A Comparison of Plenoptic Imaging System and Camera Array System

Wei-jia CAI, Ju-zheng MIAO, Cheng GU, Zhuo-ying WANG, Yun-lei YU, Ping ZHOU

ABSTRACT

The imaging method based on light field data is a hot research topic at present. By recording the location and direction of light, light field imaging can reconstruct the three-dimensional scene. From the point of view of recording method, there are two similar light field recording systems, so-called camera array system and plenoptic imaging system. In order to process data flexibly, the light field information can be represented as a 4-D plenoptic function: \( L(x, y, u, v) \). Although these two systems are similar in data recording, there exist some important differences between camera array system and plenoptic imaging system when transforming raw data into 4-D plenoptic function. Based on the imaging principle of these two different systems, this paper summarizes the distinction, which will be of great help for the researchers in this field.

INTRODUCTION

The human visual system uses binocular disparity to perceive the scene by fusing two images acquired by two eyes together. Conventional photos are two-dimensional, showing the location of captured scene but ignoring the depth information. To capture three-dimensional information of object scene, we must describe and record the light transmission. Adelson\(^1\) put forward the concept of a seven-dimensional plenoptic function in 1991, which describes the light as a function of angle, wavelength, time and location of the viewer, depicted as \( P(x, y, z, \theta, \phi, \lambda, t) \). A typical model was proposed by Levoy and Hanrahan\(^2\) in 1996, using two parallel planes to describe the location and direction information of light. Levoy and Hanrahan ignored the influence of wavelength and time in Adelson’s model, and assumed that the intensity of light is constant during propagation. Therefore, the 7-D plenoptic function is simplified into a 4-D form, showed as \( L(u, v, x, y) \), where \((u, v)\) represents the direction of light and \((x, y)\) represents the location. According to recording methods, there are two similar light field information captured systems, so-called camera array system and plenoptic imaging system. Camera array system is huge and complex, aiming to capture large amounts of light field data. The cave system designed by Q. Dai and K. Lee\(^3\) at Tsinghua University, the 128 camera system designed by Stanford University\(^4,5\) are both camera array systems. The plenoptic
imaging system is simple in structure, usually with one single hand-held camera. The Lytro camera designed by R. Ng at Stanford University and the Raytrix camera are both plenoptic camera systems. A method of decoding, calibration and rectification for lenslet-based plenoptic camera was proposed by D.G. Dansereau in 2013. Although the data captured by these two systems is similar, there are some tiny but crucial differences between them when transforming the data into 4-D function \( L(u, v, x, y) \). The difference affects the follow-up work such as digital refocusing, extraction of EPI, super-resolution reconstruction, three-dimensional profile measurement and so on. This paper compares the imaging principle and the methods of light field information extraction of these two light field imaging systems. We are particularly interested in the relation between sub-aperture images and images taken by camera array.

**LIGHT FIELD CAMERA / PLENOPTIC CAMERA**

By adding a microlens array into the original image plane (CCD plane) and moving CCD backward, a plenoptic camera can capture the light field information during an exposure time. Fig.1 shows the two-parallel-plane model, where the main lens plane represents the direction of light (depicted as \((u, v)\)), and the microlens plane represents the location of light (depicted as \((x, y)\)).

The direction and location information of light is recorded with photosensor. In this way, light field information can be expressed as \( L_F(u, v, x, y) \), where \( F \) is the distance between the main lens plane and photosensor plane. Note that Fig.1 simplifies 4-D light field into 2-D, which takes into account direction \( u \) and location \( x \) only. The \( v \) and \( y \) work in the similar ways. In the main lens plane, the range from \( u_1 \) to \( u_3 \) is the aperture of main lens. In photosensor plane, the Region A is a micro image formed with light that pass through microlens \( x_1 \), which is called Element Image. Compared with conventional photography, the sampling grid of the plenoptic camera is shorter and wider (see Fig.1(b)). In other words, the plenoptic camera provides more specificity in directional information but less specificity in spatial information, assuming a constant number of photosensor pixels. It can be proved that the element images are equidistant and have the same size, as shown in Fig.2. Different element images correspond to different coordinates \((x, y)\), which record the location information of light field. On the other hand, the different pixels in an
element image correspond to different coordinates \((u, v)\), which record the direction information of light field.

![Figure 2. 4-D information in raw image.](image)

\(A: L(5, 5, 1, 1) \ B: L(5, 6, 1, 1) \ C: L(5, 5, 1, 2) \ D: L(5, 5, 2, 1)\)

It can be seen from the above analysis that the coordinate \((u, v)\) describes the location of light passing through the main lens. As for 4-D information \(L(u, v, x, y)\), if we limit the range of integration of \((u, v)\), then we get an image formed by the light passing through some sub-aperture of the main lens. These images, in many papers, are called Sub-Aperture Images (SAI). Each SAI corresponds to a certain view of the scene. \(^9\)

\[
I(x, y) = \int_{u_0}^{u_0+\Delta u} \int_{v_0}^{v_0+\Delta v} L(u, v, x, y) \, du \, dv
\]

(1)

![Figure 3. Model of sub-aperture image.](image)

**CAMERA ARRAY SYSTEM**

Another method of recording light field information is to build a camera array. As shown in Fig.4, each camera of the array is a view point of the scene. The image taken by a certain camera is a two-dimensional sample of the three-dimensional object space from the certain viewpoint. \(^10\) The light emitted in different directions from the same object point is recorded by cameras in different locations. Therefore, the location and direction information of light field can be recorded with camera array system. The bigger the resolution of camera
is, the more the location information is. The more cameras are used, the more the direction information is. Compared with plenoptic camera system, the location resolution of camera array system is much higher. However, as shown in Fig.4, these cameras are placed in different positions, with a certain distance between each other. This “arranged in sparse” feature leads to undersample problem of direction information.

![Camera array system](image)

**Figure 4. Camera array system.**

In camera array system, the image taken by each camera is a view from a particular viewpoint. The view here is similar to sub-aperture images in plenoptic camera. However, there still exists some differences between them. Therefore, it is of importance to convert a sub-aperture image to an image taken by a camera placed in the same sub-aperture. As shown in Fig.5(a), in the plenoptic camera, the light emits from the object point \( P \), passes through the sub-aperture \( S_0 \), and passes through \( p_0 \) on the photosensor finally. Note that the sub-aperture model shown in Fig.3 is applied here, which considers the microlens array and photosensor as a virtual photosensor. Suppose that a camera is placed at \( S_1 \), the light \( PS_1 \) mentioned above will get to image point \( p_1 \). It is important to find out the offset \( \Delta \) between \( p_0 \) and \( p_1 \) in order that these two sorts of images can be interconvertible.

From the similar triangles in Fig.5(a),

\[
\frac{\Delta}{D} = \frac{h_m + h'_m}{h_m}
\]

where \( h_m \) is the object distance, \( h'_m \) is the image distance. \( D \) is the distance between these two sub-apertures (\( S_0 \) and \( S_1 \)). \( \Delta \) is the offset in the photosensor between sub-aperture image point \( p_0 \) and equivalent camera image point \( p_1 \).

Thus, the offset \( \Delta \) can be written as:

\[
\Delta = D + \frac{h'_m}{h_m} D
\]

It is worth mentioning that there exists a location offset between sub-aperture image and equivalent camera image. The center of plenoptic photosensor is \( O_0 \), while the center of equivalent camera image is \( O_1 \). The offset between two photo sensors is \( D \). Therefore, the deviation of the image points on these two images can be written as \( \Delta' \):

\[
\Delta' = \Delta - D = \frac{h'_m}{h_m} D
\]
It can be shown using Eq.(3) and Eq.(4) that $\Delta$ and $\Delta'$ are related to object distance, image distance and the distance between two sub-apertures, having nothing to do with the position of the object point in the world focal plane. Moreover, for an object point $p'$ in Fig.5(b) that is not in the world focal plane, the segment $P'S_1$ or its reverse extension line will pass through a point in the world focal plane. So its offset is still $\Delta$ and $\Delta'$. In other words, the derivation above (from Eq.(2) to Eq.(4)) actually corresponds to the light $PS_1$, and the object point can be at any point in line $PS_1$. In summary, for any object point in space, when the optical parameters (object distance, image distance, focal length, etc.) of the photographing system are determined, the deviation of the image point on the sub-aperture image and the equivalent camera image is determined uniquely. That is, all the points have the same amount of deviation. Therefore, the sub-aperture image and the equivalent camera image can be easily interconverted by shifting the offset $\Delta'$ (transformed into pixel unit). However, as the sampling range of these two images isn’t the same, pixels on the edge might not be matched. Besides, if there are more than two images captured by camera array and we want to convert them into sub-aperture images, taking central camera as the central sub-aperture is a common method, which minimizes the loss of edge pixels.

What’s more, it can be shown using Eq.(4) that $\Delta' \rightarrow 0$ when $h_m \rightarrow \infty$. That is, if a plenoptic camera focuses at infinity, then its sub-aperture images are the same to the equivalent camera images. In other words, images taken by camera array are sub-aperture images taken by plenoptic camera which focuses at infinity. Note that the term “focus” we mentioned above is a feature of the main lens. The light field data should be processed to simulate the original imaging results of the main lens. By adding together all sub-aperture images, we will get an image focusing at the original world focal plane. Moreover, if we shift and add sub-aperture images, an image that focuses at a different depth can be acquired. Actually, these “shifted sub-aperture images” can be considered as a new set of sub-aperture images. In short, sub-aperture images can be changed with post-processing: one focus depth corresponds to one set of sub-aperture images. The equivalent images taken by camera array are the sub-aperture images whose main lens focus at infinity.

**CONCLUSION**

Based on the imaging principle of different light field acquisition systems, this paper compared the plenoptic camera imaging system to the camera array imaging system. The differences of recording the 4-D light field information (location information and direction information of light) were studied in detail. This paper studied the sub-aperture image in
plenoptic system and the extraction method for it. We pointed out that the sub-aperture image is similar to the image taken by a camera placed at the same position of the sub aperture, but they are different in some ways. Finally, from the point of view of deviation correction of image points, the relationship between sub-aperture images taken by plenoptic camera and images taken by camera array was deduced. Through this paper, we have made a connection between these two light field imaging systems, which is helpful to improve the universality of the light field data processing algorithm.

REFERENCES