RESEARCH ARTICLE

Enabling Five-Degree-of-Freedom From Low-Dimensional Light Fields With Extended Light Ray Acquisition Field-of-View

HYUNMIN JUNG[1], **HYUK-JAE LEE**[2], **(Member, IEEE),**
AND CHAE EUN RHEE[3], **(Senior Member, IEEE)**

[1]Department of Smart ICT Convergence Engineering, Seoul National University of Science and Technology, Seoul 01811, South Korea
[2]Inter-University Semiconductor Research Center, Department of Electrical and Computer Engineering, Seoul National University, Seoul 08226, South Korea
[3]Department of Information and Communication Engineering, Inha University, Incheon 22212, South Korea

Corresponding author: Chae Eun Rhee (chae.rhee@inha.ac.kr)

ABSTRACT Conventional captured-image-based virtual reality (VR) systems hardly support free movement of the viewpoint, greatly reducing the user's sense of reality. A light field (LF), which generates a novel view at a free viewpoint through a combination of light rays, is a suitable approach for increasing the degree of freedom (DoF) in captured-image-based VR. Previous LF-based VR systems have increased LF dimensions or used various acquisition layouts to obtain as many light rays as possible and thereby expand the renderable view range. However, these attempts are costly and time-consuming. This paper enables 5-DoF outcomes from a low-dimensional LF by increasing the field-of-view during light ray acquisition. The method starts with the simple idea that if a three-dimensional (3D) LF is constructed with 360° images that can capture all directions, the renderable range can be widened in a cost-efficient manner. However, using light rays with a large incident angle in a 3D LF can cause visual artifacts. To cope with this problem, this paper initially defines and analyzes the visual artifacts. While these errors were simply ignored in the past, they become particularly noticeable when using light rays with large incident angles. Thus, a careful analysis of the errors and their impacts is required. Additionally, this paper proposes two schemes to deal with visual artifacts based on the error analysis: a horizontal LF and vertical depth (HLF-VD) view generation scheme that complements the structural limitation of 3D LFs and can address all types of errors, and a hybrid 3D LF connection scheme that effectively reduces connection errors by controlling the connection position of neighboring 3D LFs. The proposed system is implemented and verified in various real-world scenarios, and the experimental results show that the user's movement is naturally reflected at a level similar to that of high-dimensional LF, with minimal visual artifacts.

INDEX TERMS Free viewpoint, image-based rendering, light field, view navigation, virtual reality.

I. INTRODUCTION

A few years ago, virtual reality (VR) technology gained significant attention. However, problems such as insufficient technical maturity, lack of VR content, and motion sickness drastically reduced interest in VR. Nevertheless, with the onset of the COVID-19 pandemic, VR technology is experiencing a resurgence. The pandemic has changed the way people interact, and to avoid direct contact, telecommuting, video conferencing, and video education have become increasingly popular. VR technology can provide an all-encompassing solution in the COVID-19 era as it allows people to work,

The associate editor coordinating the review of this manuscript and approving it for publication was Michele Nappi.

meet, and play in a virtual space separated from the real world. However, the level of VR technology is still not as advanced as people expect, and there are numerous challenges that need to be addressed.

Computer graphic-based virtual reality (CG-based VR) creates virtual spaces artificially, allowing for free viewpoint movements using the three-dimensional (3D) geometry of the virtual space. Captured-image-based VR, on the other hand, has limitations on the user's viewpoint movements. Systems such as Samsung Gear 360 [1] and Facebook Surround360 [2] are examples of 360° image-based VR systems where the user's viewpoint is fixed, and only rotational view direction changes of the roll, yaw, and pitch are supported. This greatly reduces the user's sense of reality since they cannot move within the virtual space. Despite this limitation, captured-image-based VR has an advantage in creating a more natural and photorealistic view than CG-based VR. To address the issue of limited movement in captured-image-based VR, various research groups have attempted to provide free movement.

3D modeling and image-based-rendering (IBR) are traditional 3D image processing methods, and both have generally been applied to the free viewpoint systems. 3D modeling estimates the 3D geometry of a target space [3], [4], [5], which works very similarly to CG-based VR. IBR synthesizes a view at a free viewpoint based on the disparity or optical flow between images [6], [7], [8]. However, the computational complexity is quite high, and it is still difficult to guarantee robust results.

Light field (LF) [9], [10], [11] presents a novel approach where all the light rays passing through free space are captured, and some of them are recombined to generate a view. This method does not require geometry estimations, and it creates accurate and robust rendering results with low computational complexity [12], [13]. The main challenge in LF is acquiring all the light rays passing through free space. In LF, pixels in the captured image replace light rays, and only light rays passing through a specific structure such as a plane or a sphere are acquired in a limited manner. Google introduced a LF-based VR system that assumes a spherical structure [14], [15]. They developed a new shooting system to capture light rays passing through it. The user's viewpoint can move within the sphere but cannot go outside it. The recently introduced LF Unit (LFU) [16] proposes a way to expand the movement range. It constructs an LF structure for a large space by connecting and stacking small-sized 3D LFs. 3D LF, which is acquired from a line instead of a plane, is more convenient to capture light rays. However, 3D LF considers horizontal parallax only (HPO), and the available light rays are limited compared to four-dimensional (4D) LF. LFU does not provide top and bottom views due to its low vertical field of view (FOV), and it cannot create a full 360° view, which limits rotational view changes and provides only four degrees of freedom (DoF) movement.

As such, the previous approaches require significant resources and complex layouts to increase the dimension of LF and provide users with a higher degree of freedom (DoF). This is a major factor that causes a high cost in constructing LF-based systems, and has a negative impact on popularization. This paper proposes a simple approach for FOV extension, which makes it possible to support a high level of DoF in a low-dimensional LF. The contributions of this paper are as follows. First, it proposes a system that supports high 5-DoF in a relatively simple 3D LF structure. To achieve this, the FOV is expanded to acquire light rays, and a 360° camera is used instead of a normal camera. This is a simple change, but the naïve use of light rays with a large incident angle can cause various visual artifacts, which are addressed by the following contributions.

Second, this paper examines the weak points and possible errors in 3D LF that were simply ignored in previous works because artifacts are acceptable in a system with a small FOV. However, the proposed system requires light rays with large incident angles, and therefore, careful error analysis is necessary to create a high-quality view. Third, based on the in-depth analysis, the paper determines the optimal configuration to reduce the magnitude of error. The horizontal LF-based and vertical depth-based (HLF-VD) view generation scheme is used to compensate for the HPO of 3D LF and works for all types of visual artifacts. The hybrid 3D LF connection scheme plays an important role, especially when connecting multiple 3D LFs. The experimental results show that the system can render a full 360° view naturally according to the free viewpoint movement and that the tailored solutions alleviate visual artifacts.

The remainder of this paper is organized as follows. Section II introduces the background and motivation. Section III defines three types of visual artifacts and analyzes them. Section IV proposes the 3D LF system that takes into account the lack of light rays. In Section V, the experimental results are presented. The paper concludes in Section VI.

II. BACKGROUND AND MOTIVATION
A. 3D/4D LIGHT FIELD
LF refers to the concept of defining light rays and the relationships between them [9], [10], [11], [18]. The plenoptic function describes light rays using seven variables: a 3D point through which a light ray passes (x, y, z), the horizontal and vertical angles (θ, ϕ), the wavelength (λ), and the time (t) [9]. However, the representation of light rays has been simplified, and the 4D LF representation, which uses four variables, is the most widely used. In this representation, the time is fixed to a specific time point, and the wavelength is removed by considering only the recorded R, G, and B values of each light ray. In addition, the z coordinate of a 3D point is fixed by assuming a plane [11]. The 4D LF is defined as the light rays passing through a point (x, y) and the incident angles of (θ, ϕ) [19], [20]. Another type of 4D LF uses two points,
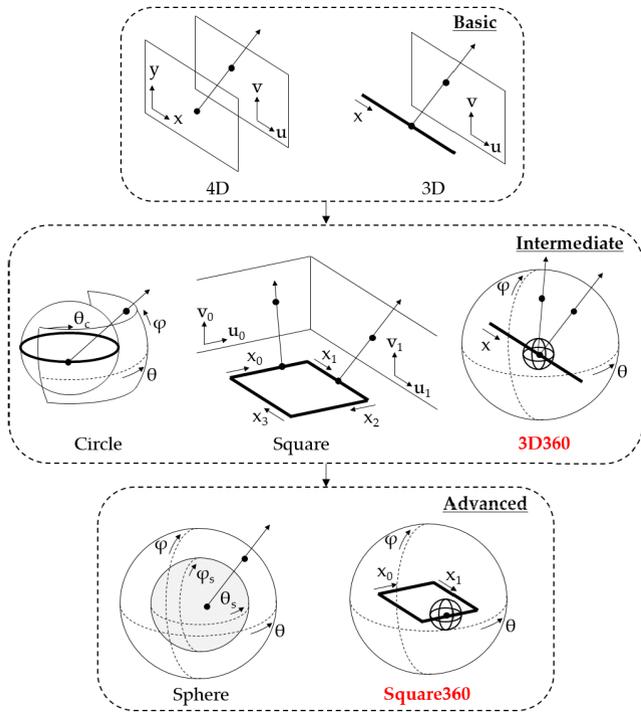
FIGURE 1. Previous LF systems and the difference of the proposed system.

FIGURE 2. The overview of the conventional stackable 3D LF system [16].

FIGURE 3. Two 3D LF connection methods included in the conventional stackable 3D LF. (a) Physical connection. (b) Non-physical connection.

(x, y) and (u, v), to represent a light ray passing through two planes [10], [11]. These two types of 4D LF representations are compatible with each other. In actual implementations, a 4D LF is constructed by capturing images from cameras that are closely placed on a plane [17], [20]. The pixels of the captured images correspond to the light rays.

Compared to 4D LF, 3D LF consists of light rays passing through a line instead of a plane, and the variable v is additionally fixed [20], [21]. In a 3D LF, a light ray is represented by a point x on the line and the horizontal and vertical incident angles (θ, ϕ) passing through it. Due to the simple structure of 3D LF, the difficulty of acquiring light rays is drastically reduced. A one-dimensional (1D) camera array is considered instead of a two-dimensional (2D) one, where a slider or dolly moving along a line is used instead of a plane-shaped camera structure or customized equipment moving on a plane. On the other hand, 3D LF considers HPO and incurs the problem of insufficient light rays in the vertical direction. 3D/4D LF is the most basic form and is used in various applications, such as depth estimations [24], [27], [39], super-resolution [23], [25], [26], [38], and denoising [28], [29].

B. LIGHT FIELD-BASED FREE VIEWPOINT SYSTEMS
Fig. 1 illustrates the different types of LF structures. The most basic forms are the 3D and 4D LFs. At the intermediate level, both circular [30], [31] and square-shaped structures [16], [32] are modified versions of the 3D LF. Due to the layout changes, light rays from four directions can be acquired, expanding the range of view rendering. However, the top and bottom views are still not generated. At the advanced level, the spherical structure is a reconstructed form of the 4D LF. It can generate views in all directions, including the top and bottom, by capturing light rays coming into the spherical surfaces [19]. However, the spherical structure usually requires specially designed devices, making the acquisition of light rays time-consuming and expensive [14], [15].

The 3D360 and Square360 structures are newly presented in this paper and use a 360° camera with the maximum FOV. 3D360 is an extended version of the basic-level 3D LF, while Square360 is an extended version of a square-shaped structure at the intermediate level. The Square360 structure supports 5-DoF outcomes close to those of a spherical structure at the advanced level but with an intermediate-level acquisition cost.

C. STACKABLE SQUARE-SHAPED SYSTEM
The approach in this paper is based on the square-shaped structure. LFU [16] introduced a stackable square-shaped structure, which dramatically expands the rendering range by stacking multiple individual squares. This section briefly reviews the stackable square-shaped system, as shown in Fig. 2. ① The bold lines on the x-z plane are 1D camera lines for composing 3D LFs. The 3D LFs are arranged in a grid, and a viewpoint moves freely on the plane. ② When a viewpoint is given, four 3D LFs surrounding the viewpoint marked in green color are selected. ③ To create a 360° view, the

rendering range to be covered by each 3D LF is assigned and pixel positions for the connection are determined. ④ Through 3D LF-based view rendering, four views are rendered. ⑤ The final 360° view is created by connecting them. This conventional system acquires rays through a camera with a narrow horizontal and vertical FOV, which cannot fill the top and bottom parts of a 360° view.

Fig. 3 presents a brief overview of the two connection methods used in step ④ of Fig. 2. Fig. 3 (a) shows the physical connection where the two 3D LFs are connected at their physical intersection. The light ray, indicated by the red dotted arrow, passes through the physical intersection and belongs to both 3D LFs, which is defined as a shared light ray. Based on this connection point, the left view is generated from the left 3D LF and the front view is generated from the front 3D LF. Fig. 3 (b) shows the non-physical connection, which takes advantage of the fact that the two 3D LFs share light rays even though they do not physically intersect. The red dotted arrow represents the shared light ray, which also belongs to both 3D LFs. The advantage of a non-physical connection is that it can secure a sufficient amount of shared light rays even when the horizontal FOV is small.

III. ANALYSIS OF VISUAL ARTIFACTS IN 3D LF VIEW RENDERING

This paper aimed to extend the field of view (FOV) to create a full 360° image, including top and bottom views. However, using a 360° camera with an extended FOV in a naïve attempt results in unexpected errors. Fig. 4 illustrates the commonly used vertical cross-section based view rendering in a 3D LF [20], [22]. Fig. 4 (a) shows a 3D LF constructed by stacking images captured along the 1D camera line in Fig. 4 (b). When a free viewpoint is selected, as shown in Fig. 4 (b), the view range is determined according to the rendering range, θ_{FOV}. Here, the light rays at the end of both sides of the view range are (x_0, θ_0) and (x_1, θ_1), respectively. The vertical cross-section in Fig. 4 (a) corresponds to the view at the selected viewpoint, and the pixel row with $\phi = \phi_0$ in Fig. 4 (b) is rendered through the θ-x cross-section with $\phi = \phi_0$ in Fig. 4 (c). Light rays belonging to the θ-x cross-section are acquired from the 1D camera line, such as L_b in Fig. 4 (d), while the actual light ray with $\phi = \phi_0$ for the viewpoint is L_a, which cannot be acquired from the 1D camera line. These two are clearly different. The 3D LF cannot acquire light rays that do not pass through the 1D camera line, like L_a. To cope with this, view rendering based on the vertical cross-section simplifies the combination of light rays by assuming that the non-acquired light rays are replaceable with similar ones obtained from the 1D camera line. The difference between the actual and replaced light rays, $\Delta\phi$, is defined as (1).

$$\Delta\phi = d_p \sin\phi \tag{1}$$

In (1), ϕ represents the vertical incident angle, and d_p is the distance between the viewpoint and the light ray acquisition point on the 1D camera line. As ϕ and d_p increase, $\Delta\phi$ also increases. When ϕ and d_p are small, the main assumption of
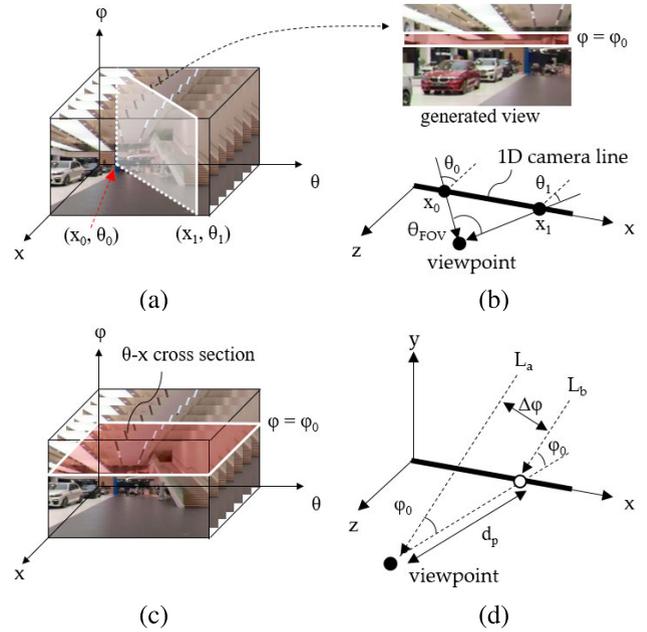
FIGURE 4. Examples of view generation in a 3D LF. (a) A constructed 3D LF and vertical cut for view generation. (b) Selected viewpoint and the generated view. (c) θ-x cross section with $\phi = \phi_0$ in the 3D LF. (d) Difference between a light ray into the 1D camera line and a light ray into the viewpoint.
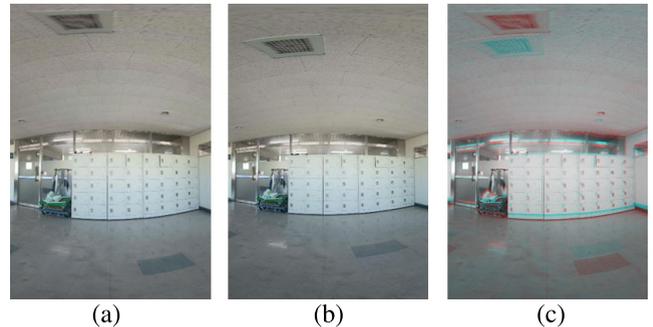
FIGURE 5. Examples of vertical parallax error. (a) A generated view from a 3D LF at a free viewpoint. (b) Actually captured view at the same viewpoint. (c) Overlay image of (a) and (b).

view rendering based on the vertical cross-section remains valid since $\Delta\phi$ is small. However, the proposed system aims to generate a full 360° view, which involves considering very large vertical angles of up to $\pm 90°$. This approach maximizes $\Delta\phi$ and results in three types of visual artifacts, which are described below.

A. VERTICAL PARALLAX ERROR
By using the replaced light rays with an error of $\Delta\phi$, the newly generated view is different from the actual view and does not reflect the vertical parallax. Figs. 5 (a), (b), and (c) show the view generated by the vertical cross-section, the actually captured view, and an overlay image of these two, respectively. The generated view in Fig. 5 (a) is photo-realistic, and visual errors are not noticeable. However,
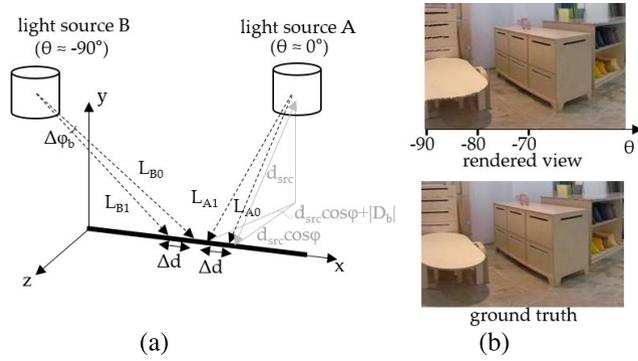
FIGURE 6. Examples of a bending error. (a) Light ray relationship in a 3D LF that causes bending error. (b) Visual artifact of bending error.

FIGURE 7. Examples of mismatching error. (a) Light ray relationship that cause mismatching error. (b) Generated view by non-physical connection and enlarged image.

Fig. 5 (c) clearly shows the difference between Figs. 5 (a) and (b). The horizontal positions of the objects are similar, but the vertical positions are not. The objects and the background appear more centered in the actual view. This tendency is evident at the top and bottom of the image because large ϕ values at the top and bottom increase $\Delta\phi$.

B. BENDING ERROR

A bending error also occurs due to view generation based on the vertical cross-section. In particular, it is caused by the difference in ϕ between light rays incident on neighboring acquisition points on the same 1D camera line. Fig. 6 (a) shows the 1D camera line, object A in front, and object B on the side. The light rays L_{A0} and L_{A1} from object A enter two adjacent points on the given camera line, whereas L_{B0} and L_{B1} enter two adjacent points from object B. The θ values of the light rays from objects A and B are close to 0° and -90°, respectively. Equations (2) and (3) define the difference in ϕ between the two neighboring light rays from the same source.

$$\Delta\phi_b = tan^{-1}\left(\frac{d_{src}\sin\phi}{d_{src}\cos\phi}\right) - tan^{-1}\left(\frac{d_{src}\sin\phi}{d_{src}\cos\phi + |D_b|}\right) \tag{2}$$

$$D_b = \sqrt{(d_{src}\cos\phi\cos\theta)^2 + (d_{src}\cos\phi\sin\theta + \Delta d)^2} - d_{src}\cos\phi \tag{3}$$

In (2), $\Delta\phi_b$ represents the difference in ϕ due to bending error as shown in Fig. 6 (a). d_{src} denotes the distance between the light source and one of the neighboring acquisition points. The two tan^{-1} terms in (2) represent the ϕ of two neighboring light rays. $d_{src}\cos\phi$ and $d_{src}\cos\phi + |D_b|$ represent the distance between the position of the light source on the x-z plane and the two neighboring acquisition points, respectively. $|D_b|$ represents the difference between the two distances and is defined by (3). In (3), given Δd as shown in Fig. 6 (a), D_b is dependent on θ. If θ is close to 0°, such as L_{A0} and L_{A1}, D_b is close to 0, and thus, $\Delta\phi_b$ is also close to 0°. On the other hand, if θ is close to $\pm90°$, as in L_{B0} and L_{B1}, D_b becomes Δd, which is the maximum value, and $\Delta\phi_b$ is also maximized. The increased $\Delta\phi_b$ causes visual artifacts, as shown in Fig. 6 (b). In the area where θ is small, there is no error in the
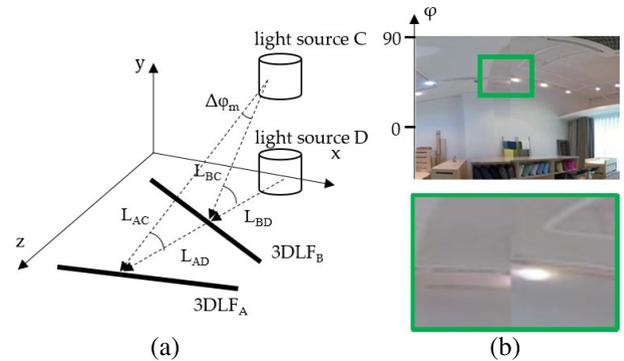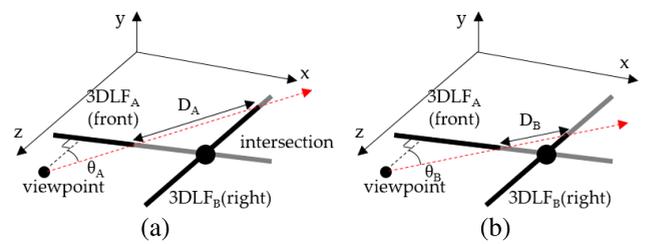
FIGURE 8. Relationship between D_m and θ in non-physical connection. (a) Case of small θ and large D_m. (b) Case of large θ and small D_m.

rendering result. However, the table with a close to -90° angle appears to be bent compared to the ground truth. Bending errors become noticeable when using light rays with large ϕ. ϕ depends on the position of the 3D LF connection, and the physical connection that requires a large ϕ is a major cause of the bending error.

C. MISMATCHING ERROR

Mismatching errors occur due to non-physical connections in the stackable 3D LF. A non-physical connection assumes that two non-intersecting 3D LFs share light rays. However, this assumption is only valid for light rays for which $\phi=0°$, strictly speaking. In Fig. 7 (a), the two bold lines represent two 3D LFs, $3DLF_A$ and $3DLF_B$. The two light rays, L_{AD} and L_{BD}, are incident from light source D coplanar with the two 3D LFs and are identical. On the other hand, the light rays, L_{AC} and L_{BC}, which are from light source C higher than light source D, have different ϕ and are thus distinctly different light rays. If the light source is close to the x-z plane and ϕ is small, the assumption that two 3D LFs share a sufficiently large number of light rays is valid. However, as ϕ increases, this assumption is violated, causing mismatch errors, as shown in Fig. 7 (b). The left half view consists of $3DLF_A$, while the right half view consists of $3DLF_B$. A discontinuity due to the difference in ϕ is clearly visible.

The mismatching error is also influenced by the position of the 3D LFs through which the shared light ray passes in the non-physical connection. Figs. 8 (a) and (b) show different connection positions. In both cases, the front and right views

TABLE 1. The relationship between visual artifacts and ϕ and θ.

Error	Vertical parallax	Bending	Mismatching
Lack of light ray in 3D LF	Noticeable with large ϕ		
3D LF connection	-	Noticeable with large θ	Noticeable with small θ

are created from $3DLF_A$ and $3DLF_B$, respectively, as indicated in bold black. The red dotted lines represent shared light rays and determine the connection position. Depending on this, the distance between the two acquisition points varies, which affects $\Delta\phi_m$, as defined by (4).

$$\Delta\phi_m = tan^{-1}\left(\frac{d_{src}\sin\phi}{d_{src}\cos\phi}\right) - tan^{-1}\left(\frac{d_{src}\sin\phi}{d_{src}\cos\phi + |D_m|}\right) \tag{4}$$

In (4), $\Delta\phi_m$ represents the difference in ϕ due to the mismatching error shown in Fig. 7. The equation is identical to (2) except for D_m, which refers to the distance between two acquisition points on different 3D LFs connected by a shared light ray. In the cases shown in Figs. 8 (a) and (b), D_m is D_A and D_B, respectively. Like (2), two tan^{-1} terms represent the ϕ of two light rays. The larger D_m is, the larger $\Delta\phi_m$ becomes, which causes the mismatching error. It is important to consider the θ value of the shared light ray. In Fig. 8 (a), the shared light ray with θ_A is used in $3DLF_A$. In contrast, the shared light in Fig. 8 (b) has a larger angle of θ_B than θ_A. If this θ becomes larger and the shared light passes through the intersection of two 3D LFs, D_m becomes 0, and there is no mismatching error. This case corresponds to the physical connection of Fig. 3 (a). However, as mentioned earlier, the use of light rays with larger θ increases the bending error.

Table 1 summarizes the relationship between the visual artifacts and ϕ and θ. All three errors defined above are caused by a lack of light rays in the 3D LF. Additionally, all three errors occur in light rays with large ϕ values. Bending and mismatching errors are especially observed in the 3D LF connection and are also affected by the θ of the light ray. Bending errors occur in light rays with large θ values, and mismatching errors occur in light rays with small θ values. Therefore, it is necessary to select an appropriate θ to effectively reduce both errors.

IV. PROPOSED LF SYSTEM WITH A LACK OF VERTICAL LIGHT RAYS

Fig. 9 depicts an overview of the proposed LF system that lacks vertical light rays. The system assumes a stackable square-shaped structure as shown in Fig. 2. To obtain light rays with the maximum FOV, a 360° camera is utilized, and 3D LFs are constructed from 360° images. In this structure, the following steps are performed: ① Given a viewpoint, four 3D LFs surrounding the viewpoint are selected. ② Next, the Hybrid 3D LF connection scheme determines the connection position and allocates the FOV range to be covered by
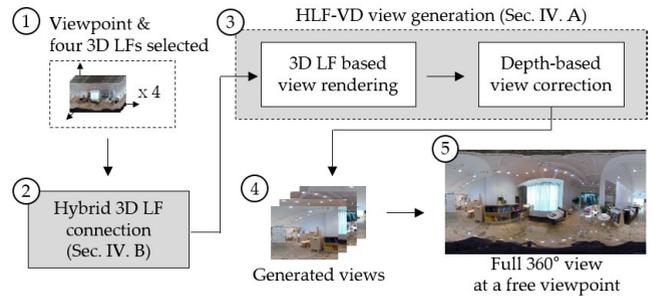
FIGURE 9. The overview of the proposed 3D LF system conscious of lack of vertical light rays.
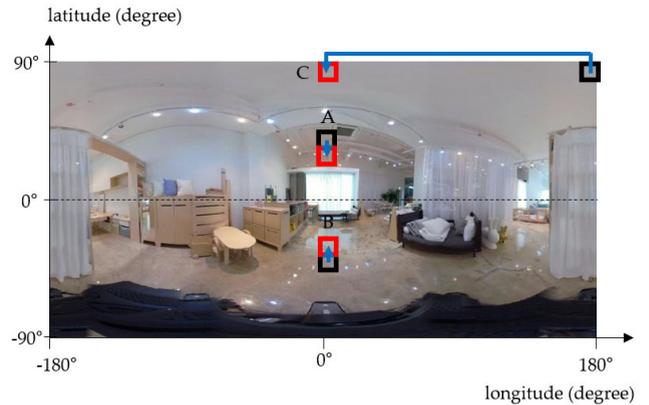
FIGURE 10. Examples of HLF-VD based view generation. Position of three pixels A, B, and C are adjusted through depth information. Black and red squares are the initial positions and the adjusted positions of the pixels, respectively.

each 3D LF out of 360°. In a Hybrid 3D LF connection, physical and non-physical connections are selectively used to avoid rapid increases in bending and mismatching errors, as explained in detail in Sec. IV-B. ③ Four 3D LFs render four views, where each view is initially created based only on the vertical cross-section of the 3D LF, and depth-based view correction is additionally applied to complement the lack of vertical light rays. This is referred to as HLF-VD view generation, which will be covered in detail in Sec. IV. A. ④ By connecting these four views, ⑤ the 360° view at the given viewpoint is created. Unlike the conventional system in Fig. 2, the proposed system generates a full 360° view and allows for 5-DoF movement.

A. HLF-VD VIEW GENERATION

HLF-VD view generation is proposed as a solution to overcome the lack of vertical light rays, and it effectively reduces all three types of errors. The HLF-VD view generation process involves creating an initial view at a free viewpoint through view rendering based on the vertical cross-section. Subsequently, depth-based correction is applied to reduce $\Delta\phi$ in (1). The relationship between ϕ on the 1D camera line and ϕ^* corrected according to the given viewpoint is defined
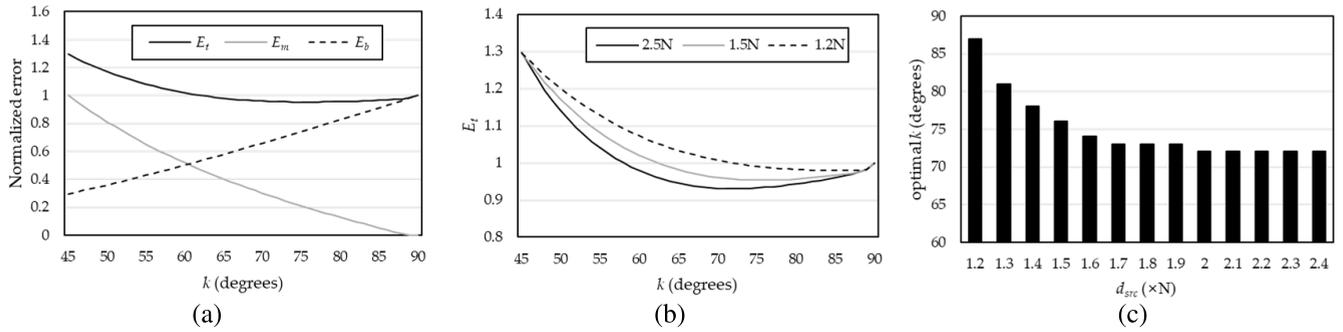
FIGURE 11. Graphs of mathematically defined errors depending on k and d_{src}. (a) E_t, E_m, and E_b when d_{src}=1.5N. (b) E_t according to k and d_{src}. (c) optimal k according to d_{src}.

in (5). In (5), the relationship between ϕ and ϕ^* is estimated from d_{src} when a viewpoint and d_p are given. The vertical position of the pixel in the initial view is adjusted based on the difference between ϕ and ϕ^*, while the horizontal position remains unchanged.

$$\phi^* = \tan^{-1} \frac{d_{src} \sin \phi}{d_{src} \cos \phi + d_p} \qquad (5)$$

Fig. 10 illustrates examples of pixel position correction. Three example pixels, A, B, and C, are shown, with their initial positions represented by black squares and their corrected positions represented by red squares. In the proposed system, the viewpoint is always behind the acquisition point, and d_p in (5) is always positive. Consequently, the absolute value of ϕ^* is always smaller than the absolute value of ϕ, and the pixel positions are adjusted towards the center direction, such as in the case of pixels A and B. For pixels with ϕ close to $\pm 90°$, they are filled with pixels located on the opposite side of the 360° image, as in the case of pixel C. This correction is applied to all pixels in the image.

In the actual implementation, pixel correction is performed based on an approximate depth for only a few steps. There are two reasons for this. First, there is a concern that incomplete depth information can compromise the photorealistic quality of the initial view. The proposed system prioritizes the creation of a photorealistic view that accurately reflects the user's movement, even if there is a slight deviation from the actual view. Second, among the errors that could not be eliminated, bending and mismatching errors can be mitigated through the Hybrid 3D LF connection presented in the next section.

B. HYBRID 3D LF CONNECTION
The proposed Hybrid 3D LF connection scheme is designed to reduce bending and mismatch errors. To minimize both types of errors, this study selectively uses physical and non-physical connections according to the viewpoint. In the Hybrid 3D LF connection, a physical connection is used by default to avoid mismatching errors. Only when light rays with θ greater than $\pm k°$ are necessary, the non-physical connection scheme is adopted using shared light rays with

$\theta = \pm k°$. In other words, this new scheme only allows light rays with $\theta \leq \pm k°$. It minimizes visual artifacts by preventing situations where bending and mismatching errors become excessively large. Mismatching and bending errors are mathematically defined in (6) to (10), and the optimal k is determined. Here, E_b and E_m refer to the bending error and mismatching error according to k and d_{src}, respectively, as defined in (7) and (9). E_t denotes the total error, which is the sum of the normalized values of the two types of errors, as defined in (10).

$$E_t(k, d_{src}) = \frac{E_m(k, d_{src})}{\max(E_m)} + \frac{E_b(k)}{\max(E_b)} \qquad (6)$$

E_m is defined in (7), where x and z refer to the position of the viewpoint on the x-z plane. $\Delta\phi_m$, defined in (4), corresponds to the mismatching in the pixel row where two views converge, as shown in Fig. 7 (b). j denotes the index of the pixel row in the generated view. The mismatching error depends on the viewpoint. Thus, the sum of $|\Delta\phi_m|$ occurring in all viewpoints and all connection pixel rows is defined as E_m.

$$E_m(k, d_{src}) = \sum_{x,z} \sum_{j=-90°}^{90°} |\Delta\phi_m| \qquad (7)$$

D_m used in the definition of $\Delta\phi_m$ in (4) depends on the form of the 3D LF. In (8) and (9), the stackable 3D LF structure shown in Fig. 2 is assumed. In (8), D_m depends on the free viewpoint (x, z) and k. N represents the length of one side of the square-shaped structure in Fig. 2. Given the viewpoint, the range of k is defined as shown in (9). When k reaches its maximum value in (9), D_m in (8) becomes zero, meaning that two 3D LFs are connected at the intersection point. As k decreases, D_m increases, reaching its maximum at $k = 45°$.

$$D_m = \frac{N/2 - (x + (N/2 - z) \tan k)}{\cos(90 - k)} \qquad (8)$$

$$k = \left[45, \tan^{-1} \frac{N/2 - x}{N/2 - z} \right] \qquad (9)$$

Next, E_b is defined in (10). Here, $\Delta\phi_b$ is obtained from (2). i and j denote θ and ϕ in the generated view, respectively.

As shown in Fig. 6 (b), bending errors prominently occur in a region where θ is close to $\pm 90°$. However, less noticeable bending errors are spread throughout the generated view. E_b is defined as the sum of $|\Delta \phi_b|$ in θ ranging from 0 to k and in the full range of ϕ. E_b depends on k because the error accumulation range is determined by k. From the equations for E_m and E_b, it is observed that as k increases, E_m decreases, and E_b increases. This is consistent with the analysis in Section III.

$$E_b(k) = \sum_{i=0}^{k} \sum_{j=-90°}^{90°} |\Delta \phi_b| \quad (10)$$

Fig. 11 depicts graphs of the defined errors with respect to k and d_{src}. Here, k varies between 45° and 90°, and d_{src} depends on the size of N and ranges from $1.2N$ to $2.5N$. Fig. 11 (a) shows the variations of E_t, E_m, and E_b with k for a fixed value of d_{src} ($1.5N$). It is observed that E_m decreases, and E_b increases as k increases. When k is 76°, E_t reaches its minimum value. As the average of E_b is relatively large compared to E_m, the optimal value of k that minimizes E_t is observed to the right of the horizontal axis. Fig. 11 (b) displays the E_t-k graph for different values of d_{src} ($2.5N$, $1.5N$, and $1.2N$). The graph becomes more convex with an increase in d_{src} because E_b is less sensitive to changes in d_{src} (since $D_b \ll d_{src}$ in (2), changes in d_{src} have relatively smaller effects on $\Delta \phi_b$ than on $\Delta \phi_m$). As a result, the optimal value of k tends to decrease as d_{src} increases. Fig. 11 (c) shows the optimal value of k with respect to d_{src}. In the given range of d_{src} ($1.2N$ to $2.5N$), the optimal k starts at 87° and converges to 72°. Since d_{src} is not fixed to a single value in the real world, the optimal value of k must be fine-tuned between 72° and 87°.

Ray-space360 [32] always uses a physical connection scheme and requires a light ray with a horizontal incident angle of up to 90°, corresponding to $k = 90°$ in Fig. 11 (a). On the other hand, LFU [16] uses light rays with $\theta = \pm 45°$ as a shared light ray, corresponding to $k = 45°$. LFU is vulnerable to mismatching errors, whereas Ray-space360 is susceptible to bending errors. The proposed method suggests using an appropriate value of k to avoid extreme exposure to both types of errors.

V. EXPERIMENTAL RESULTS
A. EXPERIMENTAL ENVIRONMENT
A GoPro Fusion camera [33], capable of capturing 360° footage, and a DollyPlus [34] were used to capture light rays. The 3D LFs were constructed by moving the dolly, equipped with the 360° camera, along the grid shown in Fig. 2. To generate HLF-VD views, the depth map was estimated using the optical flow algorithm [35]. The optical flow estimator first converts the given input image to various resolution levels and then gradually updates the optical flow estimated at low resolutions to the size of the original image.

Visual defects can easily arise due to discontinuities when connecting views in four directions, mainly for two reasons.
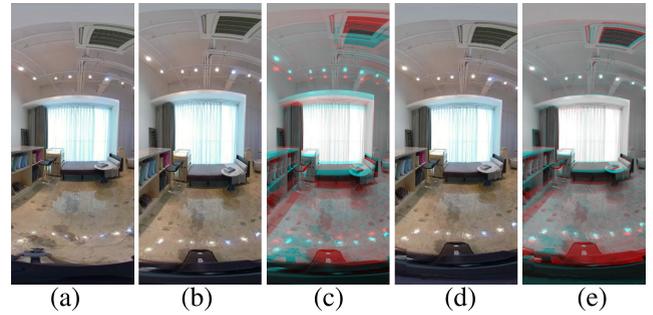
FIGURE 12. Comparison results of from HLF-VD view generation. (a) Actually captured view. (b) Initial view from 3D LF view generation. (c) Overlay image of (a) and (b). (d) Enhanced view through depth-based correction. (e) Overlay image of (a) and (d).

First, the mismatching error is not completely eliminated. The second reason is incomplete camera alignment while shooting. During light ray acquisition using a dolly, it is difficult to move accurately on an orthogonal grid. In order to mitigate the discontinuity, weighted blending is used as a post-processing technique [16]. Blending is generally used in methods of connecting different 3D LFs, and an overlapping area is required for this. The proposed hybrid connection method determines the connection point, and based on this, creates an overlapping area by rendering a slightly wider range. The additional rendering range is optional, and in this paper, 5° is used.

B. EFFECTS OF HLF-VD VIEW GENERATION
HLF-VD view generation compensates for weak vertical parallax in 3D LF. Fig. 12 compares the generated view before and after the application of depth-based correction. Fig. 12 (a) shows the actual view at a free viewpoint, while Fig. 12 (b) shows the generated view before correction. Although the unacquired light rays are replaced by the light rays obtained at the corresponding acquisition point, the result shows a photo-realistic view. Fig. 12 (c) shows an overlay image of Figs. 12 (a) and (b), where the horizontal parallax is well reflected but the vertical parallax is not. Fig. 12 (d) shows an enhanced view through depth-based correction, and Fig. 12 (e) shows an overlay image of Figs. 12 (a) and (d). Fig. 12 (e) shows that the enhanced view is relatively close to the actual captured image. In particular, the vertical parallax, which was not reflected in the initial view of Figs. 12 (b) and (c), is well reflected and still photo-realistic. There are two notable errors in Fig. 12 (d). One is the black object at the bottom, which is the dolly equipment. In many VR services, these errors in the pole are often covered by a logo or removed with a painting program. Another error is a stain on the floor. Since the rendered viewpoint is located behind the 3D LF, the floor stain is not reproducible with the given 3D LF. Nevertheless, the rendered view of Fig. 12 (d) is photo-realistic, and it is sufficient to fully experience the real space.

Fig. 13 compares the visual quality of the enhanced views and initial views in terms of the peak signal-to-noise ratio
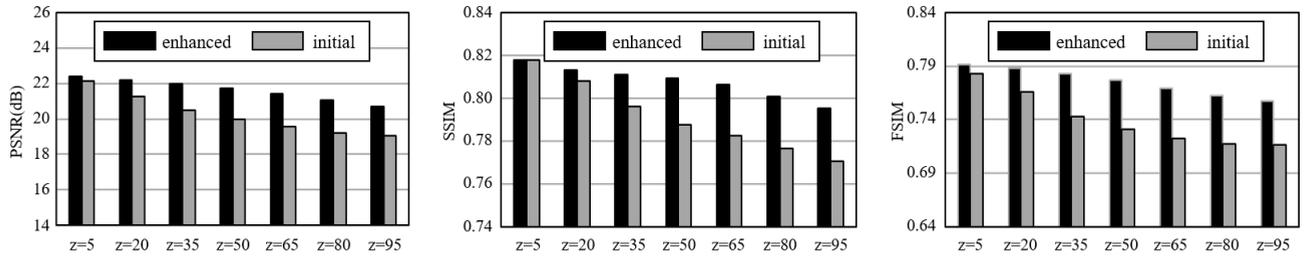
FIGURE 13. PSNR, SSIM, and FSIM comparison of initial view and enhanced view according to *z*-position.
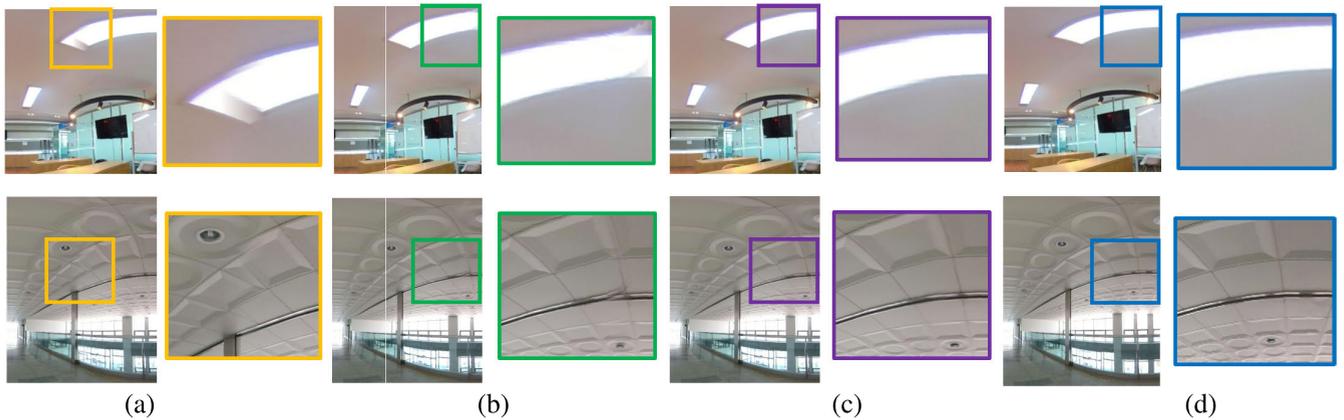
FIGURE 14. Connected part of the generated view, (a) LFU [16], (b) Ray-space360 [32], (c) proposed work, and (d) ground truth.

TABLE 2. Comparison of $E_{connect}$ of LFU, Ray-space360, and the proposed work.

Sample No.	LFU [16]	Ray-space360 [32]	Proposed work								
			k=70	k=72	k=74	k=76	k=78	k=80	k=82	k=84	k=86
1	812.6	621.8	650.1	646.0	637.9	630.1	621.5	609.8	**601.5**	602.1	621.8
2	968.0	479.6	411.0	389.1	379.0	**377.8**	397.2	410.3	424.7	440.4	477.9
3	543.2	286.6	276.7	266.1	263.4	267.8	262.5	**261.3**	264.1	268.3	285.7
4	1447.0	1303.5	1251.8	1261.0	1260.0	1256.8	1251.8	**1245.5**	1255.5	1278.5	1301.8
5	1664.9	620.6	638.4	625.4	612.9	604.9	600.0	594.1	**592.6**	600.2	620.2
6	1252.9	542.0	506.7	503.7	494.1	**486.7**	490.0	489.4	495.6	516.8	541.8

(PSNR), the structural similarity index measure (SSIM), and the feature similarity index measure (FSIM). The horizontal axis of the graphs represents the viewpoint in the *z*-direction. One side of the square-shaped 3D LF used in this experiment is 100cm. The user's movement range is limited to 100cm×100cm, and the range of *z* is from 0 to 99cm. Here, the 1D camera line is at *z*=0, and the larger the *z* of the viewpoint, the further the viewpoint is from the 1D camera line. The black and gray bars represent the measurements of the enhanced view and initial view, respectively. For all three measurements, higher values indicate better visual quality. The three graphs show that the visual error increases as *z* increases. An increase of *z* means an increase of d_p of (1), which in turn increases $\Delta\phi$. This causes a difference between the rendered view and the actual captured view, as shown in Fig. 12 (b). On the other hand, depth-based correction reduces this difference, and the enhanced view is synthesized close to the actual captured view, as shown in Fig. 12 (d). With a small *z*, the difference between the enhanced view and the initial

view is not significant. When *z* is 5cm, there is a difference of 0.22dB in the PSNR and 0.009 in the FSIM, and a difference of less than 0.001 in the SSIM. As *z* increases, the gap widens, resulting in differences of up to 1.86dB, 0.024, and 0.047 in the PSNR, SSIM, and FSIM, respectively.

When the viewpoint exceeds 99cm, it moves to the next 3D LF in the stackable 3D LF. Therefore, the view error due to the movement of the viewpoint does not deviate significantly from the level presented in the graphs of Fig. 13.

C. EFFECTS OF HYBRID 3D LF CONNECTION

Fig. 14 compares the proposed Hybrid 3D LF connection with other connection-based methods, LFU [16] and Ray-space360 [32], for two sample spaces. Figs. 14 (a), (b), and (c) show the generated views of LFU, Ray-space360, and the proposed method, respectively. Fig. 14 (d) shows the ground truth. In the first row, errors are noticeable in the lighting part of the ceiling. The LFU result of Fig. 14 (a) shows that the two
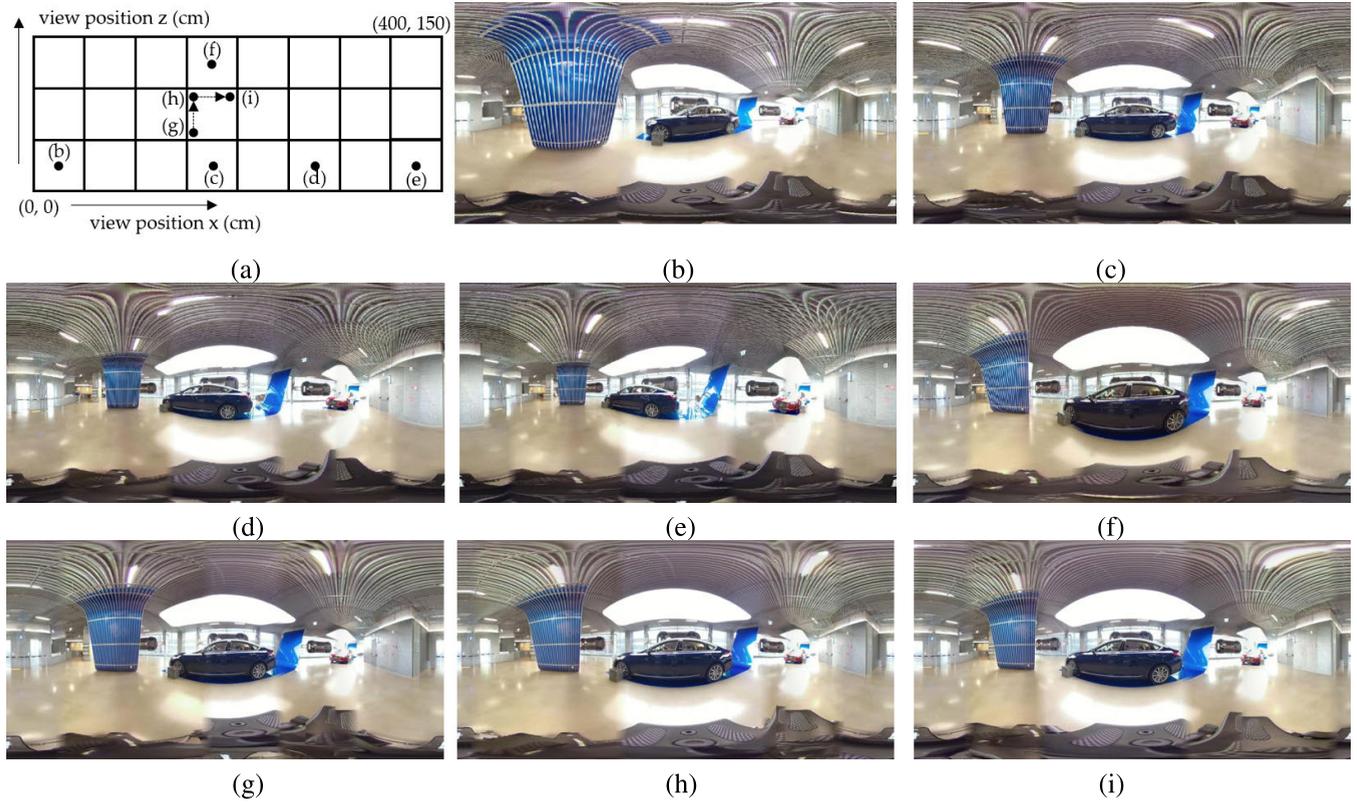
FIGURE 15. Generated 360° image of the proposed 3D LF-based VR system (Car showroom). (a) Constructed LF grid and viewpoints on it. Generated 360° view (b) at (25, 25), (c) at (175, 25), (d) at (275, 25), (e) at (375, 25), (f) at (175, 125), (g) at (155, 55), (h) at (155, 95), and (i) at (195, 95).

3D LFs are not naturally connected, which corresponds to a mismatching error. For Ray-space360, shown in Fig. 14 (b), a view is generated as if the light is bent, which corresponds to a banding error. On the other hand, the proposed method shows that the two 3D LFs are naturally connected without mismatching and banding errors, as shown in Fig. 14 (c), which is similar to the ground truth in Fig. 14 (d). Results from the second sample space in the second row show a similar tendency. As a result of LFU, the mismatching error is revealed in the ceiling pattern, and in the result of Ray-space360, the banding error appears in the bar part. The proposed method effectively reduces both types of errors.

Table 2 compares the connection error of Fig. 14 with the objective measurement, $E_{connect}$, and shows the changes according to k of the Hybrid 3D LF connection. In (11), img_0 and img_1 represent overlapping parts of the views generated from two neighboring 3D LFs. The mean-square error (*mse*) of img_0 and img_1 measures the pixel difference between the images and evaluates how well the two images are connected. $E_{connect}$ is defined as the average *mse* for N viewpoints. The smaller the $E_{connect}$, the more naturally the two images are connected.

$$E_{connect} = \frac{1}{N} \sum_{(x,z)}^{N} mse\left(img_0(x,z), img_1(x,z)\right) \quad (11)$$

Table 2 compares the $E_{connect}$ of LFU, Ray-space360, and the proposed method for six samples and 2,401 viewpoints.

The two samples shown in Fig. 14 correspond to samples 1 and 2. In Table 2, the $E_{connect}$ of LFU is always the largest in each sample because the mismatching error causes an overall pixel difference in the overlapping images. The $E_{connect}$ of Ray-space360 is relatively small because only part of the image has bending errors. Although the proposed method leads to both mismatching and bending errors, they are effectively reduced through the Hybrid 3D LF connection. The smallest $E_{connect}$ is achieved with the specific k value marked in bold for each sample. Depending on the sample, the minimum $E_{connect}$ is determined in the range of 76° to 82°.

This is analogous to the trend observed in the mathematically defined model, E_t, as shown in Fig. 11 (a). Both E_t and $E_{connect}$ are relatively high at k=45°, and decrease as k increases. The optimal range of k for E_t is between 72° to 87°, depending on d_{src}. In the experimental results presented in Table 2, the optimal values of k are 76°, 80°, and 82°. In a real-world environment, objects are typically distributed at varying distances, unlike in fixed values. For instance, samples 2 and 6 with relatively smaller optimal k values have an average d_{src} value of objects that is smaller than the other samples.

Finally, in the above comparison, the vertical parallax error is not revealed superficially. However, since the proposed method, LFC, and Ray-space360 all use the same 3D LF, the same vertical parallax error is included. As the purpose of
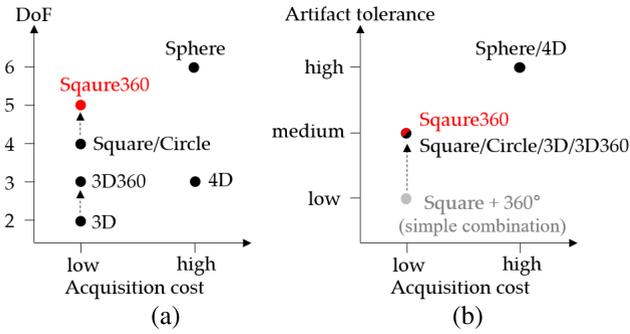
FIGURE 16. Comparison with other LF structures. (a) Light ray acquisition cost-DoF graph. (b) Light ray acquisition cost-Artifact tolerance graph.

TABLE 3. Supporting DoF according to LF structures.

	3D	4D	3D360	Circle	Square	Square360	Sphere
x	✓	✓	✓	✓	✓	✓	✓
y	-	✓	-	-	-	-	✓
z	-	-	-	✓	-	✓	✓
roll	✓	✓	✓	✓	✓	✓	✓
yaw	-	-	-	✓	✓	✓	✓
pitch	-	-	✓	-	-	✓	✓

FIGURE 17. An example of the estimated depth.

this section is to compare LF connections, the analysis of the vertical parallax error is not discussed here. It is introduced in detail in Section V-B.

D. VIEW EXPLORATION THROUGH THE PROPOSED 5-DoF VR SYSTEM

Finally, Fig. 15 shows the results of the proposed 5-DoF VR system, which includes HLF-VD view generation and the Hybrid 3D LF connection. The system is configured for the car showroom depicted in Fig. 15, and Fig. 15 (a) shows the LF grid configured in this space. The size of each square is 50cm×50cm, and the free viewpoint can move a total of 150cm×400cm. The black dots in Fig. 15 (a) represent the selected viewpoints, and the corresponding views are shown in Figs. 15 (b), (c), (d), (e), (f), (g), (h), and (i). Figs. 15 (b), (c), (d), and (e) show the corresponding view changes from the viewpoints of (25, 25) to (175, 25), (275, 25), and (375, 25), representing the case where the viewpoint moves along the x-axis while the z-axis remains fixed. In the 360° image, the right quarter represents the x-axis direction, where the red car is located. As the viewpoint moves from (b) to (c), (d), and (e), the view moves closer to the red car, and the view change for the navy-blue car in the front of the image is also well reflected. Fig. 15 (f) is a view that moves 100cm in the z-axis direction relative to the viewpoint of Fig. 15 (c). In Fig. 15 (f), the view is closer to the navy-blue car compared to that in Fig. 15 (c). Figs. 15 (g), (h), and (i) show the viewpoint movements within a single unit. The view changes within the single unit are relatively limited to a certain range, but they are still well-reflected.

As shown in Fig. 15, the viewpoint can move freely along the x-axis and the z-axis, and 360° images for each viewpoint are created well. The proposed system satisfies 5-DoF movement by supporting translational viewpoint movements of the x-axis and z-axis and rotational view changes of the roll, yaw, and pitch.

E. COMPARISON WITH OTHER LF STRUCTURES

This section compares the proposed LF structure with other LF-based structures. Firstly, Fig. 16 (a) compares the light ray acquisition cost and the supported DoF. The acquisition cost

depends on whether the acquisition layout is a plane or a line. The cost of the sphere and 4D, where the cameras are laid out on a plane, is higher than the remaining structures with a line layout. In terms of supporting DoF, Square360 and 3D360 provide a higher level of DoF compared to Square and 3D, respectively. Specifically, Square360 provides the highest level of 5-DoF among the LF structures with the line layout. Sphere provides the best 6-DoF due to the high acquisition cost. Table 3 shows the detailed DoF supported by each LF structure. Fig. 16 (b) compares the light ray acquisition cost and the degree of artifact tolerance. Sphere and 4D, with sufficient light rays, maintain high artifact tolerance. On the other hand, the remaining line-based structures are relatively vulnerable to visual artifacts due to insufficient light rays. The simple combination of the square and 360° camera shows low artifact tolerance like the gray dot, but it is improved to the level of the existing 3D-based LF structure through the proposed schemes.

F. DISCUSSION
1) INTERSECTION ALIGNMENT

The proposed system is based on a grid-shaped LF structure. The challenge with this structure is achieving precise matching at intersections. To create a grid-shaped LF, images are captured by moving the camera in the vertical and horizontal directions. It is assumed that the two images corresponding to the intersection point are identical. However, achieving precise matching during actual implementation can be difficult. To address this issue, the proposed system applies post-processing steps such as blending. However, if the distortion at the intersection point is severe, the post-processing effect may be insignificant.

2) INACCURATE DEPTH

In the depth-based approach (HLF-VD), depth accuracy plays a critical role. Fig. 17 presents an instance of the estimated optical flow utilized in this paper. Despite some inaccuracies,

the most significant error is the difference between the depth of the ceiling light and the depth of the ceiling. In the next version of the system, we plan to compensate for this by developing a rendering method that employs two 3D LFs placed in front and behind simultaneously.

3) DYNAMIC SCENE

The proposed system does not support dynamic or outdoor scenes. This is not a limitation of the proposed method, as the problem is more related to the fact that a grid-shaped LF cannot be obtained in dynamic or outdoor environments. As a complementary method, LF editing can be applied to augment dynamic objects to a static LF or change the lighting of the static LF over time [36], [37].

VI. CONCLUSION

This paper aims to achieve 5-DoF in the existing stackable 3D LF system. The contributions of this paper are as follows. Firstly, it defines and analyzes visual artifacts caused by light rays with large incident angles in a 3D LF. The proposed method increases FOV to improve DoF; however, the simple use of light rays with large incidence angles causes visual artifacts, which require careful consideration. Secondly, this paper proposes the HLF-VD view generation and hybrid 3D LF connection schemes to reduce visual artifacts based on the analysis results. The experimental results show a reduction in vertical parallax, bending, and mismatching errors compared to existing methods. Furthermore, the results in the actual space demonstrate well-reflected 360° views according to the user's free movement. The proposed system achieves the highest DoF among 3D LF-based methods and demonstrates superior performance in terms of artifacts. Consequently, the proposed method comes closest to the most complete structure, the sphere, while maintaining a low acquisition cost because it is still based on the 3D LF.

REFERENCES

[1] Samsung. *Gear 360 Camera*. Accessed: Jul. 11, 2022. [Online]. Available: https://www.samsung.com/latin_en/news/local/samsung-newgear-360-introduces-true-4k-video-and-360-degree-content/

[2] Facebook. *Surround 360 Video Camera*. Accessed: Jul. 11, 2022. [Online]. Available: https://www.facebook.com/formedia/blog/facebook-360-updates-and-introducing-surround-360

[3] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010.

[4] H.-H. Vu, P. Labatut, J.-P. Pons, and R. Keriven, "High accuracy and visibility-consistent dense multiview stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 889–901, May 2012.

[5] S. Choi, Q.-Y. Zhou, and V. Koltun, "Robust reconstruction of indoor scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5556–5565.

[6] H. R. Kaviani and S. Shirani, "An adaptive patch-based reconstruction scheme for view synthesis by disparity estimation using optical flow," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 7, pp. 1540–1552, Jul. 2018.

[7] H.-A. Hsu, C.-K. Chiang, and S.-H. Lai, "Spatio-temporally consistent view synthesis from video-plus-depth data with global optimization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 1, pp. 74–78, Jan. 2014.

[8] P. Hambarde and S. Murala, "S2DNet: Depth estimation from single image and sparse samples," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 806–817, 2020.

[9] E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision," in *Computation Models of Visual Processing*, M. S. Landy and J. A. Movshon, Eds. Cambridge, MA, USA: MIT Press, 1991, pp. 3–20.

[10] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. 23rd Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, 1996, pp. 31–42.

[11] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Proc. 23rd Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, 1996, pp. 43–52.

[12] Z. Lin and H.-Y. Shum, "A geometric analysis of light field rendering," *Int. J. Comput. Vis.*, vol. 58, no. 2, pp. 121–138, 2004.

[13] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, "Plenoptic sampling," in *Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn.*, Jul. 2000, pp. 307–318.

[14] R. S. Overbeck, D. Erickson, D. Evangelakos, M. Pharr, and P. E. Debevec, "A system for acquiring, processing, and rendering panoramic light field stills for virtual reality," *ACM Trans. Graph.*, vol. 37, no. 6, pp. 1–15, Nov. 2018.

[15] M. Broxton, J. Flynn, R. Overbeck, D. Erickson, P. Hedman, M. Duvall, J. Dourgarian, J. Busch, M. Whalen, and P. Debevec, "Immersive light field video with a layered mesh representation," *ACM Trans. Graph.*, vol. 39, no. 4, pp. 1–15, Aug. 2020.

[16] H. Jung, H.-J. Lee, and C. E. Rhee, "Flexibly connectable light field system for free view exploration," *IEEE Trans. Multimedia*, vol. 22, no. 4, pp. 980–991, Apr. 2020.

[17] B. Wilburn, N. S. Joshi, V. Vaish, E. V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High performance using large camera arrays," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 765–776, Jul. 2005.

[18] L. McMillan and G. Bishop, "Plenoptic modeling: An image-based rendering system," in *Proc. 22rd Annu. Conf. Comput. Graph. Interact. Techn.*, 1995, pp. 39–46.

[19] M. Tanimoto and H. Kurokawa, "Ray-space processing for omnidirectional FTV," *Proc. SPIE*, vol. 10666, May 2005, Art. no. 1066607.

[20] M. Tanimoto, M. P. Tehrani, T. Fujii, and T. Yendo, "Free-viewpoint TV," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 67–76, Jan. 2011.

[21] C. Kim, A. Hornung, S. Heinzle, W. Matusik, and M. Gross, "Multi-perspective stereoscopy from light fields," *ACM Trans. Graph.*, vol. 30, no. 6, pp. 1–10, Dec. 2011.

[22] H.-Y. Shum, K.-T. Ng, and S.-C. Chan, "A virtual reality system using the concentric mosaic: Construction, rendering, and data compression," *IEEE Trans. Multimedia*, vol. 7, no. 1, pp. 85–95, Feb. 2005.

[23] A. Sepas-Moghaddam, M. A. Haque, P. L. Correia, K. Nasrollahi, T. B. Moeslund, and F. Pereira, "A double-deep spatio-angular learning framework for light field-based face recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 12, pp. 4496–4512, Dec. 2020.

[24] S. Wanner and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 606–619, Mar. 2014.

[25] D. Liu, Y. Huang, Q. Wu, R. Ma, and P. An, "Multi-angular epipolar geometry based light field angular reconstruction network," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1507–1522, 2020.

[26] Y. Wang, F. Liu, K. Zhang, Z. Wang, Z. Sun, and T. Tan, "High-fidelity view synthesis for light field imaging with extended pseudo 4DCNN," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 830–842, 2020.

[27] H. Zhu, X. Sun, Q. Zhang, Q. Wang, A. Robles-Kelly, H. Li, and S. You, "Full view optical flow estimation leveraged from light field superpixel," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 12–23, 2020.

[28] S. U. Premaratne, N. Liyanage, C. U. S. Edussooriya, and C. Wijenayake, "Real-time light field denoising using a novel linear 4-D hyperfan filter," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 8, pp. 2693–2706, Aug. 2020.

[29] S. Ghosh and K. N. Chaudhury, "Artifact reduction for separable nonlocal means," *J. Electron. Imag.*, vol. 26, no. 6, pp. 1–6, Nov. 2017.

[30] S. Maesen, P. Goorts, and P. Bekaert, "Omnidirectional free viewpoint video using panoramic light fields," in *Proc. 3DTV-Conf., True Vis. Capture, Transmiss. Display 3D Video (3DTV-CON)*, Jul. 2016, pp. 1–4.

[31] X. Lyu, X. Wang, G. Zhou, and Q. Wang, "A 3DOF+ view rendering method for panoramic light field," *Proc. SPIE*, vol. 11550, Oct. 2020, Art. no. 115500B.

[32] H. Jung, H.-J. Lee, and C. E. Rhee, "Ray-space360: An extension of ray-space for omnidirectional free viewpoint," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 27–30.

[33] Gopro. *Fusion 360° Camera*. Accessed: Jul. 11, 2022. [Online]. Available: https://gopro.com

[34] Edelkrone. *Dolly Plus*. Accessed: Jul. 11, 2022. [Online]. Available: https://edelkrone.com/products/dollyplus

[35] V. Chari. *High Accuracy Optical Flow*. Accessed: Oct. 26, 2022. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/17500-high-accuracy-optical-flow

[36] W. Williem, K. W. Shon, and I. K. Park, "Spatio-angular consistent editing framework for 4D light field images," *Multimedia Tools Appl.*, vol. 75, no. 23, pp. 16615–16631, 2016.

[37] Z. Lu, X. Chen, Y. Y. Chung, and Z. Chen, "Light field editing propagation using 4D convolutional neural networks," in *Proc. IEEE Conf. Virtual Reality 3D User Interface Abstr. Workshops (VRW)*, Mar. 2020, pp. 604–605.

[38] Y. Wang, "Disentangling light fields for super-resolution and disparity estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 425–443, Jan. 2023.

[39] C. Liu, L. Shi, X. Zhao, and J. Qiu, "Adaptive matching norm based disparity estimation from light field data," *Signal Process.*, vol. 209, Aug. 2023, Art. no. 109042, doi: 10.1016/j.sigpro.2023.109042.

HYUK-JAE LEE (Member, IEEE) received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, South Korea, in 1987 and 1989, respectively, and the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 1996. From 1998 to 2001, he was a Senior Component Design Engineer with the Server and Workstation Chipset Division, Intel Corporation, Hillsboro, OR, USA. From 1996 to 1998, he was a Faculty Member with the Department of Computer Science, Louisiana Tech University, Ruston, LA, USA. In 2001, he joined the School of Electrical Engineering and Computer Science, Seoul National University, where he is currently a Professor. He is the Founder of Mamurian Design Inc., Seoul, a fabless SoC design house for multimedia applications. His current research interests include computer architecture and SoC for multimedia applications.

HYUNMIN JUNG received the B.S. degree in electric engineering from Kyung Hee University, Yongin, South Korea, in 2014, and the M.S. and Ph.D. degrees in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2016 and 2020, respectively. From 2020 to 2023, he was a Postdoctoral Researcher with the Inter-University Semiconductor Research Center, Seoul National University. In 2023, he joined the Department of Smart ICT Convergence Engineering, Seoul National University of Science and Technology, where he is currently an Assistant Professor. His research interests include immersive media, virtual reality, augmented reality, and light field processing.

CHAE EUN RHEE (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from Seoul National University, Seoul, South Korea, in 2000, 2002, and 2011, respectively. From 2002 to 2005, she was with the Digital TV Development Group, Samsung Electronics Company Ltd., Suwon, South Korea, as an Engineer, where she was involved in bus architecture and MPEG decoder development. In 2013, she joined the Department of Information and Communication Engineering, Inha University, South Korea, where she is currently an Associate Professor. Her research interests include algorithm and architecture design of video coding for HEVC and H.264/AVC, configurable video coding for real time systems, and the next generation virtual reality systems.

• • •