Received: 25 March 2024 | Revised: 8 May 2024 | Accepted: 19 May 2024

IET Image Processing

DOI: 10.1049/ipr2.13144

The Institution of Engineering and Technology WILEY

REVIEW

Light field imaging technology for virtual reality content creation: A review

Ali Khan[1] | **Md. Moinul Hossain**[1] | **Alexandra Covaci**[1] | **Konstantinos Sirlantzis**[2] | **Chuanlong Xu**[3]

[1]School of Engineering, University of Kent, Canterbury, Kent, UK

[2]Technology and Design, Canterbury Christ Church University, Canterbury, Kent, UK

[3]National Engineering Research Center of Turbo-Generator Vibration, School of Energy and Environment, Southeast University, Nanjing, China

Correspondence
Md. Moinul Hossain, School of Engineering, University of Kent, Canterbury, Kent, CT2 7NT, UK.
Email: m.hossain@kent.ac.uk

Funding information
European Regional Development Fund, Grant/Award Number: 2S05-038

Abstract

The light field (LF) imaging technique can capture 3D scene information in 4D by recording both 2D intensity and 2D direction of incoming light rays. Due to this capability, LF has shown a great interest in virtual reality (VR) and augmented reality (AR) for enhanced immersion, improved depth perception and reconstruction of realistic 3D environments. This paper presents a comprehensive review of LF imaging technology and other approaches used for VR content creation. The applications of LF technology beyond VR and AR are also discussed. The challenges and limitations of other approaches for VR content creation are examined. State-of-the-art research has focused on how VR experiences benefit from LF technology and identified the challenges to creating comfortable, immersive and realistic VR content such as (1) image size and resolution, (2) processing speed, (3) precise calibration and (4) depth reconstruction. Recommendations that can be considered for creating immersive VR content are provided to enhance user experience. These recommendations aim to contribute to developing more comfortable and realistic VR content, extending the potential applications of LF imaging technology in diverse fields.

1 | INTRODUCTION

Virtual reality (VR) is the simulation of reality that immerses the user in a virtual environment. The virtual environment can be created either based on computer-generated (CG) 3D scenes or photographically acquired content. The user experiences the sensation of being fully immersed in a 3D environment by using a VR headset even though it is not physically real. Nowadays the use of VR has become common in many applications such as healthcare, architecture, engineering and construction [1], entertainment, gaming, learning and training [2–4] due to its interactive, immersive nature and affordable consumer headsets. In addition, VR uses various visualisation techniques, including head-mounted displays (HMD), desktop screens, smartphones and VR caves [5] to deliver lifelike images, sounds and other experiences that create a fictitious or replicate a real-world scene. It also facilitates the user to mimic physical presence in the environment by allowing the user to interact with the space and any items depicted inside via specialised display screens.

Consumer headsets such as HTC VIVE and Oculus Quest have opened the market to tens of millions of customers. These headsets support a positional tracking system and enable users to experience VR more interactively where users can change the viewpoint displayed by the headset by moving their head accordingly and encounter a more immersive, natural and cosy experience. However, only CG content can be experienced in complete immersion in 3D video games and other applications such as in healthcare for educating doctors, training surgeons [6, 7] and controlling pain and anxiety [8], training industrial workers through VR content before they enter the actual workplace [9], and in education for teaching and learning [10]. Positional tracking data can be fed into CG scenes to produce motion parallax and view-dependent reflections. Besides,

List of Abbreviations: 3D, 3 dimensional; 3DOF, 3 degrees of freedom; 6DOF, 6 degrees of freedom; CG, Computer-generated; CNN, Convolutional neural networks; DCT, Discrete cosine transform; DPCA, Dual panoramic camera array; DWT, Discrete wavelet transform; EPI, Epipolar plane image; FOV, Field of view; GMM, Gaussian mixture model; GPU, Graphics processing unit; HMD, Head-mounted display; JEM, Joint exploration model; LF, Light field; LFC, Light field camera; LM, Layered meshes; MLA, Microlens array; MPI, Multi-plane image; MSI, Multi-sphere images; VR, Virtual reality.

photographic content generated by panoramic or 360° and omnidirectional stereo videos [11] can only be viewed from a single vantage point and lacks a sense of depth. Experiencing VR from one single point (with no motion parallax) can result in unpleasant side effects like headache, motion sickness and fatigue. Moreover, it is challenging for users to alter their viewpoints while experiencing VR content acquired photographically.

Recently, LF technology has demonstrated its utility across diverse industrial and entertainment applications such as industrial process monitoring [12, 13], photography and cinematography [14]. LF technology has also made significant contributions to immersive VR experiences [15, 16] and LF displays. In addition, this technology has gained attention in computer vision applications, owing to its capacity for acquiring abundant and precise visual information. This technology improves the accuracy of computer vision applications [17] including but not limited to 3D reconstruction [18], image segmentation [19], digital refocusing [20] and saliency detection [21]. Beyond these applications, LF technology is employed in medical imaging [22], as well as in the fields of autonomous vehicles and robotics [23].

However, studies show that this technology faces various challenges, depending on its applications. The common challenges encompass LF acquisition, compression, processing and 3D reconstruction. Various efforts have been devoted to tackling these challenges [14, 15, 24] such as introducing an advanced LF acquisition system and processing, and subsequently rendering the VR content [16, 25]. Various aspects of LF including shedding light related to LF representation, acquisition, depth estimation, image quality and reconstruction are discussed in [26–28]. Various review studies are also conducted focusing on the challenges associated with LF visualization and displays [29–34].

In VR, the LF technology shows a potential solution to display real-world scenes that enable users to not only rotate their heads around the scene but also able to move their heads within the VR environment due to its post-capturing refocusing capability. As a result, the user will have an immersive, comfortable and realistic VR experience and a better sense of the real-world 3D scene. Various review articles have been published addressing various aspects of LF in the VR application such as the development of LF displays [30, 33, 34], exploration of the LF evaluation, potential impact of LF technology and enhancement and transformation in VR experiences [16]. A review is also conducted on the commercial efforts that highlight the distinctive accessibility challenges and solutions for delivering more inclusive user engagement with VR [35].

This study mainly focuses on a comprehensive review of the recent developments of the LF technique, especially for VR content creation. At first, the traditional VR content generation techniques along with their methodologies, advantages and limitations are briefly discussed as shown in Figure 1. After that, the LF principle, its representation and acquisition are described. How the VR can greatly benefit from LF technology is also explored. Various challenges such as size, speed, processing time and resolution posed by LF techniques and their impact on the creation of realistic VR content are discussed. Recommendations for potential future research directions are also given at the end.

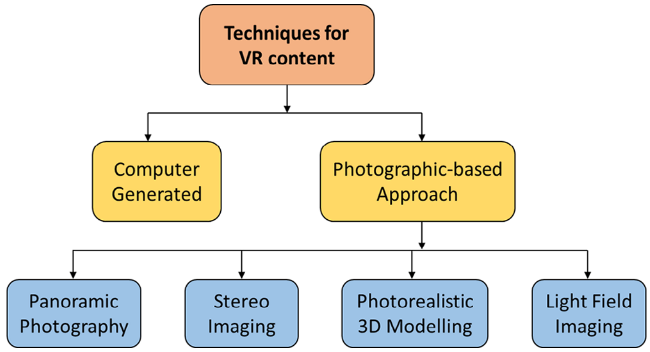

FIGURE 1 An overview of the techniques involved in VR content generation.

2 | TRADITIONAL TECHNIQUES FOR VR CONTENT CREATION

CG and photography are the common approaches to creating VR content. In CG, the entire virtual scene can be generated by gaming engines such as Unity3D [36, 37] and Unreal for different applications and 3D VR games. In the photographic-based approaches, the VR content is generated through a series of images of a scene captured by traditional cameras and the 3D reconstruction is then performed using reconstruction techniques [11, 38, 39]. This section gives an overview of the traditional techniques used for the VR content creation.

2.1 | Computer-generated approach

CG VR content refers to immersive digital experiences generated through a combination of computer graphics, 3D modelling and image rendering by utilizing computer software tools where real-world images or videos are not incorporated. The developer mainly uses 3D modelling software, animation tools and gaming engines. For example, as shown in Figure 2, the 3D scene is initially created by 3D modelling software and animation tools. Subsequently, a gaming engine is employed to simulate the constructed 3D scene along with objects and characters. The created VR content can be experienced on different platforms such as HMD, desktop platforms or mobile phones depending on the applications. Due to the interactivity of CG VR content and real-time rendering, this approach is applied in many applications. The interaction facility enables users to engage actively with the virtual environment. That also makes the user more interested in educational or training purposes and enables the delivery of training or education materials without the need for real-world exposure.

Various techniques are used to create CG VR content [36, 37, 40]. These techniques may vary depending on the applications. An example of VR content created for training and educational

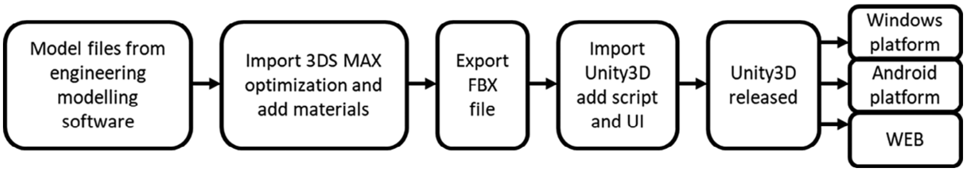

FIGURE 2 The development process of computer-generated VR content [9].

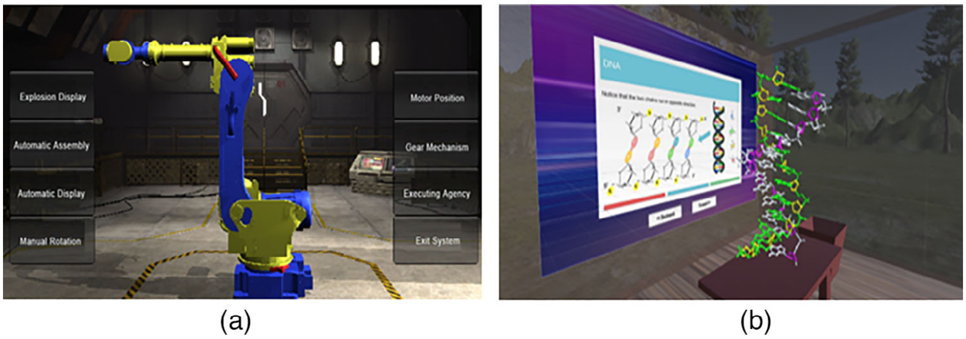

FIGURE 3 Examples of computer-generated VR content include (a) a VR training tutorial for the mechanical arm structure [9], and (b) an instructional tutorial for educational purposes on DNA molecules [10].

purposes is illustrated in Figure 3. The CG VR content enables the users to engage with simulated objects within a virtual environment, resulting in an immersive VR experience. However, most of the CG VR content is based on real-time rendering (a process that has the advantage of generating and displaying visuals in immediate response to the user's actions and movements). As the CG VR content is entirely based on a virtual environment and created based on graphics and 3D modelling software without the involvement of real-world images or videos, this approach can induce motion sickness.

2.2 | Photographic-based approach

In the photographic-based approach, traditional cameras are usually employed to capture images from the real world and then use the captured images to create VR content. This approach involves various techniques such as utilizing a single camera for capturing real-world images sequentially and then stitching them together to form 360° VR content. Also, more advanced implementation such as involving multiple cameras to facilitate the reconstruction of 3D VR content. Studies have also integrated range sensors alongside traditional cameras to capture depth information, contributing to the 3D reconstruction process. This section gives a detailed description of the photographic-based approaches involved in VR technology.

2.2.1 | Panoramic photography

The panoramic photographic-based approach employs traditional cameras to capture images or videos of a real-world scene and then reconstruct a view from that captured video. In VR, panorama or 360° video is one of the common approaches to depict real-world scenes. In this approach, the images or video is produced by rotating a single camera to capture a 360° view from a fixed point or multiple cameras to capture 360° images or videos of a scene [41]. The acquired data are aligned and stitched [42–44] together under various viewing angles to create panoramic VR content. The created panoramic content can then be translated into VR using gaming engines like Unity [45]. Figure 4 represents an example of the final output of the panoramic photography-based VR content. Rahim et al. [46] created VR content for teaching purposes to instruct students about milk powder production plants. The content is made based on 360° panoramic images of the industrial plant, including piping, instrumentation and supplemented text, accompanied by videos and animations. Tsai et al. [47] developed 360° VR content for soil and water conservation educational purposes, where the learners can view the VR content using Android phones and VR helmets.

As panoramic photography can only be viewed from one fixed position and rotate the head around the scene, it lacks motion parallax. Also, the panoramic content is captured from a single location, thus, no depth information can be achieved. Due to the absence of refocusing and motion parallax, users may experience motion sickness such as dizziness and nausea.

2.2.2 | Stereoscopic imaging

In VR, stereoscopic imaging is employed to visualize real-world scenes. This technique relies on the principle of binocular vision where a distinct view can be rendered for each eye and offers

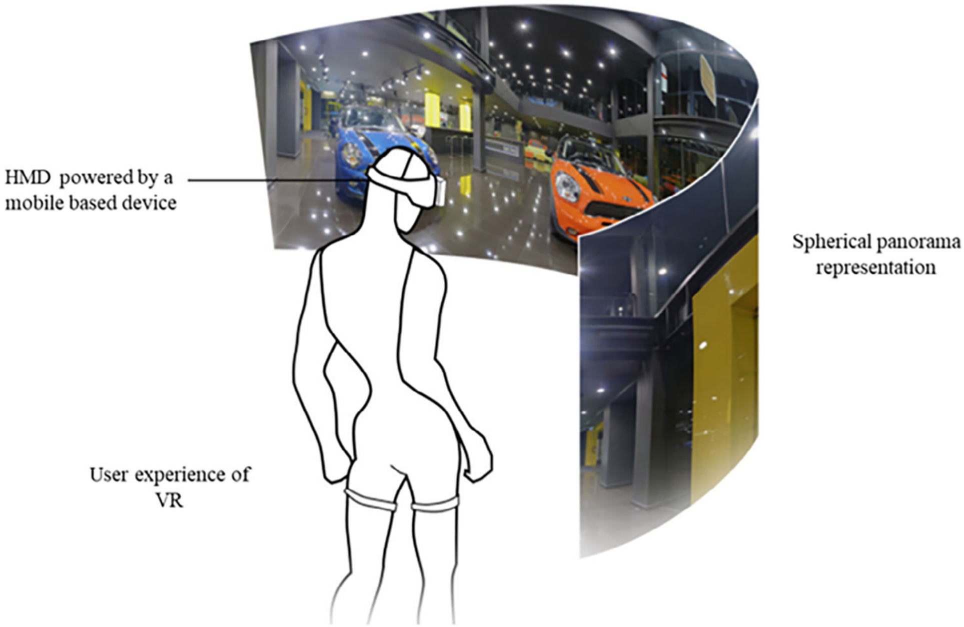

FIGURE 4 An example of a user experiencing spherical panoramic VR content using a mobile device-based head-mounted display [48].

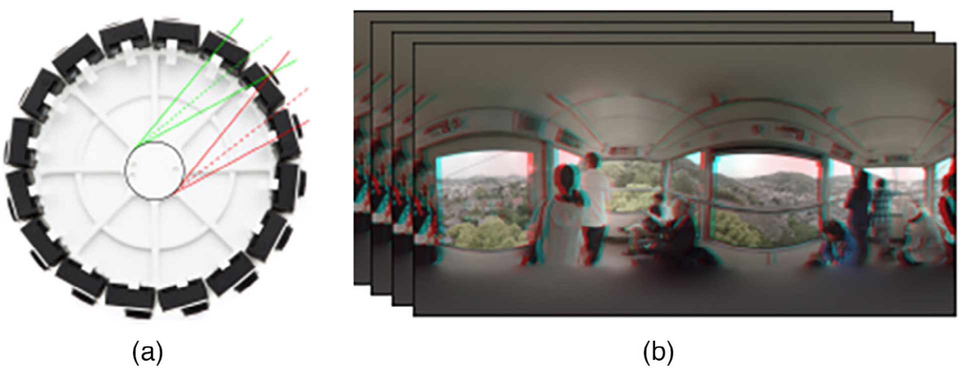

FIGURE 5 Stereoscopic panorama acquisition and rendering for VR, (a) camera rig to capture a stereo panorama, and (b) displaying the stereo panorama for VR using an anaglyphic stereo format [50].

a more immersive representation of a 360° scene and partial motion parallax compared to the single panorama. Various studies have been conducted such as Richardt et al. [49] presented a method for high-quality stereo panoramas by stabilizing and correcting input images of a static scene. An optical flow-based ray-up sampling method is used to seamlessly stitch images together. Anderson et al. [50] developed a bespoke camera system (as shown in Figure 5a) to capture videos of real-world scenes. These videos are then converted into an omnidirectional anaglyphic stereoscopic video as shown in Figure 5b. The scene can be only viewed from a single viewpoint, irrespective of any shifts in the user position (i.e. 3 degrees of freedom (DOF)). The 3DOF is not enough to provide an immersive experience. To provide 6DOF, it is important for the VR content to support head translation. Luo et al. [51] introduced Parallax360 to capture real-world scenes by spanning 360° and transforming the captured images into 3D stereoscopic perspectives. Notably, the Parallax360 incorporates support for head-motion parallax but due to the lack of refocusing ability, it is not possible to achieve an immersive VR experience.

The stereo imaging only captures partial 3D geometry of a scene [52] as depicted in Figure 6. The head motion parallax can be considered as partial motion parallax due to incomplete information on 3D geometry. In addition to motion parallax, refocusing in different depths is not possible with incomplete 3D geometrical information.

2.3 | Photorealistic-based 3D modelling

A virtual 3D scene can be created by photorealistic 3D modelling mainly extracting the textured geometry of a scene. This approach involves combining range sensors with traditional digital photography. The 2D images acquired by traditional digital

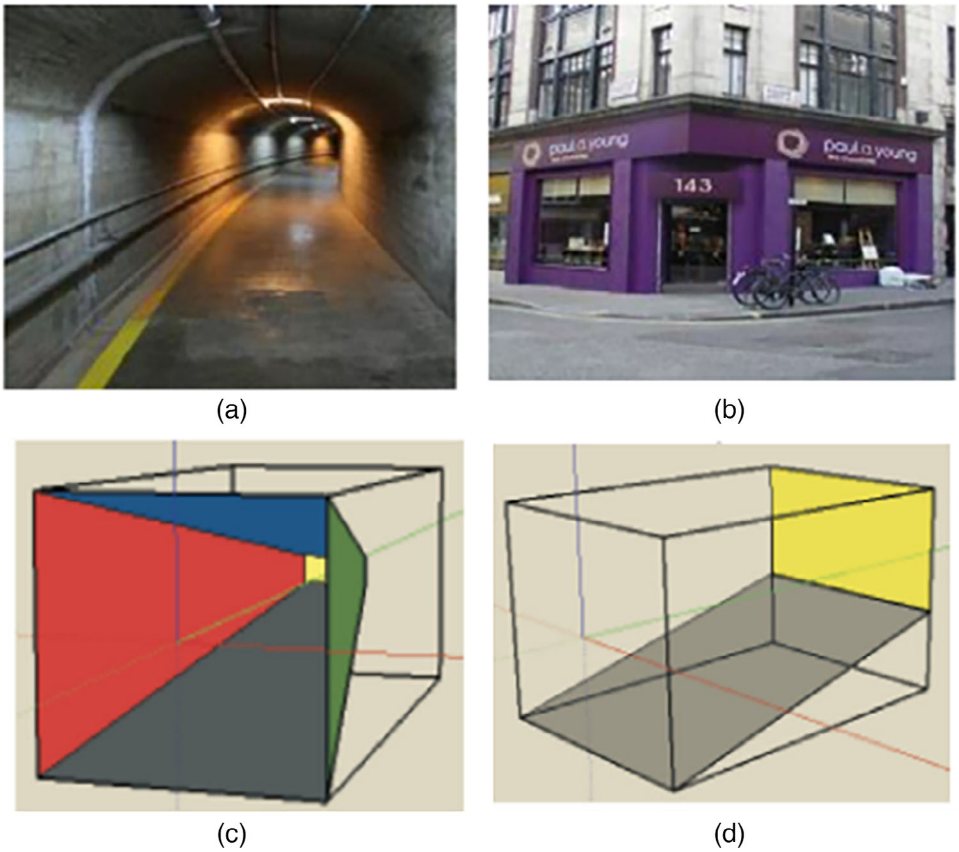

FIGURE 6 Representation of 3D geometry of a scene along with its 2D images, (a) and (b) are 2D images, (c) and (d) are their corresponding 3D geometry [52].

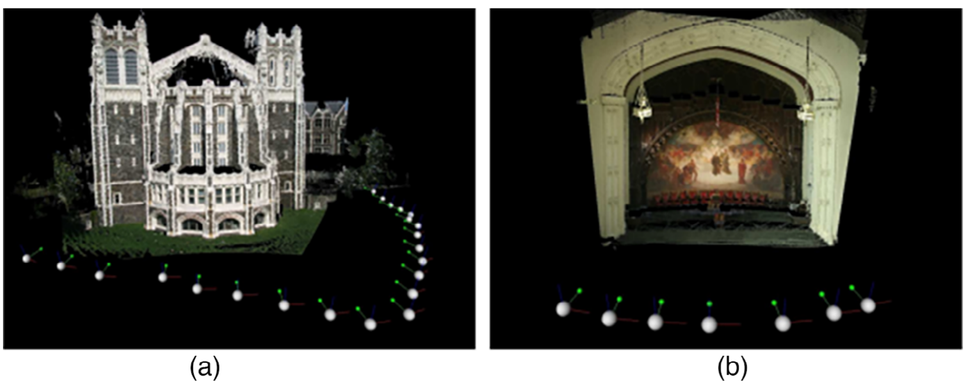

FIGURE 7 An example of reconstructed scenes, both (a) outdoor and (b) indoor. The captured images from various directions are mapped onto range data obtained using a range sensor [53].

photography are subsequently mapped onto the 3D range data through the Multiview geometry and automated 3D registration techniques. For instance, Liu et al. [53] proposed a method for modelling large-scale scenes in the photorealistic view where 3D range sensor data is combined with the 2D images obtained through conventional photography. A laser scanner such as Cyrax 2500 [54] is used to collect range data by scanning the area with an eye-safe laser beam. Figure 7a,b shows examples of outdoor and indoor 3D scenes reconstructed by photorealistic 3D modelling. Zhao et al. [55] proposed a method for modelling 3D scenes where point cloud data achieved by 3D sensors is combined with the 3D point cloud video through an Iterative Closest Point [56]. However, the photorealistic-based 3D modelling requires a very dense point cloud from the images/video sequence, also constructing such a point cloud is challenging, especially for outdoor scenes where capturing multi-view images of all objects is difficult.

2.4 | Summary of the traditional approaches

The VR content generated through the traditional approaches exhibits certain limitations [15, 24] along with the advantages. Traditional approaches to VR typically employ either a single

TABLE 1 Overview of the traditional VR content creation techniques and their advantage and limitations.

VR content generation techniques	Methodology	Advantages	Limitations
Computer-generated [36, 37, 40, 57]	3Ds Max [40], AutoCAD [57], Maya, Blender and Unity3D [36].	Users can change their viewpoint; More interactive;	Motion sicknesses such as dizziness and nausea;
Panoramic [41–47]	Images or videos are captured from a fixed point by rotating a single camera or using multiple cameras [41]; Images or videos are stitched and blended to form VR content [42, 44].	Based on real-world scenes; Creates a sense of presence in the photorealistic virtual environment;	Lack of motion parallax and refocusing; Unable to perceive the depth information of a 3D scene; Limited interactivity;
Stereo imaging [49–52, 58]	It uses binocular vision to render a view for each eye. Two cameras are used to capture two separate sets of images, each corresponding to the view of one eye [58].	Captures partial geometry of a 3D scene; Enhance VR experiences; It enables users to change their viewpoint and perceive depth;	3D geometry information is not enough to provide proper motion parallax; Unable to refocus thus lack immersive VR experiences;
Sensors and photorealistic modelling [53, 55, 59–62]	Combines 2D images with 3D range sensor data to model a virtual 3D scene [53, 55].	Better 3D point cloud to perceive the depth of objects in a 3D scene. Provide motion parallax;	Unable to handle huge outdoor scenes;

camera to capture images of a realistic environment, allowing VR view only from a fixed position or use a stereo camera to capture the same scene from two distinct perspectives. The single or stereo camera-based approach provides limited capability to achieve dynamic motion parallax and also lacks refocusing ability. However, the stereo camera-based approach offers better depth perception, restricted to achieving full 3D geometry information of the scene, resulting in limited motion parallax and the absence of refocusing ability. Moreover, the 3D modelling techniques used for the VR content generation exhibit constraints in terms of geometry, impeding the achievement of proper motion parallax and refocusing in the VR experience. An overview of the traditional VR content creation techniques, outlining their respective advantages and limitations is presented in Table 1. Based on the state-of-the-art of the traditional VR content creation, there is a demand to develop advanced technologies that are capable of capturing and generating comprehensive 3D information from a real-world environment for VR content creation. The 3D information enhances the quality of VR content significantly during the reconstruction process thus improving the overall user experience.

3 | LIGHT FIELD IMAGING TECHNOLOGY

Recently, LF technology has gained popularity in VR due to its ability to capture not only spatial information but also 2D directional information. Researchers are leveraging LF technology to reconstruct real-world 3D scenes for VR content to achieve proper immersiveness. With the inclusion of 2D directional information, LF images can be refocused at multiple depth levels and generate various angular views of the same scene, as illustrated in Figure 8. These properties of LF imaging systems contribute to the creation of VR content that closely mirrors reality. This section provides a detailed discussion of the principles, acquisition methods and applications of LF imaging in VR.

3.1 | Principle of LF

LF imaging technology captures light rays coming from the scene, which allows to recording of 2D spatial and 2D directional information of incoming light rays [26]. On the other side, conventional cameras are only limited to recording spatial data. The LF imaging used in computer vision includes image post-capture refocusing [64–66], VR [15, 24], image rendering and 3D reconstruction [67, 68] and synthesizing [69] due to its capability to provide angular information. The concept of LF was first introduced by Gershun [70] in 1936. In 1991 Adelson et al. [71] presented a 5D LF Plenoptic function, $L(x, y, z, \theta, \emptyset)$ where the (x, y, z) represent the coordinates of the light ray and the $(\theta, \emptyset)$ represent the radiance and direction of each incoming ray. Levoy et al. [72] reduced the 5D Plenoptic function to 4D $L(u, v, s, t)$ where the radiance of a light ray remains the same along its propagation direction in free space. A plane parameterization method is used to represent the LF [27, 73], the incoming light ray first intersects the (u, v) and then the (s, t) planes as shown in Figure 9. By recording the angular and directional information of incoming light rays, the LF converts 3D real-world scene information into 4D LF images.

3.2 | LF acquisition

The process of capturing LF images can be executed in various ways such as a camera array-based LF imaging system and single sensor-based LF cameras (LFC). Using a camera array-based LF imaging system, multiple cameras take images from a slightly different viewpoint of the same scene in a single shot. Where, the spatial information (s, t) can be determined by the

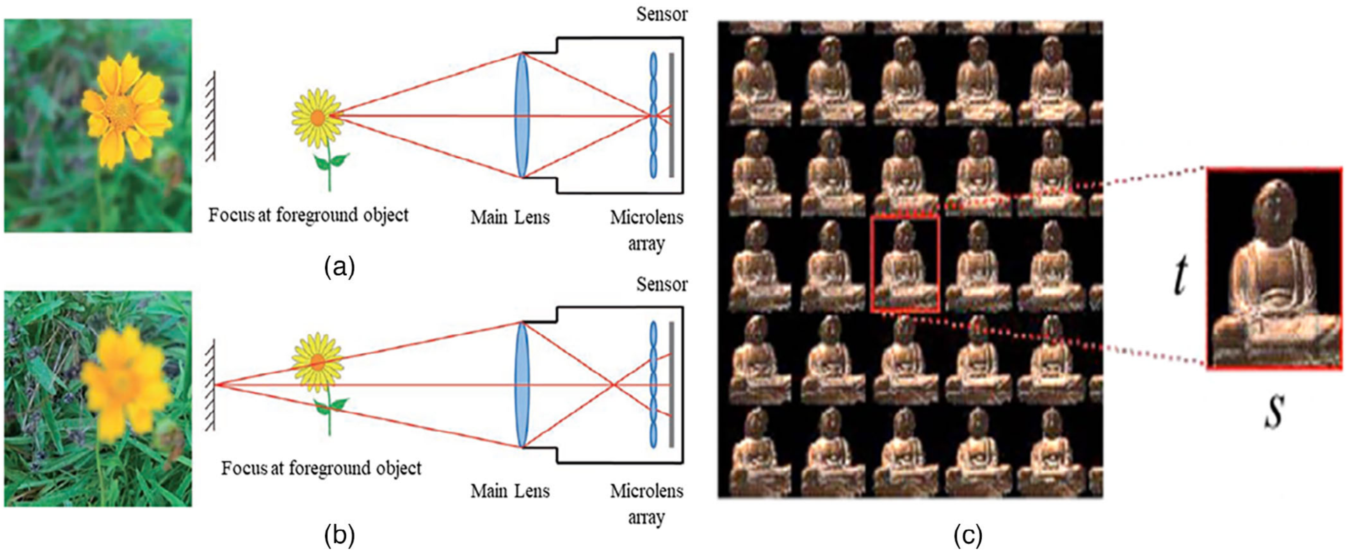

FIGURE 8 LF post-capture capabilities (a) refocusing: in the foreground, (b) refocusing: in the background [18], and (c) views of the same scene from slightly different angles [25, 63].

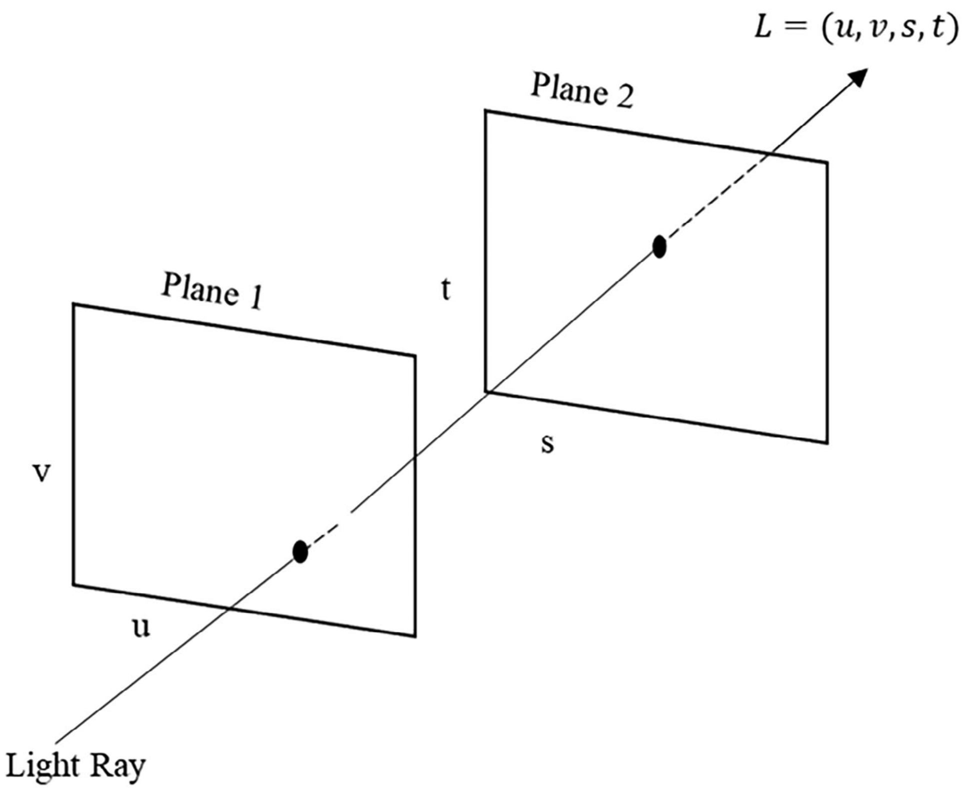

FIGURE 9 Light field representation, each light ray intersecting 2 planes at 2 positions (u,v) and (s,t) respectively, capturing angular and spatial resolution to form 4D LF $\boldsymbol{L} = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{t})$.

sensor and the directional information (u, v) can be determined by multiple cameras. The final 4D LF image can be obtained by summing all the images achieved from the multiple cameras. For instance, Wilburn et al. [74] used an 8×12 camera rig to capture high-quality LF video, as shown in Figure 10b. In this setup, each camera was equipped with a local processing board to process image data before sending it to the computer in raw form or MPEG2 video stream. To reconstruct perspectives, Zhang et al. [75] designed a 6×8 camera array-based system as shown in Figure 10c. In this system, the cameras are adjustable through a servo motor, allowing for optimal positioning for the best possible rendering results. The PiCam ultrathin camera array was

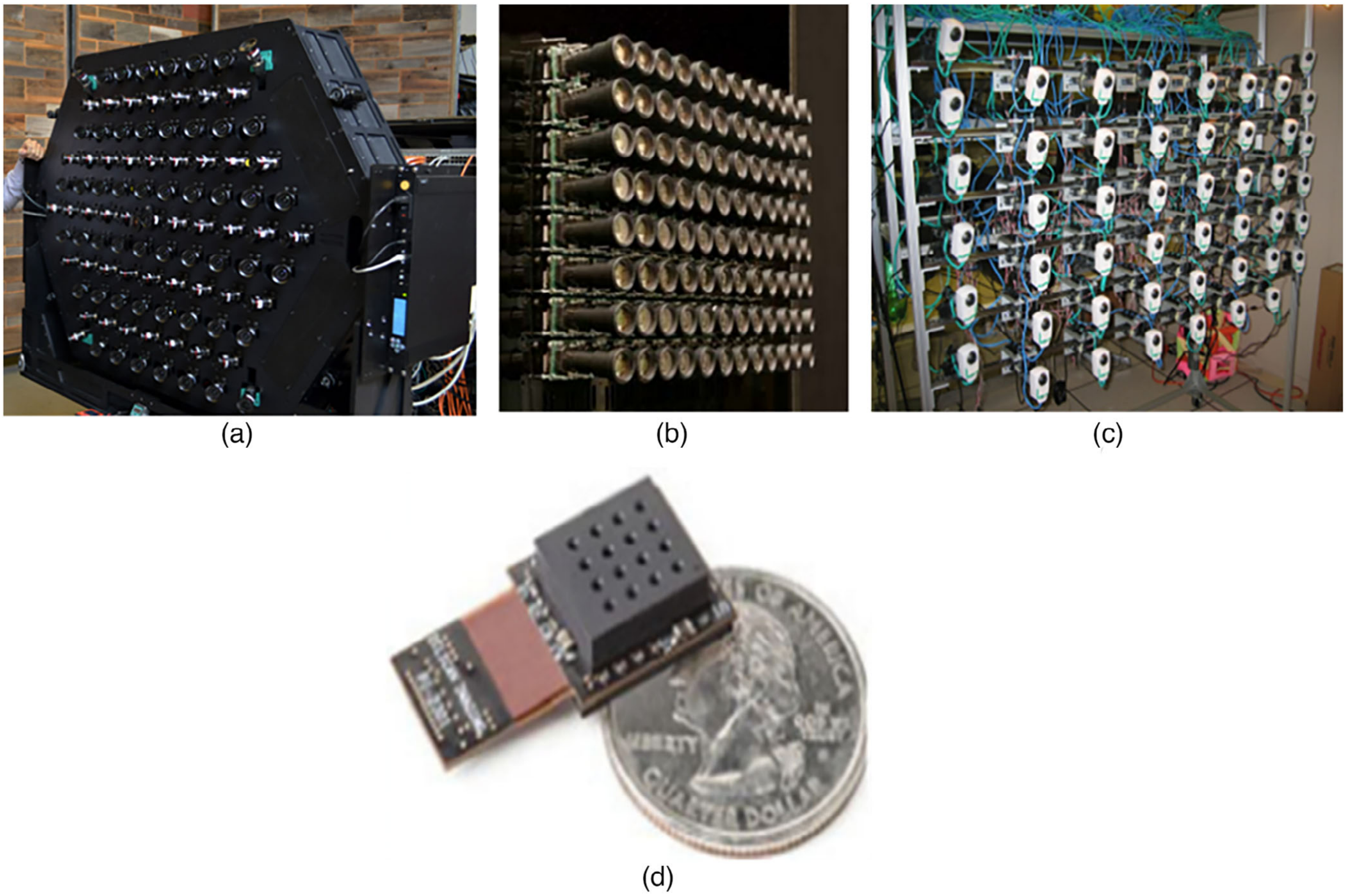

FIGURE 10 Camera array-based LF acquisition setups. (a) Lytro Immerge [14], (b) 8 × 12 camera array rig [74], (c) 6 × 8 camera array system [75], and (d) PiCam which is made up of 4 × 4 cameras [76].

FIGURE 11 Lytro LF cameras. (a) Lytro camera (Plenoptic 1.0), and (b) Lytro Illum (Plenoptic 2.0).

introduced by Venkataraman et al. [76]. PiCam is made up of 4 × 4 cameras as shown in Figure 10d and its size is very small and can be attached to a smartphone. The complete device is no bigger than a coin.

Currently, several single sensor-based LFCs are manufactured such as Lytro and Raytrix [77]. Lytro commercialised their Plenoptic cameras for the consumer market. Ng et al. [78] created the first hand-held plenoptic 1.0 as shown in Figure 11a. The Lytro camera consists of a main lens, sensor and microlens array (MLA). The main lens and the sensor are separated by an MLA. The purpose of the MLA is to split the incoming light rays and map them onto the sensor behind the corresponding microlens. The MLA functions as a tiny camera array, taking multiple pictures of the same scene from slightly different angles. Later, Lytro released its second version focused on Plenoptic 2.0 and was called as Lytro Illum camera, as shown in Figure 11b. The primary difference is that the MLA in Plenoptic 1.0 is positioned at the focal length of the main lens. In contrast, Plenoptic 2.0 is focused on the image plane of the main lens [79, 80]. The raw image size of the Lytro Illum camera is 9375 (H) × 6495 (V) pixels and the raw data is in 4D LF format, allowing for the creation of 225 sub-aperture images or different perspectives of the same scene. The size of each sub-aperture image is 625 (H) × 434 (V) pixels.

Raytrix offers various models of Plenoptic cameras for industrial and scientific applications due to their impressive effective resolution ranging from 1 to 7.25 megapixels [81]. The Raytrix cameras possess additional characteristics aimed at enhancing the depth of field (DoF) [82]. Three distinct focal lengths of MLA were introduced to enhance the DoF of the captured LF. However, the MLA with three different focal lengths creates challenges in the post-processing and manipulation of images. Each micro-image provides a different defocus blur based on the depth of the scene, adding complexity to the image processing tasks. In addition to camera array-based and MLA-based LF acquisition systems, traditional cameras are also employed for LF acquisition [83]. This approach known as time-sequential

TABLE 2 Overview of LF acquisition techniques, including their advantages and limitations.

LF acquisition techniques	Methodology	Advantages	Limitations
Camera array-based LF systems [26, 76, 84]	Multiple image sensors or cameras are used to acquire images of a scene from slightly different angles.	High spatial resolution; Wider baseline and field of view;	Complex system setup and calibration; Computational expensive;
Microlens array-based LF systems [26, 77, 78]	MLA is placed between the image sensor and the main lens to sample the light rays and acquire angular information about the scene.	Provide more angular information; Cost-effective; Easy to set up and portable;	Lower spatial resolution; Narrow baseline;
Traditional camera-based systems [26, 85, 86]	Traditional cameras are used to acquire the LF of a scene by rotating vertically or horizontally.	Higher spatial resolution; Cost-effective;	Complex acquisition process; Limited angular information;

capture, utilizes a multiple-exposure technique to capture the LF of a 3D scene. In this approach, the camera is either rotated vertically or horizontally to capture multiple images of a 360-degree 3D scene. This technique yields higher spatial resolution images by utilizing a traditional camera with a sufficiently high-resolution image sensor, although with lower angular resolution. Table 2 illustrates the overall summary of LF acquisition techniques and methodologies including their advantages and limitations.

As the LF technology can capture both the incoming light ray's spatial and the direction resolution, by adopting this technology, VR users can experience the depth of real-world scenes. The user will be able to explore in any direction and modify their viewpoint in the scene by moving their head appropriately. This technology offers enhanced depth perception and motion parallax, contributing to a more detailed understanding of real-world scenes in VR. Thus, the end effect could be a more immersive and comfortable experience for consumers.

3.3 | Light field for VR content creation

Over the past few years, various studies have been conducted to create VR content through LF imaging techniques. This section presents the different approaches used to reconstruct LF for VR content creation. These approaches are categorized into three groups based on the LF acquisition techniques such as camera array-based, microlens array-based and traditional camera-based LF acquisition systems.

3.3.1 | Camera array-based system

The camera array-based LF acquisition systems use multiple cameras to capture the LF of a 3D scene. Each camera sensor captures light rays from the same scene but from different directions, contributing to the angular information of the LF image. All images from the multiple cameras are then merged to form a 4D LF image. The multiple cameras can be arranged either vertically or on a circular/planer dome, depending on the application and the required area coverage. The vertically arranged camera rig can be rotated through a motorized mechanism to capture LF over a 360-degree area. The circular dome, on the other hand, can capture LF over 120 to 180 degrees of area. Milliron et al. [14] presented a Lytro Immerge system wherein a live song is recorded using LF and upon playback provides 6DOF. This 6DOF enables the user to freely navigate the viewing volume while maintaining flawless stereo vision and motion parallax in all directions. Lytro Immerge uses a massive camera array to capture an LF in full immersion. Yu et al. [25] presented an LF system for VR content production (called PlexVR). For LF acquisition, a dual panoramic camera array (DPCA) is employed in a circle in the form of stereoscopic camera pairs as shown in Figure 12a. Real-time stitching and data streaming are performed through a graphics processing unit (GPU). For outward scene acquisition, they have utilized the DPCA. For inward LF acquisition, they developed a dome that is composed of 140 cameras (80 static and 60 dynamic) as shown in Figure 13a. Through the DPCA, live shows like news can be captured and broadcast live in the form of 360° 3D VR. The dome can show the 360° 3D performance of humans in action as shown in Figure 13b. A small version of the dome is used to capture static objects.

Overbeck et al. [16] proposed an LF reconstruction technique and addressed the challenges of LF data size, quality and speed, also recommended that the reconstructed LF images should be of good quality to achieve a realistic view and immersive experience. A disk-based reconstruction technique is used and illustrated in Figure 14b. Where each window represents one of the camera views and is captured by the LF. For each window, the mesh is rendered. Each mesh is a tessellated version of a depth map. This disk-based technique can achieve high-quality rendering with fewer images. Tile streaming and caching techniques are proposed to reduce decoding time and fit the LF data into limited GPU memory. For data compression, a compression technique is proposed based on a modified VP9 video codec. The proposed system is limited to static images and any movement causes ghosting artefacts.

Broxton et al. [15] presented a technique for LF reconstruction by replacing a multi-plane image (MPI) scene with a

FIGURE 12 A setup for acquiring and broadcasting live VR, (a) dual panoramic camera array (DPCA), and (b) live broadcasting of 360° 3D VR [25].

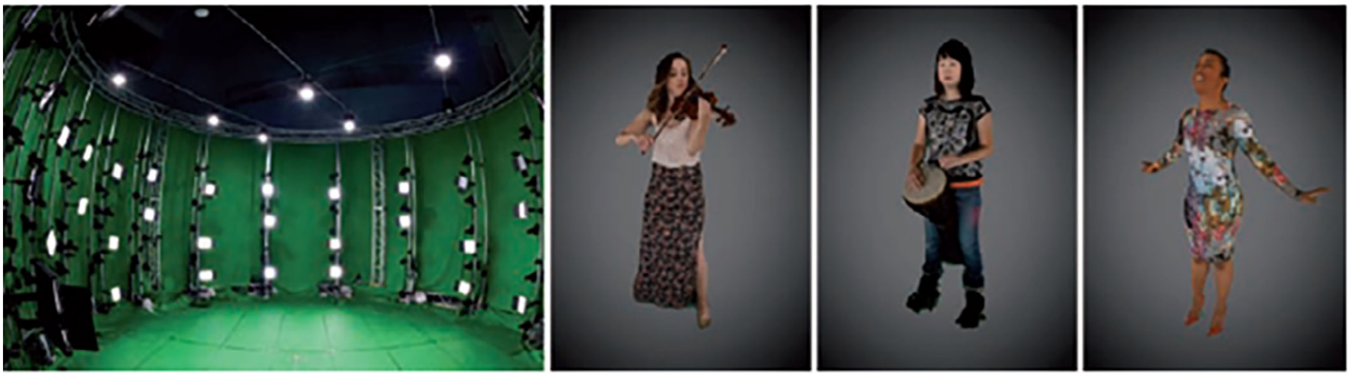

FIGURE 13 Acquiring inward-looking light fields, (a) inward light field capturing dome, comprises 140 cameras, and (b) three performances were captured with this dome and then reconstructed using light field techniques [25].

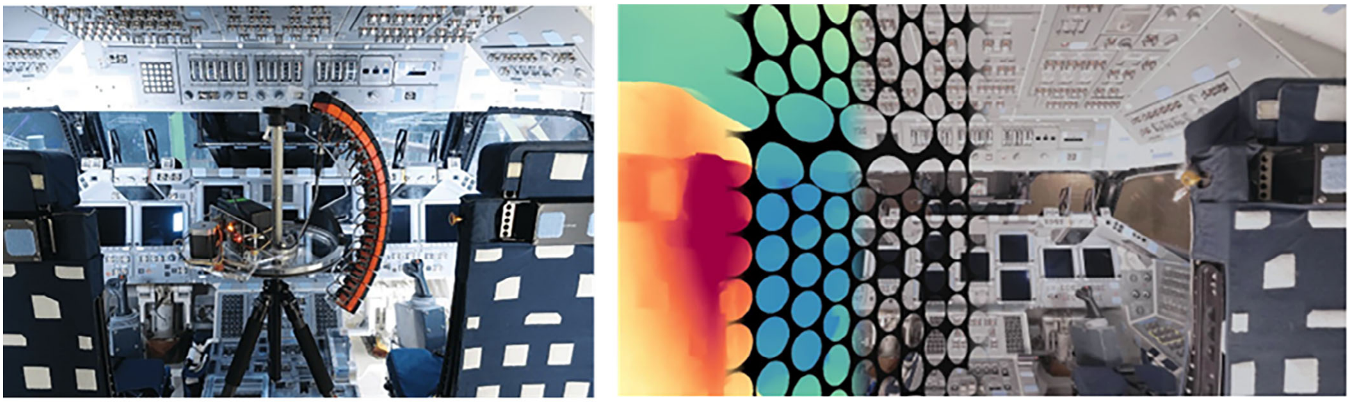

FIGURE 14 Still LF acquisition setup for VR, (a) horizontal rotating LF camera array rig, and (b) per-view depth, disk-based reconstruction, rendered image [16].

collection of sphere shells called multi-sphere images (MSI) to enable an immersive field of view (FOV) as shown in Figure 15a. The MSI encodes the scenes around the viewer as a set of concentric spheres with RGBA textures. The MSI layer is further reduced to 16 layered meshes (LM) because the MSI has more than 100 images per video frame, which is challenging to compress for efficient streaming and rendering. The LMs have the same RGBA textures and can be rendered as the same MSI. For LF acquisition, 47 action sports cameras are used and mounted on a 92 cm diameter hemisphere [84] as shown in Figure 15b. Each camera is 18 cm apart, with 120°/90° FOV. Pertuza et al. [85] utilized a single LF Lytro Illum camera to capture the LF and calculate the metric depth by utilizing focus calibration.

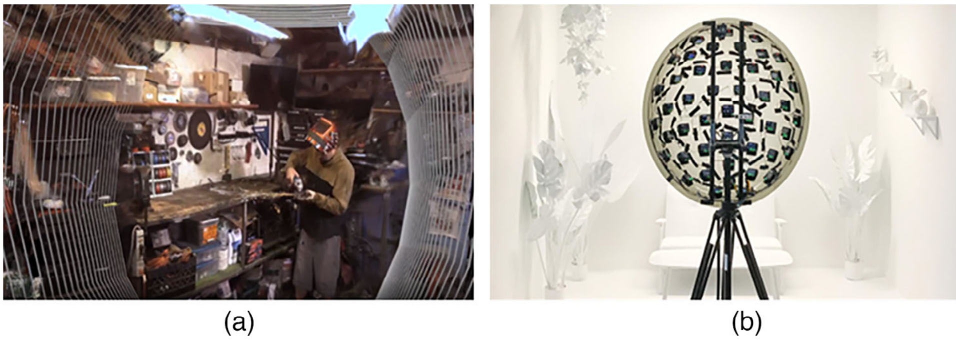

FIGURE 15 A system of acquiring and rendering LF video, (a) multi-sphere images, and (b) capture rig [15].

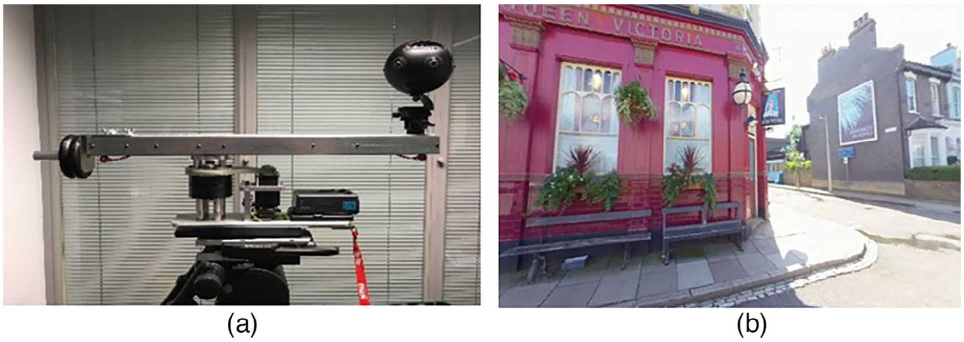

FIGURE 16 Background reconstruction using LF technology for production, (a) 360° motorised camera rig for LF acquisition, and (b) reconstructed background using Unity game engine [86, 87].

Schweiger et al. [86] proposed a technique for acquiring and rendering LF for creating realistic VR backgrounds, potentially for virtual TV production. They have utilized a 360° camera rotated in a circle with an adjustable radius between 50 and 90 cm solely in a horizontal position as illustrated in Figure 16a. The system can generate novel views for virtual camera positions within the capture circle using a dense set of source views. The Unity game engine is employed for rendering and displaying. The reconstructed background is depicted in Figure 16b. However, the limitation of their system is that the scene must remain perfectly still to avoid ghosting artefacts in synthesized views caused by moving objects and the system provides 5DoF (three for head rotation and two for movement within the plane of the capture circle). This limitation arises from capturing the LF only horizontally and neglecting the vertical viewpoints.

Although the camera array-based system provides better image resolution and larger FOV, it offers less angular resolution due to the limited number of cameras used to capture angular information. It also requires significant calibrations, especially for each camera and complex system setup. Additionally, it is expensive in terms of cost and computational requirements.

3.3.2 | Microlens array-based system

The MLA-based LF acquisition system uses a single LFC equipped with an MLA. The MLA is arranged between the main lens and the image sensor. Each microlens samples the light rays before striking the image sensor and acts as a single camera, capturing the same scene from different angles to capture directional information of the LF image. For example, a 3D point cloud-based volumetric reconstruction technique is proposed in [88] which can be used to create VR content [89–91]. In this technique, the LF images have been used. Initially, the depth map is estimated by matching sub-aperture images. The depth map is then enhanced by histogram equalization and stretching to increase the distance between adjacent depth layers. For the detection of edges in the central sub-aperture image, the Canny edge detection algorithm is utilized. The enhanced depth image and images with edge-detected images are then combined. In the end, the 3D structure of the point cloud is obtained by transforming the correspondence point plane as shown in Figure 17.

Murgia et al. [92] presented a 3D point cloud reconstruction of an object from a single Plenoptic LF image. The primary input of this technique is a depth map and a single LF image. First, the depth map is enhanced by histogram stretching to separate the depth planes better. The Sobel operator is then used to detect the edges of the LF image. The enhanced depth map and edge-detected images are combined, and masking is performed manually on the resulting image to isolate the object from the background better. In the end, the depth is computed for the points in the object. Open-source software called MeshLab has been used to process the model and produce the mesh and texture. The required steps are shown in Figure 18.

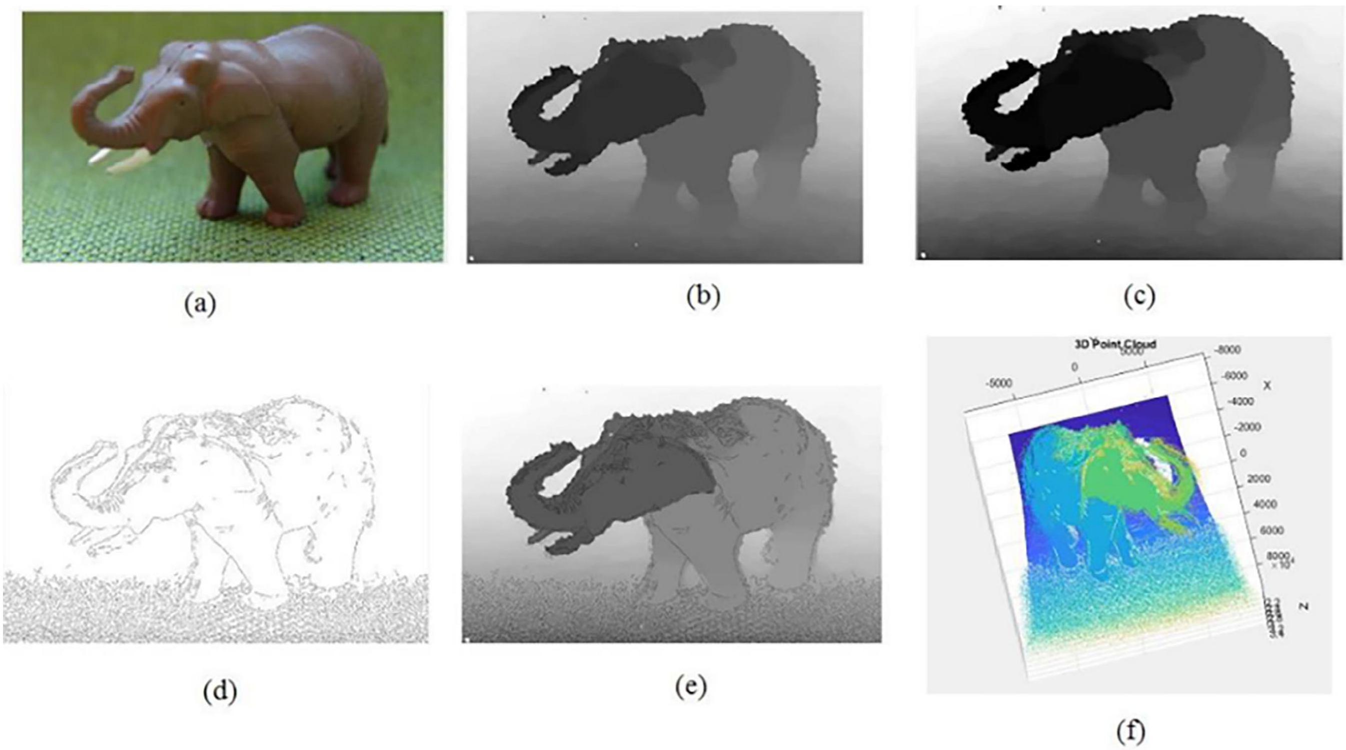

FIGURE 17 3D reconstruction from a single LF image, (a) original image, (b) depth map, (c) enhanced depth map, (d) edge detected image, (e) fussed images (c) and (d), and (f) 3D point cloud [88].

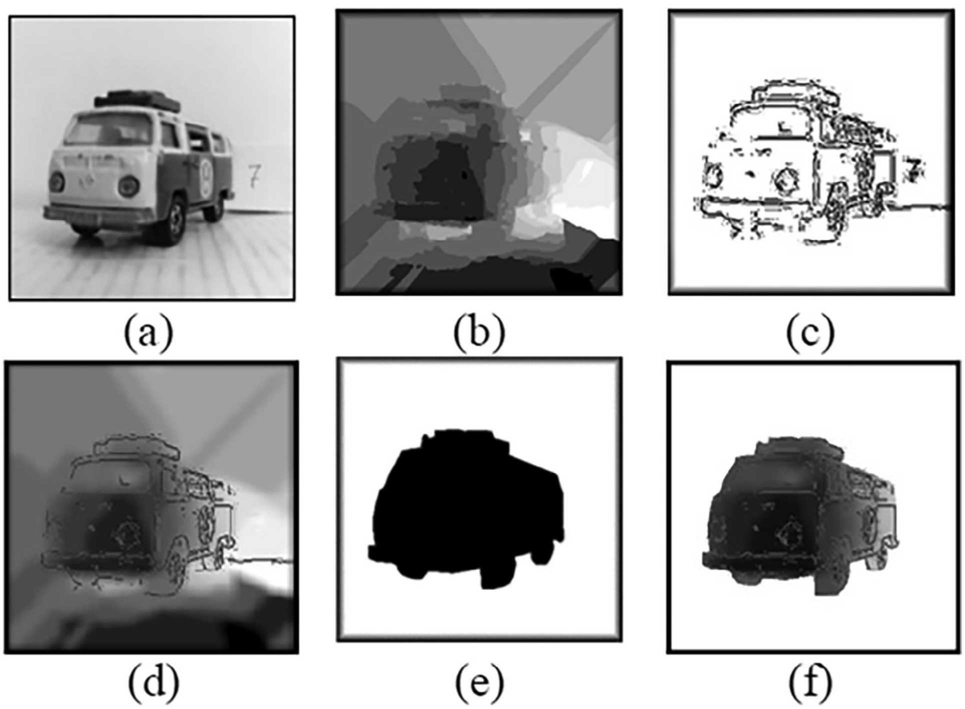

FIGURE 18 (a) Original image, (b) depth map, (c) edges detected image, (d) combined image of (b) and (c), (e) manual masking, and (f) object extracted [92].

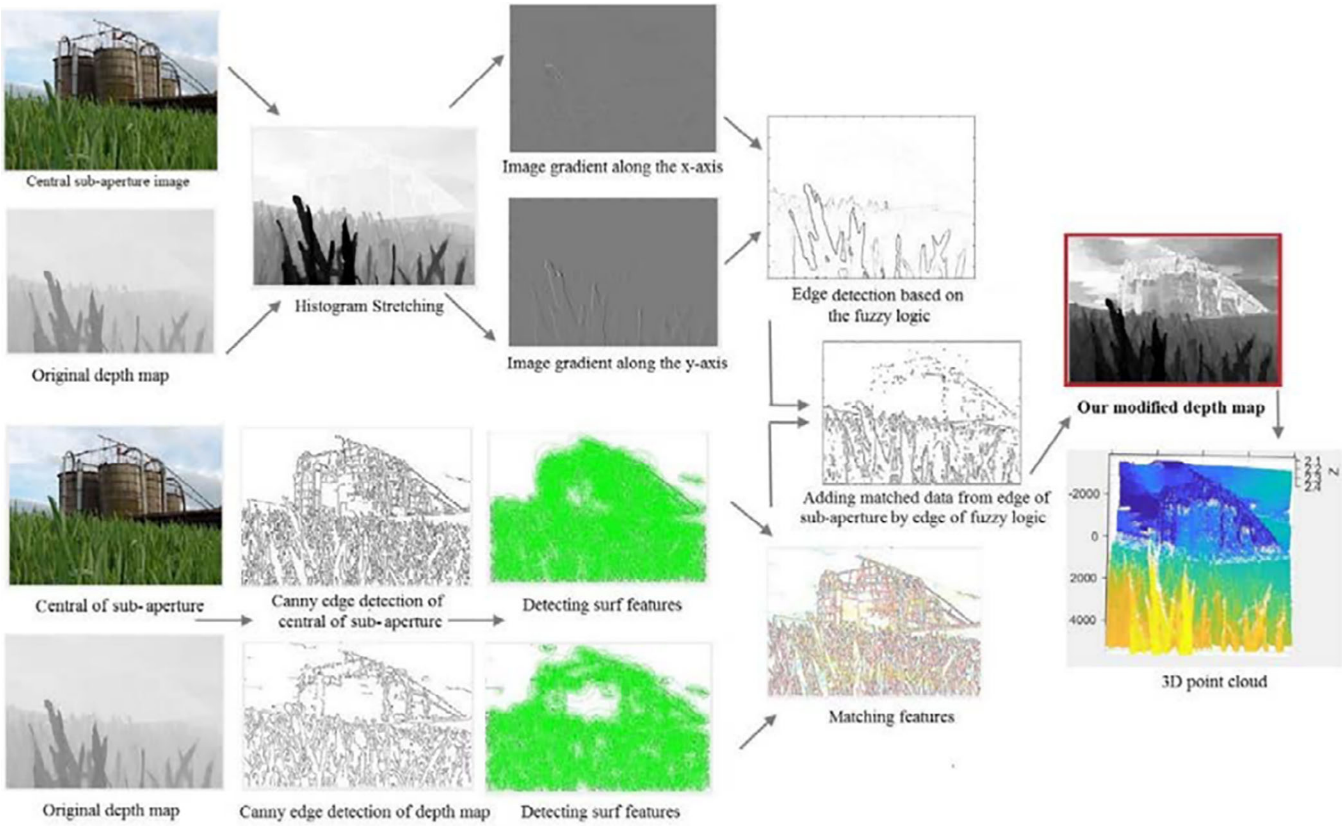

FIGURE 19 The approach used to enhance the depth map for accurate 3D point cloud estimation [93].

Farhood et al. [93] presented an enhanced 3D point cloud from a single LF image. In their proposed technique, multi-modal edge detection is considered by using feature matching and fuzzy logic. The input was the central sub-aperture LF image and the depth map. (Figure 19)

The MLA-based system provides better angular resolution and is less complex compared to the camera array-based system. Moreover, it is cost-effective and computationally less demanding. However, due to the use of a single sensor, it provides a lower spatial resolution image, which may degrade the reconstruction quality.

3.3.3 | Traditional camera-based system

The traditional camera-based LF acquisition system also known as the time-sequential capture approach utilises a single camera to capture the LF of a 3D scene using a multiple exposures technique. For instance, Debevec et al. [83] proposed an LF system to provide a motion parallax experience to users in VR. For LF acquisition a single Canon 5D Mark III DSLR camera with an 8 mm fisheye lens and a Rodeon motorised pan/tilt head has been used, as shown in Figure 20a. With the aid of a mounting rail, the camera was set 35 mm ahead from the centre of rotation. This technique offers motion parallax by taking advantage of DK2's head tracking mechanism in the HMD, which allows the user to change viewing position along with direction. The acquired images are aligned with the resolution and field of view of HMD. The dataset is compressed using the OTOY's "ORBX" codec at a ratio of 1000:1. Even though contemporary VR platforms necessitate 90 Hz for user comfort, this system reached a framerate of 75 Hz, which is sufficient for conventional real-time applications.

The traditional camera-based LF acquisition system is cost-effective as it requires only a single sensor and also provides better image resolution. However, the capturing process is time-consuming and complex which limits its application to static scenes.

3.4 | Summary of the light field for VR content creation

As the LF technology provides an accurate representation of a 3D scene by capturing both angular and spatial information, it enhances the quality of VR content significantly and improves the overall user experience. Integrating angular data into LF images facilitates immersive VR experiences, enabling users to freely explore virtual environments by moving their heads. However, current research in this domain is constrained, with most studies relying on camera array-based LF acquisition systems, as outlined in Section 3.3.1, to achieve higher

(a) (b)

FIGURE 20 Creating VR content enables shifting the user's viewpoint within the virtual environment, (a) light field capturing system, and (b) light field displayed in HMD equipped with a head tracking system [83].

TABLE 3 A comprehensive comparative analysis of different LF techniques used in VR content.

VR content generation techniques	Acquisition	Methodology	Advantages	Limitations
LF photography-based techniques [16, 86]	Multiple cameras are arranged vertically on a rig and rotated horizontally for LF acquisition.	Disk-based scene reconstruction; VP9 video codec for compression; Unity game engine is used for background reconstruction [86].	High spatial resolution; Wide field of view; Motion parallax capability;	Suitable for static environment; Lack of refocusing capability; Complex setup and calibration; High computational cost and low angular resolution;
LF video based techniques [15, 14, 25]	Multiple cameras (47–95) are used and mounted for acquisition [15, 14]. Camera array-based acquisition setups (6∼140 cameras) are used to capture inward and outward scenes [25].	MSI is used for depth representation [15]; Monte Carlo ray tracing is used for 4D integral over the 2D surface [15]; GPU is utilized for real-time processing [25].	Able to capture moving objects; High spatial resolution; Wider FOV; Motion parallax capability;	Ghosting artefacts; Unable to reconstruct thin objects; Lack of refocusing capability; Complex setup and calibrations; High computational cost and low angular resolution
3D point cloud based techniques [91, 95, 96]	A single camera is used to acquire LF.	Point cloud and 3D mesh software are used for reconstruction; Edge detection is used to achieve a depth map;	Cost-effective; Low computational cost; High angular information	Low spatial resolution; Unable to reconstruct longer distance due to narrow baseline; Limited field of view;
Traditional camera-based technique [83]	DSLR camera is used and rotated for LF acquisition.	The dataset is compressed using the OTOY's "ORBX" codec.	High spatial resolution; Motion parallax capability; Cost-effective;	Complex capturing process and calibration; Low angular resolution;

spatial resolution. Nevertheless, there exists a trade-off between spatial and angular resolutions. Enhancing spatial resolution often sacrifices angular resolution, resulting in degraded scene reconstruction quality. Table 3 illustrates a comprehensive comparative analysis of different LF techniques concerning VR content creation.

To achieve a truly immersive and cyber-sickness-free VR experience, it is crucial to replicate human vision. Human vision excels in two key aspects: motion parallax and refocusing. Presently, there is a notable gap in LF research for VR content generation regarding the lack of refocusing capability during VR experiences. To address this, future research efforts should prioritize developing LF technology that replicates the dynamic refocusing capabilities inherent in human vision for a more authentic and comfortable VR experience.

4 | CHALLENGES OF LIGHT FIELD IMAGING TECHNOLOGY

Though LF technology has the potential demand for VR content creation and offers several advantages over traditional imaging technology, it has significant challenges that must be overcome to produce realistic and immersive reconstruction of real-world VR content. In this section, several key challenges are discussed.

4.1 | Image size and storage

Applications of the LF are constrained by the volume of data. Because the large amount of data generated by a 4D LF poses problems for its storage and transmission. For example, in Lytro Illum 4D LF, each raw LF image is roughly 150 MB in size. Therefore, it is crucial to create LF compression techniques to expand its uses. LF possesses redundancy in all of its dimensions [72]. By taking advantage of that redundant information, LF data can be compressed. Several compression strategies have been proposed to use the raw LF data for correlations within and between the nearby macro pixel images (i.e. created on the image sensor behind each microlens) [94, 95]. For the LF compression, the redundancies of adjacent sub-aperture images can be reduced [96, 97]. A learning-based reconstruction approaches are proposed to encode sparsely sampled sub-aperture images at the encoder side and build all sub-aperture images at the decoder side [98, 99]. However, LF compression has made significant advancements, and certain challenges remain to be addressed appropriately. For instance, Learning-based techniques may introduce artefacts in occluded regions, thereby diminishing the overall quality of the reconstruction. Additionally, the redundancies present in adjacent sub-aperture images, intended for compression purposes, may adversely affect the quality of VR content. It is also crucial for the VR content reconstruction that the captured data keep the maximum angular resolution.

4.2 | Image resolution

There is an inherent trade-off between angular resolution and spatial resolution of the LF imaging systems [100]. Camera arrays capture high spatial resolution images but not enough angular information. The angular information is crucial for providing the 6DOF and refocusing ability. Single-sensor microlens-based cameras like Lytro Illum capture low-resolution images with high angular resolution. For example, the sensor resolution of Illum is 9375 (H) × 6495 (V) pixels. But the resolution of each sub-aperture image is 625 × 434 pixels in horizontal and vertical respectively [81, 101]. The low spatial resolution of an image can significantly affect the quality of the VR experience across various aspects, such as realism, immersion, depth perception, aliasing and blurriness. High spatial resolution images offer a more detailed and accurate representation of the virtual environment, thereby enhancing the overall realism of the VR experience. Additionally, high spatial resolution images contribute to a smoother and clearer visual experience and help to reduce issues like aliasing and blurriness.

4.3 | Speed

Another challenge of LF technology is the rendering speed for VR content creation. Specifically, it is crucial to have higher frame rates for VR content. Because a true immersion with comfort can be achieved with exceedingly high frame rates. For typical applications, a frame rate of 30–60 Hz is considered sufficient. To ensure the best comfort when viewing VR content, current VR headsets aim for a frame rate of 120–180 fps [102]. However, achieving a high frame rate can be challenging due to factors such as simulation and rendering speed. The low frame rate can have various adverse effects on the user, including simulator sickness [102], decreased presence and reduced task performance [103].

4.4 | Precise calibration and depth reconstruction

Accurate depth reconstruction in LF poses significant challenges for achieving a comfortable, immersive and realistic VR experience. Perceiving accurate depth provides users with refocusing abilities in VR environments. However, accurate depth estimation can be affected by various factors such as high-dimensional data, computational complexity and appropriate and precise calibrations. The high-dimensional LF data contains spatial-angular information on all incoming rays from different directions. Thus, processing such high-dimensional data can be computationally intensive. Additionally, depth calibration [85, 104] is crucial for accurate depth reconstruction because inconsistent calibrations can increase depth reconstruction errors. Other calibrations such as geometric [105–108] and camera lens distortion calibrations are also challenging for accurate reconstruction. For instance, geometric calibration involves estimating the intrinsic and extrinsic parameters of LFC. The geometric calibration is crucial for reconstructing novel views using the spatial-angular information of LF images. These novel views can be used to provide motion parallax or refocused perspectives for a realistic VR experience. In LFC, imperfect alignment of the MLA with the camera sensor is common [109] due to manufacturing errors, which can affect the reconstruction quality and consequently the user experience in VR.

4.5 | LF acquisition

The LF acquisition is a fundamental aspect of the LF imaging process. As discussed in Section 3.3 three types of LF acquisition systems, these acquisition systems offer certain advantages with limitations. As LF captures both spatial and directional information of a 3D scene in the form of a 4D dataset, acquiring such high-dimensional data poses significant challenges. The acquisition of high-dimensional data typically introduces a resolution trade-off between spatial and angular dimensions. Furthermore, handling such high-dimensional data increases computational complexity, thus, developing suitable algorithms is more challenging.

4.6 | Translation to VR platform

Another challenge is the translation tool for VR platforms especially when developing immersive VR content through LF imaging technology. VR content developers and researchers commonly employ game engines to simulate the reconstructed LF in VR. However, during the simulation, they encounter issues in integrating virtual depth planes or different viewing perspectives, ensuring a high frame rate and achieving a realistic experience. The absence of generalized guidelines and a tool for developers creates significant challenges when translating the real-world 3D scene after LF reconstruction. Developers could benefit from a universal tool with supporting guidance that facilitates the translation of LF imaging-based reconstructed VR content. Apart from this, achieving immersiveness and interactivity in VR poses unique challenges compared to traditional media, such as 2D games or videos. Developers must take into account user interaction, allowing users to interact with the virtual environment or objects within the reconstructed scene. In Section 5, recommendations have been made on how to address these challenges along with potential solutions.

5 | FUTURE RECOMMENDATION

In the future, to enhance the realism of VR experiences, the challenges discussed in Section 4 can be considered to improve the LF imaging technology for VR content creation. Therefore, the following recommendations have been made to the scientific community and researchers who are working on the LF imaging-based VR technology based on the challenges addressed in Section 4.

- As the LF data comprises high-dimensional data, resulting in a large image size that limits its practical applications, redundant information of the LF data can be leveraged to compress its size, making it easily transmissible and storable. Also, advanced data compression techniques [96, 97, 110] can reduce the computational complexity of the LF data. A hybrid video encoder such as the joint exploration model (JEM) can be employed to encode a sparse set of views [110]. A linear approximation can be utilized to estimate a second sparse set of views. On the decoder side, a deep learning approach can be implemented to obtain the entire LF. Image streaming of LF sub-aperture can be improved to compress the lenslet images and reduce the redundancy in lenslet images [96]. Multi-focus images from LF angular can be compressed as a sequence using discrete cosine transform (DCT) or discrete wavelet transform (DWT) [111].
- Image resolution is another crucial factor for creating immersive VR experiences. Single-sensor-based LFCs may not provide sufficient resolution for realistic VR experiences. So, the resolution issue can be addressed either on the hardware side such as by increasing the sensor size, or on the software side such as by applying super-resolution techniques [112–114] for rendering LF images. For light-field patches, the Gaussian mixture model (GMM) based super-resolution method can be considered [115]. For continuous disparity maps, epipolar plane image (EPI) based super-resolution method can be employed [116].
- Frame rate plays a vital role in ensuring a comfortable and motion-sickness-free VR experience. The higher framerate provides a better user experience with reduced simulator sickness problems. The VR content developer should consider at least a frame rate of 120 fps or higher to provide better user performance [102]. Achieving the maximum frame rate can be possible by reducing computational complexity on the translation side and also selecting the right HMDs, which can support higher frame rate VR content.
- Depth reconstruction is essential for VR content and can be achieved through accurate depth estimation. LF imaging captures spatial-angular information of light rays, providing various depth cues such as correspondence cues, defocus cues, binocular disparity, aerial perspective, and motion parallax. These cues can be leveraged for depth estimation [85, 104] and reconstruction thus facilitating virtual refocusing and allowing users to change their refocus plane within the virtual environment. Advanced focus cues technique can be utilised to estimate metric depth, which can subsequently be utilized for depth reconstruction [117]. Convolutional neural networks (CNNs) can also be considered to estimate all-in-focus images from focal stack images and then a 4D ray depth can be estimated from all the in-focus images for rendering the views.
- A universal translation tool is needed for seamlessly translating the reconstructed LF to VR. VR content developers should be provided with appropriate guidelines or software tools that can assist them in translating the reconstructed content through LF technology. The tool should be universal, enabling the development of content for available VR headsets. Additionally, developers need to create advanced techniques to enhance user comfort during VR experiences. This involves providing interaction capabilities to users, ensuring that the content is context-aware, and incorporating user-centric localization features.
- To achieve an immersive and comfortable user experience without motion sickness, VR content developers should also take the following considerations [25] such as,

Motion parallax: This feature plays a crucial role in providing 6DOF contributing to a sense of movement in VR. It enables the users to change their viewpoint and explore the 3D scene from different perspectives, enhancing the overall experience. Proper implementation of motion parallax is essential for a comfortable VR experience, aligning with users' expectations based on real-world encounters. This alignment helps to reduce visual discomfort. Motion parallax can be integrated into VR content by utilizing the angular information of LF images or the sub-aperture images. The sub-aperture images are formed by integrating the light rays coming from the same angle striking the image sensor.

Refocusing: The refocusing ability is another important aspect enabling users to refocus on different objects within the virtual scene, mimicking the natural behaviour of human vision. This capability allows users to perceive the relative depth of objects in the virtual environment. Refocusing in VR can be achieved by integrating depth planes in VR content.

Motion sickness: VR content tends to induce motion sickness in users when viewed through a VR headset. This prevents users from experiencing VR content for extended periods and limits the VR applications. Efforts can be made to investigate and identify the main factors contributing to motion sickness, allowing for the development of solutions to address this issue.

- In addition, researchers and the scientific community can explore cloud-based rendering solutions to reduce the computational demands of LF content creation and also explore developing a more affordable and accessible LF for capture devices.

By adopting the above recommendations, the LF imaging-based future VR content creation can be more efficient and capable of delivering compelling and realistic virtual experiences.

6 | CONCLUSION

This paper presents a comprehensive review of technologies utilized in the generation of VR content, with a specific focus on two primary methods: traditional approaches and LF methods. The advantages and limitations of each approach are thoroughly reviewed in this paper. The key points achieved through this review are summarized below:

- Traditional techniques for content generation encompass a range of methods such as computer-generated imagery, panoramic photography, stereo imaging, and 3D modelling. Each method is examined along with its respective limitations. For instance, panoramic content allows viewing from a fixed position with limited ability to explore the scene. Stereo imaging offers stereoscopic pairs but lacks sufficient 3D geometry for motion parallax. Current 3D reconstruction approaches are unable to handle large outdoor scenes when it comes to 3D modelling to achieve multi-view coverage of all objects in a scene.
- LF technology has shown an advantage over traditional photography when it comes to viewing real-world content in VR. It accurately captures the 3D scene by recording all light rays travelling through it, providing spatial and directional information. LF-based VR content meets most VR requirements with features like refocusing and motion parallax for a more immersive and realistic user experience.
- The LF has the potential to significantly enhance VR content and user experience by capturing angular and spatial information from real-world environments. However, existing literature highlights the deficiency in refocusing capability within current VR content research.
- The camera array-based LF acquisition systems show a trade-off between spatial and angular resolution. While these systems offer improved spatial resolution, they often lack sufficient angular resolution, resulting in suboptimal reconstruction quality.
- LF technology faces various challenges in the creation of realistic VR content, including issues related to image size and storage, image resolution, processing speed, calibration, depth reconstruction, LF acquisition, and the LF-to-VR translation platform. Overcoming these challenges is essential for achieving an immersive, realistic and comfortable VR experience.

Future efforts in LF technology should prioritize addressing these challenges to enhance the overall quality of VR content. Recommendations for future research focus on overcoming the challenges to enhance the realism of VR experiences and advancing LF imaging technology for future VR content creation.

AUTHOR CONTRIBUTIONS

Ali Khan: Writing—original draft; investigation; collection of materials; references. **Md. Moinul Hossain**: Technical discussion; re-writing; review; editing; supervision. **Konstantinos Sirlantzis**: Funding acquisition. **Alexandra Covaci**: Analysis; review. **Chuanlong Xu**: Analysis; review.

ACKNOWLEDGEMENTS

The authors acknowledge receipt of the following financial support for the research, authorship, and publication of this article: Interreg 2 Seas programme 2014–2020 co-funded by the European Regional Development Fund under subsidy contract No. 2S05-038 (MOTION project).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

As this paper constitutes a review article, there exists no dataset, model or code necessitating sharing. All data and information expounded within this review stem from antecedently published sources, each meticulously acknowledged and cited within the reference section.

ORCID

Ali Khan https://orcid.org/0000-0003-2896-2095
Md. Moinul Hossain https://orcid.org/0000-0003-4184-2397
Alexandra Covaci https://orcid.org/0000-0002-3205-2273
Konstantinos Sirlantzis https://orcid.org/0000-0002-0847-8880
Chuanlong Xu https://orcid.org/0000-0003-0083-1544

REFERENCES

1. Delgado, J.M.D., Oyedele, L., Demian, P., Beach, T.: A research agenda for augmented and virtual reality in architecture, engineering and

construction. Adv. Eng. Inf. 45, 101122 (2020). https://doi.org/10.1016/j.aei.2020.101122
2. Checa, D., Bustillo, A.: A review of immersive virtual reality serious games to enhance learning and training. Multimedia Tools Appl. 79(9), 5501–5527 (2020). https://doi.org/10.1007/s11042-019-08348-9
3. Ahir, K., Govani, K., Gajera, R., Shah, M.: Application on virtual reality for enhanced education learning, military training and sports. Augment. Hum. Res. 5(1), 7 (2019). https://doi.org/10.1007/s41133-019-0025-2
4. Katz, D., et al.: Utilization of virtual reality for operating room fire safety training: A randomized trial. Virtual Reality 27(4), 3211–3219 (2023). https://doi.org/10.1007/s10055-023-00866-0
5. Teeng, C., Lim, C.K., Tan, K.: Challenges in virtual reality system: A review. AIP Conf. Proc., 2016, 020037, Paris, France (2018). https://doi.org/10.1063/1.5055439
6. Nicholson, D.T., Chalk, C., Funnell, W.R.J., Daniel, S.J.: Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. Med. Educ. 40(11), 1081–1087 (2006). https://doi.org/10.1111/j.1365-2929.2006.02611.x
7. Basdogan, C., Sedef, M., Harders, M., Wesarg, S.: VR-based simulators for training in minimally invasive surgery. IEEE Comput. Graph. Appl. 27(2), 54–66 (2007). https://doi.org/10.1109/MCG.2007.51
8. van Twillert, B., Bremer, M., Faber, A.W.: Computer-generated virtual reality to control pain and anxiety in pediatric and adult burn patients during wound dressing changes. J. Burn Care Res. 28(5), 694–702 (2007). https://doi.org/10.1097/BCR.0B013E318148C96F
9. Lin, M., San, L., Ding, Y.: Construction of robotic virtual laboratory system based on Unity3D. IOP Conf. Ser. Mater. Sci. Eng. 768(7), 072084 (2020). https://doi.org/10.1088/1757-899X/768/7/072084
10. Safadel, P., White, D.: Effectiveness of computer-generated virtual reality (VR) in learning and teaching environments with spatial frameworks. Appl. Sci. 10(16), 5438 (2020). https://doi.org/10.3390/app10165438
11. Corbillon, X., Simon, G., Devlic, A., Chakareski, J.: Viewport-adaptive navigable 360-degree video delivery. In: *2017 IEEE International Conference on Communications (ICC)*, pp. 1–7. Paris, France (2017). https://doi.org/10.1109/ICC.2017.7996611
12. Qi, Q., Hossain, M.M., Li, J.-J., Zhang, B., Li, J., Xu, C.-L.: Approach to reduce light field sampling redundancy for flame temperature reconstruction. Opt. Express 29(9), 13094 (2021). https://doi.org/10.1364/OE.424112
13. Sun, J., et al.: Three-dimensional temperature field measurement of flame using a single light field camera. Opt. Express 24(2), 1118 (2016). https://doi.org/10.1364/OE.24.001118
14. Milliron, T., Szczupak, C., Green, O.: Hallelujah: The world's first lytro VR experience. In: *ACM SIGGRAPH 2017 VR Village*, pp. 1–2. Los Angeles, CA (2017). https://doi.org/10.1145/3089269.3089283
15. Broxton, M., et al.: Immersive light field video with a layered mesh representation. ACM Trans. Graph. 39(4), 1–15 (2020). https://doi.org/10.1145/3386569.3392485
16. Overbeck, R.S., Erickson, D., Evangelakos, D., Debevec, P.: The making of welcome to light fields VR. In: *ACM SIGGRAPH 2018 Talks*, in SIGGRAPH '18, pp. 1–2. New York, NY (2018). https://doi.org/10.1145/3214745.3214811
17. Jia, C., Shi, F., Zhao, M., Chen, S.: Light field imaging for computer vision: A survey. Front. Inf. Technol. Electron. Eng. 23(7), 1077–1097 (2022). https://doi.org/10.1631/FITEE.2100180
18. Broxton, M., et al.: Wave optics theory and 3-D deconvolution for the light field microscope. Opt. Express 21(21), 25418 (2013). https://doi.org/10.1364/OE.21.025418
19. Sheng, H., Deng, S., Zhang, S., Li, C., Xiong, Z.: Segmentation of light field image with the structure tensor. In: *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 1469–1473. Phoenix, AZ (2016). https://doi.org/10.1109/ICIP.2016.7532602
20. Georgiev, T.G., Lumsdaine, A.: Focused plenoptic camera and rendering. J. Electron. Imaging 19(2), 021106 (2010). https://doi.org/10.1117/1.3442712
21. Li, N., Ye, J., Ji, Y., Ling, H., Yu, J.: Saliency detection on light field. IEEE Trans. Pattern Anal. Mach. Intell. 39(8), 1605–1616 (2017). https://doi.org/10.1109/TPAMI.2016.2610425
22. Agus, M., et al.: An interactive 3D medical visualization system based on a light field display. Vis. Comput. 25(9), 883–893 (2009). https://doi.org/10.1007/s00371-009-0311-y
23. Bajpayee, A., Techet, A.H., Singh, H.: Real-time light field processing for autonomous robotics. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4218–4225. Madrid (2018). https://doi.org/10.1109/IROS.2018.8594477
24. Overbeck, R.S., Erickson, D., Evangelakos, D., Pharr, M., Debevec, P.: A system for acquiring, processing, and rendering panoramic light field stills for virtual reality. ACM Trans. Graph. 37(6), 1–15 (2018). https://doi.org/10.1145/3272127.3275031
25. Yu, J.: A light-field journey to virtual reality. IEEE Multimedia 24(2), 104–112 (2017). https://doi.org/10.1109/MMUL.2017.24
26. Wu, G., et al.: Light field image processing: An overview. IEEE J. Sel. Top. Signal Process. 11(7), 926–954 (2017). https://doi.org/10.1109/JSTSP.2017.2747126
27. Zhou, S., Zhu, T., Shi, K., Li, Y., Zheng, W., Yong, J.: Review of light field technologies. Vis. Comput. Ind. Biomed. Art. 4(1), 29 (2021). https://doi.org/10.1186/s42492-021-00096-8
28. Ihrke, I., Restrepo, J., Mignard-Debise, L.: Principles of light field imaging: Briefly revisiting 25 years of research. IEEE Signal Process. Mag. 33(5), 59–69 (2016). https://doi.org/10.1109/MSP.2016.2582220
29. Yin, K., He, Z., Xiong, J., Zou, J., Li, K., Wu, S.-T.: Virtual reality and augmented reality displays: Advances and future perspectives. J. Phys. Photonics 3(2), 022010 (2021). https://doi.org/10.1088/2515-7647/abf02e
30. Zhan, T., Yin, K., Xiong, J., He, Z., Wu, S.-T.: Augmented reality and virtual reality displays: Perspectives and challenges. iScience 23(8), 101397 (2020). https://doi.org/10.1016/j.isci.2020.101397
31. Simon, A., Kara, P.A.: The good news, the bad news, and the ugly truth: A review on the 3d interaction of light field displays. Multimodal Technol. Interact. 07(05), 45 (2023). https://doi.org/10.3390/mti7050045
32. Ma, Q., Cao, L., He, Z., Zhang, S.: Progress of three-dimensional light-field display. Chin. Opt. Lett. 17(11), 111001 (2019)
33. Wetzstein, G., Lanman, D., Hirsch, M., Heidrich, W., Raskar, R.: Compressive light field displays. IEEE Comput. Graph. Appl. 32(5), 6–11 (2012). https://doi.org/10.1109/MCG.2012.99
34. Li, T., et al.: Light-field displays: A view-dependent approach. In: *ACM SIGGRAPH 2020 Emerging Technologies*, Virtual Event USA, pp. 1–2 (2020). https://doi.org/10.1145/3388534.3407293
35. Dudley, J., Yin, L., Garaj, V., Kristensson, P.O.: Inclusive Immersion: A review of efforts to improve accessibility in virtual reality, augmented reality and the metaverse. Virtual Reality 27(4), 2989–3020 (2023). https://doi.org/10.1007/s10055-023-00850-8
36. Gabajová, G., Krajčovič, M., Matys, M., Furmannová, B., Burganova, N.: Designing virtual workplace using unity 3D game engine. Acta Technol. 7, 35–39 (2021). https://doi.org/10.22306/atec.v7i1.101
37. Wang, S., Mao, Z., Zeng, C., Gong, H., Li, S., Chen, B.: A new method of virtual reality based on Unity3D. In: *2010 18th International Conference on Geoinformatics*, pp. 1–5. Beijing, China (2010). https://doi.org/10.1109/GEOINFORMATICS.2010.5567608
38. Ju, R., et al.: Ultra wide view based panoramic VR streaming. In: *Proceedings of the Workshop on Virtual Reality and Augmented Reality Network*, in VR/AR Network '17, pp. 19–23. New York, NY (2017). https://doi.org/10.1145/3097895.3097899
39. Keselman, L., Iselin Woodfill, J., Grunnet-Jepsen, A., Bhowmik, A.: Intel RealSense stereoscopic depth cameras. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1–10. (2017)
40. Kuang, Y., Bai, X.: The research of virtual reality scene modeling based on unity 3D. In: *2018 13th International Conference on Computer Science & Education (ICCSE)*, Colombo Srilanka, pp. 1–3. (2018). https://doi.org/10.1109/ICCSE.2018.8468687

41. Kuna, P., Hašková, A., Borza, Ľ.: SWOT analysis of virtual reality creation equipments, *RE-SOURCE*, S24, pp. 28–32 (2022). https://doi.org/10.53349/resource.2022.iS24.a1106
42. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. Int. J. Comput. Vis. 74(1), 59–73 (2007). https://doi.org/10.1007/s11263-006-0002-3
43. Philip, S., Summa, B., Tierny, J., Bremer, P.-T., Pascucci, V.: Distributed seams for gigapixel panoramas. IEEE Trans. Vis. Comput. Graph. 21(3), 350–362 (2015). https://doi.org/10.1109/TVCG.2014.2366128
44. Szeliski, R.: Image alignment and stitching: A tutorial. Found. Trends Comput. Graph. Vis. 2(1), 1–104 (2007). https://doi.org/10.1561/0600000009
45. Sayyad, E., Sen, P., Höllerer, T.: PanoTrace: Interactive 3D modeling of surround-view panoramic images in virtual reality. In: *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10. Gothenburg Sweden (2017). https://doi.org/10.1145/3139131.3139158
46. Abdul Rahim, E., et al.: A desktop virtual reality application for chemical and process engineering education. In: *Proceedings of the 24th Australian Computer-Human Interaction Conference*, pp. 1–8. Melbourne Australia (2012). https://doi.org/10.1145/2414536.2414537
47. Tsai, H.-H., et al.: Interactive contents with 360-degree panorama virtual reality for soil and water conservation outdoor classroom. In: *2020 International Symposium on Educational Technology (ISET)*, pp. 78–82. Bangkok, Thailand (2020). https://doi.org/10.1109/ISET49818.2020.00026
48. See, Z.S., Cheok, A.D.: Virtual reality 360 interactive panorama reproduction obstacles and issues. Virtual Reality 19(2), 71–81 (2015). https://doi.org/10.1007/s10055-014-0258-9
49. Richardt, C., Tompkin, J., Wetzstein, G.: Capture, reconstruction, and representation of the visual real world for virtual reality. In: Real VR—Immersive Digital Reality: How to Import the Real World into Head-Mounted Immersive Displays. pp. 3–32. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-41816-8_1
50. Anderson, R., et al.: Jump: Virtual reality video. ACM Trans. Graph. 35(6), 1–13 (2016). https://doi.org/10.1145/2980179.2980257
51. Luo, B., Xu, F., Richardt, C., Yong, J.-H.: Parallax360: Stereoscopic 360° scene representation for head-motion parallax. IEEE Trans. Vis. Comput. Graph. 24(4), 1545–1553 (2018). https://doi.org/10.1109/TVCG.2018.2794071
52. Khan, A., Chefranov, A., Demirel, H.: Image scene geometry recognition using low-level features fusion at multi-layer deep CNN. Neurocomputing 440, 111–126 (2021). https://doi.org/10.1016/j.neucom.2021.01.085
53. Liu, L., Stamos, I., Yu, G., Wolberg, G., Zokai, S.: Multiview geometry for texture mapping 2D images onto 3D range data. In: *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, pp. 2293–2300, New York, USA (2006). https://doi.org/10.1109/CVPR.2006.204
54. Leica Geosystems. https://leica-geosystems.com/. Accessed: Jul. 28, 2023.
55. Zhao, W., Nister, D., Hsu, S.: Alignment of continuous video onto 3D point clouds. IEEE Trans. Pattern Anal. Mach. Intell. 27(8), 1305–1318 (2005). https://doi.org/10.1109/TPAMI.2005.152
56. Besl, P., McKay, H.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14, 239–256 (1992). https://doi.org/10.1109/34.121791
57. Choi, A.C.K., Chan, D.S.K., Yuen, A.M.F.: Application of virtual assembly tools for improving product design. Int. J. Adv. Manuf. Technol. 19(5), 377–383 (2002). https://doi.org/10.1007/s001700200027
58. Ainsworth, R.A., Sandin, D.J., Schulze, J.P., Prudhomme, A., DeFanti, T.A., Srinivasan, M.: Acquisition of stereo panoramas for display in VR environments. In: Proc. SPIE 7864 Three-Dimensional Imaging, Interaction, and Measurement, pp. 403–417. SPIE, Bellingham, WA (2011). https://doi.org/10.1117/12.872521
59. Zhang, Y., Xu, W., Tong, Y., Zhou, K.: Online structure analysis for real-time indoor scene reconstruction. ACM Trans. Graph. 34(5), 159:1–159:13 (2015). https://doi.org/10.1145/2768821
60. Nießner, M., Zollhöfer, M., Izadi, S., Stamminger, M.: Real-time 3D reconstruction at scale using voxel hashing. ACM Trans. Graph. 32(6), 169:1–169:11 (2013). https://doi.org/10.1145/2508363.2508374
61. Dai, A., Chang, A.X., Savva, M., Halber, M., Funkhouser, T., Niessner, M.: ScanNet: Richly-annotated 3D reconstructions of indoor scenes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu HI, USA, pp. 5828–5839. (2017)
62. Innmann, M., Zollhöfer, M., Nießner, M., Theobalt, C., Stamminger, M.: VolumeDeform: Real-time volumetric non-rigid reconstruction. In: Computer Vision—ECCV 2016. Lecture Notes in Computer Science. pp. 362–379. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-46484-8_22
63. Pichler, B.: HDR Light Field. Johannes Kepler University Linz. (2012), https://doi.org/10.13140/RG.2.2.32308.04484
64. Dayan, S.B., Mendlovic, D., Giryes, R.: Deep sparse light field refocusing. arXiv (2020). https://doi.org/10.48550/arXiv.2009.02582
65. Zhang, C., Hou, G., Zhang, Z., Sun, Z., Tan, T.: Efficient auto-refocusing for light field camera. Pattern Recognit. 81, 176–189 (2018). https://doi.org/10.1016/j.patcog.2018.03.020
66. Jayaweera, S.S., Edussooriya, C.U.S., Wijenayake, C., Agathoklis, P., Bruton, L.T.: Multi-volumetric refocusing of light fields. IEEE Signal Process. Lett. 28, 31–35 (2021). https://doi.org/10.1109/LSP.2020.3043990
67. Song, Z., Zhu, H., Wu, Q., Wang, X., Li, H., Wang, Q.: Accurate 3D reconstruction from circular light field using CNN-LSTM. In: *2020 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, (2020). https://doi.org/10.1109/ICME46284.2020.9102847
68. Perra, C., Murgia, F., Giusto, D.: An analysis of 3D point cloud reconstruction from light field images. In: 2016 IEEE Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1–6. Oulu, Finland (2016). https://doi.org/10.1109/IPTA.2016.7821011
69. Kalantari, N.K., Wang, T.-C., Ramamoorthi, R.: Learning-based view synthesis for light field cameras. ACM Trans. Graph. 35(6), 1–10 (2016). https://doi.org/10.1145/2980179.2980251
70. Gershun, A.: The light field. J. Math. Phys. 18(1–4), 51–151 (1939). https://doi.org/10.1002/sapm193918151
71. Adelson, E.H.: The plenoptic function and the elements of early vision. In: Computational Models of Visual Processing. Ch. 1. MIT Press (1991)
72. Levoy, M., Hanrahan, P.: Light field rendering. In: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pp. 31–42 (1996). https://doi.org/10.1145/237170.237199
73. Shum, H.-Y., Kang, S.B., Chan, S.: Survey of image-based representations and compression techniques. IEEE Trans. Circuits Syst. Video Technol. 13, 1020–1037 (2003). https://doi.org/10.1109/TCSVT.2003.817360
74. Wilburn, B., et al.: High performance imaging using large camera arrays. In: *ACM SIGGRAPH 2005 Papers*, in SIGGRAPH '05. pp. 765–776. New York, NY (2005). https://doi.org/10.1145/1186822.1073259
75. Zhang, C., Chen, T.: A self-reconfigurable camera array. In: ACM SIGGRAPH 2004 Sketches on—SIGGRAPH '04. p. 151. Los Angeles, CA (2004). https://doi.org/10.1145/1186223.1186412
76. Venkataraman, K., et al.: PiCam: An ultra-thin high performance monolithic camera array. ACM Trans. Graph. 32(6), 166:1–166:13 (2013). https://doi.org/10.1145/2508363.2508390
77. Raytrix 3D Light-Field Vision Plenoptic Metrology. https://raytrix.de/home-3-2-2/ Accessed 15 Sept 2023
78. Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., Hanrahan, P.: Light Field Photography with a Hand-held Plenoptic Camera. Stanford University, (2005)
79. Goldlücke, B., Klehm, O., Wanner, S., Eisemann, E.: Plenoptic cameras. In: Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality. pp. 65–77. CRC Press, Boca Raton, FL (2015)
80. Georgiev, T., Lumsdaine, A.: Resolution in plenoptic cameras. In: *Frontiers in Optics 2009/Laser Science XXV/Fall 2009 OSA Optics & Photonics Technical Digest 2009, paper CTuB3*. Optica Publishing Group, Washington, DC (2009). https://doi.org/10.1364/COSI.2009.CTuB3

81. Rerabek, M., Ebrahimi, T., (eds.) New light field image dataset. In: *8th International Conference on Quality of Multimedia Experience (QoMEX)*, Lisbon, Portugal (2016)
82. Ahmad, W., Palmieri, L., Koch, R., Sjostrom, M.: Matching light field datasets from plenoptic cameras 1.0 and 2.0. In: *2018 - 3DTV-Conference: The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1–4. Helsinki, Finland (2018). https://doi.org/10.1109/3DTV.2018.8478611
83. Debevec, P., Downing, G., Bolas, M., Peng, H.-Y., Urbach, J.: Spherical light field environment capture for virtual reality using a motorized pan/tilt head and offset camera. In: *ACM SIGGRAPH 2015 Posters*, Los Angeles (2015). https://doi.org/10.1145/2787626.2787648
84. Broxton, M., et al.: A low cost multi-camera array for panoramic light field video capture. In: *SIGGRAPH Asia 2019 Posters*, in SA '19, pp. 1–2. New York, NY (2019). https://doi.org/10.1145/3355056.3364593
85. Pertuz, S., Pulido-Herrera, E., Kamarainen, J.-K.: Focus model for metric depth estimation in standard plenoptic cameras. ISPRS J. Photogramm. Remote Sens. 144, 38–47 (2018). https://doi.org/10.1016/j.isprsjprs.2018.06.020
86. Schweiger, F., et al.: TOOLS FOR 6-DOF IMMERSIVE AUDIOVISUAL CONTENT CAPTURE AND PRODUCTION. *IBC*, BBC R&D White Paper, (2021). https://www.bbc.co.uk/rd/publications/tools-six-depth-offield-immersive-audiovisual-content-capture-production (2022)
87. Light field backgrounds for virtual reality and virtual production. BBC R&D. https://www.bbc.co.uk/rd/blog/2022-01-light-field-backgrounds-virtual-reality-production Accessed 15 Sept 2023
88. Farhood, H., Perry, S., Cheng, E., Kim, J.: 3D point cloud reconstruction from a single 4D light field image. In: Presented at the Optics, Photonics and Digital Technologies for Imaging Applications VI. 11353 (2020). https://doi.org/10.1117/12.2555292
89. Wirth, F., Quehl, J., Ota, J., Stiller, C.: PointAtMe: Efficient 3D point cloud labeling in virtual reality. In: *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1693–1698. Paris, France (2019). https://doi.org/10.1109/IVS.2019.8814115
90. Kharroubi, A., Hajji, R., Billen, R., Poux, F.: Classification and integration of massive 3D points clouds in a virtual reality (VR) environment. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Copernicus, Goettingen, Germany (2019). https://doi.org/10.5194/isprs-archives-XLII-2-W17-165-2019
91. Stets, J.D., Sun, Y., Corning, W., Greenwald, S.W.: Visualization and labeling of point clouds in virtual reality. In: *SIGGRAPH Asia 2017 Posters*, in SA '17, pp. 1–2. New York, NY (2017). https://doi.org/10.1145/3145690.3145729
92. Murgia, F., Perra, C., Giusto, D.: 3D point cloud reconstruction from single plenoptic image. Telfor J. 8(1), 26–31 (2016). https://doi.org/10.5937/telfor1601026M
93. Farhood, H., Perry, S., Cheng, E., Kim, J.: Enhanced 3D point cloud from a light field image. Remote Sens. 12(7), 1125 (2020). https://doi.org/10.3390/rs12071125
94. Conti, C., Soares, L.D., Nunes, P.: Light field coding with field-of-view scalability and exemplar-based interlayer prediction. IEEE Trans. Multimedia 20(11), 2905–2920 (2018). https://doi.org/10.1109/TMM.2018.2825882
95. Liu, D., An, P., Ma, R., Yang, C., Shen, L.: 3D holoscopic image coding scheme using HEVC with Gaussian process regression. Signal Process. Image Commun. 47, 438–451 (2016). https://doi.org/10.1016/j.image.2016.08.004
96. Dai, F., Zhang, J., Ma, Y., Zhang, Y.: Lenselet image compression scheme based on subaperture images streaming. In: *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 4733–4737. QC, Canada (2015). https://doi.org/10.1109/ICIP.2015.7351705
97. Miandji, E., Hajisharif, S., Unger, J.: A unified framework for compression and compressed sensing of light fields and light field videos. ACM Trans. Graph. 38(3), 23:1–23:18 (2019). https://doi.org/10.1145/3269980
98. Jia, C., Zhang, X., Wang, S., Wang, S., Ma, S.: Light field image compression using generative adversarial network-based view synthesis. IEEE J. Emerging Sel. Top. Circuits Syst. 9(1), 177–189 (2019). https://doi.org/10.1109/JETCAS.2018.2886642
99. Wang, J., Wang, Q., Xiong, R., Zhu, Q., Yin, B.: Light field image compression using multi-branch spatial transformer networks based view synthesis. In: *2020 Data Compression Conference (DCC)*, pp. 397–397. Snowbird, UT, USA (2020). https://doi.org/10.1109/DCC47342.2020.00047
100. Georgeiv, T., Zheng, K.C., Curless, B., Salesin, D., Nayar, S., Intwala, C.: Spatio-Angular Resolution Tradeoff in Integral Photography. Rendering Techniques 2006. 263–272 (2006): 21.
101. Vieira, A., Duarte, H., Perra, C., Tavora, L., Assuncao, P.: Data formats for high efficiency coding of Lytro-Illum light fields. In: *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pp. 494–497. Orleans, France (2015). https://doi.org/10.1109/IPTA.2015.7367195
102. Wang, J., Shi, R., Zheng, W., Xie, W., Kao, D., Liang, H.-N.: Effect of frame rate on user experience, performance, and simulator sickness in virtual reality. IEEE Trans. Vis. Comput. Graph. 29(5), 2478–2488 (2023). https://doi.org/10.1109/TVCG.2023.3247057
103. Zielinski, D.J., Rao, H.M., Sommer, M.A., Kopper, R.: Exploring the effects of image persistence in low frame rate virtual environments. In: *2015 IEEE Virtual Reality (VR)*, pp. 19–26. Arles, Camargue, Provence, France (2015). https://doi.org/10.1109/VR.2015.7223319
104. Khan, A., Hossain, M.M., Sirlantzis, K., Covaci, A., Chowdhury, W.S.: Depth estimation and validation of plenoptic light field camera. In: *2023 IEEE International Conference on Imaging Systems and Techniques (IST)*. Copenhagen, Denmark, pp. 1–6 (2023). https://doi.org/10.1109/IST59124.2023.10355697
105. Bok, Y., Jeon, H.-G., Kweon, I.S.: Geometric calibration of micro-lens-based light field cameras using line features. IEEE Trans. Pattern Anal. Mach. Intell. 39(2), 287–300 (2017). https://doi.org/10.1109/TPAMI.2016.2541145
106. Noury, C.-A., Teuliere, C., Dhome, M.: Light-field camera calibration from raw images. In: *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. Sydney, NSW, pp. 1–8 (2017). https://doi.org/10.1109/DICTA.2017.8227459
107. Sun, J., et al.: Geometric calibration of focused light field camera for 3-D flame temperature measurement. In: *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, pp. 1–6. Taipei, Taiwan (2016). https://doi.org/10.1109/I2MTC.2016.7520580
108. Sun, J., Hossain, M.M., Xu, C.-L., Zhang, B., Wang, S.-M.: A novel calibration method of focused light field camera for 3-D reconstruction of flame temperature. Opt. Commun. 390, 7–15 (2017). https://doi.org/10.1016/j.optcom.2016.12.056
109. Cho, D., Lee, M., Kim, S., Tai, Y.-W.: Modeling the calibration pipeline of the lytro camera for high quality light-field image reconstruction. In: *2013 IEEE International Conference on Computer Vision*, pp. 3280–3287. Sydney, Australia (2013). https://doi.org/10.1109/ICCV.2013.407
110. Bakir, N., Hamidouche, W., Deforges, O., Samrouth, K., Khalil, M.: Light field image compression based on convolutional neural networks and linear approximation. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 1128–1132. Athens (2018). https://doi.org/10.1109/ICIP.2018.8451597
111. Sakamoto, T., Kodama, K., Hamamoto, T.: A study on efficient compression of multi-focus images for dense Light-Field reconstruction. In: *2012 Visual Communications and Image Processing*, pp. 1–6. San Diego, CA (2012). https://doi.org/10.1109/VCIP.2012.6410759
112. Wang, Y., Liu, Y., Heidrich, W., Dai, Q.: The light field attachment: Turning a DSLR into a light field camera using a low budget camera ring. IEEE Trans. Vis. Comput. Graph. 23(10), 2357–2364 (2017). https://doi.org/10.1109/TVCG.2016.2628743
113. Boominathan, V., Mitra, K., Veeraraghavan, A.: Improving resolution and depth-of-field of light field cameras using a hybrid imaging system. In: *2014 IEEE International Conference on Computational Photography (ICCP)*, pp. 1–10. Santa Clara, CA (2014). https://doi.org/10.1109/ICCPHOT.2014.6831814
114. Yang, F., Yang, H., Fu, J., Lu, H., Guo, B.: Learning Texture Transformer Network for Image Super-Resolution. In: *2020 IEEE/CVF Conference on*

Computer Vision and Pattern Recognition (CVPR), pp. 5790–5799.Seattle, WA (2020). https://doi.org/10.1109/CVPR42600.2020.00583
115. Mitra, K., Veeraraghavan, A.: Light field denoising, light field superresolution and stereo camera based refocussing using a GMM light field patch prior. In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 22–28. Providence, RI (2012). https://doi.org/10.1109/CVPRW.2012.6239346
116. Wanner, S., Goldluecke, B.: Variational light field analysis for disparity estimation and super-resolution. IEEE Trans. Pattern Anal. Mach. Intell. 36(3), 606–619 (2014). https://doi.org/10.1109/TPAMI.2013.147
117. Huang, Z., Fessler, J.A., Norris, T.B., Chun, I.Y.: Light-field reconstruction and depth estimation from focal stack images using convolutional neural networks. In: *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8648–8652. Barcelona, Spain (2020). https://doi.org/10.1109/ICASSP40776.2020.9053586

How to cite this article: Khan, A., Hossain, M.M., Covaci, A., Sirlantzis, K., Xu, C.: Light field imaging technology for virtual reality content creation: A review. IET Image Process. 1–21 (2024). https://doi.org/10.1049/ipr2.13144