

Article

One Shot 360-Degree Light Field Capture and Reconstruction with Depth Extraction Based on Optical Flow for Light Field Camera

Youngmo Jeong, Seokil Moon, Jinsoo Jeong, Gang Li, Jaebum Cho and ByoungHo Lee *

School of Electrical and Computer Engineering, Seoul National University, Gwanak-Gu Gwanakro 1, Seoul 08826, Korea; youngmo.snu@gmail.com (Y.J.); kyo517@snu.ac.kr (S.M.); jeongfrn@gmail.com (J.J.); ligang0227@gmail.com (G.L.); chojaebum@gmail.com (J.C.)

* Correspondence: byoungHo@snu.ac.kr; Tel.: +82-2-880-7245

Received: 16 April 2018; Accepted: 19 May 2018; Published: 29 May 2018



Abstract: A system for capturing 360-degree light field information of real-existing object in one shot and then optically reconstructing is proposed. A new depth extraction algorithm for light field cameras is proposed and various camera specifications for practical use of the algorithm are employed for the analysis. With a depth extraction method that was based on optical flow for light field camera, the depth information is extracted more accurately, according to the various specifications of light field camera. For 360-degree shooting, a simple capturing system composed of two mirrors and a light field camera is used. The capturing system has an advantage of being compact and inexpensive. The locations and orientations of the two mirrors are analyzed to optimize 360-degree light field recording. Holographic display is used to optically reconstruct the captured light field information. Experimental and simulation results are presented to support the proposed system and analysis.

Keywords: light field information; three-dimensional display; image formation theory

1. Introduction

With the explosive interest in implementing and improving augmented reality (AR) and virtual reality (VR) systems, there is an increasing demand to acquire information on real three-dimensional (3D) objects to produce contents for realistic displays. Efforts to develop core applications of VR and AR displays, such as head mounted display (HMD) and head-up display (HUD), are widely researched [1–5]. In order to display natural real 3D scene, obtaining 3D information of the object is necessary. There are several different ways to acquire 3D information, such as structured light, time-of-flight, and light field [6–10]. Among them, the light field camera (LFC) is attracting much attention as a next generation camera that can record depth and color information of object, while it is similar to a conventional camera in appearance and it has an advantage in portability [10]. An LFC records four-dimensional (4D) light field (two-dimensional (2D) spatial information and 2D angular information) of given scene in the form of a 2D image. Since it was proposed as a device for taking plenoptic function, there have been a large number of studies on ways to handle and reconstruct the captured light field [11–15]. They also explored how to get more precise information from a LFC and what optical information a LFC gets exactly—in other words, where a ray starts and where it heads to [12,16,17].

According to the specifications of optical element in LFC, such as sensor size, pixel pitch, lens pitch, f -number, and position of focal plane, each pixel of LFC is magnified to a certain position on space and it records a bundle of rays having a specific range of angle of incidence. This means that light field information is sampled differently to the captured image, depending on many input parameters of the light field camera. Therefore, to transform the captured image to the exact light field, it is necessary to

define which light field the image sensor samples according to the details of the LFC. One research on the LFC has been conducted on how much information the captured 2D image has in the 4D light field domain [17]. By using this study, an image that is taken by an LFC can be easily processed to light field information. On the other hand, to reconstruct a 3D object from the recorded light field, depth information should be extracted from the light field data. Methods for extracting the depth by using integral imaging, which is an optical system that is similar to LFC, have been proposed [18,19]. These methods use stereo matching with pixel-unit accuracy. Therefore, it is impossible to extract continuous depth information and cannot reconstruct accurate 3D volume. There is also a study to extract depth map with sub-pixel accuracy in images that are taken with a LFC [20,21]. However, although this method has a high accuracy, there is a limitation that the extracted depth is a relative value between pixels, not absolute value, so that additional calibration is required to reconstruct as 3D model.

In this paper, depth extraction method based on optical flow for light field camera (DOLF) is proposed. Optical flow calculation is used to analyze light field information for several applications, such as visual odometry and the reconstruction of transparent flow surface [22,23]. In these studies, the optical flow is calculated between images made by each lens of lens array. There is also a study using optical flow for depth information in integral imaging (InIm), but it has not yet been used to extract depth information by comparing sub-images taken by an LFC [24]. Paying attention to the difference between InIm and LFC systems, we geometrically analyze the two optical systems to determine the relationship about the optical flow value of the sub-images and the actual depth of the object. The depth information of center sub-image is calculated as absolute value of length under given conditions of the LFC. To provide stability to the algorithm, maximum optical flow value where corresponding points can be located on a physically allowed area is determined, according to relative position difference between sub-images. By using this algorithm, one can extract actual depth information from a light field image that is taken by LFC under the given condition of the LFC. Consequently, the optical flow method suitable for LFC images is explained.

A 3D model of the real object can be reconstructed if there are multiple light field images captured from different location of LFCs. Each light field image taken by LFC becomes a perspective light field image from the object, which composes total light field of the object. If the location and orientation of the LFCs for perspectives are given, the 360-degree light field information of the object can be determined. To reconstruct the object by using optical or computational system, the depth and the color information should be extracted from the perspective light field. From the captured light field images, the DOLF can be used to extract depth information and digital refocusing technique can be utilized to calculate clear color images. By binding overall perspective information of depth map and color image, 3D structure of the object can be reconstructed. Therefore, we also propose an optical system that acquires 360-degree light field information of an object by using a light field camera in one-shot. Putting several depth cameras in multiple locations or while using mechanical apparatus to control orientation of subjects is a conventional way to achieve a 360-degree depth profile [25,26]. However, the systems have limitations on both cost and size, or the output results are a long way from the high-quality images. To overcome these limitations, a compact and inexpensive 360-degree light field capturing system is proposed in this paper. This system can be used as an alternative of 3D scanner. The capturing system allows for capturing not only depth and color information of front view, but also that of back and side views. The proposed system contains two plane mirrors and one LFC. Two mirrors and an object are placed within the viewing angle of the LFC to obtain 4D light field information of the object in a single shot. The locations and orientations of two mirrors are analyzed to optimize 360-degree light field recording and controlled minutely to determine the angle of pickup direction. Experiments and simulation results are presented to support the feasibility of DOLF and the capturing system. Consequently, 360-degree light field information is captured in one-shot and is optically reconstructed as the 3D object. Holographic display is used to reconstruct the wavefront of the captured objects. Experimental results prove the validity of the proposed pickup method with LFC and DOLF.

2. Light Field Camera Structure and Depth Extraction

2.1. Pickup Process of Light Field Camera Based on Geometry Optics

Figure 1 shows the geometrical structure of a light field camera. In this study, the standard plenoptic camera is used as the structure of the light field camera. When the distance between the microlens array (MLA) and main lens is a , the distance from the main lens to the focal plane is represented by b_1 where the focal plane of LFC lies at a specific location, not infinity. The microlens diameter is l_p and the focal length of the MLA is f_m . The light field information of the object is recorded by the image sensor and MLA, magnified by the main lens. Since the pickup unit is floated on the specific position, it has an advantage that all the objects in the front and rear of focal plane can be photographed. According to the modified f -number matching condition, the overall size of the image sensor and the MLA are different, and the distance between two optical components is as much as the focal length of the MLA. Because each element is located at a different distance from the main lens, the location and magnification of two elements differ from each other. By using simple geometrical optics, one can determine the magnification value for MLA, and from this, b_1 is determined with the lens formula. The distance between the main lens and the magnified sensor is $b_2 = (f^{-1} - (a + f_m)^{-1})^{-1}$. Detailed optical properties of the pickup unit depending on specifications and focal plane of LFC have been previously analyzed [17].

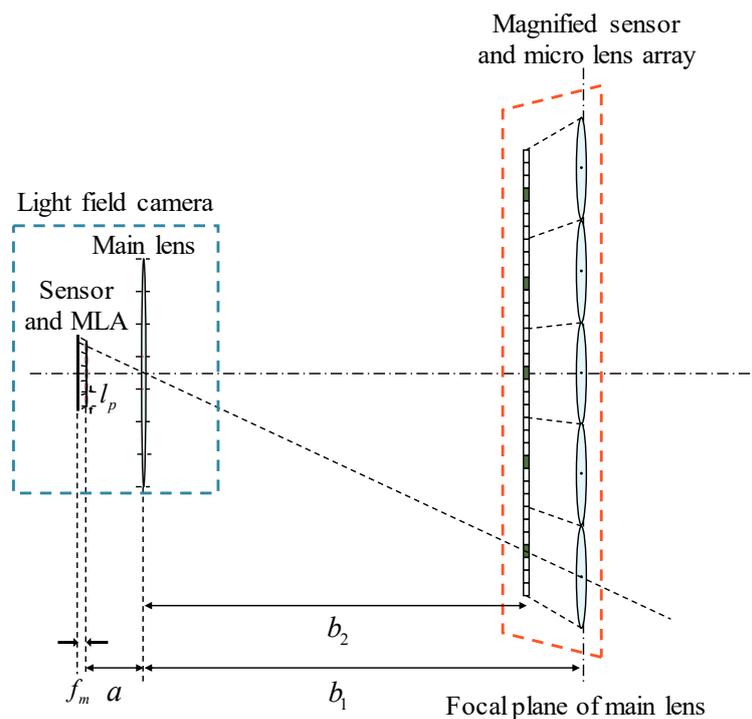


Figure 1. Geometry structure of light field camera. MLA: micro lens array.

2.2. Depth Extraction Based on Optical Flow for Light Field Camera

To obtain the depth information of an object by processing the 4D light field image, we propose depth extraction based on the optical flow for light field camera. This method extracts the local disparity between the sub-images using the optical flow method. Optical flow is a distribution of relative motion difference according to the distance between an imaging system and a subject [27]. By finding a corresponding pixel to a given pixel, the method calculates how quickly the pixel flows in the images [28]. Conventional methods of extracting the depth of the object by analyzing the elemental image is the sum of absolute difference (SAD), sum of squared differences (SSD), and sum of sum of squared differences (SSSD) while using

the minimum difference function [18,19]. These methods, however, have problems that the extracted depths are quantized and are not able to measure in a sub-pixel unit. To solve these problems, a method has been proposed for averaging the depth information that was extracted by comparing optical flows between sub-images and applying a median filter [24]. However, an optical flow method for LFC has not been suggested yet. Captured light field information of LFC also has very narrow baseline to compare the position of corresponding points [20,21]. Therefore, sub-pixel-level depth extraction algorithms are essential for the LFC. The corresponding points between the images can be found by comparing the sub-image at the very center of the light field information and periphery sub-images. According to the specifications of the optical system, the depth information of certain objects is encoded with the disparity between the sub-images calculated by the optical flow value.

Figure 2 shows the geometrical relationship between optical flow data and depth information in InIm and LFC pickup system. Optical flow (OF) means the value of optical flow between two corresponding points. The main difference of two optical systems is that, in the focused mode InIm, the axes of each microlens and image sensor are both parallel and perpendicular to the optical system. In case of pickup by the LFC, however, the MLA and the image sensor are differently located and are magnified depending on distances from the main lens to the optical elements. Since the magnification values of both optical elements are different, the microlens and the part of image sensors assigned to them are tilted out of the main optical axis depending on their positions. Because two capturing systems have different optical structures, the depth extraction method for light field camera should be reestablished. It is supposed that lens array in InIm and magnified MLA in the LFC have the same size. m_1 and m_2 are magnifications of MLA and image sensor, respectively. The variable p_p is the pixel pitch of the image sensor. As shown in checked pixels in Figure 2a, a point object in space can be repeatedly recorded on the image sensor by MLA. The core task of depth extraction is to reveal the relationship between the location of repeating points and depth information. In focused mode, InIm pickup, for example, let two checked pixels on the Figure 2a have a corresponding point relation. The rays recorded at two pixels start from the same point but pass different centers of microlenses. Because the distance between two corresponding points on image sensor is two times larger than the lens pitch, the optical flow is calculated as 2. In Figure 2a, the black pixel on a center lens image and other black pixels are supposed to be corresponding points to each other. Depending on the optical flow values, the depth of each corresponding point is determined. The position of the depth according to the calculated optical flow is represented as d_{OF} in Figure 2a, and the depth increases linearly with the value of optical flow. In case of LFC that is shown in Figure 2b, the depth of optical flow value is represented as D_{OF} . Figure 2 shows only the integer values of the optical flow for the sake of convenience though this algorithm has sub-pixel accuracy.

The relationship between depth information and optical flow in LFC image should be analyzed to determine DOLF. To solve the depth quantization of extraction algorithm, the accuracy of depth map is secured by applying a median filter or averaging process to each depth map that was obtained by comparing each sub-image with the center sub-image. Equation (1) describes the relationship between the depth information and the optical flow in an LFC. Once a sub-image to be compared and optical flow of corresponding points are determined, the location of the point object is calculated simply by triangle proportionality. Let L be the pitch of the magnified microlens and P be the pixel pitch of the magnified image sensor, i.e., $L = l_p \times m_1$ and $P = p_p \times m_2$. When the resolution of one lens image is $n_x \times n_y$, center index of lens is represented (i_c, j_c) , respectively. The distance between two optical elements can be calculated as $|b_2 - b_1|$ by lens formula. In this paper, it is assumed that optical flow only uses the result as compared with the center sub-image. The variable O_h means optical flow in the x -axis, and O_v means that in the y -axis. $O_{h(i,j)}(x,y)$ or $O_{v(i,j)}(x,y)$ denotes a 2D input function where the output is an optical flow value for a given lens index comparing the center sub-image. Note that input

variable of O_h and O_v are omitted for simplicity in Equation (1). If x and y are coordinates in the center sub-image, then the depth for the center sub-image can be obtained as:

$$D_{OF}(x, y) = \frac{b_2 - b_1}{2} \left[\frac{1}{n_x - 1} \sum_{i=1, i \neq i_c}^{n_x} \left| \frac{O_{h(i, j_c)} \cdot L}{O_{h(i, j_c)} \cdot P \cdot n_x + (i_c - i)P - O_{h(i, j_c)} \cdot L} \right| + \frac{1}{n_y - 1} \sum_{j=1, j \neq j_c}^{n_y} \left| \frac{O_{v(i_c, j)} \cdot L}{O_{v(i_c, j)} \cdot P \cdot n_y + (j_c - j)P - O_{v(i_c, j)} \cdot L} \right| \right] \quad (1)$$

The DOLF method is described in Equation (1). Using the above equation, one can generalize the case of extraction using the integral imaging technique. In integral imaging, the lens pitch and the image sensor assigned to it are assumed to have the same size. In other words, in the above equation, $P \times n$ is equal to L , which results in erasing the O -related term of the denominator. The value of $b_2 - b_1$ simply becomes the focal length of the lens array. This result is consistent with previous studies using integral imaging [24]. The above depth calculation method can be an example of the extraction method using optical flow. In the above equation, the optical flow for the center sub-image is calculated in the horizontal and the vertical directions and the values are averaged. Depth maps that were obtained from several sub-images are averaged again. However, depending on the situation and user preferences, various statistical techniques other than averaging can be adopted, such as the median filter.

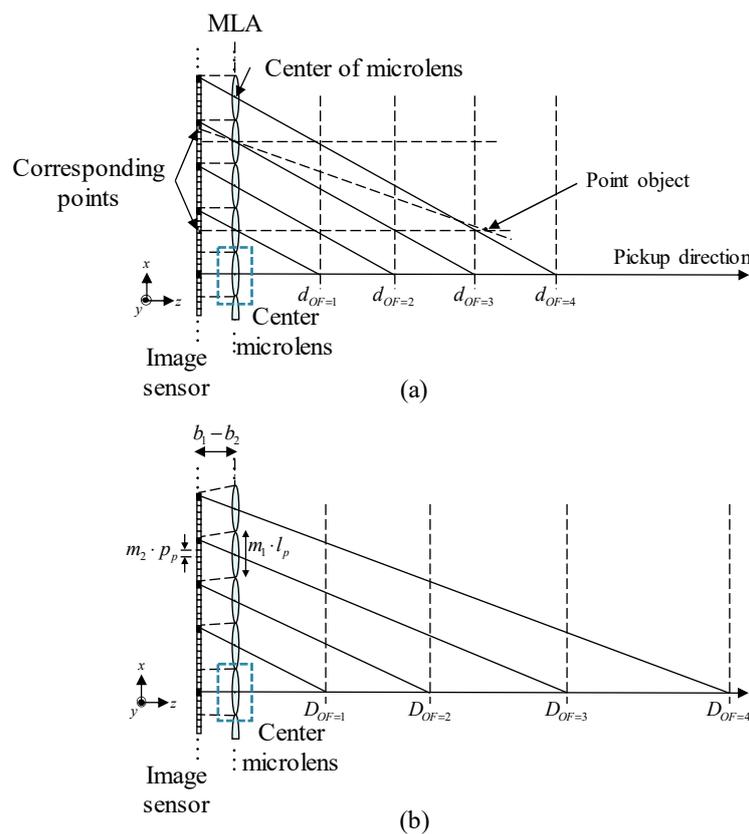


Figure 2. Depth extraction from captured light field using optical flow in (a) integral imaging and (b) light field camera.

There is another difference between two pickup methods shown in Figure 2. In 3D object pickup using focused mode, InIm, there is no theoretical limit on the maximum value of optical flow for captured light field. If the sensor format of pickup system is large enough, that is, if the overall size of the lens array and image sensor are large enough, the maximum optical flow value that can be captured also increases. This is because the central optical axes are still parallel, even though a

microlens is located at the periphery. Thus, the pixels that are located near the edges of lens images can still converge at a physically plausible location. On the contrary, a certain optical axis of microlens and the main axis of LFC converge at a physically impossible position because the optical axes of microlenses are tilted out of the main optical axis according to their positions. As the optical flow value increases, the extracted depth gradually increases and soon diverges, as shown in Figure 2b. This phenomenon occurs at a smaller optical flow value when nearer sub-image to center sub-image is used. In other words, the available optical flow range is determined according to the view interval from the center sub-image.

Figure 3a represents the depth on optical flow value by DOLF for corresponding points on center sub-image and another sub-image. Since two pixels that converge above a certain optical flow are located at the position that cannot be physically allowed, it is necessary to use a mask to neglect a range of optical flow value in forbidden area. In this paper, the mask for the specified range of optical flow value is applied to a case where the convergent depth position is completely unrealistic, that is, when the depth information is negative. Figure 3b shows the maximum optical flow value of what is available for the center sub-image in the order of the sub-images. The dotted line indicates the middle index of the sub-image. The optical flow values that are shown here are absolute values. Though, for example, the third and the fifth sub-image have the same optical flow value, the calculated data flows in different directions. Figure 4a,b shows simulated light field image and Figure 4c represents the result of depth information that was extracted from light field by optical flow calculation using Equation (1). We simulated the raw light field that was photographed by the LFC and extracted the light field information through the modified optical flow method. The specification for light field image is presented in Table 1. Figure 4d represents detailed depth information along the white dotted line in Figure 4c. Figure 4d compares the results that were obtained by DOLF and optical flow for InIm with ground truth data. To extract a precise depth, it is necessary to consider the geometrical arrangement of the light field camera, as shown in Figure 4d. This comparison is meaningful when considering that depth extraction algorithms for light field cameras extract relative depth rather than actual depth. When extracting the relative depth, it is necessary to fit the extracted value to the actual distance according to the many parameters of imaging system. In this case, the optical configuration of the light field camera must be fully considered to obtain the correct depth extraction result. As a result, the proposed depth extraction method for LFC is verified. It is noted that the data of ground truth is more discontinuous than the result of depth extraction on the center of graph because the original depth map of sub-image has been quantized after 8-bit conversion process.

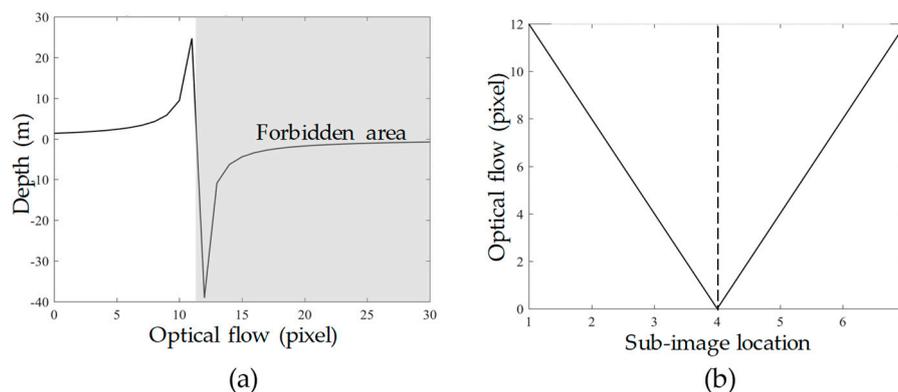


Figure 3. (a) Extracted depth of center sub-image with seventh sub-image (among seven sub-images) according to optical flow (b) Maximum optical flow on location of sub-image (among seven sub-images).

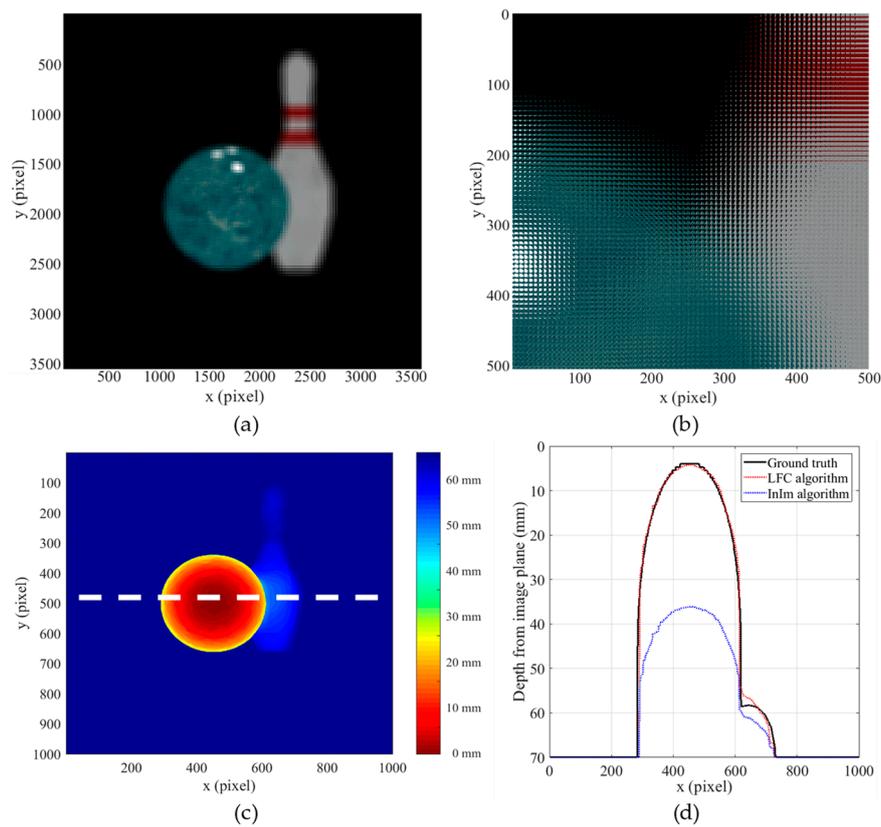


Figure 4. Depth extraction example: (a) Example of light field image; (b) magnified light field image, (c) depth extraction results; and, (d) comparison of the ground truth and extracted depth maps obtained by InIm method and light field camera (LFC) method.

Table 1. Specification for simulation and experiments.

Optical System	Specifications	Value
Microlens array	Lens pitch	38.4 μm
	Focal length	77 μm
	Number of microlenses	501 \times 501
Main lens	f -number	2
	Focal length	50 mm
	Image plane distance (a_1)	52.08 mm
Pickup system	Object distance (b_1)	1.25 m
	Sensor pitch	5.5 μm
	Total sensor resolution	3507 \times 3507
	Lens image resolution	7 \times 7

3. Simulation and Analysis for One-Shot 360-Degree Light Field Capturing

3.1. Analysis for the Proposed Light Field Capturing System

Four-dimensional light field information that is expressed as a 2D array form is simulated according to the LFC condition. From the light field image, the depth extraction adopting optical flow method is introduced. Based on these studies, we propose a system for acquiring 360-degree 3D information of object using LFC at a single capture. We choose the simplest structure to achieve our purpose. The proposed system consists of one LFC and a couple of mirrors. The light field information that was obtained by the system depends on the arrangement of viewing angle, mirrors, and objects. With two mirrors and one LFC, three perspectives surround and capture the object, according to their

locations and orientations. Since the 3 light field images need to be combined to restore 360-degree information of the object, an analysis of the location and orientation of the optical elements should be preceded. For analytical simplicity, the structure of the proposed system includes the two following assumptions. First, the two mirrors are large enough to reflect a given object. Second, three perspectives have the equal usage of viewing angle on image sensor. This is to ensure that the light field information taken at three viewpoints equally includes the angular information of the object.

Figure 5a shows the overall proposed pickup system. The LFC is oriented to the front of object and two mirrors are aligned to compose two additional perspectives, as described above. Therefore, the image sensor in LFC records three perspectives simultaneously. It is supposed that three light field images laid on sensor have same area on sensor. In other words, a third of image sensor is allocated to each perspective horizontally. One of the perspectives uses the field of view (FoV) θ , when the total FoV of LFC is 3θ , as shown in Figure 5b. Figure 5b only shows left part of the system for less complexity. The point S is the center of object and the point C means the center of main lens of LFC where the center of FoV exists. The point S' is an image of the point S that is reflected by the mirror and the distance between the mirror and the point S is l . The distance from the point S' to the main lens is represented as s . To make sure three perspectives have the same angle differences, two green dotted lines should form an angle of 120° , as shown as Ω in Figure 5b. The location and orientation of the mirror are determined by two factors, the length of perpendicular line from point S to point C , represented as d and the angle between optical axis of LFC and mirror surface, ϕ . From geometric analysis, the angle ϕ can be simply calculated as $\Omega/2-\theta/2$. The length l and s are determined by following equations, Equations (2) and (3).

$$l = \frac{\sin \theta}{2 \cos\left(\frac{\Omega+\theta}{2}\right)} d, \tag{2}$$

$$s = 2l \sin \phi + d. \tag{3}$$

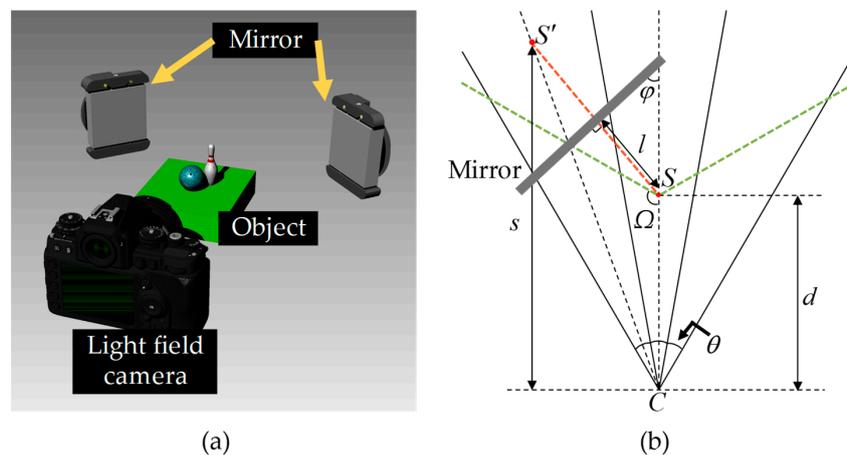


Figure 5. (a) Proposed light field pickup system and (b) geometrical figure for analysis on the proposed system.

To simultaneously capture images in S and S' , it is necessary to analyze where the focal plane of the light field camera is located. When a regular camera focuses on a specific position, the depth of field is determined based on the specification of the camera, and the information of the object can be acquired without blur in that range. On the other hand, in the case of the light field camera, the size of blur according to the distance behaves in the opposite manner to blurring mechanism of the regular camera. The blur made by light field camera appears to be small in most areas, except the near the focal plane, and the size of blur is bounded to that of the microlens in the focal plane. Analysis of blur for light field camera has been done in several researches [10,11,13]. Based on this previous research, the

blur radius of microlens is calculated according to the distance to various focal planes, which is shown in Figure 6a. The focal lengths of the main lens and the MLA are 50 mm and 0.12 mm, respectively. When the focal plane moves away from the camera, the image of the microlens becomes thick in space, and therefore, the light field information of the object that is placed near the focal plane may not be captured properly. This suggests that the method of finding the optimal focal point depends on the position of the object. According to Equations (2) and (3), the position of the mirror image can be shown in Figure 6b,c depending on the position of the object. When the object is relatively close to camera, for example, 200 mm in Figure 6b, the space that is suffered by blur is small enough to be between points S and S' . In this case, the depth of the object can be obtained properly, except when the peak of the blur covers one of the two images. However, the farther away from the peak of the blur, the greater the number of repetitions of a point on the light field image, which leads to the possibility that both of the perspectives will invade their respective regions. Therefore, in the case of Figure 6b, it is desirable that the peak be located between two objects. When the object distance is relatively large, as is the case of Figure 6c, it is impossible to place a focal plane between two objects because the width of the peak is too thick. In this case, the peak should be placed on either side of the two objects. However, as the distance from the camera increases, the number of pixels that are used for recording per unit distance decreases, which is shown as pixel per millimeter (ppm) in Figure 6b. Therefore, to record two objects in similar quality, it is desirable that the near object undergoes a larger blur. Consequently, in the case of Figure 6c, the focal point of the camera should be in front of the object. Meanwhile, the comparison between black solid lines in Figure 6b,c shows the tradeoff between the size of the photographable object and the amount of information per unit length of the recorded light field. Therefore, it should be placed at a distance from the camera that achieves the highest resolution, depending on the size of the given object.

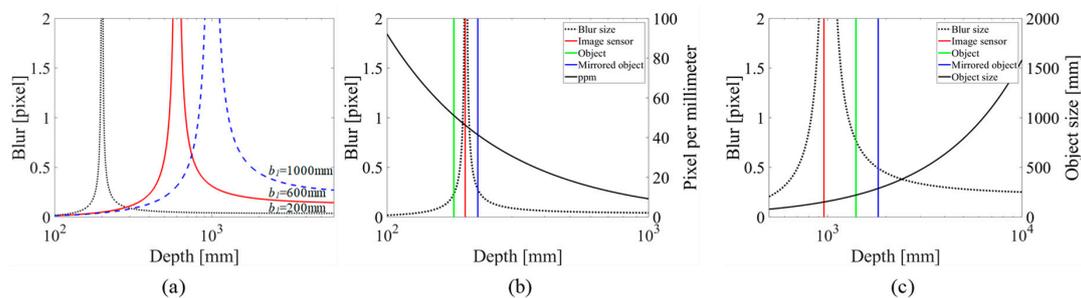


Figure 6. (a) Blur radius of microlens versus depth of scene according to various focus distances (b_1) in log scale; (b) Focal plane between the object and the mirrored image: focus distance (b_1) is 200 mm and object location is 180 mm. Solid black line represents spatial resolution of captured image in pixel per millimeter (ppm); and, (c) Focal plane in front of object: focus distance (b_1) is 1000 mm and object location is 1400 mm. Solid black line represents maximum object size to be captured.

Meanwhile, mirrors having a specific curvature can be used in the proposed capturing system. Since the mirrors in our system does not magnify images, the reflected images are located beyond mirror and occupy smaller portion of the image sensor, as compared to the center perspective light field image. To enhance light field resolution on the sensor, a pair of optical elements that magnify an object can be used, such as concave mirrors. When one uses concave mirrors to the proposed system, the location and orientation of the mirrors are determined by the distance between main lens and the object, like plane mirrors. Also, the focal length of the concave mirror adjusts the location and size of magnified image. Even though the maximum magnification is controlled by the viewing angle of the LFC, there is a possibility that the overall shape of object limits magnification value. There is another practical consideration of using concave mirrors. Because longitudinal magnification in the imaging system is not linear, in contrast to transverse direction, depth information of magnified images should be manipulated after DOLF algorithm. A shape of object should also be considered to

optimize the overall system. If the object has many wrinkles and occluded parts, some parts of object cannot be captured by an LFC. In this case, multiple mirrors, which are located at several positions toward the object, are necessary. There may happen to be another issue to use the proposed system. Though 3D scanning is the main application of the system, independent objects also can be captured on the proposed system at the same time. However, if there are more than two objects or an object that has a severely folded structure, some part of the 3D object cannot be recorded and remain as occluded parts. There are several researches about reconstruction of the occluded object in the field of integral imaging, but the occluded part to be restored is quite limited, according to the size of the object [24]. Using additional mirrors can be an effective solution to record entire space around objects. If the image sensor is large enough to spend resolution for assigning more mirrors, mirrors on the several locations can solve this limitation. If an optical configuration that makes three perspectives overlapped, the occlusion removal algorithm also can be used in this case. After sectioning three objects according to the depth of the light field image, texture information of the occlude position is reconstructed by using a combination of sub-images.

3.2. Four-dimensional Light Field Simulation Using Sub-Image Synthesis

To carry out an optical simulation for an LFC, it is necessary to define the optical elements that characterize the LFC. The 4D light field information to be captured is distributed differently, according to the focal length of the main lens, the focal length and apertures of the MLA, the pixel size of the image sensor, and the overall resolution on light field plane. When the LFC is focused at a certain depth to capture an object, the distance between the MLA and the main lens changes. According to this distance, different magnified image sensor and MLA in space capture objects. The MLA, main lens, image sensor, and the depth of focal plane determine the specifications of the magnified pickup unit. One of methods to make the 2D form of light field is to simulate a sub-image that can be extracted from the target light field. A sub-image can be created by collecting pixels in the same position in each lens image. Before taking a sub-image, the orientations of pixels in rendering software should be determined, according to the specifications of LFC [17]. From the given details of LFC, virtual cameras for sub-images are made in space as the form of Figure 1. The resolution of the camera is same as the number of microlens and the distance between pixels is lens pitch of magnified MLA. From the simulated sub-images, one can restore the light field by relocating pixels of sub-images. The center image of Figure 7a shows one of the light field images by an LFC. A total 49 sub-images, which comprise 7×7 images in horizontal and vertical direction, are merged to make light field. Resolution of each sub-image is 500×500 . The specifications of the camera used for the simulation are given in Table 1.

To verify the optical system configuration based on the above analysis, the optical simulations are performed with virtual mirrors, LFCs, and objects. Persistence of Vision Ray Tracer (POV-Ray) software is used to place the 3D information of the object and to make the light field, and the depth extraction method was performed using Matlab software. By using the LFC simulation, light field in the proposed optical capturing system is generated. Figure 6a shows the resultant light field. The specification of MLA and the main lens are also presented in Table 1. The overall condition of pickup system is also the same as the table. However, the focal plane of the main lens is slightly moved back to the camera to represent light field effectively. As shown in the Figure 6a, two additional viewpoint images are taken using a flat mirror. Because the mirror does not have any magnification, the reflected image has a smaller resolution than the front image. As mentioned above, concave mirrors can be an alternative to improve resolution of reflected light field images. With the optical flow algorithm for LFC, the depth information of the simulated light field is also extracted, as shown in Figure 7b. Note that the depth information is normalized to 8-bit information within proper depth range, according to the size of the object. To record the texture information of the light field, digital refocusing technique for LFC is used [10]. As shown in Figure 7c,d, color information of each perspective is demonstrated. Since the resolution of reflected light field has lower resolution than

that of center perspective, Figure 6c looks blurry when compared to the Figure 7d. The three pairs of resultant depth map and color information can be projected in virtual space as a form of point cloud having 3D position and color information. To combine the three light field images, the mirrored perspectives should be resized after digital refocusing. The magnification ratio between center and the others are determined by $s : d$, as shown in the Figure 5b. According to the ratio, the right/left sub-images are resized to have same scale in the results. By binding all point cloud data, the 3D object can be finally reconstructed. To transform the point cloud data to solid one, a triangular mesh is one of the alternatives to handle 3D information. Various algorithms to make mesh type object from point cloud data have already been suggested [24,29,30]. Figure 7e represents reconstructed triangular meshes of the object, which are rendered from the back of the 3D object.

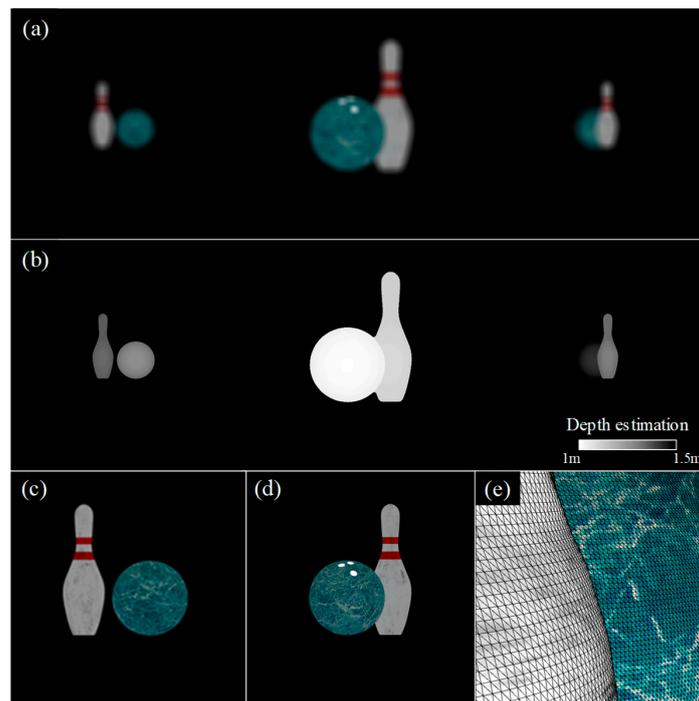


Figure 7. (a) Simulated light field image in the proposed optical setup; (b) extracted depth information of the light field; (c) refocused image of left perspective of the light field; (d) refocused image of the center perspective; and, (e) reconstructed three-dimensional (3D) triangular mesh from the color point cloud.

4. Experimental Results for the Light Field Capturing and Reconstruction

4.1. Experiments for One-Shot 360-Degree Light Field Capturing

An optical structure for 360-degree light field capturing was implemented, as shown in Figure 8a. Two plane mirrors and one light field camera were used to compose the overall system. To align each optical component, we adjusted the positions of light field camera and the stage of the subject first. The positions of two mirrors were controlled with a linear translation stage. A simple laser pointer is utilized to adjust orientation of the mirrors. With the composition of the optical elements, the position for subjects was fixed on the stage. The stage for the subject and background of the studio were wrapped with solid color sheet, which was green color in our system. To remove the background data, the chroma-key technique was used. Figure 8b,c show the captured light field data and its depth map information, respectively. The background noise was removed in the raw light field and the light field was cropped to calculate accurate depth map. The resolution of processed light field was 3864×2681 . Since the resolution of each lens image was 7×7 , the resolutions of each sub-image and depth map were 552×383 .

Figure 8d,e show left and center perspective images, which were digitally refocused to extract clear color information. In the comparison of Figure 8d,e, the center perspective had the higher resolution than the left and right one. The original resolutions of the center and left or right perspective were 1020×1020 and 510×510 , respectively. To acquire proper refocused images, involving enough depth of field was essential, according to the size of subjects. Since the f -number of the light field was 2, which provided rather low depth of field for most of objects, we used a synthetic aperture technique whose f -number is to reduce converging rays onto sensor and increase depth of field of the synthetic image. Figure 8f shows 3D color information of red box region in Figure 8e. Triangular mesh was used to illustrate the 3D structure and the location of virtual camera for rendering was adjusted to emphasize the peak of the penguin. As to the simulation results, the color point cloud was located on virtual space in the PC environment, bound altogether, and reconstructed as 3D meshes.

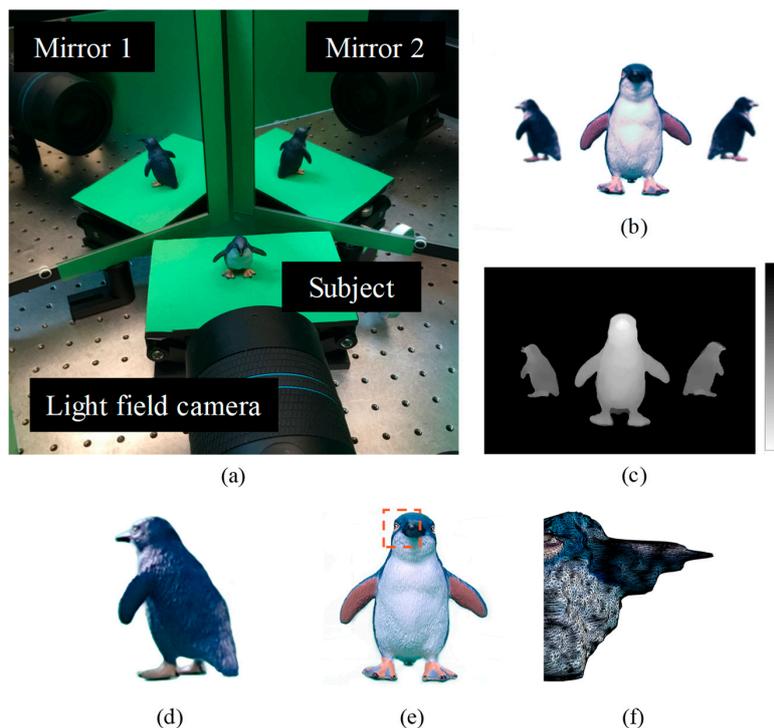


Figure 8. The 360-degree light field capturing setup and experimental results. (a) The implemented capturing system with two plane mirrors; (b) Captured light field image focused on the center perspective and processed with chroma-key technique; (c) The depth map of the light field calculated by the proposed algorithm; (d) Refocused image for left perspective; (e) Refocused image for center perspective; and, (f) Reconstructed mesh with color information.

4.2. Experiments for Optical Reconstruction of the Captured Light Field

Not only to capture 3D information with the proposed method but also to display 3D object optically, we adopted holographic display to reconstruct 3D information of the proposed capturing system. Figure 8 shows the top view of holographic display system for verifying reconstructed 3D information. Fast spatial light modulator (SLM) was used to display holographic information. The SLM, which was implemented by a digital micro-mirror device (DMD), could control the amplitude of the incident wave by toggling micro-size mirrors very quickly. The DMD was a V-9501 model of Vialux. In the principle of Fourier hologram, an amplitude hologram can be reconstructed to a complex hologram. The computer generated hologram can be made from intensity and depth information of each pixel. To enhance quality of the reconstructed hologram, random phases were wrapped on each frame of the hologram were and timely multiplexed [31]. Because the aspect ratio of DMD was

not square, reconstructed hologram in the Fourier plane shrank along horizontal direction without compensation. Anamorphic Fourier optical system, which was composed of three cylindrical lenses in Figure 9, was implemented to optically correct the distorted pixel structure and image [32]. The focal lengths of cylindrical lens 1 and 3 were 200 mm and the focal length of cylindrical lens 2 was 100 mm. The distance between adjacent cylindrical lenses was the focal length of cylindrical lens 2. Figure 10 shows the accommodation cue of the reconstructed holograms. The letter of front and rear were located on the front and rear positions of the objects.

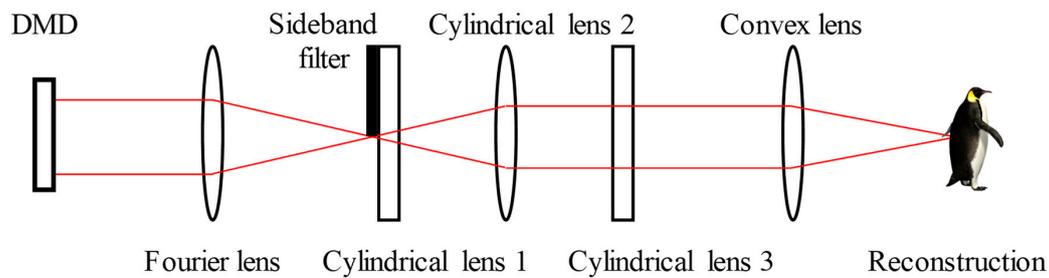


Figure 9. Top view of the holographic display system with anamorphic optical transformation model. DMD: digital micromirror device.

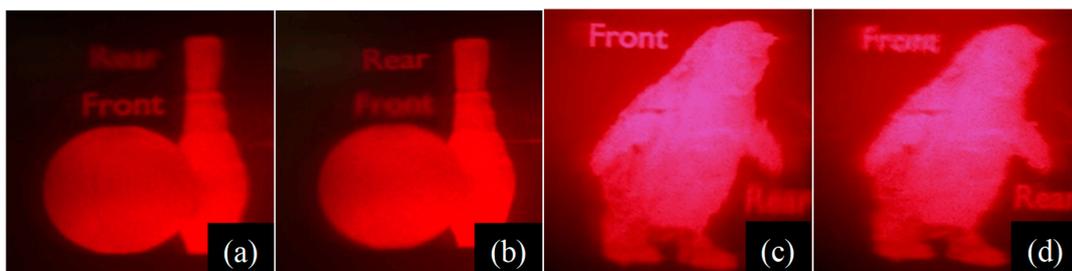


Figure 10. Optical reconstruction of captured information with holograms. Captured hologram of simulated light field focused on (a) Front (Video S1. 3D reconstruction of simulated object) and (b) Rear. Captured hologram of the captured light field focused on (c) front (Video S2. 3D reconstruction of captured object) and (d) rear.

5. Conclusions

Here, a system for capturing 360-degree light field information of one object in one shot and optically reconstructing it was proposed. To process 2D light field image to depth profile with texture, a depth extraction method based on the optical flow for light field camera (DOLF) was introduced. Maximum optical flow value and a mask for the value were introduced according to relative position difference between sub-images in light field. For 360-degree capturing, a compact and inexpensive studio that was composed of two mirrors and a light field camera was used. As a result, we captured light field information around a subject in one-shot. Experimental and simulation results were presented to support the proposed system and analysis. The resolution of the used light field was 3864×2681 and the resolution of each sub-image and depth map was 552×383 . Holographic display with anamorphic Fourier optical system and DMD was used to optically reconstruct the captured light field information.

Author Contributions: Conceptualization, Y.J. and J.C.; Methodology, Y.J.; Software, Y.J. and S.M.; Validation, Y.J., J.J. and G.L.; Formal Analysis, Y.J. and G.L.; Writing-Original Draft Preparation, Y.J.; Writing-Review & Editing, Y.J. and B.L.; Visualization, Y.J. and J.J.; Supervision, B.L.

Funding: This research was funded by Projects for Research and Development of Police Science and Technology under Center for Research and Development of Police science and Technology and Korean National Police Agency (PA-H000001).

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

Appendix A

Figure A1 introduces the derivation of Equation (1). First, Figure A1a conceptually shows the process of extracting the sub-image used for optical flow calculation from the photographed light field image. The proposed method is most commonly used in a variety of applications using lenslet arrays, for example integral photography, integral image, and light field microscopy. As can be seen from the light field image, there are five microlenses, and seven pixels per lens are assigned on the image sensor. It is assumed that the same shaded pixels have the same brightness, and two pixels of the checkered are corresponding points to each other. Sub-image creation is the process of rearranging the pixels of the lens images that are in the same position relative to the index of the lens to which they belong. Looking at the first sub-image on the far left, we can see that the leftmost pixels of each lens image are gathered together to form the first sub-image. The checkered pixels are each placed at the second pixel of the second sub-image and at the fourth pixel of the central sub-image (fourth). After the sub-images have been extracted, Figure A1b shows the process of finding the corresponding points between the two sub-images and obtaining the distance between them by using the optical flow algorithm. In general, the output of optical flow does not have its own unit, and it is necessary to fit the output value according to the situation to suit the actual situation. In this study, the optical flow values obtained from two sub-images have the pixel width in units, but there still exists area for further development. First, depending on the indices of the sub-images used, which mean their relative positions, the distances of corresponding points become different. Second, the optical flow algorithm can extract the positions of two points with sub-pixel accuracy, although the optical flow value is described as having only integers. For these two reasons, we need to analyze how the output optical flow value has a certain distance in a given situation.

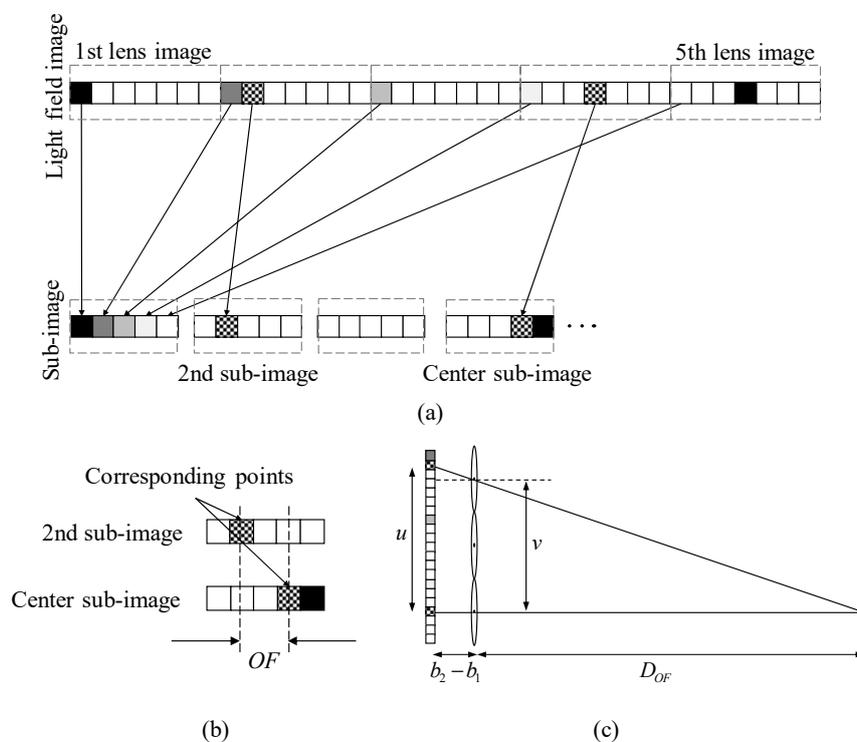


Figure A1. (a) Sub-image extraction from light field image; (b) optical flow calculation between corresponding points; and (c) schematic diagram to derive Equation (1). OF: optical flow.

The final step to derive Equation (1) is shown in Figure A1c, which is a more simplified illustration of Figure 2 in the text. Given the details of the light field camera and the information about the focal plane, we can see the size of the magnified sensor and microlens. The distance between two optical elements can also be obtained, denoted by $b_2 - b_1$. A variable v should be considered to get the size of D_{OF} . The variable v is the distance between the microlenses where the two corresponding points have belonged. From the fact that the position of each pixel in the sub-image is equal to the index of the lens image where it is included before, we can infer that v is related to the value of optical flow. For example, if the optical flow value is 2, the two pixels come from a lens image with an index difference of 2. Therefore, v is a value directly related to the optical flow, specifically, $v = L \cdot OF$, where OF means the value of optical flow and L is the pitch of the microlens considering the magnification. On the other hand, u represents the distance between two pixels on the floated image sensor. The distance between two pixels must be considered—both the distance between the lens images and the positions at which the corresponding pixels lay in the lens images. If the number of pixels allocated to one microlens is n and the size of the magnified pixel is P , then the distance caused by the optical flow difference is $P \cdot n \cdot OF$ and the term $(i_c - i)$ related to the index difference of the sub-image is added, where i represents the index of the sub-image and c represents center. As a result, by simple geometry, D_{OF} can be expressed as

$$D_{OF} = \frac{(b_2 - b_1) \cdot (O \cdot L)}{O \cdot P \cdot n + (i_c - i) \cdot P - O \cdot L'} \quad (A1)$$

which can be derived to Equation (1) by averaging on horizontal and vertical directions according to indices.

References

- Huang, F.C.; Chen, K.; Wetzstein, G. The light field stereoscope: Immersive computer graphics via factored near-eye light field displays with focus cues. *ACM Trans. Graph. (TOG)* **2015**, *34*, 60. [[CrossRef](#)]
- Lanman, D.; Luebke, D. Near-eye light field displays. *ACM Trans. Graph. (TOG)* **2013**, *32*, 220. [[CrossRef](#)]
- Hua, H.; Javidi, B. A 3D integral imaging optical see-through head-mounted display. *Opt. Express* **2014**, *22*, 13484–13491. [[CrossRef](#)] [[PubMed](#)]
- Moon, E.; Kim, M.; Roh, J.; Kim, H.; Hahn, J. Holographic head-mounted display with RGB light emitting diode light source. *Opt. Express* **2014**, *22*, 6526–6534. [[CrossRef](#)] [[PubMed](#)]
- Takaki, Y.; Urano, Y.; Kashiwada, S.; Ando, H.; Nakamura, K. Super multi-view windshield display for long-distance image information presentation. *Opt. Express* **2011**, *19*, 704–716. [[CrossRef](#)] [[PubMed](#)]
- Salvi, J.; Pages, J.; Batlle, J. Pattern codification strategies in structured light systems. *Pattern Recognit.* **2004**, *37*, 827–849. [[CrossRef](#)]
- Geng, J. Structured-light 3D surface imaging: A tutorial. *Adv. Opt. Photonics* **2011**, *3*, 128–160. [[CrossRef](#)]
- Shotton, J.; Sharp, T.; Kipman, A.; Fitzgibbon, A.; Finocchio, M.; Blake, A.; Cook, M.; Moore, R. Real-time human pose recognition in parts from single depth images. *Commun. ACM* **2013**, *56*, 116–124. [[CrossRef](#)]
- Zhu, J.; Wang, L.; Yang, R.; Davis, J.E. Reliability fusion of time-of-flight depth and stereo geometry for high quality depth maps. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 1400–1414. [[PubMed](#)]
- Ng, R.; Levoy, M.; Brédif, M.; Duval, G.; Horowitz, M.; Hanrahan, P. Light field photography with a hand-held plenoptic camera. *Comput. Sci. Tech. Rep. CSTR* **2005**, *2*, 1–11.
- Bishop, T.E.; Favaro, P. The light field camera: Extended depth of field, aliasing, and superresolution. *IEEE Tran. Pattern Anal. Mach. Intell.* **2012**, *34*, 972–986. [[CrossRef](#)] [[PubMed](#)]
- Lee, S.K.; Hong, S.I.; Kim, Y.S.; Lim, H.G.; Jo, N.Y.; Park, J.H. Hologram synthesis of three-dimensional real objects using portable integral imaging camera. *Opt. Express* **2013**, *21*, 23662–23670. [[CrossRef](#)] [[PubMed](#)]
- Levoy, M.; Ng, R.; Adams, A.; Footer, M.; Horowitz, M. Light field microscopy. *ACM Trans. Graph. (TOG)* **2006**, *25*, 924–934. [[CrossRef](#)]
- Prevedel, R.; Yoon, Y.G.; Hoffmann, M.; Pak, N.; Wetzstein, G.; Kato, S.; Schrödel, T.; Raskar, R.; Zimmer, M.; Boyden, E.S.; et al. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nat. Methods* **2014**, *11*, 727–730. [[CrossRef](#)] [[PubMed](#)]

15. Kim, J.; Jung, J.H.; Jeong, Y.; Hong, K.; Lee, B. Real-time integral imaging system for light field microscopy. *Opt. Express* **2014**, *22*, 10210–10220. [[CrossRef](#)] [[PubMed](#)]
16. Hahne, C.; Aggoun, A.; Velisavljevic, V.; Fiebig, S.; Pesch, M. Refocusing distance of a standard plenoptic camera. *Opt. Express* **2016**, *24*, 21521–21540. [[CrossRef](#)] [[PubMed](#)]
17. Jeong, Y.; Kim, J.; Yeom, J.; Lee, C.K.; Lee, B. Real-time depth controllable integral imaging pickup and reconstruction method with a light field camera. *Appl. Opt.* **2015**, *54*, 10333–10341. [[CrossRef](#)] [[PubMed](#)]
18. Martínez-Corral, M.; Javidi, B.; Martínez-Cuenca, R.; Saavedra, G. Formation of real, orthoscopic integral images by smart pixel mapping. *Opt. Express* **2005**, *13*, 9175–9180. [[CrossRef](#)] [[PubMed](#)]
19. Do, C.M.; Javidi, B. 3D integral imaging reconstruction of occluded objects using independent component analysis-based K-means clustering. *IEEE/OSA J. Disp. Technol.* **2010**, *6*, 257–262. [[CrossRef](#)]
20. Jeon, H.G.; Park, J.; Choe, G.; Park, J.; Bok, Y.; Tai, Y.W.; So Kweon, I. Accurate depth map estimation from a lenslet light field camera. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
21. Liu, F.; Hou, G.; Sun, Z.; Tan, T. High quality depth map estimation of object surface from light-field images. *Neurocomputing* **2017**, *252*, 3–16. [[CrossRef](#)]
22. Dansereau, D.G.; Mahon, I.; Pizarro, O.; Williams, S.B. Plenoptic flow: Closed-form visual odometry for light field cameras. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011.
23. Iffa, E.; Wetzstein, G.; Heidrich, W. Light field optical flow for refractive surface reconstruction. *Proc. SPIE* **2012**, *8499*, 84992H.
24. Jung, J.H.; Hong, K.; Park, G.; Chung, I.; Park, J.H.; Lee, B. Reconstruction of three-dimensional occluded object using optical flow and triangular mesh reconstruction in integral imaging. *Opt. Express* **2010**, *18*, 26373–26387. [[CrossRef](#)] [[PubMed](#)]
25. Metallo, A.; Rossi, V.; Blundell, J.; Waibel, G.; Graham, P.; Fyffe, G.; Yu, X.; Debevec, P. Scanning and printing a 3D portrait of president Barack Obama. In Proceedings of the SIGGRAPH, Studio, LA, USA, 9–13 August 2015.
26. Todoroki, H.; Saito, H. Light field rendering with omni-directional camera. *Proc. SPIE* **2003**, *5150*, 1159–1169. [[CrossRef](#)]
27. Horn, B.K.; Schunck, B.G. Determining optical flow. *Artif. Intell.* **1981**, *17*, 185–203. [[CrossRef](#)]
28. Baker, S.; Scharstein, D.; Lewis, J.P.; Roth, S.; Black, M.J.; Szeliski, R. A database and evaluation methodology for optical flow. *Int. J. Comput. Vis.* **2011**, *92*, 1–31. [[CrossRef](#)]
29. Passalis, G.; Sgouros, N.; Athineos, S.; Theoharis, T. Enhanced reconstruction of 3D shape and texture from integral photography images. *Appl. Opt.* **2007**, *46*, 5311–5320. [[CrossRef](#)] [[PubMed](#)]
30. Kim, H.; Hahn, J.; Lee, B. Mathematical modeling of triangle-mesh-modeled three-dimensional surface objects for digital holography. *Appl. Opt.* **2008**, *47*, D117–D127. [[CrossRef](#)] [[PubMed](#)]
31. Jeong, J.; Cho, J.; Jang, C.; Li, G.; Lee, B. Simple Quality Improvement Method for Holographic Display using Digital Micro-mirror Device. In Proceedings of the Imaging and Applied Optics, Heidelberg, Germany, 25–28 July 2016.
32. Kim, H.; Hwang, C.Y.; Kim, K.S.; Roh, J.; Moon, W.; Kim, S.; Lee, B.R.; Oh, S.; Hahn, J. Anamorphic optical transformation of an amplitude spatial light modulator to a complex spatial light modulator with square pixels. *Appl. Opt.* **2014**, *53*, G139–G146. [[CrossRef](#)] [[PubMed](#)]

