Urban SAR Tomography With Z-Structure Constraint of Buildings

Rui Guo , *Member, IEEE*, Yuxin Gao , Zishuai Ren, Zhao Zhang, Jinwei Xie , and Bei Yang

Abstract—Synthetic aperture radar tomography (TomoSAR) technology effectively mitigates issues such as severe layover in high-resolution synthetic aperture radar (SAR) urban imagery, presenting numerous applications in reconstructing complex 3-D urban scenes. However, prevailing TomoSAR methodologies usually overlook the geometrical feature of the targets, which are particularly pronounced in buildings in urban scenes. To improve the quality of 3-D reconstruction of building targets for large-scale urban scenes, we propose geometric-guided TomoSAR processing in this article. First, the geometrical feature referred to a Z-structure of buildings in TomoSAR point cloud is studied. Then, a Z-structure information extraction approach is proposed to provide prior information for subsequent geometric constraints. Finally, the Z-structure-constraint-based tomographic algorithm is proposed, optimizing the solution space of the original algorithms. Experiments are conducted on real SAR data, and 3-D point clouds of the entire urban scenes are obtained. The proposed algorithm demonstrates commendable performance in point cloud entropy and concentration, especially in the inversion of different floor structures.

Index Terms—Compressed sensing (CS), geometrical feature, synthetic aperture radar tomography (TomoSAR), 3-D reconstruction of buildings.

I. INTRODUCTION

SYNTHETIC aperture radar (SAR) 3-D imaging can directly obtain 3-D electromagnetic scattering of target structures and eliminate phenomena such as foreshortening, layover, and inversion, which is of great significance for applications such as 3-D environment construction [1], target interpretation [2], and urban surveying and mapping [3], [4]. Synthetic aperture radar tomography (TomoSAR), as a prominent technology in the field of SAR 3-D imaging, has aroused a research boom, whether at the level of single building or large-scale urban scene mapping.

Rui Guo, Zishuai Ren, and Bei Yang are with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: gr2003@nwpu.edu.cn; renzishuai@mail.nwpu.edu.cn; yangbei.nwpu.edu.cn@mail.nwpu.edu.cn).

Yuxin Gao is with the Jiangsu Shuguang Electro-Optics Company, Ltd., Yangzhou 225009, China (e-mail: gaoyuxin98@163.com).

Zhao Zhang is with the ZEEKR Automobile Company, Ltd., Shanghai 200232, China (e-mail: zhao.zhang7@zeekrlife.com).

Jinwei Xie is with the Nanjing Research Institute of Electronic Technology, Nanjing 210039, China (e-mail: jwxie2012@163.com).

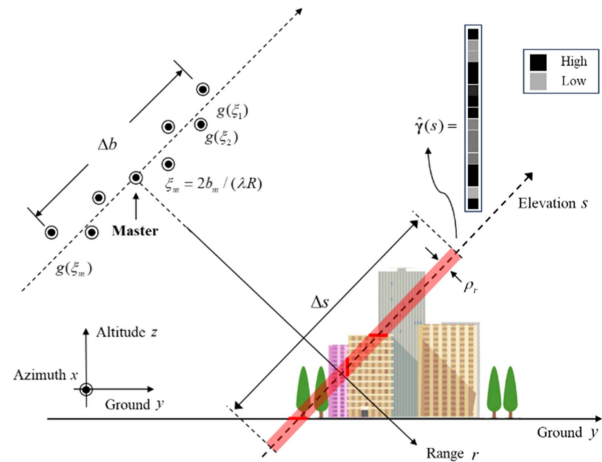Digital Object Identifier 10.1109/JSTARS.2024.3399538

Fig. 1. TomoSAR imaging geometry.

TomoSAR constructs a synthetic elevation aperture by capturing a series of M complex SAR images of the same area, forming an observation geometry, as shown in Fig. 1.

The focused imaging result is expressed as [5]

$$g_m = \int_{\Delta s} \gamma(s) \exp(-j2\pi\xi_m s)ds \qquad (1)$$

where $\gamma(s)$ is the reflectivity function along the elevation direction s. $\xi_m = b_m/(\lambda r)$ represents the Rayleigh resolution in the elevation direction, where b_m is the mth baseline relative to the master track, λ is the wavelength, r is the slant range, and Δb is the elevation aperture size. This model can be approximated as a linear equation with noise ε [6]

$$\mathbf{g} = \mathbf{\Phi} \cdot \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \qquad (2)$$

where $\mathbf{\Phi}$ is an $M \times L$ discretized measurement matrix, and its element $\Phi_{m \times l} = \exp(-j2\pi\xi_m s_l)$, where s_l ($l = 1, \dots, L$) is the lth sample of elevation s.

TomoSAR inversion is fundamentally considered as an array signal processing problem. The widely used TomoSAR imaging algorithms are spectrum estimation algorithms [5], [6] and sparse reconstruction algorithms based on compressed sensing (CS) theory [7]. The CS-based algorithms are more effective for 3-D tomography of urban scenes [8]. However, in the practical application of precise tomographic reconstruction, artifacts and aliasing resulting from issues, such as the limited number of acquired images, uneven baseline spacing, low signal-to-noise

ratio (SNR), and subpar achievable height resolution, can cause deterioration in the quality of 3-D point clouds [9], especially in the large-scale urban scene. Leveraging prior information in SAR data, such as the underlying geometrical and scattering features of the scene, stands out as an effective method to enhance multibaseline 3-D imaging technology [10], [11].

Preliminary research has been conducted on the improvement of utilizing prior information in tomographic 3-D reconstruction for the urban scene. Ding et al. [10] proposed the concept of SAR microwave visual 3-D imaging, which reduces the necessary observations for SAR 3-D reconstruction by combining information, such as microwave scattering mechanisms and visual semantics. Budillon et al. [12] applied deterministic constraints among the adjacent pixels, which consequently allow us to better identify the tower structure. Rambour et al. [8] systematically studied the TomoSAR imaging algorithm combined with spatial regularization and context information. Aghababaee et al. [13] introduced a regularization term, including prior information of buildings about the height changes in the array processing chain. Jiao et al. [14] used the local Gaussian Markov random field to explore the potential spatial connection between the scatterer and its neighborhood scatterers and constrained the position of the scatterer to the correct position. Pang et al. [15] proposed a joint sparse imaging algorithm based on constructing interest points and maximum likelihood estimation, which reduces the number of acquisitions of TomoSAR observations and suppresses scatterer outliers. Han et al. [16] proposed an algorithm that applies the extracted building façade geometry and the SNR of SAR images as prior information into an adaptive iterative shrinkage threshold algorithm, which successfully improves the single-building reconstruction.

However, current research on the prior information application is mostly limited to single-building reconstruction. When facing large-scale urban scene applications, complex issues, such as extraction of the layover regions for building targets and analysis of overlapping, should be considered.

In this article, we propose the TomoSAR processing for large-scale urban scenes based on the geometric prior information, referred to as the Z-structure. Research on the building geometry information is conducted, and the Z-structure information of typical buildings in the urban scene is proposed, including the footprint and the constraints of the building height and the roof length. Based on the preliminary TomoSAR point cloud, a Z-structure information extraction approach for the large-scale urban scene is proposed. First, buildings are extracted by point cloud filtering and cluster segmentation. Through different coordinate projection, the base and side masks for the building can be obtained separately. Then, the centerlines of building base masks are extracted to generate the footprints, and the constraint of the roof length is extracted through side masks. Based on the SAR imaging geometry, layover height maps that give the height value of each pixel in the layover region are generated. Finally, the extraction results of all the buildings are combined to generate the Z-structure information of the full scene. The preliminary point clouds are obtained by the efficient tomographic inversion method proposed in our previous work [17]. Afterward, the Z-structure information is utilized to
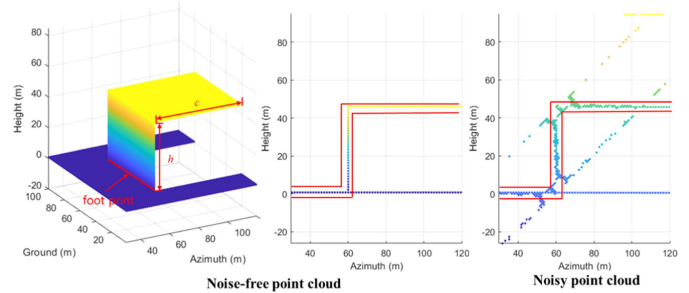

Fig. 2.   Z-structure information of the building in TomoSAR point clouds.

optimize the solution space of TomoSAR imaging for the urban scene, establishing the Z-structure-constraint-based TomoSAR. The main contributions of this article are as follows.

1) A typical building geometrical feature is analyzed, followed by an information extraction approach for the proposed Z-structure prior information.
2) Based on the Z-structure constraint, a solution space optimized TomoSAR imaging algorithm is proposed, which can be regarded as a search range adaptive method and can effectively reduce the processing of redundant information in the TomoSAR of the large-scale urban scene.
3) The proposed method outperforms the traditional method without the Z-structure constraint in point cloud concentration and accuracy [14], [16], which can better reconstruct the building façades and preserve the geometrical structure of different floors.

The rest of this article is organized as follows. The geometrical features in the 2-D SAR image and 3-D point cloud are analyzed, and the Z-structure information of building is introduced in Section II. The approach to extract the proposed Z-structure prior information is introduced in Section III. In Section IV, the Z-structure-constraint-based TomoSAR imaging algorithm is proposed. Experiments on real data are conducted in Section V to validate the effectiveness of introducing the proposed prior information to TomoSAR imaging algorithms. Finally, Section VI concludes this article.

II. GEOMETRICAL STRUCTURE FEATURES OF BUILDING

A. Analysis of the Z-Structure in Typical Buildings

For visual analysis of the building geometrical features, the building TomoSAR point cloud is generated by simulation under the noise-free condition, as shown in the first figure in Fig. 2. Parameters are set, as shown in Table I.

The underlying surface of a building contains the ground, façade, and roof [8], exhibiting an approximate "Z" structure in the azimuth–height plane outlined by the red line in the point cloud, as shown in the mid of Fig. 2. Due to the SAR side-view imaging principle, the layover of a building is distributed along the slant range direction. Therefore, the façades will face the ground–range–height plane obtained after the coordinate transformation. Since the roof and ground areas of the building are perpendicular to the façade plane, they will degenerate into lines at the top and the bottom of the façade, respectively, on the

TABLE I
PARAMETERS FOR SIMULATION

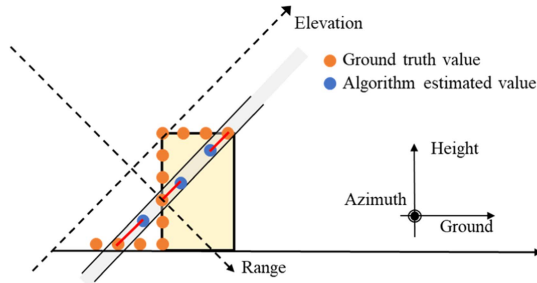| Parameter	Value
Wavelength	0.04 m
Incident angle	45°
Near slant range	500 m
Number of channels	11
Range and azimuth resolution	1 m × 1 m
Building height	50 m
Roof length	60 m
Building width	60 m

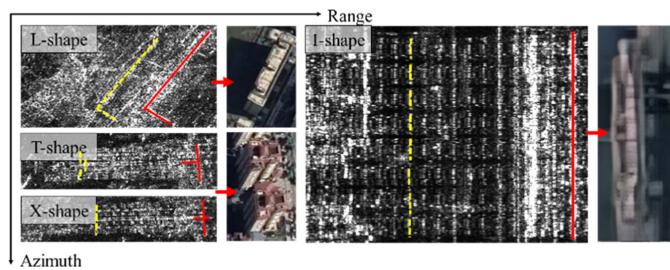Fig. 3.   Illustration of the building scatterer diffusion.

Fig. 4.   Geometrical features in the building layover region in the SAR amplitude image.

ground–height plane. Therefore, the Z-structure will not appear in the ground–height plane. In point cloud on the left of Fig. 2, the structure "Z" consists of the building roof constraint with length c, the façade height constraint with length h, and the footprint. The building footprint usually determines the position and shape of the building both in 2-D and 3-D SAR imaging.

However, due to various factors, including noise, elevation search grid settings, and the TomoSAR imaging algorithm performance, point clouds reconstructed by CS algorithms, such as orthogonal matching pursuit (OMP) [18], display typically an irregular "Z" shape. Adding noise with SNR = 40 dB to the simulated data, the building point cloud obtained by the OMP algorithm becomes the right figure in Fig. 2, which shows a scatterer diffusion phenomenon. The illustration of the scatterer diffusion is also provided in Fig. 3.

Due to the presence of noise, scatterers that are distributed on the building surfaces shown in orange may diffuse to positions represented by the blue points in Fig. 3, corresponding to scatterers outside the red box in the point cloud reconstructed by OMP after adding noise to the simulated data in Fig. 2.

Therefore, the reconstructed point clouds contain many stray points outside the Z-structure outlined by the red box, which make TomoSAR result blurring or cluttering. The Z-structure constraint aims at more accurate reconstruction of scatterers that diffuse beyond the proposed geometrical structure into their correct positions within the structure. The specific role of the Z-structure constraint in TomoSAR imaging will be discussed in Section IV.

B. Connection of the Z-structure in the SAR Amplitude Image and TomoSAR Point Cloud

To fully utilize the Z-structure prior information for the TomoSAR imaging, it is essential to connect the building area in the SAR amplitude image to that in the TomoSAR point cloud. Since the constraint based on Z-structure prior information is applied to each building in the urban scene, the subsequent analysis is conducted for individual buildings.

In the SAR amplitude image, the building information can be obtained from the building layover, as shown in Fig. 4. The geometric features of the building layover typically manifest as rectangles, parallelograms, or polygons formed by multiple parallelograms. In addition, each layover region is composed of multiple contour lines with the same direction along the direction of the building wall marked in yellow in Fig. 4, which are referred to as isoheight lines [19]. The shape of the strong angular reflection structure corresponding to different floors is similar to that of the building footprint. Therefore, the isoheight lines are typically parallel to the double-bounce lines at the bottom of the layover, and the shapes are similar to those of the building footprints marked in red.

The shape of the building footprint is consistent in 2-D and 3-D SAR imaging. The footprint $\mathbf{f}_C$ extracted from the Z-structure in TomoSAR point cloud in the *xyz* coordinate system can be converted into the *xrs* coordinate system to mark the position of the façade base in the SAR amplitude image. The coordinate directions *x*, *y*, *z*, *r*, and *s* indicate the azimuth, the ground range, the height, the slant range, and the elevation direction, respectively.

After obtaining the footprint $\mathbf{f}_C$ in the *xrs* coordinate system, the complete layover region of the building target can be extracted based on the scattering model and geometric projection. In our previous work, the relationship between the building height h and the layover length L_R has been studied [19].

Assuming that the scatterers in the building layover of the amplitude image are uniformly and completely distributed in each azimuth–range unit, for the height h of the building extracted from the vertical side of the Z-structure, the length of the building layover can be calculated as

$$L_R = \mathbf{ceil}\left(h \cdot \mathrm{ctg}\left(\theta\right) \cdot \sin\left(\theta\right) \cdot \rho_r\right) \tag{3}$$

where ρ_r represents the range resolution.

Since the building layover is composed of multiple isoheight lines arranged along the slant range direction, which are parallel to the footprint and have a similar shape, the building footprint can be regarded as the initial isoheight line located at the bottom of the building with a height of 0. For each isoheight line of the

building, there is only one fixed isoheight pixel in the azimuth direction [20], [21], [22]. By discretizing the façade height h into each pixel unit, the relationship between the height value $\mathbf{h}_P(x_P, r_P)$ of the scatterer P located at the azimuth unit x_P and the range unit r_P can be expressed as

$$\mathbf{h}_P\left(x_P, r_P\right) = h \cdot \left|r_P - r_{\mathrm{f}}\right| / L_R \tag{4}$$

where $r_{\mathbf{f}}$ represents the slant range unit position corresponding to the footprint $\mathbf{f}_C(x_P, r_{\mathbf{f}})$ at azimuth position x_P.

Then, the building layover is extracted, and the height information of each pixel in the layover is generated in the SAR amplitude image. The Z-structure constraint used in this article does not require additional extraction of building roof length in SAR amplitude images. At this point, the derivation is completed for extracting the building layover with geometry information from the 2-D amplitude image by utilizing the Z-structure information obtained from the 3-D point cloud.

III. Z-STRUCTURE INFORMATION EXTRACTION BASED ON THE PRELIMINARY TOMOSAR POINT CLOUD

A. Point Cloud Filtering and Segmentation

To extract the Z-structure information of each building in the urban scene, it is necessary to obtain independent point clouds for each building. Therefore, the point cloud filtering and segmentation steps are performed at first.

The stray scatterers including noisy scatterers are filtered out by thresholding. First, a search radius is set to iterate through each scatterer and count the number of scatterers contained in the search radius. Then, a threshold is set to filter the scatterers with the scatterer number in the neighborhood greater than the threshold, and the point cloud result filtered by thresholding is obtained. After projecting the filtered point cloud to xr coordinates, the ground points can be filtered out by density filtering. Then, the corresponding point cloud $\mathbf{P}_f$ without the ground point is obtained, which is subjected to clustering segmentation to separate and extract each building.

In this article, the density-based spatial clustering of applications with noise (DBSCAN) algorithm [23] is employed for TomoSAR point cloud segmentation, which achieves the best classification for the real data. The obtained clustering result is represented as $\mathbf{C} = \{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_k\}$, where $\mathbf{C}_k$ represents the kth extracted building.

B. Footprint Extraction

Most methods for footprint extraction are based on line segments or rectangle detection, followed by merging and grouping the extracted line segments, which may disrupt the continuity of building structures when applied to high-resolution SAR [24], [25]. Therefore, in this article, the building footprints are obtained by centerline extraction of the masks of the building bottoms based on morphological operations.

Dorninger and Pfeifer [26] first proposed façade extraction by orthogonally projecting point clouds onto the ground plane to obtain masks of the building bottoms. First, the point cloud is projected onto the xr plane, and the low-density grid units
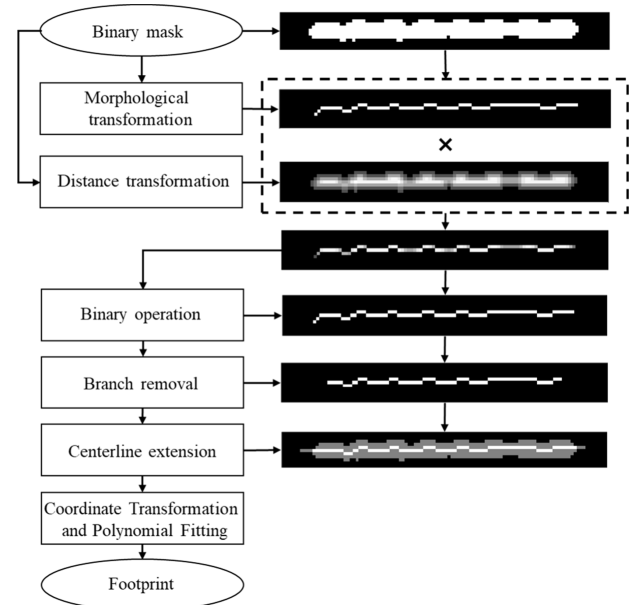

Fig. 5. Flowchart of the building footprint extraction based on morphology combined with distance transformation.

are removed, generating a binary mask of the building layover region. To obtain a more complete and continuous edge of the binary mask for centerline extraction, a sliding window of the size $w \times w$ is utilized to perform mean filtering on the projected scatterer density map of the point cloud. After filtering, the final binary mask representing the building base of the kth cluster is denoted as $\mathbf{M}_{Gk}$.

After obtaining the building base mask $\mathbf{M}_{Gk}$, this article utilizes an approach combining morphological operations with distance transformation to extract the centerline from the building layover mask. The overall processing with intermediate results is illustrated in Fig. 5.

First, the mask undergoes a morphological closing operation to further expand the area and fill in internal missing parts. Then, the skeleton $\mathbf{S}_{Gk}$ is extracted from the smoothed mask $\mathbf{B}_{Gk}$. Meanwhile, the complement of $\mathbf{B}_{Gk}$ undergoes distance transformation to obtain $\mathbf{D}_{Gk}$. Then, the skeleton $\mathbf{S}_{Gk}$ is elementwise multiplied with $\mathbf{D}_{Gk}$, and the result is binarized to obtain the initial centerline $\mathbf{f}'_{Gk}$. After removing small branches and extending, the footprint result $\mathbf{f}_{Gk}$ in the xy coordinate system is obtained. Finally, after coordinate conversion and connecting the scattered points at the bottom of the building through polynomial fitting, the footprint $\mathbf{f}_{Ck}$ in the xr coordinate system can be obtained.

C. Building Roof Length Constraint Generation

The constraint length c of the building roof is obtained by projecting the point cloud onto the xz coordinate system. By the density filtering of the projected point cloud, the binary mask of the building façade is obtained. From the building footprint $\mathbf{f}_{Gk}$ in the xy coordinate system, the average azimuth coordinate $\bar{x}_{\mathbf{f}k}$ of the centerline is obtained. Then, the constraint length of the
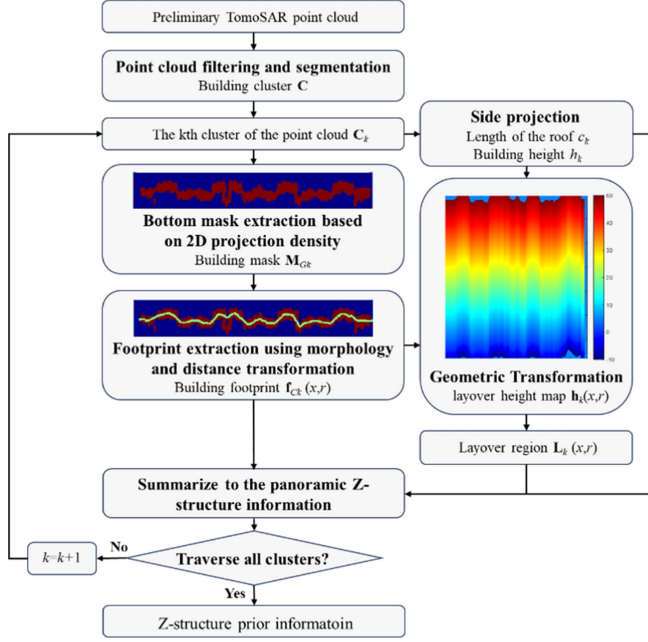
Fig. 6. Flowchart of the Z-structure prior information extraction approach.

Fig. 7. Principle diagram of the solution space constraint.

roof for the kth cluster is calculated as

$$c_k = \text{ceil}\left(\max(x_i \,|x_1, x_2, \ldots, x_N) - \bar{x}_{\text{f}k}\right) \qquad (5)$$

where x_i represents the azimuth coordinate of the binary mask in the xz coordinate system.

D. Layover Region and Height Map Generation

The binary mask of the building façade has been obtained from Section III-C yields the building height h_k. Then, the slant range L_{Rk} of the building layover can be obtained referring to (3). The building footprint $\mathbf{f}_{Ck}$ in the xr coordinate system is replicated along the opposite direction of the radar illumination by a length of L_{Rk} to generate the region of interest (ROI), which is labeled as the layover for the kth cluster. Meanwhile, the height $\mathbf{h}_k(x, r)$ is assigned to each pixel in the layover referring to (4).

E. Z-Structure Information Map Generation

After performing the steps in Sections III-B–III-D for all the clusters, the extracted information of individual buildings is aggregated into the panoramic information. First, accumulate the number of buildings for the layover pixels when the layover regions of multiple buildings overlap. Then, record the cluster number k and the pixel height $\mathbf{h}_k(x, r)$ in the height map for each pixel. After that, the panoramic Z-structure information map is obtained, including the building roof constraint length $\mathbf{c} = \{c_k | k = 1, 2, \ldots, K_{\text{cluster}}\}$, the collection $\mathbf{f}_C$ of all the building footprints, layover maps $\mathbf{L}$, and layover height map $\mathbf{h}$, where K_{cluster} is the number of building clusters.

The panoramic Z-structure information extraction approach is shown in Fig. 6.
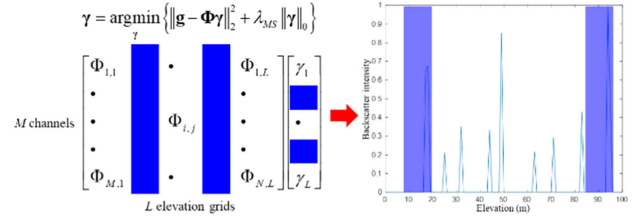
IV. 3-D RECONSTRUCTION OF URBAN SCENE BASED ON Z-STRUCTURE CONSTRAINTS

After obtaining the Z-structure prior information, the constraint can be introduced for the solution space of TomoSAR imaging algorithms to improve the building point cloud quality.

The solution space of the TomoSAR imaging algorithm corresponds to the search range in the elevation direction. Generally, the elevation search range s_{range} is smaller than the maximum ambiguity height

$$h_{\text{amb}} = \frac{\lambda R_{\min}}{2b_{\min}} \qquad (6)$$

where $R_{\min}$ is the minimum slant range and $b_{\min}$ is the minimum baseline; then, $s_{\text{range}} \ll h_{\text{amb}}$. The elevation search range is usually set based on the empirical value and remains consistent for the entire scene, which enables complete 3-D reconstruction of the urban area.

When both the high- and low-rise buildings are present in the scene, the unified search range is redundant for the low-rise buildings. Affected by the data noise and the imaging algorithm performance, the peak of the backscattering coefficient of some scatterers may erroneously appear within the redundant search range, reducing the TomoSAR imaging quality.

Therefore, the Z-structure prior information is used to exclude the solution space, where sidelobes are located, thereby improving the performance of TomoSAR imaging algorithms. The principle is illustrated in Fig. 7. The optimized solution space for TomoSAR imaging based on the Z-structure prior is determined as

$$s'_{\text{range}} = \mathcal{F}(x_P, r_P) = \Delta s_P. \qquad (7)$$

The constraint of a pixel depends on the height position of the scatterers in this pixel. For $P(x_P, r_P)$ in the xr coordinate system, the components of scatterers in the pixel and the Z-structure constraints marked in origin are shown in Fig. 8.

This article introduces the SNR index [16] to set a relaxation value for the Z-structure constraint. Given a sparse range $[\Delta_{\min}, \Delta_{\max}], \Delta_{\min}, \Delta_{\max} \in N$, the relaxation value of $P(x_P, r_P)$ is

$$\Delta_P = \Delta_{\min} + \frac{\text{SNR}(x_P, r_P) - T_{\text{SNR}}}{\text{SNR}_{\max} - T_{\text{SNR}}} \cdot (\Delta_{\max} - \Delta_{\min}) \quad (8)$$

where $\text{SNR}(x_P, r_P)$ represents the SNR value of $P(x_P, r_P)$, and $\text{SNR}_{\max}$ is the maximum SNR.

Fig. 8. Illustration of the Z-structure constraint for buildings.

Fig. 9. Flowchart of TomoSAR processing based on the Z-structure constraint for the urban scene.

The flowchart of the proposed TomoSAR processing based on the Z-structure constraint for the urban scene is shown in Fig. 9. For the urban scene, the maximum height is set as $h_{\max}>\max(\mathbf{h})$, and the minimum height is set as $h_{\min}<\min(\mathbf{h})$ based on the layover height map $\mathbf{h}$ obtained from the preliminary 3-D point cloud. The height of the SAR sensor is H, and the reference point of the scene is $O(x_0, y_0, h_0)$, where $y_0 = R_0 \cdot \sin(\theta)$. The upper limit of the ground point height satisfies $h_{\text{ground}} \in (h_0, h_{\max})$. Then, the search range Δs_P of $P(x_P, r_P)$ can be discussed with the following three cases.

1) $\mathbf{L}(x_P, r_P) = 0$. The pixel P contains no building scatterers but only a ground point P_1 in Fig. 8. Otherwise, Z-structure information of this pixel could not be correctly extracted. The search range for such pixel is set as $\Delta s_{\text{full}} = [s_{h\min}, s_{h\max}]$

$$s_{h\min} = -R_P \cdot \tan\left(\theta - \arccos\left(\frac{H - h_{\min}}{R_P}\right)\right) \tag{9}$$

Fig. 10. SAR amplitude images of (a) Emeishan and (b) Yuncheng. Corresponding optical images from Google Earth of (c) Emeishan and (d) Yuncheng.

$$s_{h\max} = R_P \cdot \tan\left(\arccos\left(\frac{H - h_{\max}}{R_P}\right) - \theta\right) \tag{10}$$

where R_P represents the slant range of P. Therefore, the elevation search range for case 1) is defined as $\Delta s_P = \Delta s_{\text{full}}$.

2) $\mathbf{L}(x_P, r_P) = 1$. The pixel P contains one building scatterers. By searching for the building index k in the layover height map, $\mathbf{h}(x, r, k)$ corresponding to the layover scatterer under k index can be obtained; thereby, the position of P in the building Z-structure can be determined. For the kth building, the temporary search range Δs_k is divided into the following cases.

a) $\mathbf{h}(x_P, r_P, k) \leq h_{\text{ground}}$. The scatterer locates at the bottom of the building, which can be treated as a ground point, and the search range is set as $\Delta s_k = \Delta s_{\text{ground}} = [s_{h\min}, s_{h\text{ground}}]$, where the maximum search range $s_{h\text{ground}}$ is fixed as

$$s_{h\text{ground}} = R_p \cdot \tan\left(\arccos\left(\frac{H - (h_{\text{ground}} + \Delta_P)}{R_p}\right) - \theta\right). \tag{11}$$

b) $h_{\text{ground}} < \mathbf{h}(x_P, r_P, k) \leq h_{\text{ceil}k}$, where $h_{\text{ceil}k} = 4/5 \cdot h_k$ represents the roof height threshold, and h_k represents the height of the kth building. In this case, P contains a ground scatterer P_1 and a façade scatterer P_2 in Fig. 8. Then, the search range is set as $\Delta s_k = \Delta s_{yk} = [s_{k\text{left}}, s_{k\text{right}}]$

$$s_{k\text{left}} = -R_P \cdot \tan\left(\theta - \arcsin\left(\frac{y_{k\text{left}}}{R_P}\right)\right) \tag{12}$$

$$s_{k\text{right}} = R_P \cdot \tan\left(\arcsin\left(\frac{y_{k\text{right}}}{R_P}\right) - \theta\right) \tag{13}$$

where

$$y_{\text{left}} = y_0 + \min(y_{\text{f}kP}) - \Delta y - \Delta_P \tag{14}$$

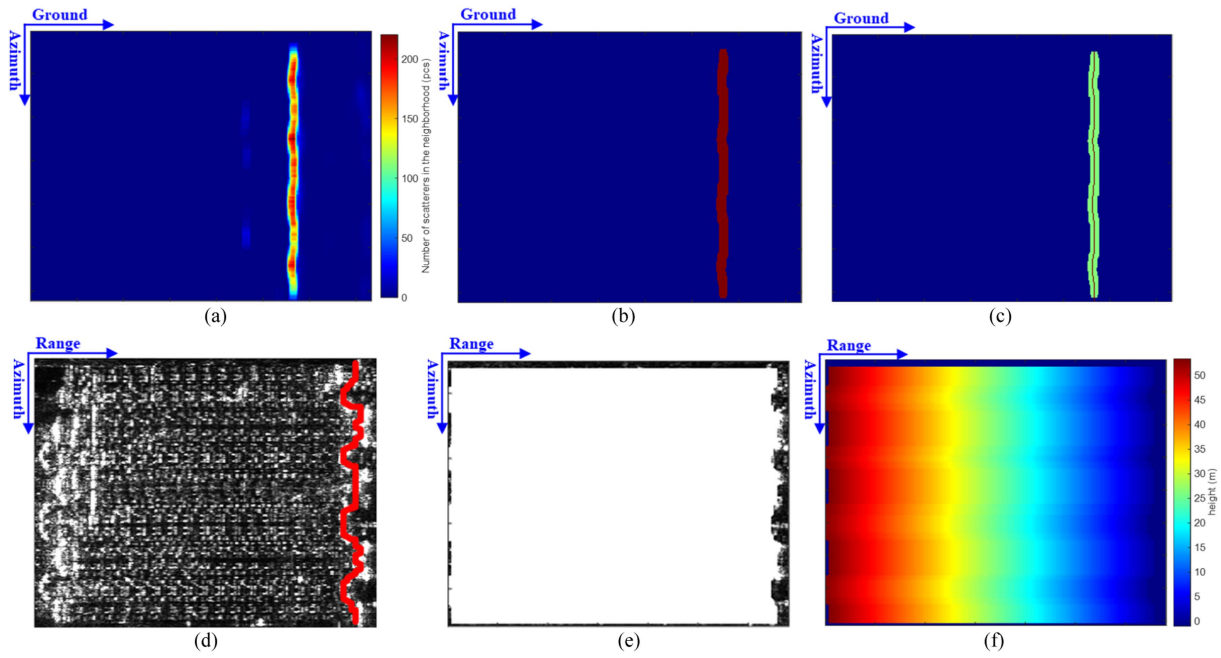$$y_{\text{right}} = y_0 + \max(y_{\text{f}kP}) + \Delta y + \Delta_P. \tag{15}$$

Fig. 11. Results of Z-structure information extraction for a single building in Yuncheng. (a) Density map of the projected 3-D point cloud after filtering. (b) Binary mask of the building base. (c) Building footprint extraction result. (d) Footprint after coordinate transformation. (e) Layover region extraction result. (f) Layover height map.
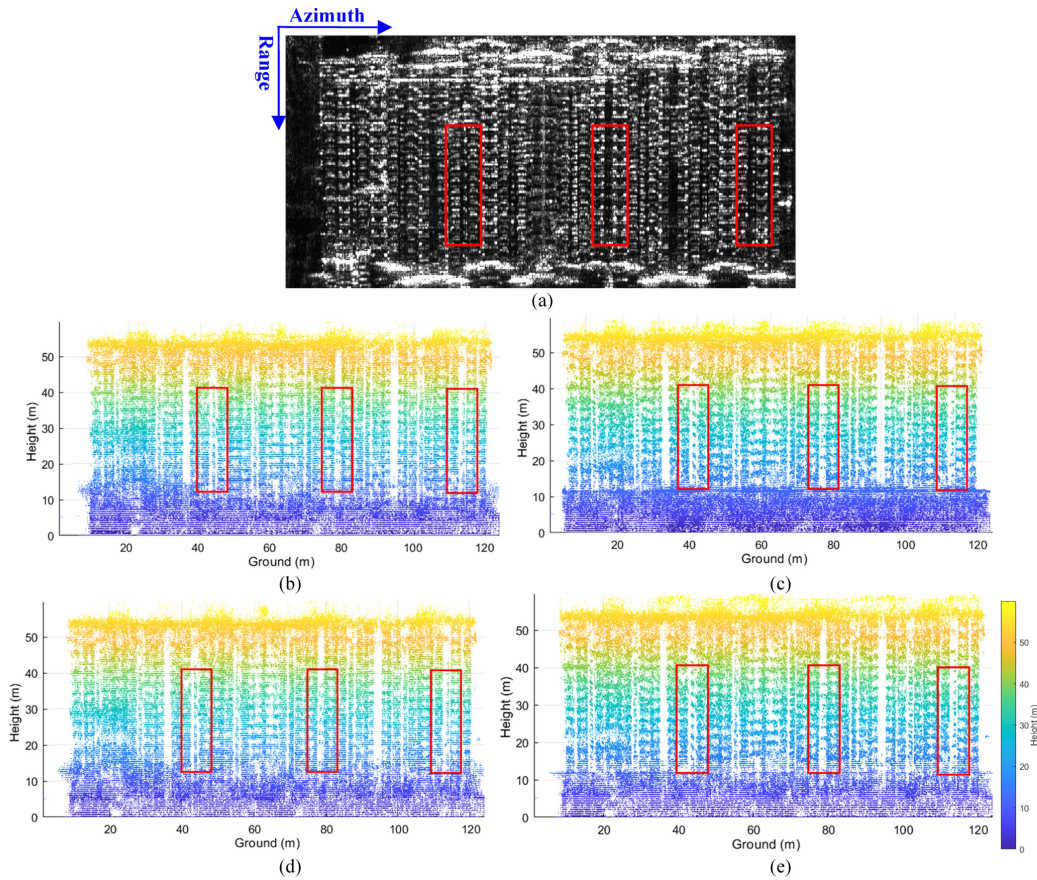

Fig. 12. (a) Amplitude image of the building in Yuncheng. TomoSAR point clouds of Yuncheng area in the *xz* plane obtained by (b) OMP, (c) Z-OMP, (d) FISTA, and (e) Z-FISTA.
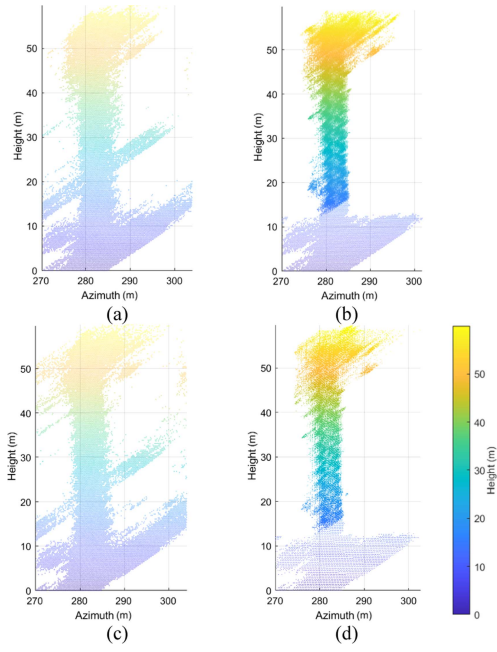
Fig. 13. TomoSAR point clouds of Yuncheng area in the yz plane obtained by (a) OMP, (b) Z-OMP, (c) FISTA, and (d) Z-FISTA.

y_{fkP} in (14) and (15) represents the azimuth position of the kth building footprint corresponding to the slant range r_p, and Δy is the preset relaxation value of the façade thickness.

c) $\mathbf{h}(x_P, r_P, k) > h_{\mathrm{ceil}k}$. The pixel P contains a ground scatterer P_1, a façade scatterer P_2, and a roof scatterer P_3 in Fig. 8. Given the length of the building roof and height, the elevation search range is $\Delta s_k = \Delta s_{ck} = [s_{k\mathrm{left}}, s_{ck}]$. The maximum search range s_{ck} is calculated as

$$s_{ck} = y_0 + \max(y_{fkP}) + c_k + \Delta_p \qquad (16)$$

where c_k represents the roof constraint length of the kth building.

Regardless of whether point P belongs to case a), b), or c) mentioned above, the elevation constraint range should include a search range $\Delta s_{\mathrm{ground}} = [s_{h\min}, s_{h\mathrm{ground}}]$ for the ground point. Moreover, for the kth building with a height of h_k, the elevation search range constraint should also be within $\Delta s_{hk} = [s_{h\min}, s_{hk}]$, where s_{hk} is calculated as

$$s_{hk} = R_P \cdot \tan\left(\arccos\left(\frac{H - (h_k + \Delta_P)}{R_P}\right) - \theta\right). \quad (17)$$

Therefore, the elevation search range of case 2) is calculated as

$$\Delta s_P = \Delta s_k \cup \Delta s_{\mathrm{ground}} \cap \Delta s_{hk}. \qquad (18)$$

3) $\mathbf{L}(x_P, r_P) > 1$. The pixel P contains the layovers of multiple buildings, including up to five scatterers in Fig. 8. From the layover height map, the height $\mathbf{h}(x_P, r_P, k)$ of P on the kth building can be obtained. For the kth building, the search range is set as Δs_k, $k = idx1, \ldots, idxn$. Then, the search range for case 3) is calculated as

$$\Delta s_P = (\Delta s_{idx1} \cup \cdots \cup \Delta s_{idxn}) \cup \Delta s_{\mathrm{ground}} \cap \Delta s_{hk\max}$$

Parameter	Area	
	Yuncheng	Emeishan
Carrier frequency	14.5 GHz	
Shortest slant range	1.17 km	1.92 km
Average Incident angle	23.94°	25.08°
Azimuth resolution	0.0734 m	0.1051 m
Range resolution	0.1499 m	0.1362 m
Number of trajectories	8	12

$$n = 1, 2, \ldots, \mathbf{L}(x_P, r_P) \qquad (19)$$

where $idxn$ represents the building index, and $\Delta s_{hk\max}$ is calculated as

$$s_{hk\max} = R_P \cdot \tan\left(\arccos\left(\frac{H - (\max(\mathbf{h}_k) + \Delta_P)}{R_P}\right) - \theta\right).$$

$$\mathbf{h}_k = h_{idx1}, \ldots, h_{idxn} \qquad (20)$$

After obtaining Δs_P, the optimized measurement matrix $\mathbf{\Phi}'_P$ based on the Z-structure constraint is generated for subsequent TomoSAR imaging algorithms.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, experiments are conducted on the SAR microwave vision 3-D imaging dataset (SARMV3D-1.0), which has been published in [27]. The data were collected by the airborne TomoSAR system equipped by Aerospace Information Research Institute, Chinese Academy of Sciences, from Yuncheng, Shanxi Province, China, in 2015, and Emeishan, Sichuan Province, China, in 2019. SAR amplitude images and optical images of the two areas are shown in Fig. 10, and parameters are given in Table II.

The preliminary TomoSAR 3-D point cloud of the entire area is obtained by the efficient TomoSAR inversion method proposed in our previous work [17]. Afterward, the TomoSAR imaging algorithms utilized in this article are the OMP algorithm [18] and the fast iterative shrinking threshold algorithm (FISTA) [28].

To exhibit the details in Z-structure prior information extraction, the proposed TomoSAR method is first applied to single buildings in Yuncheng and Emeishan. Results are evaluated by three indicators, including discrete ratio (DR) [16], TomoSAR point cloud entropy (3-D entropy) [16], and the average neighborhood height difference Δh_E [29]. A smaller DR value indicates a lower 3-D point cloud dispersion, and a smaller 3-D entropy indicates a lower 3-D point cloud confusion. Similarly, if the 3-D point cloud achieves a higher concentration and continuity in the height direction, the value of Δh_E will be smaller. Therefore, smaller values of all three indicators indicate better 3-D point cloud quality.

Afterward, the Z-structure information of the entire area is extracted in the experiments. The Z-FISTA, which achieves the best 3-D reconstruction results on single-building targets, is utilized for the reconstruction of the entire urban scene.
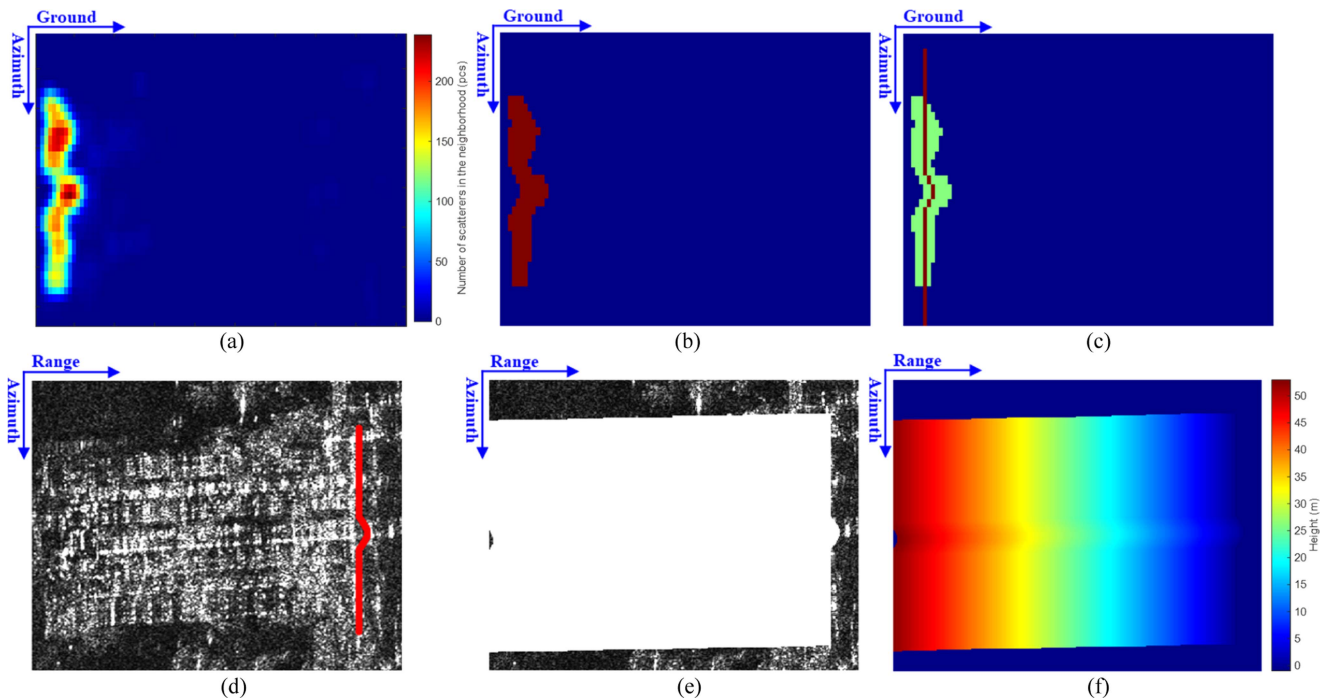
Fig. 14. Results of Z-structure information extraction for a single building in Emeishan. (a) Density map of the projected 3-D point cloud after mean filtering. (b) Binary mask of the building base. (c) Building footprint extraction result. (d) Footprint after coordinate transformation. (e) Layover region extraction result. (f) Layover height map.

A. Single-Building Experiments

Results of the Z-structure information extraction for a single building in Yuncheng are shown in Fig. 11.

After projecting the preliminary building point cloud onto the xy coordinate plane, the point cloud density is calculated. Then, the density map is filtered with a smoothing window of the size 5×5 to obtain Fig. 11(a). From the density map, the building base mask in Fig. 11(b) is extracted to identify the ROI. Utilizing the proposed building footprint extraction algorithm, the result of footprint extraction aligns well with the double-bounce line in the SAR amplitude image, as shown in Fig. 11(c) and (d).

The building layover height map is then calculated by (4) and (5) and illustrated in Fig. 11(f), which validates that the extracted layover region covers the building area almost completely. The length of the building roof constraint based on the side projection is extracted and set to 19 m.

Point clouds obtained by the original OMP algorithm and the FISTA are compared with those obtained by the OMP algorithm and the FISTA based on Z-structure prior information (referred to as Z-OMP and Z-FISTA) in Figs. 12(b)–(e) and 13(a)–(d) in different coordinate planes.

As shown in Fig. 12, compared to results in the xz plane obtained by OMP and FISTA, results of Z-OMP and Z-FISTA exhibit clearer building structures, better illustrating of the geometrical structures of each floor in the building. Particularly, in the area outlined by the red box, scatterers in point clouds of the original OMP and FISTA in Fig. 12(b) and (d) are blurry, making it difficult to separate different floors. In contrast, results

TABLE III
COMPARISON OF POINT CLOUD QUALITY OBTAINED BY DIFFERENT
ALGORITHMS FOR THE SINGLE BUILDING IN YUNCHENG

	DR	3-D entropy	Δh_E (m)
OMP	13.83%	3.8871	6.1622
Z-OMP	5.81%	3.8171	1.6720
FISTA	6.12%	3.7896	6.1576
Z-FISTA	3.85%	3.6742	1.4986

of Z-OMP and Z-FISTA in Fig. 12(c) and (e) clearly separate different floors of the building. Take the amplitude image of the building as a reference; the results based on the Z-structure constraint better maintain the floor structures in the 3-D point clouds, which are closer to that presented by the amplitude image, indicating that the scatterers are reconstructed in more accurate positions.

In the yz plane shown in Fig. 13, compared to the results of the OMP and Z-OMP algorithms in Fig. 13(a) and (b), the scatterers on the façades are more concentrated in the results of FISTA and Z-FISTA. The Z-structure of the building point cloud is clearer with fewer scattered points in the results of the Z-OMP and Z-FISTA.

To provide a more intuitive evaluation, three indicators are utilized for quantitative analysis of the results in Yuncheng area obtained by the four algorithms in Table III. The results of Z-OMP achieve even higher 3-D entropy than the original FISTA and a close DR compared with the original FISTA. TomoSAR imaging algorithms with Z-structure prior information achieve better indicator evaluation than the algorithms without
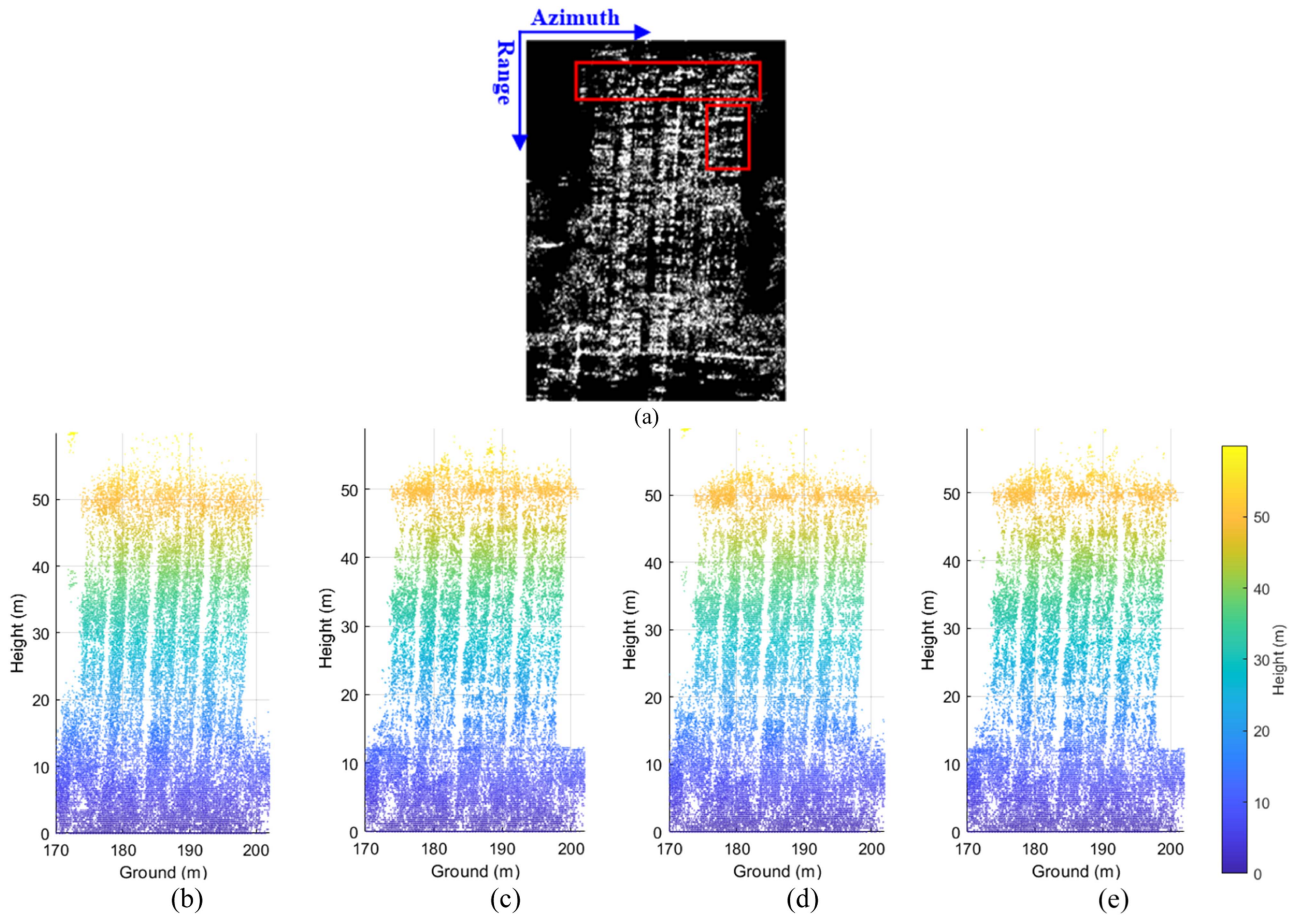
Fig. 15. (a) Amplitude image of the building in Emeishan. TomoSAR point clouds of Emeishan area in the *xz* plane obtained by (b) OMP, (c) Z-OMP, (d) FISTA, and (e) Z-FISTA.

Z-structure information, which validates the improvements of the concentration and accuracy of the point cloud by using the proposed Z-structure-based method. Among them, the Z-FISTA achieves the best result.

A single building is also selected from Emeishan area for the proposed experiment. The results of each step of Z-structure information extraction are shown in Fig. 14. The accurate result of building footprint extraction is also achieved on this building in Emeishan, which is almost consistent with the double-bounce line at the bottom of the building layover, as shown in Fig. 14(d). The extracted ROI covers almost the whole building area in the amplitude image in Fig. 14(e). Finally, a layover height map is generated. The roof length constraint of this building is set to 20 m.

Similarly, original OMP, Z-OMP, original FISTA, and Z-FISTA are utilized for comparison and analysis. The *xz* plane results are shown in Fig. 15, and the *yz* plane results are shown in Fig. 16.

The point cloud obtained by Z-OMP and Z-FISTA obtained more concentrated building roofs with less scattered points in Fig. 15(b) and (d). In terms of overall vision, the geometrical structures of different floors in the results of original FISTA and Z-FISTA are clearer than other algorithms, taking Fig. 16(a) as a reference.

TABLE IV
COMPARISON OF POINT CLOUD QUALITY OBTAINED BY DIFFERENT
ALGORITHMS FOR THE SINGLE BUILDING IN EMEISHAN AREA

	DR	3-D entropy	Δh_E (m)
OMP	12.02%	3.1718	8.4829
Z-OMP	8.29%	3.0707	6.9760
FISTA	8.63%	3.0982	6.1966
Z-FISTA	3.46%	2.9673	5.4099

In the *yz* plane, point clouds of the façade obtained by Z-OMP and Z-FISTA in Fig. 16(b) and (d) are clearer and more concentrated than those obtained by the original OMP and FISTA. However, the results of the original FISTA and Z-FISTA in Fig. 16(c) and (d) have clearer right-angle structures of the building roof, and point clouds of the roof are more complete. The concentration of the façade point clouds obtained by Z-OMP in Fig. 16(b) is even close to the original FISTA in Fig. 16(c).

For further quantitative analysis, the results in the Emeishan area of the four algorithms are compared by the DR, 3-D point cloud entropy, and average neighborhood height difference, as listed in Table IV.
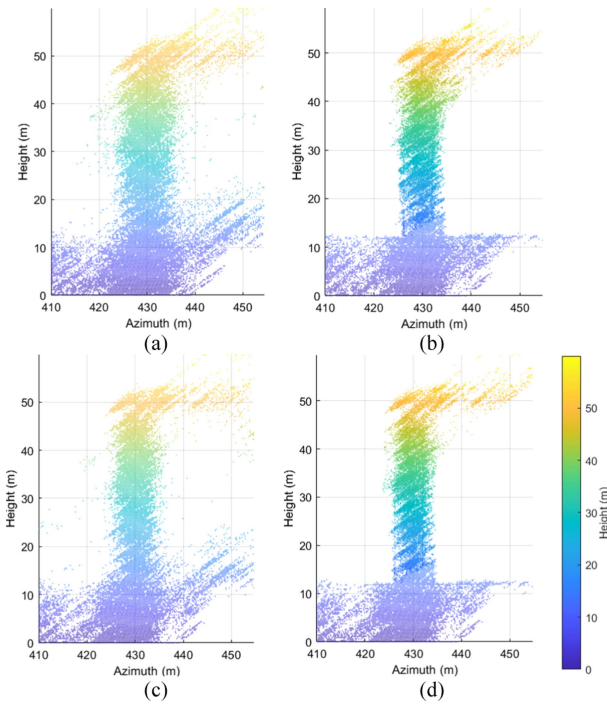
Fig. 16. TomoSAR point clouds of Emeishan area in the *yz* plane obtained by (a) OMP, (b) Z-OMP, (c) FISTA, and (d) Z-FISTA.

TABLE V
RESULT OF THE Z-STRUCTURE PRIOR INFORMATION EXTRACTION IN YUNCHENG AREA

Building index	Extracted or not	Height constraint (m)	Roof length constraint (m)	Detection rate
1	Yes	18	4	
2	Yes	28	7	
3	Yes	30	10	
4	Yes	70	16	
5	Yes	35	18	
6	Yes	30	18	
7	Yes	61	17	92.31%
8	Yes	60	19	
9	Yes	66	21	
10	Yes	65	21	
11	Yes	61	18	
12	Yes	61	19	
13	No	—	—	

Table IV implies that the quality of point clouds obtained by the TomoSAR imaging algorithms based on Z-structure constraints has been improved compared to the original algorithms. Z-OMP achieves a lower average neighborhood height difference than the original FISTA, with other indicators close to the original FISTA. The Z-FISTA achieves the lower DR, 3-D point cloud entropy, and average neighborhood height difference than the other three algorithms, achieving the best 3-D reconstruction results.

The results of 3-D reconstruction for the target buildings in Yuncheng and Emeishan both prove that the TomoSAR imaging algorithm based on Z-structure prior information can effectively suppress scattered points in the point clouds. Especially in building façades point cloud, the geometrical structures of different floors are well reconstructed. Compared to the original algorithms without the proposed prior information, higher quality 3-D point clouds are obtained with Z-structure prior information.

B. Experiments on the Large-Scale Urban Scene

The complete process of the proposed Z-structure constraint TomoSAR is conducted in the entire Yuncheng area and Emeishan area. For the entire urban area experiments, the Z-structure constraint is not applied in both nonbuilding areas and areas where the Z-structure information extraction failed.

First, the obtained Z-structure constraint prior information of Yuncheng area is shown in Table V, where the building height and the roof length are rounded up to cover the entire layover region of the buildings. The step-by-step experimental results are shown in Fig. 17. The building indexes in Table V correspond to the respective buildings in the preliminary point cloud in Fig. 17(a).

In Yuncheng area, the Z-structure prior information extraction method achieved a detection rate of 92.31% for the preliminary point cloud. All the buildings are nearly detected, and the building information is obtained based on the proposed method. Since the 13th building is inherently incomplete in the original dataset, fewer scatterers of the 13th building can be extracted from the original dataset, resulting in the low density of the 3-D point cloud. Moreover, the DBSCAN algorithm segments the buildings based on density. Therefore, it is difficult to completely extract the 13th building. As the point cloud of the 13th building is hardly formed, Z-structure information of this building is undetectable.

In Fig. 17(c) and (d), the detected buildings are projected onto the *xy* plane. Results of building footprint extraction closely match the double-bounce lines in the amplitude image of the building bases, displaying detailed geometrical structure.

The extracted layover regions are shown in Fig. 17(e), accumulating the number of building labels contained in each pixel in the layout height map. The light blue, yellow, and red areas contain information about one, two, and three buildings, respectively. Height distributions are generated for each layover region similar to the layover height map, as shown in Figs. 11(f) and 14(f). Based on the prior information from Fig. 17(c) and (e) and Table V, the Z-FISTA is utilized for the secondary TomoSAR imaging, as shown in Fig. 17(f). Compared to the preliminary results, scattered points are significantly better suppressed, with the façade point clouds more concentrated and clearer. Due to the undetected 13th building and the overlapping of the 13th building with the 12th building, partial structural loss occurred in the point cloud of the 13th building.

Z-structure information extraction and Z-FISTA 3-D reconstruction are also conducted on the entire Emeishan area. The result of Z-structure prior information extraction is shown in Table VI. Similarly, the building height and roof length constraints are rounded up. The step-by-step results are shown in Fig. 18.

In Fig. 18(c) and (d), all the building footprints are correctly extracted, aligning well with the double-bounce lines of the
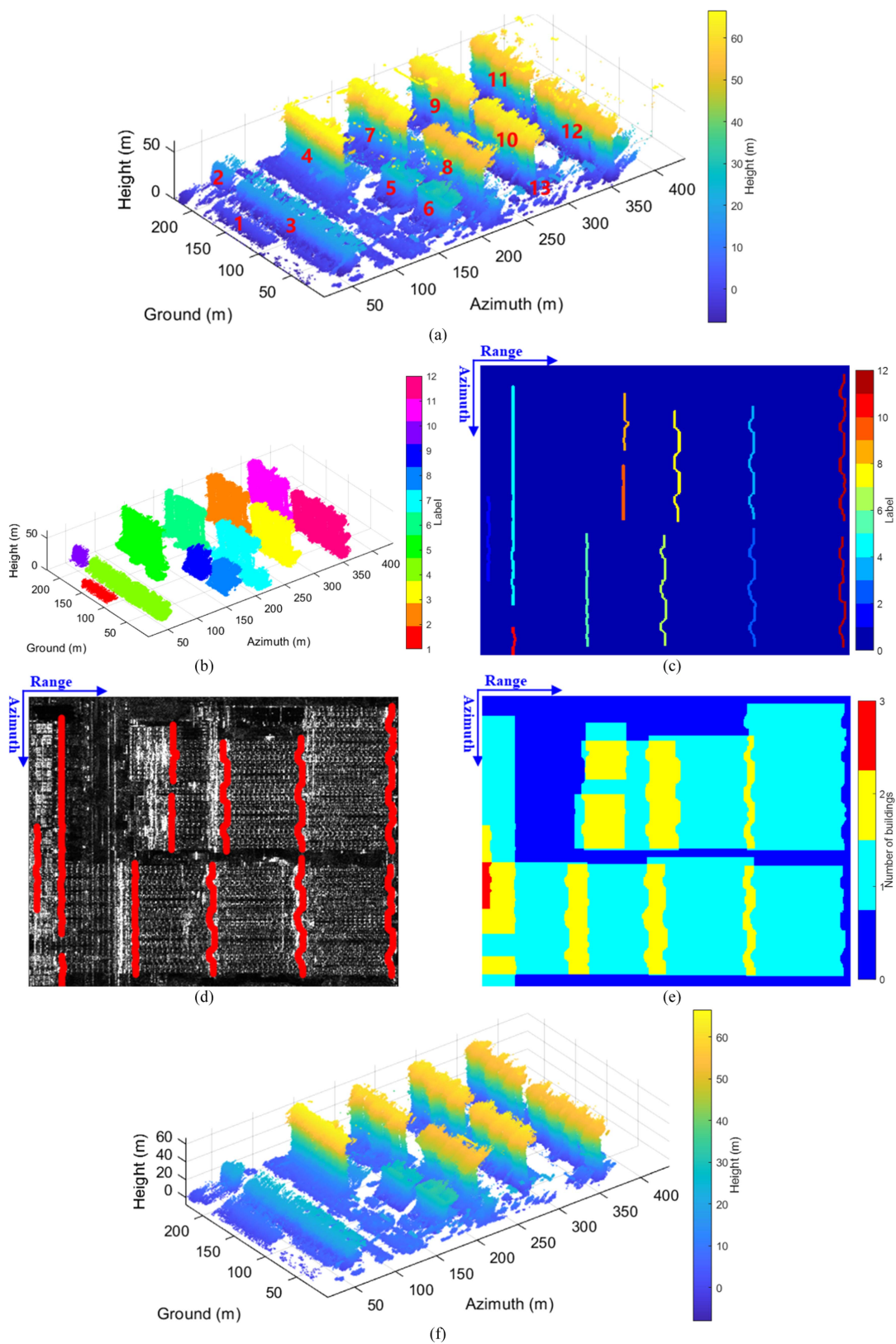
Fig. 17. Results of the TomoSAR 3-D reconstruction with the Z-structure constraint for the entire Yuncheng area. (a) Preliminary TomoSAR point cloud. (b) Building segmentation result. (c) Footprint extraction result. (d) Footprint after coordinate transformation in the amplitude image. (e) Layover region extraction result. (f) Final TomoSAR point cloud obtained by the Z-FISTA.

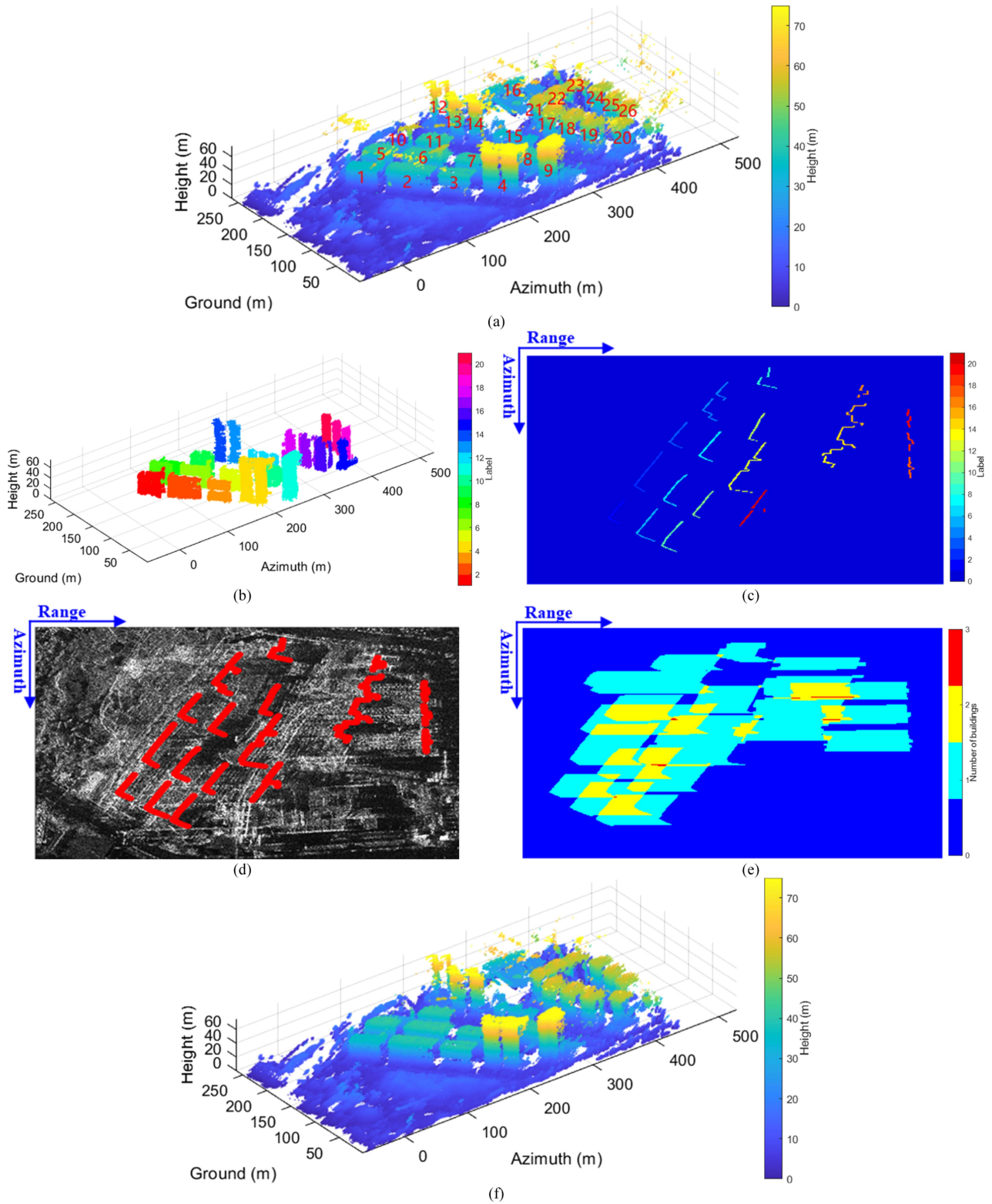(a)

(b)

(c)

(d)

(e)

(f)

Fig. 18. Results of the TomoSAR 3-D reconstruction with the Z-structure constraint for the entire Emeishan area. (a) Preliminary TomoSAR point cloud. (b) Building segmentation result. (c) Footprint extraction result. (d) Footprint after coordinate transformation in the amplitude image. (e) Layover region extraction result. (f) Final TomoSAR point cloud obtained by the Z-FISTA.

TABLE VI
RESULT OF THE Z-STRUCTURE PRIOR INFORMATION EXTRACTION IN
EMEISHAN AREA

Building index	Extracted or not	Height constraint (m)	Roof length constraint (m)	Detection rate
1	Yes	37	21	
2	Yes	40	22	
3	Yes	41	23	
4	Yes	75	27	
5	Yes	42	20	
6	Yes	41	21	
7	Yes	41	23	
8	Yes	41	16	
9	Yes	75	31	
10	Yes	25	22	
11	Yes	32	19	
12	No	—	—	
13	Yes	75	7	
14	Yes	75	9	80.77%
15	Yes	37	11	
16	No	—	—	
17	Yes	57	27	
18	Yes	57	27	
19	Yes	58	27	
20	Yes	57	22	
21	No	—	—	
22	No	—	—	
23	No	—	—	
24	Yes	55	20	
25	Yes	55	20	
26	Yes	55	20	

TABLE VII
QUANTITATIVE ANALYSIS OF 3-D RECONSTRUCTION RESULTS FOR THE ENTIRE
YUNCHENG AREA AND EMEISHAN AREA

Test area	Point cloud	DR	3-D entropy	Δh_E (m)	Time cost (s)
Yuncheng	Preliminary	8.40%	4.2994	8.6791	1133.0824
	Z-FISTA	5.83%	4.1988	6.1966	1591.2950
Emeishan	Preliminary	23.78%	4.6024	3.1108	2013.9861
	Z-FISTA	21.37%	4.2787	3.1078	2597.9799

VI. CONCLUSION

In this article, the Z-structure prior information of buildings in the urban scene was introduced, and an approach was proposed for extracting such information. Then, this prior information served as a constraint for the TomoSAR imaging algorithms to suppress the sidelobes in the redundant elevation search range and thereby improve the point cloud concentration and enhance the building structure. Finally, the TomoSAR based on the Z-structure constraint for the urban scene is established.

Single-building experiments and the large-scale scene experiments are conducted on the real SAR data. In single-building experiments, algorithms based on the Z-structure constraint achieve 3-D point cloud with clearer structures of different floors and more concentrated façade scatterers compared to the original algorithms. Among them, results of the Z-FISTA achieve the best indicators. In large-scale scene experiments, compared to the preliminary point clouds without the Z-structure constraint, the scattered points in the obtained point clouds were better suppressed, and the point clouds of the buildings were more concentrated in point clouds with the Z-structure constraint. However, some objects failed to be detected by the proposed Z-structure information extraction method, resulting in partial missing of the point clouds obtained by the subsequent Z-FISTA reconstruction.

Generally, most structures in the entire urban scene have been reconstructed. Thus, the reconstruction of these missing objects will be further studied in our future work. Meanwhile, the proposed Z-structure mainly concentrated on the regular geometric planes of building façades and roofs. More other urban structures will be considered in the further work.

building bases in the amplitude image. The building layover map is computed and generated in Fig. 18(e). Similar to Fig. 17(e), different colors of the areas represent information about a different number of buildings.

Based on the information from Fig. 18(c) and (e) and Table VI, the Z-FISTA is performed, obtaining the point cloud shown in Fig. 18(f). The 3-D point cloud obtained by the Z-FISTA has fewer scattered points compared with the preliminary point cloud. The remaining scattered points mostly concentrated in areas, where the Z-structure information is failed to be extracted. Moreover, the density of façade point clouds is higher than that of the preliminary point cloud. However, point cloud in Fig. 18(a) includes more objects in the urban scene, such as utility poles near the seventh building and the tenth building. Since these objects overlap with the building layover regions, the proposed processing cannot separate them, resulting in the partial missing.

To further validate the effectiveness of the entire urban 3-D reconstruction based on Z-structure constraints, the DR, 3-D entropy, and average neighborhood height difference are utilized to analyze the preliminary 3-D point clouds and the 3-D point clouds obtained by the Z-FISTA in the test areas in Table VII. Time costs of the original TomoSAR method and the proposed method in two areas are also compared. It can be seen that the proposed method does not introduce much time cost compared with the original method. The 3-D point clouds obtained from the Z-FISTA in both the areas achieved better indicators, with a significant improvement in the DR, indicating that the proposed processing can improve the concentration of the obtained point clouds.

REFERENCES

[1] F. S. Prol and M. M. Hoque, "A tomographic method for the reconstruction of the plasmasphere based on COSMIC/FORMOSAT-3 data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 2197–2208, 2022.

[2] A. H. Oveis, E. Giusti, S. Ghio, G. Meucci, and M. Martorella, "LIME-assisted automatic target recognition with SAR images: Toward incremental learning and explainability," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 9175–9192, 2023.

[3] Z. Yue, F. Gao, Q. Xiong, J. Wang, A. Hussain, and H. Zhou, "A novel attention fully convolutional network method for synthetic aperture radar image segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4585–4598, 2020.

[4] M. Che, A. Vizziello, and P. Gamba, "Urban change pattern exploration of megacities using multitemporal nighttime light and Sentinel-1 SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10681–10690, 2021.

[5] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.

[6] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.

[7] A. Budillon, A. Evangelista, and G. Schirinzi, "SAR tomography from sparse samples," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2009, pp. IV-865–IV-868.

[8] C. Rambour, A. Budillon, A. C. Johnsy, L. Denis, F. Tupin, and G. Schirinzi, "From interferometric to tomographic SAR: A review of synthetic aperture radar tomography-processing techniques for scatterer unmixing in urban areas," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 2, pp. 6–29, Jun. 2020.

[9] V. A. Jelalian, "Laser radar systems," in *Proc. Electron. Aerosp. Syst. Conf.*, 1980, pp. 546–554.

[10] C. Ding, X. Qiu, F. Xu, X. Liang, Z. Jiao, and F. Zhang, "Synthetic aperture radar three-dimensional imaging—From TomoSAR and array InSAR to microwave vision," *J. Radars*, vol. 8, no. 6, pp. 693–709, Dec. 2019.

[11] W. Liu, A. Budillon, V. Pascazio, G. Schirinzi, and M. Xing, "Performance improvement for SAR tomography based on local plane model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 2298–2310, 2022.

[12] A. Budillon, A. C. Johnsy, and G. Schirinzi, "Contextual information based SAR tomography of urban areas," in *Proc. Joint Urban Remote Sens. Event*, 2019, pp. 1–4.

[13] H. Aghababaee, G. Ferraioli, G. Schirinzi, and V. Pascazio, "Regularization of SAR tomography for 3-D height reconstruction in urban areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 2, pp. 648–659, Feb. 2019.

[14] Z. Jiao et al., "Urban 3D imaging using airborne TomoSAR: Contextual information-based approach in the statistical way," *ISPRS J. Photogrammetry Remote Sens.*, vol. 170, pp. 127–141, Jan. 2020.

[15] L. Pang, Y. Gai, and T. Zhang, "Joint sparsity for TomoSAR imaging in urban areas using building POI and TerraSAR-X staring spotlight data," *Sensors*, vol. 21, no. 20, pp. 1–15, Oct. 2021.

[16] D. Han, Z. Jiao, L. Zhou, C. Ding, and Y. Wu, "Geometric constraints based 3D reconstruction method of tomographic SAR for buildings," *Sci. China Inf. Sci.*, vol. 66, no. 1, pp. 1–13, Jan. 2023.

[17] R. Guo, Y. Gao, Z. Zhang, Z. Ren, S. Zhang, and M. Xing, "Efficient tomographic inversion based on refined scatterer pre-estimation," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4513305.

[18] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. IEEE 27th Asilomar Conf. Signals, Syst. Comput.*, 1993, vol. 1, pp. 40–44.

[19] R. Guo, F. Wang, B. Zang, G. Jing, and M. Xing, "High-rise building 3D reconstruction with the wrapped interferometric phase," *Sensors*, vol. 19, no. 6, pp. 1–14, Mar. 2019.

[20] X. X. Zhu and M. Shahzad, "Facade reconstruction using multiview spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3541–3552, Jun. 2014.

[21] R. Guo and X. X. Zhu, "High-rise building feature extraction using high resolution spotlight TanDEM-X data," in *Proc. IEEE 10th Eur. Conf. Synthetic Aperture Radar*, 2014, pp. 1–4.

[22] X. X. Zhu, N. Ge, and M. Shahzad, "Joint sparsity in SAR tomography for urban mapping," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1498–1509, Dec. 2015.

[23] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Discov. Data Mining*, 1996, pp. 226–231.

[24] A. Ferro, D. Brunner, L. Bruzzone, and G. Lemoine, "On the relationship between double bounce and the orientation of buildings in VHR SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 612–616, Jul. 2011.

[25] S. Auer, T. Balz, S. Becker, and B. Richard, "3D SAR simulation of urban areas based on detailed building models," *Photogrammetric Eng. Remote Sens.*, vol. 76, no. 12, pp. 1373–1384, Dec. 2010.

[26] P. Dorninger and N. Pfeifer, "A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds," *Sensors*, vol. 8, no. 11, pp. 7323–7343, Nov. 2008.

[27] X. Qiu et al., "Key technology and preliminary progress of microwave vision 3D SAR experimental system," *J. Radars*, vol. 11, no. 1, pp. 1–19, 2022.

[28] A. Beck and T. Marc, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[29] Z. Ren, Z. Zhang, Y. Gao, and R. Guo, "Three-dimensional imaging of tomographic SAR based on adaptive elevation constraint," *J. Radars*, vol. 12, no. 5, pp. 1056–1068, Aug. 2023.
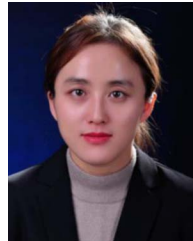
Rui Guo (Member, IEEE) received the B.S. and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2007 and 2012, respectively.

From 2012 to 2014, she was a full-time Postdoctoral Research Fellow with the Remote Sensing Technology Institute, German Aerospace Center, Weßling, Germany. She is currently an Associate Professor with the School of Automation, Northwestern Polytechnical University, Xi'an. Her research interests include radar and lidar technology, with application to remote sensing.

Yuxin Gao received the B.E. degree in transportation equipment and control engineering from Northwestern Polytechnical University, Xi'an, China, in 2023, and the master's degree in information and communication engineering from the School of Electronics and Information, Northwestern Polytechnical University, in 2024.

She is currently a Member of the Jiangsu Shuguang Electro-Optics Company, Ltd., Yangzhou, China. Her research interests include interferometric synthetic aperture radar and image processing, with application to remote sensing.

Zishuai Ren received the B.E. degree in transportation equipment and control engineering in 2022 from Northwestern Polytechnical University, Xi'an, China, where he is currently working toward the M.E. degree with the School of Automation.

His research interests include synthetic aperture radar 3-D imaging and point cloud processing.

Zhao Zhang (Student Member, IEEE) received the master's degree in information and communication engineering from the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, in 2023.

He is currently a Member of the ZEEKR Automobile Company, Ltd., Shanghai, China. His research interests include interferometric synthetic aperture radar (SAR) and tomographic SAR.

Jinwei Xie received the B.S. and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2014 and 2019, respectively.

He is currently with the Nanjing Research Institute of Electronic Technology, Nanjing, China. His research interests include tomographic synthetic aperture radar (SAR) and polarimetric interferometric SAR.

Bei Yang received the B.E. degree in automation from Northwestern Polytechnical University, Xi'an, China, in 2024, where she is currently working toward the M.E. degree in control science and engineering with the School of Automation.

Her research interests include interferometric synthetic aperture radar and image processing.