Imaging and Radiation Research (2021) Volume 4 Issue 1
doi:10.24294/irr.v4i1.1738

ORIGINAL RESEARCH ARTICLE

Focused algorithms of X-ray imaging on the basis of multi-line scanning

Ke Li[1], Bin Liu[1*], Lipeng Wang[2], Xinyu Zhang[1], Jixing Guo[3]

[1*] *Shanxi Key Laboratory of Information Detection and Processing, North University of China, Shanxi 030051, China. E-mail: liubin414605032@163.com*

[2] *Beijing Times Minxin Technology Co. Ltd, Beijing 100076, China.*

[3] *North Automatic Control Technology Institute, Shanxi 030051, China.*

ABSTRACT

In the process of X-ray transmission imaging, the mutual occlusion between structures will lead to the image information overlap, and the computed tomography (CT) method is often required to obtain the structure information at different depths, but with low efficiency. To address these problems, an X-ray focused on imaging algorithm based on multi-line scanning is proposed, which only requires the scene target to pass through the detection area along a straight line to extract multi-view information, and uses the optical field reconstruction theory to achieve the de-obscured reconstruction of the structure at a specified depth with high real-time. The results of multi-line scan and X-ray reconstruction of the target show that the proposed method can reconstruct the information of any specified depth layer, and it can perform fast imaging detection of the mutually occluded target structures and improve the recognition of the occluded targets, which has a good application prospect.

Keywords: Imaging System; Focusing Imaging; X-ray; Multi-line Scanning; De-overlapping

ARTICLE INFO

Received: 5 September 2021
Accepted: 15 October 2021
Available online: 22 October 2021

1. Introduction

In X-ray transmission imaging systems, the transmission ray information received by the detector is a superposition of the attenuation coefficients during transmission, so that the different levels of structural information cannot be effectively distinguished from the projected image[1–4]. In order to obtain information about the different levels inside the target object, conventional X-ray 3D imaging techniques and computerized tomography are usually used. The traditional X-ray 3D imaging technique, namely computed tomography (CT)[5] technique, is based on multiple angular irradiation, acquiring a large number of projection maps of the sample from multiple angles, and combining the corresponding CT reconstruction algorithm and absorption liner or phase liner imaging techniques to achieve 3D reconstruction of the sample[6]. The accurate reconstruction of samples requires complete projection data, and the low efficiency of data acquisition and reconstruction greatly limits the application of CT imaging techniques for real-time inspection. In order to simplify the CT imaging process and improve the efficiency of imaging, researchers have actively conducted research on reconstruction algorithms based on undersampled projection data, using optimization methods such as full variational and regularization for statistical reconstruction of undersampled data. Although the numerical simulation results proved that about 20 projections were sufficient to recover some relatively simple targets, the reconstruction results of actual experimental data showed that more than 100 views were required for reconstruction using regularized optimization algorithms[7,8].

Gondrom *et al.*[9] proposed a computed tomography (CL) method, which has two scanning modes: Linear and rotational. The linear mode uses an X-ray source and a detector, both of which move in opposite directions at the same time. The position of the internal point projection of the target changes with the relative motion, while the point on the focal plane is always at the same position of the detector, so that the target layer structure on the focal plane can be clearly imaged. However, in this method, one opposite motion can only achieve a clear image of the structure of the target layer at one depth, and multiple relative motions are necessary to obtain information about the multilayer structure of the target. Although the rotation type can save time and improve the imaging quality, it is not easy to detect in a limited space for targets with relatively large length and thickness, and for some specific scenarios, such as the detected object is not easy to rotate, and full-angle measurement of the detected object cannot be achieved[10,11].

To address these problems, this paper proposes a new method of focused X-ray imaging based on multilinear scanning, which can achieve clear imaging of different depth layers. The method only requires the scene target to pass through the detection area along a straight line to obtain a multi-view image, and the limited multi-view image combined with the synthetic aperture technique in optical field imaging[12–17] can achieve focused imaging of the detection target at different depths and achieve the "see-through" imaging effect of the target information on the focused surface. In order to eliminate the influence of the scattered focus region on the focused imaging and improve the quality of imaging, the best focus layer is determined by the focus degree judgment algorithm, and the image de-overlap reconstruction is realized by the image focus region extraction algorithm. The experimental results show that the proposed method can effectively use a small number of multi-view images to achieve the reconstruction of detection targets at different depths, which can be used for fast de-overlapping imaging detection of target structures and improve the recognition of targets, and has a good application prospect.

2. Focused X-ray imaging method based on multi-line scanning

Since the conventional X-ray imaging system is time-consuming to set up and not conducive to scan imaging in limited space, an improved line scan X-ray acquisition system is used to provide more views of the target to obtain better reconstruction. The system consists of an area-scan X-ray imaging system and a mechanical device that transports the inspected object at a constant speed, with a detector in the X-ray imaging system that extracts a small number of sensor lines. The different sensor lines of the area scan X-ray imaging system capture the target point at different viewpoints, and the corresponding sensor line at each moment acquires the target information and stores it. Over time, multiple multi-view images can be acquired by collecting information from multiple sensor lines, i.e., multi-view images of the target can be acquired. **Figure 1** illustrates the proposed multi-line scanning X-ray imaging system, where a, b, c, ..., n, o, p denote the acquisition regions in the target corresponding to the sensor lines, and line *n* denotes the nth sensor line, and different sensor lines acquire the target information from different angles simultaneously. Since the X-ray imaging system is in relative mo-

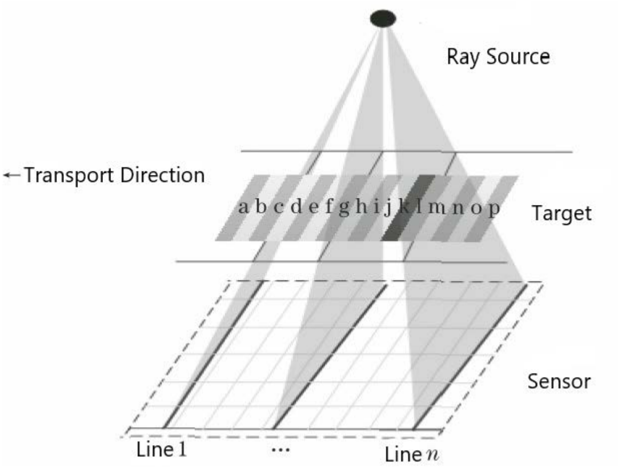


Figure 1. Multi-line scanning imaging system.

tion to the object, multiple views of a single object are collected over time based on the target object information observed from all angles.

Figure 2 shows the multi-line scanning imaging process, because the acquisition frequency of sensor lines in the imaging system should match the object motion, the acquisition period of this paper is 3 s, i.e., $\tau = 1, 2, ...$ denotes the acquisition of target information by sensor lines at 3 s, 6 s, ... moments respectively, where $\alpha_1, \alpha', \alpha_n$ denotes the first denotes, the observation angle of the 1st sensor line and the center sensor line. And observation direction of the center sensor line is perpendicular to the direction of target movement, i.e., $\alpha' = 0°$. Each sensor line of the X-ray area scan sensor scans the target at a different observation angle, i.e., multi-view imaging. For example, the viewpoint 1 image shown in **Figure 2b** is the image information of object region "k" acquired at $\tau = 1$. The reference viewpoint image shown in **Figure 2c** is the image information of object region "k" acquired at $\tau = 6$; and the viewpoint n image shown in **Figure 2d** is the image information of object region "k" acquired at $\tau = 6$; and the reference viewpoint image shown in **Figure 2e** is the image information of object region "k" acquired at $\tau = 6$. The viewpoint n image shown in **Figure 2d** is the image information of object area "k" acquired after several moments. In this way, the multi-view image information of object area "k" at different viewpoints can be acquired. As time passes, the complete image of the object in different viewpoints is acquired based on each extracted sensor line, i.e., the multi-view image.

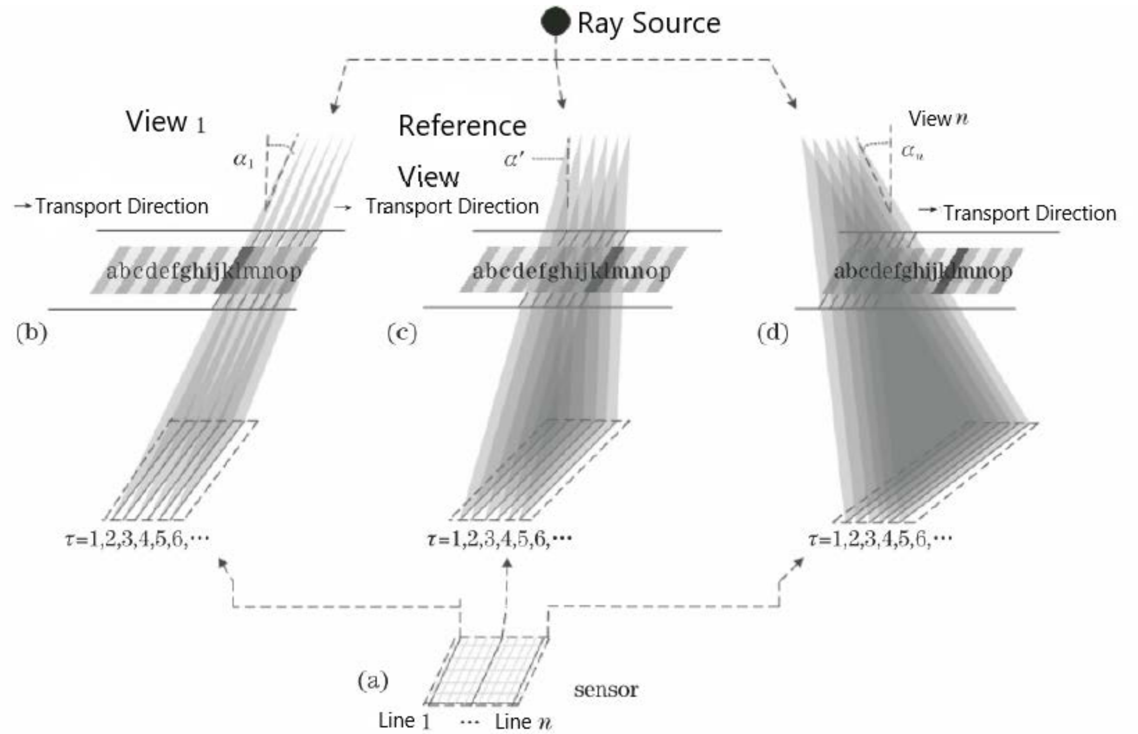


Figure 2. Multi-view imaging process.

The above acquisition system can be used to acquire images from multiple viewpoints. The image acquired by the central sensor line is the reference view, and the images acquired by the sensor lines at the same distance apart are the images under multiple views 9 of the multi-view images are extracted, and the images are shown in **Figure 3**.

The multi-view images obtained by extracting different sensor lines are shown in **Figure 3**, where the left side shows the nine extracted multi-view images, and transparency is applied to the images to show the target objects at different positions in the images. The right side shows the first view, the reference view (center view) and the image at the ninth view, respectively.

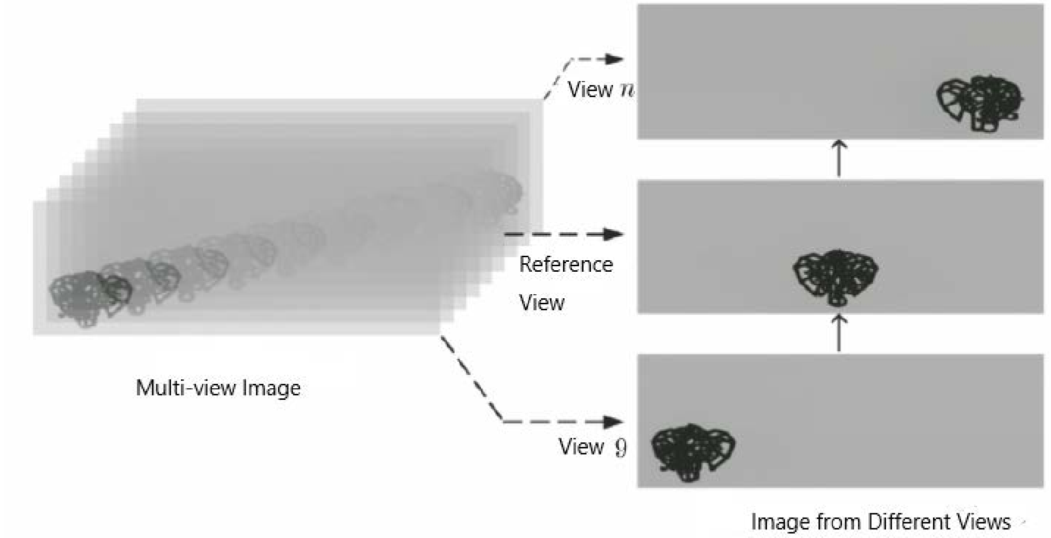


Figure 3. Multi-view images.

In the light field imaging, the multi-view imaging method is able to acquire information of the target from different directions, and the synthetic aperture focusing imaging algorithm is used to change the depth of field by virtualizing a large aperture, i.e., expanding the aperture and reducing the depth of field, so that the synthetic image has a very shallow depth of field. When the target is blocked by the foreground, the virtual camera is able to make the out-of-focus foreground blocked objects diffuse, and the target objects in the rear view stand out, achieving the effect of de-focusing the foreground and "seeing through" the target objects[18–20]. The imaging principle is as follows.

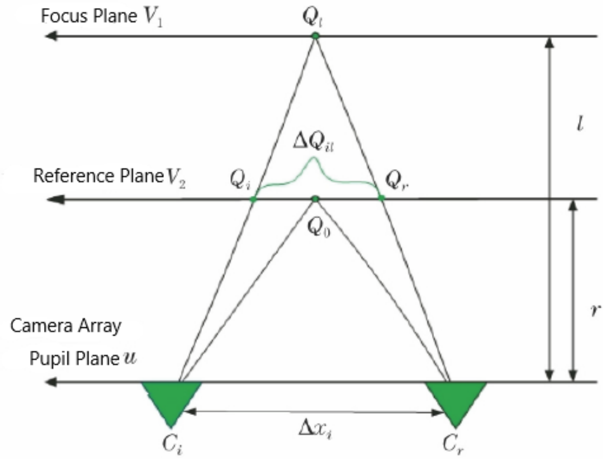


Figure 4. Schematic of synthetic aperture imaging.

As shown in **Figure 4**, a focus plane V_1 and a reference plane V_2 are defined to be parallel to the camera pupil plane, where l and r are the distances of the focus plane V_1 and the reference plane V_2 respectively, with respect to the camera. Assuming that the line array consisting of N cameras is C_1, …, C_i, …, C_N, the camera at the center of the line array is selected as the reference camera C_r, and the angle of view corresponding to this camera is the reference angle of view. Q_i is the subview image pixel point of C_i on the reference plane V_2, and i denotes the first camera i. Assuming that all cameras in the camera array are focused on the reference plane V_2, a point Q_0 will appear at the same position in each viewpoint image, and the images in each viewpoint will be superimposed and averaged to form a single image, then Q_0 pixel points will be focused in the composite image, and the image will be sharp and clear. For the pixel point Q_i on the plane V_1, the position of the image corresponding to the reference plane under the cameras C_i and C_r is Q_i and Q_r, and in synthetic aperture imaging, because the position of the image of the same object point under different cameras is not the same. Q_i pixel points are not focused on the composite image, and the image appears blurred. In order to refocus Q_l, the aberration ΔQ_{il}, where $\Delta Q_{il} = |Q_i - Q_r|$, Q_i, Q_r are the point vectors of Q_i, Q_r respectively, must be eliminated.

According to the triangle theorem, we can see that

$$\Delta Q_{il} = \Delta x_i \cdot [(l - r) / l] \tag{1}$$

where Δx_i is the distance between the pupils of the camera C_i and the pupil of the reference camera C_r;

$(l-r)/l$ is the relative depth of the refocusing plane. Then, when all cameras in the line array focus on the pixel Q_l in the V_1 plane, the aberration of the pixel Q_l with respect to the reference view can be calculated, i.e.,

$$[\Delta Q_{1l}, \cdots, \Delta Q_{il}, \cdots, \Delta Q_{Nl}]^{\mathrm{T}} = [\Delta x_1, \cdots, \Delta x_i, \cdots, \Delta x_N]^{\mathrm{T}} \cdot [(l-r)/l] \tag{2}$$

Therefore, in synthetic aperture imaging, the images of each sub-view in the reference plane must be superimposed and averaged in order to bring the target into focus. For example: According to ΔQ_{iz} panning image $I_{i,r}$, so that it focuses on the plane at depth z, because this paper uses a line array camera to acquire images, that is, in the way the target object moves in a straight line to acquire images, then the relationship equation is

$$I_{i,z}(s',t) = I_{i,z}[s + \Delta Q_{iz}(s), t] = I_{i,r}(s,t). \tag{3}$$

that is,

$$I_{i,z}(s',t) = I_{i,z}\{s + \Delta x_i \cdot [(l-r)/l], t\} = I_{i,r}(s,t) \tag{4}$$

where: $I_{i,r}$ is the image of the reference viewpoint in the synthetic aperture; $I_{i,z}$ is the image of the sub-viewpoint at the depth position z; (s,t) is the coordinate of the pixel point of the image at each viewpoint; (s',t) is the coordinate of the pixel point in the image of each viewpoint mapped to the depth position z after translation. The coordinates of are the coordinates of the pixel points mapped to the depth position.

The synthetic aperture image S_z at depth z can be obtained by averaging the images after parallax removal, i.e.,

$$S_z = \frac{1}{N}\sum_{i=1}^{N} I_{i,z}(s',t) \tag{5}$$

The synthetic aperture algorithm based on multi-view images can focus imaging on object surfaces of different depth layers, and in this paper, the central viewpoint is used as the reference viewpoint for focusing imaging, and the results are shown in **Figure 5** and **Figure 6**.

Figure 5 show the images when focusing on different target layers. It can be seen that the target layer in the focus plane is clearly imaged, and the out-of-focus occlusion is diffused, which achieves the effect of de-focusing the occlusion and "seeing through" the target object. When the target is focused at the specified depth layer, as shown in **Figure 6a**, the image is clear and the other occluded objects in the out-of-focus plane are diffuse. When focusing on other depth layers, the previous target layer is in the out-of-focus plane, the image is blurred, and the clear structure of the target layer cannot be identified, as shown in **Figure 6b**.

3. Depth resolution

In the stereo vision system, the depth resolution is the vertical resolution expressed in length, i.e., the layer resolution of the target object. The field of view is the extent of the imaging area, and in X-ray imaging refers to the maximum tensor an-

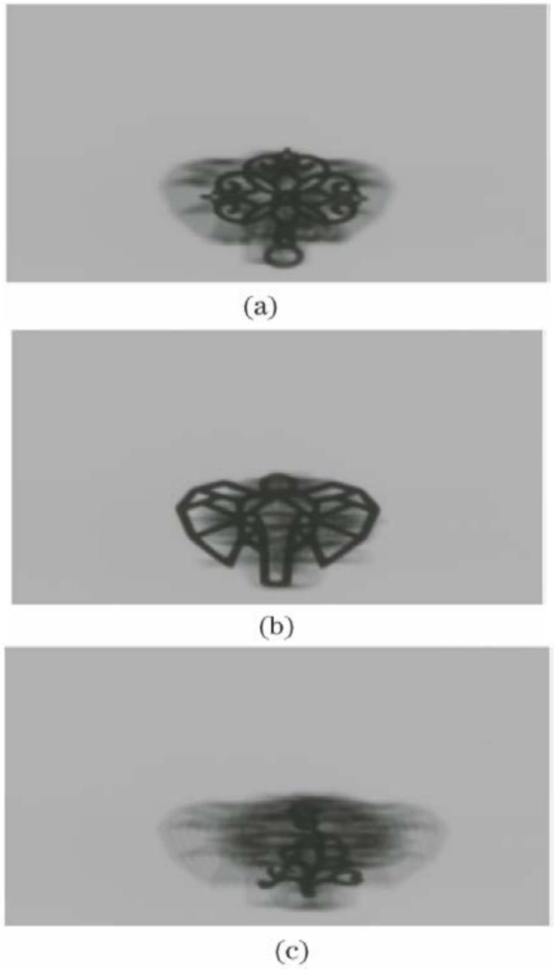
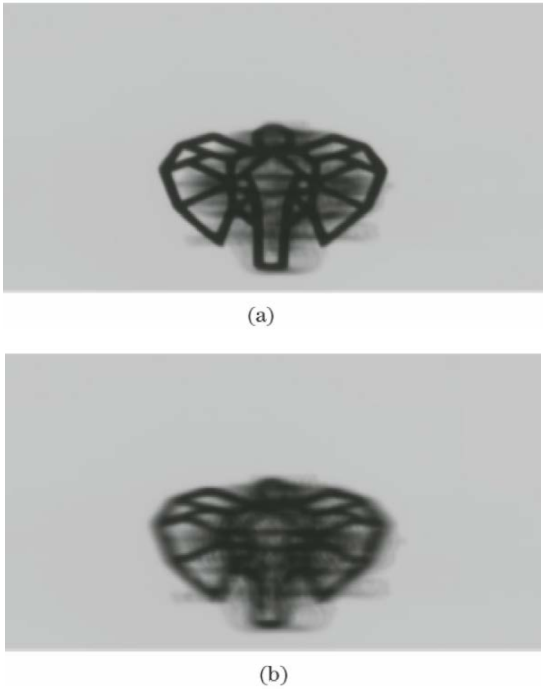


Figure 5. Images for focusing on different depth layers. (**a**) Focusing on foreground; (**b**) focusing on middle layer; (**c**) focusing on background.


(a)

(b)

Figure 6. Focus imaging on different depth layers. (**a**) Focus imaging on middle layer; (**b**) focus imaging on another layer.

gle from the ray source to the left and right sides of the detector, which reflects the maximum extent of the X-ray source to image the target object. In synthetic aperture imaging, the depth resolution affects the quality of the reconstructed image, so the factors affecting the depth resolution are investigated by analyzing two target points at different depth layers.

The process of imaging target points at different depths by the X-ray scanning imaging system is shown in **Figure 7**, where the field of view of the system at this time is the maximum field of view angle 2θ, and the ray source is located at the position of point *A*. As shown in **Figure 7a**, two points in space, object point B and object point C, are at distances *Z* and *Z* + Δ*Z* from the X-ray source respectively, and their projection points are located at projection point *O*. The distance between the source and the detector is *f*. As the target moves, object point *B* and object point *C* remain relatively stationary from left to right. The target point *B* arrives at position B_1, and the position is the far right of the X-ray imaging area, and the position of the projection point of the object point B_1 is the projection point B_2. At this time, the target point *B* is collected by the line detector on the far right. With time going by, the target object moves continuously from left to right, and the target point *C* moves to the position C_1, which is the rightmost part of the X-ray imaging area, as shown in **Figure 7b**, and the projection point position of object point C_1 is the projection point C_2.

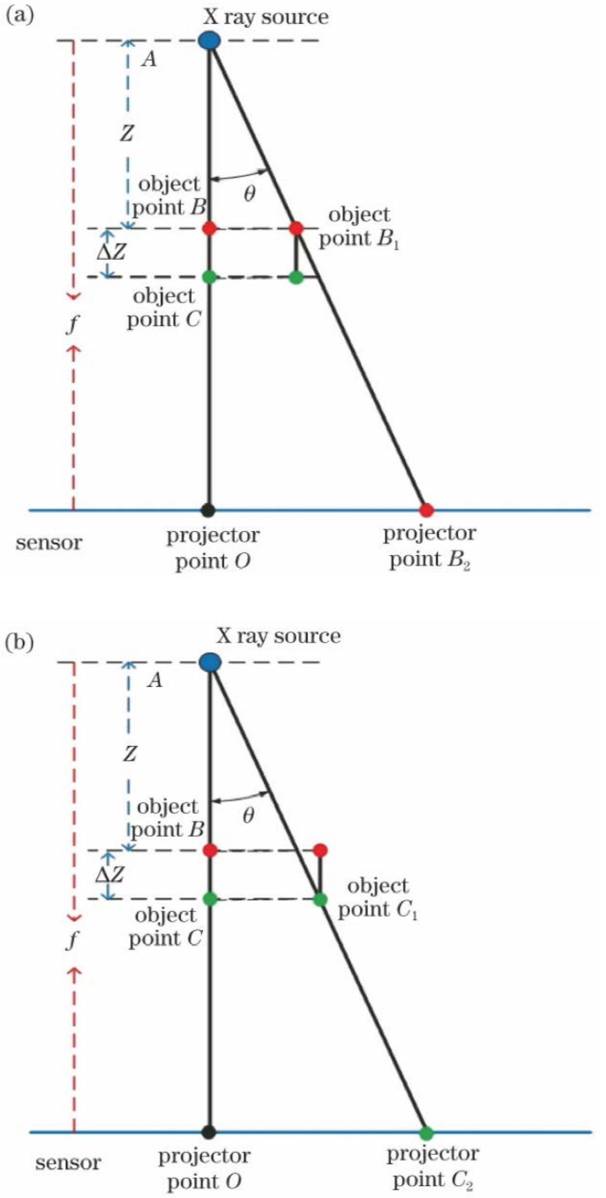


Figure 7. Imaging principle of target points with different depths in space. (**a**) Target point *B* imaged at point B_1; (**b**) target point *C* imaged at point C_1.

As shown in **Figure 7a**, when the rightmost line detector picks up the projection information of target point *B*, the information of target point *C* is not picked up by the rightmost line detector. When the target moves to the position shown in **Figure 7b**, the rightmost line detector picks up the projection information of target point *C*. Therefore, for the target point at different depth positions, its imaging in the line detector is achieved by the projection information collected by the detector at different moments. Since continuously moving target points are imaged in the adjacent line detectors, the distance of the target point movement at each moment is

$$S = \frac{Z \cdot W}{f} \tag{6}$$

where: S is the distance of the target point motion at each moment; W is the width of a pixel point; Z is the distance from the ray source to the target point; f is the distance from the ray source to the detector.

Since the field of view of the detector imaging area is 2θ, the distance L_{BB_1} of the motion of the target point B can be expressed as

$$L_{BB_1} = Z \cdot \tan\theta \tag{7}$$

Then the parallax of the target point B is

$$d_1 = \frac{L_{BB_1}}{S} = \frac{\tan\theta \cdot f}{W} \tag{8}$$

When the target point C moves to the position C_1, which is imaged on the rightmost line detector, the target point C moves at a distance L_{CC_1} of

$$L_{CC_1} = (Z + \Delta Z) \cdot \tan\theta \tag{9}$$

Similarly, the parallax of the target point C can be found as

$$d_2 = \frac{(Z + \Delta Z) \cdot \tan\theta \cdot f}{Z \cdot W} \tag{10}$$

The parallax between the target point B and the target point C is

$$\Delta d = d_2 - d_1 = \frac{\Delta Z \cdot \tan\theta \cdot f}{Z \cdot W} \tag{11}$$

Then the parallax between X-ray imaging areas ΔS is expressed as

$$\Delta S = 2 \cdot \Delta d = \frac{2 \cdot \Delta Z \cdot \tan\theta \cdot f}{Z \cdot W} \tag{12}$$

Analytically, since f, Z and W are fixed values, equation (12) can well illustrate the relationship between depth resolution ΔZ and field of view angle 2θ. When the imaging area is certain, i.e., θ is constant, for two points in space separated by ΔZ, the parallax value of the two target points increases as ΔZ increases, which also explains the larger difference in the imaging position of two target points with large depth difference. When the distance ΔZ between the two target points is certain, the larger the imaging area, i.e., the larger θ, the larger the parallax value of the two target points. Equation (12) well illustrates the layer resolution capability of the imaging system.

4. De-overlap enhancement algorithm

4.1 Focusing degree judgment

Synthetic aperture imaging is always able to achieve focused imaging of different depth layers, and for the target layer, its focused imaging process is the process of the target layer from out-of-focus to focused to out-of-focus again, which is expressed as the process of image from blurred to clear to blurred again, so it is essential to select the focused target layer, i.e., the clearest image for subsequent target region extraction. However, in the process of focusing image selection, it is difficult to intuitively determine which one is the clearest.

Based on the image formed by X-ray synthetic aperture, the traditional bokeh discrimination method is no longer applicable to the discrimination of synthetic aperture imaging because of the sharp image details in both the bokeh and focus parts due to the translation superposition of the image, so this paper proposes a sharpness discrimination method based on the image gradient without reference image to evaluate the image quality. First construct a reference image for the image through a low-pass filter. And then, the gradient image of the reference image and the image to be evaluated is extracted. And the N_1 image blocks with the most gradient information are extracted from the calculated gradient image. Finally, the structural similarity (SSIM) between the target image and the reference image is calculated to evaluate the sharpness of the image. For the image, a clear image contains more high-frequency components, which is manifested by sharp edges and large gradient value variations. Since X-ray imaging is transmission projection imaging, when imaging based on synthetic aperture algorithm, the image translation superposition is expressed as the superposition of pixel gray value, so when the image is focused, the pixel gray value of the focused area of the image changes more, i.e., the gradient value changes more, while the gradient value of the scattered area changes less than the focused area, as shown in **Figure 8**.

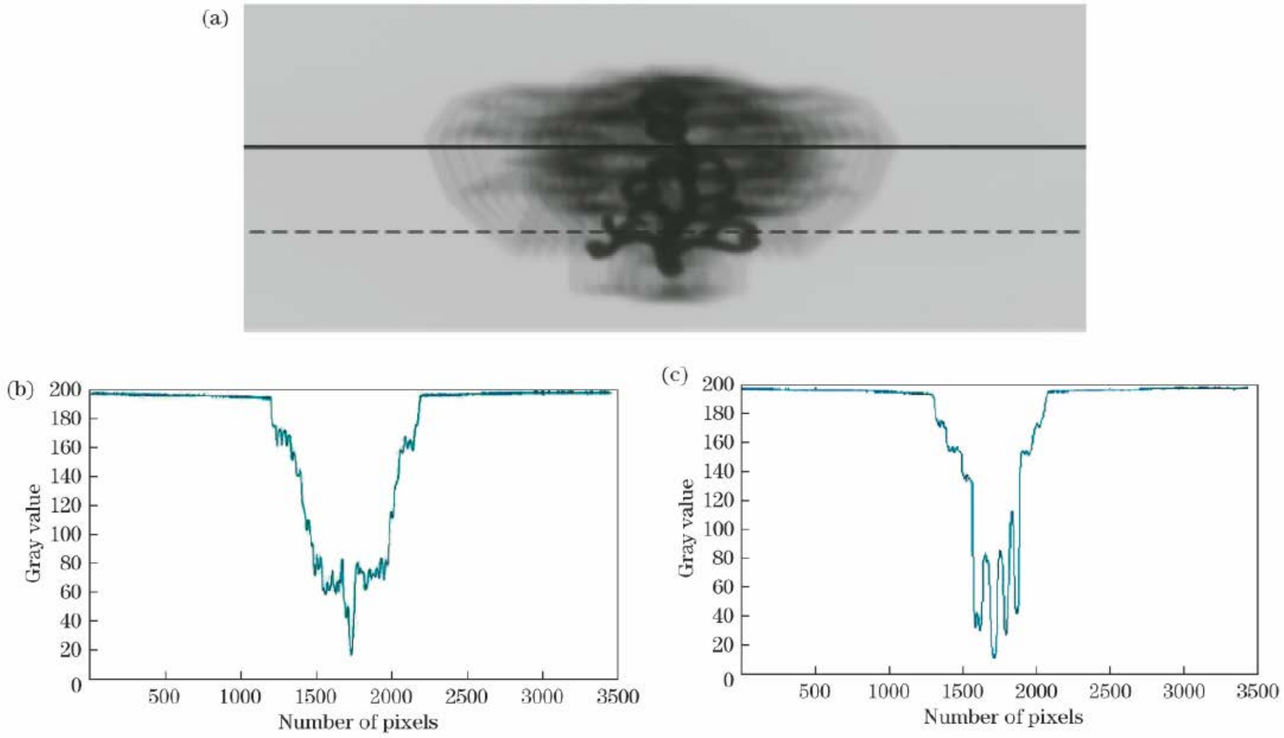


Figure 8. Focus judgment. (**a**) Synthetic aperture image; (**b**) pixel value at solid line; (**c**) pixel value at dotted line.

Figure 8 shows the variation of the grayscale values of the pixels in two lines at different locations in the image. It can be found that, the variation of the pixel gray value in the focused region of the image is larger than that in the scattered region, as shown in **Figures 8b** and **c**. Since the magnitude of the change of pixel value in the image is expressed as the change of the gradient value, the larger the magnitude of the change of pixel value becomes in the image, the larger the gradient value is, and vice versa. For the image to be evaluated and the reference image, the SSIM value of the N_1 image block with richer gradient information is calculated, and the SSIM value is larger for the clearer image to be evaluated and smaller for the fuzzier image to be evaluated.

4.2 Focusing region identification and extraction

After judging the image focus degree, the best focus image layer is determined, and then the scene target is well "highlighted", as shown in **Figure 5**, but the reconstruction effect of the specified depth layer target is not satisfactory due to the interference of the image scattering area. In order to better reconstruct the specified depth layer of the target object, it is necessary to remove the influence of the scattered focus region and achieve the segmentation and extraction of the focused and scattered focus regions. In this paper, a keying algorithm based on Bayesian theory[21] is used to extract the focus region from the image. Since the proposed method requires trilateration marking (Trimap), an automatic marking method is used to identify the target and the background[22].

Since the size of the pixel value at the focus region of the synthetic aperture image is about the same as that at that position in the multi-view image, and the size of the pixel value at the scattered focus position (other depth levels) is more different from that at that position in the multi-view image, the pixel value variance of all the multi-view images with parallax removed at that position can be calculated to determine whether the pixel point is focused or not. The evaluation of the synthetic aperture image focused at the (*s*,*t*) position is

$$v(s,t)=\frac{1}{N}\sum_{i=1}^{N}[I_{i,z}(s,t)-S_{i,z}(s,t)]^2 \tag{13}$$

where: $S_{i,z}$ (s,t) is the synthetic aperture image; $I_{i,z}$ (s,t) is the multi-view image after parallax removal; N is the number of viewpoints at that position in the image. From Equation (13), the smaller the $v(s,t)$ is, the smaller the variance of the pixel value will become, the better the focusing effect will be, and a suitable threshold can be set to mark the focusing area.

By using the automatic marking method to mark the focusing region, the image can be initially judged as focused, out-of-focus and uncertain by marking, and the Bayesian theory based on the Liu *et al*. and Chen *et al*.[23,24] is used to extract the focusing region of the image. Bayesian theory is used to estimate the values of foreground $\boldsymbol{F}$, background $\boldsymbol{D}$, and fusion coefficient $\boldsymbol{\alpha}$ to satisfy the maximized posterior probability $P(\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha}|\boldsymbol{I})$, i.e., the maximum a posteriori probability (MAP) problem, in the case of known image $\boldsymbol{I}$, using the conditional probability method. When considering the maximization problem, $P(\boldsymbol{I})$ is a constant that can be neglected, so

$$\begin{aligned}&\underset{\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha}}{\operatorname{argmax}}\,P(\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha}|\boldsymbol{I})=\\&\underset{\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha}}{\operatorname{argmax}}[P(\boldsymbol{I}|\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha})P(\boldsymbol{F})P(\boldsymbol{D})P(\boldsymbol{\alpha})/P(\boldsymbol{I})]=\\&\underset{\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha}}{\operatorname{argmax}}[L(\boldsymbol{I}|\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha})+L(\boldsymbol{F})+L(\boldsymbol{D})+L(\boldsymbol{\alpha})]\end{aligned} \tag{14}$$

where: $P(\cdot)$ is the probability function; $L(\cdot)$ is the likelihood function.

The first term in equation (14) is modeled as the error between the image $\boldsymbol{I}$ and the estimated image $\boldsymbol{I}'$, i.e.,

$$\begin{aligned}&L(\boldsymbol{I}|\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha})=\\&-\|\boldsymbol{I}-\boldsymbol{\alpha}\cdot\boldsymbol{F}-(1-\boldsymbol{\alpha})\cdot\boldsymbol{D}\|^2/2\sigma_I^2\end{aligned} \tag{15}$$

where: σ_I is the standard deviation of the noise simulation in the imaging process. The $\boldsymbol{F}$, $\boldsymbol{D}$, $\boldsymbol{\alpha}$ estimates are used to measure the $\boldsymbol{I}'$ values so that the deviation between the estimated image $\boldsymbol{I}'$ and the real image $\boldsymbol{I}$ conforms to the standard normal distribution. For the model of image $\boldsymbol{I}$, the error measurement corresponds to a Gaussian distributed error function $\boldsymbol{I}'$ with standard deviation σ_I, and the closer $\boldsymbol{I}'$ is to $\boldsymbol{I}$, the larger L $(\boldsymbol{I}|\boldsymbol{F},\boldsymbol{D},\boldsymbol{\alpha})$ is from the estimated value of $\boldsymbol{F}$, $\boldsymbol{D}$, $\boldsymbol{\alpha}$.

The second term L $(\boldsymbol{F})$ is modeled as a Gaussian distribution function with foreground-weighted mean color values $\overline{\boldsymbol{F}}$ and standard deviations σ_F, with the expression

$$L(\boldsymbol{F})=-\|\boldsymbol{F}-\overline{\boldsymbol{F}}\|^2/\left(2\sigma_F^2+2\sigma_I^2\right) \tag{16}$$

The spatial consistency of the original image is used to calculate $\bar{F}$ and σ_F in the neighborhood of the image $\boldsymbol{I}$. Since the foreground estimates are also affected by imaging noise, the image noise term is also added to the standard deviation of the Gaussian distribution to avoid regularization and most degeneracy in the optimization process. Similarly, L $(\boldsymbol{D})$ is modeled as

$$L(\boldsymbol{D})=-\|\boldsymbol{D}-\overline{\boldsymbol{D}}\|^2/\left(2\sigma_D^2+2\sigma_I^2\right) \tag{17}$$

where: $\overline{\boldsymbol{D}}$ denotes the background-weighted average color value; σ_D is the standard deviation of the noise simulation in the background $\boldsymbol{D}$.

Regarding the fusion coefficients $\boldsymbol{\alpha}$, their corresponding likelihood functions L $(\boldsymbol{\alpha})$ are modeled as Gaussian distributed error terms, i.e.,

$$L(\boldsymbol{\alpha})=-\|\boldsymbol{\alpha}-\overline{\boldsymbol{\alpha}}\|^2/2\sigma_\alpha^2 \tag{18}$$

where: $\overline{\boldsymbol{\alpha}}$ and σ_α are the mean and standard deviation of the Gaussian distribution respectively, which can be calculated by the contribution of Gaussian filtering to the domain pixels. Since the distribution of $\boldsymbol{\alpha}$ is well constrained for the MAP problem, it is important to set an appropriate standard deviation for the distribution of $\boldsymbol{\alpha}$. When the value of σ_α is larger, $\boldsymbol{\alpha}$ has less constraint on the MAP problem. The smaller the value of σ_α is, the larger the constraint of $\boldsymbol{\alpha}$ will be on the MAP problem and edge, and where the value of $\boldsymbol{\alpha}$ changes rapidly, the blurring phenomenon will occur. In order to avoid the blurring phenomenon and maintain the constraint effect of $\boldsymbol{\alpha}$ values, this paper introduces image gradient in the distribution of $\boldsymbol{\alpha}$, and when the image gradient is large, $\boldsymbol{\alpha}$ values will change greatly. The use of gradient maintains the spatial consistency of $\boldsymbol{\alpha}$ values in the smoothed region, while relaxing the constraint for the region with larger gradient. Therefore, Equation (18) can be replaced by

$$L(\boldsymbol{\alpha}) = -\| \boldsymbol{\alpha} - \overline{\boldsymbol{\alpha}}\|^2 / 2\sigma_{\boldsymbol{\alpha}}^2 \omega_g g \tag{19}$$

where: g is the normalized gradient of the pixel; ω_g is the weight of the effect of the gradient on the distribution of $\boldsymbol{\alpha}$, which is a constant. The $\boldsymbol{\alpha}$ values are iteratively optimized according to the above method to maximize the posterior probability, and then determine the front and rear views of the image to realize the front and rear view segmentation.

5. Experiment and analysis

In this paper, the multilayer scene target shown in **Figure 9a** is experimentally studied, and only three layers of targets are placed here, where the first, second and third layer targets are shown in **Figures 9b**, **c** and **d** respectively. The multi-line-scan X-ray imaging system in Section 2 was used for multi-view image acquisition. The pixel size of the X-ray imaging system detector in the experiment was 1,772 × 1,139, and the pixel size was 0.2 mm. At the maximum field of view, the image formed by the center detector column was proposed as the reference view image, and the four detector columns at the maximum interval between the left and right sides were extracted from the center detector column to form the multi-view image. After that, we use the synthetic aperture algorithm to focus the multi-view images and realize the focused images of the target layer at different depths.

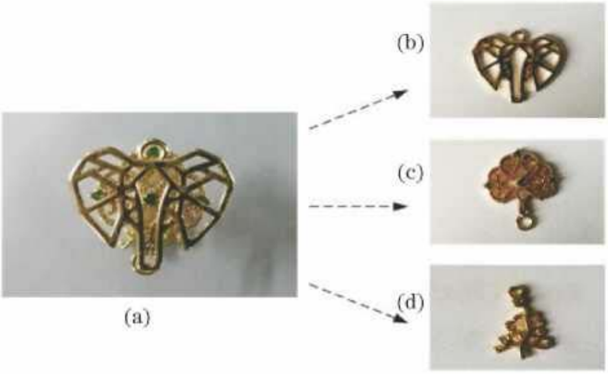


Figure 9. Multi-layer target and each layer structure. (**a**) Three-layer target; (**b**) first-layer target; (**c**) second-layer target; (**d**) third-layer target.

From the synthetic aperture imaging process, it can be seen that the focused image of the target layer changes from blurred to clear and then to blurred (as shown in **Figure 6**), so this paper uses the focus degree judgment to determine the best focused layer. **Figure 10** shows the SSIM values of image layers at different depths, and it can be seen that the three peaks in the figure correspond to the three focusing layers of the target, which correspond to the three target layers in **Figures 9b** to 9**d**.

The best focus was determined from the results of the synthetic aperture, and the best focus layer was determined for the three target layers. Since the occlusion layer is on the out-of-focus plane, a good "blur" effect is produced, as shown in **Figure 11a**. In order to exclude the influence of the out-of-focus area on the reconstruction effect and facilitate the identification of the target object, the focus area extraction algorithm is used to extract the focus area of the image, and the results are shown in **Figure 11b**.

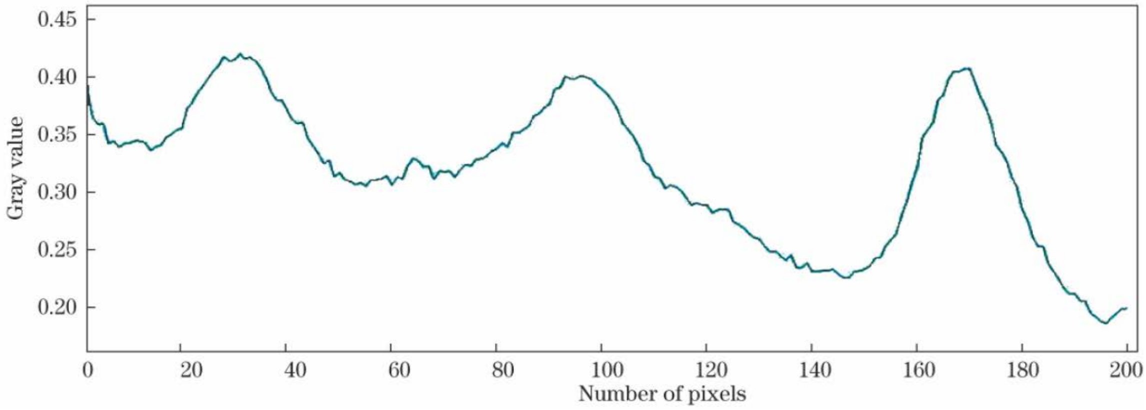


Figure 10. SSIM value of target image on each depth layer.

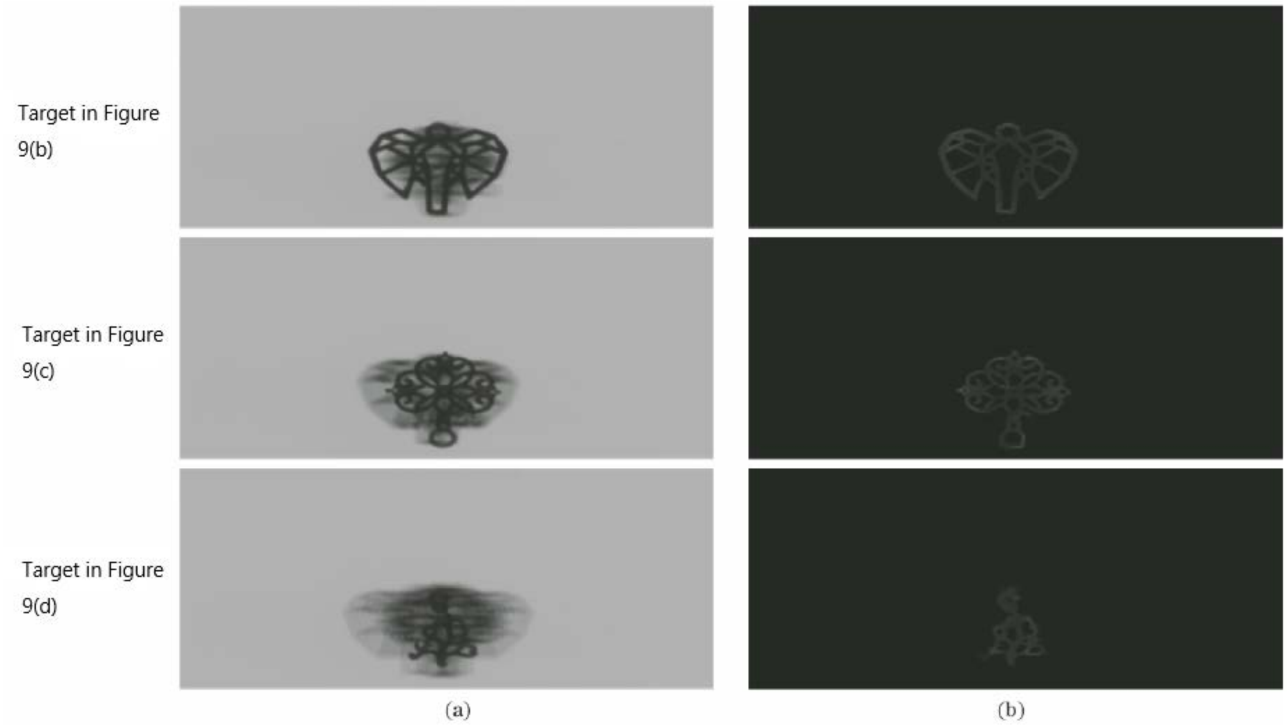


Figure 11. Extraction results of focusing area under different algorithms. (**a**) Synthetic aperture imaging algorithm; (**b**) focus area extraction algorithm.

Figure 12. False color display of focus area.

Figure 11 shows the results of the synthetic aperture imaging algorithm and the results of the aggregated region extraction algorithm, respectively. It can be seen that the target can be well extracted by the focused region extraction algorithm, and the reconstruction quality of the target can be effectively improved. In order to make the display of the target layer more intuitive, the extraction results were displayed in pseudo-color, as shown in **Figure 12**.

In this paper, the peak signal-to-noise ratio (PSNR) and SSIM are used to evaluate the quality of the synthetic aperture result map and the focus area extraction results of the target objects in the three-layer scene, as shown in **Table 1**. It can be seen that the quality of the reconstructed images using this method is much better than that of the direct refocused reconstructed images in terms of signal-to-noise ratio and structural similarity.

Table 1. Quality evaluation of imaging results

Target	**Synthetic aperture imaging algorithm**		**Focus area extraction algorithm**	
	PSNR	**SSIM**	**PSNR**	**SSIM**
Target in **Figure 9b**	11.8476	0.3535	19.0265	0.6391
Target in **Figure 9c**	12.6554	0.3303	20.6424	0.6301
Target in **Figure 9d**	11.9568	0.3656	20.3651	0.6271

6. Conclusion

A new X-ray focusing imaging method based on multi-line scanning is adopted, which only requires the detected target to pass through the detection area along a straight line, realizing multi-view imaging by using multi-line scanning imaging method, and achieving the focusing imaging of the target structure in the specified depth layer by using synthetic aperture imaging algorithm to achieve the “see-through” imaging effect. Since the scattered focus area has a large impact on the reconstruction quality of the image, the best focus image layer is identified by combining the focus degree discrimination method, and the image is segmented and extracted by using the focus area extraction algorithm to achieve a clear image at the same time, improving the reconstruction quality, and facilitating the identification of the target object. The imaging system in the proposed method is simple to build, easy to implement, and does not occupy large space resources. The proposed method achieves fast detection of target objects at different depth layers and improves the recognition of target objects at different depth layers, which can be used for the recognition and detection of complex targets in security inspection and other fields, and has a large application prospect.

Acknowledgements

The research is supported by Electronic Testing Technology National Defense Science and Technology Key Laboratory Fund (614200010401).

Conflict of interest

The authors declared no conflict of interest.

References

1. Gou L, Wang X, Cao H. The present state and future development of X-ray imaging technology. Journal of Chengdu Institute of Technology 2002; 29(2): 227–231.
2. Li B, Tang S, Yang J, *et al*. Clinical application, research progress of CT spectral imaging. Chinese Journal of Cardiovascular Rehabilitation Medicine 2015; 24(3): 343–346.
3. Zhou S, Brahme A. Development of phase-contrast X-ray imaging techniques and potential medical applications. Physica Medica, 2008; 24(3): 129–148.
4. Hao J, Zhang L, Chen Z, *et al*. Multi-energy X-ray imaging technique and its application in computed tomography. Computerized Tomography Theory and Applications 2011; 20(1): 141–150.
5. Hounsfield GN. Computerized transverse axial scanning (tomography): Part 1. Description of system. The British Jornal of Radiology 1973; 46(552): 1016–1022.
6. Qi J, Chen R, Liu B, *et al*. Grating based X-ray phase contrast CT imaging with iterative reconstruction algorithm. Acta Physica Sinica 2017; 66(5): 054202.
7. Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Physics in Medicine & Biology 2008; 53(17): 4777–4807.
8. Sidky EY, Kao CM, Pan X. Accurate image reconstruction in CT from projection data taken at few-views. Medical Imaging 2006: Physics of Medical Imaging 2006; 6142: 614229.
9. Gondrom S, Zhou J, Maisl M, *et al*. X-ray computed laminography: An approach of computed tomography for applications with limited access. Nuclear Engineering and Design 1999; 190(1/2): 141–147.
10. Wan X, Liu X, Wu Z. Review of computed laminography. CT Theory and Applications 2014; 23(5): 883–892.
11. Xu F, Helfen L, Baumbach T, *et al*. Comparison of image quality in computed laminography and tomography. Optics Express 2012; 20(2): 794.
12. Ng R. Fourier slice photography. ACM Transactions on Graphic 2005; 24(3): 735.
13. Liu Y, Liu B, Pan J. Synthetic aperture imaging algorithm via foreground removing. Acta Optica Sinica 2018; 38(6): 0611002.
14. Berry MV, Klein S. Integer, fractional and fractal Talbot effects. Journal of Modern Optics 1996; 43(10): 2139–2164.
15. Zhou Z. Research on light field imaging technology [PhD thesis]. Hefei: University of Science and Technology of China; 2012.
16. Nie Y, Xiang L, Zhou Z. Advances in light field photography technique. Journal of the Graduate School of the Chinese Academy of Science 2011; 28(5): 563–572.
17. Wilburn B, Joshi N, Vaish V, *et al*. High performance imaging using large camera arrays. ACM Transactions on Graphic 2005; 24(3): 765–776.
18. Vaish V, Wilburn B, Joshi N, *et al.* (editors). Using plane+ parallax for calibrating dense camera arrays. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2004 Jun 27–Jul 2; Washington DC. New York: IEEE; 2004. p. 8161383.
19. Yu D, Tan H. Engineering optics. Beijing: China Machine Press; 2011.
20. Vaish V, Garg G, Talvala E, *et al.* (editors). Synthetic aperture focusing using a shear-warp factorization of the viewing transform. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)-Workshops; 2005 Sept 21–23; State of California. New York: IEEE; 2006.
21. Chuang Y, Curless B, Salesin DH, *et al.* (editors). A Bayesian approach to digital matting. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2001 Dec 8–14; Hawaii State. New York: IEEE; 2001. p. 7184431.
22. Liu B, Pan Y, Yan W. Defocusing mechanism and focusing evaluation function of light field imaging. Acta Physica Sinica 2019; 68(20): 204202.
23. Levin A, Lischinski D, Weiss Y. A closed-form solution to natural image matting. IEEE Transactions on Pattern Analysis and Machine Intelligence 2008; 30(2): 228–242.
24. Chen T, Wang Y, Schillings V, *et al.* (editors). Grayscale image matting and colorization. Asian Conference on Computer Vision; 2004 Jan 27–30; Korea. Korea: ACCV; 2004. p. 1164–1169.