Displays 78 (2023) 102430

Contents lists available at ScienceDirect

Displays

journal homepage: www.elsevier.com/locate/displa

Connected without disconnection: Overview of light field metaverse applications and their quality of experience☆

Peter A. Kara [a,b,*], Roopak R. Tamboli [c], Vamsi K. Adhikarla [d], Tibor Balogh [e], Mary Guindy [d,e], Aniko Simon [f]

[a] *Department of Networked Systems and Services, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, 1111, Budapest, Hungary*
[b] *Department of Computer Science, School of Computer Science and Mathematics, Faculty of Science, Engineering and Computing, Kingston University, Penrhyn Road Campus, Kingston upon Thames, KT1 2EE, London, UK*
[c] *Department of Electrical Engineering, Indian Institute of Technology Hyderabad, 502284, Kandi, Sangareddy, India*
[d] *Faculty of Information Technology and Bionics, Pazmany Peter Catholic University, 1083, Budapest, Hungary*
[e] *Holografika, 1192, Budapest, Hungary*
[f] *Sigma Technology, 1093, Budapest, Hungary*

ARTICLE INFO

Keywords:
Light field
Metaverse
Quality of experience
Immersion

ABSTRACT

With the rapid technological advances of the recent years, the practical instances of the metaverse have become more immersive and engaging than ever before. In their most frequent forms, the 3D virtual worlds of the metaverse are enabled by virtual reality headsets. This means that the user is completely disconnected from its real environment and is fully immersed in such a virtual world. The quality of experience of virtual reality and other headset-based technologies is now definitely a hot research topic, and the findings of the relevant scientific efforts are continuously emerging. However, as a headset-free immersive 3D technology, light field visualization is greatly underinvestigated with regard to the concept of the metaverse. In this paper, we address the applications of light field metaverse, compare its advantages and disadvantages to more conventional metaverse technologies, and discuss the most important issues regarding user experience. The paper highlights the user-oriented considerations for the development of general-purpose and dedicated light field displays. Additionally, our work examines state-of-the-art display systems and the current feasibility of a light field metaverse.

1. Introduction

The term "metaverse" first appeared 30 years ago, when American writer Neal Stephenson published his science fiction novel *Snow Crash* [1]. The book describes the computer-generated universe as an urban environment, which is basically composed of the sides of a single road that runs around an artificial planet. Access to the book's metaverse is possible through a global communication network, via either a virtual reality (VR) headset or a dedicated terminal (i.e., a booth). Those connected to the metaverse appear in the forms of customized avatars, equalized in height.

The metaverse depicted by Stephenson is a familiar one today. Since 1992, computer technology and global connectivity have developed immensely, and many implementation efforts throughout the years have pointed towards the aforementioned description of a metaverse. But what exactly makes a network of 3D virtual worlds a metaverse? The Merriam-Webster dictionary defines it as "a persistent virtual environment that allows access to and interoperability of multiple individual virtual realities". There are numerous application-focused definitions as well, with gaming and entertainment in their centers. Many refer to "Second Life" – launched in 2003 – as one of the first real metaverse, although there were already notable attempts in the decade of Stephenson's book.

On the level of user equipment, personal computers, smartphones, and tablets are relevant, particularly since massively multiplayer online role-playing games (MMORPG) are often classified as instances of the metaverse concept, due to their different options for social interaction. Evidently, VR headsets are looked at as the de facto equipment for the metaverse – similarly to Stephenson's book – as they disconnect the

☆ This paper was recommended for publication by Xiongkuo Min.

* Corresponding author at: Department of Networked Systems and Services, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, 1111, Budapest, Hungary.
E-mail address: kara@hit.bme.hu (P.A. Kara).

https://doi.org/10.1016/j.displa.2023.102430
Received 30 September 2022; Received in revised form 7 March 2023; Accepted 14 March 2023
Available online 25 March 2023

users from their surroundings and immerse them in an artificial reality. Augmented reality (AR) and mixed reality (MR) also play a notable role in today's implementation efforts. Both rely on headsets – similarly to VR – and retain a portion of the real world. This latter statement means that users are not completely disconnected from their surroundings, as in the case of VR.

However, as it has been already stated, both VR, AR and MR use headsets to provide a connection to the metaverse. Headsets, in general, are not only cumbersome – like stereoscopic 3D glasses – but they also have numerous drawbacks, such as VR sickness [2–5] and general simulator sickness or cybersickness [6,7]. Of course, the fact that VR headsets disconnect the users from the real world is indeed a significant advantage in terms of immersion and the sense of presence, yet it should be noted that sensing the surroundings has its benefits – e.g., it lowers the risk of potential accidents. AR and MR solutions do enable perceptual access to the real world, but such can be counterproductive for immersion. Technically, the more a person perceives from its surroundings, the more difficult it is to immerse the individual; and the more a person's view is "enhanced" – and thus obscured – the less reliable movement and interaction becomes in the real world.

VR, AR, and XR are not the only 3D visualization technologies that benefit from the continuous stream of research efforts. Another noteworthy form of 3D imaging is light field, which is the realization of the hundred-year-old dream of Lippmann regarding integral imaging [8]. Projection-based light field technology offers glasses-free 3D visualization. It is completely different from near-eye light field displays [9–14], which, as the name suggests, place the plane of visualization near the eye. Projection-based light field displays have projectors either behind or in front of the holographic screen. If the projector array is behind the screen (i.e., the viewer is on the other side of the screen), then we call the system a back-projection display; if it is in front of the screen (i.e., the viewer is on the same side), then we call it a front-projection display. The first projection-based light field displays were created roughly two decades ago, and even though several innovative displays emerged over the years – including a tileable LED wall [15] and a quadcopter telepresence system [16] – such technology is still not widely available. In fact, real light field displays are only accessible to a handful of institutions at the time of writing this paper.

Light field displays are not to be confused with multi-view displays, which rely on so-called "sweet spots". It means that 3D visualization can be experienced only from a given set of positions, and the same perspectives may be seen from all positions. Instead, light field displays use the entire viewing area to visualize the content, and the perspective changes continuously and smoothly in accordance with the viewing angle.

Light field technology has all the potential to be utilized for the purpose of the metaverse. As no viewing gear is required and the parallax imitates real life, light field displays enable a more natural sense of 3D. However, it has not yet been considered as a metaverse technology. One could add that it was proposed by Fattal [17] earlier this year, but the presentation was narrowed down to light-field-based mobile devices.

In this paper, we address the suitability of projection-based light field displays in the context of metaverse applications. We analyze the advantages and disadvantages of such systems, and compare them to the more conventional metaverse technologies, such as VR. We detail the constraints of the technology, and the potentials that may be unlocked through research and development. We separately discuss general-purpose and dedicated displays, as they may greatly differ in terms of capabilities. Finally, the paper highlights present-day feasibility, from both a technological perspective and through the lens of Quality of Experience (QoE) [18].

The most prevalent motivations of this work are that users of a light field metaverse are not disconnected from their environment – contributing to personal safety – and other users in the vicinity – introducing potential social immersion – and that glasses-free 3D enables a more natural sense of 3D perception. Additionally, wearing viewing devices may be rather inconvenient for certain users. This is particularly relevant for the tactile sensitivity of individuals affected by autism spectrum disorder [19,20] and attention deficit hyperactivity disorder (ADHD) [21,22], and for those who suffer from anisometropia [23,24] or other medical conditions (e.g., astigmatism [25]). While wearing viewing gears may be just a minor nuisance for some, it may pose significant difficulties for others – which may even be prohibitive, isolating the individual from specific systems and services. As present and future technologies are meant to be as socially inclusive as possible, it is essential to take into consideration and to accommodate those with special needs.

The remainder of this paper is structured as follows. Section 2 briefly reviews the conventional metaverse technologies. Section 3 addresses projection-based light field displays in the context of the metaverse. Section 4 studies the applications of light field metaverse, the QoE of which is discussed in Section 5. Section 6 concludes the paper.

2. Conventional immersive metaverse technologies

In this section, we provide a concise review of the head-mounted solutions for the metaverse.

Let us begin with a brief history of head-mounted solutions [26,27]. In 1838, Sir Charles Wheatstone outlined the concept of "stereopsis" or "binocular vision", leading to the development of the first stereoscopes. In 1935, American science fiction writer Stanley Weinbaum published Pygmalion's Spectacles, in which the main character explores a fictional world using a pair of goggles. In 1956, cinematographer Morton Heilig created Sensorama, the first VR machine. This movie booth combined 3D, color video with audio, smells, and a vibrating chair to immerse the viewer in the movie. In 1960, Heilig patented the first head-mounted display. Also in 1965, Ivan Sutherland presented a paper outlining his concept of the "Ultimate Display"; a virtual world so realistic that the user would not be able to differentiate it from reality. This is widely considered to be the blueprint for modern VR. In 1968, he created the first AR headset. The headset was called "The Sword of Damocles", which displayed computer-generated graphics that enhanced the user's perception of the world. In 1977, Massachusetts Institute of Technology researchers introduced the Aspen Movie Map; a virtual tour of the streets of Aspen, which was created by using photographs taken by a car driving through the city. In 1979, the McDonnell-Douglas Corporation integrated VR into its head-mounted displays (HMDs) for military use. A head tracker in the HMD followed the pilot's eye movements to match CG images. During the 1980s and 1990s, several technologies were realized, including stereo glasses, wired gloves for tracking hands, early commercial VR goggles, virtual cockpit, VR training simulator for astronauts, VR arcades, VR gaming consoles (such as Nintendo Virtual Boy), and so on. The next impetus came with the launch of Oculus Rift in 2012, and its acquisition by Facebook in 2014. From this point, numerous companies began working on head-mounted devices. In the following, with the help of a few key examples, we explore each of these technologies.

VR is a computer-generated simulation of a realistic experience. Here, it is important to note that "computer-generated" is not to be confused with "synthetic" or "artificial", as the visualized content may be a recording of the real world. Typically, VR blocks out the real world and replaces it with a virtual world [28]. Again, the virtual world may be generated by a computer, or by interactively playing back recorded media. VR has applications in a variety of domains, such as gaming, education, entertainment, healthcare, and many more. The most common problems associated with VR headsets are headaches, nausea, eyestrain, and the possibility of running into objects or other people. The cybersickness caused by VR gears is closely connected to latency [29,30]. Despite its issues, VR is perceived as the de facto technology for the metaverse, and as authors of this manuscript, we expect it to maintain its dominance in the following years.

In the case of AR, the computer-generated content is added onto – or embedded into – the real world experience, so that both can be experienced together. Traditionally, AR was a headset-free experience, where computerized elements were added to the real world around the user. For example, AR was used to project tabletop games onto a physical surface, arrange digital furniture in a physical space, and drop animation (AR Stickers) into a camera shot. Pokemon Go is a notable example for AR experience. In most cases, AR may be viewed with an updated mobile device, equipped with a camera. Basic AR that overlays simple information over what the user is looking at operates satisfactorily with 3-degrees of freedom (DoF). However, most AR applications require 6-DoF in some form, tracking the users' physical position so that the system can maintain consistent positions for the images it projects in 3D space.

Almost all head-mounted devices offering 6-DoF are classified as MR headsets. While there is almost no confusion in the literature regarding what constitutes VR, there is hardly any agreement on what the terms AR, MR, and XR mean with regard to the reality-virtuality continuum [31]. Rauschnabel et al. recently attempted to provide a framework for AR and VR [32]. Similarly, Mann et al. introduced a taxonomy, frameworks, and conceptualization of various "realities" [28]. An extensive review of metaverse technologies, along with a discussion on the reality-virtuality continuum, appears in the work of Lee et al. [33]. For the purpose of this work, we classify smart glasses, smartphones, and similar devices as AR devices, and full-fledged HMDs as MR devices.

MR headsets are classified as optical see-through (OST) and video see-through (VST) [34,35]. In VST, a camera captures a digital video image of the real world, and transfers it to the graphics processor in real time. Then the graphics processor combines the video image feed with computer-generated images (i.e., virtual content) and displays it on the screen. Since the video processing is done before showing the content to the user, it is possible to control the brightness and contrast of both real-world and virtual elements for a seamless experience. It also has a better registration through the tracking of head movement. However, there are drawbacks, which include low resolution of reality (i.e., screens do not match the human eye resolution), a limited field of view (which is possible to increase but it is rather expensive), and eye parallax (eye-offset) due to the camera's position, which is usually at a distance from the viewers exact eye location. The VSTs can use smartphones like in the case of Samsung Gear VR, where the phone is used as a display, placed few inches away from users eyes, which is totally different from experiencing AR in handheld smartphones.

OST displays operate by using optical elements that are half transmissive and half reflective to combine real-world and virtual elements. The mirror allows a sufficient amount of light from the real world to pass through, making it possible to see the surroundings directly. Simultaneously, computer-generated images are projected onto the mirror through a display component placed overhead or on the side, which creates a perception of the combined world. OSTs show the world in real resolution and free from parallax (caused by the offset of camera position in relation to the viewer's eye). These are safer to operate since the user can see even if the power fails — making it an ideal option for military and medical purposes. However, the use of mirrors and lenses reduces the brightness and contrast of both virtual and real-world perception. Some of the recent cutting-edge MR headsets are Microsoft Hololens 2[1] and Magic Leap 2,[2] both of which are OST devices, and the Varjo XR-3,[3] which is a VST device.

[1] https://www.microsoft.com/en-us/hololens
[2] https://www.magicleap.com/device
[3] https://varjo.com/products/xr-3/

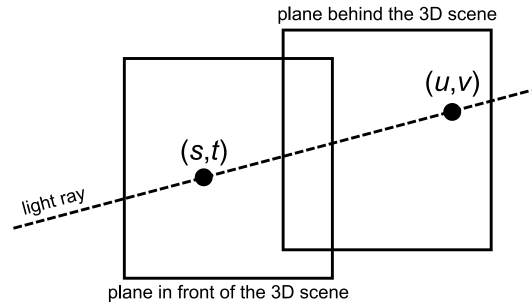

Fig. 1. Light field parametrization with parallel planes.

3. Light field as a metaverse technology

In this section, we overview light field visualization and its key performance indicators (KPIs), and address projection-based light field displays in the context of the metaverse, highlighting advantages, potentials, disadvantages, constraints, and recent advances. The section also provides a brief outlook regarding near-eye light field technology.

3.1. Operation principles

Unlike other visualization technologies of the metaverse, light field displays do not have active screens. This means that the screen of a light field display does not have active internal components. It is a curved surface built of a holographic micropattern, which transforms the incoming light rays. The light rays originate from the projector array, located either behind or in front of the screen — as explained in Section 1. The collection of transformed light rays composes the light field.

The easiest way to understand the description of each light ray is via the (s, t, u, v) coordinates of the plenoptic function. Let us imagine a finite 3D scene. Due to its finite nature, we can evidently place a plane behind the scene, and another one, parallel with the other, in front of the scene. Trivially, two coordinates identify a point on a 2D plane. If we take coordinates (s, t) on one plane, and coordinates (u, v) on the other plane, then connecting the two coordinate pairs (i.e., points) determines a line in the 3D space. The concept is illustrated on Fig. 1 [36].

It is not challenging to realize that a given (s, t, u, v) set of coordinates can only determine one specific line, and a given line can only be determined by one specific (s, t, u, v) set of coordinates, making the description free of redundancy. There are numerous alternative parameterizations as well, such as determining two points on the surface of a sphere, or using single point on a curved surface and describing the directions that originate from that point. Research efforts investigate these alternative solutions in practice, such as the angularly-continuous (s, t, ϕ) format of Cserkaszky et al. [37].

The glasses-free 3D visualization of projection-based light field displays is perceived as the light rays reach the pupils of the viewer. The 3D visual experience enabled by such displays is detailed in Section 5, as a crucial component of QoE. The QoE of light field visualization fundamentally depends on essential KPI values [38], which are reviewed in the following subsection.

3.2. Key performance indicators

There are numerous KPIs of light field imaging, describing both the system and the visualized content. In the scope of this paper, we briefly review the four KPIs that are the most relevant to the light field metaverse: spatial resolution, angular resolution, field of view (FOV), and depth budget.

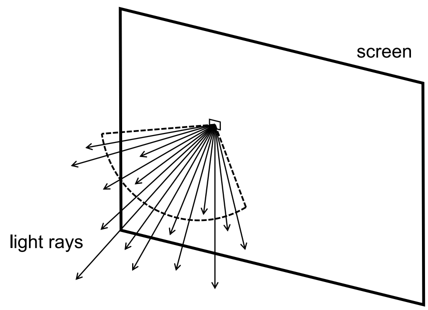

Fig. 2. Emission range for HOP displays.

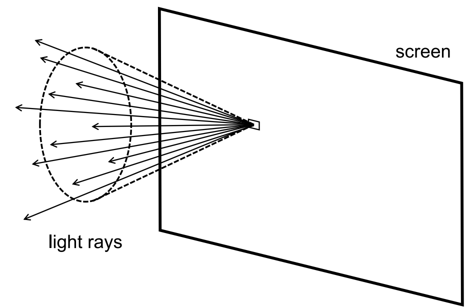

Fig. 3. Emission range for FP displays.

Spatial resolution can be approached as the 2D-equivalent resolution of visualization. It is, in fact, somewhat similar to the resolution of 2D imaging. However, certain distinctions must be made. First of all, light field visualization is not an inherently pixel-based technology. Even if we wanted to define pixels, their grid would not be uniform. Secondly, spatial resolution directly corresponds to pixels if and only if the visualized content is generated from a series of 2D views. In that case, spatial resolution is the 2D resolution of the views. For the system itself, spatial resolution is determined by the capabilities of the optical engines of the projector array. Generally, this KPI defines the sharpness of visualization, similarly to 2D technologies. However, insufficient levels of spatial resolution result in blur instead of pixellation. It needs to be noted that spatial resolution is relative, and it is the best at the plane of screen. As visualized objects or portions of the content are placed farther out of the plane, the smallest addressable volume – also known as voxel – becomes greater. Therefore, a more formalized definition of relative spatial resolution can be expressed as

$$R_S = \frac{1}{\frac{1}{R_0} + d \times tan(R_A)}, \tag{1}$$

where R_0 is the resolution in the plane of the screen, d is the distance from the plane of the screen, R_A is the angular resolution [39].

Angular resolution is generally one of the most important KPIs of light field visualization in all usage contexts. It defines the density of distinct rays. At the time of writing this paper, light field displays are typically horizontal-only parallax (HOP) systems, which means that they provide the corresponding perspective as the viewing angle changes horizontally, but not vertically. Systems that incorporate both horizontal and vertical change are known as full-parallax (FP) displays. The emission ranges – with respect to a single point on the screen – of HOP and FP displays are demonstrated in Figs. 2 and 3 [39]. Angular resolution may be conceptualized as the number of views when the input of the system is described as a series of 2D views. However, this approach depends on the FOV — which shall also be explained in this subsection. In the industry and the academia, there are two primary methods for defining angular resolution. The more general one is the degree which describes the smallest angle between two distinct rays – with respect to a given point on the screen – that the system may produce. So if the angular resolution is, for example, 0.5 degrees, then two tiny pinhole cameras looking at a point of the screen may see two distinct light rays if they enclose an angle of 0.5 degrees. This definition is illustrated in Fig. 4. Another way of describing the angular resolution is from the perspective of 2D source views. This is particularly useful if the content is indeed generated from a series of views, but this is the aforementioned case which depends on the FOV. If we have, for example, 90 source views over an FOV of 45 degrees, then the angular resolution is 2 views per degree. Note that 0.5 degrees and 2 views per degree define the same angular resolution.

The FOV of light field displays is not to be confused with the FOV of conventional 3D metaverse technologies. Generally speaking, headgear-based technologies approach the FOV from the perspective of the viewer. In the case of light field displays, FOV has nothing to do with the viewer. Instead, it describes the system. The FOV is an angle, measured from the screen of the display, in which the system creates an observable light field. A larger FOV enables more freedom of movement for both HOP and FP displays, and also allows the accommodation of more simultaneous viewers. It is important to add that while the FOV is described with a single value (i.e., the viewing angle of the display), it is actually the area on which all the emission cones – measured from the screen of the display, as shown in Fig. 2 – overlap. For the sake of simplicity, in the figures of Section 4 that aim to illustrate use cases, the valid light field of a system is not based on emission cones.

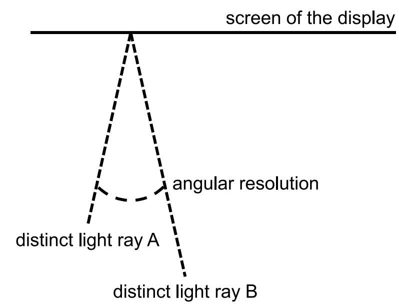

Fig. 4. Definition of angular resolution.

The final KPI that needs to be mentioned in the context of light field metaverse is the depth budget. Basically, it describes the maximum depth of visualization. Depth is the perceived distance between the plane of the screen and the front-most and back-most portions of the 3D content. As we typically refer to one-way depth as Field of Depth (FOD), depth budget is basically the twice of this measure. The FOD depends on the pixel size (S_P) – at the plane of the screen – and the angular resolution (R_A), and the relation can be formulated [40] as

$$FOD = \frac{S_P}{tan(R_A)}, \tag{2}$$

and therefore, the depth budget can be expressed as

$$DB = 2 \times \frac{S_P}{tan(R_A)}. \tag{3}$$

The KPI is called depth budget as it does not need to be fully utilized. There are, of course, certain reasons for avoiding great depth values. One is that light field visualization is always the sharpest at the plane of the screen. It is closely connected to this fact that content portions with greater depth values are more susceptible to issues related to low angular resolution. These considerations are discussed in detail in Section 5, along with the other components of the QoE of light field metaverse.

3.3. Advantages and potentials

There are many advantages of projection-based light field visualization, particularly in comparison with other 3D technologies. The first one is that the 3D contents of light field displays can be viewed without the need of any viewing device. As detailed in Section 2, such devices may have numerous shortcomings. Moreover, the lack of viewing devices also reinforces a more natural experience. After all, the whole concept behind light field imaging is that we only perceive light rays in our 3D reality, and the de facto most realistic scenario is thus the recreation of such light rays.

Again, the lack of viewing devices means the lack of the associated issues. Additionally, no viewer tracking or movement prediction [41] is required for light field visualization, as the entire system FOV is utilized all the time. This, of course, may be subject to research, as one may easily consider this to be an immense waste of resources. However, providing a continuous parallax and full visualization over the entire FOV not only nullifies the need for tracking, but also enables the accommodation of an arbitrary number of simultaneous viewers, limited only by the size of the valid viewing area (VVA). The VVA primarily depends on the FOV, as the correct 3D visualization may only be perceived within the FOV. However, it is influenced by angular resolution as well. Technically, if a viewer is too far away from the plane of the screen, then the visualized content may perceptually appear to be more 2D than 3D. Regarding the terminology of the VVA, experts often refer to the area defined by the FOV (i.e., the overlapping emission cones) as VVA, independently from the adequate viewing distance.

The systems available today are HOP displays, as stated earlier. FP displays hold much potential, as the vertical parallax component is required for a faithful representation of our world. However, for the vast majority of use cases – included the applications of the metaverse – horizontal parallax is considered to be sufficient. First of all, the eyes of a human being are separated horizontally. Secondly, movement – and thus change in viewing angle – in such contexts is usually horizontal. For instance, during an exhibition of cultural heritage, visitors can walk around the exhibited contents, and there is no notable change in vertical perspective if visualization is at more-or-less eye level. Of course, FP systems could benefit almost every single use case, and thus, they hold great potentials for the future.

Another important potential is achieving super resolution. In order to understand the phenomenon of super resolution, let us go back to Fig. 4. In this view of the projected light field, if we approach the eyes of a human being as pinhole cameras, and these cameras are fixed in both position and orientation, then the brain may enable a 3D experience if one eye collects the visual information of distinct light ray A, and the other one collects the visual information of distinct light ray B. It is necessary to highlight that the eyes do not change position (i.e., the head of the individual does not move) and orientation (i.e., the eyes do not change viewing direction), as they may both compensate the 3D experience. For super resolution, a single eye must collect the visual information of both distinct light rays A and B. If this is achieved, then the individual may focus on certain portions of the visualized content. Normally, the eyes focus on the plane of the screen, regardless of projected depth. Note that the perceived ray density fundamentally depends on the viewing distance; the farther the viewer is, the higher density is required. For the ultimate light field metaverse experience, super resolution shall be an absolute must, as it ensures that the focal behavior of the eyes of the user in the metaverse is analogous to their behavior in the real physical world.

3.4. Disadvantages

The first disadvantageous property most scientists and experts of the field think of regarding light field visualization is the cost of such display systems. It is true that at the time of writing this paper, high-performance projection-based light field displays are immensely expensive, particularly when compared to other metaverse technologies. While such expenses are expected to lessen over the upcoming years, yet, sadly, the price is not the only disadvantage of light field visualization.

Such display systems are also “expensive” when it comes to the amount of data that needs to be stored, transmitted, and processed. This issue may be mitigated by efficient compression and alternative formats, but the future realization of super resolution shall definitely increase the pressure posed by data size.

High-end light field displays also require more power to operate, and with great power, comes great responsibility towards heat management. Metaverse applications may be used continuously for extended periods of time, for which, the glasses-free nature of the technology is indeed an advantage. However, proper cooling is essential for both the computer cluster that manages the optical engines, and, of course, the optical engines themselves.

From a manufacturer’s point of view, it is extremely challenging to realize FP visualization and super resolution. More rays mean that more optical engines need to be placed in the projector array, which is not only an issue on its own due to the potential unit sizes, but it is also problematic for cooling.

From the perspective of the users of light field metaverse applications, the size and weight of such systems may be daunting. However, this may only be an issue for larger wide-baseline solutions. Additionally, it should be pointed out that all these issues are disadvantages that may be handled as the technology progresses. Yet there are constraints that are not expected to go away, and thus, should be kept in mind when designing applications of the light field metaverse.

3.5. Constraints

One major constraint of projection-based light field visualization is that such systems cannot draw virtually infinite distances in 3D. First of all, the (s, t, u, v) approach – similarly to all the other surface-based approaches – necessitates a finite 3D content space. Secondly, as mentioned earlier in the section, the more visualization leaves the plane of the screen, the more susceptible it becomes to insufficient angular density (i.e., the more effort is required towards maintaining its quality).

How does this constraint translate to the applications of light field metaverse? Let us imagine a 3D first-person shooter video game or simply any sandbox-like open-world game (either in first- or third-person view). It is common to have scenarios in open virtual areas, where the player may look towards virtually-infinite distances. Implementing such a scenario for light field visualization would result in greatly degraded contents. Today, there are two primary practices for handling the issue of having a “background” for the visualized content. The first one is simply using a plain color. This is quite common for the industrial use case of prototype review, particularly if the selected color contrasts the colors of the prototype. The other one is placing a 2D image near the plane behind the 3D scene. Of course, in this case, the image shall be immensely blurred, so it serves aesthetic purposes instead of practical ones.

The other major constraint is only applicable to front-projection displays. In such a setup, the optical engine array is on the same side of the display as the viewer. According to the designs of the state-of-the-art front-projection light field displays, it is actually possible to occlude the light rays coming from the optical engines, simply by being too close to the screen. When this happens, a shadow is cast on the screen, which should evidently be avoided. Therefore, for front-projection light field displays, the viewing distance is constrained. In essence, viewers should not approach the screen beyond a specified distance. As an example, whenever QoE research was carried out on the Holovizio C80 light field cinema system,[4] the closest viewing distance was aligned with the position of the optical engine array.

[4] https://holografika.com/c80-glasses-free-3d-cinema/

3.6. Recent advances

Let us now review the relevant advances of the past years in light field technology. Perceived quality is absolutely essential to the success of light field visualization, although practical measurements and experiments that rely on real light field displays are limited by the scarcity of such devices. Until early 2022, only 29 experiments were carried out [42]. Among the recent works are the experiments of Simon et al. [43], Guindy et al. [44], and Kara et al. [45], addressing adequate viewing distances for individuals with reduced visual capabilities, the application and assessment of conventional camera animation, and the effect of angular resolution and 3D rendering on static industrial models, respectively. Ak and Le Callet [46] reviewed the quality evaluation of light field visualization, which was also recently standardized by the IEEE [47].

A vital part of quality-related research is the creation of tools for objective quality assessment, also known as objective metrics. Their task is to estimate the subjectively perceived quality. By doing so, contents (e.g., screen contents [48]) and solutions (e.g., compression [49]) can be evaluated without the need of lengthy subjective tests with numerous test participants. Their efficiency is assessed via datasets (reference-quality and degraded contents) which also include subjective ratings [50]. Methods are classified based on the need for the reference. If it is fully needed, then it is a fully-reference (FR) metric. If it is only partially needed, then it is a reduced-reference (RR) metric. If it is not needed at all, then it is a no-reference (NR) metric. For example, Min et al. [51] created an FR metric that takes into consideration view structure matching, near-edge mean square error (MSE) and multi-view quality analysis, the RR metric of Paudyal et al. [52] measures distortion in the depth map, and the work of Meng et al. [53] proposes an NR metric based on angular–spatial characteristics. There are many NR metrics published in the scientific literature, as they are evidently the most practical – and, of course, the most challenging as well. NR metrics for light field quality assessment were proposed by Shi et al. [54], Zhou et al. [55], Shan et al. [56], Luo et al. [57], Guo et al. [58], Xiang et al. [59], Ak et al. [60], and many others. In the scientific world of QoE, FR solutions are also known as blind metrics [61–65], as they do not see what the best possible quality representation could be.

Many works were published on super resolution. However, note that in research, super resolution refers to content resolution enhancement. Furthermore, the super resolution discussed so far was defined in the angular domain. Among scientific contributions, enhancing spatial resolution is quite common as well. In fact, it is looked at as the default interpretation of term. In order to distinguish the two, the term "super resolution" usually covers spatial resolution (also denoted as "spatial super resolution" and "image super resolution"), while "angular super resolution" is the terminology that is in alignment with the super resolution discussed in this section. Such content resolution enhancement can be achieved via adaptive feature remixing [66], deformable convolution [67], deep combinatorial geometry embedding and structural consistency regularization [68], zero-shot learning [69], and many more. The recent work of Yu et al. reviews such techniques. Regarding angular resolution, techniques rely on geometry-aware networks [70], edge-aware inpainting [71], structure and scene information [72], Convolutional Neural Network (CNNs) [73,74], and several other methods. The work of Wang et al. [75] investigates spatial–angular interaction, similarly to the contribution of Cao et al. [76].

3.7. Outlook: Near-eye light field displays

As mentioned in Section 1, not every single light field display offers glasses-free 3D visualization. Yet as this is a not-so-well-known fact, this article provides a brief outlook regarding such devices. Near-eye light field displays are sometimes called "light field stereoscopes" [13], as they follow the same usage principle as the conventional stereoscopes (i.e., one display is assigned to one eye). The prototype of Lanman and Luebke [9] uses a pair of OLED panels that are covered with microlens arrays. The microlens arrays are thus situated between the eyes and the displays, and they create a virtual image perceptually behind the display. An important aspect of the implementation is that it enables the reallocation of driver electronics (e.g., can be waist-mounted). The proposed display of Zhao et al. [11] uses microlens arrays as well, but it also utilizes the structure parameters prioritization method, in order to increase the FOD. It is also possible to carry out pupil tracking [77], which is therefore analogous to some of the more conventional devices, in the sense that a portion of the physiology of the user is tracked – from which, of course, the orientation of the pupils is the most important. The work of Zhan et al. [78] relies on Pancharatnam–Berry phase lenses, which are basically half-wave plate optical elements, the crystal-axis of which is changing spatially [79,80,80]. It is possible to use active driving (via the application of voltages to switch between lens profile patterns), as well as passive driving (via an external polarization rotator). The recent work of Gao et al. [81] highlights that such displays do not result visual fatigue, which is definitely a benefit in contrast to other head-mounted 3D technologies. Still, at the end of the day, these solutions rely on head-mounted devices that are situated closely to the eyes of the individual, and thus, they disconnect the user from the real world. Additionally, they all share the disadvantages of other head-mounted technologies when it comes to use cases such as the metaverse.

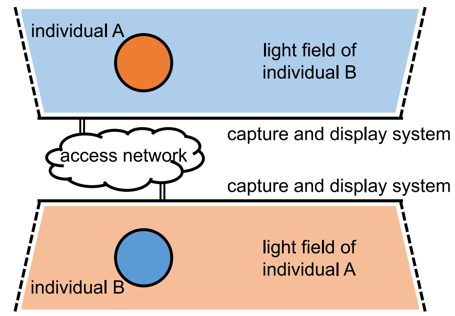

Fig. 5. Light field telepresence with 2 users.

4. Applications of light field metaverse

In this section, we study the potential applications of light field metaverse. The section aims to comprehensively cover every application, but note that the future may hold use case contexts that do not exist at the time of writing this paper – particularly enabled by the technology at hand.

4.1. Telepresence

Light field telepresence may be approached as an individual service within the metaverse, or an integral component of something more complex. The scientific community often looks at light field telepresence as the next great enabler of the sense of presence, which is also detailed in Section 5. The operational concept is analogous to a conventional laptop or smartphone videochat in the sense that a symmetric two-way telepresence solution should be equipped with both a capture system and a display system at the two ends. Examples for single-user and multi-user scenarios are provided in Figs. 5 and 6, respectively.

In today's phase of experimental and prototype solutions, the implemented telepresence systems are rather asymmetric. For example, Zhang et al. [16] introduced a levitating prototype, which is basically a quadcopter that provides a live 3D image of the individual's head.

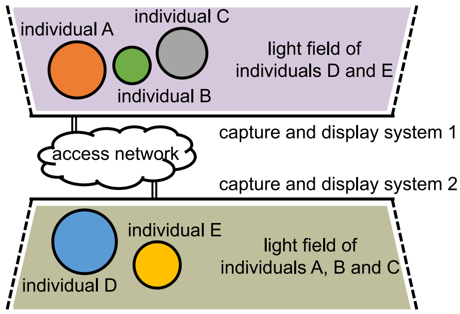

Fig. 6. Light field telepresence with 5 users.

Fig. 7. Light field telepresence system with a synthetic content [82].

A symmetric variation of this approach would mean that such a flying device is also equipped with a light field camera or a camera array to capture the presence of others. The wide-baseline solution of Cserkaszky et al. [82] technically looks like a human-sized mirror that may easily portray the entire body of an individual in a true-to-scale manner. Photographs of the prototype with a synthetic content and with actual live content are shown in Figs. 7 and 8, respectively.

Asymmetric light field telepresence may be implemented via general-purpose light field displays – and, of course, an appropriate camera system on the other end. Naturally, there are certain drawbacks of relying on existing commercial solutions. One issue is that such displays rarely replicate the true scale of the visualized individual by default. Yet this issue is not something that could not be handled on the level of visualization software (i.e., automatically zooming in or out on the captured light field, based on the dimensions of the display and the represented individual). The other notable attribute of general-purpose light field displays that does not align well with the requirements of telepresence is that they are typically landscape oriented, which means that it is not feasible to visualize an individual as shown in Figs. 7 and 8. However, it is actually adequate for waist-up representation, and multiple individuals may be visualized simultaneously. An implementation of such visualization is shown in Fig. 9.

4.2. General social activities

In the light field metaverse, there are two main approaches for representing an individual in the virtual space of social activities. One is to rely on light field telepresence, which means that the actual physical presence of the individual is captured and shared with other users of the metaverse. Another approach is to follow the concept of today's VR applications and represent the individual with a synthetic 3D avatar.

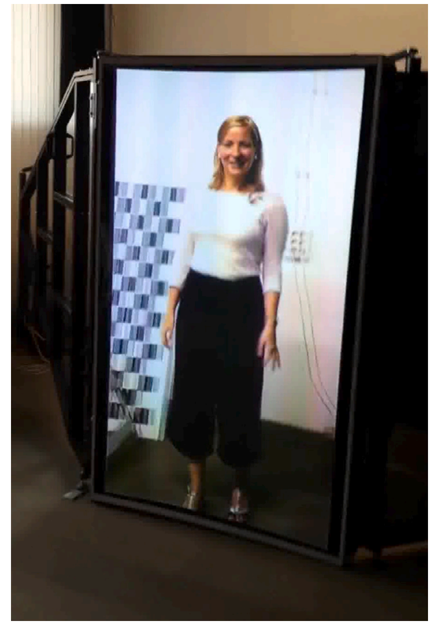

Fig. 8. Light field telepresence system during operation [82].

Fig. 9. Light field telepresence on the general-purpose HoloVizio 721RC [83].

In order to maintain the "gadget-free" nature of the light field metaverse, motion and posture should be tracked by either a dedicated tracking system or cameras. By choosing the latter, the implementation shall enable the option to seamlessly switch between a synthetic avatar and the genuine physical appearance. Of course, it is expected that the majority of users shall have a strong preference towards synthetic avatars in order to protect their privacy while engaging in the virtual world. In fact, it is quite likely that many applications would restrict representation to avatars, as the sense of anonymity may greatly contribute to certain social activities in the metaverse.

With today's motion-tracking technologies and general-purpose displays, it is quite feasible to implement avatar-based social activities for the light field metaverse. It is important to add that in the case of 3D avatars, the true-to-scale, portrait-oriented visualization of the users may be less important.

4.3. Education and cultural heritage

While many view the metaverse to be "all fun and games", the truth is that it holds immense potential for education purposes and for sharing a digital/digitized form of cultural heritage. Already in

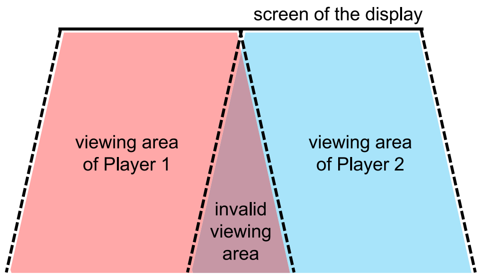

Fig. 10. Overview of split-domain light field gaming.

the current implementations of the metaverse, dedicated users invest tremendous efforts into "importing" cultural heritage into the virtual worlds. Additionally, genuine artworks are also created in these virtual environments.

Light field technology has already been explored in the context of cultural heritage. For example, the EU H2020 i-MareCulture project[5] utilized the HoloVizio C80 system (mentioned earlier for being a front-projection light field display) to visualize artifacts of underwater cultural heritage at an exhibition of the Thalassa museum in 2019.

As the example shows, conventional light field displays are appropriate for the visualization of cultural heritage. The same is expected in the context of education and training. Of course, certain instances of specialized training require dedicated displays; however, as authors of this manuscript, we do not consider specialized training to be a purpose of the metaverse. Yet education is indeed an important use of the metaverse, as virtual-community-based skill sharing may greatly benefit society.

An example for skill sharing via light field metaverse could be regarding a household activity. In such a scenario, the receiver of the online assistance could simultaneously perceive the activity that is to be mimicked in glasses-free 3D, and his or her own actions in the real world.

4.4. Gaming

There are two major aspects regarding gaming in the light field metaverse that need to be highlighted. First of all, it has already been mentioned that light field technology suffers the constraint of not being able to visualize virtually-infinite distances. However, certain game genres do not do that at all, and they are perfectly appropriate for light field visualization. Namely, real-time and turn-based strategy games tend to look towards the virtual ground from an elevated position. This camera operation makes the presentation of the game view evidently finite, as the viewed game space fits into a well-defined bounding volume. Furthermore, this is applicable to any type of isometric game genre, and it can also be extended to any game that does not necessitate the visualization of virtually-infinite distances.

A quite unique property of light field displays is that the visualized angular domain may be split. In the case of 2D split-screen gaming, a portion of the screen is allocated to one player, while another portion is allocated to the other player. In practice, during a two-player scenario, the screen is split in half either vertically or horizontally, which means that none of the players may utilize the entire screen. For four players, a player only uses a fourth of the screen. This is also rather inconvenient in the sense that it negatively affects competitive gaming, due to the potential "screen peeking", which means that a player may gain an unfair advantage by seeing the screen of the other player. While such a division of the screen is possible for light field displays as well, it is better to split the screen in the angular domain. A two-player example is shown in Fig. 10. Technically, one portion of the FOV is allocated to one player, and another portion is allocated to the other player. This means that both players may use the entire screen for gaming, yet they see two completely different views. They may also move within the viewing area. Of course, the middle of the FOV is an invalid viewing area in such a case, as a mixture of the two light fields is projected there.

From a financially-realistic point of view, it is expected that light field gaming shall emerge as arcade-like video games, before the dawn of affordable personal computers. However, it is debatable whether dedicated or general-purpose light field displays shall be used for such purpose. The advantage of using general-purpose displays is that it may house a great multitude of games, increasing flexibility and deployment efficiency. On the other hand, dedicated systems may be carefully tailored to satisfy every possible requirement towards the specific game.

4.5. Online dating

Online dating in the light field metaverse – which is practically an instance of the telepresence – should be separately addressed. It may benefit a lot from the true-to-scale realistic visualization of individuals. It is quite common in today's online dating applications that the appearance of the individual is reduced to a single 2D image, which, in many cases, may be misleading (e.g., intentionally, via photo editing). This typically causes disappointment during the first physical meeting, not to mention all the instances of the so-called "catfishing" practices.

In the case of light field visualization, individuals may be accurately portrayed in a glasses-free 3D manner. Whether it is a form of live telepresence or stored visual data, the appearance of the individual is not distorted by the constraints of 2D capture. It is feasible for both dedicated and general-purpose systems, yet the same considerations apply as in the case of telepresence.

4.6. Other applications and activities

Similarly to the state-of-the-art conventional metaverse, applications are only limited by imagination – and, of course, by technological constraints. Members of the digital society may do whatever they wish to do in the virtual world. For example, one of the candidates of the 2020 US election established an island in the life-simulation video game Animal Crossing.[6] In fact, it is possible to carry out voting in a virtual world that may have consequences in the real world.

One concern of the applications of light field metaverse may be privacy. This is already a hot topic in the conventional metaverse [84–86]. In addition, in the case of the applications of light field metaverse, anyone within the allocated FOV may perceive the content perspective of an individual. On the other hand, in the case of VR, the surrounding individuals in the real world have no information what a given person sees through the viewing device.

5. QoE of light field metaverse

In this section, we discuss the different aspects and considerations of QoE. The section also reviews the relevant findings of the scientific community.

[5] https://imareculture.eu/

[6] https://edition.cnn.com/2020/10/18/business/biden-animal-crossing-island-trnd/index.html

5.1. Immersion

Immersion is one of the most important QoE aspects of the metaverse, yet at the time of writing this paper, no subjective study has addressed it so far as a research question. One way to do that is conventional subjective ratings. On their own, such subjective results may not be completely useful, as the quantification of immersion may be subjective as well. Instead, comparisons of immersive technologies could report better the immersion achieved by light field metaverse. Another method is to involve physiological measurements, as used by the scientific literature. A great example for this is used in the cognitive experiment of Bouchard et al. [87]. It was shown in the context of VR that the sole idea of live presence induced differences in the measurements obtained through functional magnetic resonance imaging (fMRI). Such experiment should be carried out for light field visualization as well.

Evidently, there is a major consideration that a light field metaverse is not as immersive as a VR metaverse. Again, as of now, no data supports one claim or the other.

5.2. Sense of presence

The sense of presence may be classified into two distinct subsets. The first one is the sense of oneself. In the case of light field metaverse, one may naturally see oneself in exactly the same way as in everyday life. In VR, this is more complicated, particularly if a synthetic avatar is used. Although keep in mind that for light field metaverse, the individual perceives a different self-reality for avatar-based representation. Basically, the other participants of the virtual world see a synthetic avatar, while the user sees factual reality.

The other one is the sense of others. This is closely connected to the previously elaborated subset. If there are multiple users on one side of a telepresence system – as portrayed by Fig. 6 – the individual may perceive each other in the real world. On the other hand, if two users are connected via VR, even if they are next to each other in the same room, they can only see each other's digital representation.

5.3. 3D visual experience

For light field visualization, the 3D visual experience primarily depends on ray density. As explained earlier, the 3D experience requires two distinct rays to reach the two eyes, with respect to a single point on the screen. Based on the trigonometric implications of this requirement, the formalized calculation for viewing distance was tested in a subjective study on binocular disparity [88]. The results indicate that the theoretical distance – at which the two eyes are barely addressed by two distinct light rays – provides a visual experience that is halfway between 2D and 3D. For light field metaverse, it is imperative to guarantee 3D perception. Of course, the ultimate goal is to guarantee super resolution, as it is the key to the proper perception of depth.

5.4. The parallax effect

A smoothness of the parallax effect is essential to the overall visual experience, and thus, to immersion. Similarly to the 3D experience, it depends on the angular resolution of visualization. While the scientific literature indicates that artifact mitigation is possible [89], disturbances of the parallax effect should be completely avoided, as they may greatly penalize the QoE of the light field metaverse. The most immersion-breaking issue is the crosstalk effect, which fundamentally depends on the depth and the structure of the visualized content.

An example is shown in Fig. 11. The left and the right side of the figure show the same 3D object with two different ray density values. The one on the left has an angular resolution of 1 degree. Due to the great depth and the structural complexity of the visualized dodecahedron grid, both the closest and the farthest portion of the mathematical body suffers degradation. This is additional made worse by the fact that one degraded region can be seen through the other. However, the content can still be easily recognized. If we look at the right side of the figure, this is not necessarily the case. The angular resolution of 1.5 degrees already imposes severe crosstalk on the model. In the light field metaverse, such degradation not only breaks immersion, but may also perceptually exhaust and irritate the user. For long-term metaverse usage, future research should address such effect of parallax issues.

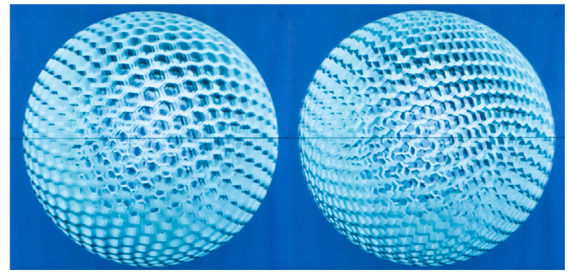

Fig. 11. Demonstration of angular resolution reduction [89].

5.5. Interaction and user interface

As detailed in Section 4, applications of light field metaverse may necessitate interaction. 3D interaction on light field displays was explored by Adhikarla et al. [90]. The authors found that using a light field interface may have a lower performance than conventional controls. In the case of various metaverse applications, conventional controls may remain dominant (e.g., in competitive gaming). Yet it should be noted that applications in which task performance is not essential may benefit from the ease of intuitive 3D interfaces. Furthermore, multi-user interaction may be necessary as well, due to the circumstances of the potential applications (e.g., multiple players using a single system).

5.6. Perceptual fatigue

For the long-term usage of the applications of light field metaverse, visualization must not – or at least, should not – exhaust the user. This is particularly important, since certain applications may be used over extended periods of time. This was already mentioned in the case of parallax degradation. However, this QoE consideration needs to be separately highlighted. Every single global statistics on personal digital device usage (i.e., computers, tablets and smartphones) show that people spend more and more time watching screens. Likewise, especially the younger generations, spend more and more time on social media. Therefore, it is expected that such devices should be usable for extended periods of time without perceptual issues and fatigue. Basically, the success of the solutions depends on it. Due to the natural, glasses-free 3D nature of light field visualization, it is expected that long-term usage shall be less exhausting than the conventional head-mounted technologies of the metaverse. Yet these are only expectations, and they need to be supported by scientific results. Measuring the perceptual fatigue of the state-of-the-art light field displays shall be one of the most important future research questions of the scientific community. If the data provides positive indications and light field displays become an affordable everyday technology, then the way shall be paved towards a world of light field metaverse – which, hopefully, shall not be some sort of a dark digital dystopia.

5.7. Audio

While this paper primarily focuses on visualization quality, it needs to be noted that audio quality shall contribute to the overall experience as well. Of course, this is applicable to only those use cases where audio is actually present – and meaningful (i.e., more than a generic background music to accompany the 3D visualization). For example, the exhibition of cultural heritage is not inherently dependent on the inclusion of audio content. However, a dominant portion of light field metaverse applications either necessitate audio or may greatly benefit from it, and thus, it is relevant to be discussed.

First of all, audio quality and visual quality are heavily intertwined, hence multimodal approaches must be considered when assessing quality [91], as well as audio–visual correspondence [92]. Audio cues also affect gaze, the tracking of which is crucial to understanding the connection between QoE and visual attention [93]. Probably the greatest audio-related contribution to the successful implementation of light field use cases in general shall be delivered by 3D sound – which is, quite fittingly, often labeled as "sound field". As light fields are enabled by projector arrays, sound field are enabled by speaker arrays. For instance, the prototype of Shin et al. [94], the design of Gupta and Abhayapala [95], and the ESPRO 2.0 of Noisternig et al. [96] are composed of 16, 18 and 350 loudspeakers, respectively. It is expected that such solutions – especially the more complex ones – will be mostly limited to dedicated visualization systems. Naturally, general-purpose light field displays can be empowered by sound field; however, the first generations of consumer-grade models are not expected to be equipped with such audio by default.

6. Conclusion

In this paper, the potential applications of light field metaverse were discussed, along with the technological advantages, disadvantages and constraints, and the consideration of QoE. Light field visualization allows the user to be "connected without disconnection"; the viewing-device-free nature of the technology enables the awareness of the physical environment, including other users of the same system. It also creates novel opportunities for certain applications, such as split-domain gaming. Although there is already a rich scientific literature of light field QoE, numerous vital research questions are still unanswered [42]. The most important of these topics include – but are not limited to – immersion, parallax degradation, and perceptual fatigue.

CRediT authorship contribution statement

Peter A. Kara: Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Roopak R. Tamboli:** Conceptualization, Methodology, Investigation, Writing – original draft. **Vamsi K. Adhikarla:** Conceptualization, Methodology, Writing – original draft, Funding acquisition. **Tibor Balogh:** Conceptualization, Methodology, Validation, Resources, Visualization, Supervision, Funding acquisition. **Mary Guindy:** Conceptualization, Resources, Funding acquisition. **Aniko Simon:** Conceptualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

The research reported in this paper is part of project no. BME-NVA-02, implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021 funding scheme. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813170. Also received funding by 2018-2.1.3-EUREKA-2018-00007 and KFI 16-1-2017-0015, NRDI Fund, Hungary.

References

[1] N. Stephenson, Snow Crash, Bantam Books, 1992.
[2] C.-J. Chao, Y.-J. Yau, C.-H. Lin, W.-Y. Feng, Effects of display technologies on operation performances and visual fatigue, Displays 57 (2019) 34–46.
[3] G. Geršak, H. Lu, J. Guna, Effect of VR technology matureness on VR sickness, Multimedia Tools Appl. 79 (21) (2020) 14491–14507.
[4] J. Iskander, M. Hossny, Measuring the likelihood of VR visual fatigue through ocular biomechanics, Displays 70 (2021) 102105.
[5] J. Hurter, C. Maraj, S. Murphy, Commercial virtual reality displays: Issues of performance and simulator sickness from exocentric depth-perception tasks, Displays 70 (2021) 102095.
[6] D. Risi, S. Palmisano, Effects of postural stability, active control, exposure duration and repeated exposures on HMD induced cybersickness, Displays 60 (2019) 9–17.
[7] M. Kaufeld, M. Mundt, S. Forst, H. Hecht, Optical see-through augmented reality can induce severe motion sickness, Displays 74 (2022) 102283.
[8] M.G. Lippmann, La photographies integrals, C. R. 146 (9) (1908) 446–451.
[9] D. Lanman, D. Luebke, Near-eye light field displays, ACM Trans. Graph. 32 (6) (2013) 1–10.
[10] L. Shi, F.-C. Huang, W. Lopes, W. Matusik, D. Luebke, Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3D computer graphics, ACM Trans. Graph. 36 (6) (2017) 1–17.
[11] J. Zhao, Q. Ma, J. Xia, J. Wu, B. Du, H. Zhang, Hybrid computational near-eye light field display, IEEE Photonics J. 11 (1) (2019) 1–10.
[12] M. Liu, C. Lu, H. Li, X. Liu, Near eye light field display based on human visual features, Opt. Express 25 (9) (2017) 9886–9900.
[13] F.-C. Huang, D.P. Luebke, G. Wetzstein, The light field stereoscope, in: SIGGRAPH Emerging Technologies, 2015, pp. 1–12.
[14] D. Teng, C. Lai, Q. Song, X. Yang, L. Liu, Super multi-view near-eye virtual reality with directional backlights from wave-guides, Opt. Express 31 (2) (2023) 1721–1736.
[15] T. Balogh, A. Barsi, P.A. Kara, M. Guindy, A. Simon, Z. Nagy, 3D light field LED wall, in: Digital Optical Technologies 2021, Vol. 11788, SPIE, 2021, pp. 1–11.
[16] X. Zhang, S. Braley, C. Rubens, T. Merritt, R. Vertegaal, LightBee: A self-levitating light field display for hologrammatic telepresence, in: CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–10.
[17] D. Fattal, Lightfield displays: A window into the metaverse, in: SPIE AR, VR, MR Industry Talks 2022, Vol. 11932, SPIE, 2022.
[18] G. Zheng, L. Yuan, A review of QoE research progress in metaverse, Displays (2023) 102389.
[19] C. Cascio, F. McGlone, S. Folger, V. Tannan, G. Baranek, K.A. Pelphrey, G. Essick, Tactile perception in adults with autism: A multidimensional psychophysical study, J. Autism Dev. Disord. 38 (1) (2008) 127–137.
[20] M. Mikkelsen, E.L. Wodka, S.H. Mostofsky, N.A. Puts, Autism spectrum disorder in the scope of tactile processing, Dev. Cogn. Neurosci. 29 (2018) 140–150.
[21] M. Panagiotidi, P.G. Overton, T. Stafford, The relationship between ADHD traits and sensory sensitivity in the general population, Compr. Psychiatry 80 (2018) 179–185.
[22] J.L. He, E. Wodka, M. Tommerdahl, R.A. Edden, M. Mikkelsen, S.H. Mostofsky, N.A. Puts, Disorder-specific alterations of tactile sensitivity in neurodevelopmental disorders, Commun. Biol. 4 (1) (2021) 1–15.
[23] S.E. Brooks, D. Johnson, N. Fischer, Anisometropia and binocularity, Ophthalmology 103 (7) (1996) 1139–1143.
[24] S. Tomaç, E. Birdal, Effects of anisometropia on binocularity, J. Pediatr. Ophthalmol. Strabismus. 38 (1) (2001) 27–33.
[25] S.A. Read, M.J. Collins, L.G. Carney, A review of astigmatism and its possible genesis, Clin. Exp. Optometry 90 (1) (2007) 5–19.
[26] The fascinating history and evolution of extended reality (XR) – Covering AR, VR and MR, 2021, https://www.forbes.com/sites/bernardmarr/2021/05/17/the-fascinating-history-and-evolution-of-extended-reality-xr--covering-ar-vr-and-mr/?sh=1beb5ed84bfd. (Accessed: September 2022).
[27] History of VR – Timeline of events and tech development, 2022, https://virtualspeech.com/blog/history-of-vr. (Accessed: September 2022).

[28] S. Mann, T. Furness, Y. Yuan, J. Iorio, Z. Wang, All reality: Virtual, augmented, mixed (x), mediated (x, y), and multimediated reality, 2018, arXiv preprint arXiv:1804.08386.

[29] J.-P. Stauffert, F. Niebling, M.E. Latoschik, Latency and cybersickness: Impact, causes, and measures. A review, Front. Virtual Real. 1 (2020) 582204.

[30] K. Brunnström, E. Dima, T. Qureshi, M. Johanson, M. Andersson, M. Sjöström, Latency impact on quality of experience in a virtual reality simulator for remote control of machines, Signal Process., Image Commun. 89 (2020) 116005.

[31] P. Milgram, F. Kishino, A taxonomy of mixed reality visual displays, IEICE Trans. Inf. Syst. 77 (12) (1994) 1321–1329.

[32] P.A. Rauschnabel, R. Felix, C. Hinsch, H. Shahab, F. Alt, What is XR? Towards a framework for augmented and virtual reality, Comput. Hum. Behav. 133 (2022).

[33] L.-H. Lee, T. Braud, P. Zhou, L. Wang, D. Xu, Z. Lin, A. Kumar, C. Bermejo, P. Hui, All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda, 2021, arXiv preprint arXiv:2110.05352.

[34] Understanding display techniques in augmented reality, 2018, https://blog.prototypr.io/understanding-display-techniques-in-augmented-reality-c258b911b5c9. (Accessed: September 2022).

[35] J.P. Rolland, R.L. Holloway, H. Fuchs, Comparison of optical and video see-through, head-mounted displays, in: Telemanipulator and Telepresence Technologies, Vol. 2351, SPIE, 1995, pp. 293–307.

[36] M. Levoy, P. Hanrahan, Light field rendering, in: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 1996, pp. 31–42.

[37] A. Cserkaszky, P.A. Kara, R.R. Tamboli, A. Barsi, M.G. Martini, L. Bokor, T. Balogh, Angularly continuous light-field format: Concept, implementation, and evaluation, J. Soc. Inf. Disp. 27 (7) (2019) 442–461.

[38] P.A. Kara, R.R. Tamboli, O. Doronin, A. Cserkaszky, A. Barsi, Z. Nagy, M.G. Martini, A. Simon, The key performance indicators of projection-based light field visualization, J. Inf. Display (2019) 81–93.

[39] T. Balogh, The HoloVizio system, in: Stereoscopic Displays and Virtual Reality Systems XIII, Vol. 6055, SPIE, 2006, pp. 279–290.

[40] Society for Information Display (SID), Information Display Measurements Standard (IDMS), 2012, https://www.sid.org/Standards/ICDM. (Accessed: January 2023).

[41] Y. Zhu, G. Zhai, X. Min, The prediction of head and eye movement for 360 degree images, Signal Process., Image Commun. 69 (2018) 15–25.

[42] P.A. Kara, R.R. Tamboli, E. Shafiee, M.G. Martini, A. Simon, M. Guindy, Beyond perceptual thresholds and personal preference: Towards novel research questions and methodologies of quality of experience studies on light field visualization, Electronics 11 (6) (2022).

[43] A. Simon, P.A. Kara, M. Guindy, X. Qiu, L. Szy, T. Balogh, One step closer to a better experience: Analysis of the suitable viewing distance ranges of light field visualization usage contexts for observers with reduced visual capabilities, in: Novel Optical Systems, Methods, and Applications XXV, Vol. 12216, SPIE, 2022, pp. 133–143.

[44] M. Guindy, A. Barsi, P.A. Kara, V.K. Adhikarla, T. Balogh, A. Simon, Camera animation for immersive light field imaging, Electronics 11 (17) (2022) 2689.

[45] P.A. Kara, M. Guindy, Q. Xinyu, V.A. Szakal, T. Balogh, A. Simon, The effect of angular resolution and 3D rendering on the perceived quality of the industrial use cases of light field visualization, in: 16th International Conference on Signal Image Technology & Internet Based Systems, SITIS, 2022.

[46] A. Ak, P. Le Callet, Quality evaluation of light fields, in: Immersive Video Technologies, Academic Press, 2022, p. 265.

[47] IEEE 3333.1.4: Recommended practice for the quality assessment of light field imaging, 2022, https://standards.ieee.org/ieee/3333.1.4/10873/. (Accessed: January 2023).

[48] X. Min, K. Gu, G. Zhai, X. Yang, W. Zhang, P. Le Callet, C.W. Chen, Screen content quality assessment: Overview, benchmark, and beyond, ACM Comput. Surv. 54 (9) (2021) 1–36.

[49] T. Ebrahimi, S. Foessel, F. Pereira, P. Schelkens, JPEG Pleno: Toward an efficient representation of visual reality, IEEE Multimedia 23 (4) (2016) 14–20.

[50] G. Zhai, X. Min, Perceptual image quality assessment: A survey, Sci. China Inf. Sci. 63 (2020) 1–52.

[51] X. Min, J. Zhou, G. Zhai, P. Le Callet, X. Yang, X. Guan, A metric for light field reconstruction, compression, and display quality evaluation, IEEE Trans. Image Process. 29 (2020) 3790–3804.

[52] P. Paudyal, F. Battisti, M. Carli, Reduced reference quality assessment of light field images, IEEE Trans. Broadcast. 65 (1) (2019) 152–165.

[53] C. Meng, P. An, X. Huang, C. Yang, D. Liu, Full reference light field image quality evaluation based on angular-spatial characteristic, IEEE Signal Process. Lett. 27 (2020) 525–529.

[54] L. Shi, W. Zhou, Z. Chen, J. Zhang, No-reference light field image quality assessment based on spatial-angular measurement, IEEE Trans. Circuits Syst. Video Technol. 30 (11) (2019) 4114–4128.

[55] W. Zhou, L. Shi, Z. Chen, J. Zhang, Tensor oriented no-reference light field image quality assessment, IEEE Trans. Image Process. 29 (2020) 4070–4084.

[56] L. Shan, P. An, C. Meng, X. Huang, C. Yang, L. Shen, A no-reference image quality assessment metric by multiple characteristics of light field images, IEEE Access 7 (2019) 127217–127229.

[57] Z. Luo, W. Zhou, L. Shi, Z. Chen, No-reference light field image quality assessment based on micro-lens image, in: 2019 Picture Coding Symposium, PCS, IEEE, 2019, pp. 1–5.

[58] Z. Guo, W. Gao, H. Wang, J. Wang, S. Fan, No-reference deep quality assessment of compressed light field images, in: 2021 IEEE International Conference on Multimedia and Expo, ICME, IEEE, 2021, pp. 1–6.

[59] J. Xiang, G. Jiang, M. Yu, Y. Bai, Z. Zhu, No-reference light field image quality assessment based on depth, structural and angular information, Signal Process. 184 (2021) 108063.

[60] A. Ak, S. Ling, P. Le Callet, No-reference quality evaluation of light field content based on structural representation of the epipolar plane image, in: 2020 IEEE International Conference on Multimedia & Expo Workshops, ICMEW, IEEE, 2020, pp. 1–6.

[61] X. Min, K. Gu, G. Zhai, J. Liu, X. Yang, C.W. Chen, Blind quality assessment based on pseudo-reference image, IEEE Trans. Multimed. 20 (8) (2017) 2049–2062.

[62] X. Min, G. Zhai, K. Gu, Y. Liu, X. Yang, Blind image quality estimation via distortion aggravation, IEEE Trans. Broadcast. 64 (2) (2018) 508–517.

[63] B. Appina, S.V.R. Dendi, K. Manasa, S.S. Channappayya, A.C. Bovik, Study of subjective quality and objective blind quality prediction of stereoscopic videos, IEEE Trans. Image Process. 28 (10) (2019) 5027–5040.

[64] W. Sun, X. Min, G. Zhai, K. Gu, H. Duan, S. Ma, MC360IQA: A multi-channel CNN for blind 360-degree image quality assessment, IEEE J. Sel. Top. Sign. Proces. 14 (1) (2019) 64–77.

[65] A.K.R. Poreddy, P.A. Kara, R.R. Tamboli, A. Simon, B. Appina, CoDIQE3D: A completely blind, no-reference stereoscopic image quality estimator using joint color and depth statistics, Vis. Comput. (2023) 1–11.

[66] K. Ko, Y.J. Koh, S. Chang, C.-S. Kim, Light field super-resolution via adaptive feature remixing, IEEE Trans. Image Process. 30 (2021) 4114–4128.

[67] Y. Wang, J. Yang, L. Wang, X. Ying, T. Wu, W. An, Y. Guo, Light field image super-resolution using deformable convolution, IEEE Trans. Image Process. 30 (2020) 1057–1071.

[68] J. Jin, J. Hou, J. Chen, S. Kwong, Light field spatial super-resolution via deep combinatorial geometry embedding and structural consistency regularization, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2260–2269.

[69] Z. Cheng, Z. Xiong, C. Chen, D. Liu, Z.-J. Zha, Light field super-resolution with zero-shot learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 10010–10019.

[70] J. Jin, J. Hou, H. Yuan, S. Kwong, Learning light field angular super-resolution via a geometry-aware network, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, no. 07, 2020, pp. 11141–11148.

[71] X. Liu, M. Wang, A. Wang, X. Hua, S. Liu, Depth-guided learning light field angular super-resolution with edge-aware inpainting, Vis. Comput. 38 (8) (2022) 2839–2851.

[72] J. Yang, L. Wang, L. Ren, Y. Cao, Y. Cao, Light field angular super-resolution based on structure and scene information, Appl. Intell. (2022) 1–17.

[73] D.-M. Kim, H.-S. Kang, J.-E. Hong, J.-W. Suh, Light field angular super-resolution using convolutional neural network with residual network, in: 2019 Eleventh International Conference on Ubiquitous and Future Networks, ICUFN, IEEE, 2019, pp. 595–597.

[74] G. Liu, H. Yue, J. Yang, A coarse-to-fine convolutional neural network for light field angular super-resolution, in: Artificial Intelligence: Second CAAI International Conference, CICAI 2022, Beijing, China, August 27–28, 2022, Revised Selected Papers, Part I, Springer, 2022, pp. 268–279.

[75] Y. Wang, L. Wang, J. Yang, W. An, J. Yu, Y. Guo, Spatial-angular interaction for light field image super-resolution, in: European Conference on Computer Vision, Springer, 2020, pp. 290–308.

[76] Y. Cao, L. Wang, L. Ren, J. Yang, Y. Cao, View position prior-supervised light field angular super-resolution network with asymmetric feature extraction and spatial-angular interaction, Neurocomputing 518 (2023) 206–218.

[77] C. Jang, K. Bang, S. Moon, J. Kim, S. Lee, B. Lee, Retinal 3D: Augmented reality near-eye display via pupil-tracked light field projection on retina, ACM Trans. Graph. 36 (6) (2017) 1–13.

[78] T. Zhan, Y.-H. Lee, S.-T. Wu, High-resolution additive light field near-eye display by switchable Pancharatnam–Berry phase lenses, Opt. Express 26 (4) (2018) 4863–4872.

[79] E. Hasman, V. Kleiner, G. Biener, A. Niv, Polarization dependent focusing lens by use of quantized Pancharatnam–Berry phase diffractive optics, Appl. Phys. Lett. 82 (3) (2003) 328–330.

[80] N.V. Tabiryan, S.V. Serak, D.E. Roberts, D.M. Steeves, B.R. Kimball, Thin waveplate lenses: New generation in optics, in: Liquid Crystals XIX, Vol. 9565, SPIE, 2015, pp. 116–124.

[81] J. Gao, S. Ma, T. Hong, P. Liang, 70-2: A design for near eye light field display, in: SID Symposium Digest of Technical Papers, Vol. 53, no. 1, Wiley Online Library, 2022, pp. 941–944.

[82] A. Cserkaszky, A. Barsi, Z. Nagy, G. Puhr, T. Balogh, P.A. Kara, Real-time light-field 3D telepresence, in: 2018 7th European Workshop on Visual Information Processing, EUVIP, IEEE, 2018, pp. 1–5.

[83] P.A. Kara, A. Cserkaszky, M.G. Martini, A. Barsi, L. Bokor, T. Balogh, Evaluation of the concept of dynamic adaptive streaming of light field video, IEEE Trans. Broadcast. 64 (2) (2018) 407–421.

[84] Y. Wang, Z. Su, N. Zhang, R. Xing, D. Liu, T.H. Luan, X. Shen, A survey on metaverse: Fundamentals, security, and privacy, IEEE Commun. Surv. Tutor. (2022).
[85] B. Falchuk, S. Loeb, R. Neff, The social metaverse: Battle for privacy, IEEE Technol. Soc. Mag. 37 (2) (2018) 52–61.
[86] R. Di Pietro, S. Cresci, Metaverse: Security and privacy issues, in: 2021 Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications, TPS-ISA, IEEE, 2021, pp. 281–288.
[87] S. Bouchard, S. Dumoulin, J. Talbot, A.-A. Ledoux, J. Phillips, J. Monthuy-Blanc, G. Labonté-Chartrand, G. Robillard, M. Cantamesse, P. Renaud, Manipulating subjective realism and its impact on presence: Preliminary results on feasibility and neuroanatomical correlates, Interact. Comput. 24 (4) (2012) 227–236.
[88] P.A. Kara, R.R. Tamboli, A. Cserkaszky, A. Barsi, A. Simon, A. Kusz, L. Bokor, M.G. Martini, Objective and subjective assessment of binocular disparity for projection-based light field displays, in: 2019 International Conference on 3D Immersion, IC3D, IEEE, 2019, pp. 1–8.
[89] P.A. Kara, A. Cserkaszky, A. Barsi, T. Papp, M.G. Martini, L. Bokor, The interdependence of spatial and angular resolution in the quality of experience of light field visualization, in: 2017 International Conference on 3D Immersion, IC3D, IEEE, 2017, pp. 1–8.
[90] V.K. Adhikarla, J. Sodnik, P. Szolgay, G. Jakus, Exploring direct 3D interaction for full horizontal parallax light field displays using leap motion controller, Sensors 15 (4) (2015) 8642–8663.
[91] X. Min, G. Zhai, J. Zhou, M.C. Farias, A.C. Bovik, Study of subjective and objective quality assessment of audio-visual signals, IEEE Trans. Image Process. 29 (2020) 6054–6068.
[92] X. Min, G. Zhai, J. Zhou, X.-P. Zhang, X. Yang, X. Guan, A multimodal saliency model for videos with high audio-visual correspondence, IEEE Trans. Image Process. 29 (2020) 3805–3819.
[93] X. Min, G. Zhai, K. Gu, X. Yang, Fixation prediction through multimodal analysis, ACM Trans. Multimed. Comput. Commun. Appl. (TOMM) 13 (1) (2016) 1–23.
[94] M. Shin, F.M. Fazi, P.A. Nelson, F.C. Hirono, Controlled sound field with a dual layer loudspeaker array, J. Sound Vib. 333 (16) (2014) 3794–3817.
[95] A. Gupta, T.D. Abhayapala, Three-dimensional sound field reproduction using multiple circular loudspeaker arrays, IEEE Tran. Audio Speech Lang. Process. 19 (5) (2010) 1149–1159.
[96] M. Noisternig, T. Carpentier, O. Warusfel, ESPRO 2.0–Implementation of a surrounding 350-loudspeaker array for 3D sound field reproduction, in: Audio Engineering Society Conference: UK 25th Conference: Spatial Audio in Today's 3D World, Audio Engineering Society, 2012.