Stereo SAR for Building Imaging

Yun Lin, Yanping Wang, Junqi Tian, Jiankun Dong, Wenjie Shen, and Wen Jiang

Radar Monitoring Technology Laboratory, School of Artificial Intelligence and Computer Science, North China University of Technology, Beijing 100144, China

Keywords: SAR, stereo SAR, 3D reconstruction, polar coordinate transformation

Abstract

Structural health monitoring is essential for building safety. While SAR provides all-weather, non-contact imaging, it is often affected by geometric distortions like layover and foreshortening, making it difficult to extract accurate 3D structural information from complex targets like buildings. Inspired by stereo vision, we propose a stereo SAR mode that acquires two images via a single rotation. By transforming Cartesian to polar coordinates, the disparity is constrained to the angular direction, significantly simplifying the matching process. We derive the nonlinear relationship between height and disparity and apply Newton’s iterative method for accurate 3D reconstruction. Real data collected by a millimetre-wave radar system validate the effectiveness of the proposed approach.

1. Introduction

Structural health monitoring of buildings and infrastructure is essential to prevent catastrophic failures and ensure public safety. In recent decades, radar remote sensing, particularly synthetic aperture radar (SAR), has emerged as a promising tool for structural health monitoring due to its potential for millimetre-level deformation monitoring, penetration capability, and all-weather operation. When deployed as a ground-based SAR system, it enables continuous and uninterrupted monitoring of distributed targets such as slopes, dams, and building facades.

Operating in the microwave band, SAR can penetrate clouds, rain, and fog, making it an all-weather, day-and-night imaging system. However, despite its success in slope deformation monitoring, the application of ground-based SAR in building structural health monitoring has been limited. This is mainly due to inherent SAR imaging artifacts such as layover and foreshortening, which make it difficult to retrieve accurate 3D structural information from complex vertical structures like buildings.

Existing SAR 3D imaging techniques primarily include Tomographic SAR (TomoSAR) and Circular SAR (CSAR). While effective in their respective domains, both face challenges for ground-based continuous monitoring.

In the field of TomoSAR, which synthesizes an aperture in the elevation direction from multiple passes, significant advances have been made. For instance, Shi (2020) proposes a nonlocal filtering method for SAR tomographic inversion to address the challenge of limited baselines, validated using five bistatic interferograms from TanDEM-X. Subsequently, Ding et al. (2024) introduced a Multi-Master TomoSAR 3D imaging technique to enhance reconstruction performance, which is verified with TerraSAR-X data. To alleviate the computational burden, Qian (2022) develops a novel deep learning-based super-resolution TomoSAR inversion approach, termed γ-Net, evaluated on six baselines of TerraSAR-X data. Despite these algorithmic improvements, TomoSAR remains constrained by orbital mechanics in spaceborne implementations, leading to long revisit times unsuitable for continuous monitoring. Moreover, it is rarely implemented in ground-based systems due to the high cost and complexity of vertical rail deployment.

Regarding CSAR, which observes targets from various aspect angles to extract 3D information, recent research includes the optical-flow-inspired 3D reconstruction method by Palm et al. (2020) and the probabilistic framework for joint height and anisotropy estimation by Zhang et al. (2022). However, since CSAR requires the sensor to move around the target, it is primarily confined to airborne platforms and is likewise impractical for permanent or continuous monitoring.

Inspired by binocular stereo vision, this paper proposes an efficient stereo SAR imaging mode. By performing only a single rotation of the linear track around its central point, two SAR images from different range-Doppler planes are acquired. The difference in geometric distortions between these images is used to accurately reconstruct the 3D structure of the target. The imaging geometry of SAR is fundamentally different from that of optical cameras. The disparity between images acquired at different rotation angles lacks a strict epipolar geometry as in optical stereo, meaning the correspondence search is inherently two-dimensional. This paper derives an analytical expression for image offset as a function of rotation angle and proposes a polar coordinate transformation that effectively reduces the inherently 2D correspondence search to a 1D problem. In the resulting images, the disparity occurs only in the horizontal (angular) dimension, analogous to epipolar-rectified images in optical stereo, which greatly simplifies matching.

Furthermore, owing to the unique imaging geometry of range-Doppler-based SAR, the relationship between target height and image offset is nonlinear, in contrast to the linear depth-disparity relation found in optical stereo systems. This paper derives the exact expression that relates 3D coordinates to image offsets. Subsequently, Newton's iteration method is employed to accurately retrieve the target height, thereby enabling the reconstruction of full 3D coordinates.

Experimental validation with real building data, gathered by a prototype millimetre-wave stereo SAR system, confirms the method's effectiveness.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-3-2026-427-2026 |

2. Imaging Geometry and Signal Model

The imaging geometry of the rotating-track stereo SAR is illustrated in Figure 1. A Cartesian coordinate system *XOY* is defined with the *XOY* plane representing the ground plane and the *Z*-axis pointing vertically upward. A linear rail is placed on the ground plane, with its centre located at the origin *O*. The rail rotates around the origin *O*, starting from the positive *OX*-axis, with counterclockwise rotation defined as positive. The rotation angle is denoted as θ. The radar is mounted on the linear rail and observes the area in front of it.

The rail is first rotated to an angle θ_1 and held stationary. The radar then moves along the rail, transmitting and receiving signals at uniform time intervals to acquire the SAR echoes. Subsequently, the rail is rotated again to a new angle θ_2, and the radar repeats the motion along the rail while transmitting and receiving echo signals.

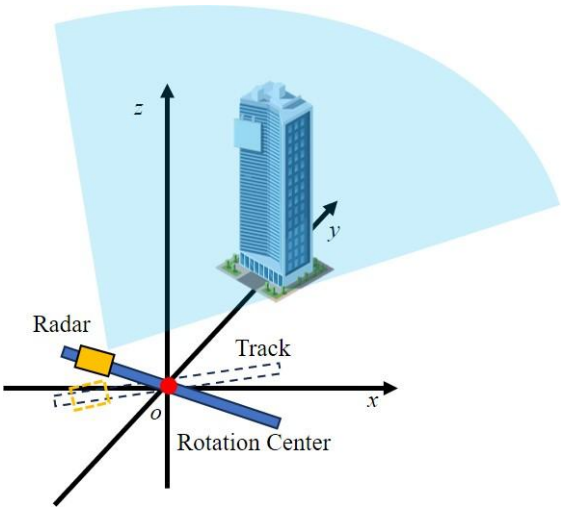

Figure 1. Imaging Geometry of rotating-track stereo SAR

Assume that the radar transmits a frequency-modulated continuous wave (FMCW) signal. The echo signal from a point target, after dechirping (mixing), can be expressed as:

$$s(t,\tau) \approx \sigma \cdot \exp\left\{-j\frac{4\pi\left(K(t-T_p/2)+f_c\right)}{C}R(\tau)\right\} \quad (1)$$

where σ denotes the scattering coefficient of the target, t is the fast time, τ is the slow time, T_p is the pulse width, K is the chirp rate, f_c is the centre frequency of the transmitted signal, C is the speed of light, and R denotes the time varying distance from the radar to the target.

The frequency of the beat signal in Equation (1) is proportional to the time delay. Applying the Fourier transform along the fast-time dimension yields the signal in the frequency domain:

$$s(f,\tau) \approx \sigma \cdot \exp\left(-j\frac{4\pi f_c}{C}R(\tau)\right)\mathrm{sinc}\left(T_pK\left(\frac{f}{K}+\frac{2R}{C}\right)\right) \quad (2)$$

Defining $t' = -f/K$, Equation (2) can be rewritten as:

$$s(t',\tau) \approx \sigma \cdot \exp\left(-j\frac{4\pi f_c}{C}R(\tau)\right)\mathrm{sinc}\left(T_pK\left(t'-\frac{2R}{C}\right)\right) \quad (3)$$

Equation (3) represents the signal after pulse compression in the range direction for a single point target. The echo from the entire scene is the superposition of echoes from all scattering centres.

3. Proposed Method

3.1 Main Principle

This method leverages the geometric distortion differences between two SAR images to accurately reconstruct the 3D structural information of the target. The expression for image disparity is first derived.

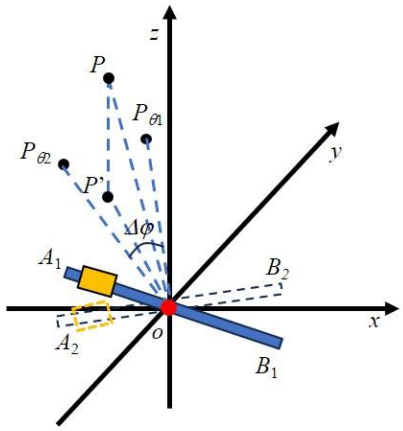

Figure 2. Disparity and height

As illustrated in Figure 2, data from two different rotation angles are both imaged with reference to the *XOY* plane. As a result, the disparity between the two images depends only on the Z-coordinate of the target. When z=0, no image disparity occurs.

The backprojection (BP) algorithm is used to generate the complex-valued images. The BP imaging expression is given by:

$$g(x,y) = \int s\left(t' = \frac{2R_{xy}(\tau)}{C}, \tau\right)\exp\left\{j\frac{4\pi f_c}{C}R_{xy}(\tau)\right\}d\tau \quad (4)$$

where R_{xy} denotes the distance from the radar to pixel (x, y). Since the images from different rotation angles exhibit significant decorrelation, and phase information is discarded in subsequent processing, and only the amplitude images, denoted as $G_\theta(x,y)$, are used.

As shown in Fig. 2., consider a point target P in the 3D scene. Its projected image point in $G_{\theta_1}(x,y)$ is denoted as P_{θ_1}, and in $G_{\theta_2}(x,y)$ as P_{θ_2}. According to the range-Doppler principle of SAR imaging, the distances from the origin O to P, P_{θ_1}, and P_{θ_2} are all equal. Therefore, the disparity between the projected points P_{θ_1} and P_{θ_2} under different rotation angles occurs only along the angular direction. By transforming the Cartesian coordinate system *XOY* into polar coordinates (θ, R), the disparity is constrained to the angular direction.

This coordinate transformation results in image disparity behaviour similar to that of epipolar-rectified images in binocular vision, where disparity occurs only horizontally, thereby significantly simplifying the matching process.

The relationship between disparity and height is derived as follows. Let P' be the vertical projection of P onto the XOY plane, and define $\angle A_1OP = \varphi_1$, $\angle A_1OP' = \varphi_1'$, $\angle B_2OP = \varphi_2$, $\angle B_2OP' = \varphi_2$.

From the range-Doppler imaging geometry, it follows that $\angle A_1OP = \angle A_1OP_{\theta_1}$, $\angle B_2OP = \angle B_2OP_{\theta_2}$. The angular disparity $\angle P_{\theta_1}OP_{\theta_2}$ can be expressed as:

$$\Delta\varphi = (\varphi_1 + \varphi_2) - (\varphi_1' + \varphi_2') \tag{5}$$

where

$$\varphi_1' + \varphi_2' = \pi + \theta_1 - \theta_2 \tag{6}$$

Let the elevation angle be $\angle POP' = \alpha$. Based on the geometric relations, we have:

$$\cos\varphi_1 = \cos\alpha\cos\varphi_1' \tag{7}$$
$$\cos\varphi_2 = \cos\alpha\cos\varphi_2' \tag{8}$$

There are five variables: φ_1 , φ_2 , φ_1' , φ_2' and α , three of which are redundant. Retaining φ_1 and α , the other variables can be expressed in terms of these two.

From Equation (7), we obtain:

$$\varphi_1' = \operatorname{acos}(\cos\varphi_1 / \cos\alpha) \tag{9}$$

From Equation (6), it follows that:

$$\varphi_2' = \pi + \theta_1 - \theta_2 - \varphi_1' \tag{10}$$

Substituting Equations (8), (9), and (10) into Equation (5) yields:

$$\begin{aligned}\Delta\varphi = {} & \varphi_1 \\ & + \arccos\left(\cos\alpha\cos\left(\pi + \theta_1 - \theta_2 - \operatorname{acos}(\cos\varphi_1 / \cos\alpha)\right)\right) \\ & - (\pi + \theta_1 - \theta_2)\end{aligned} \tag{11}$$

Equation (11) establishes a relationship between the angular disparity $\Delta\varphi$ and the elevation angle α , but it is difficult to express α explicitly in terms of $\Delta\varphi$. Given $\Delta\varphi$, Newton’s iterative method is applied to accurately determine α , which is then used to compute the 3D coordinates of the target. The iteration formula is:

$$\cos\alpha_{n+1} = \cos\alpha_n + \frac{\Delta\varphi_o - \Delta\varphi(\cos\alpha_n)}{\Delta'\varphi(\cos\alpha_n)} \tag{12}$$

where the subscript n denotes the iteration index, and $n = 0$ corresponds to the initial value (set to 0). $\Delta\varphi_o$ denotes the observed disparity. The derivative $\Delta'\varphi$ is given by:

$$\Delta'\varphi(\cos\alpha) = -\frac{(\cos\varphi_2' + \sin\varphi_2' \cdot \cot\varphi_1')}{\sin\varphi_2} \tag{13}$$

Once α is obtained, the 3D coordinates (x, y, z) of point P can be calculated as follows:

$$\begin{cases} x = -R_p\cos\alpha\cos(\varphi_1' + \theta_1) \\ y = R_p\cos\alpha\sin(\varphi_1' + \theta_1) \\ z = R_p\sin\alpha \end{cases} \tag{14}$$

where R_p represents the distance from a pixel in the image to the origin, and φ_1' denotes the angle formed by the pixel, the origin O, and point A_1 in the master image.

3.2 Workflow

The overall workflow of the method is illustrated in Figure 3. It consists of the following five steps:

Step 1: Backprojection (BP) Imaging. SAR data acquired at two different rotation angles are reconstructed into images in a unified ground coordinate system, resulting in complex images $G_{\theta_1}(x,y)$ and $G_{\theta_2}(x,y)$.

Step 2: Image Coordinate Transformation. Both amplitude images are transformed from the Cartesian coordinate system to the polar coordinate system, yielding $G'_{\theta_1}(\theta,R)$ and $G'_{\theta_2}(\theta,R)$.

Step 3: Image Matching. The two images are matched along the horizontal (angular) direction to estimate the image disparity $\Delta\varphi$. Matching can be performed using methods such as Normalized Cross-Correlation (NCC) or Daisy. This paper employs the NCC method.

Step 4: Elevation Angle Estimation. The elevation angle α is derived from the disparity $\Delta\varphi$ using Newton’s iterative method.

Step 5: 3D Coordinate Reconstruction. The 3D coordinates (x, y, z) of the target are calculated based on the elevation angle α .

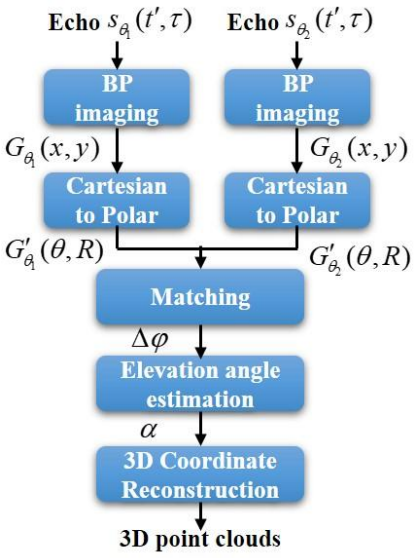

Figure 3. Flow chart

3.3 Theoretical Accuracy

The angular resolution has the expression as follows:

$$\Delta_a = \frac{\lambda}{2L} \tag{15}$$

where λ denotes wavelength, and L denotes the length of the linear rail. Matching accuracy is proportional to angular resolution. Assume the Matching accuracy is ρ_a, according to Equation (13), then the height estimation accuracy is about the following,

$$\begin{aligned} \rho_z &= \left| \frac{dz}{d\cos\alpha} \cdot \frac{d\cos\alpha}{d\Delta\varphi} \right| \cdot \rho_a \\ &= \left| \frac{\sin\varphi_2 \cdot \cot\alpha}{\left(\cos\varphi_2' + \sin\varphi_2' \cdot \cot\varphi_1'\right)} \right| \cdot R \cdot \rho_a \end{aligned} \tag{16}$$

4. Experimental Results

A prototype stereo SAR system is developed for validation, employing a TI AWR1843BOOST millimetre-wave radar. The radar is mounted on a linear rail, with the rail assembly fixed onto a turntable capable of precise angular increments. The experimental parameters are summarized in Table 1, and the observation scene with the radar setup is illustrated in Figure 4.

Parameter	Value
Radar Waveform	FMCW
Start frequency	77GHz
Chirp Rate	10 MHz/μs
Bandwidth	2198.68 MHz
Sample Rate	5000 ksps
Samples in Chirp	1024
Periodicity	2 ms
No of Frames	60000
Track Length	1.2 m
Rotation Angle	-9.4°, -9.4°

Table 1. Experimental parameters

Figure 5 shows a pair of SAR images in the Cartesian coordinate system, acquired at two different rotation angles, where the disparity is visible but complex in direction. The same image pair is transformed into polar coordinates and displayed in Fig. 6. After this transformation, the disparity is constrained to the horizontal (angular) direction, significantly simplifying the subsequent matching process.

The disparity map obtained using the NCC matching method is presented in Fig. 7. The polar coordinate images had an angular pixel spacing of 0.05°. The matching process for an image size of 887 × 2401 pixels required only 6.15 seconds in a MATLAB implementation on a standard 2.40 GHz CPU. This demonstrates a substantial efficiency improvement over the height-assumption matching strategy used in Zhang (2022), which involves an exhaustive search over a predefined height range and took over 5 minutes for the same task, highlighting the benefit of our one-dimensional matching strategy. The disparity map was masked using a coherence threshold of 0.7 and an image intensity threshold of -40 dB to filter out unreliable pixels.

Figure 8 shows the convergence of Newton’s iterative method for height estimation. The residual error between the measured disparity and the disparity calculated from the estimated height using Equation (13) converges to a negligible level after only six iterations, confirming the stability and efficiency of our inversion approach.

The final 3D point cloud reconstructed by the proposed method is presented in Fig. 9. A qualitative comparison with a reference LiDAR point cloud is shown in Figs. 10 and 11, indicating good agreement. The theoretical height accuracy can be estimated using Eq. (16). Assuming a slant range R = 60 m, a height z = 40 m, and a matching accuracy of one angular pixel (0.05°), the theoretical height error is approximately 0.26 m. The actual reconstructed building height was 43.43 m, which is in close agreement with the laser measurement, with a Root Mean Square Error (RMSE) of 0.28 m against the reference data. This result validates both the accuracy and practical utility of our method.

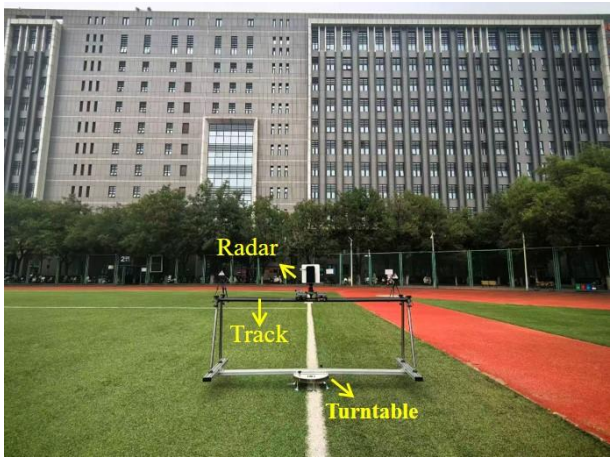

Figure 4. Experiment scene

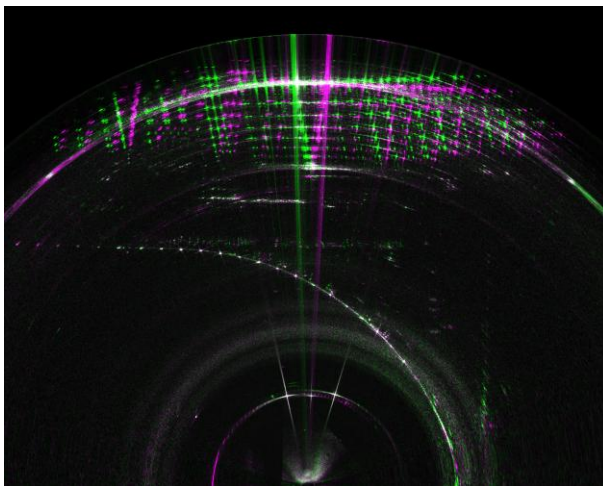

Figure 5. Image pair in cardisian coordinate (purple for θ_1, green for θ_2)

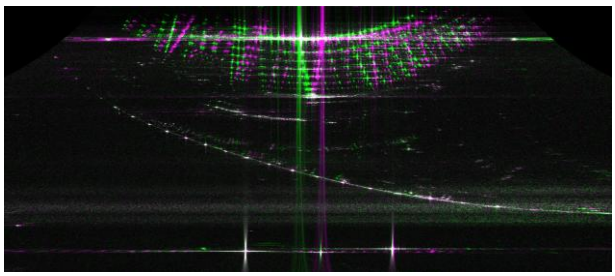

Figure 6. Image pair in polar coordinate (purple for θ_1, green for θ_2)

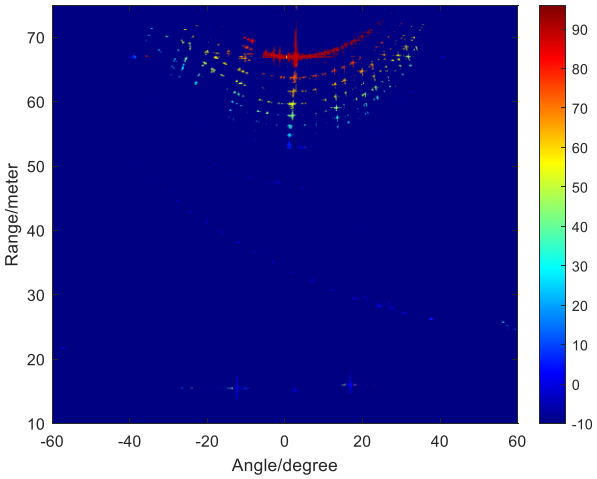

Figure 7. Disparity map with the proposed method

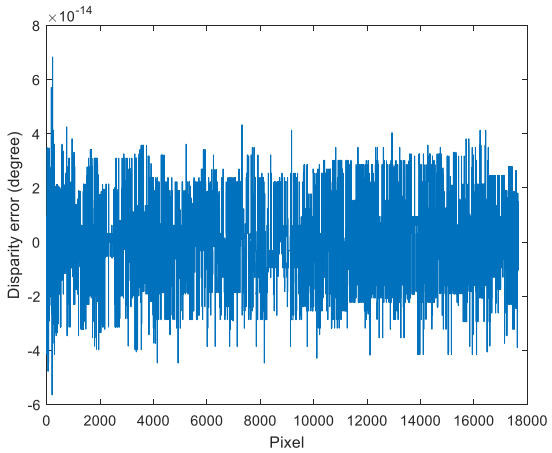

Figure 8. Disparity error after 6 iterations with Newton’s iteration

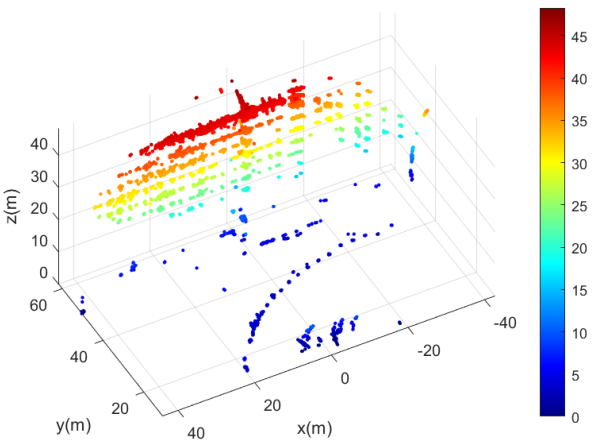

Figure 9. Reconstructed 3D point clouds with the proposed method

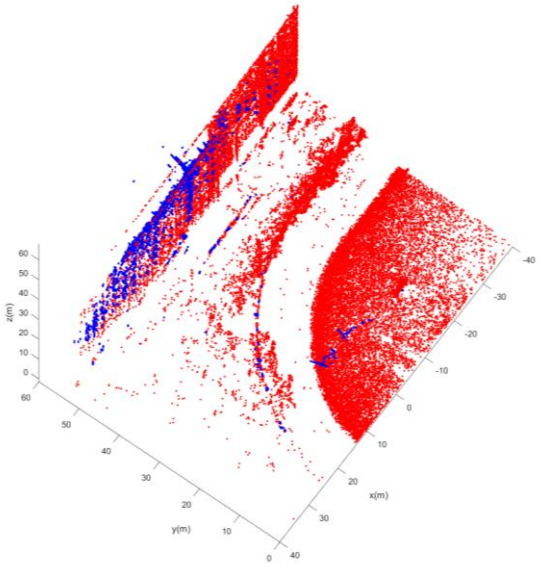

Figure 10. Result compared with Lidar (perspective view)

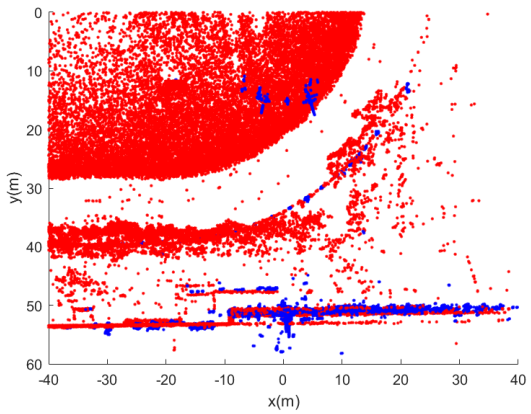

Figure 11. Result compared with Lidar (top view)

5. Conclusions

This paper proposes a novel and efficient stereo SAR imaging mode. The core innovation is a polar coordinate transformation that rectifies the disparity to a single dimension, dramatically reducing matching complexity. This, coupled with Newton's iteration to resolve the inherent nonlinearity height-disparity relationship, enables accurate 3D coordinate retrieval. The effectiveness of this approach is validated through real-data experiments using a developed prototype system, demonstrating its practical utility for 3D building imaging. Based on this 3D reconstruction foundation, our future research will target the inversion of 3D structural deformation using the stereo SAR technique.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China (Grants 62371005 and U25B6001), the R&D Program of Beijing Municipal Education Commission (No. 110052972508-20), and the Youth Research Special Project of NCUT (Project No. 2025NCUTYRSP010).

References

Y. Shi, R. Bamler, Y. Wang, and X. X. Zhu, “SAR Tomography at the Limit: Building Height Reconstruction Using Only 3–5 TanDEM-X Bistatic Interferograms,” IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 11, pp. 8026–8037, Nov. 2020, doi: 10.1109/TGRS.2020.2986052.

Z. Ding et al., “Multi-Master TomoSAR 3-D Imaging: Theoretical Complement and Performance Extension,” IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1–18, 2024, doi: 10.1109/TGRS.2024.3433033.

K. Qian, Y. Wang, Y. Shi, and X. X. Zhu, “γ-Net: Superresolving SAR Tomographic Inversion via Deep Learning,” IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–16, 2022, doi: 10.1109/TGRS.2022.3164193.

S. Palm and U. Stilla, “3-D Point Cloud Generation From Airborne Single-Pass and Single-Channel Circular SAR Data,” IEEE Transactions on Geoscience and Remote Sensing, vol. 59, pp. 8398–8417, 2020, doi: 10.1109/TGRS.2020.3041320.

H. Zhang, Y. Lin, F. Teng, and W. Hong, “A Probabilistic Approach for Stereo 3D Point Cloud Reconstruction from Airborne Single-Channel Multi-Aspect SAR Image Sequences,” Remote Sensing, vol. 14, no. 22, Art. no. 22, Jan. 2022, doi: 10.3390/rs14225715.