Enhanced depth of field and resolution three-dimensional integral imaging by variable spatial filtering and intermediate-view reconstruction technique

JUEZHEN HUANG, PENG GE*

School of Physics and Opto-electronic, South China University of Technology Guangzhou, 510640 Guangzhou, China

*Corresponding author: scpge@scut.edu.cn

In this paper, we propose a novel three-dimensional (3D) integral imaging system to simultaneously improve the depth of field (DOF), resolution, and image quality of reconstructed images by variable spatial filtering and intermediate-view reconstruction technology (IVRT). In the proposed method, the camera performs element images acquisition on a 3D scene with objects of different depths through a 2D grid plane. The reconstructed slice image and block matching algorithm are used to extract the depth of the element images. To improve the sharpness of depth, the Laplace operator is used to perform variable depth filtering on objects of different depths, and depth-enhanced all-filtering element images are obtained through simple pixel fusion. IVRT is applied to all-filtering element images to obtain more element images to reconstruct a resolution-enhanced 3D image. According to the energy of gradient (EOG) value and the Tenengrad value, the reconstruction image quality evaluation of the proposed method is improved by 7.63 and 4.81 times compared with the traditional method, respectively. By the proposed method of generating all-filtering element images and an IVRT in 3D integral imaging system, the experimental results demonstrate that the 3D reconstructed image has extended depth of field, enhanced resolution and improved image quality.

Keywords: integral imaging, depth of field, variable filtering.

1. Introduction

With the development of science and technology, 3D image display technology has become a hot topic in recent years. In the coming era of 5G networks, integrated imaging is one of the most promising 3D display technologies [1–8]. It can provide full-color, full-parallax and continuous-view 3D integrated images without special glasses equipment assistance and coherent light source. And it is widely used in 3D TV, visual display and film-television media industries [9].
Integral imaging technology is a three-dimensional display technology proposed by Lippmann, the Nobel Prize winner in 1908. This integral imaging system is mainly divided into two stages: pickup and reconstruction. In the pickup part, it uses a micro-lens array to pick-up light fields of target objects in a three-dimensional space scene. Each micro-lens obtains a two-dimensional image, called element image (EI). And finally we obtain an element image array (EIA) with different parallax. In the reconstruction stage, according to the principle of a reversible optical path, the pick-up two-dimensional element images and micro-lens array are used to reconstruct a three-dimensional scene on the imaging plane to produce a three-dimensional integrated image. Although 3D integrated imaging technology has its own unique advantages, it also has some defects, such as narrow observation field, low resolution and short depth of field for 3D reconstruction.

The limited depth of field is one of the main defects of 3D integrated imaging technology. In order to improve this defect, many scholars have proposed many methods [10–16]. Among them, a synthetic aperture integrated imaging system [10, 11] that pickups high-resolution element images by moving the camera on a two-dimensional grid is proposed. Although it improves the depth of field and resolution of the SAIL system, the camera lens still limits the depth of field. When the camera focuses on the object plane, the focus target is clear, but objects of other depths will be out of focus, and the clarity and details will become blurred or even lost. MARTÍNEZ-CUENCA et al. [12] proposed the technique of using the amplitude modulation sensor array to improve the depth of field. JANG and JAVIDI [13] proposed the use of lens array methods with different focal lengths and aperture sizes to increase the depth of field, while using the mobile array lens technology (MALT) to overcome resolution degradation. The multi-layer display device proposed by Kim et al. [14] replaced a single display to generate multiple central depth planes to enhance the depth of field. Zhang et al. [15] used a negative lens array to enhance the depth of field in a direct projection imaging system. Although these methods can improve the depth of field of the II system, the manufacture of these devices is expensive and cumbersome, and it cannot effectively enhance the resolution. Xing et al. [16] proposed a depth-enhanced integral imaging system based on spatial filtering, which proved that spatial filtering can effectively improve the image quality of deep objects. However, the spatial filtering is uniform and immutable, and the filtering intensity of the image cannot be increased correspondingly according to the change in the depth of the object, which makes objects with different depths unable to obtain better filtering effects.

At the same time, in the reconstructed image of integral imaging, the method of improving the resolution has been extensively studied [17–19]. In integrated imaging, the resolution of the reconstructed image can be improved by overlapping enough element images on the reconstructed image plane. Jang and Javidi [17] proposed a moving array-lens technology to improve the resolution of the reconstructed 3D image. In this system, the sampling rate of the element image can be improved by the rapid movement of the micro-lens array and the rapid acquisition of the element image by the time
multiplexing scheme, so as to improve the image resolution. But, such high frequency movement will generate mechanical noise and will be more difficult to implement in actual implementation. Therefore, PARK et al. [20] proposed a resolution-enhanced integrated imaging method based on the intermediate-view reconstruction technology. It allows the number of element images to be simply increased. By using only a limited number of original element images, many intermediate element images (IEIs) can be digitally synthesized as needed to reconstruct a three-dimensional image with enhanced resolution. It only needs to use a limited number of original element images to synthesize many intermediate element images to expand the element image array and reconstruct a three-dimensional image with enhanced resolution.

Thus, in this paper, in order to further improve the imaging quality of integrated imaging 3D display, while improving the depth of field and resolution of the reconstructed image, we propose a novel integral imaging system with variable spatial filtering and intermediate-view reconstruction technology (IVRT). In the proposed method, we first use the synthetic aperture method to collect images of three-dimensional objects, and use the bicubic interpolation method to increase the resolution of element images to facilitate image processing [21]. The depth of the element image and its image are extracted by using segmentation and matching methods [22]. Then, we apply reasonable depth filtering to the extracted depth objects, and fuse the filtered objects together to form an all-filtering element image. This filtering method is applied to all element images to form an all-filtering element image array. In addition, the IVRT is used to generate more all-filtering element images. Finally, the computational integral imaging technology [23–25] is used to reconstruct the 3D integrated image with improved resolution, depth of field and image visualization. In order to prove the feasibility of the proposed system, we conducted experiments with test objects, and discussed and evaluated the results of the experiments. Figure 1 depicts a block diagram of the proposed method, which consists of five processes.

2. Proposed method

2.1. Depth image extraction method of element image

After pickup 3D objects by synthetic aperture integral imaging (SAII), we carry out depth extraction of on the element image to obtain a depth image, and finally obtain the extraction image of each object. In SAII, we use slice matching to extract 3D object depth from element images. Among them, when we reconstruct the 3D image, the reconstructed image with different focusing effects will be obtained depending on the reconstructed distance. We call it a slice image. The object at the reconstructed distance
position will be clear, while the object at different depth distance will be relatively fuzzy. Therefore, according to this principle, we use the new block matching algorithm to obtain the depth image of the element image as shown in Fig. 2. We first select the block image with the same size and position from the reference element image and the slice image, as shown in the blue block and the red block in Fig. 2. Then, we carry out a new block matching algorithm for these two images to minimize the calculation value of the matching error between the block image of the reference element image and the block image of a group of slice images, so as to determine the depth position of the current block image. Here, we measure the accuracy of block matching by using the sum of absolute differences (SAD):

\[
SAD(x, y, z) = \sum_{i=1}^{b} \sum_{j=1}^{b} \left| B_{\text{ref}}(x + i, y + j) - B_z(x + i, y + j) \right|
\]  

(1)

where \( b \) represents the size of the block image, while \( B_{\text{ref}} \) and \( B_z \) represent the reference element image and the slice image, respectively.

According to this matching algorithm, we can select the \( z \) value from the minimum independent variable of \( SAD(x, y, z) \) results, and extract the depth of each image \((x, y)\). The mathematical formula can be expressed as:

\[
\hat{z}(x, y) = \arg\min SAD(x, y, z)
\]  

(2)

Next, the depth map can be obtained by calculating the depth of all the blocks in the reference image. Finally, we extract the images of objects at various depths to obtain the images of each object with depth information, which can be used in the following variable filtering process to improve the quality of the depth of field.
2.2. Generation of all-filtering element images

In the next step, after obtaining the image of each object with depth information in the element image, we need to filter it from different degrees. Due to the limitations of the camera’s aperture and focal length, when we collect element images with depth, the focused objects will be clearer while other objects will be blurred because they are far from the focal plane. This blur will become more and more serious as the distance increases. The obtained image display will also become smooth, and the image recognition will be reduced, which will affect the image quality. Therefore, the Laplace filter is used to filter the images of objects with different depths to appropriate degrees. The farther away the object is from the focused plane, the stronger the filtering intensity is needed.

Figure 3 depicts a scheme of using variable Laplace filtering to filter EI with different depth of different intensities to obtain all-filtering EI. The star is in the plane of focus, and the heart is in the out-of-focus plane. After extracting their depth object images, according to their depth information, we perform slight filtering, medium filtering and strong filtering on the star, circle and heart, respectively, so that they can obtain more reasonable clearing effects. Finally, we merge them together to generate an all-filtering EI. Here, the Laplace matrix operator $La$ we use is:

$$La = D \times \frac{1}{25} \times \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 24 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$ \hspace{1cm} (3)$$

where $D$ is the intensity parameter of the filter operator.

![Fig. 3. Schematic diagram of the variable filtering method.](image-url)
The La operator can detect the texture in the element image to obtain the texture image. We add the pixel value of the original EI and the texture image to obtain the filtered EI by the corresponding $D$ intensity. The filter matrix $Fla$ used is:

$$Fla = D \times \frac{1}{25} \times \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 24 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We change the degree of filtering by changing the value of $D$. The value of $D$ determines the strength of the image contour information. The optimal value of $D$ is affected by many factors. First, the original physical limitations of the camera, the focal depth limitation of the lens and the size of the CCD will affect the quality of the image and the blur degree of out-of-focus objects. Then, different objects will have different blur effects due to different defocus distances. It is difficult to determine that the same $D$ value has the best filtering effect on different objects. Finally, users have different requirements for the visual effect of the filtering effect according to their own needs. However, the principle that the value of $D$ must follow is that the farther the object is from the focal plane, the larger the value of $D$ can be to achieve better results. Therefore, the distance from the focus is the biggest factor affecting the $D$ value. In order to determine the value of $D$ and illustrate the effectiveness of the proposed method, we adopt the linear growth method. First, we assume that $L_n$ is the distance between the focused object and the out-of-focus object, $D_n$ is the $D$ value of each out-of-focus object, and $n$ is sorted from near to far with the out-of-focus distance, indicating the $n$-th out-of-focus object. We adjust the value of $D$ to make the defocused object at the distance of $L_1$ achieve the filtering effect we want most. That value of $D$ is $D_1$. Then, the $D$ value of out-of-focus objects can be determined as:

$$D_n = \frac{L_n}{L_1} \times D_1, \quad n = 1, 2, 3, ...$$

In addition, in order to obtain a better visual effect and make the image quality more coordinating, we can perform appropriate filtering called $D_0$ on the focused object.

By applying this method to all EIs, an all-filtering EIA can be obtained. In this way, the method of variable Laplace filtering according to the depth of the object can alleviate the problem of losing the image stripes and contours due to the smooth blur of the defocused image, improve the clarity and recognition of the image, and enhance the depth of field and visual effect of the reconstructed image.

### 2.3. Intermediate-view reconstruction technology (IVRT)

In this paper, we adopted IVRT to further improve the resolution of the reconstructed object image and optimize the imaging quality [20, 26, 27]. That is to say, we use the
all-filtering method to improve the depth of field and clarity of the reconstructed image, while IVRT increases the resolution of the image. In fact, the resolution of the reconstructed image largely depends on the number of element images collected from the 3D object. In general, the larger the number of element images, the higher the resolution of the image. Therefore, we can enhance the resolution by increasing the number of collected element images. In order to avoid physical limitations that may limit the number of EIs and the number of lenses, we used IVRT to calculate the required number of IEIs based on the original number of EIs, so as to increase the total number of EIs.

Figure 4 shows the IEI synthesis process using IVRT from the captured EIs in the horizontal and vertical directions. Finally, we can use volume calculation integrated imaging reconstruction technology for digital reconstruction.

3. Experimental results

The experiment is carried out by the software 3ds Max. Figure 5 shows the experimental setup. In the experiment, we use a simulated camera with focal length of 50 mm and sensor size of 35 mm to capture 3D light field scenes of three 3D moons at different depths. The three moons are separated from the camera array by 90, 95, and 100 mm, respectively. The camera array is composed of a 5 × 5 scale, and each camera can pickup element images with a size of 640 × 480 pixels. The distance between the horizontal and vertical directions of the camera array is 1 mm, and the target distance for pickup is 90 mm on the nearest moon. In order to improve the resolution of the initial element image and facilitate observation and calculation, the bicubic interpolation method is used to expand the element image by 5 times, that is, each element image is enlarged to 3200 × 2400 pixels.

The images of the moon at three different depths were extracted and filtered with appropriate intensity. According to the definition and determination method of $D$ value,
Fig. 5. The experiment setup: (a) 5 × 5 camera array, (b) the captured 3D scene, and (c) top view of the capture stage.

Fig. 6. (a) The all-filtering element image obtained by simple fusion of the filtered depth image and (b) the unfiltered element image.
we can perform $D_0 = 10$, $D_1 = 20$ and $D_2 = 40$ Laplace filtering for the three moons at 90, 95 and 100 mm, respectively. Figure 6 shows the all-filtering element image and the unfiltered element image obtained by simple fusion of the filtered depth image. Obviously, the image quality of the all-filtering element image is better than that of the non-filtered element image. According to this method, we can obtain a $5 \times 5$ all-filtering element image array.

In order to further improve the image resolution, we used IVRT to expand the number of element images and extend the $5 \times 5$ EIA to $9 \times 9$ EIA. Finally, we use CIIR technology to perform 3D reconstruction on the $5 \times 5$ non-filtered EIA, the $5 \times 5$ all-filtering EIA, and the $9 \times 9$ all-filtering EIA at a distance of 90 mm. Figure 7 shows these three reconstructed images. Figure 8 shows three enlarged images of the reconstructed image obtained by the proposed method and the traditional method. It can be seen from this that the image reconstruction of the $5 \times 5$ and $9 \times 9$ all-filtering EIA does not change much. However, compared to the reconstructed image of the traditional $5 \times 5$ non-filtered element images, the stripes and outlines information on the lunar surface of the other two reconstructed images is richer, more legible, and the depth effect is better. This proves that the proposed method can significantly improve the depth of field, resolution and visualization quality of the reconstructed image.

![Images](image_url)

Fig. 7. Reconstructed object images at 90 mm by different EIA: (a) $5 \times 5$ non-filtered EIA (traditional method), (b) $5 \times 5$ all-filtering EIA, and (c) $9 \times 9$ all-filtering EIA (proposed method).
In order to further verify the imaging quality of the proposed method, we use the
energy of gradient (EOG) function \[28\] and the Tenengrad function \[29\] to evaluate
the quality of the 3D reconstructed image. Here, the EOG function is the sum of squares
of the differences between the gray values of adjacent pixels in the \(x\) direction and the
\(y\) direction as the gradient value of each pixel. The sum of all pixel gradient values as
the definition of evaluation function value is shown below:

\[
F = \sum_{x} \sum_{y} \left\{ \left[ f(x+1, y) - f(x, y) \right]^2 + \left[ f(x, y+1) - f(x, y) \right]^2 \right\} \tag{6}
\]

The Tenengrad function uses the Sobel operator to extract the gradient values of
the pixels in the horizontal direction and the vertical direction. It is defined as the sum
of the squares of the pixel gradients and sets a threshold \(T\) for the gradient to adjust
the sensitivity of the function. The expression is as follows:

\[
F = \sum_{x} \sum_{y} \left[ G(x, y) \right]^2, \quad G(x, y) > T \tag{7}
\]

where \(G(x, y)\) is the gradient at the pixel \((x, y)\):

\[
G(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)} \tag{8}
\]
where $G_x(x, y)$ and $G_y(x, y)$ are the gradient values of the pixels in the horizontal and vertical directions.

The Table shows the summary table of the reconstructed image quality assessment results calculated by using the EOG function and the Tenengrad function in the traditional non-filtering method, 5 × 5 all-filtering method and 9 × 9 all-filtering method. From the evaluation summary of the EOG function, the evaluation quality of the proposed method can be improved by 7.63 times over the traditional method, which is 6.08% higher than the all-filtering method without IVRT technology. On the other hand, from the results of the Tenengrad function, the image quality of the proposed method can be improved by 4.81 times in the traditional method, which is 6.22% higher than the 5 × 5 full filtering method. In summary, the simulation results verify that the method of using the variable full-filter element image and IVRT can significantly improve the resolution, depth of field, image visualization quality and image evaluation quality of the 3D reconstruction image, and optimize the imaging effect.

### 4. Conclusions

In this paper, we proposed a novel 3D integrated imaging system that can simultaneously improve the depth of field, resolution and image quality by combining the variable full filtering method and the intermediate-view reconstruction technology. In the proposed method, we obtain the depth object image by using the slice matching to extract the three-dimensional object depth of the element image, and then combine the variable Laplace filtering to obtain the all-filtering element image array. The result is that the image of each depth of the element image becomes clear and visible, and the effect of the depth of field is enhanced. Then, the IVRT technology is used to extend the all-filtering element image array, and the CIIR technology is used to obtain a reconstructed image with improved depth of field, resolution and image quality. Finally, according to the calculation of the EOG and the Tenengrad image evaluation function, the reconstructed image quality obtained by the proposed method is 7.63 and 4.81 times higher than that of the traditional method, respectively. From the experimental results and analysis, we prove the feasibility of the proposed method.

---

Acknowledgements – This work was supported by the Natural Science Foundation (61805082), the Key Technologies R&D Program of Guangzhou City (201704020038, 201704020182, 201803030008), and the Fundamental Research Funds for the Central Universities (2019NS125).
References


[29] Liang X., Analysis and improvement on digital refocusing sharpness evaluation function of light field photography, Electro-Optic Technology Application, Issue 6, 2015, p. 56.

Received September 10, 2020
in revised form December 1, 2020