Terahertz holographic detection for high-resolution sensing of axially diverse targets

Huang, Haochong; Yang, Meihui; Li, Zhijie; Panezai, Spozmai; Zheng, Zhiyuan

DOI
10.1088/1361-6501/add1ff

Publication date
2025

Document Version
Final published version

Published in
Measurement Science and Technology

IOP Publishing Measurement Science and Technology
Meas. Sci. Technol. **36** (2025) 055210 (20pp) https://doi.org/10.1088/1361-6501/add1ff

Terahertz holographic detection for high-resolution sensing of axially diverse targets

Haochong Huang[1,3,*], Meihui Yang[1], Zhijie Li[1], Spozmai Panezai[2] and Zhiyuan Zheng[1,3,*]

[1] School of Science, China University of Geosciences (Beijing), Beijing 100083, People's Republic of China
[2] TU Delft | University of Technology, Mekelweg 5, 2628 CD Delft, The Netherlands
[3] Frontiers Science Center for Deep-time Digital Earth, China University of Geosciences (Beijing), Beijing 100083, People's Republic of China

E-mail: hchhuang@cugb.edu.cn, zhyzheng@cugb.edu.cn and 2119230032@email.cugb.edu.cn

Received 12 September 2024, revised 24 April 2025
Accepted for publication 29 April 2025
Published 15 May 2025

Abstract

For the focused sensing of targets at different axial depths, the synchronized display of their three-dimensional morphology is challenging. Terahertz waves exhibit selective penetration of some dielectrics and non-polar substances, which produces constructive guidelines for the targets. Here, a high-quality, superior-resolution, and depth-free sensing approach is demonstrated by a terahertz holographic mode. This work proposes and develops a novel stereoscopic-field method of axial multi-depth targets based on terahertz digital holography. The feasibility of the pre-propagating diffraction decomposition strategy is investigated, and the mutual disturbance from the stacking interference patterns is effectively curbed. The optimized holograms are propagated to a pre-screening plane where exclusive mask filters of samples are shaped by using Otsu's method. The signals are reconstructed by classical auto-focusing criteria and the automated merging process can promise the performance of the targets in one stereoscopic-field. Meanwhile, the implementation of extending the field of view method, global and discrete optimization, and the phase retrieval and unwrapping algorithm also contribute to the imaging performance. Furthermore, two representative setups are applied and the positive experimental results indicate the universality and robustness of stereoscopic-field terahertz depth-free holography technology.

Supplementary material for this article is available online

Keywords: terahertz, holographic sensing, pre-propagating diffraction decomposition, stereoscopic-field, depth-free

1. Introduction

In the study of stereoscopic-field sensing technology, accurately measuring object shapes across varying distances is essential for evaluating their characteristics effectively. The traditional two-dimensional (2D) display is limited by the biased geometric assumptions of axial multi-depth [1]. Differently, three-dimensional (3D) scenes can present full parallax, large depth of field (DOF), and smooth scene transition without discontinuity and have promising applications in remote sensing [2], lidar ranging [3], deep-sea exploration [4], geographical mapping [5], and environmental monitoring [6]. The 3D imaging detection includes technologies such as time-of-flight imaging [7, 8], light-field cameras [9, 10], structured light [11, 12], computed tomography [13, 14], metalens

* Authors to whom any correspondence should be addressed.

depth sensors [15], etc. It can not only obtain 2D scenes but also record the depth and shape features that are reflected in the amplitude and phase of the light field [16]. Holography is considered a promising technology among the various methods discussed earlier, particularly in the realm of 3D scenes with vast potential applications [17–19].

The invention of holography enables the recording, storage, and visualization of depth information, which is compressed for efficient representation. It allows the utilization of the light naturally scattered in the physical targets, which are the light field intensity and depth information [20]. Besides, the fast technological development of image sensors, micromanufacturing, and computer algorithms have substantially advanced the digital reproduction process [21]. Digital holography encodes phase and amplitude variations in the propagation beam through an interference pattern and utilizes a signal sensor for data capture [22, 23]. The computerized reconstruction enables the integration of signal processing algorithms conveniently as well as improves image quality effectively. This technique is particularly prevalent in the infrared spectrum, where digital holography is widely applied [24, 25]. This full-field, real-time, and non-destructive holographic imaging has considerable application prospects in microscopic observation [26–28], contour measurement [29–31], particle determination [32, 33], etc.

Terahertz radiation, located between microwave and infrared radiation, refers to the component of electromagnetic waves with a frequency ranging from 0.1 to 10 THz and corresponding to wavelengths between 30 μm and 3 mm [34–36]. It has excellent penetration capability in non-polar materials [37], low single photon energy [38] and fingerprint absorption spectrum, etc [39]. In more than twenty years of terahertz imaging development [40, 41], the stereoscopic-field perception has become available with the improvement of sources and array detectors [42]. The introduction of digital holography with terahertz band can make full use of the advantages of both [43–45]. Meanwhile, the update of computer algorithms also keeps pace and is accompanied by novel resolving approaches and modalities [46]. In terms of improving the quality of terahertz holograms and enhancing the accuracy of the reconstruction, a large number of methods have been implemented, such as auto-focusing [47], phase retrieval and unwrapping [48, 49], etc. However, the disturbance of the axial multi-depth targets (AMDTs) is extremely exaggerated, particularly for the varying phase objects or the scenes with extended depth. Achieving high-resolution imaging with exceptional quality in these contexts remains elusive.

Advancements in stereoscopic-field imaging technologies rely on DOF expansion techniques, which enhance imaging quality by maintaining clear focus within a specific range. Previous studies have demonstrated that liquid lenses can extend the DOF in zoom microscopes, significantly improving imaging performance [50]. However, this approach is limited to the visible light spectrum and requires multi-frame scanning and image fusion, leading to a complex imaging process and high hardware costs. Traditional 3D deconvolution techniques, although capable of single-frame 3D reconstruction, depend on the precise modeling and experimental calibration of the point spread function, resulting in high computational complexity and stringent requirements for experimental conditions and hardware precision [51, 52]. In contrast, this paper proposes the pre-propagation diffraction (PPD) method, based on terahertz digital holography, offering broad-spectrum adaptability, particularly suitable for non-contact 3D reconstruction in the submillimeter waveband. The PPD method eliminates the need for multi-frame scanning and depth layer presetting, enabling multi-target separation and depth-independent 3D reconstruction. By leveraging the physical wavefront propagation model, PPD effectively reduces artifacts and depth estimation errors, demonstrating enhanced adaptability and interference resistance in complex multi-target scenarios. It significantly improves imaging efficiency and hardware adaptability, making it compatible with various terahertz platforms.

Driven by advancements in deep learning, convolutional neural networks (CNNs) have been introduced into the field of holographic imaging, significantly enhancing imaging speed and DOF expansion capabilities. For example, Wu *et al* proposed the holographic imaging using deep learning for extended focus method, which leverages deep learning models for automatic focusing and phase recovery. However, this method highly depends on large-scale data training, leading to potential performance degradation on unseen data, while its results lack physical interpretability [53]. In contrast, the PPD Method proposed in this paper is entirely based on the physical propagation model, requiring no data training, thereby ensuring high transparency and interpretability of results. It is particularly suitable for multi-target automatic reconstruction and industrial applications. Additionally, Yi *et al* introduced the Mask-Regional CNN algorithm, which is frequently used for object detection and segmentation in 2D images. By employing a segmentation-first, reconstruction-later strategy, this method mitigates the defocusing problem. However, its performance is limited by data distribution biases and category recognition accuracy [54]. By comparison, the PPD method automatically performs multi-target separation and 3D reconstruction during the imaging process, eliminating the need for prior target classification or depth layer settings, thus significantly improving automation and computational efficiency.

Traditional structure tensor methods provide a solution for automatic focusing and multi-depth reconstruction by performing depth estimation based on image feature values and gradient analysis. However, this approach requires frame-by-frame evaluation of reconstructed images at different depths, relying heavily on image processing and feature extraction algorithms, which exhibit significant limitations in environments with weak edge features or high noise

levels [47]. Additionally, in holographic decomposition techniques, commonly used automatic focusing algorithms typically involve a multi-step processing workflow, including back propagation, image segmentation, and morphological operations. The sequential execution of these steps substantially reduces imaging efficiency and automation levels. In contrast, the PPD Method proposed in this paper directly analyzes depth information based on the physical wavefront propagation model, enabling automatic target separation and depth estimation without requiring multiple reconstructions or frame-by-frame data acquisition. The PPD method demonstrates greater robustness and environmental adaptability in high-noise environments and densely distributed target scenarios, significantly improving imaging speed and automation performance, while excelling in complex multi-target reconstruction tasks.

Building on these considerations, this paper proposes a stereoscopic-field terahertz depth-independent sensing holography (STDSH) technology, incorporating the PPD method to achieve multi-target depth-independent imaging and automatic reconstruction. This pioneering model enables full-focusing on the AMDT with a single-shot. The proposed approach introduces a PPD plane with a binarized mask and an automatic merging algorithm, significantly expanding the field of view (FOV). Intensity calibration and subpixel displacement correction ensure the accuracy and stability of image reconstruction. In this study, we assume that the objects under analysis do not overlap. This assumption is made because the segmentation method we applied is based on the separation of connected regions, which inherently assumes non-overlapping objects. We recognize that addressing overlapping objects is a challenging issue, particularly when it comes to accurately segmenting overlapping regions. The segmentation of these overlapping regions requires careful consideration of additional factors, such as image depth information, grayscale gradients, and other optical properties, all of which introduce complexities into the segmentation process. Since these challenges were not the primary focus of our current research, this topic has not been discussed in detail in the manuscript. However, we acknowledge that overlapping objects represent an important direction for future work. Tackling this issue may require the application of more advanced techniques, such as deep learning algorithms or physics-based models, which could improve the ability to detect and segment overlapping regions effectively. In future studies, we plan to investigate these approaches further, which may provide a more robust solution for handling overlapping objects in terahertz imaging and other similar applications. Experimental results demonstrate that the proposed method exhibits strong robustness and environmental adaptability across different hardware configurations, offering a highly efficient and versatile solution for terahertz stereoscopic-field sensing technologies. To the best of our knowledge, this represents a novel contribution to the field, showcasing remarkable innovation and potential for broad applications.

2. Setup and samples

The Gabor in-line configuration was adopted in different terahertz hardware systems to practice the proposed STDSH method. The samples are directly illuminated by the incident terahertz wave, and their morphology and vertical variations are captured using an imager, obviating the need for supplementary lenses. Schematic diagrams and parameter comparisons of the different systems are presented in figure 1 and SM Table 1, respectively. (SM Table 1 is provided as an attachment).

Based on the continuous-wave terahertz digital holographic systems, there were different hardware devices involved in experiments. As shown in figure 1(a), the CO_2 laser pumped source and pyroelectric detector system were introduced to sense the natural plants and artificiality, and the quantum cascade laser (QCL) and microbolometer system were built for photographing the diverse structure as shown in figure 1(b). The photodetector is equipped with a rotary chopper (not shown in figure 1(a)), which is used in the high-power setup to modulate the optical signal. It periodically interrupts the light beam to accommodate the dynamic response range of the pyroelectric detector. According to the specialty of the devices, accessories, and samples in systems, a comprehensive comparison of the apparatus in different systems is illustrated as follows: (i) In terms of terahertz source selection, the CO_2 laser pumped light source was used in the pyroelectric detection system. The generated terahertz wave holds a wavelength of 118.83 μm and a frequency of 2.52 THz. (ii) The QCL laser exhibited a relatively higher dominant frequency of 3.1 THz and a shorter wavelength of 97 μm. The reflective-type module was built for the high-power system and the transmission-type was designed for the other one. The off-axis parabolic mirrors (PM) pack with a focal length ratio of 1:2 (PM1:PM2) in figure 1(a) was placed to reduce the propagation attenuation in air. Correspondingly, two high-resistance silicon single-crystal convex lenses with an effective focal length of 13 mm were used in the low-power system, as shown in figure 1(b). Functionally, the original beams were collimated and expanded after passing through each wave modulation device. (iii) Concerning the matching detectors, a large size sensor pyroelectric detector (Pyrocam IV) with a square array of 25.6 mm × 25.6 mm and a pixel pitch of 80 μm was installed in the high-power system to obtain more high-frequency information. And in the low power setup, to collect enough wave signals, a 2D electric displacement stage was used for the synthetic aperture process. A rectangular array microbolometer (IRV) with a field range of 7.52 mm × 5.64 mm corresponding to an element space of 23.5 μm was fixed. (iv) Samples with different properties were detected such as a branch of natural asparagus fern and a metal pin were used in the pyroelectric detection system while a steel needle with a hole and a silicon structure were photographed in the microbolometer system. The geometric center positions of the samples in different systems diverged more or

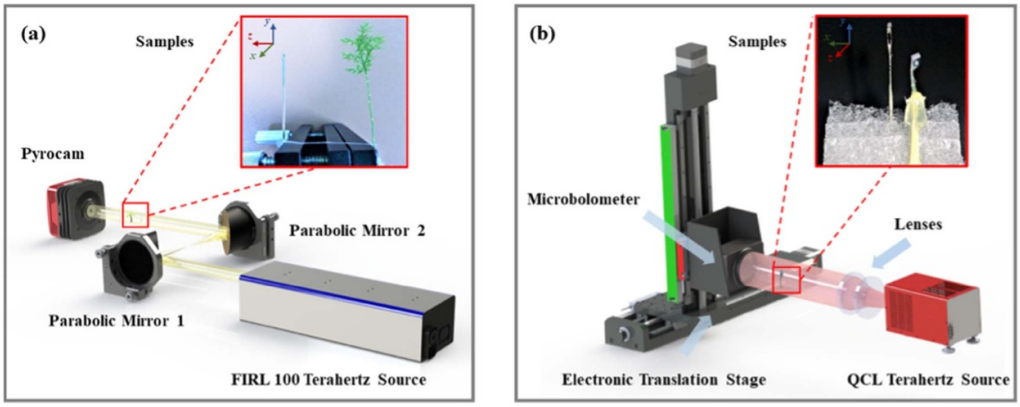

Figure 1. Layout diagrams and samples in experiments: (a) CO_2 laser pumped source and pyroelectric detector system with a branch of natural asparagus fern and a metal pin, (b) QCL and microbolometer system with a steel needle with a hole and a silicon structure.

less. All the samples were fixed and spaced at a certain distance along the direction of illumination, in other words, the depth direction.

3. Principle and methods

Herein, the outline of the STDSH method is elaborated. The full procedure can be divided into two processes i.e. recording and reconstruction. The solutions corresponding to practical problems encountered in different systems are presented as shown in the flowchart. The involved processes including the Terahertz data denoising optimizations, expanded FOV, Pre-propagating diffraction decomposition, Edge averaging blurred transition, Auto-focusing phase retrieval and unwrapping, and Stereoscopic-field automated merging are thoroughly displayed. Before introducing these sub-method modules, a brief overview is depicted as follows:

(1) First, for the system in figure 1(a), the background images and holograms are pre-processed through global denoising optimization, as shown in the flowchart. For the system in figure 1(b), both global and discrete denoising optimizations are applied, with the global optimization method providing supplementary enhancement, while the discrete optimization, primarily based on interval acquisition and consistent frame looping extracted from the time series, is the main approach used for the holograms. The expanded FOV [54] method is also used to process the synthetic holograms and background images. Both systems normalize the holograms based on the background images. All the pre-processing operations of the images are shown in blue in figure 2.
(2) Normalized holograms are pre-propagated to a diffraction plane where exclusive masks are designed by using Otsu's method [55]. The masks squared with blue lines are multiplied with the orange wavefront image to obtain the respective complex amplitude images of AMDT. This filtering process is indicated by green arrows and the resulting images are marked as wavefront 1 and wavefront 2 in figure 2.
(3) Next, the distance of each independent object can be determined by using auto-focusing algorithms [56] as shown by the purple arrows. In the digital reconstruction, the processes of phase retrieval [57] and unwrapping [58] are shown as red and brown labels. Last but not the least, the mask filters are applied to the reconstructed planes to filter the extra noise resulting from the masks themselves. The results of automated merging are shown in the recording plane.

3.1. Terahertz data denoising optimization

The terahertz data includes holograms and background images. Terahertz digital hologram $H(x,y)$ is encoded by the interference of the modulated object wave $O(x,y)$ with reference $R(x,y)$ that directly illuminates sensors. And the distributions can be expressed as:

$$H(x,y) = |O(x,y)|^2 + |R(x,y)|^2 + O(x,y)^* R(x,y) + O(x,y) R(x,y)^* \quad (1)$$

where $O(x,y)^*$ and $R(x,y)^*$ represent the corresponding conjugate term, and (x,y) are the coordinates at the detector plane. Besides, the correspondence between the holograms and light intensity $I(x,y)$ and amplitude $A(x,y)$ can be clarified by the following equation:

$$|H(x,y)| = I(x,y) = A^2(x,y). \quad (2)$$

In practice, captured terahertz images typically exhibit low contrast and poor signal-to-noise ratio (SNR). In the high-power pyroelectric detector system, defective pixels and negative pixel blocks in the conversion chip result in reduced sensitivity. Additionally, the rotary chopper introduces significant interference with the image signal quality. To overcome these issues, the system employs a global denoising optimization approach. Statistical analysis indicates that thermal noise caused by the pyroelectric effect typically follows a Gaussian

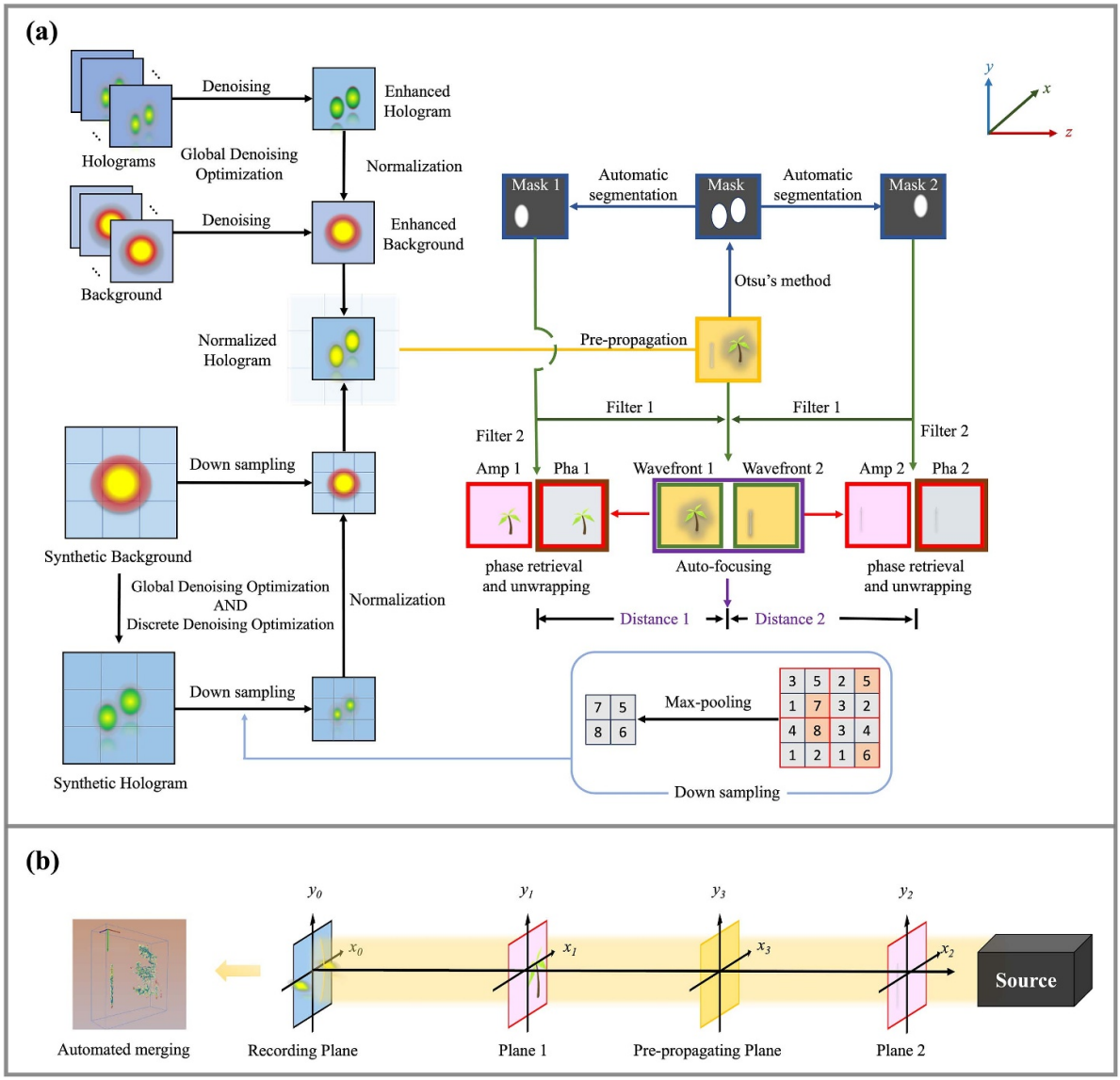

Figure 2. The flowchart of the proposed STDSH method. (a) Schematic diagrams illustrating the experimental principles of the pyroelectric detector system and the microbolometer system. (b) Schematic diagram of the beam propagation plane in the experimental optical system. (plane 1 corresponds to the natural asparagus fern and silicon structure, and plane 2 corresponds to the metal pin and steel needle with a hole).

distribution, and global denoising optimization using a mean filter can effectively mitigate this problem. The mean filter is applied to replace each pixel in the image with the average intensity value of its neighboring pixels, effectively smoothing the image. This process reduces random noise, such as thermal noise, while preserving the overall image structure. By averaging the pixel values over a defined window, the mean filter improves the image's SNR and ensures that important image features are retained for further processing. As shown in figure 2, both the background images and holograms undergo preprocessing during the denoising step. In this step, the raw data values of the acquired terahertz images are replaced with their corresponding statistical average intensity values, which significantly improves the image's SNR. This global denoising approach, using the mean filter, plays a key role in enhancing the quality and stability of the images for subsequent analysis.

During the up-sampling process for terahertz image optimization, although this step does not directly enhance the image resolution or quality, it effectively reduces the noise and blur typically associated with processing low-resolution images by introducing additional pixels. We employed bilinear interpolation as the up-sampling technique. The principle of bilinear interpolation involves estimating the pixel values in the enlarged image based on the values of neighboring pixels in the original image. This approach helps fill the gaps between existing pixels while avoiding artifacts that may arise from excessive sharpening. For the up-sampling process, a sampling factor (F) of 2 was applied in both horizontal and vertical directions, effectively increasing the image size. This method enriches the pixel data, providing more information for subsequent algorithmic processing, thereby facilitating the denoising process. The subsequent denoising step can focus

more effectively on extracting meaningful signals from the newly generated pixels, ultimately improving overall image quality and reducing noise. Additionally, during the optimization process, we introduce the use of mean squared error (MSE) and correlation coefficient (CC) functions to assess the denoising effectiveness [49] (see in particular, section 3.1 of this reference). MSE quantifies the average squared difference between the pixel values of the denoised image and the original reference, thus capturing intensity-level discrepancies. In contrast, CC evaluates the linear correlation between the denoised and original images, providing a statistical measure of their structural similarity. The combined use of these two metrics ensures a comprehensive evaluation of both local accuracy and global structural preservation. By optimizing both MSE and CC during training, we aim to maintain fine image details while suppressing noise, thereby enhancing both perceptual quality and quantitative performance. These functions not only account for image quality but also contribute to improving acquisition efficiency by reducing the need for repeated measurements.

In low-power miniaturized QCL systems, vibrations from the compressor may induce resonance effects, leading to system noise and instability, which can negatively impact image quality. To mitigate these issues and achieve high-quality hologram reconstruction, the system utilizes both global and discrete denoising optimization methods. Through experimental analysis of our low-power setup, we determined that the jitter perturbations exhibit a characteristic temporal period of 2 s. Over this period, the microbolometer, operating at a sampling frequency of 8.5 Hz, captures 17 images. As a result, the set of images captured by the microbolometer is organized into consecutive subsets, each containing 17 images. Within each subset, the images are indexed from $i = 1$ to $i = 17$, corresponding to specific time intervals within the jitter period. To address the challenges posed by vibration-induced noise, we employ a discrete acquisition process, where we select images with a fixed index, such as $i = 1$, from each subset. By choosing only the images with this specific index across all subsets, we extract a sequence of frames digitized at consistent intervals throughout the jitter period. This approach effectively forms a loop of temporally consistent frames, stabilizing the data by minimizing image position shifts and background noise caused by the jitter of the terahertz spot. This step can be described as extracting a consistent loop of frames from the time series, meaning that frames exhibiting a high degree of temporal consistency are chosen, further stabilizing the data.

In the experiment, we employed the max-pooling method for down sampling to reduce the number of data points (pixels) while preserving key features of the image. This approach selectively samples important areas of the image, discarding less relevant information, ensuring that critical image details are retained. The goal is to optimize the image processing pipeline, improving computational efficiency without compromising the integrity of essential features.

The first step in the max-pooling down sampling process involves selecting data points. To achieve this, the image is divided into smaller subregions. Each subregion is then evaluated based on its spatial consistency and significance within the overall image. Regions containing important features, such as edges, textures, and other structural elements crucial to image quality, are prioritized. In contrast, areas with minimal variation or fewer details are discarded, as their contribution to the overall image quality is minimal. Next, to reduce the data volume, we apply the max-pooling operation within each selected subregion. In max-pooling, the pixel value of each local region is replaced by the maximum value within that region. This method effectively lowers the image resolution while retaining the most prominent features of the image. By selecting the maximum value, we ensure that the most critical details, such as edges or high-contrast areas, are preserved. Although the resolution is reduced, this method ensures that the core features needed for subsequent processing remain intact. Despite the resolution reduction, we carefully ensure the preservation of important image features, such as object boundaries and high-contrast regions, by selecting sub-images that emphasize these key elements. These sub-images form the foundation for the subsequent synthetic aperture stitching process. By focusing on the most information-rich areas of the image, max-pooling not only alleviates the computational burden but also ensures the integrity of essential image features. In addition to reducing data volume, the max-pooling process helps to improve the SNR. By eliminating redundant or less informative data points, the overall clarity of the image is enhanced, allowing subsequent denoising algorithms to focus on the most important features. This reduction significantly improves the final image quality and stability, ensuring that only the most relevant details are retained. Although the system also utilizes a global mean filter for optimization, its enhancement effect is relatively modest. The primary improvements in image quality and stability are achieved through the discrete optimization process described above.

3.2. Expanded FOV

Lying in a relatively long-wavelength band (Terahertz and visible light), the transmitted terahertz beams always generate vigorous diffraction effect. However, most existing terahertz detectors have nearly an order of magnitude fewer pixels than visible light. The limitation of the terahertz hardware makes it unprocurable to collect high-frequency fringes outside the FOV, which plays a vital role in discriminating the key particulars of targets. In the pyroelectric detector system, the detection area of the detector in use is 25.6 mm $\times$ 25.6 mm, which is more than four times larger than the typical detector with a detection area of 12.8 mm $\times$ 12.8 mm. The calculation function of imaging resolution is shown as follows:

$$R = \frac{\lambda \times d}{N \times \Delta}, \tag{3}$$

where R is the theoretical resolution, λ is the incident wavelength, d represents the distance between detector and object plane, N is the number of pixels, and Δ is the pixel size. As for the insufficient sensing area of the microbolometer,

an expanded FOV method is applied to realize the congruent objective of expanding the recording surface. Light singles at the periphery of central images are captured by shifting the microbolometer on a 2D electric translation stage. The schematic diagram shown in figure 2 can picture this process intuitively.

The scanning path follows a grid pattern, preserving the rectangular format and aspect ratio of the original image. During the sequential acquisition of the nine sub-holograms, it is essential to ensure an overlap between adjacent sub-holograms, meaning the movement distance should be smaller than the lateral or vertical dimensions of the detector. The step size between adjacent sub-images is carefully set to guarantee that any artifacts in the data overlap and stitching process are smoothly corrected through intensity calibration. A formula of this method is shown below:

$$H_{\text{new}} = H_{\text{ori}} \times \frac{I_{\text{cen}}}{I_{\text{per}}}\,. \tag{4}$$

In the equation, H_{new} represents the sub-holograms after intensity calibration, H_{ori} denotes the original sub-holograms, I_{cen} indicates the total light intensity of the overlapping region in the center hologram, and I_{per} refers to the total light intensity of the overlapping region in the peripheral sub-holograms. According to the intensity distributions in central images, the matrix parameters of the other sub-images can be calibrated. After intensity registration, there is no obvious intensity 'jump' in synthetic images. In addition, the micro-displacement error that cannot coincide with real pixel usually causes stitching faults. It can be eliminated by applying the subpixel registration algorithm [59].

As shown in figure 2(b), the planes in this figure correspond to the targets identified in figure 1. Specifically, the y_1 plane corresponds to the natural asparagus fern and the silicon structure, whereas the y_2 plane corresponds to the metal pin and the steel needle with a hole. In addition, Amp1 and Pha1 in figure 2(a) correspond to the y_1 plane in figure 2(b), and Amp2 and Pha2 correspond to the y_2 plane in figure 2(b). Moreover, the images before (background images) and after placing samples (holograms) are both recorded for the normalization which can overcome the influences of light spot inhomogeneity and sensor finiteness. The denoised, enhanced and normalized terahertz synthetic hologram as shown in the flowchart is then prepared for the propagating reconstruction.

3.3. Pre-propagating diffraction decomposition

The numerical propagation of the obtained holograms is implemented according to the angular spectrum integral equation. The reproduced complex amplitude distribution $U(\mu,\varepsilon)$ of the wavefield at variable depth can be expressed as:

$$U(\mu,\varepsilon) = FT^{-1}\left\{FT\left[R(x,y)H(x,y)\right]G(f_\mu,f_\varepsilon)\right\}, \tag{5}$$

where FT and FT^{-1} stand for the Fourier transform and inverse Fourier transform, $G(f_\mu,f_\varepsilon)$ denotes the spatial transfer function, expressed as follows:

$$G(f_x,f_y) = \exp\left[j\frac{2\pi}{\lambda}Z\sqrt{1-(\lambda f_\mu)^2-(\lambda f_\varepsilon)^2}\right], \tag{6}$$

spectively, f_μ and f_ε are the spatial frequency at corresponding coordinate positions, Z is the propagation distance, and λ represents the employed wavelength. The spatial frequencies f_x and f_y represent the variations in the x and y directions of the image, and their formulas are given as:

$$f_x = \frac{x}{\lambda Z},\ f_y = \frac{y}{\lambda Z}, \tag{7}$$

where x and y are the spatial coordinates along the x and y axes, respectively. Meanwhile, the light field distribution can also be expressed with amplitude and phase as follows:

$$U(\mu,\varepsilon) = A(\mu,\varepsilon)\times\exp(i\varphi), \tag{8}$$

where

$$A(\mu,\varepsilon) = \sqrt{\{\text{Re}[U(\mu,\varepsilon)]\}^2+\{\text{Im}[U(\mu,\varepsilon)]\}^2}, \tag{9}$$

$$\varphi(\mu,\varepsilon) = \arctan\left\{\frac{\text{Im}[U(\mu,\varepsilon]}{\text{Re}[U(\mu,\varepsilon)]}\right\}. \tag{10}$$

Define the amplitude $A(\mu,\varepsilon)$ and the phase $\varphi(\mu,\varepsilon)$ of the complex amplitude $U(\mu,\varepsilon)$, with Im and Re representing the imaginary and real parts, respectively. By combining formulas (5), (9), and (10), the amplitude $A(\mu,\varepsilon)$ and the phase $\varphi(\mu,\varepsilon)$ at the reconstructed position can then be deduced.

When performing direct propagation reconstruction of holograms with axially compressed depth information, the overlapping diffraction orders can significantly affect the accuracy. To address this issue, a PPD decomposition method is proposed to filter out the interference noise caused by AMDT. The decomposition process is achieved using mask filters, which are automatically identified based on grayscale thresholding using Otsu's method [55]. This method determines the optimal threshold by maximizing the variance between two classes: foreground and background. In our implementation, Otsu's method is applied to segment the reconstructed hologram, distinguishing the diffraction noise from useful signal regions. Specifically, the Otsu algorithm determines the optimal threshold T by dividing an image into foreground and background components. In this algorithm, the grayscale values of foreground pixels are less than the threshold T, while the grayscale values of background pixels are greater than or equal to T. The goal of the algorithm is to maximize the difference in grayscale variance between the foreground and the background. When this difference reaches its maximum, the selected threshold T is considered the optimal threshold. By maximizing the grayscale difference between the foreground and the background, this method achieves optimal image segmentation.

In this study, we apply the Otsu algorithm to segment the hologram into the background and the sample body. First, the image is converted to grayscale, and its grayscale histogram is generated. The histogram illustrates the distribution of pixels at different grayscale levels, providing us with an initial basis for assessing the grayscale difference between the foreground and the background. Next, the Otsu algorithm calculates the mean grayscale values and between-class variance for the foreground and background at various thresholds based on the pixel counts provided by the grayscale histogram for each level. The between-class variance reflects the difference between the foreground and background, and the algorithm determines the optimal threshold T by maximizing this variance. Specifically, the Otsu algorithm iterates through all possible thresholds and selects the one that maximizes the between-class variance as the optimal segmentation threshold. This process maximizes the grayscale difference between the foreground and the background, thereby ensuring the best possible segmentation. Once the optimal threshold T is determined, the Otsu algorithm classifies the pixel values of the hologram into foreground and background regions based on this threshold. The grayscale values of the foreground are less than T, and those of the background are greater than or equal to T. Through this segmentation process, the resulting binary mask effectively isolates the sample body from the background, providing clear sample information for subsequent image reconstruction and analysis. By applying the Otsu algorithm, we can automatically select the optimal threshold T, we use it to generate binary masks that isolate high-intensity noise components, which are then filtered out during the reconstruction process.

The complex amplitude at an appropriate PPD plane is expressed as $U(m,n)$, and the pixel value in this diffractogram is represented with κ. The numerical determination of the mask $K(m,n)$ can be expressed as:

$$K(m,n) = \left\{ \begin{array}{ll} 1 & \kappa \geqslant T \\ 0 & \kappa < T \end{array} \right\}. \tag{11}$$

The Otsu method is first employed to process the global mask $U(m,n)$. The core principle of the Otsu method is to maximize the inter-class variance of the image grayscale distribution, thereby automatically determining the optimal binarization threshold T. This divides the grayscale levels into two classes: the target region and the background, converting $K(m,n)$ into a binary mask. After binarization, the connected component labeling algorithm is applied to the binary mask to identify all connected regions in the target area [60]. Each connected region is treated as a potential independent target, and a corresponding sub-mask $K(m,n)_i$ is generated for each region. The pixel values of each sub-mask $K(m,n)_i$ are defined as follows:

$$K(m,n)_i = \left\{ \begin{array}{ll} 1, & \text{if } (m,n) \in \text{Region } i, \\ 0, & \text{otherwise} \end{array} \right. . \tag{12}$$

Obtained masks are broken down to corresponding fragments which can be expressed as K_1, K_2, …, K_p to match with the involved target modules. By multiplying the pre-propagating result $U(m,n)$ with these dissected masks, the appearance characteristics of each target at the PPD plane can be described as follows:

$$\begin{array}{l} U(m,n)_1 = K(m,n)_1 \times U(m,n) \\ U(m,n)_2 = K(m,n)_2 \times U(m,n) \\ \quad\vdots \qquad\qquad\quad \vdots \\ U(m,n)_p = K(m,n)_p \times U(m,n) \end{array} . \tag{13}$$

Through this process, the global mask $K(m,n)$ is divided into a series of independent sub-masks $K(m,n)_1$, $K(m,n)_2$, …, $K(m,n)_p$. This ensures the accurate identification and separation of distinct independent targets, even in the case of complex multi-target structures. Once individual masks for each target are obtained, an edge-blurring operator is applied to reduce the sharpness of the mask edges. This process, which involves the introduction of an edge-averaging blurred transition, effectively mitigates the impact of sharp boundaries. By smoothing the edges, we achieve a gradual transition that minimizes the disruptive effects of abrupt edges, thus enhancing the overall robustness and precision of target identification in challenging scenarios.

By integrating the target pre-decomposition process described in formula (13), this method effectively simplifies the detection of AMDT, enabling accurate reconstruction of different targets across multiple depth planes. After applying the PPD method, conventional direct reconstruction methods are substituted with phase retrieval and phase unwrapping reconstruction techniques to accurately restore the real light field of AMDT. These techniques significantly enhance the precision of image reconstruction by precisely processing the phase information of the diffractograms. Phase retrieval effectively restores missing phase information, while phase unwrapping resolves the artifacts caused by phase wrapping, ensuring clear reconstruction of targets at different depths. Comparative experimental results demonstrate that this approach not only significantly enhances the precision of target separation and the quality of the reconstructed images but also greatly optimizes the final stereoscopic-field imaging performance.

3.4. Auto-focusing phase retrieval and unwrapping

For the reconstruction of AMDT after decomposition and independence, the auto-focusing algorithm has been adopted in perceiving a single target position. In this work, Tenenbaum's gradient criterion (TEG) and Fourier spectrum function (FSF) are employed [56, 61]. For phase data, the phase unwrapping based on the least square method [62] is adopted to reconstruct the continuous phase distribution, while uncovering the vesture of the actual profile with a truncated phase.

Moreover, in order to mitigate the twin image artifact resulting from the in-line construction and enhance the accuracy of the auto-focusing technique for precise distance estimation, the classic phase retrieval algorithm is applied to every generated pattern $U(m,n)_p$. The main process of the iteration process can be figured out as follows:

(i) The amplitude and phase values in the PPD plane are presented as wavefront distributions in the initial loop. The complex image $U(m,n)_p$ is propagated to the corresponding object plane by using the angular spectrum algorithm as introduced above and the result can be expressed as $U_0(\mu_0,\varepsilon_0)$.

(ii) Meanwhile, the constraint of positive absorption is imposed on this position. Based on the classical light field theory, the intensity of the terahertz wave after diffraction will decrease rather than remain the same or even increase. Therefore, the value of amplitude will be replaced with the maximum normalized parameter, and the phase is changed to zero when a violation signal occurs at the pro-propagating plane. The modified distribution with these limiting conditions is further denoted as $U_0'(\mu_0,\varepsilon_0)$.

(iii) Later, the updated image is propagated back to the initial plane where the phase value is extracted and the original amplitude value is retained, then forming a new distribution $U'(m,n)_p$ for the continued iteration. During the retrieval process, an output amplitude without any violated values is considered a sign of iterative convergence.

The phase retrieval method is capable of significantly enhancing image quality by eliminating noise, primarily originating from defocused twin images and inter-layer fusion images. These noise sources do not contribute to an increment in propagating energy within the physical space where free-space propagation takes place, and it is expected that the absorption of the light-propagating medium remains positive. In accordance with this criterion, a constraint enforcing positive absorption is incorporated into the iterative algorithm, thereby mitigating non-compliant noise during each iteration, and ultimately enhancing imaging quality through multiple iterations.

In addition, the reconstruction process is often accompanied by diffraction noise originating from binary masks, which can adversely affect image definition. Consequently, the obtained masks in formula (12) are subsequently employed to refilter the additional interference. The precise purpose of incorporating 'Filter 2' within the flowchart is to accomplish this objective.

3.5. Stereoscopic-field automated merging

In the last, high-definition reconstructed images from different axial views are automatically merged. The threshold $T*$ is selected based on the auto-focusing curve generated for the samples by the TEG and FSF algorithms. The auto-focusing algorithm produces a curve that reflects the focus quality at different positions. The threshold $T*$ is determined by analyzing the local maximum of the curve, which corresponds to the optimal focus position. Values higher than $T*$ are considered in-focus regions and are set to 1, while values below $T*$ are considered out-of-focus regions and are set to 0. When constructing the 3D image, we aim to use focused images, so the value of $T*$ must lie within a certain range around the maximum focus value (focus position); otherwise, the image will be defocused and the resulting synthetic image will be unclear. Therefore, the threshold $T*$ is selected empirically, typically by visually observing the position corresponding to the maximum focus value, which generally aligns with the human eye's observation. In this experiment, the value of $T*$ was set to 0.9. The threshold selection can be further refined using techniques such as mean square error minimization to ensure the capture of the clearest focused regions. The formula is as follows:

$$I^*(d) = K^*(d) \times I(d) = \left\{ \begin{array}{cc} I(d) & \text{if } K(d) \geqslant T^*, K^*(d) = 1 \\ 0 & \text{if } K(d) < T^*, K^*(d) = 0 \end{array} \right\}, \tag{14}$$

where $I(d)$ and $I^*(d)$ represent original and new complex amplitude distributions at stereoscopic location d, $K(d)$ and $K^*(d)$ are the values of the auto-focusing curve before and after updating. The position of the slices ranges within the retrieve distance scale and these generated masks are overlaid on them. To this, the high-quality, super-resolution, and depth-free holographic detection of AMDT can be realized in stereoscopic-field.

4. Experimental results

Two sets of experiments were conducted on AMDT. To enhance the intuitive presentation of imaging results, this article introduces two distinct sections, namely A. Pyroelectric Detection System and B. Microbolometer Detection System, which correspond to different systems. The following outlines the details of each section.

4.1. Pyroelectric detection system

In the denoising process of the pyroelectric detection system, fifty frames of background images and holograms were recorded for global optimization. Beyond the corresponding number, the MSE and CC criteria reached a relatively stable state, and the rate became almost constant. For comparison, figures 3(a) and (c) depict the unprocessed data, consisting of single images, whereas figures 3(b)–(d) show the optimized images obtained after noise reduction. Significant improvements in contrast were observed for the white dashed circles, with the unwanted smeary pattern distributions in (a) and (c) successfully removed in the optimized images of (b) and (d). Clearly, the mean filter effectively suppresses the inherent Gaussian thermal noise. The normalized hologram, obtained by dividing the optimized hologram (figure 3(d)) by the optimized background image (figure 3(b)), is shown in figure 3(e). Subsequently, the focusing values of the reconstructed images based on this optimized and normalized hologram were plotted. As shown in figure 4(f), the retrieval distance search was conducted from 40 mm to 70 mm, but only one main peak was distinguished as the focal distance representation in both the red FSF and green curves TEG. Two focus peaks should have been observed in this context; however, consecutive subpeaks can be attributed to stacking diffraction between AMDT

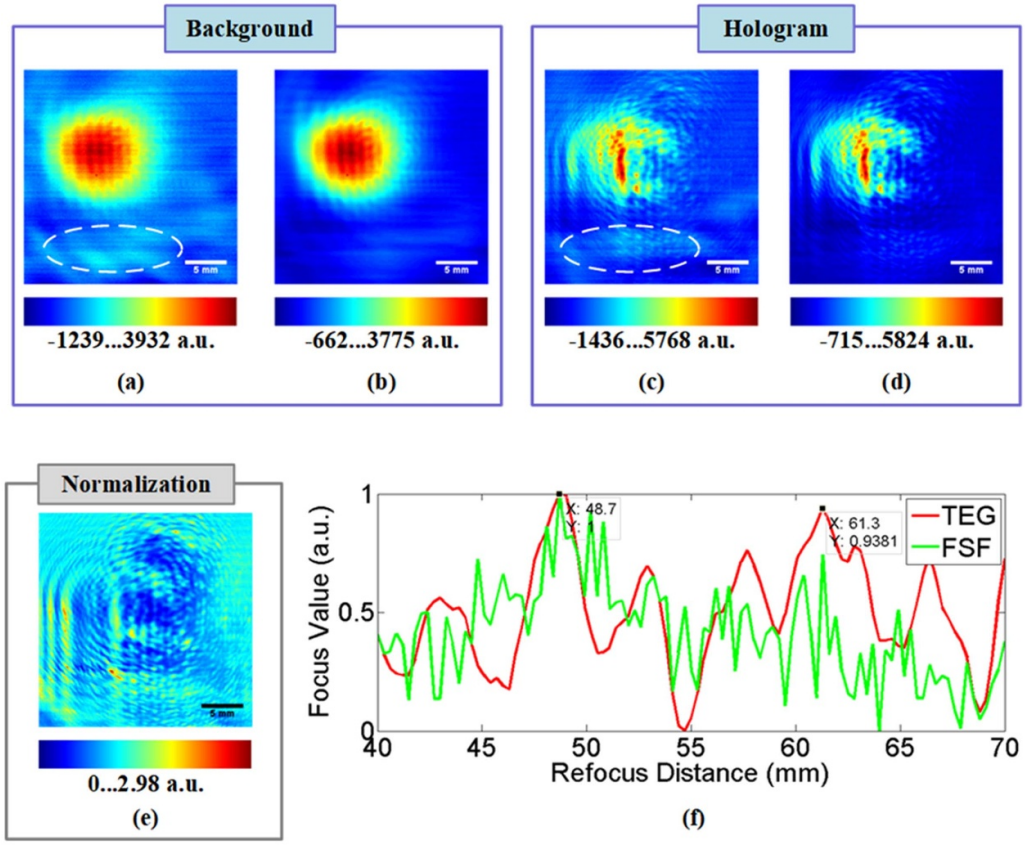

Figure 3. Global optimization of the terahertz images in the pyroelectric detection system. (a) and (b) Original and denoising background image. (c) and (d) Original and denoising hologram. (e) Normalized hologram. (f) Auto-focusing curves: the red curve from TEG and the green curve from FSF.

or hindrance caused by twin imaging. The conclusion is that ambiguous results from a direct focus approach can lead to inaccurate localization. This paper introduces advancements in pre-propagation decomposition reconstruction techniques to significantly improve target localization precision.

The specific process in the pyroelectric detection system is depicted in figure 4. First, the optimized and normalized hologram of figure 3(e) was pre-propagated, and the distance of 55 mm from the detection plane was identified as a preferred distance for pre-propagating. The selection of the pre-propagation distance was based on a specific calculation. Specifically, we first perform iterative reconstruction for each individual sample to determine its reconstruction distance. Then, we take the midpoint of the reconstruction distances of the two samples as the pre-propagation distance. Mathematically, this can be expressed as $\frac{X_p+X_p}{2}$. Where X_p and X_q are the reconstruction distances determined for each sample through iterative reconstruction. This approach ensures that the selected pre-propagation plane provides a balanced and optimal condition for the reconstruction process, making it more suitable for subsequent segmentation and analysis. The resulting amplitude and phase distributions of the diffractogram are shown in figures 4(a) and (b), respectively. According to formula (3), the theoretical resolution of this pyroelectric detector system is approximately 255 μm. Then, a binarized mask is created using Otsu's method, as illustrated in figure 4(c), based on the amplitude diffraction pattern. The exclusive masks depicting various targets, as illustrated in figures 4(d) and (e), were derived from figure 4(c) through the utilization of automated disassembling algorithms. Concurrently, these exclusive masks underwent edge mean blurring transition to effectively eliminate the impact of sharp edges.

By multiplying the complex amplitude images presented in figures 4(a) and (b) with the corresponding exclusive masks depicted in figures 4(d) and (e) individually, we successfully obtained the unattached distributions of each target module. These distributions are free from any stacking noise, as demonstrated in figure 4, which is appropriately labeled as 'decomposition'. Furthermore, figures 4(f) and (g) showcases the resulting product of the metal pin target when the left mask of (d) is applied. Additionally, figures 4(h) and (i) exhibits the images of the asparagus target acquired through the application of the right mask (e).

Utilizing the precisely defined amplitude distributions of segmented diffractograms within the framework of 'decomposition', an Auto-focusing algorithm was employed to determine the physical distances from the PPD plane to the sample planes. The corresponding outcomes are displayed in figures 4(j) and (k), wherein the magenta curve denotes the result obtained using the TEG function, while the cyan curve is derived from the FSF method. In contrast to the cluttered peaks

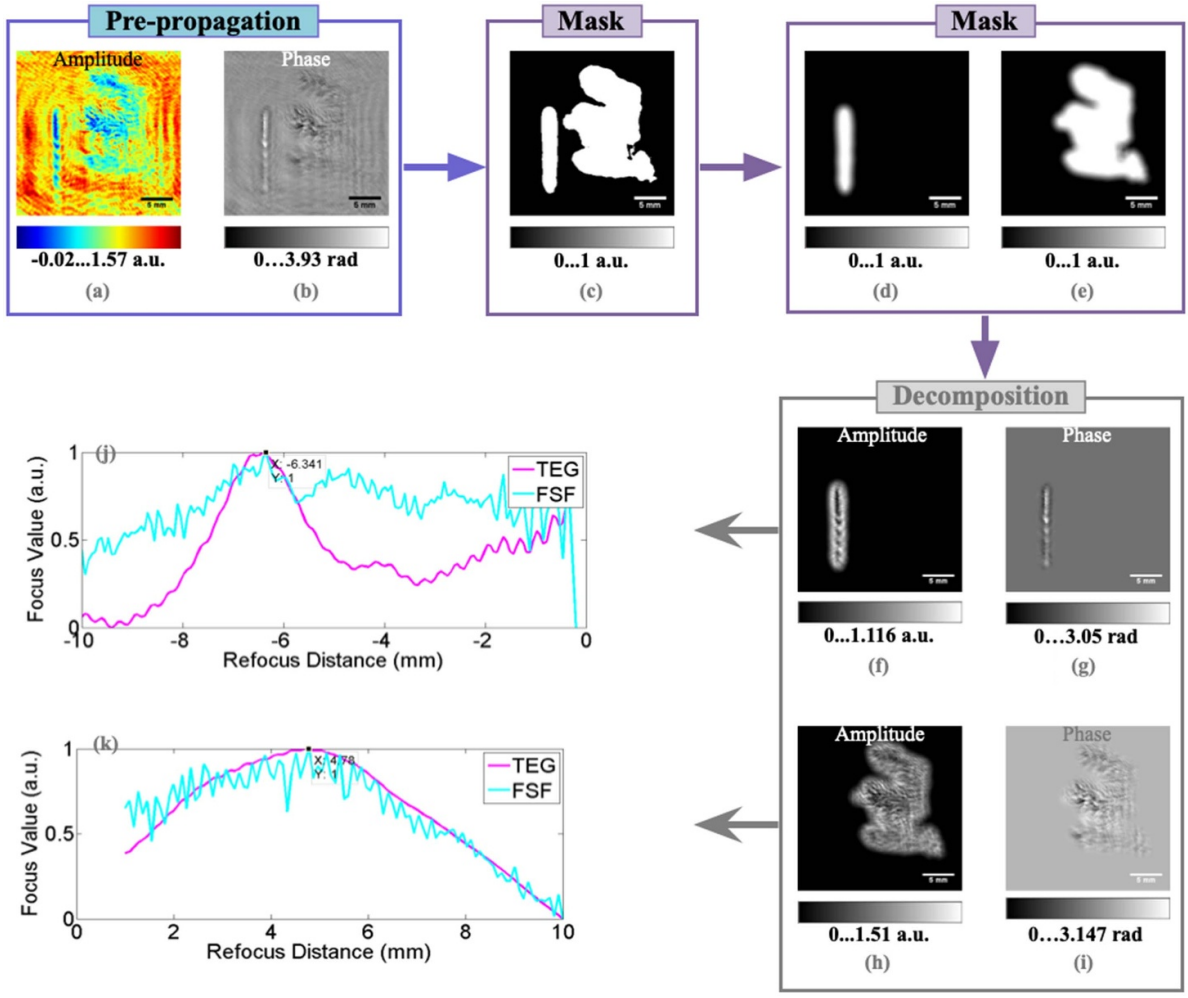

Figure 4. PPD decomposition in the pyroelectric detection system: Amplitude and phase distributions (a) and (b) at PPD plane. (c) Binary mask. (d) and (e) Smoothed and exclusive split masks. Segmented images with (f) and (g) left filter and (h) and (i) right filter. (j) and (k) are the auto-focusing results.

observed in figure 3(f), these curves exhibit elegant singular zenith points. The maximum focusing value of the metal pin occurs at a refocusing distance of $\Delta z = -6.341$ mm, whereas the position of the asparagus manifests at a distinct point of $\Delta z = 4.780$ mm. The direction of propagation in reconstruction aligns with the opposite optical path, consistent with the orientation of the z-axis in the flowchart. Consequently, the position of the asparagus plane (P1) can be determined as 50.220 mm (55.000 mm–4.780 mm) in front of the detector, while the metal pin plane (P2) is situated at 61.341 mm (55.000 mm + 6.341 mm). It is worth noting that both the critical points of the TEG and the FSF criteria happen at a consistent location which ensures the credence of the results with the pre-propagating decomposition method.

The complex amplitude results with different methods are subsequently exhibited with zoomed views and cross-sections. As a comparison, the traditional direct reconstruction with the distances of 50.220 mm (P1) and 61.341 mm (P2) from the detector is implemented as shown in figure 5. The asparagus sample produces a more delicate profile than the metal pin sample at the shorter distance, while at another position, the opposite is true, which agrees with the for-real arrangement of the target location. Besides, a quite poor SNR overwhelms plentiful details, and the disfigured diffraction information causes the samples indistinguishable. The fact that the samples can only be reconstructed finer at their respective planes makes the confocal imaging of the AMDT impossibility.

The direct reconstructed results obtained after applying the pre-propagating decomposition method are summarized in figure 6, where (a) and (b) presents the amplitude and phase distributions of the asparagus sample at a distance of $\Delta z = 4.780$ mm from the PPD plane, and (c) and (d) presents the reconstructed images of the metal pin sample at $\Delta z = -6.341$ mm. In comparison to the direct reconstruction in figure 5, the pre-propagating decomposition prevents the reconstruction of complex field information of the sample from the obstruction of disturbance. The clearness of the images in figure 6 is distinguished from the chaotic noise in figure 5. This improvement turns out that the pre-propagating

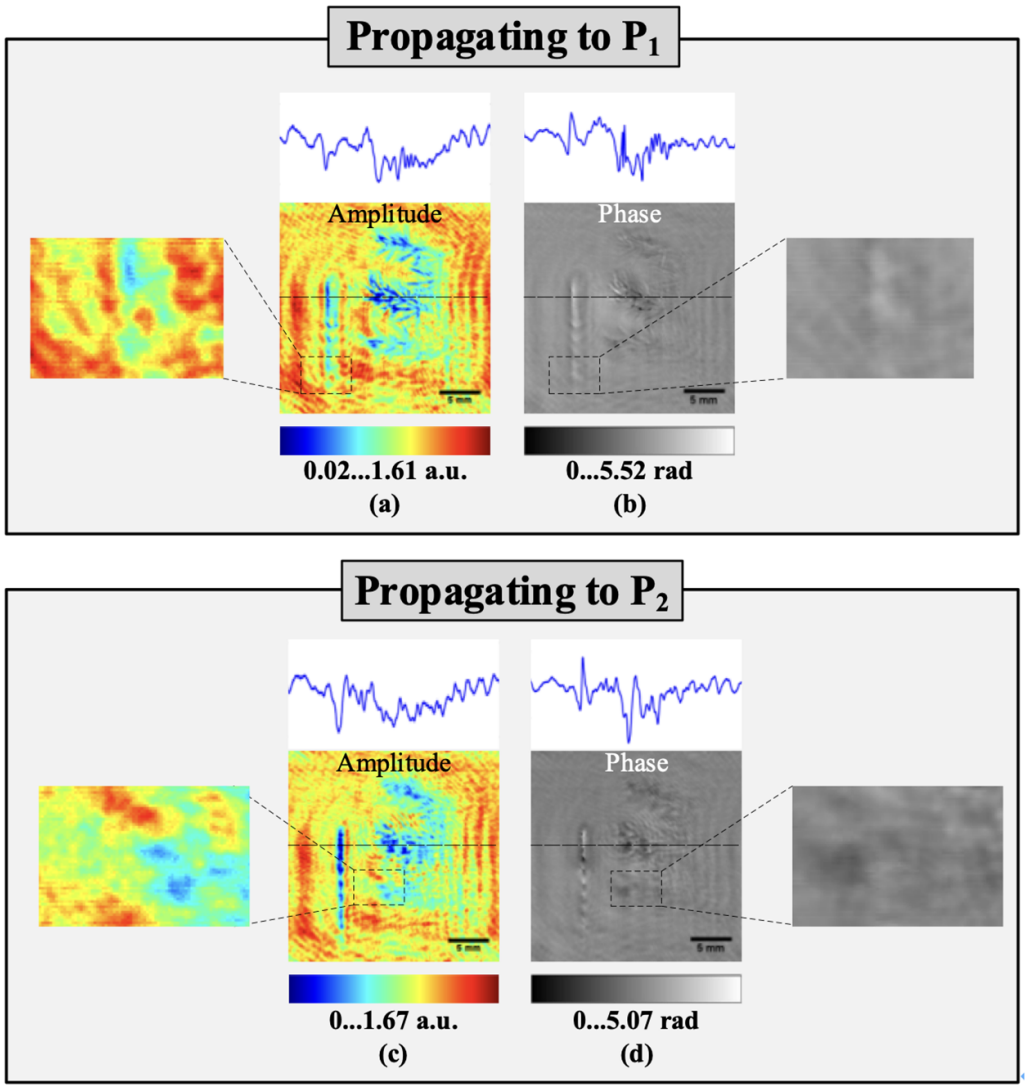

Figure 5. The direct reconstruction of normalized hologram (figure 4(e)): (a) and (b), amplitude and phase distributions of metal pin at the propagating plane of 50.220 mm and (c) and (d), amplitude and phase distributions of asparagus sample at the propagating plane of 61.341 mm.

decomposition process can boost the application of holography to reconstruct thick sample information. By zooming an arbitrary region in figure 6, it can be noticed that the existence of the defocused twin image leads to the failure to differentiate crucial details.

To accurately restore the real light field of AMDT after applying the pre-propagating decomposition method, the conventional direct reconstruction is substituted with phase retrieval and unwrapping reconstruction techniques. After fifty successive iterative propagation between the PPD plane and target planes, the resulting amplitude and phase distributions of both samples are displayed in figure 7. It can be seen that the image qualities had been enhanced considerably, neither diffraction information nor the twin image is preserved. As shown in the zoomed views, compared with the other previous approaches, the SNR and quality are substantially improved. The integral appearance of the asparagus sample and bristle phylloclade has been shown clearly. There are about ten branches in clusters of 4–5 mm in length, and the scale of the metal pin becomes measurable. In addition, profile curves along the white cross-sections in the individual images, have shown the highest concentration that fluctuates with sample presence in the images, making the selection of focus positions more credible.

A comprehensive comparison is subsequently presented. The phase retrieval reconstruction without PPD decomposition had been carried out. With phase retrieval iteration, figures 8(a) and (d) exhibits the complex amplitude distributions at the propagation plane of 50.220 mm and 61.341 mm. It can be seen that samples can only be reconstructed at their respective planes. Even if the twin image is swept away, the desire for synchronous reconstruction is still a serious shortcoming beyond neglect. In contrast it, figures 8(e) and (f) shows the synthetic images of direct reconstructed results after applying the PPD decomposition method (as shown in figure 6). Due to the absence of phase retrieval method,

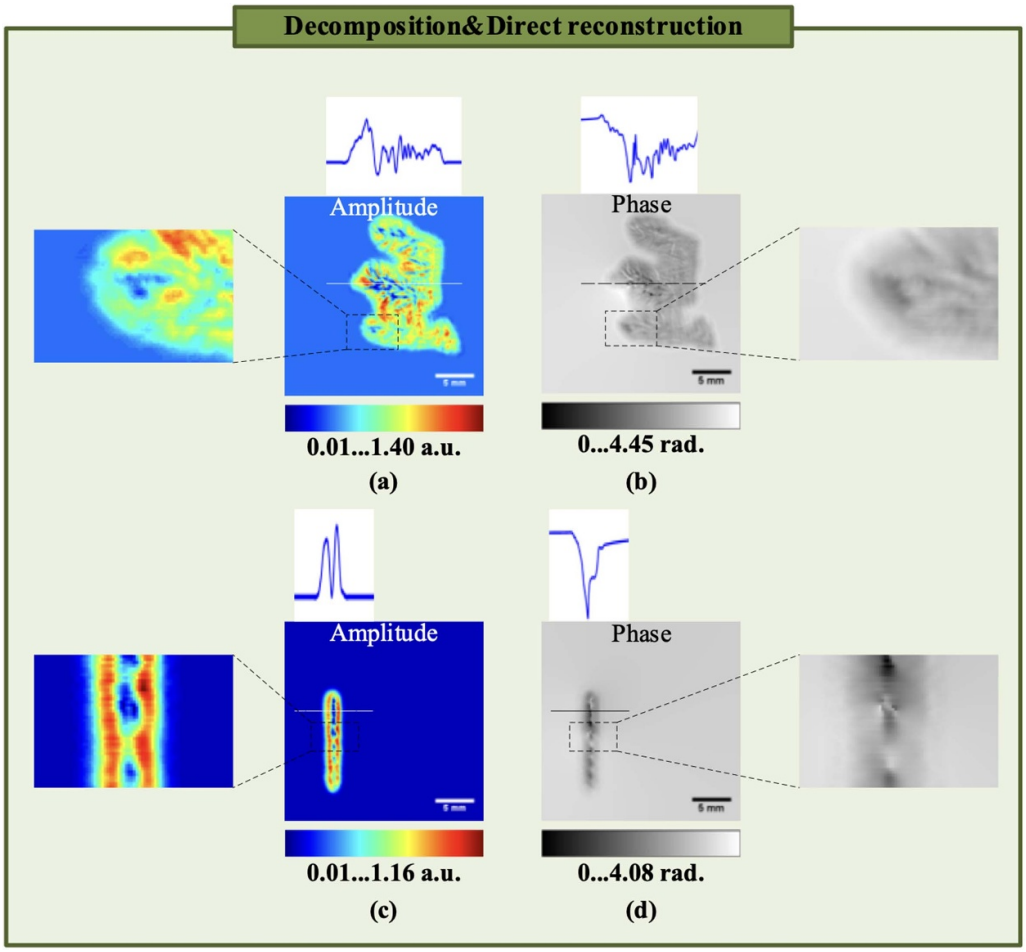

Figure 6. Direct reconstruction after PPD decomposition: the amplitude and phase profile at the distance of (a) and (b) $\Delta z = 4.780$ mm for the asparagus sample and (c) and (d) $\Delta z = -6.431$ mm for the metal pin sample from PPD plane.

the artifacts from the twin image are overlapping on sample images. The synthetic amplitude and phase images that are reconstructed by the proposed STDSH methods (as shown in figure 7) are shown in figures 8(g) and (h). Both the disturbance of diffraction information and twin images are erased.

To compare the reproduced effect with different approaches more intuitively, the pixel intensity curves of the cross-sections are extracted and shown in figures 8(i) and (j). It can be observed that highly concentrated peaks are registered in the results of the proposed STDSH technique, revealing also the high-resolution imaging of the AMDT. Besides, the noise of masks is re-filtered by applying themselves to corresponding results again. Taken together, the STDSH model proposed in this paper is feasible for sensing the AMDT, and each procedure in the method is indispensable.

4.2. Microbolometer detection system

The flexibility and applicability of the aforementioned STDSH model have been validated using other self-developed instruments, such as the microbolometer detection system depicted in figure 2. To practice the expanded FOV method, there is a 2D electric displacement stage employed in the sampling process. To ensure an overlapping pixel area, the moving steps are set to be less than the actual size of the chip sensor in horizontal and vertical directions. Enhanced synthetic aperture operation is introduced to record enough high-frequency fringes at supplementary sensing area.

Figures 9(a) and (b) shows the splicing terahertz background image and hologram which were taken before and after placing objects. The size of the images collected before merging is 320 $\times$ 240 pixels, while eight sub-images are spliced with the center to obtain a patulous scale of 580 $\times$ 580 pixels. Besides, a rhythmic vibration is brought by the refrigerating machine with 0.5 Hz, yet the corresponding captured maximum frequency of the camera is 8.5 Hz. Discrete down-sampling pattern can avert this fluctuating noise which is distinct from the global optimization in the pyroelectric detection system. The collection interval of seventeen frames among the sequent sub-images is carried out in this discrete optimization method. Figure 9(c) shows the normalized hologram which is used for the evaluation of reconstructed distances by auto-focusing algorithms. In figure 9(f), the red curve calculated by the TEG criterion produces a maximum point at the position of 27.4 mm, while the peak of the green curve from FSF appears at 16.2 mm. Other than that, the occurrence of other sub-peaks suggests that there is some inevitable deviation with automated searching.

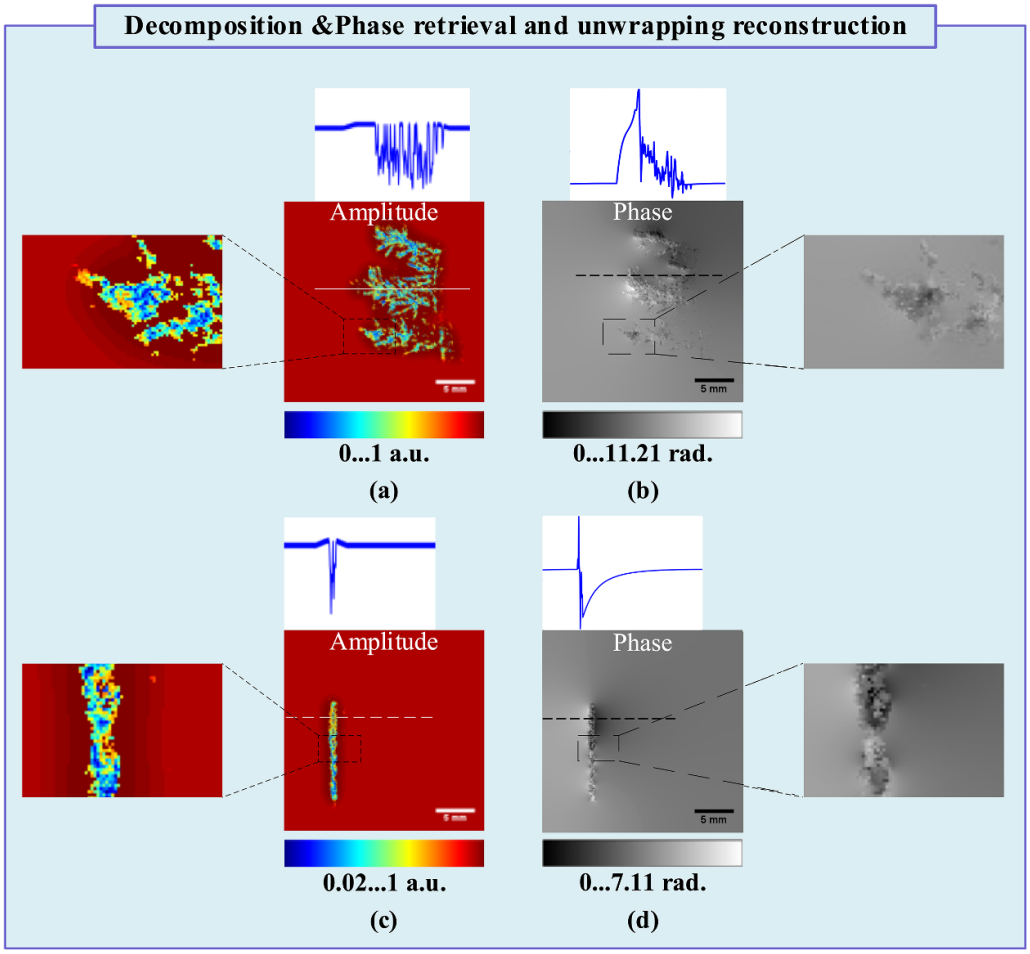

Figure 7. Phase retrieval and unwrapping reconstruction after PPD decomposition: the amplitude and phase images at a distance of $\Delta z = 4.780$ mm for asparagus sample (a) and (b) and $\Delta z = -6.341$ mm for metal pin sample (c) and (d) from the PPD plane.

Instead of any biased refocus distance in figure 9(d), the enhanced and normalized hologram is refrozen at the PPD plane with a distance of 21.8 mm. Based on equation (3), the expanded FOV method can improve the resolution of the microbolometer system from ∼426 μm to ∼155 μm corresponding to the pixel numbers of 580 and a wavelength of 97 μm. The pre-propagating patterns are shown in figures 10(a) and (b). The threshold mask as shown in figure 10(c) is designed from the amplitude result (a). Figures 10(d) and (e) is the exclusive masks that belong to different target sections. Discrete amplitude and phase results of different targets are presented in the 'decomposition' frame. Figures 10(f) and (g) pertains to a steel needle with a hole, and figures 10(h) and (i) holds the morphology of a silicon structure. The amplitude images were used as the basis for the adjudication of reconstructed distance. Retrieving from the PPD plane to the hologram plane, figures 10(j) and (k) displays the measured value of each propagation that is associated with variable depth. Both the magenta curve outlined by TEG function and the cyan curve of FSF criterion reach a peak position of $\Delta z = -5.503$ mm in the pinhole module. For the silicon structure, the quantitative calculation based on the FSF criterion function approximately happens at sub-maximum values. Therefore, the position that coincides with the peak value of TEG function is defined as the credible reconstructed arguments. Hence, the misfit of 6.707 mm in the cyan curve is discarded while the distance of $\Delta z = 5.630$ mm is retained for reconstruction.

With cross-sectional profile along the white dotted lines of the pictures in figure 11, the results of direct reconstruction by propagating the discrete targets to a distance of 16.297 mm (21.8 mm–5.503 mm) are shown in figures 11(a) and (b). The reproduced image at a distance of 27.430 mm (21.8 mm + 5.63 mm) is shown in figures 11(c) and (d). It can be discerned in the zoomed views that the hole of steel needle and a hockey stick shape of the silicon structure can be hunted explicitly. Although the twin image overlapping is still present, the sharpness of involved edges is reconstructed well which indicates that the PPD decomposition made great contributions to improving reconstructed image quality.

Figure 12 shows the reconstructed results of the proposed STDSH method. With the cross-sectional profile along the white dotted lines and zoomed views, figures 12(a) and (b) is the amplitude and phase distributions by propagating the

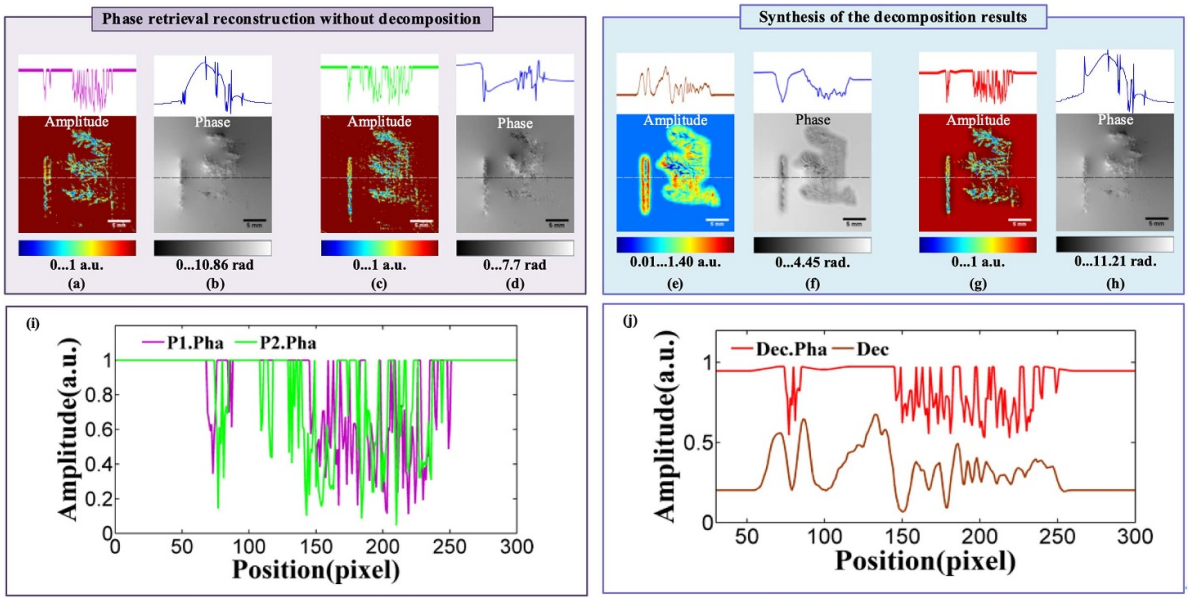

Figure 8. Comparison of different reconstructed results. Phase retrieval reconstruction without decomposition: the amplitude and the phase images at a propagation plane of (a) and (b) 50.220 mm and (c) and (d) at 61.341 mm. Synthesis of reconstruction results following the implementation of the PPD decomposition technique: amplitude and phase images obtained using the conventional direct reconstruction approach (e) and (f), alongside the phase retrieval and unwrapping reconstruction method (g) and (h). (i) and (j) The comparison of the cross-sections in above images. (P1.Pha and P2.Pha represent the focusing curves of amplitude images reconstructed by phase retrieval and unwrapping for the asparagus and pinhole samples at their respective propagation distances. Dec represents the focusing curve of the amplitude image reconstructed for the asparagus and pinhole samples after decomposition using the PPD method. Dec.Pha is the focusing curve of the synthesized image obtained after phase retrieval reconstruction of the decomposed asparagus and pinhole samples using the PPD method.).

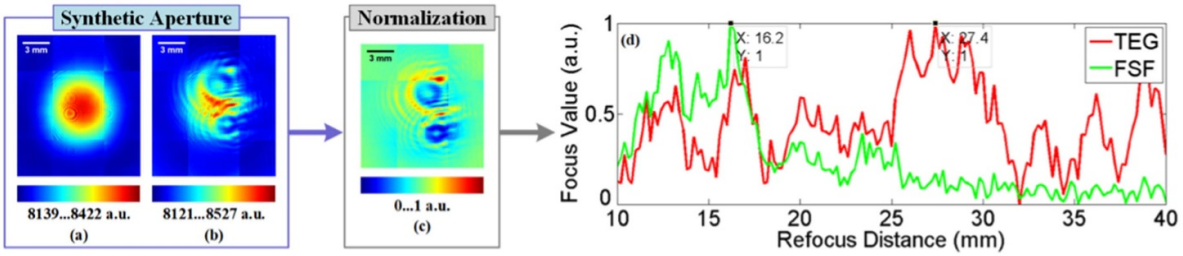

Figure 9. Discrete optimization and enhancement of the images in the microbolometer system. (a) Splicing background image. (b) Splicing hologram. (c) Normalized hologram. (d) Auto-focusing curves: the red curve from TEG and the green curve from FSF.

discrete targets to the distance of $\Delta z = -5.503$ mm and figures 12(b) and (c) shows the reconstruction results at a distance of $\Delta z = 5.630$ mm from the PPD plane. In comparison with the direct reconstruction without phase retrieval (as shown in figure 11), both the clarity and quality of these images are polished considerably. In reconstruction, the cooperation of PPD decomposition method and the phase retrieval and unwrapping technology can not only wipe out the disturbance from different diffraction orders but also remove the twin images and phase fault.

The direct and phase retrieval reconstructed results of the silicon arm microstructure are presented in figures 13(a) and (d). The zoomed views in 'Pha. Rec', show blurred features of the out-of-focus samples. The 'synthesis' box is utilized to present a uniform visualization of the reproduced discrete images in figures 11 and 12. The results of figures 13(e) and (f) carry twin artifacts due to the in-line holographic setup. On the contrary, the realistic distributions in figures 13(g) and (h) preserve the edges and most of the objective details with less blurring and suppression of twin artifacts. The intensity profiles along the cross-sectional white dash line of the reconstructed amplitude images are as shown in figures 13(i) and (j) where the intensity profile of the images reconstructed by the proposed STDSH method shows well defined peaks. Besides, the reconstructed images are filtered again with smoothed masks to suppress the excess noise. The results in the microbolometer detection system are consistent with the performance capacity of the pyroelectric system and demonstrated the ability of terahertz depth-free holography to sense the details in stereoscopic-field.

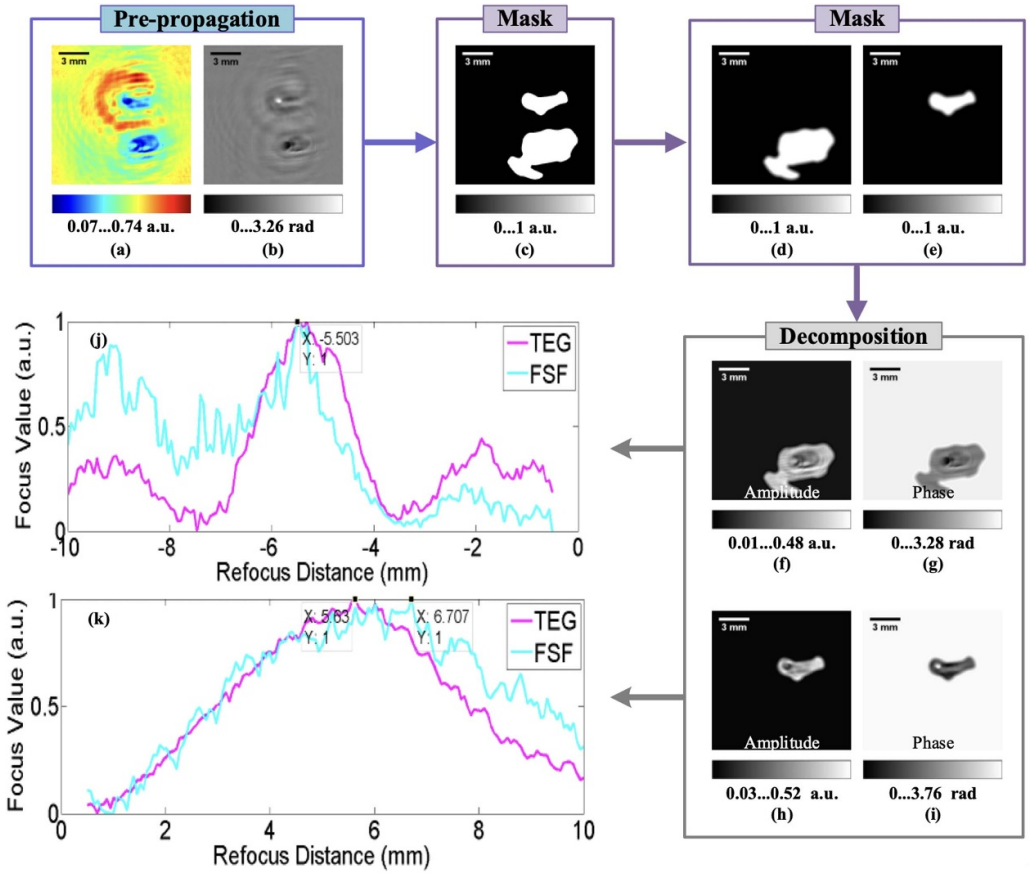

Figure 10. PPD decomposition in the microbolometer system. Amplitude and phase distributions (a) and (b) at PPD plane. (c) Binary mask. (d) and (e) Masks after decomposition and smoothing. Decomposition images after (f) and (g) left filter and (h) and (i) right filter. (j) and (k) Auto-focusing curves.

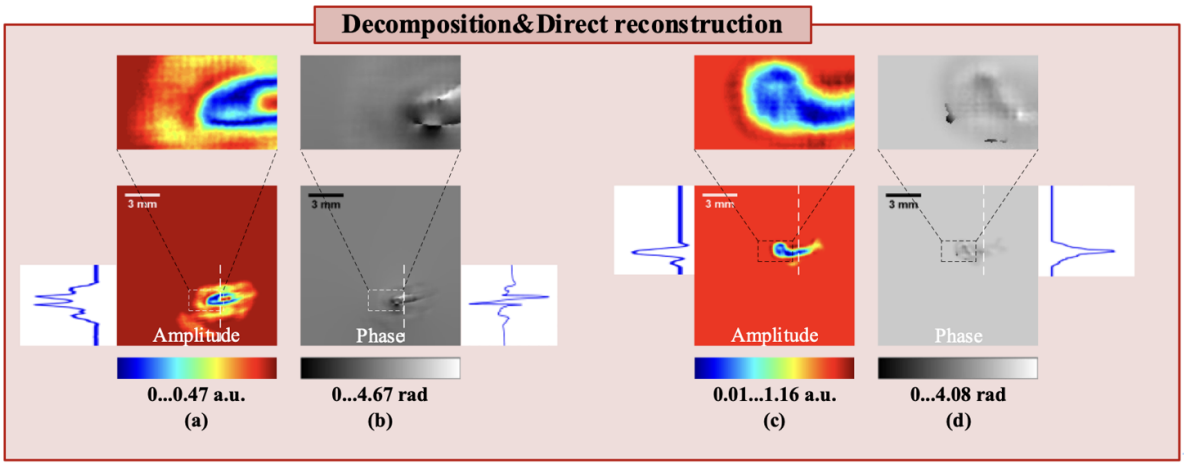

Figure 11. Direct reconstruction after PPD decomposition: amplitude and phase images at a propagation plane of (a) and (b) 16.297 mm and (c) and (d) 27.430 mm.

The simultaneous presentation of the AMDT in different systems is illustrated in figure 14. A total of 100 pre-propagated images were meticulously chosen with a precise step size of 0.2 mm. Base on formula (14), the amplitude distributions, excluding the target focusing position, are selectively harnessed by incorporating an imperceptible mask with zero values. This technique effectively enhances the emphasis on the focused image. Finally, the stereoscopic-field confocal distribution of different targets can be obtained by automatically synthesizing the acquired images using an automated merging algorithm.

The vivid image of the asparagus fern and a metal pin with various directional visions are observed, as depicted in figure 14(a). The results for a steel needle and a silicon

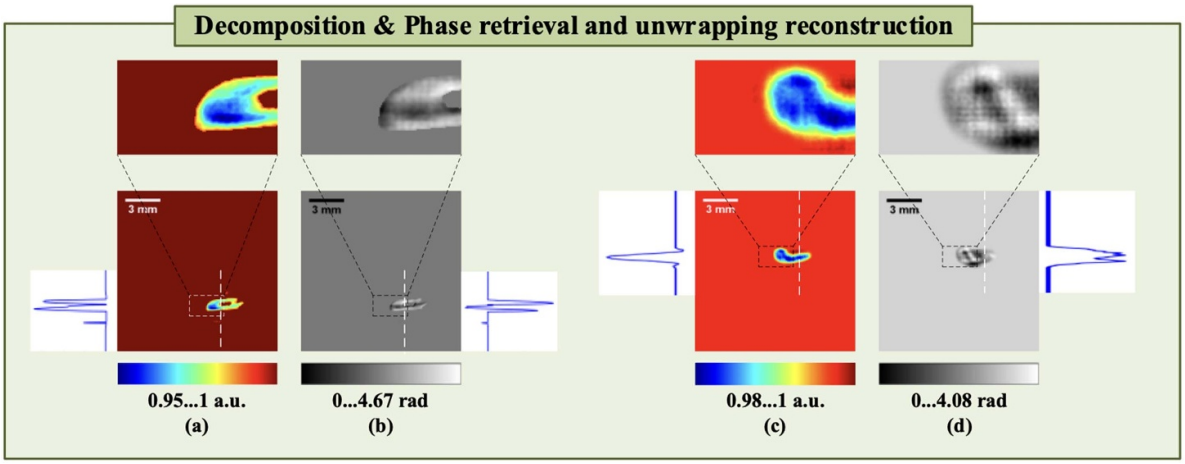

Figure 12. Phase retrieval and unwrapping reconstruction after PPD decomposition: (a) and (b) amplitude and phase images at a propagation plane of $\Delta z = -5.503$ mm and (c) and (d) $\Delta z = 5.630$ mm.

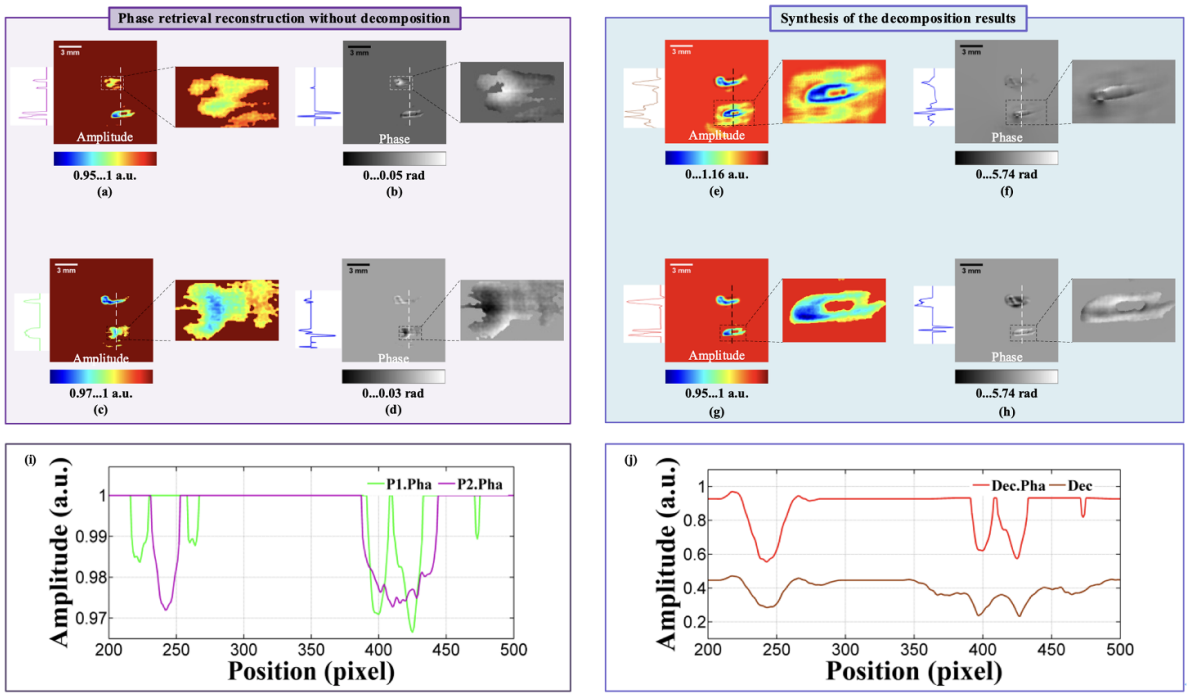

Figure 13. Reconstructed results with different methods. Phase retrieval reconstruction without decomposition: complex amplitude distributions at the distance of (a) and (b) $\Delta z = 5.630$ mm and (c) and (d) $\Delta z = -5.503$ mm. Synthesis of reconstruction results following the implementation of the PPD decomposition technique: amplitude and phase images obtained using the conventional direct reconstruction approach (e) and (f), alongside the phase retrieval and unwrapping reconstruction method (g) and (h). (i) and (j) The comparison of cross-sections in above images. (P1.Pha and P2.Pha represent the focusing curves of amplitude images reconstructed by phase retrieval and unwrapping for the asparagus and pinhole samples at their respective propagation distances. Dec represents the focusing curve of the amplitude image reconstructed for the asparagus and pinhole samples after decomposition using the PPD method. Dec.Pha is the focusing curve of the synthesized image obtained after phase retrieval reconstruction of the decomposed asparagus and pinhole samples using the PPD method.).

structure can be seen in figure 14(b). In summary, the automatic synthesis process can be perceived as the reconfiguration of slices, upon which a mask is applied. This process relies on the determination of the autofocusing, enabling the complete elimination of out-of-focus reconstructions in multi-depth targets by preserving the focused image and segregating the out-of-focus image. Furthermore, this process enables the conversion of the original 2D plane display into a 3D perspective, accurately quantifying the practical distance in all directions.

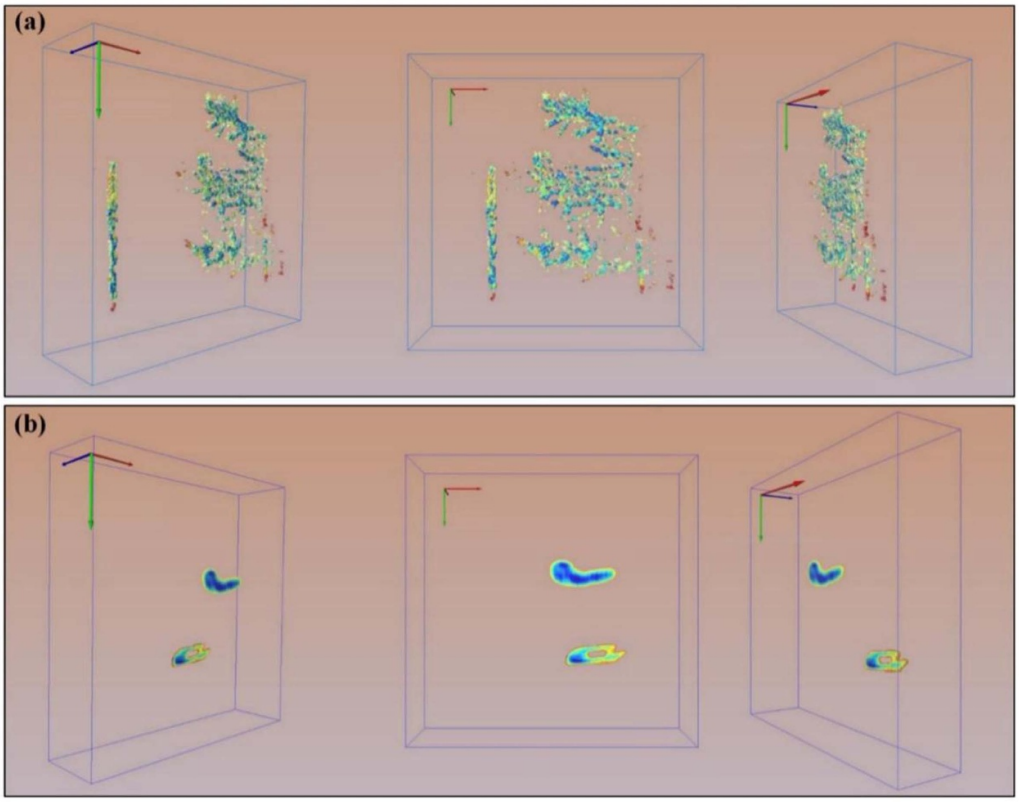

Figure 14. Stereoscopic-field vision of AMDT from various directions: (a) asparagus fern and a metal pin. (b) steel needle with a hole and a silicon structure.

5. Conclusions

The proposed STDSH method successfully reconstructed the stereoscopic-field scene of AMDT and extended the imaging spatial depth. For the critical challenge of mutual disturbance from the stacking interference patterns, pre-propagating diffraction decomposition based on the binarization masks and phase retrieval propagation algorithm can filter the redundant information by the iterations between objects and reproduced discrete images. Global and discrete optimization can effectively eliminate background clutter as well as the SNR of normalized holograms. The implication of the expanded FOV method enables sufficient high-frequency details in a precise acquisition with intensity calibration and subpixel registration. In the reconstructed process, the phase retrieval and unwrapping skills can improve the accuracy of auto-focusing. The automatic stereoscopic-field synthesis truly achieves the depth-free perception of AMDT. Different comparative terahertz layouts were built and the positive experimental verification confirms that the proposed STDSH method is flexible and robust in sensing AMDT with only a single exposure.

The study conducted in this work offers a rapprochement for the comprehensive application of terahertz wave and digital holography technology which not only gains the advantage of high-quality and superior resolution but also is viable to apply in other complex systems with a release of hardware limitations. We think that the STDSH strategy has no restrictions on incident waves which can also be applied to the other bands. The proposed technique can be useful for real-time monitoring and support the application fields, such as medical imaging technology, wavefront sensing, immersive visualization, and remote sensing of 3D environments. Future research will focus on addressing the limitations, such as target distance variability and overlap. Techniques like synthetic aperture, resolution enhancement, and machine learning are expected to improve performance.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

This work is supported by the National Natural Science Foundation of China (61805214, 42072087), Open Fund of State Key Laboratory of Infrared Physics (SITP-NLIST-YB-2025-12), Piesat Information Technology remote sensing interdisciplinary research Project (HTHT202202) the Fundamental Research Funds for the Central Universities (2-9-2024-203). Young Elite Scientists Sponsorship Program by Bast (BYESS2020037), Frontiers Science Center for Deep-time Digital Earth (2652023001).

Author contributions

Conceptualization, H H and Z Z; Data curation H H and M Y; Investigation H H and Z L; Methodology H H and S P; Writing—review and editing, H H and M Y Supervision Z Z All authors have read and agreed to the published version of the manuscript. (H H, M Y and Z L contributed equally to this work).

Conflict of interest

The authors declare no conflicts of interest.

Disclaimer/Publisher's Note

The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

ORCID iD

Haochong Huang https://orcid.org/0000-0002-0897-6695

References

[1] Hall M L, McCall P, Towers C E and Towers D P 2021 Improved three-dimensional localization of multiple small objects in close proximity in digital holography *Appl. Opt.* **60** A285–A95
[2] Zhang Y, Zheng H, Wang Y, Wang R and Guo L 2021 Investigation on THz EM wave scattering from oil-covered sea surface: exploration for an approach to probe the thickness of oil film *IEEE Trans. Geosci. Remote Sens.* **59** 1827–35
[3] Zhao X, Sun P, Xu Z, Min H and Yu H 2020 Fusion of 3D LIDAR and camera data for object detection in autonomous vehicle applications *IEEE Sens. J.* **20** 4901–13
[4] Levin L A *et al* 2019 Global observing needs in the deep ocean *Front. Mar. Sci.* **6** 1–32
[5] Zhang X, Liu L, Wu C, Chen X, Gao Y, Xie S and Zhang B 2020 Development of a global 30m impervious surface map using multisource and multitemporal remote sensing datasets with the Google Earth Engine platform *Earth Syst. Sci. Data* **12** 1625–48
[6] Tmušić G *et al* 2020 Current practices in UAS-based environmental monitoring *Remote Sens.* **12** 1001
[7] Sarbolandi H, Plack M and Kolb A 2018 Pulse based time-of-flight range sensing *Sensors* **18** 1679
[8] Qiu Y, Du S, Li L and Xu Y 2020 A simulation study of time-of-flight mass spectrometry with an angular chirp enhanced delay cavity *Int. J. Mass Spectrom.* **456** 116387
[9] Lyu N, Zuo J, Zhao Y and Zhang C 2021 Layer-resolving terahertz light-field imaging based on angular intensity filtering method *Sensors* **21** 7451
[10] Fan Q *et al* 2022 Trilobite-inspired neural nanophotonic light-field camera with extreme depth-of-field *Nat. Commun.* **13** 1–10
[11] Wätzel J, Sherman E Y and Berakdar J 2020 Nanostructures in structured light: photoinduced spin and orbital electron dynamics *Phys. Rev. B* **101** 1–10
[12] Jana K, Okocha E, Møller S H, Mi Y, Sederberg S and Corkum P B 2022 Reconfigurable terahertz metasurfaces coherently controlled by wavelength-scale-structured light *Nanophotonics* **11** 787–95
[13] Guan S, Khan A A, Sikdar S and Chitnis P V 2020 Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal *IEEE J. Biomed. Health Inform.* **24** 568–76
[14] Coupland J and Lobera J 2008 Optical tomography and digital holography *Meas. Sci. Technol.* **19** 070101
[15] Guo Q, Shi Z, Huang Y-W, Alexander E, Qiu C-W, Capasso F and Zickler T 2019 Compact single-shot metalens depth sensors inspired by eyes of jumping spiders *Proc. Natl Acad. Sci. USA* **116** 22959–65
[16] Chen N, Zuo C, Lam E Y and Lee B 2018 3D imaging based on depth measurement technologies *Sensors* **18** 1–38
[17] Gao H, Fan X, Xiong W and Hong M 2021 Recent advances in optical dynamic meta-holography *Opto-Electron. Adv.* **4** 210030
[18] Huang H, Li Z, Zhang Q, Hui J, Zheng Z, Sun D and Panezai S 2024 Infrared digital holography *IEEE Trans. Instrum. Meas.* **73** 4508337
[19] Huang H, Huang H, Zheng Z and Gao L 2025 Insights into the infrared crystal phase characteristics based on deep learning holography with attention residual network *J. Mater. Chem. A* **13** 6009–19
[20] Kozacki T, Mikuła-Zdańkowska M, Martinez-Carranza J and Idicula M S 2021 Single-shot digital multiplexed holography for the measurement of deep shapes *Opt. Express* **29** 21965
[21] Goodman J W and Lawrence R W 1967 Digital image formation from electronically detected holograms *Appl. Phys. Lett.* **11** 77–79
[22] Jaroslavski L P and Merzlyakov N S 1979 Information display using the methods of digital holography *Comput. Graph. Image Process.* **10** 1–29
[23] Shi L, Li B, Kim C, Kellnhofer P and Matusik W 2021 Towards real-time photorealistic 3D holography with deep neural networks *Nature* **591** 234–9
[24] Zhai Y, Huang H, Sun D, Panezai S, Li Z, Qiu K, Li M, Zheng Z and Zhang Z 2024 End-to-end infrared radiation sensing technique based on holography-guided visual attention network *Opt. Lasers Eng.* **178** 108201
[25] Li Z, Huang H, Sun D, Zheng Z, Wang F, Panezai S, Xing J, Yang Y and Qiu K 2024 Deep learning infrared holography with transformer for crystal material characterization *Cryst. Growth Des.* **24** 6851–64
[26] Takaki Y and Ohzu H 1999 Fast numerical reconstruction technique for high-resolution hybrid holographic microscopy *Appl. Opt.* **38** 2204–11
[27] Huang H, Yuan E, Zhang D, Sun D, Yang M, Zheng Z, Zhang Z, Gao L, Panezai S and Qiu K 2023 Free field of view infrared digital holography for mineral crystallization *Cryst. Growth Des.* **23** 7992–8008
[28] Talapatra S and Katz J 2012 Three-dimensional velocity measurements in a roughness sublayer using microscopic digital in-line holography and optical index matching *Meas. Sci. Technol.* **24** 024004
[29] Hossain M M, Sheoran G, Mehta D S and Shakher C 2007 Contouring of diffused objects by using digital holography *Opt. Lasers Eng.* **45** 684–9
[30] Yuan E, Huang H, Zhang D, Zheng Z, Sun D, Dong J and Qiu Y 2024 Real-time three-dimensional monitoring and analysis of perovskite solar cell using short wavelength infrared digital holography *Appl. Phys. Lett.* **125** 234102
[31] Qiu Y, Huang H, Zhai Y, Zheng Z and Sun D 2025 High-fidelity infrared U-shaped residual network digital

holography with attention module for crystal growth observation *Opt. Lett.* **50** 1905
[32] Murata S and Yasuda N 2000 Potential of digital holography in particle measurement *Opt. Laser Technol.* **32** 567–74
[33] Kang W, Zheng Z and Huang H 2025 Advancements in digital holography for crystalline material characterization: a review (Invited) *Cryst. Res. Technol.* **60** 2400264
[34] Sun Z, Li C, Gao X and Fang G 2015 Minimum-entropy-based adaptive focusing algorithm for image reconstruction of terahertz single-frequency holography with improved depth of focus *IEEE Trans. Geosci. Remote Sens.* **53** 519–26
[35] Zhang X C, Shkurinov A and Zhang Y 2017 Extreme terahertz science *Nat. Photon.* **11** 16–18
[36] Tonouchi M 2007 Cutting-edge terahertz technology *Nat. Photon.* **1** 97–105
[37] Jepsen P U, Cooke D G and Koch M 2011 Terahertz spectroscopy and imaging—Modern techniques and applications *Laser Photonics Rev.* **5** 124–66
[38] Santavicca D F, Reulet B, Karasik B S, Pereverzev S V, Olaya D, Gershenson M E, Frunzio L and Prober D E 2010 Energy resolution of terahertz single-photon-sensitive bolometric detectors *Appl. Phys. Lett.* **96** 1–4
[39] Huang H *et al* 2025 Holographic characterization of typical silicate minerals by terahertz time-domain spectroscopy *Appl. Clay Sci.* **267** 107720
[40] Uerboukha H I G and Allappan K A N 2018 Toward real-time terahertz imaging *Adv. Opt. Photon.* **10** 843–938
[41] Huang H, Liu Z, Ruggiero M T, Zheng Z, Qiu K, Li S, Zhang Z and Zhang Z 2025 Terahertz geoscience: tHz time-domain spectroscopy for mineral materials *Cryst. Growth Des.* (https://doi.org/10.1021/acs.cgd.4c01423)
[42] Locatelli M, Ravaro M, Bartalini S, Consolino L, Vitiello M S, Cicchi R, Pavone F and De Natale P 2015 Real-time terahertz digital holography with a quantum cascade laser *Sci. Rep.* **5** 1–7
[43] Shi X, Wang D, Rong L, Zhao J and Wang Y 2020 Phase contrast imaging based on continuous-wave terahertz digital holography *Guangdian Gongcheng/Opto-Electron. Eng.* **47** 42–54
[44] Wan M, Healy J J and Sheridan J T 2020 Terahertz phase imaging and biomedical applications *Opt. Laser Technol.* **122** 105859
[45] Huang H *et al* 2019 Continuous-wave terahertz high-resolution imaging via synthetic hologram extrapolation method using pyroelectric detector *Opt. Laser Technol.* **120** 105683
[46] Nan Y, Huang X, Member S, Gao X, Guo Y J and Abstract A 2021 3-D terahertz imaging based on piecewise constant doppler algorithm and step- frequency continuous-wave signaling *IEEE Trans. Geosci. Remote Sens.* **59** 6771–83
[47] Ren Z, Chen N and Lam E Y 2017 Automatic focusing for multisectional objects in digital holography using the structure tensor *Opt. Lett.* **42** 1720–3
[48] Latychevskaia T and Fink H-W 2007 Solution to the twin image problem in holography *Phys. Rev. Lett.* **98** 233091
[49] Huang H, Wang D, Rong L, Panezai S, Zhang D, Qiu P, Gao L, Gao H, Zheng H and Zheng Z 2018 Continuous-wave off-axis and in-line terahertz digital holography with phase unwrapping and phase autofocusing *Opt. Commun.* **426** 612–22
[50] Liu C, Jiang Z, Wang X, Zheng Y, Zheng Y-W and Wang Q-H 2022 Continuous optical zoom microscope with extended depth of field and 3D reconstruction *PhotoniX* **3** 20
[51] Latychevskaia T, Gehri F and Fink H-W 2010 Depth-resolved holographic reconstructions by three-dimensional deconvolution *Opt. Express* **18** 22527–44
[52] Ma M, Liang W, Qin F, Guan Q, Zhong X, Deng H and Wang Z 2024 Extending the depth-of-field of computational ghost imaging: computational refocus via *in situ* point spread function estimation *Appl. Phys. Lett.* **124** 021106
[53] Wu Y, Rivenson Y, Zhang Y, Wei Z, Günaydin H, Lin X and Ozcan A 2018 Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery *Optica* **5** 704–10
[54] Yi F, Jeong O, Moon I and Javidi B 2021 Deep learning integral imaging for three-dimensional visualization, object detection, and segmentation *Opt. Lasers Eng.* **146** 106695
[55] Otsu N 1979 A tlreshold selection method from gray-level histograms *IEEE Trans. Syst. Man Cybern.* **9** 62–66
[56] Huang H, Wang D, Rong L, Zhou X, Li Z and Wang Y 2015 Application of autofocusing methods in continuous-wave terahertz in-line digital holography *Opt. Commun.* **346** 93–98
[57] Hu J, Li Q and Zhou Y 2016 Support-domain constrained phase retrieval algorithms in terahertz in-line digital holography reconstruction of a nonisolated amplitude object *Appl. Opt.* **55** 379–86
[58] Li Y, Xiao W, Pan F and Rong L 2015 Phase unwrapping method based on multiple recording distances for digital holographic microscopy *Opt. Commun.* **346** 38–42
[59] Huang H, Rong L, Wang D, Li W, Deng Q, Li B, Wang Y, Zhan Z, Wang X and Wu W 2016 Synthetic aperture in terahertz in-line digital holography for resolution enhancement *Appl. Opt.* **55** A43–A48
[60] Bhattacharya P 1996 Connected component labeling for binary images on a reconfigurable mesh architecture *J. Syst. Arch.* **42** 309–13
[61] Bshouty N H and Tamon C 1996 On the Fourier spectrum of monotone functions *J. ACM* **43** 747–70
[62] Zebker H A and Lu Y 1998 Phase unwrap algorithms for radar interferometry: residue-cut, least-squares, and synthesis algorithms *J. Opt. Soc. Am. A* **15** 586–98