Laser & Photonics Reviews

LASER & PHOTONICS REVIEWS

www.lpr-journal.org

RESEARCH ARTICLE

Inverse Synthetic Aperture Macroscopic Fourier Ptychographic Imaging of Moving Targets

Sheng Li[1,2,3] | Bowen Wang[1,2,3] | Haitao Guan[1,2,3] | Jiasong Sun[1,2,3] | Kunyao Liang[1,2,3] | Zihao Pei[1,2,3] | Min Zeng[1,2,3] | Shang Bao[1,2,3] | Chunting Xu[1,2] | Yuzhen Zhang[3] | Qian Chen[3] | Chao Zuo[1,2,3]

[1]Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu, China | [2]Smart Computational Imaging Research Institute (SCIRI), Nanjing University of Science and Technology, Nanjing, Jiangsu, China | [3]Jiangsu Key Laboratory of Visual Sensing & Intelligent Perception, Nanjing, Jiangsu, China

Correspondence: Bowen Wang (wangbowen@njust.edu.cn) | Chao Zuo (zuochao@njust.edu.cn)

Received: 14 October 2025 | **Revised:** 13 January 2026 | **Accepted:** 18 January 2026

Keywords: inverse synthetic aperture imaging | macroscopic fourier ptychography | super-resolution

ABSTRACT

High-resolution imaging of distant, dynamic targets is fundamentally limited by finite-aperture diffraction. Unlike radar, optical detectors record only intensity and lack the temporal resolution for direct phase measurement, hindering straightforward extension of synthetic aperture techniques to the optical domain. While Macroscopic Fourier ptychography (FP) offers a promising computational alternative, its reliance on mechanical scanning or engineered illumination restricts its use to cooperative, static scenes. In this paper, we present Inverse Synthetic Aperture Macroscopic Fourier ptychography (ISAMFP), which, to the best of our knowledge, is the first practical realization of inverse synthetic aperture imaging for diffusely reflecting targets within the Fourier ptychography framework. By transforming the natural target motion from a detrimental factor into an effective source of angular diversity, ISAMFP eliminates the need for mechanical scanning or illumination modulation and achieves resolution enhancement far beyond the aperture limit using intensity-only measurements. Experimental validations on USAF resolution targets, unmanned aerial vehicles, and complex multi-target scenes highlight both the robustness and versatility of ISAMFP, establishing it as a powerful tool for non-interferometric synthetic aperture imaging of moving, non-cooperative targets, with broad implications for surveillance, defense, and remote sensing.

1 | Introduction

The ability to achieve high-resolution imaging beyond the diffraction limit of a finite aperture has been of vital importance for gaining new insight and understanding in astronomy, remote sensing, and space science [1, 2]. Current high-resolution observations typically rely on a large single-aperture systems, such as the Five-hundred-meter Aperture Spherical Radio Telescope (FAST) [3], the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) [4], and the Hubble Space Telescope (HST) [5], to acquire data at higher resolution and greater distances. However, driven by the increasing demand for higher resolution, the aperture size continues to grow, and the manufacturing complexity and cost increase rapidly, becoming a major bottleneck to further improvements in detection capability.

With advances in radar technology and signal processing, we have witnessed the continuous progress of synthetic aperture imaging technologies over the past decades. As one of the most representative super-resolution approaches, synthetic aperture radar (SAR) has been widely applied in the microwave and millimeter-wave bands [6], where target or platform motion is exploited to form a virtual large aperture and thereby significantly enhancing image resolution. However, for super-resolution imaging in the nearinfrared and even visible bands, the concept of synthetic aperture continues to face a fundamental challenge: the limited temporal response of optical detectors prevents direct phase measurement [7, 8], thereby constraining the development of optical synthetic aperture techniques.

Macroscopic Fourier ptychographic imaging is a super-resolution method based on frequency-domain stitching [9–11], where sub-aperture intensity images of the target are recorded from different angles and combined through a ptychographic algorithm to simultaneously achieve phase retrieval and synthetic aperture formation [12, 13]. Notwithstanding the need for sub-aperture image processing, it is well suited for long-term super-resolution imaging of distant targets in astronomy, remote sensing, and military observation due to its unique advantages of non-interferometric technique, active illumination, controllable spectrum extension, and robustness to complex environments [14, 15].

Despite all the advantages, Macroscopic FP still suffers from several technical challenges that hinder its practical application and widespread adoption for real-time long-term high-resolution imaging of distant targets. Generally, high-quality and robust FP reconstructions rely heavily on sufficient spectrum overlap to ensure stable phase retrieval [16–18]. Even minor deviations in acquisition conditions may have an effect on the reconstruction results [19], leading to artifacts or resolution degradation. On the other hand, Macroscopic FP often requires sequential scanning to acquire a large number of low-resolution images, which incurs time overhead and severely limits its applicability to dynamic scenes [20]. Various strategies have been proposed to improve the temporal resolution of Macroscopic FP, such as total variation regularization [21], camera arrays [22, 23], and learning-based approaches [24]. These methods involve complex system configurations and lack validation in long-range applications. Therefore, overcoming the limitations of acquisition delay, temporal efficiency, and stability remains an open quest for Macroscopic FP imaging of dynamic and non-cooperative targets.

To this end, we present ISAMFP (Inverse Synthetic Aperture Macroscopic Fourier ptychography), an efficient and robust inverse synthetic aperture (ISA) imaging system embedded in the Fourier ptychographic framework. Inverse synthetic aperture radar (ISAR) exploits target displacement and attitude variations to synthesize an equivalent large aperture [25], with the received echoes processed through range pulse compression, phase history formation, motion compensation, and coherent imaging algorithms to achieve high-resolution reconstruction. While coherent phase retrieval via heterodyne detection enables ISAR, heterodyne-based approaches impose stringent requirements on source stability, reference beam control, and long-term coherence maintenance, thereby limiting their practicality in long-range scenarios. Here, we introduce the concept of "motion-driven synthetic aperture" from ISAR into Macroscopic FP, where natural target motion replaces external scanning or illumination modulation to provide multi-angle measurements, and iterative ptychographic updates recover the complex field from sub-aperture intensity recordings. By jointly applying a frequency-domain and SIFT-flow registration strategy, frame-wise translations, minor rotations, and slight magnification variations are estimated to align each sub-aperture to the correct position in the frequency domain. Leveraging motion trajectories, ISAMFP achieves up to a four-fold resolution enhancement beyond the diffraction limit of the single aperture. Quantitative metrics including structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), contrast curves, and line profiles validate the feasibility and robustness of ISAMFP in scenarios ranging from programmed motion, video-rate UAV flights, to multi-target experiments. As a result, we demonstrate the potential of integrating inverse synthetic aperture concepts with Macroscopic FP to realize high-speed, high-resolution, and artifact-free imaging, pushing the non-interferometric super-resolution imaging application to non-cooperative dynamic targets in the far field.

2 | Working Principle of ISAMFP

2.1 | Comparison of ISAMFP and Conventional Macroscopic FP

To provide a clear illustration of the working principle of ISAMFP, we compare ISAMFP with conventional scanning-based macroscopic Fourier ptychographic imaging techniques. As shown in Figure 1a, beam scanning changes the illumination angle at the source side to acquire sub-aperture images, whereas aperture scanning varies the receiving angle to sample different spectral components at the detector side. In contrast, ISAMFP keeps the optical system fixed and adopts a transceiver configuration similar to ISA, where angular diversity is introduced by the inherent motion of the target.

The ISAMFP system, as shown in Figure 1b, employs a 650-nm fiber laser and a Celestron 120EQ3M-R5 telescope to deliver coherent illumination. The beam is expanded and collimated to provide low-divergence illumination for long-range operation. On the detection side, an Imaging Source DMK-33U×183 sensor (5472 × 3648, 2.4 μm) is mounted on the telescope for stable alignment and backscattered signal collection. As shown in Figure 1c, the target motion naturally provides successive frames for acquisition, while Figure 1d illustrates the subsequent high-resolution reconstruction, where each frame is located in the frequency plane, registered, and synthesized.

2.2 | ISAMFP Forward Imaging Model

The forward model of ISAMFP is based on a simplified optical configuration: a laser beam is collimated and expanded by a telescope to form an approximately planar illumination field over the target region [11]. The complex amplitude of the illumination can be expressed as: $U_0(x,y) = A_0 O(x,y) e^{jk_0 z}$, where $k_0 = 2\pi/\lambda$ denotes the wavenumber, A_0 is the amplitude, z represents the corresponding propagation direction, and $o(x,y)$ denotes the spatial reflectivity function of the target. Due to the presence of surface roughness, the reflected wavefront undergoes strong phase modulation under diffuse reflection conditions. To model the propagation modulation, a random scattering layer is introduced to describe the surface scattering characteristics. Following illumination at the initial target position and back-propagation to the imaging plane, the resulting complex amplitude can be

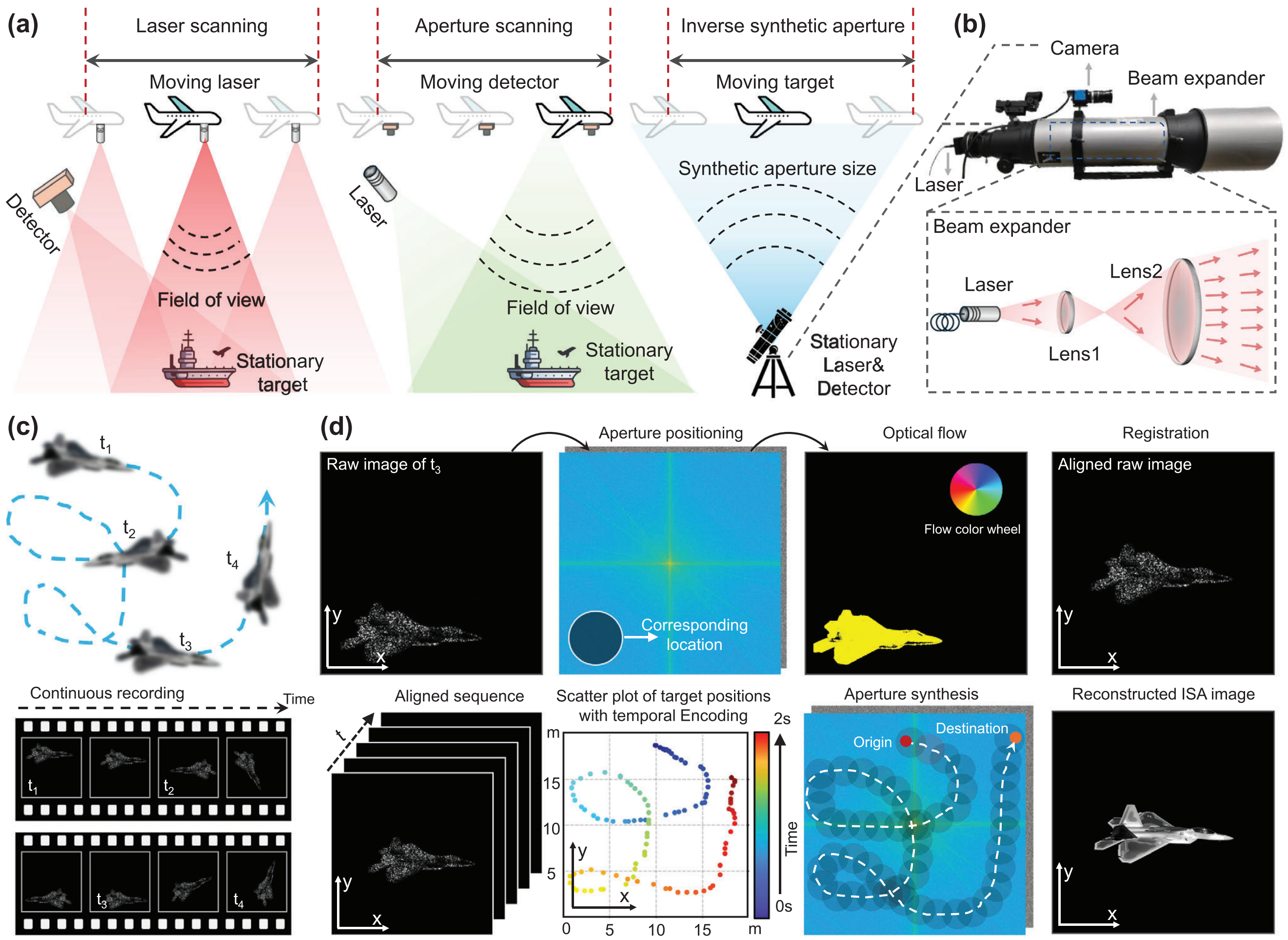

FIGURE 1 | ISAMFP concept and system overview. (a) Schematics of beam-scanning, aperture-scanning, and ISA acquisition. (b) Optical layout of the system. (c) Target motion trajectories with recorded frames. (d) Registration results and reconstructed high-resolution images.

written as:

$$U_0(x, y) = [A_0\, O(x, y)\, \zeta(x, y)] * h_{z_0}(x, y) \tag{1}$$

For a diffuse scattering surface, the multiplicative phase term can be decomposed into a deterministic and a random component (see Note S5). Specifically, the complex modulation can be expressed as:

$$\zeta(x, y) = a(x, y)\, e^{i\varphi_{\mathrm{det}}(x,y)}\, e^{i\varphi_{\mathrm{ran}}(x,y)} \tag{2}$$

where $a(x, y)$ denotes the surface reflectance amplitude, $\phi_{\mathrm{det}}(x, y)$ represents the deterministic phase related to the macroscopic object structure (e.g., geometry and curvature), and $\phi_{\mathrm{ran}}(x, y)$ corresponds to the random phase induced by microscopic surface roughness, which gives rise to diffuse scattering. This decomposition allows the deterministic phase to carry structural information, while the random phase spreads the spectral content across both x and y directions, enabling resolution gain in two dimensions even under unidirectional motion. $h_{z_0}(x, y)$ is the coherent point-spread function (PSF) from the object plane to the (x, y) plane. Previous studies have suggested that without a scattering layer, the frequency spectrum of the target remains approximately invariant across positions. However, for rough surfaces commonly encountered in practice, the illumination angle continuously changes as the target moves. Due to random surface variations, light reflected from different directions carries distinct phase modulations [26], causing the spectral content to vary with motion. The moving modulation breaks the strict spectral invariance, and introduces additional frequency components, which facilitate further frequency-domain expansion.

An illumination angle corresponding to the initial position J_1 is considered, where coherent light propagating through the simulated scattering layer acquires a random phase $\varepsilon_{initial}$. Following propagation through the imaging lens, the light reaches the sensor plane, and the intensity is recorded as:

$$I(x, y) = \left| \mathcal{F}^{-1}\left\{ P_i(k_x, k_y)\, \mathcal{F}[\, O(x, y)\, \varepsilon_{\mathrm{initial}}(x, y)\,] \right\} \right|^2 \tag{3}$$

where $P_i(k_x, k_y) = P_0\left(\frac{k_x - \Delta k_{x,i}}{\gamma_i}, \frac{k_y - \Delta k_{y,i}}{\gamma_i}\right)$ denotes the pupil function for the i-th frame, where $(\Delta k_{x,i}, \Delta k_{y,i})$ is the center frequency shift of the corresponding sub-aperture (set to 0 if absent), and γ_i is the spectral scaling factor. In the forward model, γ_i accounts for the equivalent change in aperture size induced by the spectral scaling under longitudinal motion. However, during reconstruction, the pupil update is still performed using the physical (actual) aperture size. This modulation uniformly redistributes the high- and low-frequency components of the target, effectively encoding high-frequency information across the entire spectral plane. The recorded intensity $I(x', y')$ is considered the first frame. If the target moves with an instantaneous velocity along a given direction, the position of the target in the subsequent frame J_2 can be expressed as:

$$J_2 = J_1 + v\, t_{frame}, \qquad v = (v_x, v_y)^{\top} \tag{4}$$

Equivalently, $x_2 = x_1 + v_x\, t_{frame}, y_2 = y_1 + v_y\, t_{frame}$. Let $\Delta x = v_x\, t_{frame}$ and $\Delta y = v_y\, t_{frame}$. Under fixed-beam illumination, the object-plane field at the new position is modeled as

$$U_0^{(2)}(x, y) = A_0\, O(x - \Delta x,\ y - \Delta y)\, \varepsilon_2(x, y) \tag{5}$$

where $\varepsilon_2(x, y) = e^{j\phi_2(x,y)}$ denotes the angle-dependent multiplicative phase corresponding to the new position J_2.

Notably, motion-induced changes in the viewing angle shift the object spectrum relative to the pupil, so each frame contributes a different portion of the synthetic aperture. Random phase modulation provides additional diversity that stabilizes the reconstruction. With irregular target motion, the recorded intensity stack can be written as:

$$\mathcal{I} = [\, I_i(x', y')\,]_{i=1}^{N} \tag{6}$$

$$I_i(x', y') = \left| \mathcal{F}^{-1}\left\{ P_i(k_x, k_y)\, \widetilde{O\varepsilon_i}\big(k_x - \Delta k_{x,i},\ k_y - \Delta k_{y,i}\big) \right\} \right|^2 \tag{7}$$

where $\widetilde{O\varepsilon_i}(k_x, k_y) = \mathcal{F}\{O(x, y)\, \varepsilon_i(x, y)\}$. By using intensity frames that sample different spectral regions, the reconstruction proceeds with a Fourier ptychographic update scheme.

2.3 | ISAMFP Reconstruction Algorithm

Accurate and efficient target localization and registration are crucial for effective reconstruction in the proposed method. Previous studies have reported that for transmissive targets or smooth reflective surfaces [27], translational motion only shifts the target position in sub-aperture images without introducing additional spectral components, thereby failing to achieve effective frequency extension. While target rotation can introduce spectral rotation and a restricted degree of spectral extension [28], the improvement remains limited. In contrast, diffuse targets with rough surfaces, which constitute the majority of real-world objects, act as dynamic scattering layers [29–32], where motion effectively alters the illumination direction relative to the detector and provides additional spectral components for aperture synthesis (see Notes S1 and S2). In the application scenario, the ISAMFP system continuously captured image sequences of a moving UAV, as shown in Figure 2a,b. The temporal evolution of target positions is also reflected in the frequency domain, necessitating registration (Figure 2c). Different types of motion (rotation, translation, and magnification, the magnification and 3D motion see Note S6) affect the spectrum differently, as illustrated in Figure 2d, thereby requiring distinct registration strategies (see Note S3). Following precise registration [33, 34], the dataset (Figure 2e) was subjected to ptychographic reconstruction, where the spectral extension corresponds to the UAV's flight trajectory. From a spatial-domain perspective, the target's motion within the illumination region can be regarded as a convolution process between the incident wavefront and the target function. As shown in Figure 2g, the reconstruction achieves improvements in both resolution and signal-to-noise ratio (SNR) compared to the single-aperture image. We further compared the reconstruction with multi-frame averaging in terms of SSIM. While the SSIM of the averaging result increases slowly with frame number and eventually levels off, ISAMFP consistently achieves higher values and reveals finer structural details, demonstrating the effectiveness of the proposed method.

In the frequency domain, different target positions provide access to spectral information under varying illumination angles. The corresponding variation primarily arises from the angular modulation induced by the laser interacting with the rough surface. Therefore, under Macroscopic imaging conditions, we adopt the field-recovery and spectral-sampling mechanisms of ptychography to perform synthetic aperture, enabling high-resolution reconstruction of the target. We initialize the object amplitude $O_{initial}$ and the system aperture function P. Since the raw sub-aperture frames contain speckle, a direct start of reconstruction may converge to local minima. We therefore first average the frames to suppress noise and then take the square root as the initial amplitude. The result is upsampled to the high-resolution grid to provide a stable starting point.

$$O_{\text{init}}(x', y') = \uparrow M \left\{ \sqrt{\frac{1}{N} \sum_{i=1}^{N} I_i(x, y)} \right\} \tag{8}$$

where $I_i(x, y)$ denotes the recorded image of the i^{th} frame (with a total of N frames), and $\uparrow M$ represents the up sampling M factor. For each recorded frame, the target complex amplitude function is multiplied by the random phase modulation corresponding to the illumination angle of that frame, yielding a modulated complex wavefront. The resulting wavefront is then numerically propagated according to the forward imaging model, producing the observed low-resolution intensity image I_i^g under the current numerical aperture constraints:

$$I_i^g(x, y) = \left| \mathcal{F}^{-1}\{P_i(k_x, k_y)\, \mathcal{F}[\, O(x, y)\, \varepsilon_i(x, y)]\} \right|^2 \tag{9}$$

The recorded sub-aperture images are used to update the intensity of the corresponding generated sub-aperture complex amplitudes:

$$\phi_i^{update}(x, y) = \phi_i(x, y) \frac{\sqrt{I_i(x, y)}}{\sqrt{I_i^g\left(\left\lceil \frac{x}{M} \right\rceil, \left\lceil \frac{y}{M} \right\rceil\right)}} \tag{10}$$

where $I_i(x, y)$ is the i^{th} frame intensity image, $I_i^g(x, y)$ represents the high-resolution image initially estimated from the sub-aperture image acquired at the current position, M denotes the

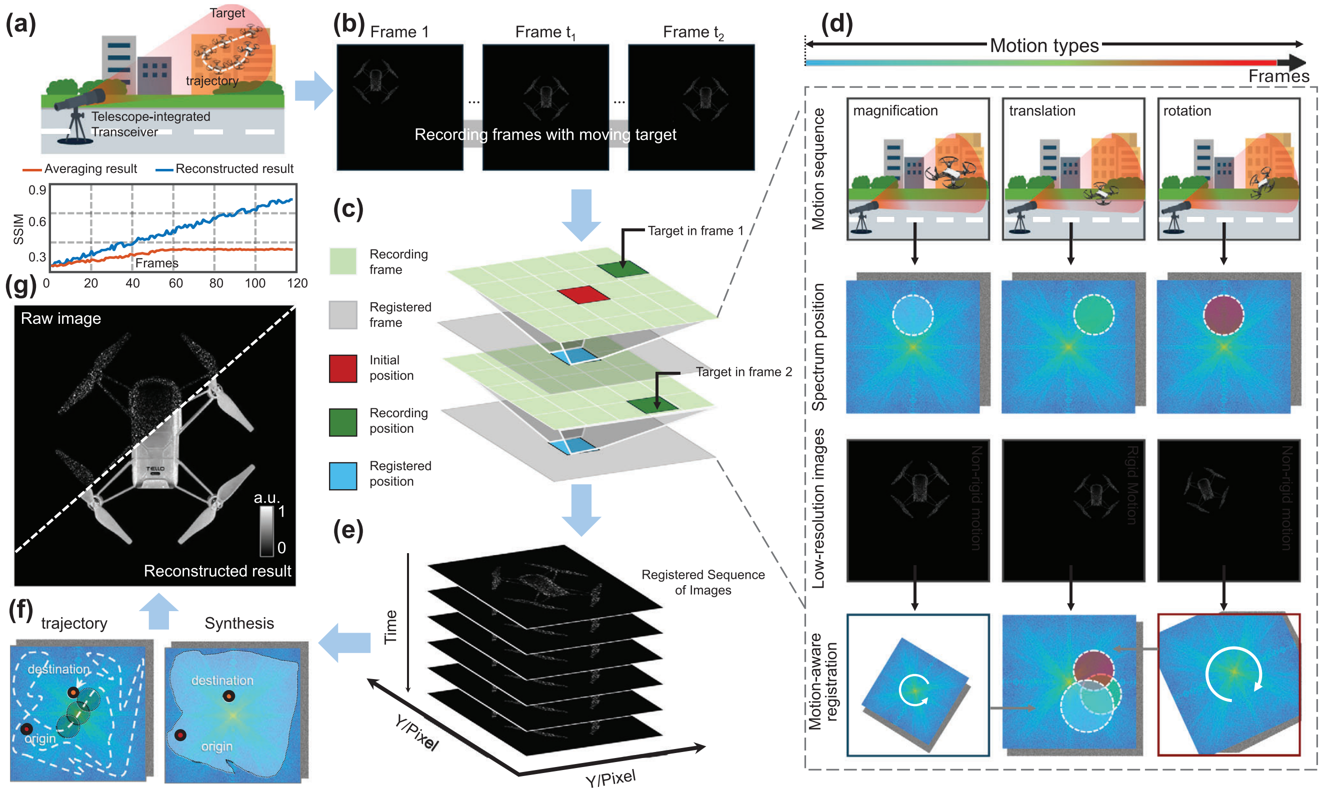

FIGURE 2 | ISAMFP imaging of a UAV target. (a) Continuous acquisition. (b) Captured image sequence. (c) Target registration process. (d) Comparison of registration. (e) Registered raw image sequence. (f) Frequency-plane positions. (g) Reconstructed result compared with the raw image.

up sampling factor [35]. The complex amplitude recovery of the target can be expressed as:

$$O^{\text{new}}(k_x, k_y) = O^{\text{old}}(k_x, k_y) + \frac{\overline{P(k_x, k_y)}\left[\Psi_i^{\text{new}}(k_x, k_y) - \Psi_i^{\text{old}}(k_x, k_y)\right]}{(1-\alpha_o)\,|P(k_x, k_y)|^2 + \alpha_o \max\limits_{k_x, k_y} |P(k_x, k_y)|^2} \tag{11}$$

where $\Psi_i^{\text{old}}(k_x, k_y)$ denotes the predicted complex field in the frequency plane for the i_{th} view before ($\Psi_i^{\text{new}}(k_x, k_y)$ is after) the amplitude replacement step of the current iteration. $O^{\text{old}}(k_x, k_y)$ is the object estimates before ($O^{\text{new}}(k_x, k_y)$ is after) the update, $\overline{(\cdot)}$ denotes complex conjugation, $\alpha_o \in (0, 1)$ is the weighting parameter that balances local vs. global information in the denominator.Equations (9)–(11) are applied sequentially to all registered frames within each iteration, where the corresponding sub-aperture parameters $\{(\Delta k_{x,i}, \Delta k_{y,i}), \gamma_i\}$ are provided by the motion-aware registration. After cycling through the full dataset, the object estimate O is refined and used to initialize the next iteration. The iterations terminate when the relative change of O or the data-fidelity error falls below a preset threshold, or when a maximum number of iterations is reached; the final high-resolution reconstruction is obtained as $\hat{O}(x, y) = \mathcal{F}^{-1}\{\hat{O}(k_x, k_y)\}$, whose resolution is determined by the synthesized frequency coverage accumulated along the target trajectory.

3 | Result

3.1 | Structured Motion Validation

To quantitatively evaluate the super-resolution capability of ISAMFP results, a USAF resolution target was mounted on a translation stage to simulate target motion, as illustrated in Figure 3a. The ISAMFP system was positioned 5.82 m away to record the sub-aperture images, coupled with a 75 mm focal length lens (F/16) for image acquisition. The USAF resolution target was mounted on an *XY* dual-axis translation stage, enabling a 13 × 13 array scan with 1.6 mm aperture spacing. A circular translation stage was further employed to generate a non-Cartesian trajectory, testing reconstruction capability under more practical scanning conditions (Figure 3b). Due to the 0.005 mm repeatability precision of the used stage (HGTA04100), slight positional errors may occur as the stage may not precisely reach the target positions during scanning. As a solution, the joint registration method was used to obtain the actual target positions (orange triangles). Compared to the theoretical stage positions (blue circles), the actual positions obtained through registration exhibited higher accuracy, thereby eliminating the precision errors introduced by the stage.

Due to the diffraction limit and coherent speckle effects, the resolvable line pairs were reduced to Group -1, Element 6. Figure 3d compares the PSNR and SSIM values for the single aperture image, multi-frame image average (averaged over the

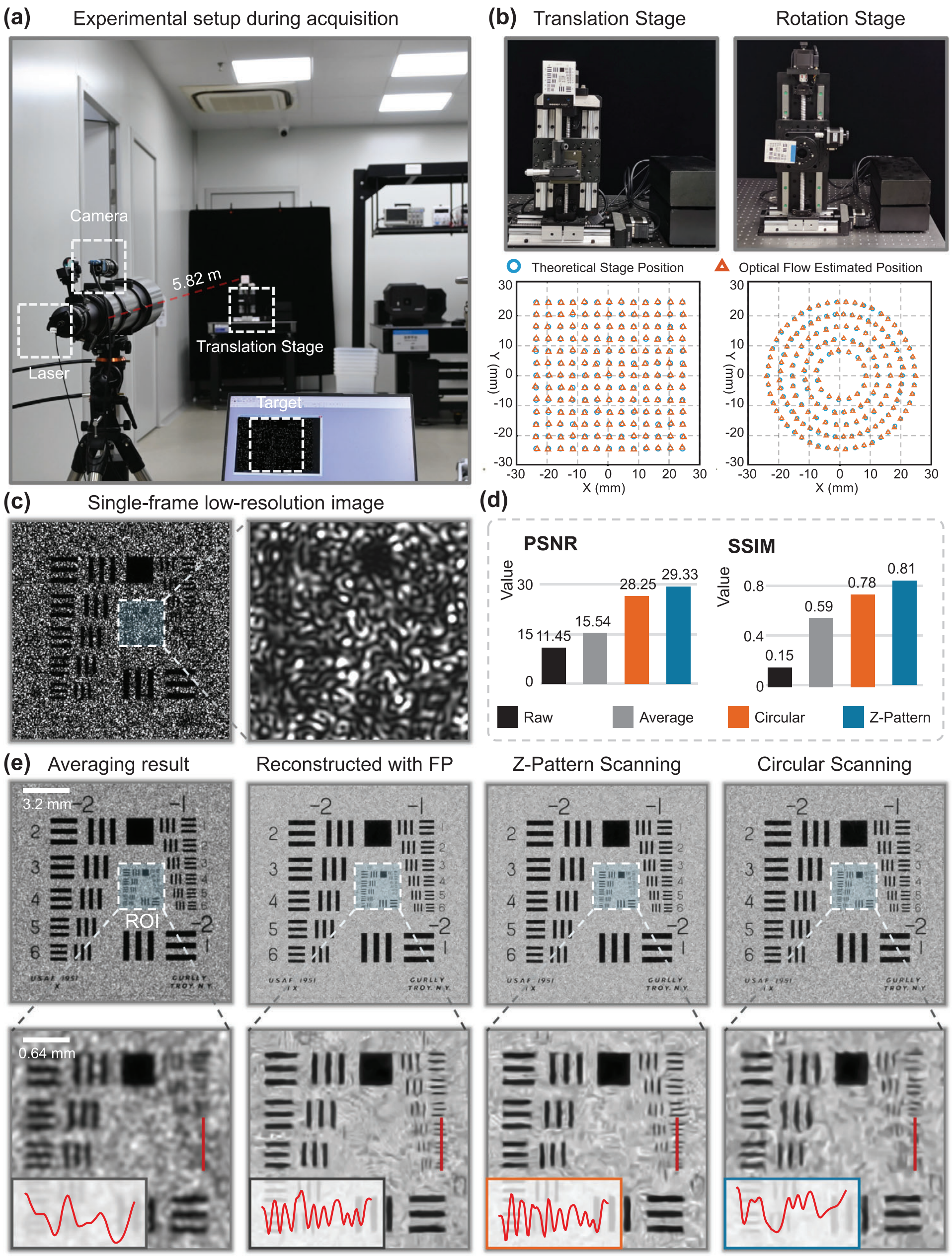

FIGURE 3 | Validation of ISAMFP under structured trajectories. (a) Experimental setup of ISAMFP. (b) Target image sequence acquired with *XY* and circular scanning. (c) Single-aperture image and zoom. (d) PSNR and SSIM comparison. (e) Visual comparison of reconstructions.

number of images scanned by the XY dual-axis stage), XY dual-axis stage scan reconstruction, and circular scan reconstruction (with the ground truth selected as the reconstruction result of aperture-scanning FP). Figure 3e provides a comparative imaging result for different schemes.

The reconstructed results demonstrate that the multi-frame averaging improves the resolvable line pairs to Group 0, Element 6 compared to the single-aperture result. The traditional FP using Z-axis scanning for a stationary resolution target with a 13 × 13 array and BM3D denoising achieved a high-quality reconstruction, improving the resolvable line pairs to Group 1, Element 6. The Z-axis scan reconstruction result closely approximated that of traditional FP, achieving a Group 1, Element 6 resolution enhancement and a relative fourfold resolution improvement over the single-aperture result, thus demonstrating the resolution enhancement capability of ISAMFP. The circular scan's spectrum coverage is smaller than that of the Z-axis scan, leading to a reduction in the high-frequency information that can be reconstructed. As a result, the resolvable line pairs increased to Group 1, Element 3, achieving a threefold resolution improvement compared to the single-aperture case, thereby validating the effectiveness and advantages of the ISAMFP system in dynamic target imaging.

3.2 | Dynamic UAV Experiment

To further examine the robustness of ISAMFP, we evaluated the imaging performance in scenarios involving irregularly moving targets. A miniature UAV (DJI Tello) was selected as the irreg-

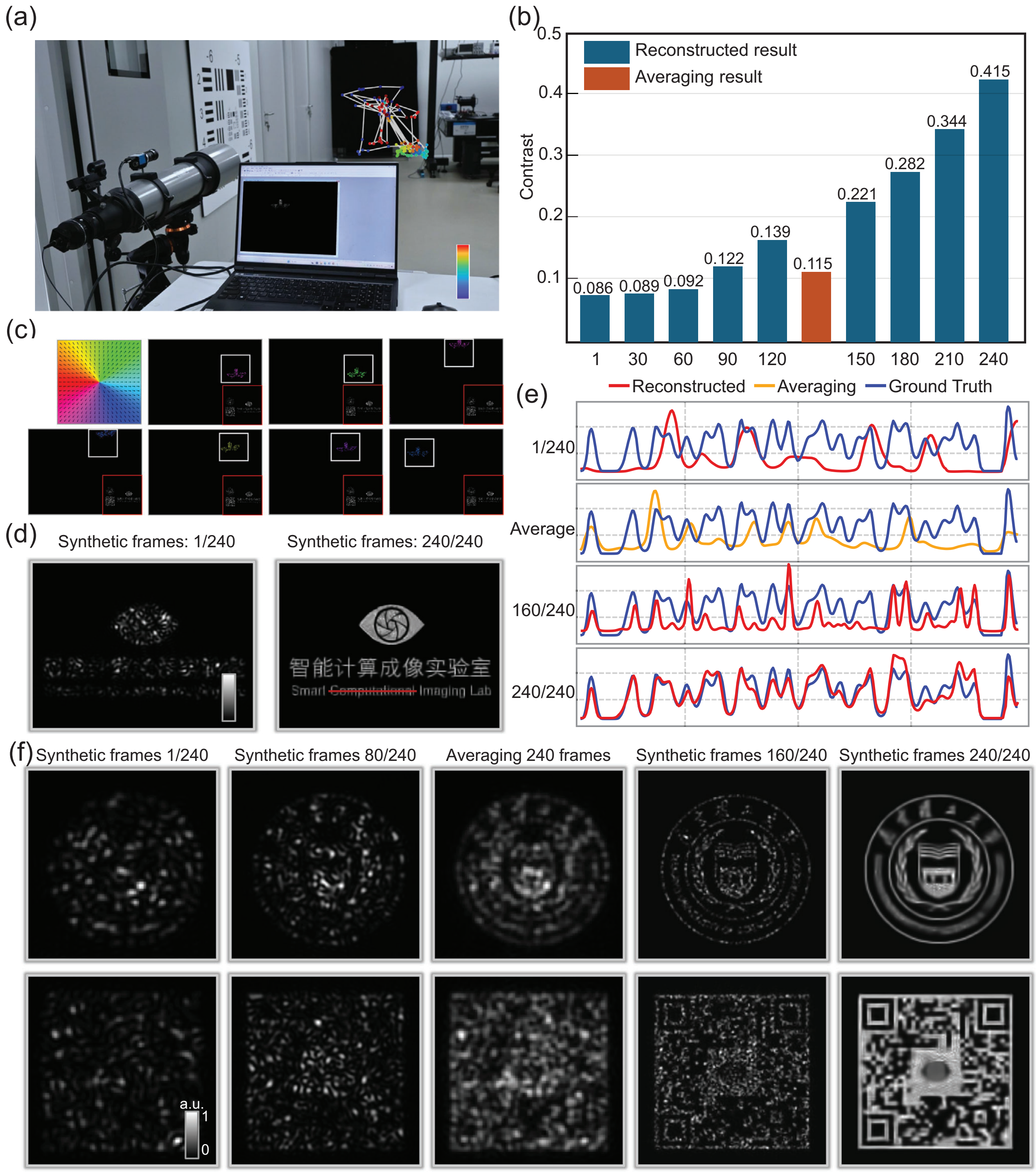

FIGURE 4 | ISAMFP imaging of a flying UAV. (a) Scene capture during ISAMFP imaging. (b) The registration of UAV frames. (c) Contrast evolution with the increasing number of recorded frames. (d) Sub-aperture images and denoised reconstructions of UAV markings. (e) Line profiles corresponding to (d). (f) Reconstruction comparison for QR code region.

ularly moving target, and continuous recordings were captured at 48 fps (1920 × 1080 pixels) within the illuminated region, as illustrated in Figure 4a. Following frame registration, the sub-aperture images were aligned to the corresponding synthetic aperture positions for reconstruction (Figure 4b). To quantitatively evaluate the reconstruction performance, we compute the image contrast C from the bright and dark bars of the image. The average intensity in the bright and dark regions is denoted by $\bar{w}$ and $\bar{b}$, respectively. The contrast is then defined as

$$C = \frac{\bar{w} - \bar{b}}{\bar{w} + \bar{b}} \cdot \bar{w} \cdot (1 - \bar{b}) \tag{12}$$

The quantitative results in Figure 4c demonstrate that the image contrast of a single frame (0.086) gradually improves with increasing frame number, ultimately reaching 0.415. For comparison, the simple average of 160 frames yields a contrast of only 0.115 (orange bar), which is lower than the corresponding ISAMFP reconstruction. Figure 4d presents representative sub-aperture images of the marked pattern together with denoising reconstruction. The corresponding line profiles are compared with aperture-scanning FP (as shown in Figure 4e), which serves as the benchmark. Figure 4f further presents the reconstruction of another marked region, validating the consistency of ISAMFP in resolving fine structural details across different areas. As the number of recorded sub-aperture frames increases, the target resolution progressively improves, ultimately enabling recognition of the QR code pattern. The dynamic imaging process and reconstruction details are provided in Visualization 1. These results confirm that ISAMFP can deliver high-resolution reconstructions of a single dynamic target, laying the foundation for evaluating its performance in more complex multi-target scenarios.

3.3 | Multi-Target Scenario

Finally, ISAMFP was applied to reconstruct a complex dynamic scene involving two independently moving targets. Representative low-resolution frames acquired during the experiment are shown in Figure 5a, while Figure 5b illustrates the motion trajectories and relative distance variations of the two targets, confirming independent motion along non-overlapping paths. Figure 5c quantifies the contrast of the high-resolution reconstructions as a function of frame count, showing a steady increase with additional frames. The corresponding reconstruction results of the two targets are presented in Figure 5d, where progressively finer structural details become resolvable as the number of frames increases. The multi-target imaging results demonstrate that ISAMFP can effectively integrate multi-frame information in complex scenarios, thereby enabling reconstruction of structural details inaccessible to single-frame imaging. The random phase introduced by diffuse scattering couples the high-frequency information from uniaxial motion into the 2D spectrum, thereby enabling lateral motion of the target to simultaneously enhance the resolution along both horizontal and vertical directions (see Notes S4 and S5). Notably, the reconstruction performance varies with trajectory coverage and spectral redundancy. For the same number of frames, the taxi target, with its narrower spectral coverage, exhibits limited resolution improvement, and its contrast saturates at 0.292 once the motion ceases. In contrast, the bus target, benefiting from broader trajectory coverage, achieves higher detail resolution, with the contrast continuing to increase and reaching 0.394 at 180 frames (provided in Visualization 2). This comparison highlights the critical role of motion-induced spectral coverage in determining the reconstruction quality of ISAMFP, underscoring its adaptability and applicability in multi-target, complex dynamic scenarios. In cases of target trajectory overlap or occlusion, the proposed method can handle such scenarios by temporarily excluding overlapping frames, effectively preventing incorrect merging of spectral information. This approach aligns with occlusion estimation strategies in multi-object tracking (MOT), where occluded targets are either excluded or their positions inferred during the occlusion [36–38], ensuring accurate tracking and reconstruction without identity confusion.

4 | Discussion and Conclusion

To conclude, we have reported ISAMFP, a motion-driven inverse synthetic aperture Fourier ptychographic imaging technique that achieves high-resolution imaging of dynamic and non-cooperative targets directly from naturally acquired frame sequences of target motion. ISAMFP employs the registration strategy to precisely estimate motion parameters. The subsequent Fourier ptychographic reconstruction "regenerates" the lost high-frequency components, resulting in up to four-fold resolution enhancement beyond the single-aperture diffraction limit without mechanical scanning. Moreover, owing to its non-interferometric and intensity-based nature, the ISAMFP method avoids the stringent requirements of heterodyne detection and long-term phase stability that are typically necessary in inverse synthetic aperture laser radar systems. Finally, by simply employing higher-frame-rate imaging sensors in conjunction with camera arrays, the spectral coverage speed could, in principle, be further increased to accommodate complex dynamic scenarios. Experimental results suggest that ISAMFP is expected to offer new insights for long-range, non-cooperative, high-resolution imaging of dynamic targets, advancing the frontier of target tracking and high-resolution remote sensing.

Nonetheless, several limitations warrant further investigation. The maximum achievable resolution of our prototype is jointly determined by the theoretical upper bound of the recoverable bandwidth and the bandwidth that can be practically realized in a given scenario. In principle, for a single acquisition (one sub-aperture), the usable image-plane spatial frequencies are limited by the coherent diffraction cutoff and sensor sampling, which together define the single-frame baseline resolution. The resolution gain after multi-frame synthesis arises from the enlarged frequency-domain coverage, whose upper bound is governed by the effective synthetic numerical aperture accumulated during the observation (set by system geometry and target motion, equivalently the angular-span coverage), as well as the spectral overlap between adjacent measurements; insufficient angular span or inadequate overlap weakens the constraints at high frequencies and leads to saturation of the resolution improvement. In practice, the realizable bandwidth can be further reduced by several factors. First, errors in sub-aperture localization and spectral-scaling estimation make frequency-domain stitching increasingly sensitive to misregistration at high frequencies. Second, illumination conditions and target scattering properties jointly determine speckle contrast and signal-to-noise ratio, and strong inter-frame decorrelation or low SNR can bury high-frequency information, particularly for long-range imaging and rough surfaces. Third, ISAMFP inherently depends on natural motion and diffuse scattering to redistribute spectral content; extreme motion irregularity, high-speed dynamics, or sparse sampling may reduce spectral redundancy and induce phase inconsistency between frames, degrading reconstruction robustness. Introducing adaptive confidence-based fusion or local spectral weighting could mitigate these effects. The current implementation mainly targets rigid or quasi-rigid motion; for large-field non-rigid or deformable targets [39, 40], local motion inconsistencies and registration errors may accumulate and cause spatially varying artifacts. Moreover, under multi-target occlusion or overlapping trajectories, the registration accuracy—already constrained by motion-estimation precision [41]—may further

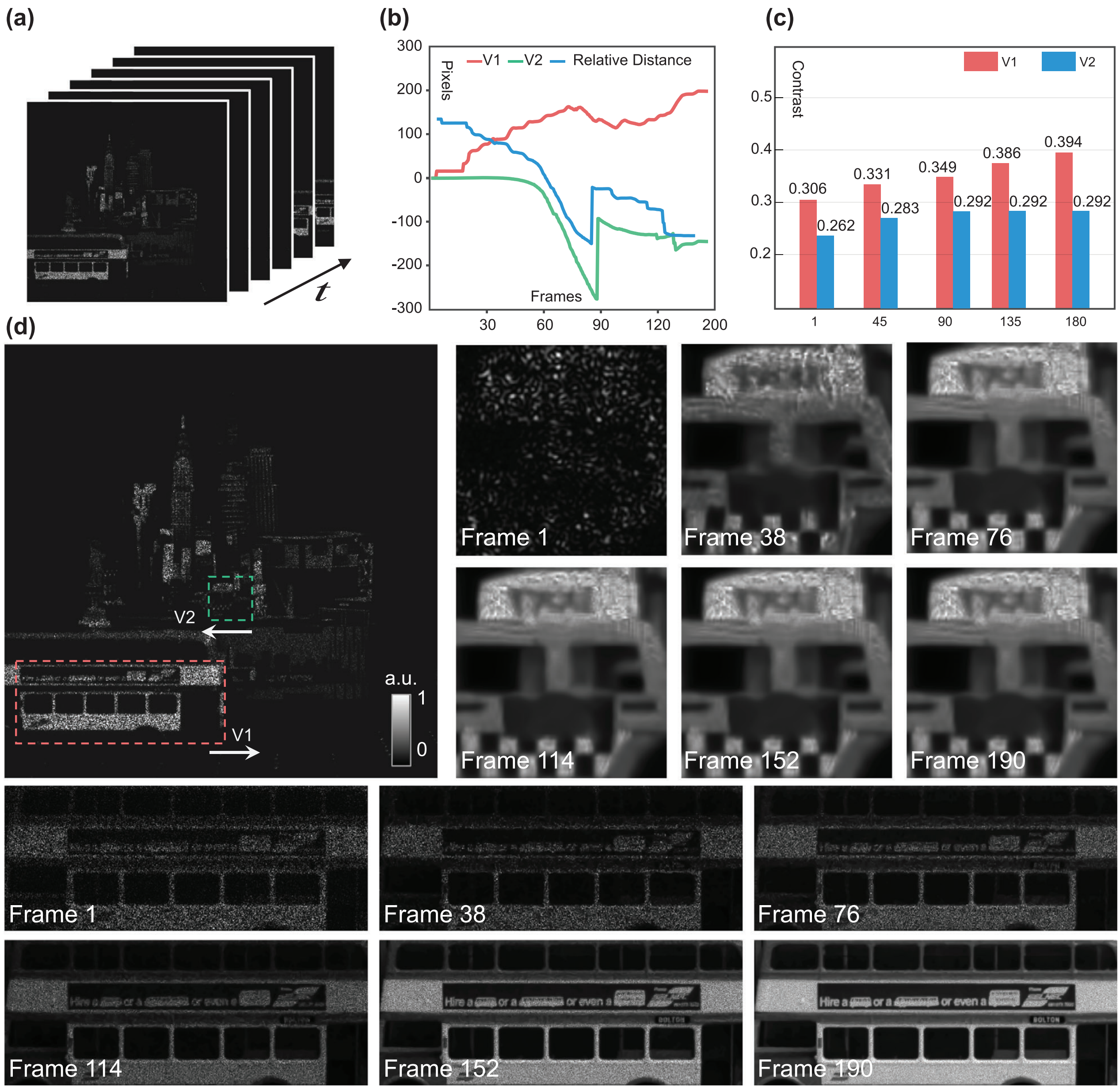

FIGURE 5 | ISAMFP imaging of two independently moving targets. (a) Typical low-resolution frames captured during the experiment. (b) Motion trajectories and relative distance variations of two independent targets. (c) The contrast improvement of the reconstructed result with increasing frame numbers. (d) Reconstructions of resolvable details on both targets as a function of frame count.

deteriorate. Finally, for highly complex [42], multi-plane, or low-contrast targets, deviations from the assumed scattering model, and limited illumination diversity can lead to incomplete spectral recovery, motivating additional calibration, localized windowing, or multi-polarization acquisition to better preserve reconstruction fidelity.

Future directions will focus on enhancing the robustness and scalability of ISAMFP through adaptive motion estimation and learning-based registration [43], which can dynamically compensate for irregular or non-rigid motion. Incorporating sliding-window or block-wise reconstruction strategies [44] could further alleviate drift accumulation and improve stability in large-scale or deformable scenes. Hardware and software co-design, such as integrating high-frame-rate sensors and synchronized camera arrays [45], will enable faster spectral coverage and near-real-time implementation. Beyond the current framework, integration of coded illumination or spatial light modulation with motion-driven sampling could further increase information throughput and extend ISAMFP toward real-time, video-rate super-resolution imaging. More broadly, the concept of converting uncontrolled target motion into an imaging advantage offers a new avenue for developing passive synthetic aperture systems in remote sensing, surveillance, astronomical observation, and autonomous navigation, extending the scope of computational imaging toward more adaptive and motion-aware paradigms.

Acknowledgements

This work was supported by National Natural Science Foundation of China (U21B2033, 62075096, 62227818, 62361136588), National Key

Research and Development Program of China (2022YFB2804603, 2022YFA1205002, 2024YFE0101300), Fundamental Research Funds for the Central Universities (30925020211), State key Laboratory of Extreme Environment Optoelectronic Dynamic Measurement Technology and Instrument (2025-SYSJJ-04), Leading Technology of Jiangsu Basic Research Plan (BK20192003), "333 Engineering" Research Project of Jiangsu Province (BRA2016407), Open Research Fund of Jiangsu Key Laboratory of Visual Sensing & Intelligent Perception (SJCGKFJJ20251).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. E. H. K. Stelzer, "Beyond the Diffraction Limit?" *Nature* 417, no. 6891 (June 2002): 806–807, https://www.nature.com/articles/417806a.

2. C. Zuo, J. Qian, S. Feng, et al., "Deep Learning in Optical Metrology: A Review," *Light: Science & Applications* 11, no. 1 (Feb. 2022): 39, https://www.nature.com/articles/s41377-022-00714-x.

3. D. Li and Z. Pan, "The Five-Hundred-Meter Aperture Spherical Radio Telescope Project," *Radio Science* 51, no. 7 (July 2016): 1060–1064, https://ieeexplore.ieee.org/abstract/document/7773907.

4. Q. Gao, J.-R. Shi, H.-L. Yan, et al., "The Lithium Abundances from the Large Sky Area Multi-Object Fiber Spectroscopic Telescope Medium-Resolution Survey. I. The Method," *The Astrophysical Journal* 914, no. 2 (June 2021): 116, https://doi.org/10.3847/1538-4357/abf841.

5. R. Nan, "Five Hundred Meter Aperture Spherical Radio Telescope (FAST)," *Science in China Series G* 49, no. 2 (Apr. 2006): 129–148, https://doi.org/10.1007/s11433-006-0129-9.

6. P. W. Vachon, "Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation," *Geomatica* 59, no. 3 (Sept. 2005): 353–355, https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=11951036&v=2.1&it=r&id=GALE%7CA674713152&sid=googleScholar&linkaccess=abs.

7. L. Li, S. Wang, F. Zhao, et al., "Single-Shot Deterministic Complex Amplitude Imaging with a Single-Layer Metalens," *Science Advances* 10, no. 1 (Jan. 2024): eadl0501, https://www.science.org/doi/full/10.1126/sciadv.adl0501.

8. J. R. Fienup, "Phase Retrieval Algorithms: A Personal Tour [Invited]," *Applied Optics* 52, no. 1 (Jan. 2013): 45–56, https://opg.optica.org/ao/abstract.cfm?uri=ao-52-1-45.

9. J. Holloway, M. S. Asif, M. K. Sharma, et al., "Toward Long-Distance Subdiffraction Imaging Using Coherent Camera Arrays," *IEEE Transactions on Computational Imaging* 2, no. 3 (Sept. 2016): 251–265.

10. S. Dong, R. Horstmeyer, R. Shiradkar, et al., "Aperture-Scanning Fourier Ptychography for 3D Refocusing and Super-Resolution Macroscopic Imaging," *Optics Express* 22, no. 11 (2014): 13 586–13 599.

11. J. Holloway, Y. Wu, M. K. Sharma, O. Cossairt, and A. Veeraraghavan, "SAVI: Synthetic Apertures for Long-Range, Subdiffraction-Limited Visible Imaging Using Fourier Ptychography," *Science Advances* 3, no. 4 (2017): e1602564.

12. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-Field, High-Resolution Fourier Ptychographic Microscopy," *Nature Photonics* 7, no. 9 (2013): 739–745.

13. Y. Shu, J. Sun, J. Lyu, et al., "Adaptive Optical Quantitative Phase Imaging Based on Annular Illumination Fourier Ptychographic Microscopy," *PhotoniX* 3, no. 1 (2022): 1–15.

14. R. Zhou, Y. Guo, Q. Zhang, et al., "TurbFPNet: Neural Far-Field Turbulent Fourier Ptychography with a Camera Array," *Optica* 12, no. 7 (July 2025): 1068–1078, https://opg.optica.org/optica/abstract.cfm?uri=optica-12-7-1068.

15. G. Zheng, C. Shen, S. Jiang, P. Song, and C. Yang, "Concept, Implementations and Applications of Fourier Ptychography," *Nature Reviews Physics* 3, no. 3 (2021): 207–223.

16. S. Zhang, A. Wang, J. Xu, T. Feng, J. Zhou, and A. Pan, "FPM-WSI: Fourier Ptychographic Whole Slide Imaging via Feature-Domain Backdiffraction," *Optica* 11, no. 5 (May 2024): 634–646, https://opg.optica.org/optica/abstract.cfm?uri=optica-11-5-634.

17. S. Li, B. Wang, K. Liang, Q. Chen, and C. Zuo, "Far-Field Synthetic Aperture Imaging via Fourier Ptychography with Quasi-Plane Wave Illumination," *Advanced Photonics Research* 4, no. 10 (2023): 2300180, https://onlinelibrary.wiley.com/doi/abs/10.1002/adpr.202300180.

18. N. Chen, Y. Wu, C. Tan, L. Cao, J. Wang, and E. Y. Lam, "Uncertainty-Aware Fourier Ptychography," *Light: Science & Applications* 14, no. 1 (July 2025): 236, https://www.nature.com/articles/s41377-025-01915-w.

19. R. Wang, L. Yang, Y. Lee, et al., "Spatially-Coded Fourier Ptychography: Flexible and Detachable Coded Thin Films for Quantitative Phase Imaging with Uniform Phase Transfer Characteristics," *Advanced Optical Materials* 12, no. 15 (2024): 2303028, https://onlinelibrary.wiley.com/doi/abs/10.1002/adom.202303028.

20. B. Wang, W. Chen, J. Qian, S. Feng, Q. Chen, and C. Zuo, "Single-Shot Super-Resolved Fringe Projection Profilometry (SSSR-FPP): 100,000 Frames-Per-Second 3D Imaging with Deep Learning," *Light: Science & Applications* 14, no. 1 (Feb. 2025): 70, https://www.nature.com/articles/s41377-024-01721-w.

21. J. Wu, F. Yang, and L. Cao, "Resolution Enhancement of Long-Range Imaging with Sparse Apertures," *Optics and Lasers in Engineering* 155 (Aug. 2022): 107068, https://www.sciencedirect.com/science/article/pii/S0143816622001208.

22. C. Wang, M. Hu, Y. Takashima, T. J. Schulz, and D. J. Brady, "Snapshot Ptychography on Array Cameras," *Optics Express* 30, no. 2 (Jan. 2022): 2585–2598, https://opg.optica.org/oe/abstract.cfm?uri=oe-30-2-2585.

23. S. Li, B. Wang, H. Guan, G. Zheng, Q. Chen, and C. Zuo, "Snapshot Macroscopic Fourier Ptychography: Far-Field Synthetic Aperture Imaging via Illumination Multiplexing and Camera Array Acquisition," *Advanced Imaging* 1, no. 1 (2024): 011005, https://www.researching.cn/articles/OJc9c0865cda6fb188.

24. B. Wang, S. Li, Q. Chen, and C. Zuo, "Learning-Based Single-Shot Long-Range Synthetic Aperture Fourier Ptychographic Imaging with a Camera Array," *Optics Letters* 48, no. 2 (Jan. 2023): 263–266, https://opg.optica.org/ol/abstract.cfm?uri=ol-48-2-263.

25. S. M. Beck, J. R. Buck, W. F. Buell, et al., "Synthetic-Aperture Imaging Laser Radar: Laboratory Demonstration and Signal Processing," *Applied Optics* 44, no. 35 (Dec. 2005): 7621–7629, https://opg.optica.org/ao/abstract.cfm?uri=ao-44-35-7621.

26. P. Song, R. Wang, L. Loetgering, et al., "Ptycho-Endoscopy on a Lensless Ultrathin Fiber Bundle Tip," *Light: Science & Applications* 13, no. 1 (July 2024): 168, https://www.nature.com/articles/s41377-024-01510-5.

27. R. Jiang, D. Shi, and Y. Wang, "Long-Range Fourier Ptychographic Imaging of the Object in Multidimensional Motion," *Optics Communications* 575 (Jan. 2025): 131307, https://www.sciencedirect.com/science/article/pii/S0030401824010447.

28. M. A. Chan, C. J. Pellizzari, and C. A. Metzler, "Inverse Synthetic Aperture Fourier Ptychography," (July 2025), http://arxiv.org/abs/2507.03733.

29. J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications* (SPIE Press Bellingham, Washington, 2020).

30. H. Zhang, S. Jiang, J. Liao, et al., "Near-Field Fourier Ptychography: Super-Resolution Phase Retrieval via Speckle Illumination," *Optics Express* 27, no. 5 (2019): 7498–7512.

31. P. Song, S. Jiang, T. Wang, et al., "Synthetic Aperture Ptychography: Coded Sensor Translation for Joint Spatial-Fourier Bandwidth Expan-

sion," *Photonics Research* 10, no. 7 (July 2022): 1624, https://opg.optica.org/abstract.cfm?URI=prj-10-7-1624.

32. Y. Luo, Y. Zhao, J. Li, et al., "Computational Imaging without a Computer: Seeing through Random Diffusers at the Speed of Light," *eLight* 2, no. 1 (Jan. 2022): 4, https://doi.org/10.1186/s43593-022-00012-4.

33. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient Subpixel Image Registration Algorithms," *Optics Letters* 33, no. 2 (Jan 2008): 156–158, https://opg.optica.org/ol/abstract.cfm?uri=ol-33-2-156.

34. C. Liu, J. Yuen, and A. Torralba, "SIFT Flow: Dense Correspondence across Scenes and Its Applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, no. 5 (May 2011): 978–994, https://ieeexplore.ieee.org/abstract/document/5551153.

35. A. Maiden, D. Johnson, and P. Li, "Further Improvements to the Ptychographical Iterative Engine," *Optica* 4, no. 7 (July 2017): 736–745, https://opg.optica.org/optica/abstract.cfm?uri=optica-4-7-736.

36. Q. Liu, D. Chen, Q. Chu, et al., "Online Multi-Object Tracking with Unsupervised Re-Identification Learning and Occlusion Estimation," (Jan. 2022), arXiv:2201.01297 [cs], http://arxiv.org/abs/2201.01297.

37. P. Bergmann, T. Meinhardt, and L. Leal-Taixe, "Tracking Without Bells and Whistles," (2019), 941–951, https://ieeexplore.ieee.org/document/9010033.

38. "Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics | EURASIP Journal on Image and Video Processing," https://link.springer.com/article/10.1155/2008/246309.

39. J. Zhang, W. Wei, K. Yang, et al., "Atmospheric Turbulence-Resilient Long-Range Fourier Ptychography," (July 2025), http://arxiv.org/abs/2507.03943.

40. M. Cao, S. Zheng, L. Wang, Z. Chen, D. Brady, and X. Yuan, "Sparse Transformer for Ultra-Sparse Sampled Video Compressive Sensing," (Sept. 2025), http://arxiv.org/abs/2509.08228.

41. Z. Li, P. Wang, H. Zhang, Z. Li, and X. Yuan, "Motion-Aware Reconstruction for Video Snapshot Compressive Imaging," in *2025 IEEE International Conference on Image Processing (ICIP)* (Sept. 2025), 2002–2007, https://ieeexplore.ieee.org/abstract/document/11084499.

42. J. B. Senanayaka and C. A. Metzler, "Fundamental Limits of Long Range Fourier Ptychography: On the Importance of Capturing Data in a Snapshot," in *2025 IEEE Conference on Computational Imaging Using Synthetic Apertures (CISA)* (June 2025), 1–4, https://ieeexplore.ieee.org/document/11091570.

43. A. Dosovitskiy, P. Fischer, E. Ilg, et al., "FlowNet: Learning Optical Flow With Convolutional Networks," (2015), 2758–2766, https://ieeexplore.ieee.org/document/7410673.

44. J. Wu, Y. Guo, C. Deng, et al., "An Integrated Imaging Sensor for Aberration-Corrected 3D Photography," *Nature* 612, no. 7938 (Dec 2022): 62–71, https://www.nature.com/articles/s41586-022-05306-8.

45. J. B. Senanayaka and C. A. Metzler, "Snapshot Synthetic Aperture Imaging with Boiling Speckle," (Sept. 2025), http://arxiv.org/abs/2509.21682.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting File 1: lpor70930-sup-0001-SuppMat.pdf

Supplemental Video 1: lpor70930-sup-0002-VideoS1.mp4

Supplemental Video 2: lpor70930-sup-0003-VideoS2.mp4