remote sensing

MDPI

Article

RCRC-Guided Low-Tubal-Rank Tensor Recovery for Sparse Aperture Terahertz Video SAR Imaging

Xu Wang, Jiahao Zhang, Tongxu Hu, Jialu Liu, Jin Li *, Rui Min, Zongjie Cao and Yiming Pi

The School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
* Correspondence: lijin@uestc.edu.cn

Highlights

What are the main findings?

- An RCRC-aware sensing operator is integrated with low-tubal-rank tensor recovery and pulsewise residual phase estimation for variable-curvature THz video SAR.
- Shared pulse masks and controlled ablations are used to examine geometry modeling and residual phase estimation within the TNN-regularized framework.

What are the implications of the main findings?

- A 40% pulse retention ratio is evaluated as a candidate sparse observation operating point, whereas the 30% case represents a severe robustness test.
- Simulation and measured data experiments evaluate the integrated framework under incomplete slow-time observations.

Abstract

THz-SAR provides high-resolution imaging but is highly sensitive to trajectory perturbations, while video SAR further requires efficient reconstruction of consecutive sub-aperture frames from incomplete measurements. Under variable curvature motion, forward-model errors introduce geometric and phase inconsistencies that weaken the interframe correlation exploited by tensor recovery. This article develops an integrated radius-of-curvature change rate (RCRC)-guided low-tubal-rank recovery framework for sparse aperture THz video SAR. An RCRC-aware variable curvature trajectory operator models deterministic motion-dependent effects, a tensor nuclear-norm prior exploits interframe redundancy, and pulsewise unit modulus phase factors compensate for residual perturbations. The resulting joint problem is solved by alternating proximal-gradient, tensor singular value thresholding, and closed-form phase updates. Controlled point target simulations and measured data experiments are conducted at pulse retention ratios of 60%, 40%, and 30%. A single global slow-time mask is applied to the shared pulse stream before sub-aperture slicing, ensuring consistent pulse omissions across overlapping frames. Comparisons with three literature baselines evaluate the integrated method, while two internal ablations examine the effects of RCRC modeling and residual phase estimation within the adopted TNN-regularized formulation. The 40% case preserves the principal scattering structure and is treated as a candidate sparse observation operating point, whereas the degradation at 30% identifies a more severe robustness boundary. Because the sparse observations are generated retrospectively from full-rate measurements, the experiments validate reconstruction under missing pulses but do not demonstrate physical reductions in PRF, transmitted pulses, or collected raw data volume.

Check for updates

Academic Editor: John Trinder

Received: 23 September 2026
Revised: 26 October 2026
Accepted: 27 October 2026
Published: 29 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: THz-SAR; video SAR; VCT; RCRC; low-tubal-rank tensor recovery; sparse aperture imaging; slow-time sparse observation; residual phase correction

1. Introduction

THz-SAR has attracted increasing attention because the combination of submillimeter wavelengths and wide signal bandwidths enables high-resolution electromagnetic imaging and provides considerable potential for high-frame-rate observation of dynamic scenes [1–3]. However, the short wavelength that provides such high spatial resolution also makes THz-SAR considerably more sensitive to platform motion and trajectory perturbations than conventional microwave SAR, such that small mechanical vibrations, residual motion errors, and uncompensated phase perturbations can lead to appreciable image defocusing and sidelobe degradation [4–6].

This sensitivity becomes particularly important when the radar platform follows a nonlinear or nonideal trajectory for which conventional straight track or constant curvature approximations may introduce errors in the range history, wavenumber mapping, and phase compensation [7–10]. To address this problem, our previous work developed a VCT THz-SAR imaging model in which the RCRC is explicitly incorporated into the wavenumber domain imaging process, enabling more accurate modeling of the spatially variant phase history under nonideal curvilinear motion [11]. Nevertheless, the RCRC-aware VCT method primarily addresses accurate image formation for an individual aperture and does not provide a mechanism for reconstructing intentionally omitted slow-time pulses or exploiting interframe redundancy across overlapping sub-apertures. Although overlapping frame formation increases the computational and representation costs of video reconstruction, the underlying physical pulses are acquired only once in the shared master echo stream.

Video SAR generates a temporally ordered sequence of high-resolution images by continuously forming overlapping sub-apertures, thereby enabling persistent scene observation and dynamic target analysis but at the cost of sustained full-rate acquisition and repeated processing of highly overlapping apertures and repeated image formation [12–14]. The challenge is even more pronounced for THz video SAR, where high frame rates and fine spatial resolution must be achieved while maintaining accurate phase coherence among consecutive frames in the presence of platform perturbations [15,16]. Therefore, reducing the number of required measurements without destroying the geometric and temporal coherence of the image sequence is a key issue for practical high-resolution THz video SAR.

Compressive sensing and sparsity-driven radar imaging provide a natural approach to enabling reconstruction from incomplete or sparsely retained slow-time measurements by exploiting the sparse or compressible structure of radar scenes and have been extensively investigated for SAR and ISAR reconstruction from incomplete measurements [17–20]. However, sparse reconstruction becomes considerably more difficult when incomplete measurements are accompanied by motion-induced phase errors because inaccuracies in the observation operator directly violate the data consistency model on which sparse recovery relies [21–24]. Joint sparsity and low-rank priors have consequently been introduced into sparse aperture radar imaging to exploit complementary structural information and improve reconstruction robustness under incomplete observations [25,26].

For video SAR, the strong redundancy among consecutive frames provides an additional source of prior information, and multi-frame image sequences have therefore been modeled as low-rank tensors to improve reconstruction when only a subset of the available slow-time pulses is retained while jointly exploiting spatial and temporal corre-

lations [27,28]. More recently, low-tubal-rank models based on the tensor singular value decomposition (t-SVD) have provided a more structured representation of temporal correlation, with the tubal rank and tensor nuclear norm offering effective tools for recovering multidimensional data from incomplete observations [29,30]. These tensor-based approaches, however, implicitly rely on the assumption that corresponding scene structures remain sufficiently coherent across frames; when the forward sensing model does not accurately represent the platform trajectory, frame-dependent phase distortions may reduce the apparent interframe correlation and consequently weaken the effectiveness of the low-rank or low-tubal-rank prior [31,32].

In parallel, model-driven deep unfolding and deep-prior approaches have been developed for sparse SAR imaging, autofocus, and video SAR reconstruction, improving reconstruction efficiency and allowing data-adaptive priors to be incorporated into iterative imaging frameworks [33–37]. Although these learning-based methods provide powerful reconstruction capabilities, the fundamental observation model remains critical: a learned or handcrafted prior may be unable to compensate reliably for a deterministic mismatch between the assumed sensing geometry and the actual variable curvature trajectory.

Furthermore, even after deterministic trajectory-dependent errors have been modeled, residual phase perturbations caused by imperfect motion measurements, mechanical vibration, or other unmodeled effects may remain and can further degrade high-resolution SAR focusing [38,39]. Existing autofocus and joint imaging approaches estimate such phase errors either independently or simultaneously with the image. To the best of our knowledge, their coordinated use within an RCRC-aware multi-frame model, together with a shared pulse sparse observation protocol, has not been systematically investigated for THz video SAR [40,41]. Consequently, a key unresolved issue is how to simultaneously preserve the physical fidelity of variable-curvature THz-SAR sensing and exploit the temporal redundancy of video SAR when only incomplete measurements are available.

Motivated by this observation, this article proposes an RCRC-guided low-tubal-rank tensor recovery framework for sparse aperture THz video SAR imaging in which the deterministic variable curvature geometry is embedded directly into the multi-frame forward model rather than treated as a postprocessing correction. By reconstructing the video tensor under an RCRC-aware sensing operator, geometrically coherent frames are obtained before the low-tubal-rank constraint is imposed, allowing the temporal redundancy among adjacent sub-apertures to be more effectively exploited. A unit modulus residual phase model is further incorporated to accommodate microscopic phase perturbations that are not fully described by the deterministic RCRC model, and the resulting optimization problem is solved through alternating image tensor and residual phase updates. This study focuses on reconstruction quality under incomplete observations, rather than real-time image delivery.

The main contributions of this work are summarized as follows.

(1) A geometry-consistent multi-frame observation model is formulated for sparse slow-time THz video SAR under variable curvature motion. The previously developed RCRC-aware VCT operator is embedded into the frame-dependent sensing process and combined with a single pulse selection mask applied to the shared master pulse stream before overlapping sub-apertures are formed. Consequently, an omitted physical pulse is absent from every frame that references it, avoiding artificial information reuse across overlapping frames.

(2) A unified reconstruction model couples the RCRC-aware observation operator with tensor nuclear norm regularization and pulsewise unit modulus residual phase variables. Within this formulation, controlled ablations isolate the complementary roles of RCRC geometry modeling and phase correction, while the low-tubal-rank prior exploits multi-frame redundancy.

(3) A practical alternating solver is derived for the joint image tensor and residual phase problem. The image update combines a proximal gradient step with tensor singular value thresholding, whereas the pulsewise phase update admits a closed-form unit modulus solution with an explicit safeguard for vanishing correlations. The same stopping criteria and reconstruction protocol are used for the complete method and its component ablations.

(4) A controlled evaluation protocol is established using point target simulations and measured THz-SAR data at pulse retention ratios of 60%, 40%, and 30%. Three literature baselines, LRT-AM, JLRS-T, and ORTP, are compared with constant curvature and no residual phase ablations to examine the effects of geometry modeling and residual phase estimation within the TNN-regularized framework. The 40% case is examined as a candidate sparse observation operating point, whereas 30% is treated as a severe robustness test. Because the masks are applied retrospectively to full-rate measurements, the experiments evaluate reconstruction from missing pulses rather than claiming a demonstrated reduction in acquisition PRF or raw data volume.

The remainder of this article is organized as follows. Section 2 reviews related work on variable curvature THz-SAR imaging, tensor-based video SAR recovery, and autofocus imaging. Section 3 develops the RCRC-aware multi-frame THz video SAR observation model. Section 4 formulates the RCRC-guided low-tubal-rank recovery model with residual phase correction. Section 5 derives the corresponding alternating optimization algorithm. Section 6 presents the experimental results and discussion, and Section 7 concludes this article.

2. Related Work

2.1. Geometry-Aware SAR Imaging Under Nonideal Trajectories

Accurate trajectory modeling is fundamental to high-resolution SAR image formation when the radar platform deviates from an ideal linear or circular path. For nonlinear flight trajectories, geometry-aware imaging methods have been developed to explicitly account for the actual platform motion during range history modeling and image formation [7,8]. In parallel, two-dimensional autofocus and motion compensation approaches have been investigated to mitigate residual spatially variant or frequency-dependent phase errors that remain after conventional motion correction [9,10].

These methods substantially improve imaging fidelity under nonideal platform motion, but their primary objective is accurate reconstruction within an individual aperture. In our previous work [11], the RCRC was introduced into a VCT THz-SAR imaging model. By incorporating the local variation of the curvature radius into the wavenumber domain formulation, the RCRC-based model provides a more accurate description of the range history and trajectory-dependent phase evolution than constant curvature approximations.

However, directly extending framewise geometry-aware imaging to video SAR does not exploit the redundancy among overlapping sub-apertures. More importantly, when consecutive frames are reconstructed under an inaccurate trajectory model, the resulting geometric and phase inconsistencies affect not only the focusing quality of each frame but also the correlation among frames. Therefore, for tensor-based video SAR reconstruction, trajectory modeling should be incorporated into the multi-frame sensing process rather than treated solely as an independent framewise imaging step.

2.2. Tensor-Based Reconstruction for Video SAR

Sparse and low-rank regularization has been extensively used to reconstruct radar images from incomplete measurements [17–23]. For video SAR, the multi-frame structure provides additional redundancy that can be exploited beyond conventional pixelwise

sparsity. Pu et al. formulated video SAR imaging as a low-rank tensor recovery problem, enabling correlated information among adjacent frames to contribute jointly to image reconstruction [27]. Joint low-rank and sparse tensor models were subsequently introduced to separate temporally correlated scene components from localized variations [28].

Low-tubal-rank modeling further provides a structured representation of the temporal dependence of video SAR sequences. The ORTP framework in [29] formulates video SAR reconstruction using a low-tubal-rank prior, while the underlying t-SVD and tensor nuclear norm theories provide effective tools for representing and recovering multidimensional low-rank data [30–32]. Compared with independent frame reconstruction, these tensor formulations can substantially reduce the required measurements by exploiting shared information across frames.

A common characteristic of these approaches, however, is that the tensor prior is imposed under an assumed forward observation model. If this model does not adequately describe the actual trajectory-dependent phase history, the recovered frames may exhibit deterministic interframe inconsistencies. From the viewpoint of tensor recovery, such inconsistencies are particularly unfavorable because they reduce the apparent temporal correlation and distribute the signal energy over additional tensor components. Consequently, a low-rank or low-tubal-rank prior may be forced to explain model mismatch that originates from the sensing geometry rather than from the intrinsic scene structure.

This issue is particularly relevant to the problem considered in this article. The objective is therefore not simply to introduce a stronger tensor regularizer but to apply the low-tubal-rank prior to frames reconstructed under a physically consistent RCRC-aware forward model. In this manner, the tensor prior is used to exploit true interframe redundancy rather than to compensate for deterministic variable curvature modeling errors.

2.3. Joint Image Reconstruction and Phase Error Estimation

Joint image reconstruction and phase error estimation provide another important line of work closely related to the proposed framework. Sparsity-driven methods jointly estimate the radar image and unknown phase errors by embedding phase correction into the reconstruction model [21]. Entropy-constrained, sparse Bayesian, and related optimization methods have subsequently been developed for sparse aperture imaging and autofocus under incomplete measurements [24–26]. Other formulations estimate residual phase errors directly through optimization-based autofocus criteria [40,41].

More recently, model-driven deep networks and unfolded optimization methods have been introduced to combine physical observation models with trainable regularization or iterative structures. Representative studies include joint low-rank and sparse recovery networks [33], deep SAR imaging and motion compensation [34], and learning-based autofocus methods [35,38,39]. Similar ideas have also been extended to video SAR through tensor recovery combined with unfolded or deep prior architectures [36,37].

Although these methods improve reconstruction or autofocus performance, most do not explicitly distinguish deterministic trajectory-induced phase variation from residual microscopic phase perturbations within an RCRC-aware multi-frame model. If these two effects are represented by a single unknown phase term, the phase estimator may be required to compensate simultaneously for predictable geometric mismatch and genuinely residual perturbations.

The formulation adopted in this article separates these two roles. The RCRC-dependent VCT operator models the deterministic macroscopic variation of the platform trajectory, whereas the unit modulus residual phase term accounts only for the remaining phase perturbations after geometric correction. The resulting framework therefore combines three established components: physics-based RCRC-aware sensing, low-tubal-rank

regularization for interframe redundancy, and blind residual phase correction for unmodeled perturbations. This separation enables the tensor prior and residual phase estimator to operate on a multi-frame sequence whose dominant deterministic geometric variations have already been accounted for.

3. RCRC-Aware THz Video SAR Observation Model

3.1. Multi-Frame Variable Curvature Geometry

Consider a THz-SAR system that continuously acquires a sequence of overlapping sub-apertures, each corresponding to one video SAR frame. Let $t = 1, \cdots, T$ denote the frame index and η denote slow time within the t-th sub-aperture. The two-dimensional reflectivity distribution of the t-th frame is denoted by

$$\mathbf{X}_t = \mathcal{X}(:,:,t) \in \mathbb{C}^{M\times N}, \tag{1}$$

where M and N are the numbers of cross-range and range pixels, respectively. The complete video SAR sequence is represented by the third-order tensor

$$\mathcal{X} \in \mathbb{C}^{M\times N\times T}. \tag{2}$$

The overall imaging geometry and multi-frame data organization are illustrated in Figure 1. The continuous synthetic aperture is partitioned into a sequence of overlapping sub-apertures, with each sub-aperture corresponding to one video SAR frame. For the t-th frame, the acquired echo samples are indexed by the slow-time pulse p and fast-time frequency q, forming the framewise echo matrix $\mathbf{Y}_t \in \mathbb{C}^{P\times Q}$. The corresponding scene is discretized into $M \times N$ spatial cells indexed by (m, n), giving the reflectivity image $\mathcal{X} \in \mathbb{C}^{M\times N\times T}$. Stacking the framewise echo matrices and reflectivity images along the frame dimension yields the echo tensor $\mathcal{Y} \in \mathbb{C}^{P\times Q\times T}$. and the image tensor $\mathcal{X} \in \mathbb{C}^{M\times N\times T}$, respectively. The RCRC-aware forward model developed in the following subsection establishes the frame-dependent mapping between X_t and $\mathbf{Y}_t$ under variable curvature motion.

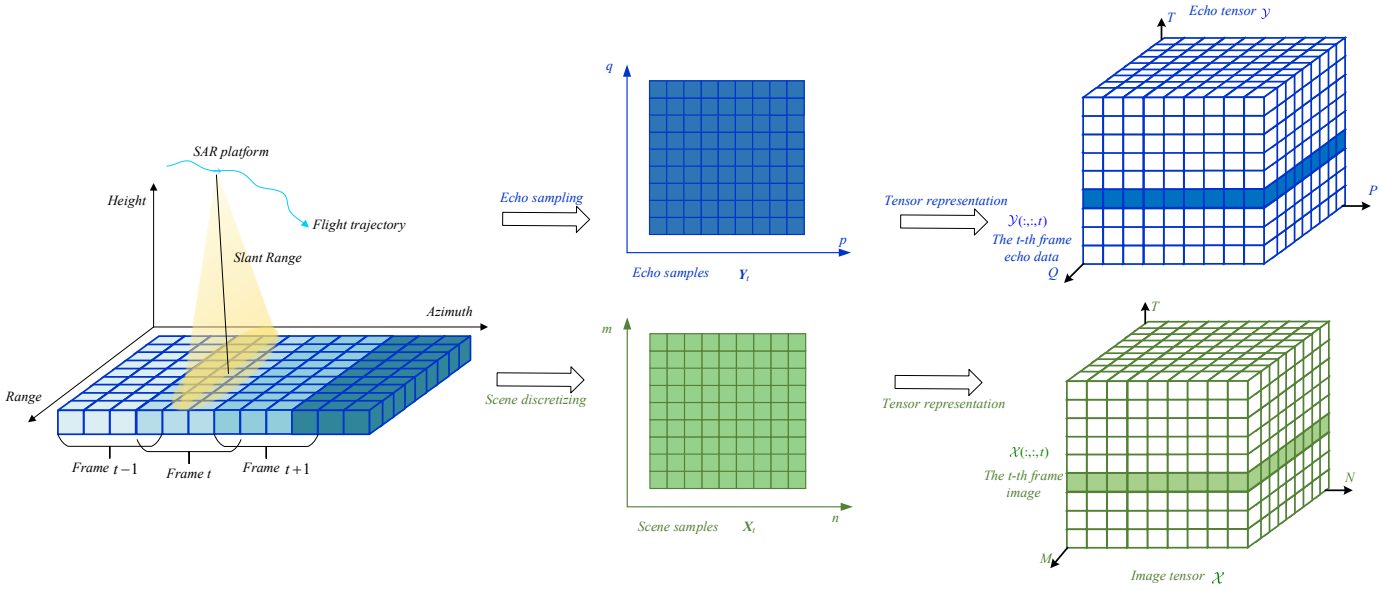


Figure 1. Geometry configuration, echo sampling, and multi-frame tensor organization of VCT THz video SAR.

Let the projected radar trajectory in the t-th sub-aperture be

$$\mathrm{r}_t(\eta) = [x_t(\eta), y_t(\eta)]^{\mathrm{T}}, \tag{3}$$

and let the spatial position of the (m,n)-th pixel be

$$\mathrm{s}_{m,n} = [x_m, y_m]^{\mathrm{T}}. \tag{4}$$

The instantaneous slant range is

$$R_t(\eta; m, n) = \|\mathrm{r}_t(\eta) - \mathrm{s}_{m,n}\|_2. \tag{5}$$

Under ideal circular motion, the curvature radius is assumed to be constant. Under variable curvature motion, this assumption is no longer valid. Within a short sub-aperture, the curvature radius can be approximated by a first-order model

$$\rho_t(\eta) = \rho_{0,t} + \rho_{r,t}\eta, \tag{6}$$

where $\rho_{0,t}$ is the curvature radius at the center of the t-th sub-aperture and $\rho_{r,t} = \left.\frac{d\rho_t(\eta)}{d\eta}\right|_{\eta=\eta_t}$ denotes the RCRC evaluated at the center of the t-th sub-aperture. The RCRC sequence $\boldsymbol{\rho}_r = [\rho_{r,1}, \rho_{r,2}, \cdots, \rho_{r,T}]$ characterizes the deterministic macroscopic trajectory variation across video frames.

3.2. RCRC-Aware VCT Forward Operator

For the t-th frame, the baseband echo at slow-time index p and fast-time frequency index q can be modeled as

$$Y(p,q,t) = \sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n,t)\phi_{\mathrm{VCT}}(p,q,m,n,t;\rho_{r,t}) + N(p,q,t), \tag{7}$$

where $N(p,q,t)$ denotes measurement noise and $\phi_{VCT}(\cdot)$ is the VCT-dependent sensing kernel. This kernel contains the RCRC-dependent slant range, propagation phase, range frequency modulation, wavenumber domain interpolation, and phase compensation terms. A representative phase component can be written as

$$\begin{aligned}\phi_{VCT}(p,q,m,n,t;\rho_{r,t}) = w_a(p)w_r(q) \quad & \times \exp\left\{-j\pi K_r\left(\tau(q) - \tfrac{2R_t(p;m,n;\rho_{r,t})}{c}\right)^2\right\} \\ & \times \exp\left\{-j4\pi f_c \tfrac{R_t(p;m,n;\rho_{r,t})}{c}\right\},\end{aligned} \tag{8}$$

where c is the speed of light, $w_a(p)$ and $w_r(q)$ are slow and fast time signal envelopes, f_c is the carrier frequency, $R_t(p;m,n;\rho_{r,t})$ denotes the RCRC-dependent range history, K_r is the range chirp FM rate, τ is the fast time, and $\tau(q)$ denotes the fast time at the q-th sampling. Therefore, changes in the local curvature radius and its rate of variation directly modify the phase history associated with each image pixel.

In addition to the propagation phase in (8), the complete VCT sensing kernel incorporates the RCRC-dependent range frequency modulation, wavenumber domain mapping and interpolation, and space-variant phase compensation terms derived from the variable curvature geometry [11]. These operations are collectively represented by the framewise RCRC-aware VCT forward operator

$$\mathbf{y}_t = \mathcal{A}_{\mathrm{VCT},t}(\mathbf{x}_t;\rho_{r,t}) + \mathbf{n}_t, \tag{9}$$

where $\mathbf{x}_t = \mathrm{vec}(\mathbf{X}_t)$ is the vectorized reflectivity distribution of the t-th frame, $\mathbf{y}_t = \mathrm{vec}(\mathbf{Y}_t)$ denotes the vectorized echo of the t-th frame, $\mathbf{n}_t$ is the vectorized noise of the t-th frame, and $\mathcal{A}_{\mathrm{VCT},t}$ is the RCRC-aware VCT forward operator.

It is worth emphasizing that $\mathcal{A}_{\mathrm{VCT},t}$ is not identical for all video frames. Because consecutive sub-apertures generally have different local curvature parameters, the forward operator varies with t through $\rho_{r,t}$. Nevertheless, all framewise operators are constructed according to the same RCRC-aware geometric model. Stacking the observations of all T frames give

$$\mathcal{Y} = \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r) + \mathcal{N}, \tag{10}$$

where $\mathcal{A}_{\mathrm{RCRC}}$ applies the corresponding framewise VCT operator to each sub-aperture and $\mathcal{N} \in \mathbb{C}^{M \times N \times T}$ denotes the noise tensor.

Unlike conventional video SAR tensor recovery formulations in which the sensing operator is usually constructed from an idealized or geometry-independent observation model, (10) explicitly preserves the frame-dependent variable curvature geometry. Consequently, the subsequent low-tubal-rank constraint is imposed on an image sequence reconstructed under a physically consistent RCRC-aware sensing model rather than being used to compensate for deterministic trajectory model mismatch.

3.3. Sparse Aperture Observation Model

To evaluate and handle incomplete measurements along the slow-time dimension, we formulate a sparse observation model. Note that sparse observation is defined relative to the full-PRF recorded measurement rate, which is standard in the compressed sensing radar literature.

To ensure that the missing measurements correspond to genuine physical pulse omissions across the entire observation period, the sparse aperture sampling mask is applied directly to the global, shared slow-time pulse stream, namely the master raw echo matrix prior to sub-aperture slicing, rather than independently generated for each frame. Consequently, if a specific physical pulse is omitted, it is consistently absent across all overlapping video SAR frames that reference it, thereby establishing a rigorous test for multi-frame tensor recovery.

To evaluate reconstruction from incomplete slow-time observations, only a subset of the fully recorded echo measurements is retained for reconstruction. Let Ω_t denote the set of observed measurements in the t-th sub-aperture, and let $\mathcal{P}_{\Omega_t}$ denote the corresponding sampling operator. The sparse aperture observation of the t-th frame is then given by

$$\mathbf{y}_{t,\Omega_t} = \mathcal{P}_{\Omega_t}[\mathcal{A}_{\mathrm{VCT},t}(\mathbf{x}_t; \rho_{r,t})] + \mathbf{n}_{t,\Omega_t}, \tag{11}$$

where $\mathbf{y}_{t,\Omega_t}$ denotes the observed echo samples and $\mathbf{n}_{t,\Omega_t}$ denotes the corresponding measurement noise.

For the complete video SAR sequence, the framewise observations can be stacked into a third-order measurement tensor. Accordingly, the multi-frame sparse aperture observation model is expressed as

$$\mathcal{Y}_\Omega = \mathcal{P}_\Omega \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r) + \mathcal{N}_\Omega, \tag{12}$$

where $\mathcal{P}_\Omega$ denotes the multi-frame sampling operator obtained by stacking the framewise sampling operators $\{\mathcal{P}_{\Omega_t}\}_{t=1}^{T}$ and $\mathcal{N}_\Omega$ denotes the noise tensor restricted to the observed entries.

The slow-time observation ratio is defined as

$$\gamma_t = \frac{|\Omega_t|}{\left|\Omega_t^{full}\right|}, \tag{13}$$

where $|\Omega_t|$ and $\left|\Omega_t^{full}\right|$ denote the numbers of retained and fully sampled measurements, respectively. For a common slow-time observation ratio across all frames, $\gamma_t = \gamma$ holds for $t = 1, \cdots, T$.

Let S denote the sub-aperture length; s is the step size, then the number of frames is $T = \lfloor (P - S)/s \rfloor + 1$, and the overlap ratio is $\varpi_{ov} = 1 - s/P$. The expanded tensor representation contains $P \times S \times T$ entries. Compared with the original echo matrix, the expansion factor is $\xi = (S \times T)/Q$. This expansion is introduced by the overlapping video-SAR representation and should not be interpreted as the number of physically acquired samples. The same physical pulse may be referenced by multiple adjacent frames, but it is acquired only once in the raw echo domain.

Under this formulation, sparse aperture observation acquisition is introduced through the sampling operator $\mathcal{P}_\Omega$, while the physical sensing process is described independently by the RCRC-aware operator $\mathcal{A}_{\mathrm{RCRC}}$. This separation allows the effects of trajectory-dependent sensing geometry and incomplete measurements to be explicitly distinguished in the subsequent tensor recovery formulation.

3.4. Pulsewise Residual Phase Model

The RCRC-aware VCT model accounts for deterministic macroscopic phase variation induced by the nonideal variable curvature trajectory. However, practical THz-SAR measurements may still contain residual phase errors caused by mechanical vibration, air path fluctuation, thermal drift, imperfect trajectory measurements, and other unmodeled small-scale motion effects. These residual errors are not explicitly represented by the RCRC model and therefore need to be estimated from the measured data. To account for them, we introduce a unit modulus phase tensor

$$\mathbf{\Phi}_t(p, q) = \exp[j\theta_t(p, q)], \ |\mathbf{\Phi}_t(p, q)| = 1, \tag{14}$$

where $\mathbf{\Phi}_t(p, q)$ and $\theta_t(p, q)$ denote the residual phase at the p-th slow-time sample and q-th fast-time frequency sample. The unit modulus constraint ensures that the residual phase term modifies only the phase of the predicted echo without introducing additional amplitude scaling.

Accordingly, the complete observation model can be written as

$$\mathcal{Y}_\Omega = \mathcal{P}_\Omega[\mathbf{\Phi} \odot \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r)] + \mathcal{N}_\Omega, \tag{15}$$

where $\mathbf{\Phi}$ denotes the residual phase tensor and $\odot$ represents element-wise multiplication.

However, estimating an independent phase variable for every fast-time sample introduces a large number of additional unknowns and may reduce the identifiability of the phase correction problem under highly undersampled measurements. In practical SAR autofocus, the dominant residual phase variation is generally associated with the slow-time dimension. Therefore, this work adopts a pulsewise residual phase model in which the phase is assumed to be approximately common across the fast-time samples belonging to the same pulse. Specifically,

$$\phi_t(p) = \exp[j\theta_t(p)], \ |\phi_t(p)| = 1, \tag{16}$$

and the corresponding phase factor is broadcast along the fast-time dimension. The complete observation model used in the proposed method therefore becomes

$$\mathcal{Y}_\Omega = \mathcal{P}_\Omega\left[\mathbf{\Phi}_{pw} \odot \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r)\right] + \mathcal{N}_\Omega, \tag{17}$$

where $\mathbf{\Phi}_{pw}$ denotes the pulsewise residual phase tensor obtained by broadcasting $\phi_t(p)$ along the fast-time dimension.

The pulsewise formulation substantially reduces the number of phase unknowns while retaining the dominant slow-time-dependent phase variation. More importantly, it separates the two different sources of phase distortion in the proposed framework. The RCRC-aware VCT operator models the deterministic macroscopic phase variation caused by the variable curvature trajectory, whereas $\mathbf{\Phi}_{pw}$ estimates the remaining residual phase perturbations that are not explained by the geometric model.

This separation allows the low-tubal-rank prior to be imposed on a video SAR sequence whose dominant trajectory-dependent phase variation has already been compensated by the RCRC-aware forward model while the remaining phase inconsistency is jointly estimated during tensor recovery.

4. RCRC-Aware Low-Tubal-Rank Tensor Recovery Formulation

4.1. Low-Tubal-Rank Prior for Multi-Frame THz Video SAR

In THz video SAR, consecutive sub-apertures observe largely overlapping spatial regions, and scene reflectivity generally exhibits strong correlation across adjacent frames. After the deterministic trajectory-dependent phase variation is incorporated into the RCRC-aware forward model, the resulting multi-frame reflectivity images can therefore be jointly represented as a third-order tensor

$$\mathcal{X} = stack(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_T) \in \mathbb{C}^{M\times N\times T}, \tag{18}$$

Under the sub-aperture configuration of sub-aperture length $S = 3000$, step size $s = 1000$ and overlap ratio $\varpi_{ov} \approx 67\%$, each pulse is shared by at most $S/s = 3$ consecutive frames, meaning frames separated by a lag $\Delta t \geq 3$ share no raw pulses. Consequently, the complex THz SAR scene exhibits heterogeneous temporal correlation: dominant human-made structures (e.g., building edges, roads, and strong scatterers) maintain strong cross-frame structural persistence due to continuous antenna footprint illumination, forming a pronounced low-tubal-rank subspace. Conversely, distributed clutter (e.g., farmland) exhibits local correlation only within adjacent frames sharing pulses ($\Delta t \leq 2$) while rapidly decorrelating into independent speckle realizations for $\Delta t \geq 3$.

Let

$$\overline{\mathcal{X}} = fft(\mathcal{X}, [], 3) \tag{19}$$

denote the discrete Fourier transform of $\mathcal{X}$ along the third mode, and let $\overline{\mathbf{X}}^{(k)}$ denote the kth frontal slice of $\overline{\mathcal{X}}$. The tubal rank of $\mathcal{X}$ is defined as

$$rank_t(\mathcal{X}) = \max_k rank(\overline{\mathbf{X}}^{(k)}). \tag{20}$$

Direct minimization of the tubal rank is nonconvex and computationally difficult. Therefore, the tensor nuclear norm (TNN) is adopted as its convex surrogate:

$$\|\mathcal{X}\|_{\mathrm{TNN}} = \frac{1}{T}\sum_{k=1}^{T}\left\|\overline{\mathbf{X}}^{(k)}\right\|_*, \tag{21}$$

where $\|\cdot\|_*$ denotes the matrix nuclear norm. Equivalently,

$$\|\mathcal{X}\|_{\mathrm{TNN}} = \frac{1}{T}\sum_{k=1}^{T}\sum_{i}\sigma_i(\overline{\mathbf{X}}^{(k)}), \tag{22}$$

where $\sigma_i(\cdot)$ denotes the ith singular value. Minimizing the TNN promotes a compact tubal rank representation and enables the recovery process to exploit the spatial–temporal redundancy among consecutive THz video SAR frames.

It is worth emphasizing that the low-tubal-rank constraint is imposed on the multi-frame image tensor rather than the measured echo tensor. Moreover, this prior is introduced to exploit inter-frame redundancy after the deterministic trajectory variation has been explicitly incorporated into the RCRC-aware sensing model. Therefore, the tensor prior complements the physics-based RCRC model instead of compensating for deterministic trajectory mismatch. Based on this structural prior and the observation model derived in Section 3, the RCRC-aware tensor recovery formulation is established in the following subsection.

4.2. RCRC-Aware Tensor Recovery Under Sparse-Aperture Sampling

Based on the RCRC-aware observation model established in Section 3, the multi-frame THz video SAR reconstruction can be formulated as an inverse problem under incomplete measurements. We first present the phase-free reconstruction formulation before introducing joint residual phase estimation. Define the composite sensing operator as

$$\mathcal{H}_{\Omega,\boldsymbol{\rho}_r} = \mathcal{P}_\Omega \circ \mathcal{A}_{\mathrm{RCRC}}(\cdot;\boldsymbol{\rho}_r). \tag{23}$$

Accordingly, the observation model is compactly expressed as

$$\mathcal{Y}_\Omega = \mathcal{H}_{\Omega,\boldsymbol{\rho}_r}(\mathcal{X}) + \mathcal{N}_\Omega. \tag{24}$$

Under sparse aperture sampling, direct inversion of (24) becomes ill-conditioned or underdetermined because only a subset of the full measurements is available. To exploit the potentially compressible multi-frame structure discussed in Section 4.1, we introduce TNN regularization into the reconstruction problem. Therefore, the reconstruction is regularized by the tensor nuclear norm, leading to

$$\hat{\mathcal{X}} = \underset{\mathcal{X}}{\operatorname{argmin}} \frac{1}{2}\left\|\mathcal{H}_{\Omega,\boldsymbol{\rho}_r}(\mathcal{X}) - \mathcal{Y}_\Omega\right\|_F^2 + \lambda\|\mathcal{X}\|_{TNN} \tag{25}$$

where the first term enforces consistency between the reconstructed echoes and the available measurements, whereas the second term promotes the low-tubal-rank structure of the multi-frame image tensor. The regularization parameter $\lambda > 0$ controls the tradeoff between measurement fidelity and inter-frame structural redundancy.

For fixed trajectory parameters $\boldsymbol{\rho}_r$, the RCRC-aware forward operator remains linear with respect to the reflectivity tensor $\mathcal{X}$. Hence, the deterministic trajectory-dependent phase history is explicitly embedded in the sensing operator rather than being absorbed into tensor regularization. This distinction is important because the TNN term is used only to exploit the intrinsic correlation among consecutive SAR frames, while the RCRC-aware operator accounts for the frame-dependent propagation geometry.

4.3. Joint Tensor Recovery with Blind Residual Phase Correction

Although the formulation in (25) explicitly accounts for the deterministic trajectory-dependent phase history through the RCRC-aware sensing operator, practical THz SAR

measurements may still contain residual phase perturbations caused by unmodeled platform motion, vibration, propagation fluctuations, and measurement imperfections. As described in Section 3.4, such residual phase errors are modeled by pulsewise unit modulus phase factors. Ignoring these residual phase perturbations may introduce phase inconsistency among the measurements and consequently degrade coherent image reconstruction. Let

$$\phi = \{\phi_t(p)\},\ |\phi_t(p)| = 1, \tag{26}$$

denote the unknown pulsewise residual phase factors for the multi-frame measurements. Because the residual phase is assumed to remain constant across the fast-time samples within each transmitted pulse, a broadcasting operator $\mathcal{B}(\cdot)$ is introduced such that

$$[\mathcal{B}(\phi)]_{p,q,t} = \phi_t(p),\ \forall q. \tag{27}$$

Accordingly, the RCRC-aware sparse aperture observation model with residual phase perturbation becomes

$$\mathcal{Y}_\Omega = \mathcal{P}_\Omega[\mathcal{B}(\phi) \odot \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r)] + \mathcal{N}_\Omega, \tag{28}$$

Combining the observation model in (28) with the low-tubal-rank prior introduced in Section 4.1 leads to the following joint image and residual phase recovery problem:

$$\begin{aligned} (\hat{\mathcal{X}}, \hat{\phi}) = \quad & \underset{\mathcal{X},\phi}{\operatorname{argmin}} \tfrac{1}{2}\|\mathcal{P}_\Omega[\mathcal{B}(\phi) \odot \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r)] - \mathcal{Y}_\Omega\|_F^2 + \lambda\|\mathcal{X}\|_{TNN} \\ & s.t.\ |\phi_t(p)| = 1,\ \forall t, p. \end{aligned} \tag{29}$$

The data fidelity term enforces consistency between the available sparse aperture measurements and the echoes predicted by the RCRC-aware physical model after residual phase modulation. The TNN regularizer exploits the spatial–temporal redundancy of consecutive THz video SAR frames, whereas the unit modulus constraint ensures that ϕ represents a pure phase perturbation without modifying the echo amplitude. In this formulation, the deterministic phase evolution associated with the varying curvature trajectory is modeled explicitly by $\mathcal{A}_{\mathrm{RCRC}}$, while ϕ accounts only for residual phase errors that are not captured by the deterministic trajectory model.

Compared with assigning an independent phase variable to every slow-time/fast-time sample, the pulsewise parameterization significantly reduces the number of nuisance variables and improves the identifiability of the joint reconstruction problem, which is particularly important under sparse aperture measurements.

Problem (29) is jointly nonconvex because the image tensor and residual phase variables are multiplicatively coupled. Nevertheless, when ϕ is fixed, the problem reduces to a TNN-regularized tensor reconstruction problem with respect to $\mathcal{X}$; when $\mathcal{X}$ is fixed, the residual phase factors can be estimated pulse by pulse under the unit modulus constraint. This block-wise structure enables an efficient alternating optimization strategy, which is developed in Section 5.

5. Alternating Optimization for Joint Tensor and Residual Phase Recovery

5.1. Proximal Gradient Image Tensor Update

To solve the joint recovery problem in (29), an alternating optimization strategy is adopted. At the k-th outer iteration, the residual phase variables are first fixed at $\phi^{(k)}$, and the image tensor is updated while retaining both the RCRC-aware physical model and the low-tubal-rank regularization.

For notational convenience, define the phase-conditioned sensing operator

$$\mathcal{H}_{\phi^{(k)}}(\mathcal{X}) = \mathcal{P}_\Omega[\mathcal{B}(\phi) \odot \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r)]. \tag{30}$$

With $\phi^{(k)}$ fixed, the image tensor subproblem becomes

$$\mathcal{X}^{(k+1)} = \underset{\mathcal{X}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathcal{H}_{\phi^{(k)}}(\mathcal{X}) - \mathcal{Y}_\Omega \right\|_F^2 + \lambda \|\mathcal{X}\|_{TNN}, \tag{31}$$

which consists of a smooth quadratic data fidelity term and a nonsmooth TNN regularizer. Define

$$f_k(\mathcal{X}) = \frac{1}{2} \left\| \mathcal{H}_{\phi^{(k)}}(\mathcal{X}) - \mathcal{Y}_\Omega \right\|_F^2. \tag{32}$$

Because the RCRC-aware forward operator is linear with respect to $\mathcal{X}$ for fixed trajectory parameters, the gradient of the data fidelity term is

$$\nabla f_k(\mathcal{X}) = \mathcal{H}_{\phi^{(k)}}^H \left[\mathcal{H}_{\phi^{(k)}}(\mathcal{X}) - \mathcal{Y}_\Omega \right] \tag{33}$$

where $\mathcal{H}_{\phi^{(k)}}^H$ denotes the Hermitian adjoint of the phase-conditioned sensing operator. More explicitly,

$$\mathcal{H}_{\phi^{(k)}}^H(\mathcal{Z}) = \mathcal{A}_{RCRC}^H \left(\mathcal{B}\left(\phi^{(k)}\right)^* \odot \mathcal{P}_\Omega^H(\mathcal{Z}); \boldsymbol{\rho}_r \right), \tag{34}$$

where $(\cdot)^*$ denotes complex conjugation and $\mathcal{P}_\Omega^H$ is the adjoint sampling operator.

A gradient step is then performed as

$$\mathcal{Z}^{(k)} = \mathcal{X}^{(k)} - \mu_k \nabla f_k(\mathcal{X}^{(k)}), \tag{35}$$

where $\mu_k > 0$ denotes the step size.

The low-tubal-rank prior is subsequently imposed through the proximal mapping

$$\mathcal{X}^{(k+1)} = prox_{\mu_k \lambda \|\cdot\|_{TNN}} \left(\mathcal{Z}^{(k)} \right). \tag{36}$$

The proximal operator in (36) is implemented by tensor singular value thresholding under the t-SVD framework, as described in Section 5.2.

5.2. TNN Proximal Mapping via Tensor Singular Value Thresholding

The proximal mapping in (36) corresponds to the tensor nuclear norm regularization and can be efficiently evaluated using tensor singular value thresholding (t-SVT). Specifically, let

$$\tau_k = \mu_k \lambda. \tag{37}$$

Then, the proximal step can be equivalently written as

$$\mathcal{X}^{(k+1)} = \underset{\mathcal{X}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathcal{X} - \mathcal{Z}^{(k)} \right\|_F^2 + \tau_k \|\mathcal{X}\|_{TNN}. \tag{38}$$

According to the t-SVD framework, this problem can be decomposed into independent matrix singular value thresholding operations in the Fourier domain. First, the intermediate tensor $\mathcal{Z}^{(k)}$ is transformed along the third mode as

$$\overline{\mathcal{Z}}^{(k)} = fft(\mathcal{Z}^{(k)}, [], 3). \tag{39}$$

Let $\overline{\mathbf{Z}}^{(k,l)}$ denote the lth frontal slice of $\overline{\mathcal{Z}}^{(k)}$, where $l = 1, \cdots, T$. The matrix singular value decomposition of each frontal slice is

$$\overline{\mathbf{Z}}^{(k,l)} = \mathbf{U}_l^{(k)} \mathbf{\Sigma}_l^{(k)} \mathbf{V}_l^{(k)H}, \tag{40}$$

where $\mathbf{\Sigma}_l^{(k)}$ contains the singular values of the l-th Fourier domain frontal slice. The singular value shrinkage operator is defined as

$$\mathcal{S}_{\tau_k}\left(\mathbf{\Sigma}_l^{(k)}\right) = diag\left[\left(\sigma_i^{(k,l)} - \tau_k\right)_+\right], \tag{41}$$

where $(a)_+ = \max(a, 0)$. Accordingly, the thresholded frontal slices are obtained as

$$\overline{\mathbf{X}}^{(k+1,l)} = \mathbf{U}_l^{(k)} \mathcal{S}_{\tau_k}\left(\mathbf{\Sigma}_l^{(k)}\right) \mathbf{V}_l^{(k)H}, \tag{42}$$

Finally, the updated multi-frame image tensor is recovered by applying the inverse Fourier transform along the third mode:

$$\mathcal{X}^{(k+1)} = ifft(\overline{\mathcal{X}}^{(k+1)}, [], 3). \tag{43}$$

The t-SVT operation suppresses weak singular components while preserving the dominant correlated structures shared by consecutive THz video SAR frames. Consequently, the inter-frame redundancy is incorporated directly into each image tensor update without requiring a pre-specified tensor rank. After obtaining $\mathcal{X}^{(k+1)}$, the residual phase variables are updated with the image tensor fixed, as described in Section 5.3.

5.3. Closed-Form Pulsewise Residual Phase Update

After updating the image tensor, the residual phase factors are estimated with $\mathcal{X}^{(k+1)}$ fixed. Define the RCRC-aware predicted echo tensor before residual phase modulation as

$$\xi^{(k+1)} = \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}^{(k+1)}; \boldsymbol{\rho}_r). \tag{44}$$

Because the TNN regularizer is independent of the residual phase variables, the phase update subproblem reduces to

$$\begin{aligned} \phi^{(k+1)} = \quad & \underset{\phi}{\operatorname{argmin}} \tfrac{1}{2} \left\| \mathcal{P}_\Omega\left[\mathcal{B}(\phi) \odot \xi^{(k+1)}\right] - \mathcal{Y}_\Omega \right\|_F^2 \\ & s.t.\ |\phi_t(p)| = 1. \end{aligned} \tag{45}$$

Because the residual phase is modeled as pulsewise constant over the fast-time dimension, the optimization in (45) can be decomposed into independent scalar subproblems. Let

$$\Omega_{t,p} = \{p | (p, q, t) \in \Omega\} \tag{46}$$

denote the set of available fast-time samples associated with the p-th pulse of the t-th frame. The corresponding phase factor is obtained by solving

$$\phi_t^{(k+1)} = \arg\min_{|\phi|=1} \frac{1}{2} \sum_{q \in \Omega_{t,p}} \left| \phi E_{p,q,t}^{(k+1)} - Y_{p,q,t} \right|^2. \tag{47}$$

Expanding the squared error and omitting the terms independent of ϕ yields

$$\phi_t^{(k+1)}(p) = \arg\min_{|\phi|=1} \mathrm{Re}\left\{ \phi^* c_t^{(k+1)}(p) \right\}, \tag{48}$$

where

$$c_t^{(k+1)}(p) = \sum_{q \in \Omega_{t,p}} Y_{p,q,t} E_{p,q,t}^{(k+1)*} \tag{49}$$

is the complex correlation between the measured and predicted echoes for the corresponding pulse. To ensure numerical stability in practical implementations, degenerate cases where a pulse contains no retained observations or where the correlation coefficient $c_t(p)$ approaching zero are handled by a safeguarding rule: if the energy of the available fast-time samples for pulse p is below a small numerical tolerance, the phase update is bypassed, and the phase factor is retained from the previous iteration, thereby preventing division by zero or numerical divergence under extreme undersampling. Therefore, the unit modulus constrained problem admits the closed-form solution with a deterministic tie-breaking rule for the degenerate case

$$\phi_t^{(k+1)}(p) = \begin{cases} \frac{c_t^{(k+1)}(p)}{\left|c_t^{(k+1)}(p)\right|}, & \left|c_t^{(k+1)}(p)\right| > \varepsilon_c, \\ \phi_t^{(k)}(p), & \left|c_t^{(k+1)}(p)\right| \le \varepsilon_c, \end{cases} \tag{50}$$

where ε_c is a small threshold used to determine whether a value is numerically zero.

Thus, rather than estimating an independent phase factor for every slow-time/fast-time sample, all available fast-time measurements within the same pulse contribute jointly to a single residual phase estimate through coherent correlation. This pulsewise update substantially reduces the number of nuisance parameters and improves the robustness of phase estimation under noisy and slow-time sparse observation. The resulting phase estimate is subsequently used in the next image tensor update, completing one alternating reconstruction cycle.

Based on the image tensor update and the closed-form pulsewise phase estimation derived above, the proposed reconstruction is implemented using an alternating optimization strategy. At each outer iteration, the current residual phase estimate is first fixed, and the multi-frame image tensor is updated through a proximal gradient step followed by tensor singular value thresholding. The newly reconstructed image tensor is subsequently used to update the pulsewise residual phase factors in closed form. These two blocks are alternated until a prescribed stopping criterion is satisfied.

5.4. Overall Alternating Reconstruction Algorithm

In the absence of prior residual phase information, the pulsewise phase factors can be initialized as

$$\phi_t^{(0)}(p) = 1\,, \forall t, p, \tag{51}$$

corresponding to zero initial residual phases. An initial image tensor $\mathcal{X}^{(0)}$ can be obtained from the available measurements using the RCRC-aware adjoint reconstruction. The complete reconstruction procedure is summarized in Algorithm 1.

At the k-th iteration, the image tensor is first updated according to

$$\nabla f_k\left(\mathcal{X}^{(k)}\right) = \mathcal{H}_{\phi^{(k)}}^H\left[\mathcal{H}_{\phi^{(k)}}(\mathcal{X}^{(k)}) - \mathcal{Y}_\Omega\right], \tag{52}$$

and

$$\mathcal{X}^{(k+1)} = t\mathrm{SVT}_{\mu_k\lambda}\left(\mathcal{Z}^{(k)}\right). \tag{53}$$

Algorithm 1 RCRC-Aware Joint Tensor and Residual Phase Reconstruction

Input: slow-time sparse observation. $\mathcal{Y}_\Omega$, sampling set Ω, RCRC parameters $\boldsymbol{\rho}_r$, regularization parameter λ, step size $\{\mu_k\}$, tolerance ε, and maximum iteration number $K_{\max}$.

Initialization: initialize $\mathcal{X}^{(0)}$ and $\phi^{(0)}$; set $k = 0$.

1: *repeat*
2: Compute the data-fidelity gradient using (33).
3: Perform the gradient step using (35).
4: Apply the TNN proximal mapping using (36).
5: Predict the RCRC-aware echoes using (44).
6: *for each observed pulse* (t, p) *with* $\Omega_{t,p} \neq \varnothing$ *do.*
7: Compute the pulsewise correlation using (49).
8: Update the residual phase using (50).
9: *end for*
10: Set $k \leftarrow k + 1$.
11: until $\left\|\mathcal{X}^{(k+1)} - \mathcal{X}^{(k)}\right\|_F / \left\|\mathcal{X}^{(k)}\right\|_F < \varepsilon$ or $k \geq K_{\max}$.

Output: reconstructed multi-frame image tensor $\hat{\mathcal{X}} = \mathcal{X}^{(k)}$ and estimated pulsewise residual phase $\hat{\varphi} = \varphi^{(k)}$.

The updated image tensor is then projected through the RCRC-aware forward model, after which the pulsewise residual phase factors are updated as

$$\phi_t^{(k+1)}(p) = \frac{\sum\limits_{q \in \Omega_{t,p}} Y_{p,q,t} E_{p,q,t}^{(k+1)*}}{\left| \sum\limits_{q \in \Omega_{t,p}} Y_{p,q,t} E_{p,q,t}^{(k+1)*} \right|}, \tag{54}$$

for pulses containing available measurements.

The iterations are terminated when the relative change in the reconstructed image tensor satisfies

$$\frac{\left\|\mathcal{X}^{(k+1)} - \mathcal{X}^{(k)}\right\|_F}{\left\|\mathcal{X}^{(k)}\right\|_F} < \varepsilon, \tag{55}$$

or when the maximum number of iterations $K_{\max}$ is reached.

For fixed $\phi^{(k)}$, the gradient of the quadratic data fidelity term is Lipschitz continuous with constant

$$L_k = \left\| \mathcal{H}_{\phi^{(k)}}^H \mathcal{H}_{\phi^{(k)}} \right\|_2. \tag{56}$$

Therefore, with an appropriate step size, e.g., $0 < \mu_k < 1/L_k$, the proximal gradient image update provides a descent step for the fixed-phase subproblem. In addition, the pulsewise phase update in (54) is the exact minimizer of the corresponding unit modulus scalar subproblem.

5.5. Descent Property and Convergence Discussion

To characterize the numerical behavior of the proposed alternating reconstruction, consider the objective function

$$\begin{aligned} F(\mathcal{X}, \phi) \quad &= \tfrac{1}{2} \| \mathcal{P}_\Omega[\mathcal{B}(\phi) \odot \mathcal{A}_{\mathrm{RCRC}}(\mathcal{X}; \boldsymbol{\rho}_r)] - \mathcal{Y}_\Omega \|_F^2 \\ &\quad + \lambda \| \mathcal{X} \|_{TNN}, \end{aligned} \tag{57}$$

subject to $|\phi_t(p)| = 1$.

The proximal gradient image update satisfies

$$F\left(\mathcal{X}^{(k+1)}, \phi^{(k+1)}\right) \leq F\left(\mathcal{X}^{(k)}, \phi^{(k)}\right) - \frac{1-\mu_k L_k}{2\mu_k}\left\|\mathcal{X}^{(k+1)} - \mathcal{X}^{(k)}\right\|_F^2. \tag{58}$$

Hence, the image tensor update does not increase the joint objective.

For fixed $\mathcal{X}^{(k+1)}$, the pulsewise residual phase update derived in Section 5.3 gives the exact minimizer of each unit modulus scalar phase subproblem. Consequently,

$$F\left(\mathcal{X}^{(k+1)}, \phi^{(k+1)}\right) \leq F\left(\mathcal{X}^{(k+1)}, \phi^{(k)}\right). \tag{59}$$

Combining (58) and (59) yields

$$F\left(\mathcal{X}^{(k+1)}, \phi^{(k+1)}\right) \leq F\left(\mathcal{X}^{(k)}, \phi^{(k)}\right) \tag{60}$$

for each alternating iteration. Because the objective function is nonnegative, the resulting objective value sequence is monotonically nonincreasing and bounded below and therefore converges to a finite limit.

Because the joint recovery problem remains nonconvex owing to the multiplicative coupling between the image tensor and the residual phase variables, the above descent property does not imply global optimality. Nevertheless, the monotonic descent property provides a stable numerical basis for the proposed alternating reconstruction.

6. Experiments

To comprehensively validate the proposed RCRC-guided low-tubal-rank tensor recovery framework with blind residual phase correction, this section presents a controlled point target simulation followed by experiments using measured 220 GHz airborne THz-SAR data. Section 6.1 introduces the experimental system parameters, the slow-time sparse observation protocol, the six comparison and ablation configurations, and the implementation details of large-scale tensor block processing. Section 6.2 uses a controlled point target array simulation to evaluate the azimuth-focusing performance, residual phase estimation accuracy, third-mode low-tubal-rank prior, and iterative convergence of the proposed method. Section 6.3 presents full-scene reconstruction results of the measured THz video SAR data under different slow-time observation ratios, together with enlarged regions of interest (ROIs) and azimuth profiles of representative strong scatterers. Section 6.4 provides an in-depth discussion based on sampling ratio evolution curves (20–70%), full-sequence frame-wise stability curves (Frames 1–98), and quantitative evaluation tables.

6.1. Experimental Setup and Comparison Protocol

The proposed method is evaluated using measured THz-SAR data acquired by an airborne THz-SAR system together with simulated data configured using the same parameters. The radar operates at a carrier frequency of 220 GHz with a bandwidth of 900 MHz. The main system and trajectory parameters are summarized in Table 1. The full aperture contains $P = 100,000$ slow-time pulses and 8192 fast-time range samples. By partitioning the aperture into sliding sub-apertures with a length of $S = 3000$ pulses and a sliding step of $s = 1000$ pulses, corresponding to an aperture overlap ratio of $\varpi_{ov} \approx 67\%$, a continuous video SAR sequence consisting of $T = 98$ frames is constructed. The complete echo data for each frame have dimensions of $p \times q = 3000 \times 8192$, and the reconstructed 3-D video SAR image tensor has dimensions of $8192 \times 8192 \times 98$.

Table 1. THz SAR measured parameters.

Parameter	Value
Carrier Frequency	220 GHz
Bandwidth	900 MHz
PRF	10 kHz
T_p	12.5 μs
F_s	480 MHz
Reference slant range	5263 m
Initial radius	5042 m
Synthetic aperture time	0.3 s
Flight altitude	≈1000 m
Flight velocity	≈70 m/s
Mean RCRC	≈6.1 m/s
Number of range samples P	8192
Number of slow-time samples P	100,000
Sub-aperture length S	3000
Step size s	1000
Overlap ratio ϖ_{ov}	≈67%
Number of video frames T	98
Slow-time observation ratios ϖ_{ob}	60%, 40%, 30%

To provide a controlled comparison with the fully sampled reference and evaluate the algorithm's robustness, we perform retrospective sparse sampling on the fully acquired echo data. Specifically, the slow-time sampling mask Ω is generated and applied to the master slow-time dimension of the full-aperture echo before the overlapping sliding windows are constructed. Thus, the missing pulses are synchronized across all video frames, avoiding any frame-independent redundancy artifacts. The full echo is first acquired at the original PRF, and only the pulses indexed by the sampling set Ω are used for reconstruction. Therefore, the retrospective experiment itself does not reduce the actual PRF, transmitted pulses, or collected raw data volume. Physical pulse skipping is a possible future implementation of the prescribed observation pattern. At a fixed number of range samples per pulse, retaining a fraction of the pulses would proportionally reduce the number of uniquely acquired samples. This sample–count relationship is not an assessment of encoded storage or downlink volume, which also depends on quantization, coding, and metadata. Hardware implementation and comparisons with raw data compression are not evaluated in this study.

To compare the integrated method with existing tensor-based video SAR reconstruction baselines and to isolate the roles of trajectory modeling and residual phase estimation within a common TNN-regularized framework, six representative reconstruction configurations were evaluated. (1) Proposed: the proposed RCRC-guided low-tubal-rank tensor recovery method with blind residual phase correction, which combines geometry-aware physical modeling with multi-frame structural regularization. (2) Ablation experiment 1 (AE1): constant curvature TNN recovery with residual phase estimation, which replaces the RCRC-aware imaging operator with a constant curvature operator while retaining the same TNN regularization and residual phase estimation procedure as the proposed method. It is used to isolate the contribution of RCRC modeling. (3) AE2: RCRC-guided

TNN recovery without residual phase estimation, which retains the RCRC-aware imaging operator and TNN regularization but disables pulsewise residual phase estimation by fixing all phase factors to unity. It quantifies the contribution of the residual phase correction module. (4) Comparison method 1 (CM1): this method [27] models the multiframe video SAR images as a low-rank tensor and recovers the target tensor using an alternating-minimization strategy. It exploits interframe redundancy but does not explicitly account for the RCRC-dependent imaging geometry or pulsewise residual phase errors. (5) CM2: this method [28] represents the video SAR image sequence as the sum of a low-rank tensor and a sparse tensor, which characterize the temporally correlated scene component and sparse prominent scatterers, respectively. The resulting joint recovery problem is solved using a tensor alternating direction method of multiplier (ADMM) scheme. (6) CM3: this method [29] reconstructs the video SAR tensor by iteratively estimating orthogonal rank-one basis tensors and their associated weight tensors. The observation residual is used as the stopping criterion, eliminating the need for a prespecified tubal rank value.

For the proposed method, the regularization parameter λ is empirically set to 0.01 and fixed for all experiments. The proximal gradient step size is set to $\mu = 0.9/L$. The relative change tolerance is set to $\varepsilon = 10^{-4}$, and the maximum number of outer iterations is $K_{\max} = 50$. The iteration is terminated once either the relative change criterion is satisfied or the maximum iteration number is reached.

In practical implementation, to prevent excessive memory consumption when handling large-scale image dimensions (8192 × 8192 per frame), the reconstruction is executed via a sliding temporal window of $T_{sub} = 14$ consecutive frames combined with spatial sub-block partitioning (using 512 × 512 blocks with boundary overlap). Furthermore, in the t-SVT step, the frontal slices in the Fourier domain are decoupled from one another. Truncated SVD is employed to retain only the dominant singular components, with GPU-based parallel acceleration. These partitioning and parallelization strategies significantly reduce the peak memory consumption from 45 GB to 6.8 GB. All experiments are conducted on a workstation equipped with an Intel Core i9-14900KF CPU, 128 GB RAM, and Nvidia GeForce RTX 4080 Super GPU. All reconstructions in this study are performed offline using recorded echo data. The sliding temporal window limits the size of the active reconstruction problem and its memory requirements; it does not establish bounded end-to-end latency under continuous acquisition. The temporal sampling rate of the video sequence should therefore be distinguished from the computational throughput of the reconstruction algorithm.

6.2. Controlled Point Target Simulation and Algorithmic Mechanism Validation

To evaluate the algorithm's ability to suppress geometric mismatch and phase disturbances under controlled conditions with accurate ground truth, a 3 × 3 point target array was first placed within the imaging scene, with an adjacent target spacing of 50 m. The reconstruction tensor dimensions were set to 512 × 512 × 98, with all remaining parameters kept consistent with those of the measured data.

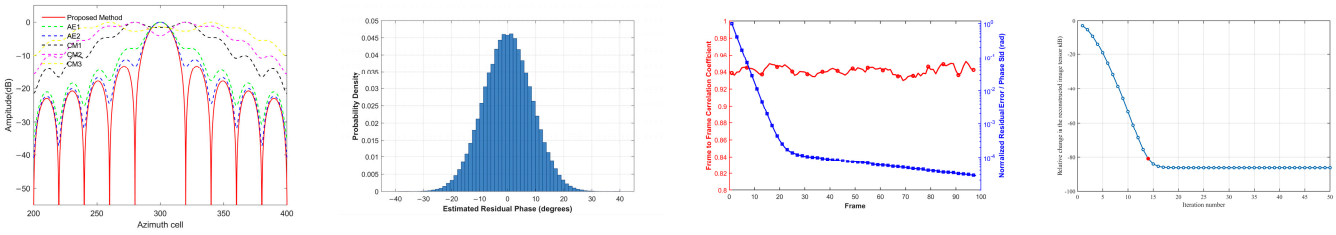
Figure 2 presents the two-dimensional reconstruction results of the central point target obtained through the six methods under a 40% slow-time observation ratio. The corresponding one-dimensional azimuthal profile cuts are shown in Figure 3a. As observed, the conventional tensor-based baseline methods CM1, CM2, and CM3, which lack the RCRC variable curvature geometric operator and phase compensation, exhibit severe azimuthal mainlobe broadening, peak splitting, and degraded symmetry under 40% undersampling. Among the ablation variants, AE1, which does not use RCRC, introduces deterministic quadratic and higher-order phase mismatches due to the constant curvature approximation.

Consequently, the point target mainlobe becomes noticeably broadened, accompanied by an elevated first sidelobe. Although AE2, which excludes residual phase correction, eliminates the macroscopic geometric drift and restores the cross-shaped orthogonal profile through the RCRC operator, the uncompensated pulse-level phase fluctuations still cause sidelobe energy leakage, with the first sidelobe reaching approximately -11.2 dB. In contrast, the complete proposed method achieves the sharpest two-dimensional impulse response and the narrowest mainlobe, and the first azimuthal sidelobe is effectively suppressed below -13.26 dB.



(**a**) proposed (**b**) AE1 (**c**) AE2 (**d**) CM1 (**e**) CM2 (**f**) CM3

Figure 2. Simulation results for point target reconstruction under the slow-time observation ratios of 40%.



(**a**) (**b**) (**c**) (**d**)

Figure 3. (**a**) Simulation azimuth slices of point targets. (**b**) Histograms of the estimated residual phases. (**c**) The singular value spectrum of the recovered tensor along the third mode and frame to frame correlation coefficients. (**d**) Iteration number.

Because the pulse-level phase updates rely on the coherent correlation between the observed echoes and the echoes predicted by the RCRC physical model, global phase drift and spatial translation ambiguity are effectively eliminated. The statistical histogram shows that the estimated residual phases follow a compact, zero-mean, symmetric Gaussian-like distribution, with the vast majority concentrated within the interval $[-15°, +15°]$. To rigorously verify whether the pulsewise estimator recovers the true perturbation rather than merely fitting noise, we evaluated its performance in a controlled simulation where the ground truth phase perturbation was explicitly known. Specifically, a zero-mean Gaussian random slow-time residual phase with a standard deviation of $\sigma_\phi = 8.6°$ was injected into the echo. By comparing the estimated phase against the known ground truth perturbation, the root mean square error is reduced to 1.15°, achieving a high correlation coefficient greater than 0.98. This result demonstrates that the phase estimator accurately captures the unmodeled microscopic platform jitter rather than overfitting noise.

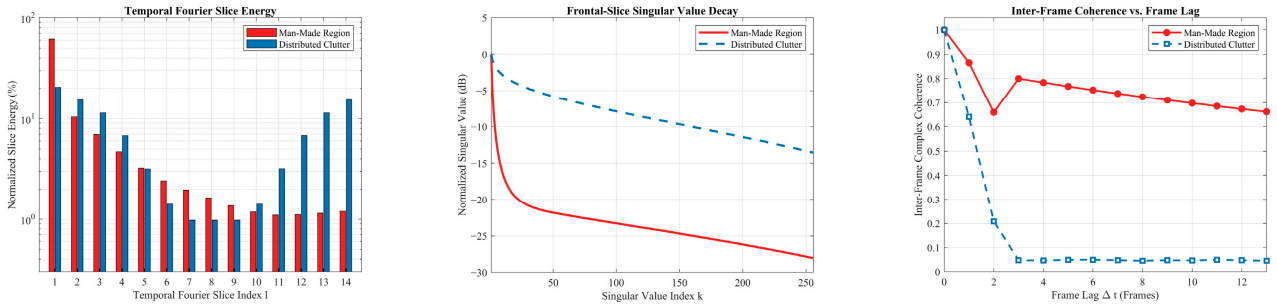
After RCRC-based geometric alignment and phase compensation, the normalized singular-value spectrum of the recovered tensor along the third mode exhibits a sharp exponential decay over the first 20 orders, spanning four orders of magnitude. This directly confirms the strong low-tubal-rank compressibility of the video sequence after geometric correction. Meanwhile, the inter-frame correlation coefficient remains stably high at approximately 0.94 throughout the full 98-frame sequence, demonstrating strong spatiotemporal coherence under the $\varpi_{ov} \approx 67\%$ aperture overlap ratio.

The relative change in the reconstructed image tensor decreases monotonically with the number of iterations. It falls below -80 dB at the 14th iteration and then converges stably to a floor of approximately -86 dB after the 18th iteration. This numerically verifies the stability of the monotonic descent property discussed in Section 5.5.

6.3. Comparison of Measured THz Video SAR Reconstructions Under Multiple Observation Ratios

In the experiment using measured airborne data at 220 GHz, the performance of the algorithms was further evaluated in complex distributed scene environments and on high-resolution structures with strong scattering characteristics.

Figure 4 contrasts a $512 \times 512 \times 14$ human-made region against a distributed farmland clutter region. As shown in Figure 4a,b, the frontal slice singular value spectrum of the human-made region exhibits a steep decay (dropping over -20.9 dB at rank $k = 30$), whereas the clutter region decays much more gradually, reflecting a nearly full-rank background. Furthermore, the inter-frame complex coherence in Figure 4c clearly demonstrates that while distributed clutter decorrelates rapidly to the noise floor $\approx$0.05 when $\Delta t \geq 3$ (beyond the pulse-sharing boundary), the human-made region maintains a stable high-coherence plateau due to continuous footprint illumination and structural angular persistence. This verifies that the TNN prior primarily captures the multi-frame redundancy of dominant human-made geometries while acting as a robust regularizer for background clutter.



(**a**) (**b**) (**c**)

Figure 4. (**a**) Temporal Fourier slice energy; (**b**) singular value spectrum of frequency domain slices; (**c**) inter-frame complex coherence versus frame lag of human-made region and distributed clutter.

Under the full observation condition, the effects of missing measurements are excluded. Figures 5–8 first presents the representative reference scenes of Frames 1, 31, and 81 under slow-time observation ratio $\varpi_{ob} = \{60\%, 40\%, 30\%\}$, which contain clear building edges, road networks, and farmland textures. When the slow-time observation ratio decreases to 60% and 40%, the full-scene reconstruction results and enlarged local ROI as shown in the Figure 9 views show that CM1, CM2, and CM3 suffer from severe azimuthal stripe blurring and defocused ghost artifacts. Building contours and road boundaries are overwhelmed by undersampling-induced aliasing noise. Because RCRC compensation is absent, AE1 exhibits evident inter-frame geometric drift and defocusing in sub-aperture frames far from the reference center. Although AE2 preserves the overall geometric structure, it has a higher background clutter floor and reduced detail contrast. In contrast, the proposed method preserves the building contours, road shadow contrast, and sharp focusing of strong point targets at both 60% and 40% observation ratios. When the observation ratio is further reduced to the extreme level of 30%, all methods experience noticeable degradation because of the extensive loss of support in the wav number domain. Nevertheless, the proposed method still retains the most identifiable contours of the main scene.

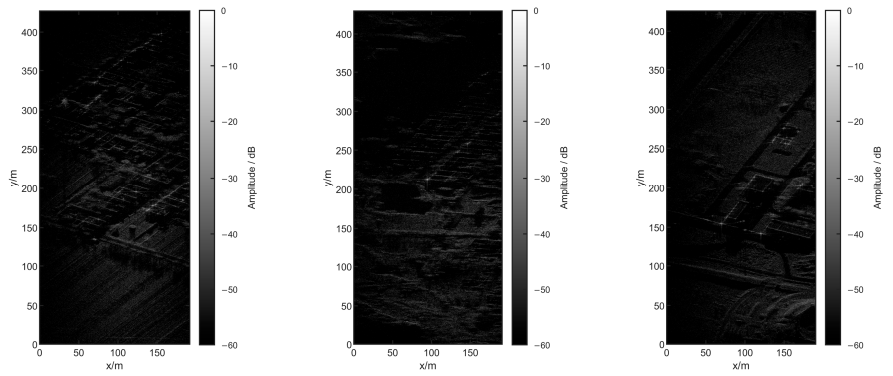


Figure 5. Representative reference scenes under full slow-time observation. (**a**) Frame 1. (**b**) Frame 31. (**c**) Frame 81.

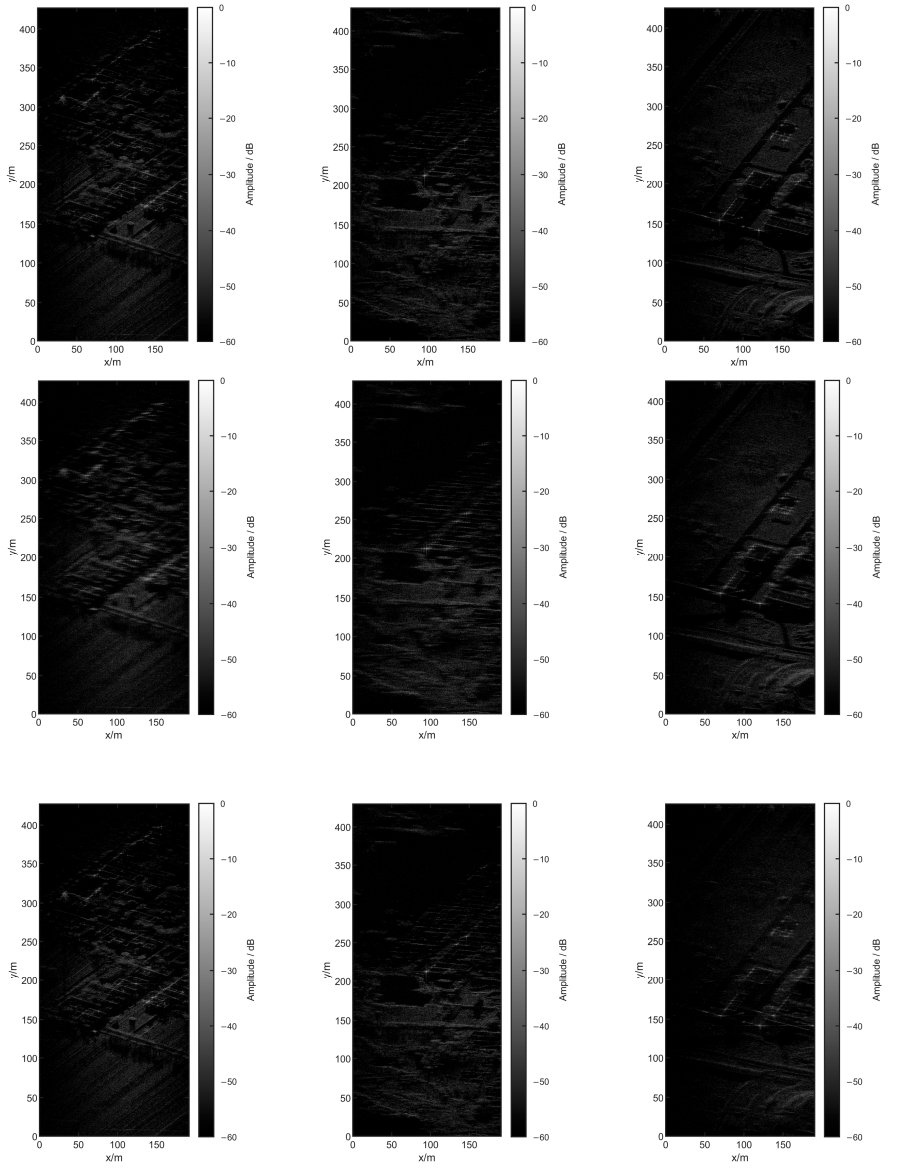


Figure 6. *Cont.*

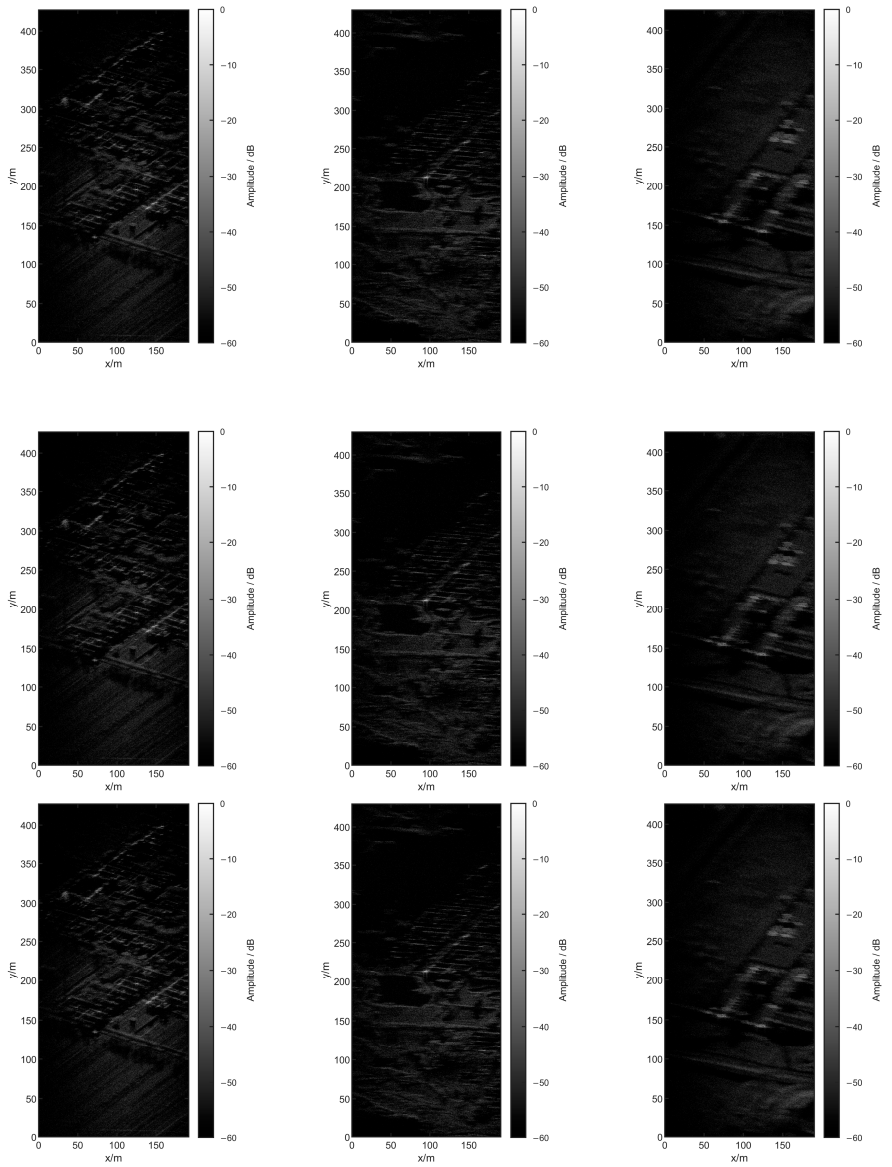


Figure 6. Representative reference scenes under 60% observation ratio. Columns 1–3 correspond to the 1st, 31st, and 81st frames, respectively. Rows 1–3 correspond to the proposed method, AE1, AE2, CM1, CM2, and CM3, respectively.

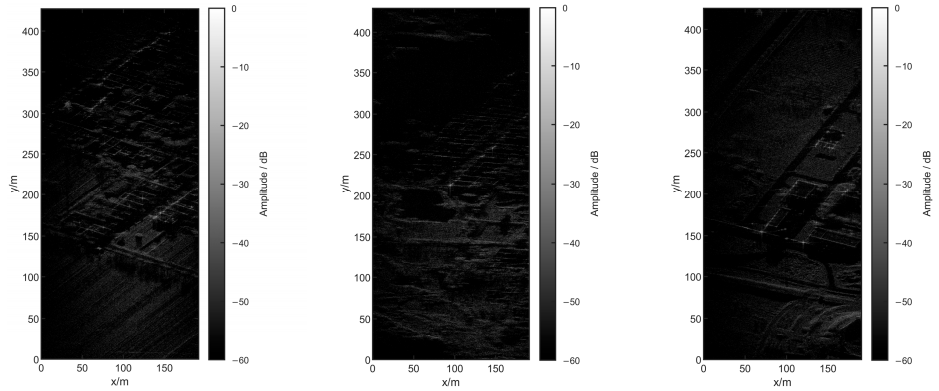


Figure 7. *Cont.*

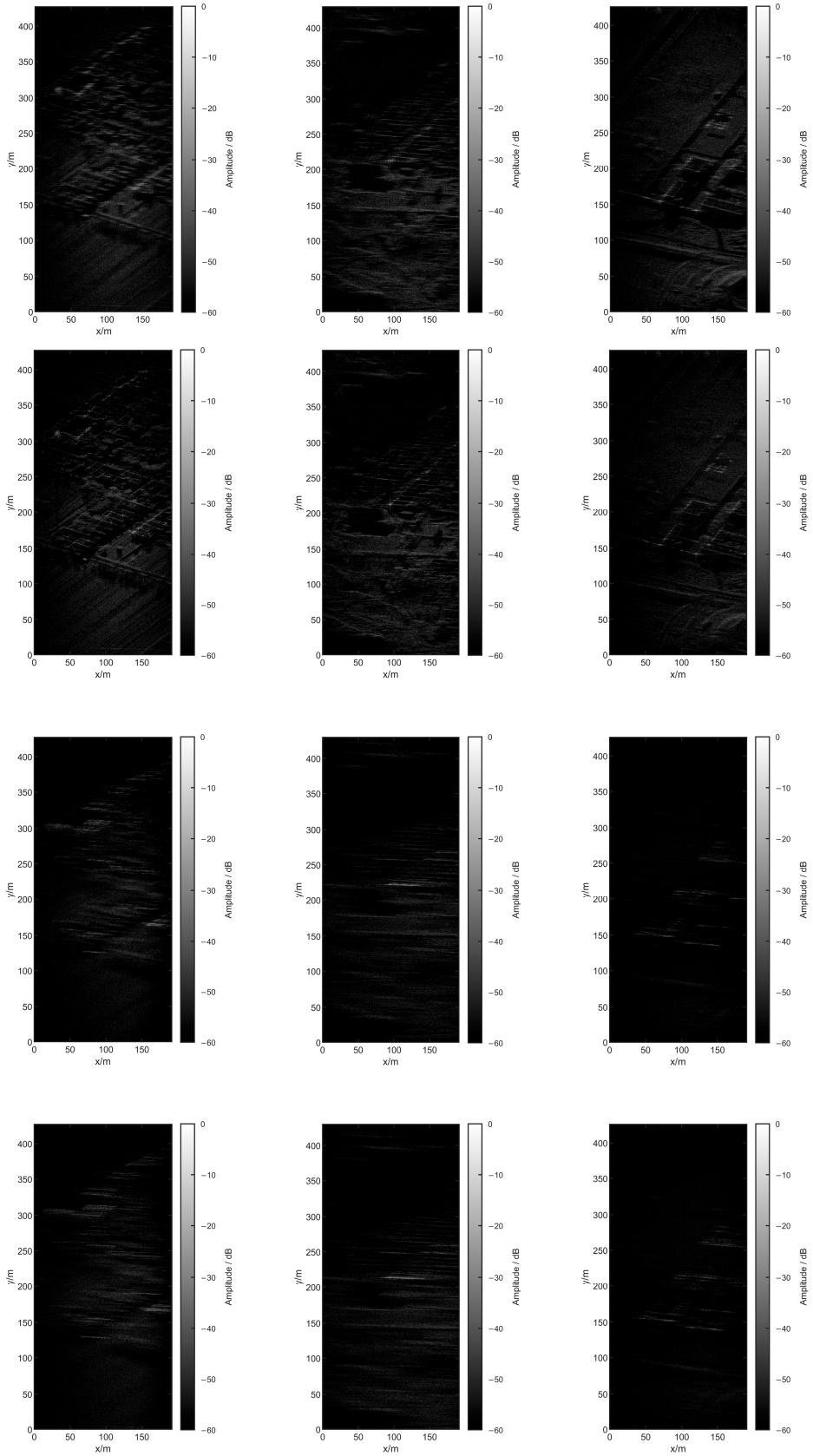


Figure 7. *Cont.*

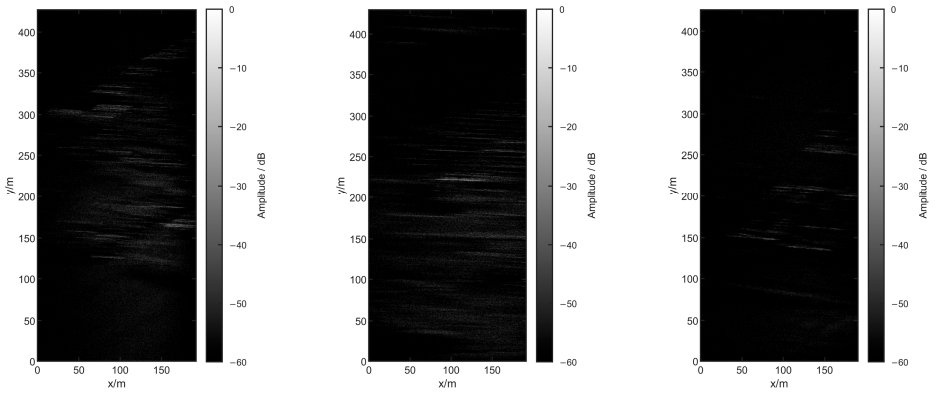


Figure 7. Representative reference scenes under 40% observation ratio. Columns 1–3 correspond to the 1st, 31st and, 81st frames, respectively. Rows 1–3 correspond to the proposed method, AE1, AE2, CM1, CM2, and CM3, respectively.

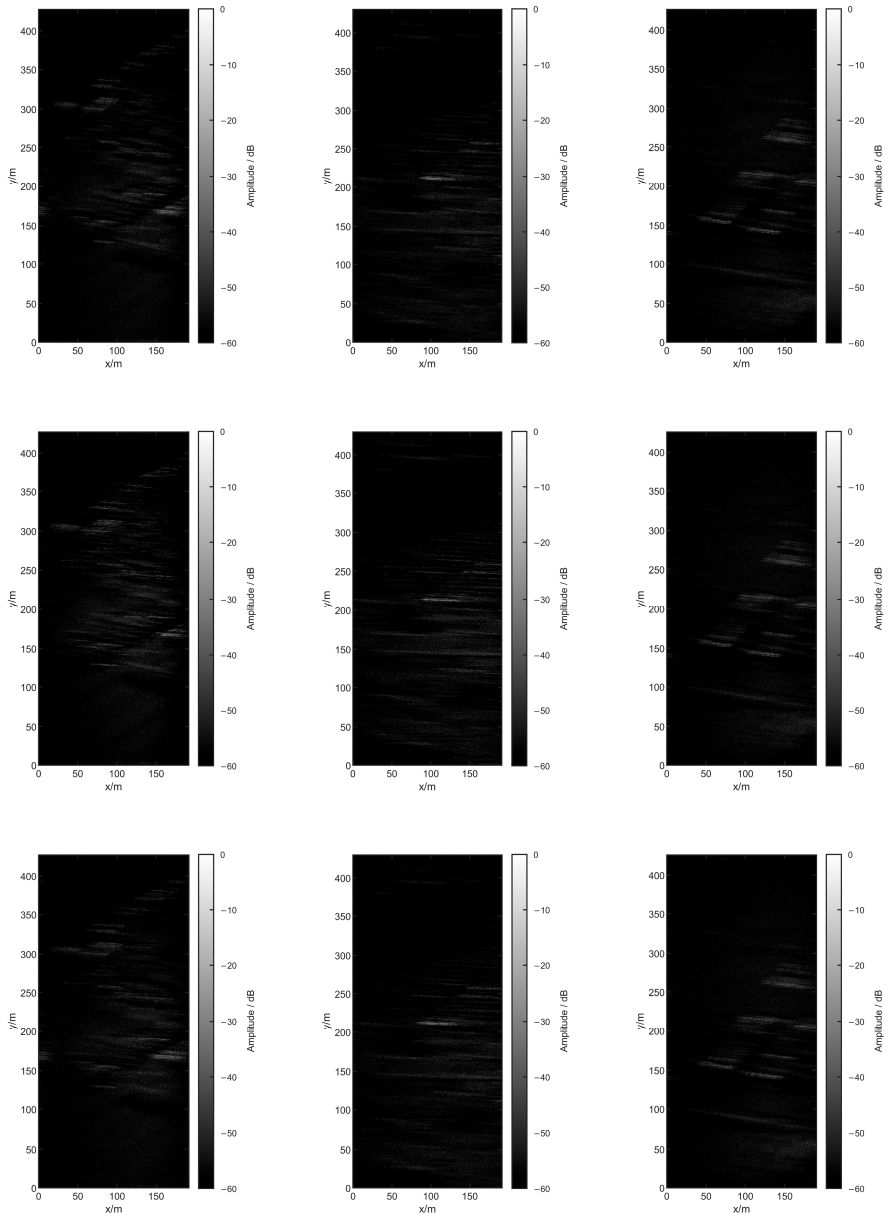


Figure 8. *Cont.*

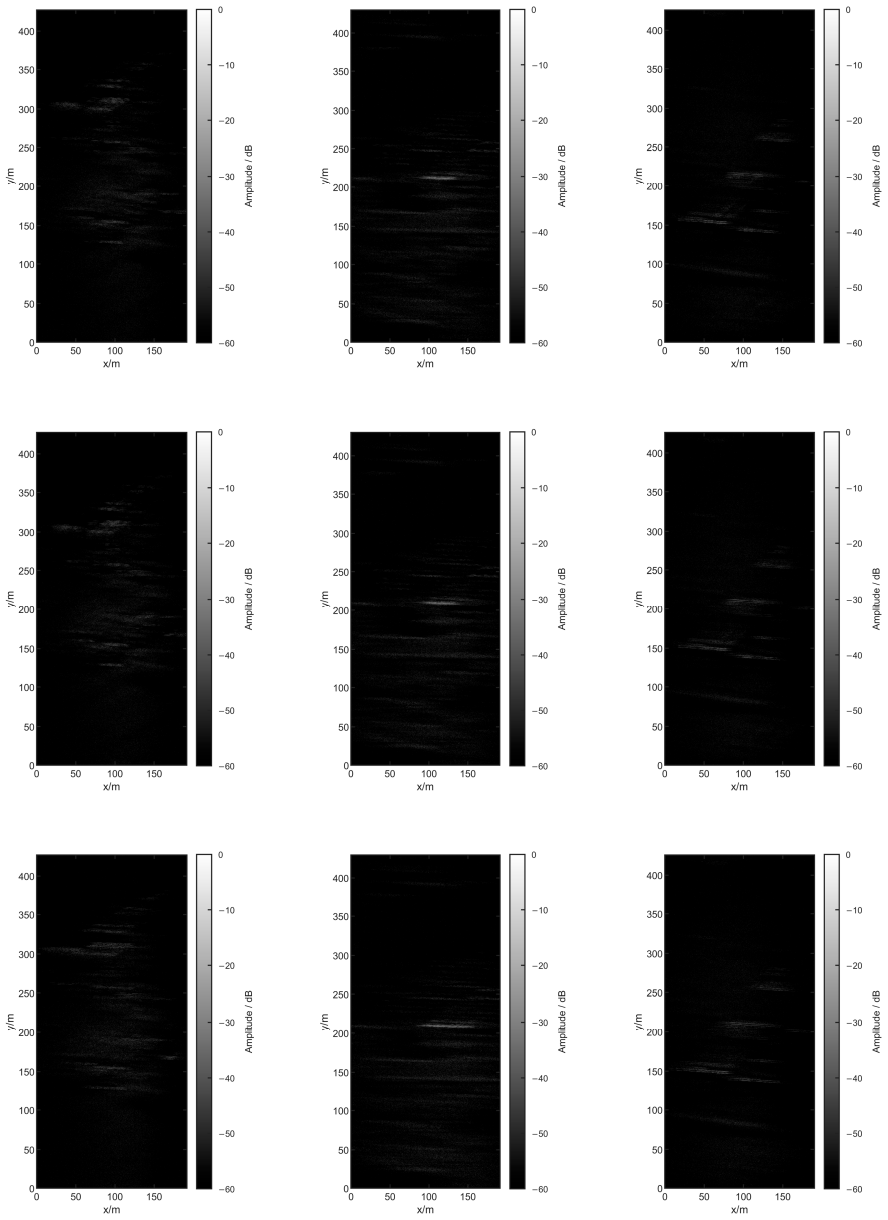


Figure 8. Representative reference scenes under 30% observation ratio. Columns 1–3 correspond to the 1st, 31st, and 81st frames, respectively. Rows 1–3 correspond to the proposed method, AE1, AE2, CM1, CM2, and CM3, respectively.

One-dimensional azimuthal profiles as shown in Figure 10 were extracted from three representative strong scatterers in the ROIs. Both the overall comparison among the six methods and the internal ablation comparison among Proposed, AE1, and AE2 clearly demonstrate that CM1–CM3 exhibit severely broadened mainlobes, with the sidelobes almost merging into the mainlobes. AE1 shows significant asymmetric mainlobe distortion and high sidelobe shoulders. Although AE2 achieves a narrower mainlobe, it still exhibits local sidelobe elevation above -5 dB in regions such as around sample index 6500. In contrast, the proposed method maintains symmetric and sharply defined mainlobes for all three measured characteristic points while consistently suppressing the near-in sidelobes to below approximately -12 dB.

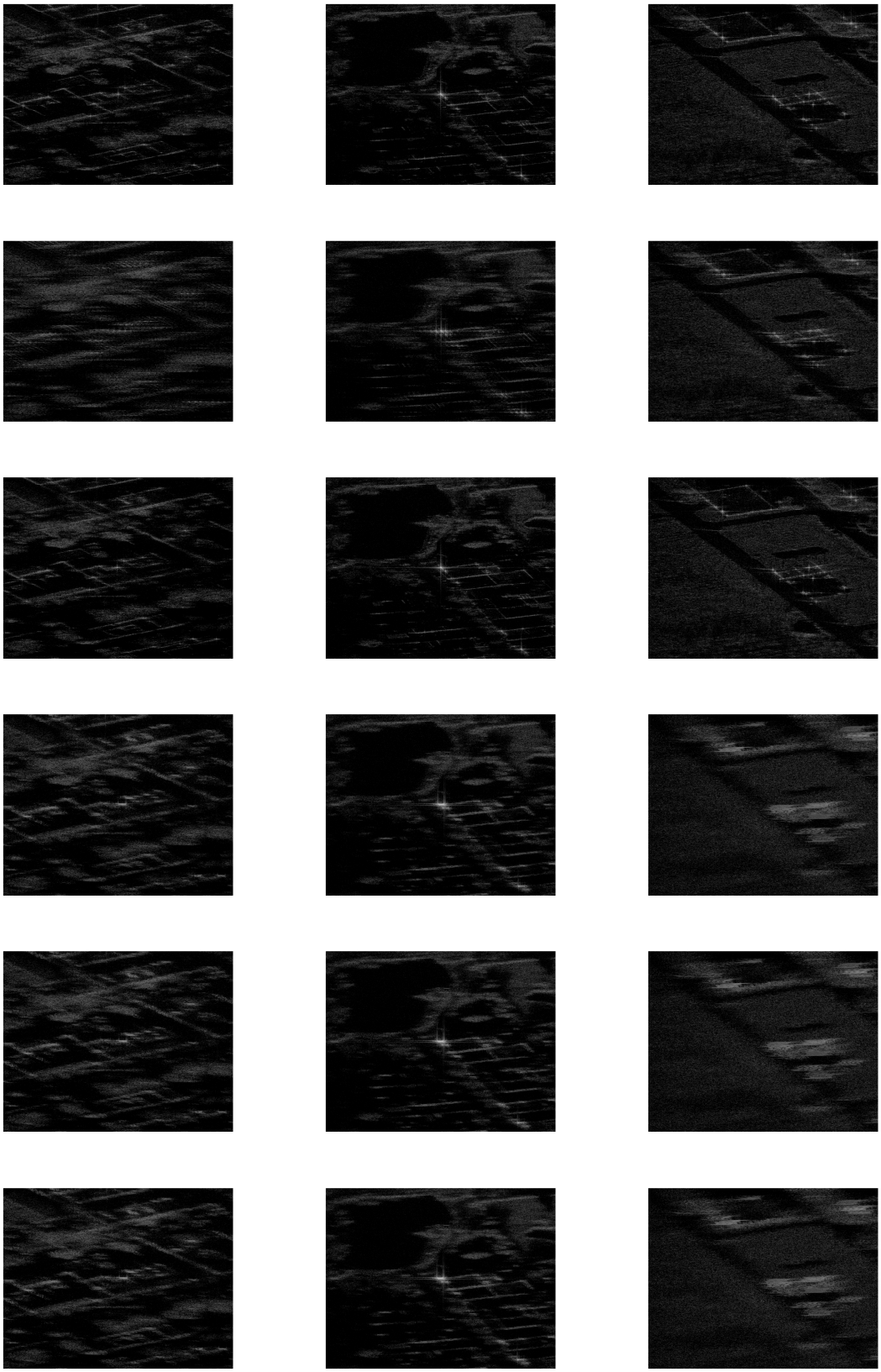


Figure 9. ROIs under 60% observation ratio. Columns 1–3 correspond to the 1st, 31st, and 81st frames, respectively. Rows 1–6 correspond to the proposed method, AE1, AE2, CM1, CM2, and CM3, respectively.

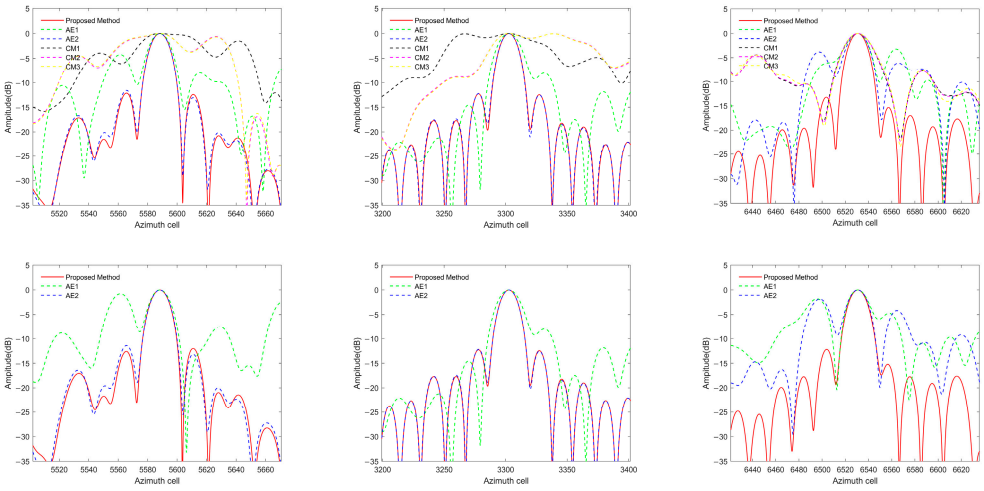


Figure 10. Azimuthal profile of the strong scatterer in the ROIs.

6.4. Quantitative Performance and Discussion

Because airborne measured THz-SAR data do not provide an absolute ground truth reflectivity map, the fully sampled RCRC-aware VCT reconstruction, which has been validated, is adopted as the pseudo-reference benchmark $\mathcal{X}_{ref}$ for quantitative evaluation. It should be objectively noted that because this reference is generated using the RCRC forward model, external baseline methods that are independent of the geometric model, namely CM1, CM2, and CM3, are inevitably penalized by deterministic geometric model mismatch when pixel-wise errors such as NRMSE and PSNR are computed. In addition to errors caused by incomplete observation recovery, these methods are therefore also affected by the geometric mismatch. Accordingly, NRMSE and PSNR are mainly used to measure the data consistency and focusing preservation of each method relative to the fully sampled physical reference. To eliminate model bias and provide a fairer evaluation, we jointly analyze the external baseline comparisons with the internal component ablation experiments involving AE1 and AE2.

The NRMSE and PSNR are used to evaluate reconstruction fidelity. The NRMSE is defined as

$$NRMSE = \sqrt{\frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\sum_{t=0}^{T-1}\left|\hat{\mathcal{X}}(m,n,t)-\mathcal{X}(m,n,t)\right|^2}{E}}, \tag{61}$$

where $\hat{\mathcal{X}}$ denotes the estimate of the video tensor $\mathcal{X}$ obtained by the video imaging algorithm and

$$E = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\sum_{t=0}^{T-1}|\mathcal{X}(m,n,t)|^2. \tag{62}$$

Meanwhile, PSNR is also employed as a performance metric and is defined as

$$PSNR = 20\log_{10}\left(\frac{E_{\max}(\mathcal{X}(m,n,t))}{NRMSE}\right). \tag{63}$$

The structure similarity (SSIM) is used to evaluate the estimation performance. SSIM is a perceptual metric that interprets image degradation as a perceptual change in structural information. It has been widely applied in video compression quality assessment. SSIM is defined as

$$SSIM = \frac{(2\mu_{x_1}\mu_{x_2}+c_1)(2\delta_{x_1x_2}+c_2)}{(\mu_{x_1}^2+\mu_{x_2}^2+c_1)(\delta_{x_1}^2+\delta_{x_2}^2+c_2)}. \tag{64}$$

where x_1 and x_2 denote vectors of the reconstructed tensor $\overline{\mathcal{X}}$ and the initial tensor $\mathcal{X}_0$, respectively; μ_{x_1} and μ_{x_2} respect the mean value of x_1 and x_2, respectively; $\delta_{x_1}^2$ and $\delta_{x_2}^2$ are the variance of x_1 and x_2, respectively; $c_1 = (a_1 L)^2$ and $c_2 = (a_2 L)^2$ denote two variables to stabilize the division with a weak denominator; L is the dynamic range of the entropy values; and $a_1 = 0.01$ and $a_2 = 0.03$ are chosen by default. The resultant SSIM index is a decimal value between -1 and 1, and value 1 is only reachable in the case in which the focus quality of the reconstructed video is the same as that of the original video.

The first row of Figure 11 shows the PSNR, NRMSE, and SSIM curves as the slow-time observation ratio varies from 20% to 70%. In the PSNR curves, all methods exhibit an increasingly steep upward trend as the observation ratio increases. The proposed method consistently achieves the highest PSNR across the entire sampling range, reaching approximately 24.9 dB at 40% observation ratio and 34.2 dB and 42.6 dB at 60% and 70%, respectively. It significantly outperforms the next-best methods, AE2 and AE1, while CM1 performs the worst. In the NRMSE curves, the proposed method exhibits the fastest decline, decreasing rapidly from 0.2587 at 20% observation ratio to 0.1342, 0.0460, and 0.0175 at 40%, 60%, and 70%, respectively. Under full-scene evaluation with global logarithmic normalization, when the observation ratio decreases to 40%, the inherent amplitude attenuation bias caused by the t-SVT singular-value soft-thresholding with a fixed parameter $\lambda = 0.01$, together with partial sidelobe energy leakage, increases the pixel-wise NRMSE to 0.1342, corresponding to a PSNR of 24.9 dB. Nevertheless, because the spatial geometry and mainlobe focusing of the dominant scattering structures are preserved, the SSIM metric, which is relatively insensitive to uniform amplitude shrinkage, remains as high as 0.85 at the 40% observation ratio knee point.

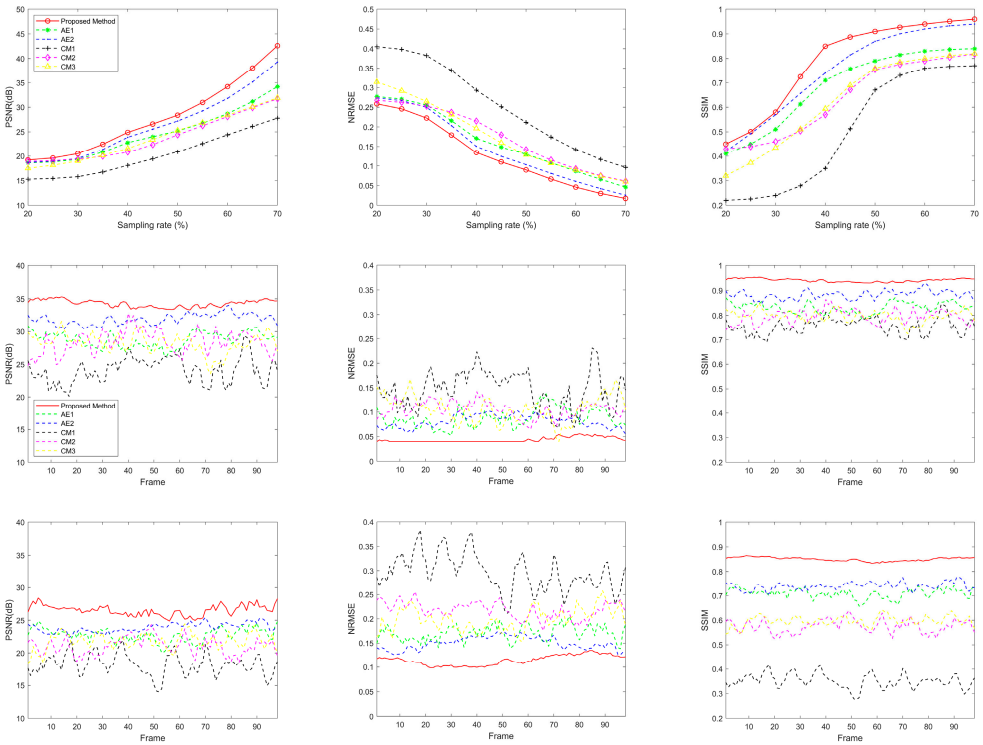


Figure 11. PSNR, NRSME, and SSIM performances comparison with other methods. The first row shows the performance metrics as functions of the slow-time observation ratio. The second and third rows show the performance metrics as functions of the frame index under 60% and 40% observation ratios, respectively.

In the SSIM curves, a characteristic pattern is observed, consisting of a steep rise at low observation ratios, a knee-point around 40%, and plateau saturation at higher observation

ratios. When the observation ratio is below 35%, undersampling aliasing disrupts the overall structural information, and the SSIM values of all methods remain below 0.75. At the critical 40% knee point, the proposed method is the first to reach an SSIM of 0.85, and it then stabilizes within a high plateau range of 0.91–0.96 over the 50–70% interval. In contrast, AE1 and CM1–CM3, which lack RCRC compensation, achieve SSIM upper bounds of only 0.77–0.84 even at 70% observation ratio.

The second and third rows of Figure 11 show the frame-wise PSNR, NRMSE, and SSIM curves for frames 1–98 at the 60% and 40% observation ratio. the proposed method maintains an extremely stable frame-wise PSNR within 34.0–35.2 dB, an NRMSE close to the 0.04–0.06 floor at the 60% observation ratio, and an SSIM consistently around 0.94–0.95. At 40%, its frame-wise PSNR remains stable within 25–28 dB, while the SSIM stays around 0.85 with minimal frame-to-frame variation. In contrast, AE1 and CM1–CM3, which lack RCRC compensation, achieve SSIM upper bounds of only 0.77–0.84 even at 70% observation ratio. This demonstrates that simply increasing the number of sampling pulses cannot eliminate deterministic geometric mismatch.

Benefiting from the physically consistent warm-start initialization, the algorithm converges within $12 \sim 16$ iterations in practice, typically 16 iterations at the 40% observation ratio, whereas geometry-mismatched baselines (CM1–CM3 and AE1) require 26–38 iterations to reach the stopping criterion. The computational efficiency and memory consumption of the proposed method and the comparison methods are shown in Table 2, the average runtime of the proposed method is approximately 64.2 s per frame 104.9min for the entire 98-frame sequence), which is only marginally higher than AE2 (61.5 s/$frame$) due to the lightweight $O(\Omega)$ closed-form phase update, and is notably faster than the matrix-unfolding baseline CM1 (112.5 s/$frame$) and the dual-tensor ADMM baseline CM2 (128.4 s/$frame$). With a sub-aperture step of 1000 pulses and a PRF of 10 kHz, adjacent frame centers are separated by 0.1 s, corresponding to a nominal frame rate of 10 frames/s. In contrast, the reported average reconstruction time of 64.2 s per output frame corresponds to an effective throughput of approximately 0.0156 frames/s. The processing time is therefore approximately 642 times the frame interval. The current implementation is offline and does not demonstrate real-time or near-real-time reconstruction. Sliding-window processing limits the active tensor size, but cannot prevent backlog growth when acquisition proceeds faster than reconstruction. The pulse-retention ratio specifies the fraction of available slow-time pulses used for reconstruction; it is not a compression ratio measured in encoded bits. No equal-data-volume comparison with BAQ or other raw-data codecs has been conducted. Accordingly, the present results support offline reconstruction from incomplete observations, but do not establish a rate–distortion advantage or a preferred storage/downlink strategy.

The 40% observation ratio can be regarded as a potential operating point for future sparse-pulse THz video SAR systems, provided that sufficient slow-time and wavenumber-domain coverage is retained. The 30% case is included mainly as a more severe stress test. The proposed RCRC prior assumes locally smooth trajectory variation and is not intended to handle arbitrary high-frequency platform jitter or abrupt maneuvering. If severe aliasing or quality-gated rejection removes large continuous regions of the wavenumber support, the remaining observations may be insufficient for accurate recovery.

Regarding the identifiability of the pulsewise residual phase, because the phase alignment relies on coherent correlation against spatially anchored predicted echoes generated by the RCRC model, global phase drift and spatial shift ambiguities are avoided. Statistical evaluation of the estimated residual phases shows that they remain tightly bounded within a zero-centered interval (predominantly $\pm\pi/6$), confirming that the estimator captures genuine unmodeled microscopic jitters rather than over-fitting the data.

Table 2. Comparison of computational efficiency and memory consumption.

Method	Forward Operator	Low-Rank Tensor Decomposition	Convergence Iterations	Peak Memory/GB	Average Runtime per Frame/s	Total Runtime for the Sequence/min
Proposed	RCRC-aware VCT operator	Parallel truncated t-SVT in the frequency domain	16	6.8	64.2	104.9
AE1	Constant-curvature wavenumber-domain operator	Parallel truncated t-SVT in the frequency domain	26	6.5	78.6	128.4
AE2	RCRC-aware VCT operator	Parallel truncated t-SVT in the frequency domain	18	6.8	61.5	100.5
CM1	Conventional constant-curvature geometry operator	Multimode matrix-unfolding SVD	30	14.2	112.5	183.8
CM2	Conventional constant-curvature geometry operator	Low-rank + sparse dual-tensor ADMM	36	15.6	128.4	209.7
CM3	Conventional constant-curvature geometry operator	Orthogonal rank-one tubal-basis iteration	38	8.5	74.2	121.2

The internal ablations examine the effects of RCRC-aware trajectory modeling and residual phase estimation while retaining TNN regularization. They therefore assess these components within the adopted tensor-recovery formulation. Because TNN is neither removed nor replaced in these ablations, the experiments do not isolate its incremental benefit over framewise regularization.

7. Conclusions

This article developed an integrated RCRC-guided low-tubal-rank tensor recovery framework for sparse-aperture THz video SAR imaging under variable curvature trajectories. Controlled point target simulations and measured data experiments indicate that RCRC physical modeling and residual phase estimation make distinct and complementary contributions to geometric focusing and phase error correction, whereas low-tubal-rank tensor regularization provides robust exploitation of interframe redundancy compared with existing literature baselines. The resulting joint problem is solved by alternating im-age-tensor and residual phase updates.

Controlled point target simulations and measured data experiments were con-ducted at pulse retention ratios of 60%, 40%, and 30%. A single slow-time mask was applied to the shared master pulse stream before overlapping sub-apertures were formed, ensuring that an omitted physical pulse was absent from every frame that referenced it. Comparisons with LRT-AM, JLRS-T, and ORTP, together with constant curvature and no residual phase ablations, indicate that RCRC modeling and residual phase estimation make distinct and complementary contributions to reconstruction quality. In particular, the 40% pulse retention case preserves the principal scene structures under the tested retrospective protocol, whereas the 30% case is more appropriately interpreted as a severe robustness test: the marked degradation observed at this ratio shows that the proposed method does not guarantee high-fidelity recovery when the retained slow-time support becomes insufficient.

Several limitations should be emphasized. The measured data were originally acquired at the full 10 kHz PRF and were sparsified retrospectively; consequently, the present experiments validate reconstruction from incomplete slow-time observations but do not by themselves demonstrate reductions in transmitted pulses, acquisition PRF, or collected raw data volume. A dedicated comparison against framewise regularization is required

to establish the incremental benefit of the low-tubal-rank prior. The present experiments evaluate the integrated method but do not establish the independent contribution of TNN. Moreover, the measured data metrics are calculated relative to a fully sampled RCRC-aware pseudo-reference rather than unknown ground truth reflectivity, and the first-order RCRC model assumes locally smooth trajectory variation within each sub-aperture. The current implementation is computationally intensive and intended for offline reconstruction. Although temporal windowing and spatial partitioning reduce the active memory requirements, the achieved throughput remains substantially below the acquisition frame rate. No comparison with conventional raw data compression has been performed; therefore, no superiority in storage efficiency, downlink efficiency, or rate distortion performance is claimed. Future work will therefore investigate hardware-level pulse skipping and reduced or staggered PRF acquisition, more complex trajectory and residual phase models, adaptive sampling and parameter selection, and computationally efficient blockwise implementations for large-scale THz video SAR reconstruction.

Author Contributions: Methodology, X.W., J.Z., and J.L. (Jin Li).; Software, X.W.; Validation, X.W. and T.H.; Formal analysis, X.W. and J.L. (Jialu Liu); Data curation, X.W. and J.Z.; Writing—original draft, X.W.; Writing—review & editing, J.L. (Jin Li) and Z.C.; Visualization, X.W. and J.Z.; Supervision, R.M. and Y.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are not publicly available due to restrictions on data sharing. The data may be available from the corresponding author upon reasonable request, subject to the applicable data-use restrictions.

Acknowledgments: The authors would like to thank all individuals who provided technical support and assistance with the THz-SAR experiments.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

THz-SAR	Terahertz synthetic aperture radar
RCRC	Radius-of-curvature change rate
VCT	Variable-curvature trajectory
NRMSE	Normalized root mean square error
PSNR	Peak signal-to-noise ratio
t-SVD	Tensor singular value decomposition
t-SVT	Tensor singular value thresholding
TNN	Tensor nuclear norm
ADMM	Alternating direction method of multiplier
AE	Ablation experiment
CM	Comparison method
ROI	Region of interest
SSIM	Structure similarity

References

1. Fan, L.; Wang, H.; Yang, Q.; Wang, Y.; Deng, B.; Xiao, H. High Frame-Rate and Low-Latency Video SAR Based on Robust Doppler Parameters Estimation in the Terahertz Regime. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5207016. [CrossRef]
2. Fan, L.; Wang, H.; Yang, Q.; Deng, B. High-Quality Airborne Terahertz Video SAR Imaging Based on Echo-Driven Robust Motion Compensation. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 2001817. [CrossRef]
3. Tang, B.; Wang, H.; Yang, Q.; Wang, R.; Yuan, Z.; Yi, J.; Deng, B. High-Order Rotational-Migration Correction and ISAR Imaging in the THz Band Considering Impact of Scattering Intensity. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5109611. [CrossRef]

4. Yuan, Z.; Chen, X.; Deng, B.; Yi, J.; Tang, B.; Wan, J.; Wang, H.; Wang, R. Terahertz ISAR Imaging With Nonrigid Vibration Compensation and Sidelobe Suppression of Space Targets. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5110516. [CrossRef]
5. Fan, L.; Wang, H.; Yang, Q.; Deng, B. THz-ViSAR-Oriented Fast Indication and Imaging of Rotating Targets Based on Nonparametric Method. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5217515. [CrossRef]
6. Tu, C.; Dong, Z.; Chen, X.; Hou, Z.; Ji, Y.; Li, W.; Dong, J. A Modified 2-D Autofocus Approach for 2-D Frequency-Dependent Residual Motion Error in Ultrahigh-Resolution Airborne SAR. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5221516. [CrossRef]
7. Frey, O.; Magnard, C.; Ruegg, M.; Meier, E. Focusing of Airborne Synthetic Aperture Radar Data From Highly Nonlinear Flight Tracks. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1844–1858. [CrossRef]
8. Chen, J.; Xing, M.; Xia, X.-G.; Zhang, J.; Liang, B.; Yang, D.-G. SVD-Based Ambiguity Function Analysis for Nonlinear Trajectory SAR. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 3072–3087. [CrossRef]
9. Mao, X.; He, X.; Li, D. Knowledge-Aided 2-D Autofocus for Spotlight SAR Range Migration Algorithm Imagery. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5458–5470. [CrossRef]
10. Mao, X.; Ding, L.; Zhang, Y.; Zhan, R.; Li, S. Knowledge-Aided 2-D Autofocus for Spotlight SAR Filtered Backprojection Imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9041–9058. [CrossRef]
11. Wang, X.; Hu, T.; Zhang, J.; Liu, J.; Li, J.; Min, R.; Cao, Z.; Pi, Y. A Wavenumber-Domain Imaging Algorithm for Terahertz Synthetic Aperture Radar With Variable-Curvature Trajectory. *IEEE Trans. Geosci. Remote Sens.* **2026**, *64*, 5208719. [CrossRef]
12. Huang, X.; Xu, Z.; Ding, J. Video SAR Image Despeckling by Unsupervised Learning. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10151–10160. [CrossRef]
13. Zhong, C.; Ding, J.; Zhang, Y. Joint Tracking of Moving Target in Single-Channel Video SAR. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5212718. [CrossRef]
14. Zuo, F.; Li, J.; Hu, R.; Pi, Y. Unified Coordinate System Algorithm for Terahertz Video-SAR Image Formation. *IEEE Trans. Terahertz Sci. Technol.* **2018**, *8*, 725–735. [CrossRef]
15. Song, X.; Yu, W. Processing Video-SAR Data With the Fast Backprojection Method. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2838–2848. [CrossRef]
16. Cheng, Y.; Ding, J.; Sun, Z.; Zhong, C. Processing of Airborne Video SAR Data Using the Modified Back Projection Algorithm. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5238013. [CrossRef]
17. Potter, L.C.; Ertin, E.; Parker, J.T.; Cetin, M. Sparsity and Compressed Sensing in Radar Imaging. *Proc. IEEE* **2010**, *98*, 1006–1020. [CrossRef]
18. Patel, V.M.; Easley, G.R.; Healy, D.M., Jr.; Chellappa, R. Compressed Synthetic Aperture Radar. *IEEE J. Sel. Top. Signal Process.* **2010**, *4*, 244–254. [CrossRef]
19. Alonso, M.T.; Lopez-Dekker, P.; Mallorqui, J.J. A Novel Strategy for Radar Imaging Based on Compressive Sensing. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 4285–4295. [CrossRef]
20. Cetin, M.; Stojanovic, I.; Onhon, O.; Varshney, K.; Samadi, S.; Karl, W.C.; Willsky, A.S. Sparsity-Driven Synthetic Aperture Radar Imaging: Reconstruction, Autofocusing, Moving Targets, and Compressed Sensing. *IEEE Signal Process. Mag.* **2014**, *31*, 27–40. [CrossRef]
21. Onhon, N.Ö.; Cetin, M. A Sparsity-Driven Approach for Joint SAR Imaging and Phase Error Correction. *IEEE Trans. Image Process.* **2012**, *21*, 2075–2088. [CrossRef]
22. Qiu, W.; Zhou, J.; Fu, Q. Jointly Using Low-Rank and Sparsity Priors for Sparse Inverse Synthetic Aperture Radar Imaging. *IEEE Trans. Image Process.* **2020**, *29*, 100–115. [CrossRef]
23. Zeng, C.; Zhu, W.; Jia, X.; Yang, L. Sparse Aperture ISAR Imaging Method Based on Joint Constraints of Sparsity and Low Rank. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 168–181. [CrossRef]
24. Zhang, C.; Zhang, S.; Liu, Y.; Li, X. Joint Structured Sparsity and Least Entropy Constrained Sparse Aperture Radar Imaging and Autofocusing. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6580–6593. [CrossRef]
25. Zhang, S.; Liu, Y.; Li, X. Fast Sparse Aperture ISAR Autofocusing and Imaging via ADMM Based Sparse Bayesian Learning. *IEEE Trans. Image Process.* **2020**, *29*, 3213–3226. [CrossRef]
26. Wang, Y.; Dai, F.; Liu, Q.; Hong, L.; Lu, X. Sparse Aperture Autofocusing and Imaging Based on Fast Sparse Bayesian Learning From Gapped Data. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5101716. [CrossRef]
27. Pu, W.; Wang, X.; Wu, J.; Huang, Y.; Yang, J. Video SAR Imaging Based on Low-Rank Tensor Recovery. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 188–202. [CrossRef]
28. An, H.; Wu, J.; Teh, K.C.; Sun, Z.; Li, Z.; Yang, J. Joint Low-Rank and Sparse Tensors Recovery for Video Synthetic Aperture Radar Imaging. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5214913. [CrossRef]
29. Pu, W.; Wu, J.; Huang, Y.; Yang, J. ORTP: A Video SAR Imaging Algorithm Based on Low-Tubal-Rank Tensor Recovery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 1293–1308. [CrossRef]
30. Zhang, Z.; Aeron, S. Exact Tensor Completion Using t-SVD. *IEEE Trans. Signal Process.* **2017**, *65*, 1511–1526. [CrossRef]

31. Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; Yan, S. Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 925–938. [CrossRef]
32. Zhou, P.; Lu, C.; Feng, J.; Lin, Z.; Yan, S. Tensor Low-Rank Representation for Data Recovery and Clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 1718–1732. [CrossRef]
33. An, H.; Jiang, R.; Wu, J.; Teh, K.C.; Sun, Z.; Li, Z.; Yang, J. LRSR-ADMM-Net: A Joint Low-Rank and Sparse Recovery Network for SAR Imaging. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5232514. [CrossRef]
34. Pu, W. Deep SAR Imaging and Motion Compensation. *IEEE Trans. Image Process.* **2021**, *30*, 2232–2247. [CrossRef]
35. Pu, W. SAE-Net: A Deep Neural Network for SAR Autofocus. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5220714. [CrossRef]
36. Li, M.; Wu, Z.; Huo, W.; Wu, J.; Zhang, J. Video SAR Image Reconstruction Based on Sparse Tensor Recovery and Unfolded Transformer. *IEEE Trans. Circuits Syst. Video Technol.* **2026**, *36*, 4678–4692. [CrossRef]
37. Li, M.; Jiang, L.; Liu, X.; Huo, W.; Wu, J.; Zhang, J. Unsupervised VideoSAR Sparse Imaging Using Physically Interpretable Video Tensor Decomposition-Based Deep Prior. *IEEE Trans. Geosci. Remote Sens.* **2026**, *64*, 5205018. [CrossRef]
38. Li, M.; Wu, J.; Huo, W.; Li, Z.; Yang, J.; Li, H. STLS-LADMM-Net: A Deep Network for SAR Autofocus Imaging. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5226914. [CrossRef]
39. Liu, Z.; Yang, S.; Gao, Q.; Feng, Z.; Wang, M.; Jiao, L. AFnet and PAFnet: Fast and Accurate SAR Autofocus Based on Deep Learning. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5238113. [CrossRef]
40. Liu, K.-H.; Wiesel, A.; Munson, D.C. Synthetic Aperture Radar Autofocus via Semidefinite Relaxation. *IEEE Trans. Image Process.* **2013**, *22*, 2317–2326. [CrossRef]
41. Zhang, T.; Liao, G.; Li, Y.; Gu, T.; Zhang, T.; Liu, Y. A Two-Stage Time-Domain Autofocus Method Based on Generalized Sharpness Metrics and AFBP. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5205413. [CrossRef]

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.