of error. Considering LFDs, the associated ROI matrix facilitates the different camera motions. Practically, the inverse of the ROI matrix is used to evaluate the display rays and to transform them into the world space. This can be applied to individual rays, where the same coordinate system includes all objects and light rays in the scene.

In our earlier works [1–3], we simulated some of the camera animations mentioned earlier, including cinematography, simulation, and realistic physical camera animations. Cinematography animations included pan, tilt, zoom in/out, dolly in/out, truck, and pedestal animations. For simulation camera animations, first-person and third-person cameras were implemented. As for the realistic physical camera animations, a collision camera, a suspension camera, and a falling camera were implemented, while taking into account the physical properties of the camera and the scene; hence, the realistic physical simulation. The aforementioned camera animations were tested and visualized on the HoloVizio C80 LFD. Further tests were performed to measure the plausibility of the visualized content, including subjective tests, as well as objective metrics performed specifically for the realistic physical simulations.

There is a lot to discuss regarding LF camera animation due to its relevance to practical utilization contexts and also due to the fact that research on the topic has only recently begun. In this paper, we organize this discussion into the topics of use cases, visual content, quality assessment, and LF hardware. Each of these aspects are approached from the perspectives of implications, limitations, potentials, and future research efforts. Therefore, our core contribution is distributed among 16 subsections.

6. Implications

6.1. Use Cases

Due to the ability of LF cameras to capture and store 3D information related to scenes, LF camera animations can be effectively integrated in many applications. These include—but are not limited to—medical imaging, telepresence, resource exploration, prototype review, training and education, gaming, digital signage, cinematography, cultural heritage exhibition, air traffic control, and driver assistance systems [142,143]. For some use cases, using a static LF camera may be sufficient, such as in the case of telepresence, where the camera is placed in front of the individual(s) to capture either the entire body [84] or a portion of it (e.g., just the head [83]). In the case of medical imaging, the content itself may be static, but its visualization may allow interaction for better diagnostics accuracy (e.g., using rotation and zoom). Furthermore, a static camera was used to investigate the various interactions on LFDs by means of a theater model simulation, depicting real theaters in their angularly selective nature [144]. Furthermore, static LF cameras can be used in digital signage and cultural heritage exhibitions. Considering the latter, an LFD was

installed in a museum via the i-MARECULTURE project [145] to enable a visually attractive, state-of-the-art representation of cultural heritage. Nonetheless, for a more engaging—yet rendering-wise challenging—experience, dynamic camera motion can be used, which may be applicable for the vast majority of use cases. For example, it should be noted that in time, medical LF videos may become viable as well.

6.2. Visual Content

The effect of using LF camera animations is rather case-dependent, hence, affecting the resulting visualized content. Considering static cameras, the final output depends on the position of the entity (e.g., character or object) under consideration. Regarding dynamic camera animations, our previous work [3] indicates that first-person cameras have the worst visual quality, followed by tilt, zoom, dolly, pan, pedestal, and third-person cameras. On the contrary, pan and truck movements scored the best visual results on LFDs. Thus, it may be preferable to avoid first-person and zoom-in camera animations if possible. The instance of zoom is illustrated in Figure 5. The dashed lines in the figure characterize the symmetrical regions of LF visualization (i.e., behind and in front of the plane of the screen). The example is exhibited for one side of visualization, but it is, of course, applicable to the opposite side as well. As shown in the figure, while normal camera zoom places the 3D object for camera capture in the sharp region (i.e., visualization is sharp), the zoomed-in camera results in blurriness. On the other hand, blurriness resulting from zoom-out on LFDs is not considered an issue since the same occurs to objects far from observers in real life; however, the same does not apply to the zoom-in camera animations, since the entity with the main camera focus is blurred, causing severe irritation with the possibility of hindering the visualization of the entire 3D scene [22,146].

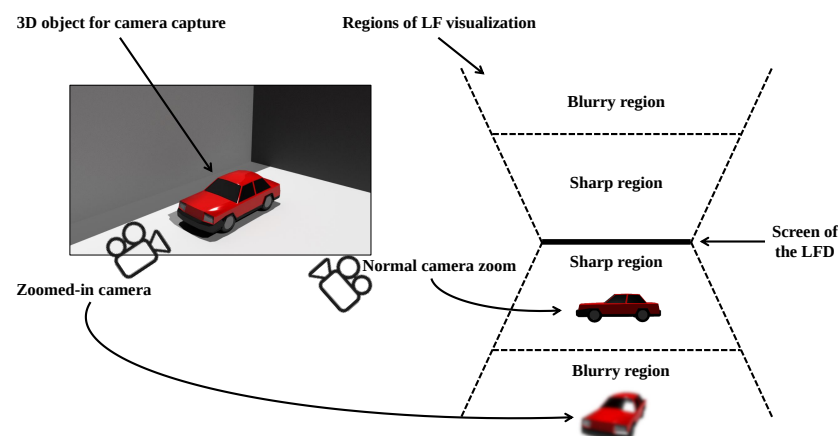


Figure 5. The impact of zoom on visualization sharpness.

6.3. Quality Assessment

Quality assessment may serve different goals, depending on the research question of the study. As mentioned earlier, it primarily focuses on perceptual thresholds (e.g., whether test participants can distinguish different representations) and personal preference (e.g., whether test participants prefer one representation over another). In a subjective test, either a single test participant is present at a time (which is the dominant methodology within the scientific community), or multiple test participants observe the same visualization simultaneously (which is plausible for the investigation of inter-user effects). A great portion of the experiments in the scientific literature rely on static contents (i.e., objects and scenes), shown from the perspective of a static camera.

Considering realistic physical LF camera animation simulations, three objective metrics were devised for assessment: (i) a collision metric, detecting collisions between the camera and scene elements; (ii) a blurry region metric, keeping track of the number of scene elements rendered outside the sharp region of LFD; and (iii) an occlusion region metric, used in the case of third-person cameras to detect scene elements obstructing the main entity from the camera [1,3].

6.4. Capture and Display Hardware

According to Wetzstein [11], LF acquisition can be classified into three main categories: (i) multiple sensors, (ii) temporal multiplexing, and (iii) spatial and frequency multiplexing. Table 3 characterizes the different types of LF acquisition. Due to their ability to capture multiple images with broad range of distances, using multiple sensors generates high spatial resolution LFs by enabling efficient depth reconstruction. On the other hand, the significant weight of the camera setup may result in issues of portability. The portability of camera arrays can be generally difficult, especially if they are used outdoors, due to their reliance on power supply, as well as the usage of complicated transmission lines [147]. Constant synchronization and calibration should be maintained across the different cameras within the system, as well as managing the storage and processing of the huge amount of captured data. 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 LF acquisition [148]. A major limitation of using multiple cameras is the limited view resolution, where the physical dimensions (i.e., size of the camera) and limitations (e.g., constrained speed and degrees of freedom of the rig) restrict the gaps between the cameras when being placed beside one another, along with the possibility of a self-capture approach [10,11,149,150]. Not only do these methods present numerous challenges, but they are also primarily designed for static cameras. When it comes to dynamic camera animations—such as the case of a car chase scene—this type of acquisition method faces even greater challenges and, in some instances, significant limitations in its application. Accordingly, an alternative to using multiple cameras for wide-baseline capture is temporal multiplexing, achieved by a single sensor. Unlike using multiple sensors, temporal multiplexing has the advantage of significant reduction in costs and complexity where a single camera is used, requiring less calibration. On the other hand, the scene is required to be static, hence, the utilization of this method is rather limited, less universal, and thus, less practical. In other words, this method is rendered invalid in terms of dynamic camera animations. Unlike the techniques utilizing multiple sensors, spatial and frequency multiplexing methods overcome the portability issue, while generating high-density views. While dynamic scenes can be captured efficiently by means of spatial multiplexing, a trade-off occurs between the angular and spatial sampling rates. Accordingly, compared to the previous methods, reduction in spatial resolution is noticeable [10,11,149]. Moreover, in most cases, the baseline between the captured views is small due to the small distance between the microlenses in the microlens array setup. Thus, the majority of the methods within this category are classified as narrow-baseline solutions. It is noteworthy that all acquisition methods encounter greater challenges with wide-baseline capture systems. Consequently, the use of virtual cameras to capture wide-baseline LFs is favored, particularly for scenarios involving dynamic camera animations [1,3].

Table 3. LF acquisition types and methods.

LF Acquisition Type	Definition	Acquisition Methods	Examples
Multiple sensors	Camera arrays for wide-baseline capture	- Linear camera setup - Arc camera setup - 2D grid camera setup - Spherical camera setup	[66] [84] [151] [152]
Temporal multiplexing	Uses a single camera instead of multiple cameras for wide-baseline capture	- Camera on turntable or rotating camera while reorienting - Programmable aperture photography - Extension of integral photography - Rotation of a planar mirror - Lensless LF camera	[9,153,154] [155] [17,156,157] [158] [159]
Spatial and frequency multiplexing	Uses a single camera to create LF images by means of spatial or frequency multiplexing	- Parallax barriers - Integral photography - External lens arrays - Array of planar, tilted mirrors or mirrored spheres - Lens arrays and a single sensor in related compound imaging systems - Combining a lens array and a flatbed scanner in a lenslet-based architecture - Plenoptic cameras - Frequency multiplexing	[7] [8] [160–162] [163–166] [167–170] [171] [156,172] [173]

An alternative way to categorize LF capture systems is by using the baseline as a criterion, where a temporally moving single-camera or a camera array configuration is considered wide-baseline capture. In the case of HOP wide-baseline capture, cameras are placed horizontally in a linear (e.g., the LF transmission of Balogh and Kovacs [66]) or an arc (e.g., the telepresence system of Cserkaszy et al. [84]) manner. For FP wide-baseline systems, cameras are arranged as a 2D grid (e.g., a 64-camera setup arranged in an 8×8 grid [151]) or spherically (e.g., a spherical LF camera using the Gaussian blending method for vision reconstruction [152]). On the other hand, LF cameras (e.g., a plenoptic camera) are considered narrow-baseline, as their baseline is limited by the aperture size [156,174]. Not only do plenoptic cameras capture the light intensity in the scene but also the exact direction traversed by light in space [17]. An example for commercial plenoptic cameras are Raytrix cameras [172]. There were many other attempts to implement LF cameras, as well as to enhance the content captured by LF cameras [120,175,176].

With the aid of LFDs, the ability to perceive 3D contents without the need of visual gears is possible, which, in turn, enhances the user experience. Motion parallax in LFDs allows users to view the scene from multiple perspectives as they move sideways within the VVA. This is a major advantage of LFDs, as even with the usage of a static camera, the scene can still be viewed from many angles. Accordingly, in many applications, using camera animations can be avoided, which simplifies the rendering process.

7. Limitations

7.1. Use Cases

Real-time rendering for LFDs is an important yet complex task that needs to be achieved in order for LFDs to extend their usage in many applications, including essential fields such as time-critical medical visualization, telepresence, and gaming. For visually plausible contents, high ray density (i.e., angular resolution) is necessary. If the angular resolution is not sufficient, then different forms of degradation may occur [78,79,177], such as the interference of neighboring perspectives and the sudden jumps between adjacent views as the observer changes viewing position and angle.

There are major limitations associated with the sharp rendering region of LFDs. One significant issue is that the more depth a content has, the more it is affected by insufficient

angular resolution (i.e., the portions of the content that are the closest to the observer(s) and those that are farthest away are the most susceptible to low angular resolution). While this may be somewhat compensated by content and display characteristics, there are certain use-case-related camera perspectives that are less feasible. In the context of gaming, games in which the camera looks towards a well-defined surface may be implemented without notable difficulties, as the space of visualization is evidently constrained. For example, in the case of a strategy game where the camera looks towards the ground, the ground is basically a barrier; no visualized object can be farther away from the camera than the plane of the ground. However, in an open-world action game or first-person shooter, the camera may look towards extreme distances, which cannot be properly encompassed by LF—as already emphasized by the theoretical framework. Thus far, no universal solution has been proposed to overcome this specific limitation. This, of course, does not necessarily mean that first-person games are not plausible at all. For instance, one may implement a first-person action game in which levels are conventional mazes, where the maximum length of a corridor is constrained—thus, limiting the maximum depth of visualization. However, first-person perspective in general is very challenging as of now, and even if it is implemented with certain constraints, the camera movements and other factors may still hinder the perceived quality of LF visualization.

7.2. Visual Content

As stated in Section 5, realistic LF camera animations can sometimes be exceedingly challenging due to the portability- and cost-related issues. Accordingly, creating virtual contents for LFDs is generally a viable solution to overcome the problems associated with realistic LF contents. Such may be implemented by using the ROI of the LFD—determining the region within the virtual scene that needs to be displayed.

Even though virtual LF capture can solve numerous issues related to real capture systems, this is not applicable to using a first-person camera, as blurriness may occur in the resulting visuals due to the inflicted zooming effect—pushing the rendered content into blurry regions. In other words, the optical limitations of LFDs may gravely reduce the quality of the rendered content when using first-person cameras. Additionally, significant camera motions (i.e., rigorous, strong motions) may further deteriorate the visualized content, causing dizziness and loss of focus [2,3].

Moreover, capturing the details of an entity by zooming in with real LF cameras is technically impossible (i.e., no solution exists for it or even proposed for it at the time of writing this paper, and no theoretical framework supports it) since such requires an optical change in the focal lens of the cameras, which, in turn, requires a change in the lenses or the baseline, with a possible system re-calibration for the entire LF capture system. One possible solution is to scale the volume of the ROI in a way to simulate the zooming effect without the need to change the capture system [23]. However, this may result in the degradation of the visualized content as it enters the blurry regions of LFDs.

Considering LFDs, some contents are impossible to display, even when captured virtually. As an example, let us take the case of visualizing very detailed tiny objects—such as molecules—captured by virtual static LF cameras. In this case, the need to zoom in to visualize the content more clearly is evident. However, doing so results in deteriorated visualizations due to severe blurriness. Another attempt is to visualize the content at a distance while scaling it up to avoid zooming in, but then the spatial resolution is affected, leading again to blurriness. Furthermore, adding dynamic LF cameras to this situation may further degrade the quality of visualization, as it may introduce dizziness and fatigue through the camera motion itself.

7.3. Quality Assessment

One of the major factors to consider when assessing the quality of LF camera animation is the display on which the test stimulus is visualized. Hence, the limitations of LFDs affect the quality assessment of the captured camera animations.

Subjective tests on LF QoE are typically performed in controlled laboratory environments. Control is necessary in order to maintain the validity of the collected ratings and user input. It extends to every relevant parameter of the environment, such as lighting conditions and the isolation of external distractions (e.g., external sources of sound and light). While the size of the accommodating environment (i.e., the size of the room where the test takes place) is not always essential to many research efforts on visualization quality, for LFDs, it is particularly important for two reasons. First of all, in the case of high-end displays, the investigation of viewing distance may be limited, as a high angular resolution may support a maximum viewing distance greater than the size of the actual environment (i.e., the corresponding dimension of the room is smaller). Secondly, the measurement methodology may allow free movement within the VVA. In this case, a similar limitation may apply, as the environment may not be able to fully accommodate the entire VVA.

Another major limitation of assessing LF animations is the fact that the assessed animations are mostly virtual ones, since capturing real LF camera animations by means of LF capture systems is almost impossible at the time of writing this paper.

7.4. Capture and Display Hardware

Considering wide-baseline LF capture systems, the extensive weight of the camera arrays poses a major problem on portability. Accordingly, for dynamic motions, capturing rigorous movements for wide-baseline LFs is practically impossible, in addition to being overly expensive to deploy camera rigs accommodating the huge camera array setup. Additionally, more issues arise per each use case. Let us consider a case where a camera array setup is being used to shoot a car chase scene while other cars on the street are moving by. Not only does the setup poses the threat of outweighing the weight of the vehicle, but also there is a possibility of the setup colliding with other cars or road entities by accident since it consumes a huge space, exceeding the dimensions of the car itself.

Due to the angularly selective nature of LF visualization, the perspective of the content depends on the viewing angle. This is very useful in most cases. For instance, simulating a theater model on LFDs [144] conveys an experience similar to that of a real theater with many simultaneous spectators. However, in some cases, having a visualized content that depends on the viewing angle can result in certain issues. In essence, two observers looking at the same screen are rather likely to see different contents. These are the scenarios where using a 3D multiview display may actually be beneficial, as every single observer perceives the same perspective.

Regarding the sharp rendering regions of the display hardware, one may overcome this limitation by issuing constraints regarding content depth, as well as by creating a system that has an inherently lower depth budget. However, this is rather counterproductive, as it negatively affects the realism and 3D nature of visualization. On the other hand, it may actually prove beneficial in certain contexts. For example, in the case of a third-person camera, the entity of interest is visualized in a notably sharper manner than its environment, creating a sense of focus.

8. Potentials

8.1. Use Cases

Since LFs are the closest to simulating the HVS understanding and interpretation of the real world, such displays have great potential in a variety of applications. Regarding medical imaging, crucial information is sometimes missed using 2D methods, such as multiplanar reformations (MPR). This issue is overcome by means of various solutions of 3D visualization, offering more detailed information in many fields by creating high-quality 3D volumetric renderings [178–181]. Similarly to other 3D technologies, LFs have great potential in many aspects of the medical field, among which is radiology [182]. The ability of LF cameras to capture the complete information of the scene is beneficial for the medical field, in addition to using dynamic cameras such as orbiter cameras [183] to capture a full

image of the analyzed medical content (e.g., organ). On the other hand, zooming in on LFDs is limited, as it results in deteriorated content visualization.

Due to the ability to convey more information compared to 2D displays, LFDs can be used for specialized training. Furthermore, it is suitable to use for educational purposes (e.g., higher education), where the internal structures of complex machines or even the human body can be visualized [142,184]. Accordingly, for such cases, high 3D quality should be conveyed to encompass all the complete details. Depending on the use case, a static camera or dynamic camera can be used, all of which should take into account capturing and displaying the corresponding information with a certain aspect of detail while avoiding blurriness.

Unlike the case of using smartphones and digital dashboards for driver assistance applications [185–188], LFD utilization can be extended to include driving simulations. Since the ultimate goal of LFDs is to provide a window to the 3D world—achieving realistic 3D perception—LFDs are excellent tools to perform driving simulations, mimicking the real world. In that case, LF camera animations simulating that of a camera attached to a vehicle should be investigated. Additionally, LF visualization technology may be directly used in actual vehicles as a 3D windshield [189].

8.2. Visual Content

LF camera animation can greatly affect the plausibility of the visualized content. As previously mentioned, certain camera animations can lead to the deterioration of the final output, such as first-person cameras and zoom-in camera motions. One possible solution to overcome the first-person camera issue is to render the scene from the third-person camera perspective, and then cull the part of the scene facing the observer until the main entity is removed, hence, giving the perception of the first-person camera with a little bit of a zoom-out effect to avoid blurriness. An alternative approach is to use transparent rendering instead of culling, but this takes more time to process.

The angularly selective nature of LFDs (i.e., the perception of visual information depends on the viewing location) can be an issue for some applications. This could be partially solved by moving the camera in a way to reveal all information captured from all directions for all users. As an example, let us take a side-standing playing card where either the front or the back of the card is displayed to the user, depending on the current location. In such a case, the camera can be rotated around the card to capture it from the opposite direction, such that the other missing information is conveyed to the user; hence, ending up with all users perceiving the same information. This could be integrated in many applications, such as prototype reviewing and medical imaging, where an orbiter camera [183] could be used. The utilization of an orbiter camera in medical use cases is illustrated in Figure 6.

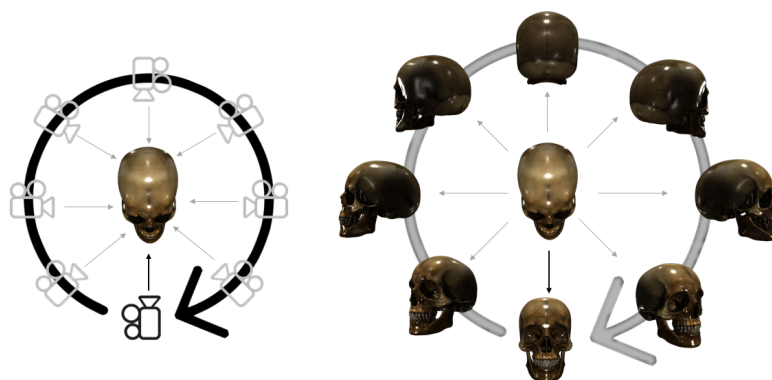


Figure 6. Orbiter camera for medical use cases.

It is possible to take advantage of the angularly selective nature of LF visualization. For example, in the case of split-domain gaming [59,143,190], the VVA is split into different

domains in which the different user perspectives (i.e., player interfaces) may be perceived; thus, everyone only perceives their own interface. Such a solution may benefit from camera animation (e.g., the panning camera of turn-based and real-time strategy games); however, the core concept is fully implementable without any camera animation. Conventional split-screen gaming (vertical and horizontal division) and LF split-domain gaming are compared in Figure 7. The primary advantage of the latter is that unlike in the case of split-screen gaming, the entire screen is allocated to both players simultaneously. The overlapping region (also known as the separation zone) is considered to provide invalid LF, as the two user perspectives interfere with each other. Evidently, any viewing position outside the FOV is similarly considered to provide invalid LF.

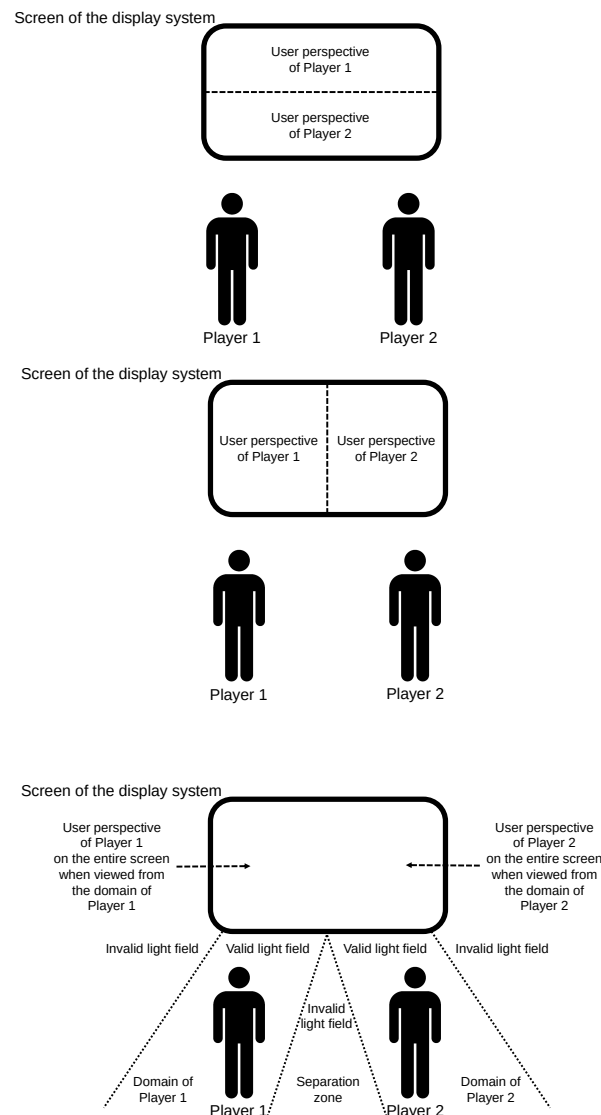


Figure 7. Vertical split-screen gaming (**top**), horizontal split-screen gaming (**middle**), and LF split-domain gaming (**bottom**).

8.3. Quality Assessment

Possibly the greatest advantage of LF camera animation for quality assessment is the broadening of investigated contents and scenarios. For example, an earlier study on cinematic setup [69] was conducted by using static contents, and LF videos in various works [75,191,192] exclusively used static cameras.

In order to assess the quality of interactive use cases such as gaming, LF camera animation may be necessary. While this may fundamentally depend on the investigated gaming genre, it is important to explore such challenging forms of LF visualization, as the primary goals of LF quality assessment not only include the subjective evaluation of LF visualization but also the proposals for degradation mitigation and resource optimization (i.e., perceptual coding).

8.4. Capture and Display Hardware

Most of the limitations of capture LF systems are related to wide-baseline implementations due to their issues regarding portability and cost. In contrast, dynamic LF narrow-baseline cameras can be used in research contexts and use cases without having considerations for such limitations. Furthermore, since most limitations are more concerned with the display systems rather than the capture ones, different camera motions may be investigated for narrow-baseline systems without directly relying on LFDs for content validation. For example, it is possible to take a certain view for all captured scenes in a dynamic camera movement and then display them consecutively as one video to evaluate the captured content without the need of LFDs.

A notable potential of camera animation is enhancing the captured scene with pre-defined changes in the perspective, which may be essential for multiple aspects of technology utilization, such as story telling. When using LFDs, two types of motion should be considered: (i) LF camera animations of the content and (ii) the movements of the spectators within the VVA of the LFD. It is possible to have them both simultaneously, such as in the case of digital signage. In certain scenarios, the direction of viewers' movement may be well defined (e.g., LF unit of digital signage placed along a one-way road). If the display hardware provides a smooth horizontal parallax via high angular resolution, the camera animation may exploit the context to make visualization more impressive and appealing (e.g., designing content that takes into account the estimated change in viewer perspective).

9. Future Research Efforts

9.1. Use Cases

In order to extend their usage in various fields, real-time rendering for LFs is essential. For visually plausible content, multiple views need to be rendered for a single scene. This further necessitates the utilization of compression techniques—exploiting the spatial coherence between neighboring views. In addition to spatial coherence, temporal coherence needs to be investigated between the same views of the scene when using dynamic cameras. For example, Figure 8 illustrates the truck LF camera movement, capturing 30 views of the classroom scene [20], arranged in a 5×6 grid. As shown in Figure 8, the spatial coherence between consecutive views is indicated by yellow arrows, while the temporal coherence between the same views across the different frames is indicated by the brown arrow.

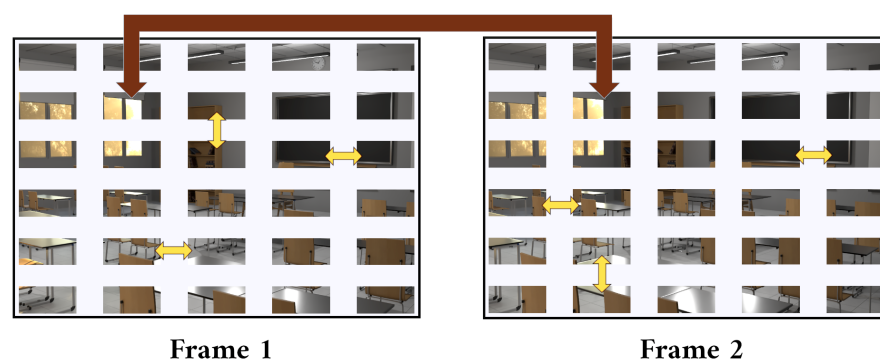


Figure 8. Types of image compression across two consecutive frames captured using a virtual LF camera with truck movement [20].

LFDs offer great potential for digital signage. They are widely used across the globe for commercial and other purposes. In the case of using LFDs for digital signage, the glasses-free 3D nature of the visualization technology may add great value to the use case, as such visuals may easily grab the attention. For this use case, a simple static LF camera is sufficient to create the desired visual effect without further complications. However, for a more lively experience, dynamic LF camera motions should be further investigated for digital signage, while studying their relations with the movements of vehicles and pedestrians.

Even though different attempts succeeded in implementing an LF telepresence by using prototype LFDs [83,84], dynamic camera animations—mimicking that of a handheld camera to simulate mobile video calls—should be investigated.

Angular super resolution may play a significant role in many of the potential use cases. For example, it may greatly enhance the 3D visual experience of LF cinematography and digital signage. In certain professional contexts, it may also benefit task efficiency; however, this is yet to be investigated.

9.2. Visual Content

As stated before, the general rule is that the more the parameters of the content match the capabilities of the display, the less degradation visualization suffers. Therefore, capturing display-specific LF contents (i.e., the content is created specifically for a given LFD) may avoid the potential losses in quality. Accordingly, the same content should be visualized on different LFDs with LFs matching that of the capturing system; including systems with narrow- and wide-baselines, large and small screens, as well as with deteriorated and enhanced spatial resolutions. This large pool of the different LFD categories is able to generalize the plausibility of visual content. In this context, research may be initiated by using static LF cameras, and then followed by dynamic LF cameras.

Another factor to consider with respect to visual content is the lighting conditions of the scene or object. This is particularly important for wide-baseline real LF capture systems, as light is perceived from different positions and angles of the scene for a single viewing position. As an example, let us consider the case of specular lighting on a sphere, captured by a wide-baseline camera system. In the case of static cameras, the light would be perceived differently with respect to each camera, while in the case of dynamic cameras, this would change drastically as the camera position greatly affects the lighting conditions in scene. The relation between camera position and the lighting of the scene should be further investigated for the wide-baseline LF capture systems; both statically and dynamically.

In the context of LFDs, spatial resolution is measured by the display's capability to produce fine details for the rendered content [193], attainable by increasing the angular resolution of the display. This is achieved by increasing the number of rendered views for a single scene. For more complex scenes, this could be burdensome due to the huge data storage required. Although compression methods for LFs have long been investigated (as shown in Section 4), the possibility of rendering the animated entities in a scene while ignoring the background should be further studied.

Possibly the greatest drawback of an LF content with angular super resolution is the immense size, as the angular density needs to be high enough to satisfy the criteria specified in Section 4. This characteristic, of course, not only affects storage requirements but also transmission and processing, not to mention the difficulties of capturing such via real cameras. On the other hand, creating LF contents in angular super resolution via rendering is a much more straightforward task, since the camera is virtual.

A major aspect to consider when dealing with visual content on any display is the Turing test. Originally, the well-known Turing test was suggested by Alan Turing in 1950 [194] under the name of "Imitation Game" to check whether a machine could perfectly imitate human beings. The idea of the test was to check the ability of humans to differentiate a human-to-human conversation from a human-to-computer one—or rather the ability of a computer to simulate a human conversation. In the context of visualization technologies, the Turing test was proposed to be adopted for 3D displays. This basically signifies the

ultimate goal of LF visualization, which is to become indistinguishable from the real world [195]. Hamilton et al. [196] created a framework for an LFD-related Turing test, in which certain visual characteristics must be reached to satisfy human visual acuity [197,198]. In order to achieve the goal of the framework, a tremendous amount of data—capturing not only the various views of the scene but also trillions of pixels to obtain the necessary resolution values—is required. This poses a challenge on the LF capture system, specifically for dynamic camera movements. Accordingly, LF camera animation should be taken into account with respect to the Turing test. In other words, the additional requirements related to dynamic camera motions compared to static cameras must be satisfied in order to successfully pass the Turing test for LFDs. Generally, LF camera animations should be able to capture the scene as realistically as possible. For example, motion blur should be accounted for in the animations to simulate real life, in addition to other factors, specifically of the HVS. In order for LF visualization to pass the visual Turing test, both image resolution and angular density must reach sufficiently high levels. Looking at the long-term evolution of the technology, LFDs aim to provide perceptually perfect replicas of the real world. The first milestone of such a grand scientific achievement is evidently of a static nature; however, the subsequent goals will eventually cover camera animation as well, which is expected to add new layers of difficulty via many of the considerations elaborated in this paper.

9.3. Quality Assessment

The vast majority of studies on LFDs thus far focus on perceptual thresholds and personal preference [59]. The future research efforts shall address immersion, interaction, HCI, viewing conditions, perceptual fatigue, and many more. It is important to highlight that for most research efforts not only the content was static but the camera as well. Even if the test stimuli were LF videos, those contents completely avoided camera animation.

Many of the general future research directions are applicable to LF camera animation. Immersion is relevant, as issues related to specific forms of camera animation may break the immersion instead of enhancing it. Interaction and HCI are particularly important for gaming but also essential for many other engaging use cases where camera animation is utilized. The viewing conditions can be greatly diverse among the different usage contexts, ranging from the fixed perspective of cinematography (i.e., the viewer watches the entire content from a given position) to highly mobile ones such as digital signage (i.e., the individual moves by the apparatus). Finally, perceptual fatigue is of key importance, especially considering effects such as dizziness or the loss of focus. After all, numerous contexts may necessitate extended periods of uninterrupted usage—particularly professional ones.

There are numerous important research questions associated with angular super resolution. First of all, the different perceptual thresholds must be exhaustively studied at different viewing distances and other viewing conditions. It is also vital to investigate the phenomenon of shifting between such high angular resolution and lower values and how it may affect the various content types. It is possible that one may need to rely on trade-offs in order to achieve angular super resolution. In such a case, the impact of the trade-offs should be examined both independently and in conjuncture with the achieved angular density.

Regarding the standardization of LF QoE, according to the current IEEE recommendation [60], the viewing distance threshold is defined as

$$VDT = \frac{ID}{\tan(AR)}, \quad (1)$$

where ID is the average interpupillary distance, and AR is the angular resolution of LF visualization. This threshold defines the maximum theoretical distance at which two distinct light rays with respect to a single point on the screen of the display may address the two pupils—hence, enabling proper 3D perception of the visualized content without the need for movement (e.g., sideways movement in the case of HOP visualization). Equation (1)

accounts for the average interpupillary distance as 6.5 cm. However, in the context of angular super resolution, the interpupillary distance should be replaced by pupil size, since two distinct light rays with respect to a single point on the screen address only one pupil in such case. In subjective tests that aim to study angular super resolution, various lighting conditions and display brightness values should be investigated, as the size of the pupil typically varies between 2 mm and 8 mm, depending on the intensity of the light. The novel results that are to be obtained by subjective studies may provide the foundations of new standards of LF QoE.

9.4. Capture and Display Hardware

Considering wide-baseline LF capture systems, the in-portability and cost of the camera setup further complicates its usage in dynamic movements. This problem may be tackled by using virtual cameras instead of using real cameras, simulating realistic camera movements [1,3]. On that note, more LF camera animations need to be investigated virtually and tested on LFDs.

Regarding real LF capture systems, future work including the manufacture of specialized rigs—supporting multiple cameras in the case of wide-baseline—should be considered. This should take into account the different camera array setups, including the horizontal placement of cameras in a line and an arc, spherical camera arrays, and 2D camera grids. The weight and dimensions of the rig should be considered specifically for dynamic camera motions: the impact of weight on camera motion; the effect of rig dimension (particularly extreme sizes) on camera motion; and camera motions that should be avoided for large-scale solutions.

As the vast majority of LFDs at the time of writing this paper are HOP, it would be important to investigate the differences in the perceived quality of camera animation between HOP and FP solutions. As of now, camera animation has been solely implemented for HOP visualization, and thus, no data regarding the visualization on FP systems are available.

Furthermore, in order to achieve angular super resolution for dynamic LF contents via real cameras, the capture system must adhere to the requirements enforced by angular super resolution (as detailed in Section 4), which, in turn, creates a particularly great challenge. In essence, for camera arrays, the number of cameras along a line or arc is constrained by the physical size of the cameras.

Regarding feasible super resolution for LFDs, Table 4 examines the characteristics of different LFDs with regards to angular resolution and the associated maximum viewing distance for angular super resolution. If we replace the average interpupillary distance with the diameter of the pupil—which ranges between 2 mm and 4 mm in bright light and between 4 mm and 8 mm in the dark [199–202]—Equation (1) for viewing distance threshold can be reformulated for angular super resolution as follows:

$$VDT_{SR} = \frac{PD}{\tan(AR)}, \quad (2)$$

where PD is the diameter of the pupil. For this analysis, we used 8 mm for the threshold calculation, as any viable viewing distance should be smaller than what is listed in the table.

Note that the table does not contain LFDs that are glasses-based (e.g., AR LFD [203]) and those that do not have their angular resolution values precisely specified (e.g., those that simply state that “hundreds of views are supported” [204,205]). The angular resolution values are expressed in the degree format, which means that the lower values indicate higher density for distinct light rays. As explained earlier, the feasible viewing distance for the utilization of the LFD in a given context should be smaller than the maximum viewing distance for super resolution, as it assumes the darkest lighting conditions. If the diameter of the pupil was taken as 4 mm, then the values in the table would be halved. Even with 8 mm, the maximum viewing distance does not exceed 1 m. While such displays may be used in single-user scenarios, there are other factors that may impact feasibility. For example, in the case of the HoloVizio C80 [206], as the large-scale LFD is a front-projection

display, viewing the screen from the distance specified in Table 4 would potentially result in invalid LF, as the observer's body would block the light rays coming from the projector array. Additionally, FOV and screen dimensions are listed in the table as well, since they are also crucial to feasibility. For instance, if the viewing distance threshold is low, than a low FOV can severely limit the VVA—and thus, the maximum number of simultaneous observers and their mobility—which is already constrained. Regarding screen dimensions, large screens are not necessarily made to be observed from a close distance and may come with other limitations, as exhibited previously. Based on the available information and the analysis above, we can conclude that LFD solutions at the time of writing this paper are not feasible for use cases that incorporate angular super resolution.

Table 4. LFD characteristics for angular super resolution.

LFD	Angular Resolution	FOV	Screen Dimensions	Maximum Viewing Distance for Super Resolution
Lume Pad 2 [207]	10.75°	86°	12.4"	4.21 cm
HoloVizio 80WLT [208]	1°	180°	30"	45.83 cm
HoloVizio 640RC [62]	0.8°	100°	72"	57.29 cm
Looking Glass Portrait [209]	0.58°	58°	7.9"	79.03 cm
Looking Glass Go [210]	0.58°	160° overall / 58° optimal	60"	79.03 cm
Looking Glass 65" [211]	0.53°	53°	65"	86.48 cm
Looking Glass 32" Spatial Display [212]	0.53°	53°	32"	86.48 cm
Looking Glass 16" Spatial Display [213]	0.53°	150° overall / 53° optimal	16"	86.48 cm
HoloVizio 722RC [214]	0.5°	70°	72"	91.67 cm
HoloVizio C80 [206]	0.5°	40°	140"	91.67 cm

As for spatial super resolution, generally, 2D images captured by conventional cameras have significantly higher resolution compared to LF images captured by LF cameras. In efforts to solve the super resolution problem for LF images, multiple solutions are suggested, including the usage of CNNs [215], subspace projection [216], and others. Most solutions, however, consider static images without addressing LF camera animations. Consequently, this shall pose greater challenges, primarily due to the huge amounts of captured data.

10. Limitations of the Work

One major limitation of our work is that at the time of writing this paper, achievements and findings related to LF camera animation are rather scarce, as the first steps towards the investigated topic have only been recently taken. This is partially due to the fact that currently only a handful of research institutions have access to real LFDs, which greatly limits the potential output. Furthermore, while LF camera animation is indeed relevant for future use cases, many research efforts target more fundamental issues that are yet to be tackled. Another limitation of our work is the lack of available data on LF camera animation with visual contents captured by real cameras. Achieving the associated results in the future may grant a more sophisticated perspective—empowered by objective data and subjective ratings—for such a review.

11. Conclusions

In this paper, the implications, limitations, potentials, and future research efforts of LF camera animation were discussed. We conclude that there are numerous research efforts to

be conducted, as the investigation of the topic only began recently. Some of the limitations are severe—particularly related to visualization blurriness—which may impact a great number of use cases and deployment potentials. The creation of content databases for research (i.e., LF videos with camera animation) is necessary for the efficient implementation of relevant subjective studies for quality assessment. Most of such content is expected to be captured by virtual cameras due to the challenges related to real capture systems. Eventually, real contents shall be necessitated as well, which requires the development of novel capture solutions.

Section 9 is dedicated to the relevant future research efforts. As additional guidance to what is already detailed in the section, the authors of this work would like to highlight the importance of continuous, iterative testing on LFDs, which may provide valuable feedback regarding the feasibility of LF camera animation in various use cases. However, as access to LFDs at the time of writing this paper is rather limited, alternative solutions may be explored (i.e., visualization of LF content on different display technologies). However, it should be noted that while such research may indeed be viable, the obtainable findings may be somewhat misleading, due to the inherent differences in the resulting visuals. Moreover, testing should aim to address extended periods of usage as well, since slighter perceptual impacts may accumulate over time. Finally, datasets dedicated to specific LFD models (i.e., the characteristics of the LF dataset match the characteristics of the LFD) may be preferable for numerous experiments, and the creation of novel datasets is encouraged to accommodate commercially available LFDs.

Author Contributions: Conceptualization, M.G. and P.A.K.; methodology, M.G. and P.A.K.; validation, M.G. and P.A.K.; investigation, M.G. and P.A.K.; resources, M.G. and P.A.K.; writing—original draft preparation, M.G. and P.A.K.; writing—review and editing, M.G. and P.A.K.; visualization, M.G.; supervision, P.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors would like to thank Holografika Ltd. and Tibor Balogh in particular for supporting this work.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AFR	Adaptive feature remixing
CNN	Convolutional neural network
CTU	Coding tree unit
EPI	Epipolar plane image
FOV	Field of view
FP	Full parallax
HCI	Human–computer interaction
HOP	Horizontal-only parallax
HLRA	Homography-based low-rank approximation
HVS	Human visual system
GAN	Generative adversarial network
LF	Light field
LFCNN	Light field convolution neural network
LFD	Light field display
LF-DFnet	Light field—deformable convolution network
LF-IINet	Light field—intra–inter view interaction network

MPI	Multiplane image
MPR	Multipolar reformatations
POV	Point of view
QoE	Quality of experience
ROI	Region of interest
SAI	Sub-aperture images
SADN	Spatial-angular-decorrelated network
VVA	Valid viewing area

References

- Guindy, M.; Barsi, A.; Kara, P.A.; Balogh, T.; Simon, A. Realistic physical camera motion for light field visualization. In Proceedings of the Holography: Advances and Modern Trends VII, Online, 19–29 April 2021; SPIE: Bellingham, WA, USA, 2021; Volume 11774, pp. 70–77.
- Guindy, M.; Kara, P.A.; Balogh, T.; Simon, A. Perceptual preference for 3D interactions and realistic physical camera motions on light field displays. In Proceedings of the Virtual, Augmented, and Mixed Reality (XR) Technology for Multi-Domain Operations III, Orlando, FL, USA, 3–7 April 2022; SPIE: Bellingham, WA, USA, 2022; Volume 12125, pp. 156–164.
- Guindy, M.; Barsi, A.; Kara, P.A.; Adhikarla, V.K.; Balogh, T.; Simon, A. Camera animation for immersive light field imaging. *Electronics* **2022**, *11*, 2689. [\[CrossRef\]](#)
- Gershun, A. The light field. *J. Math. Phys.* **1939**, *18*, 51–151. [\[CrossRef\]](#)
- da Vinci, L. *The Notebooks of Leonardo da Vinci*; Richter, J.P., Ed.; Courier Corporation: Chelmsford, MA, USA, 1970; Volume 2.
- Faraday, M. LIV. Thoughts on ray-vibrations. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1846**, *28*, 345–350. [\[CrossRef\]](#)
- Ives, F.E. Parallax Stereogram and Process of Making Same. US725567A, 14 April 1903.
- Lippmann, G. Epreuves reversibles donnant la sensation du relief. *J. Phys. Theor. Appl.* **1908**, *7*, 821–825. [\[CrossRef\]](#)
- Levoy, M.; Hanrahan, P. Light field rendering. In Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, New Orleans, LA, USA, 4–9 August 1996; pp. 31–42.
- Balram, N.; Tošić, I. Light-field imaging and display systems. *Inf. Disp.* **2016**, *32*, 6–13. [\[CrossRef\]](#)
- Wetzstein, G. Computational Plenoptic Image Acquisition and Display. Ph.D. Thesis, University of British Columbia, Vancouver, BC, Canada, 2011.
- Wu, G.; Masia, B.; Jarabo, A.; Zhang, Y.; Wang, L.; Dai, Q.; Chai, T.; Liu, Y. Light field image processing: An overview. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 926–954. [\[CrossRef\]](#)
- McMillan, L.; Bishop, G. Plenoptic modeling: An image-based rendering system. In Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles, CA, USA, 6–11 August 1995; pp. 39–46.
- Shum, H.Y.; Kang, S.B.; Chan, S.C. Survey of image-based representations and compression techniques. *IEEE Trans. Circuits Syst. Video Technol.* **2003**, *13*, 1020–1037. [\[CrossRef\]](#)
- Adelson, E.H.; Bergen, J.R. *The Plenoptic Function and the Elements of Early Vision*; Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology: Cambridge, MA, USA, 1991; Volume 2.
- McMillan, L.; Bishop, G. Plenoptic modeling: An image-based rendering system. In *Seminal Graphics Papers: Pushing the Boundaries*; Association for Computing Machinery: New York, NY, USA, 2023; Volume 2, pp. 433–440.
- Ng, R.; Levoy, M.; Brédif, M.; Duval, G.; Horowitz, M.; Hanrahan, P. *Light Field Photography with a Hand-Held Plenoptic Camera*; Technical Report; Stanford University, Stanford, CA, USA, 2024. Available online: <https://hci.stanford.edu/cstr/reports/2005-02.pdf> (accessed on 26 July 2024).
- Ijsselstein, W.A.; Seuntjens, P.J.; Meesters, L.M. Human factors of 3D displays. In *3D Videocommunication: Algorithms, Concepts and Real-Time Systems in Human Centred Communication*; Wiley Library: Hoboken, NJ, USA, 2005; pp. 217–233.
- Kara, P.A.; Tamboli, R.R.; Cserkaszky, A.; Barsi, A.; Simon, A.; Kusz, A.; Bokor, L.; Martini, M.G. Objective and subjective assessment of binocular disparity for projection-based light field displays. In Proceedings of the 2019 International Conference on 3D Immersion (IC3D), Brussels, Belgium, 11 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–8.
- Guindy, M.; Adhikarla, V.K.; Kara, P.A.; Balogh, T.; Simon, A. CLASSROOM: Synthetic high dynamic range light field dataset. In Proceedings of the Applications of Digital Image Processing XLV, San Diego, CA, USA, 22–24 August 2022; SPIE: Bellingham, WA, USA, 2022; Volume 12226, pp. 153–162.
- Sung, K.; Shirley, P.; Baer, S. *Essentials of Interactive Computer Graphics: Concepts and Implementation*; CRC Press: Boca Raton, FL, USA, 2008.
- Darukumalli, S.; Kara, P.A.; Barsi, A.; Martini, M.G.; Balogh, T. Subjective quality assessment of zooming levels and image reconstructions based on region of interest for light field displays. In Proceedings of the 2016 International Conference on 3D Imaging (IC3D), Liege, Belgium, 13–14 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–6.
- Darukumalli, S.; Kara, P.A.; Barsi, A.; Martini, M.G.; Balogh, T.; Chehaibi, A. Performance comparison of subjective assessment methodologies for light field displays. In Proceedings of the 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Limassol, Cyprus, 12–14 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 28–33.
- Magnor, M.; Girod, B. Data compression for light-field rendering. *IEEE Trans. Circuits Syst. Video Technol.* **2000**, *10*, 338–343. [\[CrossRef\]](#)

25. Girod, B.; Chang, C.L.; Ramanathan, P.; Zhu, X. Light field compression using disparity-compensated lifting. In Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003, Proceedings, (ICASSP'03), Hong Kong, China, 6–10 April 2003; Volume 4, pp. IV–760.
26. Jagmohan, A.; Sehgal, A.; Ahuja, N. Compression of lightfield rendered images using coset codes. In Proceedings of the The Thirtieth-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 9–12 November 2003; IEEE: Piscataway, NJ, USA, 2003; Volume 1, pp. 830–834.
27. Chen, W.C.; Bouguet, J.Y.; Chu, M.H.; Grzeszczuk, R. Light field mapping: Efficient representation and hardware rendering of surface light fields. *ACM Trans. Graph. (TOG)* **2002**, *21*, 447–456. [[CrossRef](#)]
28. Zhu, X.; Aaron, A.; Girod, B. Distributed compression for large camera arrays. In Proceedings of the IEEE Workshop on Statistical Signal Processing, St. Louis, MO, USA, 28 September–1 October 2003; IEEE: Piscataway, NJ, USA, 2003; pp. 30–33.
29. Li, Y.; Sjöström, M.; Olsson, R.; Jennehag, U. Efficient intra prediction scheme for light field image compression. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 539–543.
30. Li, Y.; Sjöström, M.; Olsson, R.; Jennehag, U. Scalable coding of plenoptic images by using a sparse set and disparities. *IEEE Trans. Image Process.* **2015**, *25*, 80–91. [[CrossRef](#)]
31. Perra, C. Lossless plenoptic image compression using adaptive block differential prediction. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, Australia, 19–24 April 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1231–1234.
32. Li, Y.; Olsson, R.; Sjöström, M. Compression of unfocused plenoptic images using a displacement intra prediction. In Proceedings of the 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Seattle, WA, USA, 11–15 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–4.
33. Conti, C.; Nunes, P.; Soares, L.D. HEVC-based light field image coding with bi-predicted self-similarity compensation. In Proceedings of the 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Seattle, WA, USA, 11–15 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–4.
34. Monteiro, R.; Lucas, L.; Conti, C.; Nunes, P.; Rodrigues, N.; Faria, S.; Pagliari, C.; Da Silva, E.; Soares, L. Light field HEVC-based image coding using locally linear embedding and self-similarity compensated prediction. In Proceedings of the 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Seattle, WA, USA, 11–15 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–4.
35. Liu, D.; Wang, L.; Li, L.; Xiong, Z.; Wu, F.; Zeng, W. Pseudo-sequence-based light field image compression. In Proceedings of the 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Seattle, WA, USA, 11–15 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–4.
36. Jiang, X.; Le Pendu, M.; Farrugia, R.A.; Guillemot, C. Light field compression with homography-based low-rank approximation. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 1132–1145. [[CrossRef](#)]
37. Chen, J.; Hou, J.; Chau, L.P. Light field compression with disparity-guided sparse coding based on structural key views. *IEEE Trans. Image Process.* **2017**, *27*, 314–324. [[CrossRef](#)] [[PubMed](#)]
38. Zhao, Z.; Wang, S.; Jia, C.; Zhang, X.; Ma, S.; Yang, J. Light field image compression based on deep learning. In Proceedings of the 2018 IEEE International conference on multimedia and expo (ICME), San Diego, CA, USA, 23–27 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
39. Dib, E.; Le Pendu, M.; Guillemot, C. Light field compression using Fourier disparity layers. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 3751–3755.
40. Huang, X.; An, P.; Chen, Y.; Liu, D.; Shen, L. Low bitrate light field compression with geometry and content consistency. *IEEE Trans. Multimed.* **2020**, *24*, 152–165. [[CrossRef](#)]
41. Chen, Y.; An, P.; Huang, X.; Yang, C.; Liu, D.; Wu, Q. Light field compression using global multiplane representation and two-step prediction. *IEEE Signal Process. Lett.* **2020**, *27*, 1135–1139. [[CrossRef](#)]
42. Liu, D.; Huang, X.; Zhan, W.; Ai, L.; Zheng, X.; Cheng, S. View synthesis-based light field image compression using a generative adversarial network. *Inf. Sci.* **2021**, *545*, 118–131. [[CrossRef](#)]
43. Tong, K.; Jin, X.; Wang, C.; Jiang, F. SADN: Learned light field image compression with spatial-angular decorrelation. In Proceedings of the ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 23–27 May 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1870–1874.
44. Jin, P.; Jiang, G.; Chen, Y.; Jiang, Z.; Yu, M. Perceptual Light Field Image Coding with CTU Level Bit Allocation. In Proceedings of the International Conference on Computer Analysis of Images and Patterns, Limassol, Cyprus, 25–28 September 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 255–264.
45. Kawakami, M.; Tsutake, C.; Takahashi, K.; Fujii, T. Compressing Light Field as Multiplane Image. *ITE Trans. Media Technol. Appl.* **2023**, *11*, 27–33. [[CrossRef](#)]
46. Shi, J.; Xu, Y.; Guillemot, C. Learning Kernel-Modulated Neural Representation for Efficient Light Field Compression. *arXiv* **2023**, arXiv:2307.06143.
47. Magnor, M.A.; Endmann, A.; Girod, B. Progressive Compression and Rendering of Light Fields. In Proceedings of the VMV, Saarbrücken, Germany, 22–24 November 2000; Citeseer: Princeton, NJ, USA, 2000; pp. 199–204.

48. Aggoun, A. A 3D DCT compression algorithm for omnidirectional integral images. In Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, Toulouse, France, 14–19 May 2006; IEEE: Piscataway, NJ, USA, 2006; Volume 2.
49. Dong, X.; Qionghan, D.; Wenli, X. Data compression of light field using wavelet packet. In Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No. 04TH8763), Taipei, Taiwan, 27–30 June 2004; IEEE: Piscataway, NJ, USA, 2004; Volume 2, pp. 1071–1074.
50. Chang, C.L.; Zhu, X.; Ramanathan, P.; Girod, B. Light field compression using disparity-compensated lifting and shape adaptation. *IEEE Trans. Image Process.* **2006**, *15*, 793–806. [\[CrossRef\]](#)
51. Aggoun, A. Compression of 3D integral images using 3D wavelet transform. *J. Disp. Technol.* **2011**, *7*, 586–592. [\[CrossRef\]](#)
52. Kundu, S. Light field compression using homography and 2D warping. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 1349–1352.
53. Conti, C.; Kovács, P.T.; Balogh, T.; Nunes, P.; Soares, L.D. Light-field video coding using geometry-based disparity compensation. In Proceedings of the 2014 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), Budapest, Hungary, 2–4 July 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 1–4.
54. Jin, X.; Han, H.; Dai, Q. Image reshaping for efficient compression of plenoptic content. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 1173–1186. [\[CrossRef\]](#)
55. Dai, F.; Zhang, J.; Ma, Y.; Zhang, Y. Lenselet image compression scheme based on subaperture images streaming. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 4733–4737.
56. Vieira, A.; Duarte, H.; Perra, C.; Tavora, L.; Assuncao, P. Data formats for high efficiency coding of Lytro-Illum light fields. In Proceedings of the 2015 International Conference on Image Processing Theory, Tools and Applications (IPTA), Orleans, France, 10–13 November 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 494–497.
57. Li, L.; Li, Z.; Li, B.; Liu, D.; Li, H. Pseudo-sequence-based 2-D hierarchical coding structure for light-field image compression. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 1107–1119. [\[CrossRef\]](#)
58. Shao, J.; Bai, E.; Jiang, X.; Wu, Y. Light-Field Image Compression Based on a Two-Dimensional Prediction Coding Structure. *Information* **2024**, *15*, 339. [\[CrossRef\]](#)
59. Kara, P.A.; Tamboli, R.R.; Shafiee, E.; Martini, M.G.; Simon, A.; Guindy, M. Beyond perceptual thresholds and personal preference: Towards novel research questions and methodologies of quality of experience studies on light field visualization. *Electronics* **2022**, *11*, 953. [\[CrossRef\]](#)
60. IEEE P3333.1.4-2022; Recommended Practice for the Quality Assessment of Light Field Imaging. IEEE Standards Association: Piscataway, NJ, USA, 2023. Available online: <https://standards.ieee.org/ieee/3333.1.4/10873/> (accessed on 26 July 2024).
61. Balogh, T. The holovizio system. In Proceedings of the Stereoscopic Displays and Virtual Reality Systems XIII, San Jose, CA, USA, 15–19 January 2006; SPIE: Bellingham, WA, USA, 2006; Volume 6055, pp. 279–290.
62. Balogh, T.; Kovács, P.T.; Barsi, A. Holovizio 3D display system. In Proceedings of the 2007 3DTV Conference, Kos, Greece, 7–9 May 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 1–4.
63. Megyesi, Z.; Barsi, A.; Balogh, T. 3D Video Visualization on the Holovizio System. In Proceedings of the 2008 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, Istanbul, Turkey, 28–30 May 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 269–272.
64. Balogh, T.; Kovács, P. Holovizio: The next generation of 3D oil & gas visualization. In Proceedings of the 70th EAGE Conference and Exhibition-Workshops and Fieldtrips, Rome, Italy, 9–12 June 2008; European Association of Geoscientists & Engineers: Utrecht, The Netherlands, 2008.
65. Balogh, T.; Kovács, P.T.; Dobrányi, Z.; Barsi, A.; Megyesi, Z.; Gaál, Z.; Balogh, G. The Holovizio system—New opportunity offered by 3D displays. In Proceedings of the TMCE, Kusadasi, Turkey, 21–25 April 2008; pp. 79–89.
66. Balogh, T.; Kovács, P.T. Real-time 3D light field transmission. In Proceedings of the Real-Time Image and Video Processing, Brussels, Belgium, 16 April 2010; SPIE: Bellingham, WA, USA, 2010; Volume 7724, pp. 53–59.
67. Balogh, T.; Nagy, Z.; Kovács, P.T.; Adhikarla, V.K. Natural 3D content on glasses-free light-field 3D cinema. In Proceedings of the Stereoscopic Displays and Applications XXIV, Burlingame, CA, USA, 3–7 February 2013; SPIE: Bellingham, WA, USA, 2013; Volume 8648, pp. 103–110.
68. Kovács, P.T.; Balogh, T. 3D light-field display technologies. In *Emerging Technologies for 3D Video: Creation, Coding, Transmission and Rendering*; Wiley Library: Hoboken, NJ, USA, 2013; pp. 336–345.
69. Kara, P.A.; Cserkaszy, A.; Darukumalli, S.; Barsi, A.; Martini, M.G. On the edge of the seat: Reduced angular resolution of a light field cinema with fixed observer positions. In Proceedings of the 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), Erfurt, Germany, 31 May–2 June 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
70. Cserkaszy, A.; Kara, P.A.; Tamboli, R.R.; Barsi, A.; Martini, M.G.; Bokor, L.; Balogh, T. Angularly continuous light-field format: Concept, implementation, and evaluation. *J. Soc. Inf. Disp.* **2019**, *27*, 442–461. [\[CrossRef\]](#)
71. Tamboli, R.; Vupparaboina, K.K.; Ready, J.; Jana, S.; Channappayya, S. A subjective evaluation of true 3D images. In Proceedings of the 2014 International Conference on 3D Imaging (IC3D), Liege, Belgium, 9–10 December 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 1–8.

72. Tamboli, R.R.; Appina, B.; Channappayya, S.; Jana, S. Super-multiview content with high angular resolution: 3D quality assessment on horizontal-parallax lightfield display. *Signal Process. Image Commun.* **2016**, *47*, 42–55. [\[CrossRef\]](#)
73. Tamboli, R.R.; Appina, B.; Channappayya, S.S.; Jana, S. Achieving high angular resolution via view synthesis: Quality assessment of 3D content on super multiview lightfield display. In Proceedings of the 2017 International Conference on 3D Immersion (IC3D), Brussels, Belgium, 11–12 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–8.
74. Ahar, A.; Chlipala, M.; Birnbaum, T.; Zaperty, W.; Symeonidou, A.; Kozacki, T.; Kujawinska, M.; Schelkens, P. Suitability analysis of holographic vs light field and 2D displays for subjective quality assessment of Fourier holograms. *Opt. Express* **2020**, *28*, 37069–37091. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Dricot, A.; Jung, J.; Cagnazzo, M.; Pesquet, B.; Dufaux, F.; Kovács, P.T.; Adhikarla, V.K. Subjective evaluation of Super Multi-View compressed contents on high-end light-field 3D displays. *Signal Process. Image Commun.* **2015**, *39*, 369–385. [\[CrossRef\]](#)
76. Cserkaszkzy, A.; Barsi, A.; Kara, P.A.; Martini, M.G. To interpolate or not to interpolate: Subjective assessment of interpolation performance on a light field display. In Proceedings of the 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Hong Kong, China, 10–14 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 55–60.
77. Kovács, P.T.; Lackner, K.; Barsi, A.; Balázs, Á.; Boev, A.; Bregović, R.; Gotchev, A. Measurement of perceived spatial resolution in 3D light-field displays. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 768–772.
78. Kovács, P.T.; Bregović, R.; Boev, A.; Barsi, A.; Gotchev, A. Quantifying spatial and angular resolution of light-field 3-D displays. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 1213–1222. [\[CrossRef\]](#)
79. Kara, P.A.; Guindy, M.; Xinyu, Q.; Szakal, V.A.; Balogh, T.; Simon, A. The effect of angular resolution and 3D rendering on the perceived quality of the industrial use cases of light field visualization. In Proceedings of the 2022 16th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Dijon, France, 19–21 October 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 600–607.
80. Tamboli, R.R.; Kara, P.A.; Cserkaszkzy, A.; Barsi, A.; Martini, M.G.; Jana, S. Canonical 3D object orientation for interactive light-field visualization. In Proceedings of the Applications of Digital Image Processing XLI, San Diego, CA, USA, 19–23 August 2018; SPIE: Bellingham, WA, USA, 2018; Volume 10752, pp. 77–83.
81. Adhikarla, V.K.; Jakus, G.; Sodnik, J. Design and evaluation of freehand gesture interaction for light field display. In Proceedings of the Human-Computer Interaction: Interaction Technologies: 17th International Conference, HCI International 2015, Los Angeles, CA, USA, 2–7 August 2015; Proceedings, Part II 17; Springer: Berlin/Heidelberg, Germany, 2015; pp. 54–65.
82. Adhikarla, V.K.; Sodnik, J.; Szolgay, P.; Jakus, G. Exploring direct 3D interaction for full horizontal parallax light field displays using leap motion controller. *Sensors* **2015**, *15*, 8642–8663. [\[CrossRef\]](#)
83. Zhang, X.; Braley, S.; Rubens, C.; Merritt, T.; Vertegaal, R. LightBee: A self-levitating light field display for hologrammatic telepresence. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, 4–9 May 2019; pp. 1–10.
84. Cserkaszkzy, A.; Barsi, A.; Nagy, Z.; Puhr, G.; Balogh, T.; Kara, P.A. Real-time light-field 3D telepresence. In Proceedings of the 2018 7th European Workshop on Visual Information Processing (EUVIP), Tampere, Finland, 26–28 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–5.
85. Shafiee, E.; Martini, M.G. Datasets for the quality assessment of light field imaging: Comparison and future directions. *IEEE Access* **2023**, *11*, 15014–15029. [\[CrossRef\]](#)
86. Vaish, V.; Adams, A. *The (New) Stanford Light Field Archive*; Computer Graphics Laboratory, Stanford University: Stanford, CA, USA, 2008; Volume 6, p. 3.
87. Rerabek, M.; Yuan, L.; Authier, L.A.; Ebrahimi, T. [ISO/IEC JTC 1/SC 29/WG1 Contribution] EPFL Light-Field Image Dataset 2015; p. 3. Available online: <https://www.epfl.ch/labs/mmspg/downloads/epfl-light-field-image-dataset/> (accessed on 26 July 2024).
88. Shekhar, S.; Kunz Beigpour, S.; Ziegler, M.; Chwesiuk, M.; Paleń, D.; Myszkowski, K.; Keinert, J.; Mantiuk, R.; Didyk, P. Light-field intrinsic dataset. In Proceedings of the British Machine Vision Conference 2018 (BMVC), British Machine Vision Association, Newcastle, UK, 3–6 September 2018.
89. Tamboli, R.R.; Reddy, M.S.; Kara, P.A.; Martini, M.G.; Channappayya, S.S.; Jana, S. A high-angular-resolution turntable data-set for experiments on light field visualization quality. In Proceedings of the 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), Cagliari, Italy, 29 May–1 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–3.
90. Ellahi, W.; Vigier, T.; Le Callet, P. Analysis of public light field datasets for visual quality assessment and new challenges. In Proceedings of the European Light Field Imaging Workshop, Borovets, Bulgaria, 4–6 June 2019.
91. Static Planar Light-Field Test Dataset. Available online: <https://www.iis.fraunhofer.de/en/ff/amm/dl/lightfielddataset.html> (accessed on 26 July 2024).
92. Guillo, L.; Jiang, X.; Lafruit, G.; Guillemot, C. *ISO/IEC JTC1/SC29/WG1 & WG11; Light Field Video Dataset Captured by a R8 Raytrix Camera (with Disparity Maps)*; International Organisation for Standardisation: Geneva, Switzerland, 2018. Available online: <http://clim.inria.fr/Datasets/RaytrixR8Dataset-5x5/index.html> (accessed on 26 July 2024).
93. Dansereau, D.G.; Girod, B.; Wetzstein, G. LiFF: Light field features in scale and depth. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8042–8051.

94. Moreschini, S.; Gama, F.; Bregovic, R.; Gotchev, A. CIVIT datasets: Horizontal-parallax-only densely-sampled light-fields. In Proceedings of the European Light Field Imaging Workshop, Borovets, Bulgaria, 4–6 June 2019; Volume 6, pp. 1–4.
95. Zakeri, F.S.; Durmush, A.; Ziegler, M.; Bätz, M.; Keinert, J. Non-planar inside-out dense light-field dataset and reconstruction pipeline. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1059–1063.
96. Gul, M.S.K.; Wolf, T.; Bätz, M.; Ziegler, M.; Keinert, J. A high-resolution high dynamic range light-field dataset with an application to view synthesis and tone-mapping. In Proceedings of the 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), London, UK, 6–10 July 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
97. Yue, D.; Gul, M.S.K.; Bätz, M.; Keinert, J.; Mantiuk, R. A benchmark of light field view interpolation methods. In Proceedings of the 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), London, UK, 6–10 July 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
98. Rerabek, M.; Ebrahimi, T. New light field image dataset. In Proceedings of the 8th International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, Portugal, 6–8 June 2016.
99. Wanner, S.; Meister, S.; Goldluecke, B. Datasets and benchmarks for densely sampled 4D light fields. In Proceedings of the VMV, Saarbrücken, Germany, 3–6 September 2013; Volume 13, pp. 225–226.
100. Mousnier, A.; Vural, E.; Guillemot, C. Partial light field tomographic reconstruction from a fixed-camera focal stack. *arXiv* **2015**, arXiv:1503.01903.
101. Honauer, K.; Johannsen, O.; Kondermann, D.; Goldluecke, B. A dataset and evaluation methodology for depth estimation on 4D light fields. In Proceedings of the Computer Vision—ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, 20–24 November 2016; Revised Selected Papers, Part III 13; Springer: Berlin/Heidelberg, Germany, 2017; pp. 19–34.
102. Sabater, N.; Boisson, G.; Vandame, B.; Kerbiriou, P.; Babon, F.; Hog, M.; Gendrot, R.; Langlois, T.; Bureller, O.; Schubert, A.; et al. Dataset and pipeline for multi-view light-field video. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 30–40.
103. Ahmad, W.; Palmieri, L.; Koch, R.; Sjöström, M. Matching light field datasets from plenoptic cameras 1.0 and 2.0. In Proceedings of the 2018-3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), Helsinki, Finland, 3–5 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–4.
104. Kim, C.; Zimmer, H.; Pritch, Y.; Sorkine-Hornung, A.; Gross, M.H. Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans. Graph.* **2013**, *32*, 1–12. [\[CrossRef\]](#)
105. Hu, X.; Wang, C.; Pan, Y.; Liu, Y.; Wang, Y.; Liu, Y.; Zhang, L.; Shirmohammadi, S. 4DLFVD: A 4D light field video dataset. In Proceedings of the 12th ACM Multimedia Systems Conference, Istanbul, Turkey, 28 September–1 October 2021; pp. 287–292.
106. Srinivasan, P.P.; Wang, T.; Sreelal, A.; Ramamoorthi, R.; Ng, R. Learning to synthesize a 4D RGBD light field from a single image. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2243–2251.
107. Wang, T.C.; Zhu, J.Y.; Hiroaki, E.; Chandraker, M.; Efros, A.A.; Ramamoorthi, R. A 4D light-field dataset and CNN architectures for material recognition. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Proceedings, Part III 14; Springer: Berlin/Heidelberg, Germany, 2016; pp. 121–138.
108. The Plenoptic 2.0 Toolbox: Benchmarking of Depth Estimation Methods for MLA-Based Focused Plenoptic Cameras. Available online: https://zenodo.org/records/3558284#_YeXpMHrP2Hs (accessed on 26 July 2024).
109. Kiran Adhikarla, V.; Vinkler, M.; Sumin, D.; Mantiuk, R.K.; Myszkowski, K.; Seidel, H.P.; Didyk, P. Towards a quality metric for dense light fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 58–67.
110. Viola, I.; Ebrahimi, T. VALID: Visual quality assessment for light field images dataset. In Proceedings of the 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), Cagliari, Italy, 29 May–1 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–3.
111. Shi, L.; Zhao, S.; Zhou, W.; Chen, Z. Perceptual evaluation of light field image. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 41–45.
112. Schambach, M.; Heizmann, M. A Multispectral Light Field Dataset for Light Field Deep Learning. *IEEE Access* **2020**, *8*, 193492–193502. [\[CrossRef\]](#)
113. Zizien, A.; Fliegel, K. LFDD: Light field image dataset for performance evaluation of objective quality metrics. In Proceedings of the Applications of Digital Image Processing XLIII, Online, CA, USA, 24 August–4 September 2020; SPIE: Bellingham, WA, USA, 2020; Volume 11510, pp. 671–683.
114. Paudyal, P.; Battisti, F.; Sjöström, M.; Olsson, R.; Carli, M. Towards the perceptual quality evaluation of compressed light field images. *IEEE Trans. Broadcast.* **2017**, *63*, 507–522. [\[CrossRef\]](#)
115. Shan, L.; An, P.; Liu, D.; Ma, R. Subjective evaluation of light field images for quality assessment database. In Proceedings of the Digital TV and Wireless Multimedia Communication: 14th International Forum, IFTC 2017, Shanghai, China, 8–9 November 2017; Revised Selected Papers 14; Springer: Berlin/Heidelberg, Germany, 2018; pp. 267–276.
116. Nava, F.P.; Luke, J. Simultaneous estimation of super-resolved depth and all-in-focus images from a plenoptic camera. In Proceedings of the 2009 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, Potsdam, Germany, 4–6 May 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 1–4.

117. Lim, J.; Ok, H.; Park, B.; Kang, J.; Lee, S. Improving the spatail resolution based on 4D light field data. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 1173–1176.
118. Georgiev, T.; Chunev, G.; Lumsdaine, A. Superresolution with the focused plenoptic camera. In Proceedings of the Computational Imaging IX, San Francisco, CA, USA, 23–27 January 2011; SPIE: Bellingham, WA, USA, 2011; Volume 7873, pp. 232–244.
119. Liang, C.K.; Ramamoorthi, R. A light transport framework for lenslet light field cameras. *ACM Trans. Graph. (TOG)* **2015**, *34*, 1–19. [[CrossRef](#)]
120. Bishop, T.E.; Favaro, P. The light field camera: Extended depth of field, aliasing, and superresolution. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *34*, 972–986. [[CrossRef](#)] [[PubMed](#)]
121. Mitra, K.; Veeraraghavan, A. Light field denoising, light field superresolution and stereo camera based refocussing using a GMM light field patch prior. In Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, RI, USA, 16–21 June 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 22–28.
122. Wanner, S.; Goldluecke, B. Variational light field analysis for disparity estimation and super-resolution. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *36*, 606–619. [[CrossRef](#)] [[PubMed](#)]
123. Rossi, M.; Frossard, P. Graph-based light field super-resolution. In Proceedings of the 2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), Luton, UK, 16–18 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
124. Rossi, M.; El Gheche, M.; Frossard, P. A nonsmooth graph-based approach to light field super-resolution. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 2590–2594.
125. Alain, M.; Smolic, A. Light field super-resolution via LFBM5D sparse coding. In Proceedings of the 2018 25th IEEE international conference on image processing (ICIP), Athens, Greece, 7–10 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 2501–2505.
126. Farag, S.; Velisavljevic, V. A novel disparity-assisted block matching-based approach for super-resolution of light field images. In Proceedings of the 2018-3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), Helsinki, Finland, 3–5 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–4.
127. Fan, H.; Liu, D.; Xiong, Z.; Wu, F. Two-stage convolutional neural network for light field super-resolution. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1167–1171.
128. Wang, Y.; Liu, F.; Zhang, K.; Hou, G.; Sun, Z.; Tan, T. LFNet: A novel bidirectional recurrent convolutional neural network for light-field image super-resolution. *IEEE Trans. Image Process.* **2018**, *27*, 4274–4286. [[CrossRef](#)] [[PubMed](#)]
129. Wang, Y.; Yang, J.; Wang, L.; Ying, X.; Wu, T.; An, W.; Guo, Y. Light field image super-resolution using deformable convolution. *IEEE Trans. Image Process.* **2020**, *30*, 1057–1071. [[CrossRef](#)]
130. Zhang, S.; Lin, Y.; Sheng, H. Residual networks for light field image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 11046–11055.
131. Farrugia, R.A.; Guillemot, C. Light field super-resolution using a low-rank prior and deep convolutional neural networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 1162–1175. [[CrossRef](#)]
132. Liu, G.; Yue, H.; Wu, J.; Yang, J. Intra-inter view interaction network for light field image super-resolution. *IEEE Trans. Multimed.* **2021**, *25*, 256–266. [[CrossRef](#)]
133. Mo, Y.; Wang, Y.; Xiao, C.; Yang, J.; An, W. Dense dual-attention network for light field image super-resolution. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *32*, 4431–4443. [[CrossRef](#)]
134. Zhang, S.; Chang, S.; Lin, Y. End-to-end light field spatial super-resolution network using multiple epipolar geometry. *IEEE Trans. Image Process.* **2021**, *30*, 5956–5968. [[CrossRef](#)] [[PubMed](#)]
135. Van Duong, V.; Huu, T.N.; Yim, J.; Jeon, B. Light field image super-resolution network via joint spatial-angular and epipolar information. *IEEE Trans. Comput. Imaging* **2023**, *9*, 350–366. [[CrossRef](#)]
136. Yoon, Y.; Jeon, H.G.; Yoo, D.; Lee, J.Y.; Kweon, I.S. Light-field image super-resolution using convolutional neural network. *IEEE Signal Process. Lett.* **2017**, *24*, 848–852. [[CrossRef](#)]
137. Wang, Y.; Wang, L.; Yang, J.; An, W.; Yu, J.; Guo, Y. Spatial-angular interaction for light field image super-resolution. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part XXIII 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 290–308.
138. Ko, K.; Koh, Y.J.; Chang, S.; Kim, C.S. Light field super-resolution via adaptive feature remixing. *IEEE Trans. Image Process.* **2021**, *30*, 4114–4128. [[CrossRef](#)] [[PubMed](#)]
139. Brown, B. *Cinematography: Theory and Practice: Image Making for Cinematographers and Directors*; Taylor & Francis: Milton Park, Oxfordshire, 2016.
140. Schell, J. *The Art of Game Design: A Book of Lenses*; CRC Press: Boca Raton, FL, USA, 2008.
141. Callenbach, E. *The Five C's of Cinematography: Motion Picture Filming Techniques Simplified by Joseph V. Mascelli*; Silman-James Press: West Hollywood, CA, USA, 1966.
142. Kara, P.A.; Barsi, A.; Tamboli, R.R.; Guindy, M.; Martini, M.G.; Balogh, T.; Simon, A. Recommendations on the viewing distance of light field displays. In Proceedings of the Digital Optical Technologies 2021, Online Only, Germany, 21–26 June 2021; SPIE: Bellingham, WA, USA, 2021; Volume 11788, pp. 166–179.

143. Kara, P.A.; Simon, A. The Good News, the Bad News, and the Ugly Truth: A Review on the 3D Interaction of Light Field Displays. *Multimodal Technol. Interact.* **2023**, *7*, 45. [\[CrossRef\]](#)
144. Guindy, M.; Barsi, A.; Kara, P.A.; Balogh, T.; Simon, A. Interaction methods for light field displays by means of a theater model environment. In Proceedings of the Holography: Advances and Modern Trends VII, Online Only, Czech Republic, 19–29 April 2021; SPIE: Bellingham, WA, USA, 2021; Volume 11774, pp. 109–118.
145. iMARE CULTURE. 2020. Available online: <https://imareculture.eu/> (accessed on 26 July 2024).
146. Rotter, P. Why did the 3D revolution fail?: The present and future of stereoscopy [commentary]. *IEEE Technol. Soc. Mag.* **2017**, *36*, 81–85. [\[CrossRef\]](#)
147. Pei, Z.; Li, Y.; Ma, M.; Li, J.; Leng, C.; Zhang, X.; Zhang, Y. Occluded-object 3D reconstruction using camera array synthetic aperture imaging. *Sensors* **2019**, *19*, 607. [\[CrossRef\]](#) [\[PubMed\]](#)
148. Xu, Y.; Maeno, K.; Nagahara, H.; Taniguchi, R.i. Camera array calibration for light field acquisition. *Front. Comput. Sci.* **2015**, *9*, 691–702. [\[CrossRef\]](#)
149. Goldlücke, B.; Klehm, O.; Wanner, S.; Eisemann, E.; Cameras, P. *Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality*; CRC Press: Boca Raton, FL, USA, 2015; pp. 67–79. Available online: <http://www.crcpress.com/product/isbn/9781482243819> (accessed on 26 July 2024).
150. Cserkaszkzy, A.; Kara, P.A.; Tamboli, R.R.; Barsi, A.; Martini, M.G.; Balogh, T. Light-field capture and display systems: Limitations, challenges, and potentials. In Proceedings of the Novel Optical Systems Design and Optimization XXI, International Society for Optics and Photonics, San Diego, CA, USA, 19–23 August 2018.
151. Yang, J.C.; Everett, M.; Buehler, C.; McMillan, L. A real-time distributed light field camera. *Render. Tech.* **2002**, *2002*, 2.
152. Popovic, V.; Afshari, H.; Schmid, A.; Leblebici, Y. Real-time implementation of Gaussian image blending in a spherical light field camera. In Proceedings of the 2013 IEEE international conference on industrial technology (ICIT), Cape Town, Western Cape, South Africa, 25–28 February 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 1173–1178.
153. Gortler, S.J.; Grzeszczuk, R.; Szeliski, R.; Cohen, M.F. The lumigraph. In Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, New York, NY, USA, 4–9 August 1996; SIGGRAPH'96; pp. 43–54.
154. Taguchi, Y.; Agrawal, A.; Ramalingam, S.; Veeraraghavan, A. Axial light field for curved mirrors: Reflect your perspective, widen your view. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 499–506.
155. Liang, C.K.; Lin, T.H.; Wong, B.Y.; Liu, C.; Chen, H.H. Programmable aperture photography: Multiplexed light field acquisition. In *ACM Siggraph 2008 Papers*; ACM, Inc.: New York, NY, USA, 2008; pp. 1–10.
156. Adelson, E.H.; Wang, J.Y. Single lens stereo with a plenoptic camera. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 99–106. [\[CrossRef\]](#)
157. Okano, F.; Arai, J.; Hoshino, H.; Yuyama, I. Three-dimensional video system based on integral photography. *Opt. Eng.* **1999**, *38*, 1072–1077. [\[CrossRef\]](#)
158. Ihrke, I.; Stich, T.; Gottschlich, H.; Magnor, M.; Seidel, H.P. Fast incident light field acquisition and rendering. *J. WSCG* **2008**, *16*, 25–32.
159. Zhang, C.; Chen, T. Light field capturing with lensless cameras. In Proceedings of the IEEE International Conference on Image Processing, Genoa, Italy, 11–14 September 2005; IEEE: Piscataway, NJ, USA, 2005; Volume 3.
160. Georgeiv, T.; Zheng, K.C.; Curless, B.; Salesin, D.; Nayar, S.; Intwala, C. Spatio-Angular Resolution Tradeoffs in Integral Photography. In *Symposium on Rendering*; Akenine-Moeller, T., Heidrich, W., Eds.; The Eurographics Association: London, UK, 2006.
161. Ueda, K.; Koike, T.; Takahashi, K.; Naemura, T. Adaptive integral photography imaging with variable-focus lens array. In Proceedings of the Stereoscopic Displays and Applications XIX, San Jose, CA, USA, 27–31 January 2008; SPIE: Bellingham, WA, USA, 2008; Volume 6803, pp. 443–451.
162. Ueda, K.; Lee, D.; Koike, T.; Takahashi, K.; Naemura, T. Multi-focal compound eye: Liquid lens array for computational photography. In Proceedings of the ACM SIGGRAPH 2008 New Tech Demos, New York, NY, USA, 11–15 August 2008; SIGGRAPH'08.
163. Unger, J.; Wenger, A.; Hawkins, T.; Gardner, A.; Debevec, P.E. Capturing and Rendering with Incident Light Fields. *Render. Tech.* **2003**, *2003*, 1–10.
164. Levoy, M.; Chen, B.; Vaish, V.; Horowitz, M.; McDowall, I.; Bolas, M. Synthetic aperture confocal imaging. *ACM Trans. Graph. (ToG)* **2004**, *23*, 825–834. [\[CrossRef\]](#)
165. Lanman, D.; Crispell, D.; Wachs, M.; Taubin, G. Spherical catadioptric arrays: Construction, multi-view geometry, and calibration. In Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06), Chapel Hill, NC, USA, 14–16 June 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 81–88.
166. Taguchi, Y.; Agrawal, A.; Veeraraghavan, A.; Ramalingam, S.; Raskar, R. Axial-cones: Modeling spherical catadioptric cameras for wide-angle light field rendering. *ACM Trans. Graph.* **2010**, *29*, 172. [\[CrossRef\]](#)
167. Ogata, S.; Ishida, J.; Sasano, T. Optical sensor array in an artificial compound eye. *Opt. Eng.* **1994**, *33*, 3649–3655.
168. Tanida, J.; Kumagai, T.; Yamada, K.; Miyatake, S.; Ishida, K.; Morimoto, T.; Kondou, N.; Miyazaki, D.; Ichioka, Y. Thin observation module by bound optics (TOMBO): Concept and experimental verification. *Appl. Opt.* **2001**, *40*, 1806–1813. [\[CrossRef\]](#)
169. Tanida, J.; Shogenji, R.; Kitamura, Y.; Yamada, K.; Miyamoto, M.; Miyatake, S. Color imaging with an integrated compound imaging system. *Opt. Express* **2003**, *11*, 2109–2117. [\[CrossRef\]](#)

170. Hiura, S.; Mohan, A.; Raskar, R. Krill-eye: Superposition compound eye for wide-angle imaging via grin lenses. *IPSI Trans. Comput. Vis. Appl.* **2010**, *2*, 186–199. [CrossRef]
171. Yang, J.; Lee, C.; Isaksen, A.; McMillan, L. A Low-Cost Portable Light Field Capture Device. In Proceedings of the Siggraph Conference Abstracts and Applications, New Orleans, LA, USA, 23–28 July 2000.
172. Raytrix: 3D Light Field Vision. Available online: <https://raytrix.de/> (accessed on 26 July 2024).
173. Veeraraghavan, A.; Raskar, R.; Agrawal, A.; Mohan, A.; Tumblin, J. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.* **2007**, *26*, 69. [CrossRef]
174. Hahne, C.; Aggoun, A.; Velisavljevic, V.; Fiebig, S.; Pesch, M. Baseline and triangulation geometry in a standard plenoptic camera. *Int. J. Comput. Vis.* **2018**, *126*, 21–35. [CrossRef]
175. Georgiev, T.; Intwala, C. Light Field Camera Design for Integral View Photography, Adobe System. 2006. Available online: <https://www.tgeorgiev.net/IntegralView.pdf> (accessed on 26 July 2024).
176. Jeon, H.G.; Park, J.; Choe, G.; Park, J.; Bok, Y.; Tai, Y.W.; So Kweon, I. Accurate depth map estimation from a lenslet light field camera. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1547–1555.
177. Kara, P.A.; Kovacs, P.T.; Vagharshakyan, S.; Martini, M.G.; Barsi, A.; Balogh, T.; Chuchvara, A.; Chehaibi, A. The effect of light field reconstruction and angular resolution reduction on the quality of experience. In Proceedings of the 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Naples, Italy, 28 November–1 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 781–786.
178. Höhne, K.H.; Fuchs, H.; Pizer, S.M. *3D Imaging in Medicine: Algorithms, Systems, Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; Volume 60.
179. Chan, S.; Conti, F.; Salisbury, K.; Blevins, N.H. Virtual reality simulation in neurosurgery: Technologies and evolution. *Neurosurgery* **2013**, *72*, 154–164. [CrossRef]
180. Ferrosi, P.; Tringali, G.; Acerbi, F.; Schiariti, M.; Broggi, M.; Aquino, D.; Broggi, G. Advanced 3-dimensional planning in neurosurgery. *Neurosurgery* **2013**, *72*, 54–62. [CrossRef]
181. Langdon, W.B.; Modat, M.; Petke, J.; Harman, M. Improving 3D medical image registration CUDA software with genetic programming. In Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation, Vancouver, BC, Canada, 12–16 July 2014; pp. 951–958.
182. Cserkaszy, A.; Kara, P.A.; Barsi, A.; Martini, M.G. The potential synergies of visual scene reconstruction and medical image reconstruction. In Proceedings of the Novel Optical Systems Design and Optimization XXI, San Diego, CA, USA, 19–23 August 2018; SPIE: Bellingham, WA, USA, 2018; Volume 10746, pp. 19–25.
183. Robinson, M.S.; Brylow, S.; Tschimmel, M.; Humm, D.; Lawrence, S.; Thomas, P.; Denevi, B.W.; Bowman-Cisneros, E.; Zerr, J.; Ravine, M.; et al. Lunar reconnaissance orbiter camera (LROC) instrument overview. *Space Sci. Rev.* **2010**, *150*, 81–124. [CrossRef]
184. Yan, Z.; Wang, C.; Yan, Z.; Wang, F. Research Summary on Light Field Display Technology Based on Projection. In Proceedings of the 2020 International Conference on Machine Learning and Computer Application, Shangri-La, China, 11–13 September 2020; IOP Publishing: Bristol, UK, 2020; Volume 1682.
185. Diewald, S.; Möller, A.; Roalter, L.; Kranz, M. DriveAssist—A V2X-Based Driver Assistance System for Android. In *Mensch & Computer 2012—Workshopband: Interaktiv Informiert—Allgegenwärtig und Allumfassend!?* Oldenbourg Verlag: Munich, Germany, 2012.
186. Olaverri-Monreal, C.; Jizba, T. Human factors in the design of human–machine interaction: An overview emphasizing V2X communication. *IEEE Trans. Intell. Veh.* **2016**, *1*, 302–313. [CrossRef]
187. Xu, T.; Jiang, R.; Wen, C.; Liu, M.; Zhou, J. A hybrid model for lane change prediction with V2X-based driver assistance. *Phys. A Stat. Mech. Its Appl.* **2019**, *534*, 122033. [CrossRef]
188. Hirai, T.; Murase, T. Performance evaluations of PC5-based cellular-V2X mode 4 for feasibility analysis of driver assistance systems with crash warning. *Sensors* **2020**, *20*, 2950. [CrossRef] [PubMed]
189. Kara, P.A.; Wippelhauser, A.; Balogh, T.; Bokor, L. How I met your V2X sensor data: Analysis of projection-based light field visualization for vehicle-to-everything communication protocols and use cases. *Sensors* **2023**, *23*, 1284. [CrossRef]
190. Kara, P.A.; Tamboli, R.R.; Adhikarla, V.K.; Balogh, T.; Guindy, M.; Simon, A. Connected without disconnection: Overview of light field metaverse applications and their quality of experience. *Displays* **2023**, *78*, 102430. [CrossRef]
191. Kara, P.A.; Cserkaszy, A.; Martini, M.G.; Barsi, A.; Bokor, L.; Balogh, T. Evaluation of the concept of dynamic adaptive streaming of light field video. *IEEE Trans. Broadcast.* **2018**, *64*, 407–421. [CrossRef]
192. Kara, P.A.; Tamboli, R.R.; Cserkaszy, A.; Martini, M.G.; Barsi, A.; Bokor, L. The viewing conditions of light-field video for subjective quality assessment. In Proceedings of the 2018 International Conference on 3D Immersion (IC3D), Brussels, Belgium, 5 December 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–8.
193. Kovács, P.T.; Boev, A.; Bregovic, R.; Gotchev, A. Quality Measurement Of 3D Light-Field Displays. In Proceedings of the Eight International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM-2014, Chandler, AZ, USA, 30–31 January 2014. Available online: <https://researchportal.tuni.fi/publications/quality-measurement-of-3d-light-field-displays> (accessed on 26 July 2024).
194. Turing, A.M. Computing machinery and intelligence. *Mind* **1950**, *59*, 433–460. [CrossRef]
195. Banks, M.S.; Hoffman, D.M.; Kim, J.; Wetzstein, G. 3D Displays. *Annu. Rev. Vis. Sci.* **2016**, *2*, 397–435. [CrossRef]

196. Hamilton, M.; Wells, N.; Soares, A. On Requirements for Field of Light Displays to Pass the Visual Turing Test. In Proceedings of the 2022 IEEE International Symposium on Multimedia (ISM), Naples, Italy, 5–7 December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 86–87.
197. Hopper, D.G. 1000 X difference between current displays and capability of human visual system: Payoff potential for affordable defense systems. In Proceedings of the Cockpit Displays VII: Displays for Defense Applications, Orlando, FL, USA, 24–28 April 2000; SPIE: Bellingham, WA, USA, 2000; Volume 4022, pp. 378–389.
198. Curry, D.G.; Martinsen, G.L.; Hopper, D.G. Capability of the human visual system. *Cockpit Displays X* **2003**, 5080, 58–69.
199. Ellis, C. The pupillary light reflex in normal subjects. *Br. J. Ophthalmol.* **1981**, 65, 754–759. [[CrossRef](#)] [[PubMed](#)]
200. Walker, H.K.; Hall, W.D.; Hurst, J.W. *Clinical Methods: The History, Physical, and Laboratory Examinations*; Butterworth-Heinemann: Oxford, UK, 1990. Available online: https://www.acpjournals.org/doi/10.7326/0003-4819-113-7-563_2 (accessed on 26 July 2024).
201. Atchison, D.A.; Markwell, E.L.; Kasthurirangan, S.; Pope, J.M.; Smith, G.; Swann, P.G. Age-related changes in optical and biometric characteristics of emmetropic eyes. *J. Vis.* **2008**, 8, 29. [[CrossRef](#)] [[PubMed](#)]
202. Bradley, M.M.; Miccoli, L.; Escrig, M.A.; Lang, P.J. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* **2008**, 45, 602–607. [[CrossRef](#)] [[PubMed](#)]
203. Sluka, T.; Kvasov, A.; Kubes, T. Digital Light-Field. 2021. Available online: <https://creal.com/app/uploads/2022/04/CREAL-White-Paper-Digital-Light-field.pdf> (accessed on 26 July 2024).
204. ELF-SR1 Spatial Reality Display - Sony Pro. Available online: https://pro.sony/ue_US/products/spatial-reality-displays/elf-sr1#TEME502131AllYouNeedIsYourEyes-elf-sr1 (accessed on 26 July 2024).
205. ELF-SR2 Spatial Reality Display - Sony Pro. Available online: https://pro.sony/ue_US/products/spatial-reality-displays/elf-sr2 (accessed on 26 July 2024).
206. HoloVizio C80 Glasses-Free 3D Cinema System. Available online: <https://holografika.com/c80-glasses-free-3d-cinema/> (accessed on 26 July 2024).
207. Lume Pad 2. Available online: <https://www.leiainc.com/lume-pad-2> (accessed on 26 July 2024).
208. HoloVizio 80WLT Full-Angle 3D Displaying. Available online: <https://holografika.com/80wlt/> (accessed on 26 July 2024).
209. Looking Glass Portrait. Available online: <https://lookingglassfactory.com/looking-glass-portrait> (accessed on 26 July 2024).
210. Looking Glass Go. Available online: <https://lookingglassfactory.com/looking-glass-go> (accessed on 26 July 2024).
211. Looking Glass 65". Available online: <https://lookingglassfactory.com/looking-glass-65> (accessed on 26 July 2024).
212. Looking Glass 32" Spatial Display. Available online: <https://lookingglassfactory.com/looking-glass-32> (accessed on 26 July 2024).
213. Looking Glass 16" Spatial Display. Available online: <https://lookingglassfactory.com/16-spatial-oled> (accessed on 26 July 2024).
214. HoloVizio 722RC Large-Scale 3D Displaying. Available online: <https://holografika.com/722rc/> (accessed on 26 July 2024).
215. Yeung, H.W.F.; Hou, J.; Chen, X.; Chen, J.; Chen, Z.; Chung, Y.Y. Light field spatial super-resolution using deep efficient spatial-angular separable convolution. *IEEE Trans. Image Process.* **2018**, 28, 2319–2330. [[CrossRef](#)]
216. Farrugia, R.A.; Galea, C.; Guillemot, C. Super resolution of light field images using linear subspace projection of patch-volumes. *IEEE J. Sel. Top. Signal Process.* **2017**, 11, 1058–1071. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.