Automatic Detection and Reconstruction of 2-D/3-D Building Shapes From Spaceborne TomoSAR Point Clouds

Muhammad Shahzad, *Student Member, IEEE*, and Xiao Xiang Zhu, *Senior Member, IEEE*

Abstract—Modern spaceborne synthetic aperture radar (SAR) sensors, such as TerraSAR-X/TanDEM-X and COSMO-SkyMed, can deliver very high resolution (VHR) data beyond the inherent spatial scales of buildings. Processing these VHR data with advanced interferometric techniques, such as SAR tomography (TomoSAR), allows for the generation of four-dimensional point clouds, containing not only the 3-D positions of the scatterer location but also the estimates of seasonal/temporal deformation on the scale of centimeters or even millimeters, making them very attractive for generating dynamic city models from space. Motivated by these chances, the authors have earlier proposed approaches that demonstrated first attempts toward reconstruction of building facades from this class of data. The approaches work well when high density of facade points exists, and the full shape of the building could be reconstructed if data are available from multiple views, e.g., from both ascending and descending orbits. However, there are cases when no or only few facade points are available. This usually happens for lower height buildings and renders the detection of facade points/regions very challenging. Moreover, problems related to the visibility of facades mainly facing toward the azimuth direction (i.e., facades orthogonally oriented to the flight direction) can also cause difficulties in deriving the complete structure of individual buildings. These problems motivated us to reconstruct full 2-D/3-D shapes of buildings via exploitation of roof points. In this paper, we present a novel and complete data-driven framework for the automatic (parametric) reconstruction of 2-D/3-D building shapes (or footprints) using unstructured TomoSAR point clouds particularly generated from one viewing angle only. The proposed approach is illustrated and validated by examples using TomoSAR point clouds generated using TerraSAR-X high-resolution spotlight data stacks acquired from ascending orbit covering two different test areas, with one containing simple moderate-sized buildings in Las Vegas, USA and the other containing relatively complex building structures in Berlin, Germany.

Index Terms—Building footprint, building reconstruction, clustering, dynamic city models, TerraSAR-X, tomographic SAR (TomoSAR) inversion, four-dimensional (4-D) point cloud.

I. INTRODUCTION

AUTOMATED methods for the detection and reconstruction of buildings are becoming increasingly important due to the ever-increasing growth of urban population, giving rise to a wide range of potential applications in numerous fields. For instance, 2-D building footprints (or outlines) can be used in urban energy modeling [1], disaster management [2], and devising emergency responses [3]. Additionally, extruding these footprints using height information leads to the development of 3-D building models. Such models are valuable for various applications, e.g., virtual city modeling [4], urban planning [5], and analyzing electromagnetic reflections for placement of telecommunication antennas [6], etc.

Modern spaceborne synthetic aperture radar (SAR) sensors, such as TerraSAR-X/TanDEM-X and COSMO-SkyMed, are able to provide very high resolution (VHR) data beyond the inherent spatial scales (on the order of 1 m) of buildings, constituting an invaluable data source for detailed urban mapping on a global scale. Moreover, due to the coherent imaging nature of SAR, these sensors are able to provide day/night global coverage in virtually all weather conditions. Moreover, with image acquisitions taken at different times, they are also uniquely capable of imaging the dynamics of the illuminated area down to the scale of centimeters or even millimeters from space. These encouraging attributes have drawn the attention of many remote sensing analysts, and therefore, various techniques have been developed, which make use of VHR SAR imagery for 2-D/3-D reconstruction of man-made structures in particular buildings [7]–[11].

Although much research effort has been put in the area, the problem of building reconstruction remains a challenging task due to the inherent characteristics of SAR images, such as speckle, layover/foreshortening, etc. Moreover, complex building structures and high variability of objects appearing in SAR images also render difficulties in the detection of buildings, particularly in dense urban areas where buildings are compactly packed, and smaller buildings are often occluded by higher buildings. Therefore, prior knowledge with certain regularization (geometric) constraints (e.g., vertical facades [10] and rectangular building shapes [9]) is often incorporated for realistic and automatic reconstruction.

SAR tomography (TomoSAR) is an advanced interferometric technique that aims for real 3-D SAR imaging [12]–[16]. It resolves multiple/dominant scatterers within one resolution element (i.e., able to retrieve more than one scatterer per
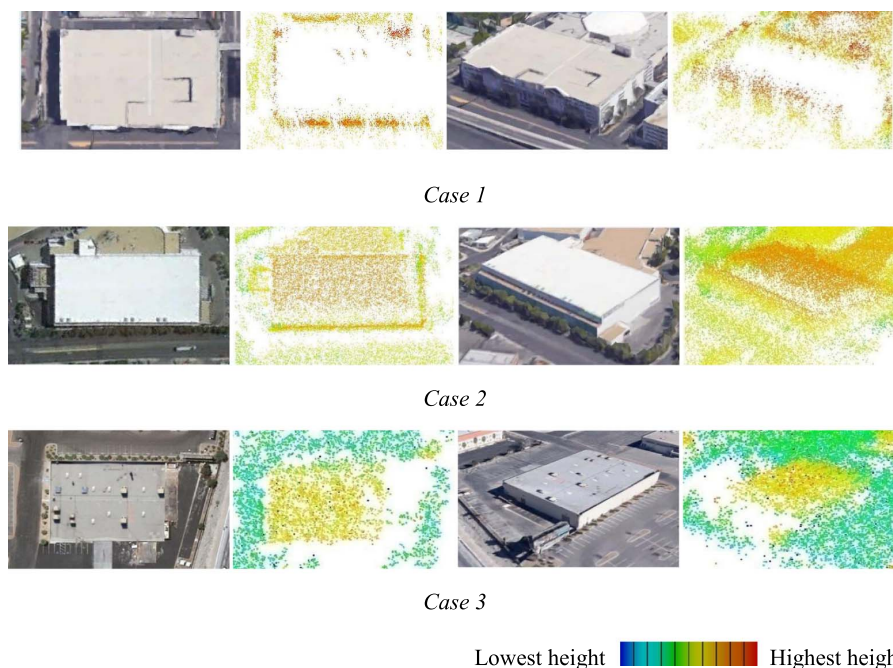
Fig. 1. Three different cases over simple buildings commonly occurring in TomoSAR point clouds. Heights of the points are color coded. (Optical images: Google).

pixel) via spectral analysis by building up a synthetic aperture in the elevation direction to localize each scatterer in 3-D [15], [16]. Likewise, building up a synthetic aperture in the temporal domain enables 4-D space–time) focusing of SAR images [17]–[19]. The technique is referred to as D-TomoSAR, which combines the strengths of both TomoSAR and persistent scatterer interferometry [17]–[21]. It is capable of retrieving elevation and deformation information (linear, seasonal, etc.) even of multiple scatterers inside a single SAR image pixel [15], [18], [22], [23]. Geocoding high density of scatterers, retrieved from TomoSAR, into world coordinates produces high-quality TomoSAR point clouds, containing not only the 3-D positions of the scatterer location but also the estimates of seasonal/temporal deformation, making them very attractive for generating 4-D city models from space.

Object reconstruction from spaceborne TomoSAR point clouds was started not until recently [24], [25]. These point clouds have point density in the range of $600\,000-1\,000\,000$ points/km^2 using TerraSAR-X meter resolution data [16], [27]. Similar statistics using CosmoSkyMed data are reported in [27]. TomoSAR point clouds, however, are associated with some special considerations that are worth to mention [24], [25].

1) They deliver moderate 3-D positioning accuracy on the order of 1 m [28].
2) Few number of images and limited orbit spread render the location error of TomoSAR points highly anisotropic, with an elevation error typically at least one order of magnitude higher than in range and azimuth.
3) Due to the coherent imaging nature, temporally incoherent objects such as trees cannot be reconstructed from multipass spaceborne SAR image stacks.

4) Ghost scatterers may be generated due to multiple scattering that appears as outliers far away from a realistic 3-D position [29].
5) Side-looking SAR geometry enables strong scatterer reflections from dihedral and trihedral structures or metallic parts resulting in higher density of points on man-made objects, particularly building facades and bridges. Naturally, this urges for the systematic reconstruction of building footprints via facade point analysis.

As depicted over smaller [25] and larger areas [24], facade reconstruction turns out to be an appropriate first step in detecting and reconstructing building shapes from these point clouds when dense points on the facade are available. In particular, when data from multiple views, e.g., from both ascending and descending orbits, are available, the full shape of buildings can be reconstructed using extracted facade points. However, there are cases when no or only few facade points are available. This usually happens for lower height buildings and renders the detection of facade points/regions very challenging. Moreover, problems related to the visibility of facades mainly facing the azimuth direction (i.e., having normals pointing toward the flight direction) can also cause difficulties in deriving the complete structure of an individual building. These problems motivate us to reconstruct full 2-D/3-D shapes of building footprints via facade–roof analysis. In this regard, based on different object contents illuminated by side-looking SAR, the following three cases could be derived using data acquired from one incidence angle, e.g., in this case, image stacks from the ascending orbit only (see Fig. 1).

Case 1) *Higher density of facade points present with no or very few roof points*: In this case, the complete

2-D/3-D building shapes could only be fully reconstructed by adding points from multiple incidence angles. The solution to this case is demonstrated in [24], where 3-D facade models have been reconstructed for high-rise buildings using one incidence angle only, and in [25], where the full shape of the building was derived by prior fusion of two point clouds [30] from ascending and descending stacks.

Case 2) *Higher density of facade points present together with the existence of relatively higher density of roof points*: This case allows for the reconstruction of the full shape of the building footprints from a single data stack by making use of both facade and roof points. Thus, the side of the building visible to the sensor could be reconstructed as the first step using facade points, and later, the other side of the building could be completed by exploiting the available roof points.

Case 3) *No or very few facade points available but enough roof points exist*: This case particularly appears for lower height buildings, rendering the detection of facade points/regions very challenging. This motivates us to obtain the complete 2-D/3-D footprint of these buildings via roof point analysis only using conventional techniques as applied by the light detection and ranging (LiDAR) community. Although these techniques are very much developed, still, their adaptation to TomoSAR point clouds is not straightforward due to different object contents illuminated by side-looking SAR together with problems related to less and varying point density and much less positioning accuracies of TomoSAR point clouds in comparison to airborne LiDAR.

In this paper, we propose a novel data-driven approach that systematically allows for the automatic reconstruction of 2-D/3-D building shapes (or footprints) using unstructured TomoSAR point clouds generated from one incidence angle only. The approach proposes new methods and aims at finding a more general and systematic solution toward the automatic reconstruction of the whole city area. This paper essentially presents solutions for the latter two cases (i.e., case 2 and case 3) by extending (or utilizing) the solution provided for case 1 in [24]. The innovative contributions specific to the approach proposed in this paper are as follows.

1) A hybrid approach based on the region-growing procedure and the energy minimization framework is proposed to automatically extract building roof/facade points directly from 3-D TomoSAR points. Seed points are extracted using the reconstructed facade information, and later, a minimum-height-constraint surface-normal-based region-growing approach is adopted to detect probable building regions in case 2. The problem of extracting roof points in case 3 (i.e., having no facade information) from the remaining set of points is then resolved by computing local height and planar features and formulating the whole detection problem into an energy minimization

framework. Graph cuts are later employed to globally extract roof/building points.

2) A recursive angular-deviation-based approach is presented to smooth/refine the initial coarse building polygons obtained using alpha shapes (generalization of convex hulls). The smoothed boundaries yield better visually appealing building shapes.

3) A novel facade–roof fusion procedure is proposed in this paper. The developed method is robust and fuses the legitimate *facade-Polygon* pair together by interpreting the refined/smoothed polygon of each building as a graph. A series of operations taking part in the reduction of graph (or smoothed polygon) includes identification of points on the graph that are nearest to all the associated/paired reconstructed facades, computation of longest and shortest paths on the graph using these identified points, simplification of shortest paths by representation as line segments, and rotation/projection of certain line segments (simplified shortest paths that are identified as positive paths) to match all the reconstructed facades belonging to the same graph (i.e., smoothed building polygon).

4) An effective and robust procedure is developed for rectilinear identification of building polygons. To this end, the dominant direction of each building is first determined based on the weighted-line-segment approach. Angular deviations at each adjacent or consecutive vertex of the polygon are computed from the dominant direction, and histogram analysis is then carried out for rectilinear building footprint identification. Later, rectilinear constraints are added to the identified building polygons.

5) The approach presented in [24] may fail to reconstruct building facades facing north or south because of missing measurements caused by the high inclination angle of the TerraSAR-X orbit, i.e., near-polar orbit (see Fig. 2). This paper inherently provides a solution to this problem by exploiting roof points in determining the complete shape/footprint of the building (see Fig. 3).

6) Finally, this paper presents the first demonstration of automatic reconstruction of 2-D/3-D building shapes from this class of data. Moreover, the developed methods are applicable not only to TomoSAR point clouds, as demonstrated in this paper, but also to unstructured 3-D point clouds generated from a different sensor with a similar configuration (i.e., oblique geometry) and with both low and high point densities.

The aforementioned contributions allow for a completely automatic (but parametric) reconstruction of 2-D/3-D building shapes from TomoSAR point clouds in larger areas.

The remainder of this paper is structured as follows. Section II provides the basic idea of the whole approach. Section III presents the proposed methodology for detection/extraction of building points. Section IV then presents in detail the processing steps that are employed for reconstruction of building footprints. In Section V, the experimental results obtained from the TomoSAR point cloud generated from a TerraSARX
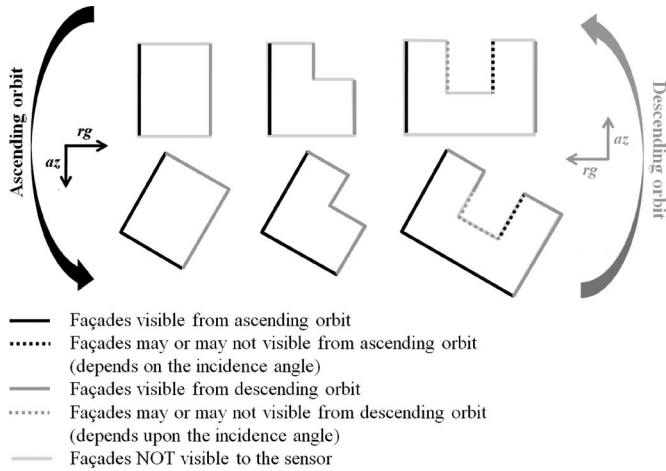
Fig. 2. Illustration of facade visibility problems in data available from single- and/or multiview perspective (i.e., data from both ascending and descending orbits). Six polygons depict the 2-D building footprints or outlines (top view) of three buildings in two different orientations. Due to side-looking geometry, not all facades are visible to the sensor. If we consider data available from the ascending orbit only, it can be easily inferred that there is no possibility to reconstruct the shape of the footprint for any building as oriented in the first row. In the second row, only the first building could be reconstructed if we impose certain strict constraints (e.g., symmetric or rectangular-shaped footprint assumption). However, such a constraint is not always feasible due to a variety of building shapes, and therefore, for the latter two buildings (for which two facades are visible), the complete shape of the footprint could only be reconstructed if either the data from multiple views are available or enough roof points exist, which could be exploited together with the reconstructed facades (see Fig. 3). If data from the descending orbit are also incorporated, it could be seen that the building footprints in the second row could be almost completely reconstructed. However, even in this case, it is still difficult to reconstruct building footprints in the first row due to lack of data on facades that are not visible to the sensor.
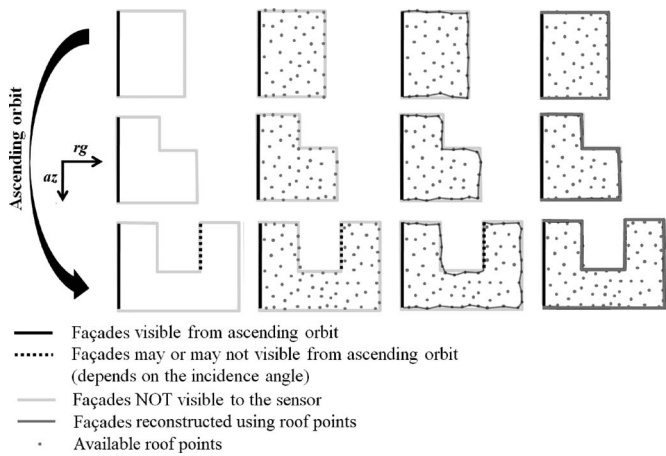

Fig. 3. Illustration of the solution to the facade visibility problems from the single-view perspective, as proposed in this paper. The first column contains the same three buildings oriented as in the first row of Fig. 2. Assuming that enough facade points available, the facades shown by black polylines could be reconstructed. Moreover, if enough roof points, which are represented by gray dots in the second column, are available, the complete shape of the building footprints could be reconstructed by fusing the rough shape obtained by connecting roof points occurring at the building boundary together with the reconstructed facades, as shown in the third column. Finally, if the building shape is rectilinear, constraints could be added to yield geometrically correct and visually aesthetic building shapes (fourth column).

high-resolution spotlight data stack (ascending orbit only) are presented and validated. Finally, in Section VI, conclusions are drawn, and future perspectives are discussed.
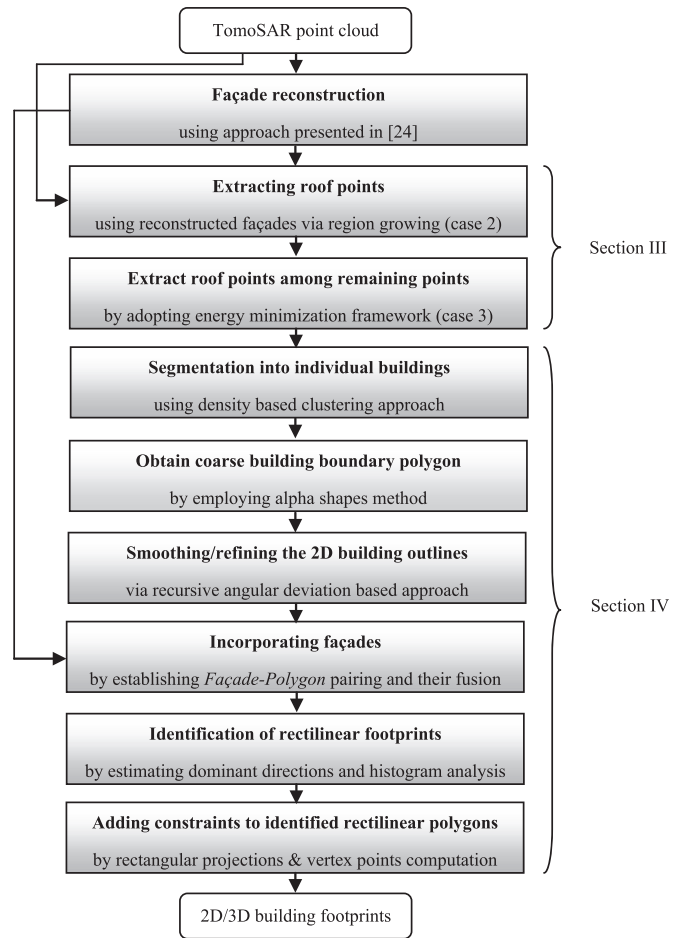

Fig. 4. Block diagram of the proposed approach.

II. METHODOLOGICAL OVERVIEW OF THE RECONSTRUCTION PROCEDURE

Fig. 4 shows the block diagram of the processing steps involved in the complete methodology. The whole processing begins by first identifying the probable building regions. This is accomplished by incorporating information pertaining to facades as prior knowledge (i.e., regions corresponding to higher point density indicate probable facade regions). Thus, building facade points are first extracted, segmented to points belonging to individual facades, and further reconstructed. Detailed processing schemes are described in [24] and [25]. The reconstructed facades are used to select seed points, from which a minimum-height-constraint surface-normal-based region-growing procedure is adopted to extract available roof points. Then, roof points without the support of facade points (i.e., case 3) are further extracted from the remaining points by formulating the extraction problem into an energy minimization framework. Automatic segmentation of detected/extracted building points is then obtained by clustering points belonging to individual buildings. Later, boundary/outline polygons (or footprints) are reconstructed and refined/smoothed for each individually segmented building cluster. Afterward, robust fusion of legitimate *facade-Polygon* pairs is carried out to improve the geometrical accuracy of the refined footprints. Finally, after the identification of rectilinear
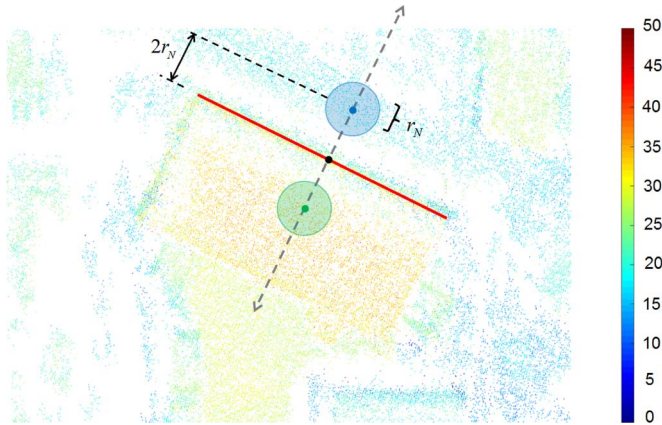
Fig. 5. Graphical illustration of the seed point selection procedure. The top view of 3-D TomoSAR points of a small region containing one building is depicted. The height (in meters) of each TomoSAR point is color coded. The red line shows the reconstructed facade segment with the black point as its midpoint. The green and blue points show the selected points in two opposite orthogonal directions with respect to the reconstructed facade (depicted as gray dotted lines). r_N is the cylindrical radius used to determine local neighbors (i.e., points within shaded circles in respective colors). The local mean height of the blue and green points is computed, denoted as m_b and m_g, respectively. Seed point s_p is chosen as the maximum height point among the local neighbors of blue points if $m_b > m_g$ or among the neighbors of green points if $m_g > m_b$, where the latter is true in the depicted case.

footprints, rectangular constraints are inserted to yield geometrically correct and visually aesthetic building shapes.

Next, we explain in detail the procedures of the proposed approach in the following two dedicated sections.

III. BUILDING DETECTION

A. Selection of Seed Points

The reconstructed facades provide direct hints for the desired building regions. The idea is then to search in the nearby vicinity of the reconstructed facade to select seed points and later employ a surface-normal-based region-growing algorithm to extract existing roof points. In this regard, the seed points are selected by first computing the midpoint of the reconstructed facade segment and then analyzing the local height statistics in the two opposite directions perpendicular to the reconstructed segment. Fig. 5 provides the graphical illustration of the seed selection procedure.

B. Region-Growing Procedure

Extracted seed points from the above procedure are then used in the region-growing procedure to extract existing roof points based on the similarity of their surface normals (robustly estimated [31]), i.e., points having an angular difference between their surface normals less than $\theta_{normals}$ are added to the grown cluster. However, using surface normals only could also add ground points during the growing procedure. An example of such a case will be where both the roof and the neighboring ground are flat. A minimum height constraint is therefore needed and is incorporated in the growing procedure to restrict the addition of points below a certain height $h_{\min}$, which is adaptively computed for each seed point as $m_p +$ (height of $\mathbf{s}_p - m_p) \cdot fac$, where $m_p = m_b$ if $m_g > m_b$ or

$m_p = m_g$ if $m_b > m_g$ (see Fig. 5). The factor fac (empirically set to 0.55 in this work) can be adjusted between 0 and 1 to adjust $h_{\min}$.

The above procedure extracts the existing roof points only for those buildings whose facades can be determined/reconstructed. However, as already mentioned, still, there exist cases when no or very few points are available on the building facades, rendering the detection of facade points/regions very difficult. Moreover, if data are acquired from one viewing angle, e.g., ascending orbit only, the facades of lower buildings could get fully (or partly) occluded due to the presence of nearby higher building structures. As a consequence, the corresponding roof points will be misdetected. To cope with the aforementioned issue, we sequentially formulate the building detection problem among the remaining points into a simple energy minimization framework to extract those building/roof points where no information pertaining to facades could be exploited (i.e., resolve case 3).

C. Formulation of Energy Equation

The detection/extraction of building roof points from the remaining point cloud can be formulated in terms of an energy minimization problem as follows [32]–[35]:

$$E(l) = \sum_{p_i \in \mathbf{P}} D_{p_i}(l_{p_i}) + \sum_{(p_i, p_j) \in \mathbf{N}} w_{p_{ij}} \delta\left(l_{p_i} \neq l_{p_j}\right) \quad (1)$$

where $\mathbf{P} = \{p_i | i = 1, \dots, n\}$ denotes the set of remaining n 3-D points; $\mathbf{N}$ denotes the set containing pairs of neighboring points (p_i, p_j); l_{p_i} represents the label assigned to point p_i, i.e., roof or nonroof point; $\delta(\cdot)$ is an indicator function; and $w_{p_{ij}}$ is the weight of each pair of neighboring points (p_i, p_j) and is defined as $w_{p_{ij}} = \exp(-\|\mathbf{p}_i - \mathbf{p}_j\|)$. The second summation term in (1) ensures spatial smoothness, i.e., it favors consistent labeling between the neighboring point pairs, whereas the first term in (1) denotes the data discrepancy term that measures the suitability/unsuitability of a particular labeling l assigned to the set of points $\mathbf{P}$. It is defined as [32]

$$D_{p_i}(l_{p_i}) = \begin{cases} (1 - h_{p_i}) + \eta \cdot r_{p_i}, & \text{if } l_{p_i} = \text{building roof point} \\ h_{p_i} + \eta \cdot (1 - r_{p_i}), & \text{if } l_{p_i} \neq \text{building roof point} \end{cases}$$
$$(2)$$

where r_{p_i} is the plane residual distance feature computed for each point p_i by locally fitting the robust RANSAC plane among its local neighbors, and h_{p_i} is the differential height of point p_i obtained after subtracting the height of the underlying terrain, which is approximated by fitting a cubic polynomial surface to the nonbuilding/ground points via the robust least absolute residual (LAR) method, as explained in the following section. Both features h_{p_i} and r_{p_i} are normalized to the scale of 0–1 by adopting the forms [32]: $h_{p_i} = \min(1, h_{p_i}/\varepsilon)$ and $r_{p_i} = \min(1, r_{p_i}/r_N)$, where ε is the tuning factor adjusting the sensitivity of the height feature [i.e., it ensures that all points having relative heights greater than ε provide a minimum data discrepancy term in (1)], and r_N is the radius size used to extract local neighbors v_c. η is the relative importance factor for the feature r_{p_i}, i.e., η defines the relative importance of r_{p_i}

with respect to h_{p_i}. The value of η less than 1 (e.g., $\eta = 0.5$ used in this work) means more importance is given to the height feature. This is reasonable since r_{p_i} essentially depicts the degree of planarity and is lower for both building roofs and flat ground surfaces. Thus, to differentiate between points on a flat planar ground segment, e.g., parking area and on the flat roof, their heights become the only discriminative factor in the labeling process and, therefore, should be given relatively higher importance. Moreover, planar objects with higher height are more probable to be part of the building structure than to the ground surface.

D. Approximation of Terrain Height

The remaining point cloud $\mathbf{P}$ mentioned in Section III-C may contain roof points belonging to buildings for which no or very few facade points are available. To adopt the above energy-based formulation, the height of the underlying terrain surface required in (2) is approximated by fitting a cubic polynomial surface to the nonbuilding/ground points via the robust LAR method. Ground points are extracted via successive reduction of nonground points in the remaining point cloud $\mathbf{P}$. This is done by adopting the following sequence of steps.

- Determine the local height difference of each point by taking the difference between the maximum and minimum height of points among its neighbors.
- Identify those points whose local height jump is greater than 5 m. These higher jump points are referred to as transition points.
- Cluster these transition points and, for each transition cluster having at least ten points, begin a region-growing procedure (similar to that explained in Section III-B).
- Probable ground points are then extracted by removing all the grown regions from the set of remaining points $\mathbf{P}$.

It is worth mentioning here that the grown regions can also be incorporated into the set of extracted roof points, as depicted in [36]. However, due to gaps in the data and localization errors of TomoSAR, it is still possible that few buildings remain undetected. Formulating the problem into the energy minimization framework helps us to detect these buildings.

E. Minimization via Graph Cuts

The above energy formulation in (1) is solved (minimized) via graph-cut-based optimization library using the $\alpha\beta$-swap move algorithm [34], [35], [37], [38]. The minimum energy corresponds to the labeling l such that higher planar points are detected as building roof points. Combining them with the set of roof points extracted in Section III-B via facade information completes the extraction procedure.

IV. BUILDING SHAPE/FOOTPRINT RECONSTRUCTION

A. Segmentation Into Individual Buildings

The extracted building points are segmented such that each cluster represents points from an individual building. This is done by means of density connectivity [39], i.e., two points

are considered to be directly density connected to each other if one point is in the neighborhood vicinity of the other point. If the two points are not directly connected to each other, still, they can be density connected to each other if there is a chain of points between them such that they are all directly density connected. Thus, starting from a point, all points that are density connected to each other are clustered into a single group representing an individual building. These clustered points are then removed, and the procedure is repeated for the remaining points until all the points are assigned to a particular cluster.

B. Coarse Building Footprint

The reconstruction of building shapes is initially obtained by employing alpha shapes (or α-shape, i.e., generalization of convex hull) around each segmented building [40]. This results in vertices that describe the coarse 2-D polygonal boundary of the building footprint. The reconstructed shape depends on a particular value of α, which has to be carefully chosen since it controls the model complexity. For instance, an overlarge α could make it difficult to follow concave polygonal shapes, e.g., an L-shaped building. In [41], it is recommended that α be chosen as twice the mean Euclidean point distance among the building roof points to produce a reliable building shape, including smaller structures.

C. Refinement of Alpha Shape Vertices by Recursively Analyzing Angular Deviations

The alpha shape method provides good initial estimates of building outlines. However, due to lesser point density of TomoSAR points, alpha shapes only define the coarse outline of an individual building. The resulting polygons are therefore irregular and need to be refined/regularized.

If we denote $V_{alpha} = \{\mathbf{V}_{i=1,\ldots,N}\}$ as a set containing N matrices of building polygons returned by the alpha shape algorithm and $\mathbf{V}_j$ with $(j \in i)$ as the matrix containing 2-D vertices of the initial alpha polygon of the jth building, then inspired by the work in [41] (although different), the recursive procedure provided in Table I is adopted to refine the coarse reconstructed building footprints returned by the alpha shape algorithm.

The procedure in Table I begins by computing the angular deviations at each vertex point of the alpha polygon as

$$\beta_j = \begin{cases} \theta_j, & \text{if } \theta_j \leq 90° \\ 180 - \theta_j, & \text{if } \theta_j > 90° \end{cases} \text{ with } \theta_j = \cos^{-1}\left(\frac{\mathbf{dv}_j \cdot \mathbf{dv}_{j+1}}{\|\mathbf{dv}_j\|\|\mathbf{dv}_{j+1}\|}\right) \tag{3}$$

where "$\cdot$" denotes the dot product, and $\mathbf{dv}_j$ is the direction vector computed at each edge formed by connecting two consecutive vertices $\mathbf{v}_j$ and $\mathbf{v}_{j+1}$ of the polygon $\mathbf{V}_{prev}$ (initialized to $\mathbf{V}_j$). Steps 4 and 5 ensure that all those vertices (or edges) having angular deviations less than the threshold θ_{ang} are removed. $\mathbf{V}_{new}$ and $\mathbf{V}_{prev}$ are then compared, and the process repeats itself if any vertex is removed in the current recursive iteration, i.e., $\mathbf{V}_{new}$ and $\mathbf{V}_{prev}$ do not contain the same number of elements. Finally, the process terminates when there is no further removal of vertices.

TABLE I
REFINEMENT OF ALPHA SHAPE POLYGONS

Given: Alpha shape polygon of jth building $\mathbf{V}_j$.

1. Initialize: $\mathbf{V}_{prev} = \mathbf{V}_j$

2. **while** (true)
3. 　　Compute angular deviation matrix $\boldsymbol{\beta}$
4. 　　Identify indices in $\boldsymbol{\beta}$ greater than θ_{ang}
5. 　　Extract vertices that corresponds to the identified indices in $\mathbf{V}_{prev}$ and assign them to $\mathbf{V}_{new}$
6. 　　**if** number of elements in $\mathbf{V}_{prev}$ equals number of elements in $\mathbf{V}_{new}$
7. 　　　　break
8. 　　**else**
9. 　　　　$\mathbf{V}_{prev} = \mathbf{V}_{new}$
10. 　　**end if**
11. **end loop**
12. $\mathbf{V}_{j_{refined}} = \mathbf{V}_{new}$

D. Incorporating Reconstructed facades

To improve the geometrical accuracy of the footprints, the reconstructed facades are fused with the refined building polygons. For this purpose, the facade associated to each refined building polygon is categorized into the following two types:

- type-I facade: facade fully or partly inside the refined polygon;
- type-II facade: facade lying completely outside but associated to the refined polygon.

The aforementioned two facade types are fused with the refined building polygon in slightly different manners, as will be explained later.

a) Identification of legitimate facade-polygon pair (facade-polygon pairing): To achieve fusion of reconstructed facades with the refined building polygons, the foremost task is to identify the association of each facade to its respective building polygon.

Identification of type-I facades is easily achieved by checking if the endpoints of the reconstructed facades lie inside the polygon. Thus, if both or at least one of the facade endpoints lie inside the building polygon, it is categorized as a type-I facade.

To identify facades of type II, the following procedure is adopted.

1) First, the midpoint of the reconstructed facade is computed, and two points are chosen in opposite directions orthogonal to the reconstructed facade at a distance d from the midpoint.
2) Compute intersections of $line_1$ and $line_2$ with all the building polygons. Here, $line_1$ denotes the line segment formed by connecting the midpoint to one of the chosen points, and similarly, $line_2$ is the line segment formed by connecting the midpoint to the other opposite point.
3) If there exists an intersection of $line_1$ or $line_2$ with any of the building polygons, the facade is assigned to the polygon with which the intersection occurs. In case there

are more than one line–polygon intersections or both $line_1$ and $line_2$ intersect with different polygons, the facade is assigned to the polygon having the intersection point nearest to it.

Implementation-wise, steps 1–3 are performed in a recursive manner. That is, d is initialized to 1 m, and steps 1–3 are carried out. If there exists no line–polygon intersection (i.e., the facade is not assigned to any polygon), the procedure repeats itself, but this time, d is incremented by 1 m. The recursion stops if either the facade is assigned to any polygon or distance d exceeds a certain threshold, which is set to a fixed value of 20 m in this work. Thus, a facade is only associated/paired to any building polygon if it lies at a distance less than 20 m; otherwise, it is regarded to have no polygon associated to it (i.e., categorized to case 1).

b) Fusion of reconstructed facades: Similar to earlier notation, let us denote $V_{refined} = \{\tilde{\mathbf{V}}_{i=1,\ldots,N}\}$ as a set containing N matrices of refined building polygons, with $\tilde{\mathbf{V}}_j = \{\mathbf{v}_{k=1,\ldots,m}\}$ $(j \in i)$ being the matrix containing 2-D vertices of the jth refined polygon having m vertices and $f_{r=1,\ldots,s}$ being the corresponding s number of (paired) reconstructed facades. Now, if the building polygon, formed by connecting vertices of $\tilde{\mathbf{V}}_j$, is interpreted as a graph, then we may define a path P_t for any particular facade f_t as a path consisting of a polygonal chain of vertices that connect two points $\mathbf{v}_{ta}$ and $\mathbf{v}_{tb}$ lying on the graph/polygon. That is, the polygonal segment comprising of all the points of polygon $\tilde{\mathbf{V}}_j$ within the interval $[\mathbf{v}_{ta}, \mathbf{v}_{tb}]$ defines path P_t. $\mathbf{v}_{ta}$ and $\mathbf{v}_{tb}$ denote points on the building polygon, which are nearest to the two endpoints of the particular reconstructed facade f_t. Since, in our case, the polygon is nonintersecting (or simple), it thus renders only two distinct paths to exist, which are referred to as $P_{t\_shortest}$ and $P_{t\_longest}$ [see Fig. 6(b)]. If the path length of P_t is denoted as $PL_{\mathbf{v}_{ta}\mathbf{v}_{tb}}$, then P_t is $P_{t\_shortest}$ only if $PL_{\mathbf{v}_{ta}\mathbf{v}_{tb}} < TL/2$, where TL is the total path length (i.e., perimeter) of the polygon.

$P_{t\_shortest}$ is further classified into two types: positive path P^+ and negative path P^-. If we denote the set of points on the polygons that are nearest to the endpoints of all facades of the same building other than f_t as $\mathbf{K}_t$ (i.e., $\mathbf{K}_t$ contains points nearest to the endpoints of facades $f_{r'=1,\ldots,t-1,t+1,\ldots,s}$ such that $r' \cup t = s$), then $P_{t\_shortest}$ of the reconstructed facade f_t is defined to be P^+ if the set of points $\mathbf{P}_{t\_shortest}$ belonging to path $P_{t\_shortest}$ does not contain any element of $\mathbf{K}_t$, i.e., $\mathbf{P}_{t\_shortest} \cap \mathbf{K}_t = \emptyset$. Thus, all facades whose paths are identified as positives are incorporated in the fusion process, whereas facades having negative paths are not considered any further. A sample illustration of the concept of a positive path is presented in Fig. 6(c).

Table II provides the complete procedure to incorporate facades of both types with the refined building polygon.

Fig. 7 practically illustrates the procedure of fusing both types of facades. Note that there lie some differences in the computation of $\mathbf{v}'_{ta}$ and $\mathbf{v}'_{tb}$ for type-I and type-II facades (steps 8, 9, and 17 in Table II). The reason for this is due to the fact that point density on building roofs is quite varying and can contain gaps in between. This could lead to *under reconstruct*
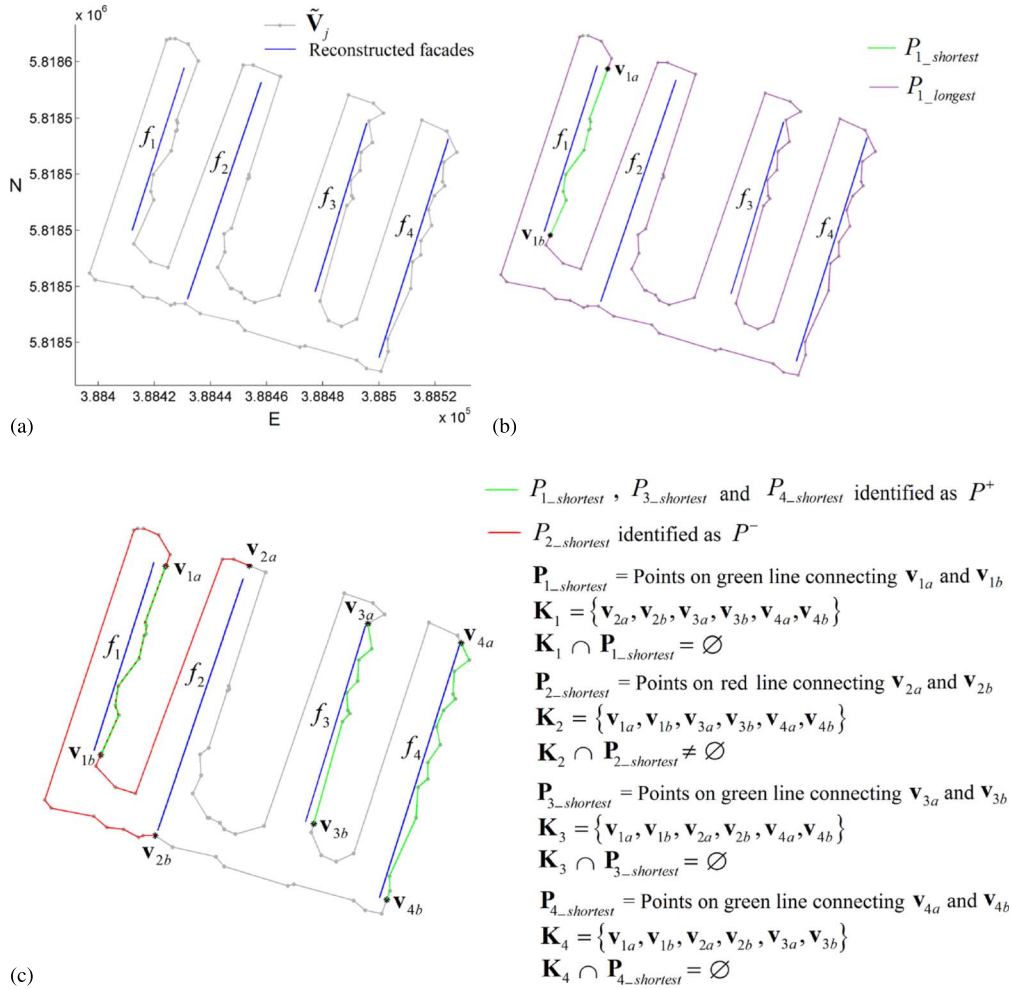
Fig. 6. Procedure describing the concept of $P_{t\_shortest}$ and P^+. (a) Refined polygon of one particular jth building $\tilde{\mathbf{V}}_j$ having vertices $\mathbf{v}_{k=1,\ldots,m}$ with four reconstructed facades f_1, f_2, f_3, and f_4 that are to be incorporated. (b) Sample illustration of the concept of shortest and longest paths associated for a particular facade f_1. $\mathbf{v}_{1a}$ and $\mathbf{v}_{1b}$ denote the closest points on the polygon/graph $\tilde{\mathbf{V}}_j$ to the two endpoints of facade f_1, respectively. (c) Sample illustration of the concept of positive path P^+. $P_{2\_shortest}$ is identified as P^- as there exist points in $\mathbf{K}_2$ that are also present in $\mathbf{P}_{2\_shortest}$.

of the building footprint, i.e., part of the building roof region could not be reconstructed due to unavailability of points. The presence of type-II facades implicitly validates this plausible phenomenon, and therefore, fusion of refined polygons by fully incorporating the reconstructed facades (of type II only) results in improved overall accuracy of reconstruction. Doing the same for type-I facades, on the other hand, may affect the footprint polygon in the presence of facades belonging to inner building structures. Thus, only the orientation of type-I facades is essentially incorporated by the proposed procedure (steps 8 and 9 in Table II). In addition, steps 12–15 in Table II also pose a condition $C1$ for type-I facades such that they do not take part in the fusion process if the change in area of the polygon after incorporating the particular facade is greater than the certain fraction a_f (fixed to 0.15 in this work) of the previous polygonal area. Thus, using condition $C1$ together with the method of type-I facade fusion, it is ensured that facades belonging to the inner structures of the building do not interfere during the fusion procedure, or in other words, only facades that are exterior and define the building outlines are utilized.

E. Identification of Rectilinear Footprints

The next step in the reconstruction procedure is to identify if the building is composed of two or more than two dominant directions. If the building polygon is composed of only two dominant directions orthogonal to each other, rectilinear constraints are added to derive geometrically correct and better visually looking building shapes.

a) Estimation of principal direction: The decision of identifying a rectilinear building is based on its estimated dominant/principal direction. For this purpose, building polygons belonging to case 2 (i.e., having one or more reconstructed facades associated to each polygon), the principal direction is easily determined by assigning it to the direction vector computed by subtracting the endpoints of the longest reconstructed facade paired to it. For case-3 building polygons, the principal direction is directly estimated from the polygon itself. Since no facade is associated to them, a weighted method based on polygonal edge lengths is employed to estimate the two orthogonal principal directions of the building. The basic idea is to give weight to each edge of the polygon according to its

TABLE II
PROCEDURE TO FUSE FACADES OF BOTH TYPES

Given: Refined polygon of jth building $\tilde{\mathbf{V}}_j$ & s reconstructed façades $f_{r=1,\dots,s}$ belonging to the same jth building

1. $\mathbf{V}_{prev} = \tilde{\mathbf{V}}_j$

2. **for** $t = 1$ to s (i.e., total number of reconstructed façades associated to this building polygon)

3. Determine points $\mathbf{v}_{ta}$ and $\mathbf{v}_{tb}$ by computing points on the polygon that are nearest to the two endpoints of the façade f_t

4. Build a matrix $\mathbf{K}_t$ specific to the façade f_t that contains points on the polygon that are nearest to the endpoints of all other façades

5. Determine $P_{t\_shortest}$

6. **if** $P_{t\_shortest}$ is also P^*

7. **if** façade is identified as type I

8. Compute midpoint of the line segment L_t formed by connecting the two points $\mathbf{v}_{ta}$ and $\mathbf{v}_{tb}$

9. Determine two new points $\mathbf{v}'_{ta}$ and $\mathbf{v}'_{tb}$ by projecting $\mathbf{v}_{ta}$ and $\mathbf{v}_{tb}$ onto another line parallel to the respective façade f_t but passing through the midpoint of L_t (see Figure 7(c))

10. Replace vertices of $\mathbf{V}_{prev}$ within $P_{t\_shortest}$ (i.e., points in $\mathbf{P}_{t\_shortest}$) by $\mathbf{v}'_{ta}$ and $\mathbf{v}'_{tb}$ and store the result in the matrix $\mathbf{V}_{new}$

11. Compute area of old $\mathbf{V}_{prev}$ and new $\mathbf{V}_{new}$ polygons denoted as $A_{\mathbf{V}_{prev}}$ and $A_{\mathbf{V}_{new}}$ respectively

12. Check the condition $C1$: $\left| A_{\mathbf{V}_{prev}} - A_{\mathbf{V}_{new}} \right| \le a_f A_{\mathbf{V}_{prev}}$

13. **if** $C1$ is satisfied

14. $\mathbf{V}_{prev} = \mathbf{V}_{new}$

15. **end if**

16. **else**

17. Assign $\mathbf{v}'_{ta}$ and $\mathbf{v}'_{tb}$ to the respective endpoints of the reconstructed façade (see Figure 7(c))

18. Replace vertices of $\mathbf{V}_{prev}$ within $P_{t\_shortest}$ (i.e., points in $\mathbf{P}_{t\_shortest}$) by $\mathbf{v}'_{ta}$ and $\mathbf{v}'_{tb}$ & store the result in the matrix $\mathbf{V}_{new}$

19. $\mathbf{V}_{prev} = \mathbf{V}_{new}$

20. **end**

21. **end if**

22. **end loop**

relative length (with respect to total polygon length) and the angular deviation it makes with a particular direction vector $\mathbf{dv}_q$. $\mathbf{dv}_q$ is a candidate for one of the two orthogonal principal directions and is rotated within a certain interval to minimize the following objective function [42]:

$$PD = \sum_{i=1}^{n} g_1(l_i) g_2\left(\varphi_i(\beta_i, \phi_{\mathbf{dv}_q})\right) \quad (4)$$

where n is the total number of vertices of the polygon, and $\beta_i (0 \le \beta_i \le 90°)$ is the angular deviation of each edge l_i with respect to the direction vector $\mathbf{dv}_q$. β_i is computed similar to (3) with the difference that the two direction vectors are $\mathbf{dv}_j$ and $\mathbf{dv}_q$ instead of direction vectors of consecutive edges $\mathbf{dv}_j$ and $\mathbf{dv}_{j+1}$. $\phi_{\mathbf{dv}_q}$ is the anticlockwise rotation angle that $\mathbf{dv}_q$ makes with the unrotated coordinate system. $\varphi_i(\cdot)$ is a function that maps the angular deviations β_i to one of the two orthogonal directions (or axes) as defined by direction vector $\mathbf{dv}_q$ and its corresponding normal vector. It is computed as

$$\varphi_i\left(\beta_i, \phi_{\mathbf{dv}_q}\right) = \begin{cases} \beta_i, & \text{if } \beta_i \le 45° \\ 90 - \beta_i, & \text{if } \beta_i > 45°. \end{cases} \quad (5)$$

Both $g_1(\cdot)$ and $g_2(\cdot)$ are the weighting functions. $g_1(\cdot)$ assigns weight to each edge based on its relative length with

respect to the overall length of the polygon edges. It is constructed such that edges with longer lengths contribute less in (4) as compared with shorter edge lengths. The following linear function is used to describe $g_1(\cdot)$:

$$g_1(l_i) = 1 - \frac{l_i}{\sum\limits_{i=1}^{n} l_i}. \quad (6)$$

Similarly, $g_2(\cdot)$ assign weights to each edge based on its φ_i value. The assignment of weight is directly proportional to φ_i, i.e., lower weight is given to an edge with lower φ_i, inferring that edges close to one of the two orthogonal directions are given less weight as compared with those that are deviating. Since the span of φ_i for each edge is defined to be within the interval $[0, 45°]$, $g_2(\cdot)$ is thus computed by adopting the following linear function:

$$g_2\left(\varphi_i(\beta_i, \phi_{\mathbf{dv}_q})\right) = \frac{\varphi_i}{45}. \quad (7)$$

The solution of (4) is obtained by rotating $\mathbf{dv}_q$ within the interval $[0, 90°]$. An optimum (or minimum) $\hat{\phi}_{\mathbf{dv}_q}$ is found by comparing PD for each $\phi_{\mathbf{dv}_q}$ value. The direction vector and its corresponding normal vector associated to the optimum (or minimum) $\hat{\phi}_{\mathbf{dv}_q}(= \arg\min_{\phi_{\mathbf{dv}_q}}(PD))$ thus describe the two orthogonal principal directions.

b) Identification procedure: Once the principal/dominant directions are determined, the following procedure is adopted for identification of rectilinear buildings.

- Determine angular difference $\beta_i (0 \le \beta_i \le 90°)$ of all the edges of the polygon with respect to the dominant/principal directions.
- Compute the histogram of these angular differences.
- Find the edges whose angular differences are within the bin intervals $[0°, 20°]$ and $[70°, 90°]$.
- Identify the polygon to be rectilinear if the total sum of the lengths of these edges is more than a certain fraction L_f of the total polygonal length TL. (L_f is fixed to 0.75 in our work, i.e., 75% of the total polygonal length TL).

F. Addition of Rectilinear Constraints

Subsequently, rectilinear constraints are added to the identified building polygons to yield much better (visually appealing) geometric building shapes. The following steps are performed to obtain rectilinear building footprint.

- Classify each edge of the building polygon such that it belongs to one of the two orthogonal principal axes based on its angular deviation (i.e., an edge is associated to that principal direction with whom the angular difference is less).
- Merge all adjacent edges that share the same class, i.e., associated to the same principal direction.
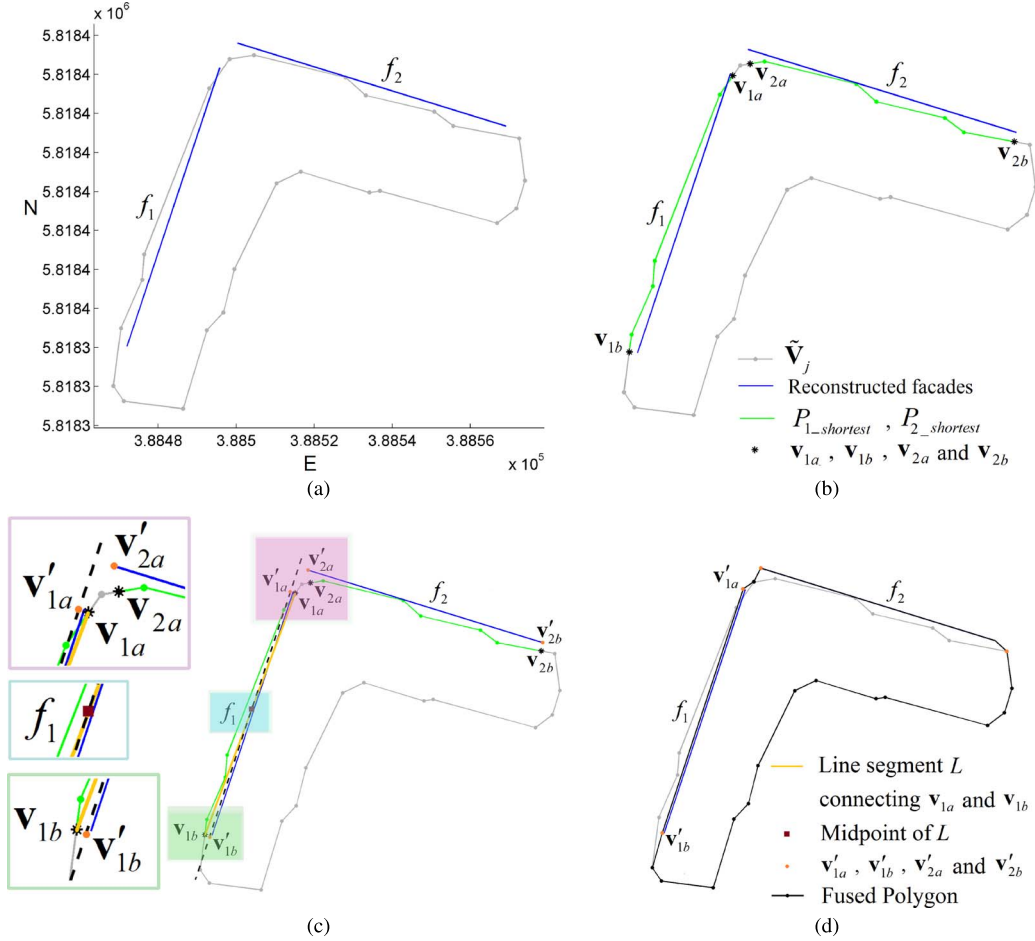
Fig. 7. Fusion of both type-I and type-II facades. (a) Refined polygon of one particular jth building $\tilde{\mathbf{V}}_j$ with two reconstructed facades f_1 and f_2 with types I and II, respectively. (b) Shortest paths (also P^+) determined for both facades. (c) Procedure of computing points $\mathbf{v}'_{ta}$ and $\mathbf{v}'_{tb}$. Note that for the type-I facade, $\mathbf{v}'_{1a}$ and $\mathbf{v}'_{1b}$ are computed by projecting $\mathbf{v}_{1a}$ and $\mathbf{v}_{1b}$ onto the black dotted line, which is parallel to the reconstructed facade f_1 and passes through the midpoint of line segment L formed by connecting points $\mathbf{v}_{1a}$ and $\mathbf{v}_{1b}$ (steps 8 and 9 in Table II), whereas for the other facade having type II, $\mathbf{v}'_{2a}$ and $\mathbf{v}'_{2b}$ are simply the endpoints of facade f_2 (step 17 in Table II). All vertices of $\tilde{\mathbf{V}}_j$ on $P_{1\_shortest}$ and $P_{2\_shortest}$ (i.e., points in $\mathbf{P}_{1\_shortest}$ and $\mathbf{P}_{2\_shortest}$) are replaced by vertices $\mathbf{v}'_{1a}$, $\mathbf{v}'_{1b}$ and $\mathbf{v}'_{2a}$, $\mathbf{v}'_{2b}$, respectively. (d) Resulting polygon in black after fusing facades f_1 and f_2 with types I and II, respectively.

- Apply rectilinear transformation to every merged polygonal edge by projecting it onto its corresponding principal axis/vector.
- Computing intersection (or vertex) points between the consecutive vertices.

Fig. 8 illustrates the mechanism of practically adding rectilinear constraints to the building polygon using the aforementioned procedure. Note that the depicted polygon contains no associated facade and, therefore, belongs to case 3. For case-2 polygons, the associated facades are first fused to the refined polygon, and if identified as a rectilinear, constraints are added to complete the reconstruction procedure.

V. EXPERIMENTAL RESULTS AND VALIDATION

A. Data Sets

To validate our approach, we tested the proposed algorithms on two different data sets. One is composed of TomoSAR point clouds generated from a stack of 25 TerraSAR-X high-resolution spotlight images covering approximately

$(900 \times 600 \approx)0.54$ km^2 area in the city of Las Vegas, USA. It contains 0.48 million points and consists of moderate-sized buildings with relatively simple geometry. To test the generality of the proposed algorithms, the other data set is composed of TomoSAR point clouds of complex building structures produced from a stack of 102 TerraSAR-X high-resolution spotlight images covering around $(1750 \times 900 \approx)$ 1.5 km^2 area in the city of Berlin, Germany. The number of points in the Berlin data set is approximately 0.52 million. Both TomoSAR point cloud data sets are generated from images taken from an ascending orbit using the Tomo-GENESIS software developed in the German Aerospace Center (DLR) [43], [44].

B. Reference Footprints

For the area of interest in Las Vegas, the reference footprints for the Las Vegas data set were acquired from CyberCity3D [45]. These footprints are highly precise with positional accuracy up to +/−15 cm and are generated using automated and
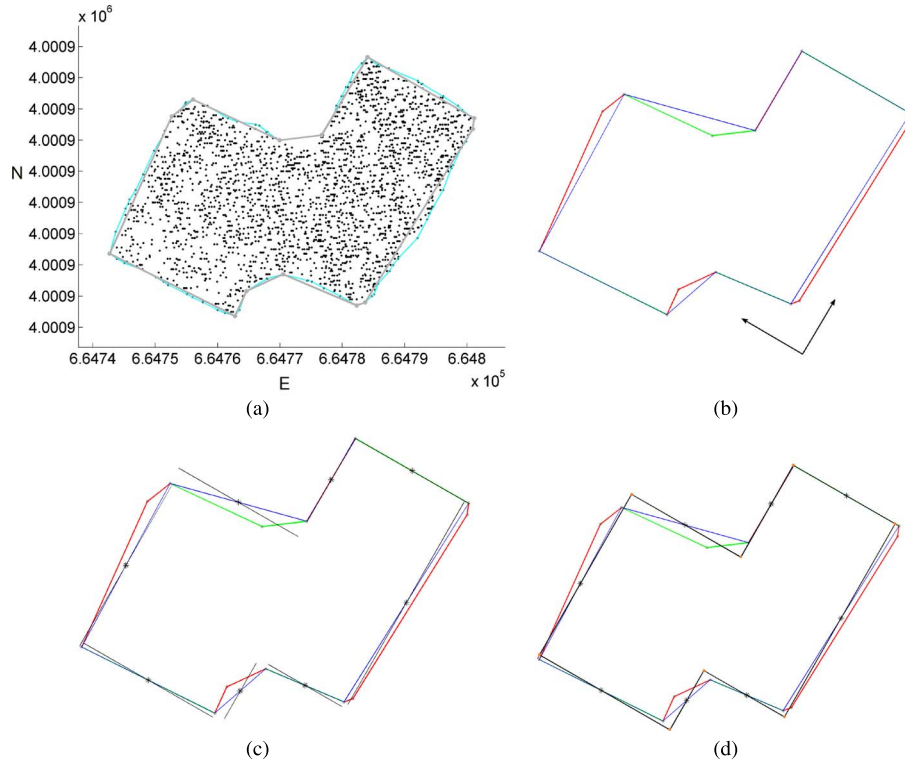
Fig. 8. Procedure to add rectilinear constraints to any one particular building polygon. (a) Black dots indicate the extracted building points, the cyan polygon represents the initial coarse outline obtained using the alpha shape algorithm, and the gray polygon depicts the refined/smoothed polygon to which rectilinear constraints are added. (b) Each edge of the refined gray polygon of (a) is segmented to belong to one of the two estimated dominant directions indicated by black arrows based on their angular deviation (i.e., an edge is associated to that principal direction with whom the angular difference is less). Segmented edges are shown in red and green. The blue polygon then depicts the reduced refined polygon of (a) by merging all adjacent edges that belong to the same principal axis. (c) Each edge of the blue polygon is then rotated/projected around its midpoint onto its corresponding principal axis. (d) Finally, the intersection (vertex) points among the adjacent edges of the projected blue polygon edges of (c) are computed for rectilinearization.
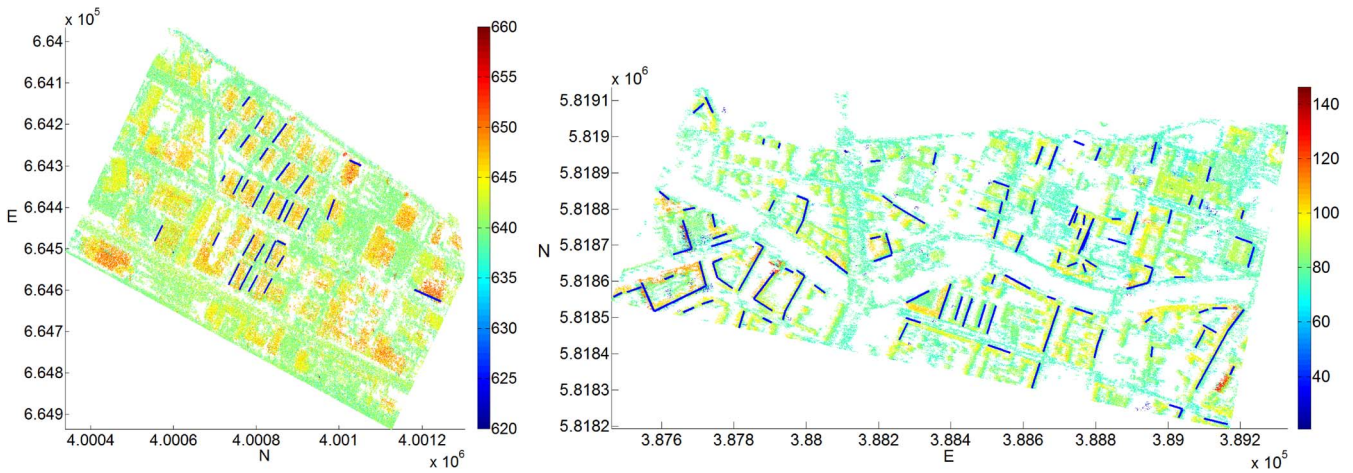

Fig. 9. Data sets. Top view of the 3-D TomoSAR points in UTM coordinates of the area of interest in (left) Las Vegas, USA and (right) Berlin, Germany. Blue lines depict the reconstructed facade segments (longer than 10 m). The height of TomoSAR points is color coded [unit: m].

semiautomated photogrammetry-based techniques with data source derived from aerial, oblique, or satellite stereo imagery.

For the Berlin data set, we compared our building extraction results to reference polygons downloaded from OpenStreetMap (OSM) [46]. Based on the concept of crowd sourcing, which involves crowd or community to effectively and efficiently fulfill a task at hand, OSM with around two million registered users (as of today and also rapidly growing) is considered to

be the most successful Volunteered Geographic Information (VGI) project [47], [48]. The OSM database contains multitude of building footprints represented as polygons with an ordered list of nodes/vertices (i.e., pairs of UTM or Latitude/Longitude coordinates according to the WGS 84 coordinate system) and is updated every day. The data are free to download and comes under the open license Open Data Commons Database License (ODbL). Since it is a VGI project, the data quality
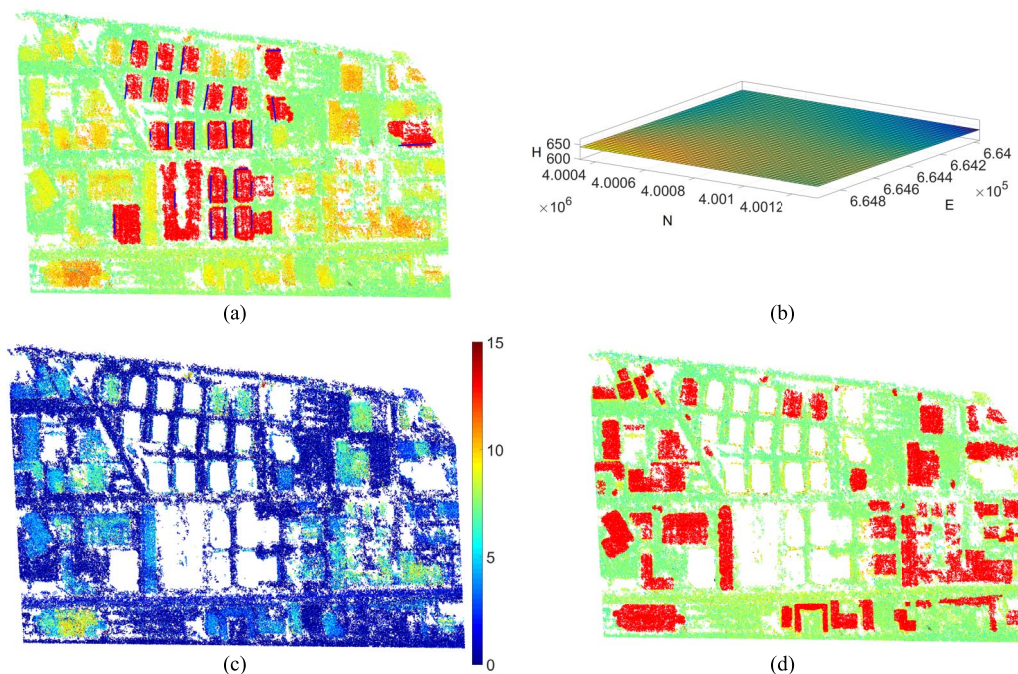
Fig. 10. Building extraction procedure. (a) Red points depict building points extracted via utilizing reconstructed facades. (b) Approximated terrain surface from the remaining set of points via successive reduction of building points and polynomial surface fitting, as explained in Section III-D. (c) Differential height computed via subtracting the terrain height from the height of each point. (d) Further building points (in red) are then extracted using the energy minimization framework with tuning parameters: $\varepsilon = 10$ m, $r_N = 5$ m.
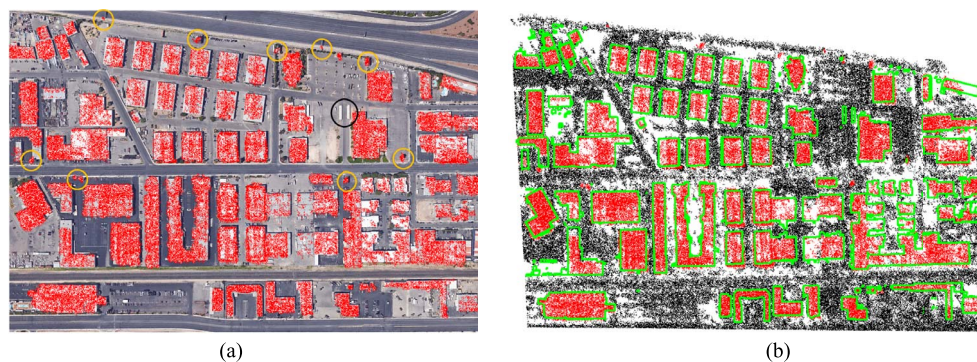

Fig. 11. Results of building extraction: Las Vegas. (a) Extracted building points in red from Fig. 10(a) and (d) are overlaid onto the optical image (Google) of the area of interest. Yellow circles indicate extracted points originating from sources such as advertisement boards and monuments. Large black circle encloses two buildings that remain undetected due to lower relative heights. (b) Finally, reference footprints (in green) overlaid onto the extracted building points. Red points are building points, whereas black points are nonbuilding points.

may vary from region to region. To this end, the first investigations regarding OSM data quality were carried out for roads [48], followed by an assessment of other attributes present in the database, e.g., lines [50] and polygonal objects [51]. Recently, the building footprints have also been evaluated for their completeness [52] and correctness [49] for various cities in Germany. The analysis of OSM data with surveying data sets reveals fairly precise positioning accuracy varying within 4 m [47], [49].

C. Results: Extraction of Building Points

Fig. 9 shows the result of applying facade reconstruction procedure over both data sets. Seed points are selected from each reconstructed facade, and the region is grown using a

surface-normal-based similarity measure with $\theta_{normals}$ set to 15°. Fig. 10(a) shows the grown region using seeds from the reconstructed facades. Later, among remaining points, terrain is approximated, and building points are extracted by adopting an energy minimization procedure, as previously explained in Section III-C [see Fig. 10(b)–(d)]. Figs. 11 and 12 show the final extracted building points in both data sets overlaid onto the optical images.

D. Validation: Extraction of Building Points

Figs. 11(b) and 12(b) show the reference polygons overlaid onto the extracted building points of the test area in Las Vegas and Berlin, respectively. It can be visually seen that the extracted building points fit very well to these reference polygons.

(a)

(b)

Fig. 12. Results of building extraction: Berlin. (a) Extracted roof points in red are overlaid onto the optical image (Google) of the area of interest. (b) Red and black points depict building and nonbuilding points, respectively. The overlaid green polygons are reference buildings downloaded from OSM [46]. Blue polygons are manually extracted buildings not present in OSM data. Gray polygons are newly constructed buildings that are not present in our data set, whereas magenta polygons are buildings that do not actually exist but are present in OSM data. Both gray and magenta polygons are not included in the evaluation.

For the Berlin data set, we found out that few buildings are missing in the OSM data set, and therefore, by analyzing the detected buildings from TomoSAR point clouds and validating using optical data, we completed few missing buildings in the OSM data set; polygons are depicted as blue polygons in Fig. 12(b). The performance of the (detection) extraction procedure in both data sets is then assessed by employing the evaluation metrics [53], [54] given in

$$
\left.
\begin{array}{c}
\text{Completeness (\%)} : comp = 100 \times \left(\frac{TP}{TP+FN} \right) \\[2mm]
\text{Correctness (\%)} : corr = 100 \times \left(\frac{TP}{TP+FP} \right) \\[2mm]
\text{Quality(\%)} : Q = \frac{comp \times corr}{comp+corr-comp \times corr} = \frac{TP}{TP+FP+FN}
\end{array}
\right\}
\tag{8}
$$

where

- TP (true positives) represents the number of detected building points inside the reference building polygons;
- FN (false negatives) represents the number of detected nonbuilding points inside the reference building polygons;
- FP (false positives) represents the number of detected building points outside the reference building polygons; and
- TN (true negatives) represents the number of detected nonbuilding points outside the reference building polygons.

TABLE III
EVALUATION STATISTICS OF THE DETECTION ALGORITHM

Datasets	Las Vegas	Berlin
TP	168065	295367
FN	10874	16269
FP	26966	50834
TN	228551	154420
Completeness (%) - $comp$	93.923	94.779
Correctness (%) - $corr$	86.173	85.316
Quality (%) - Q	81.622	81.487

The aforementioned metrics assess the overall performance of the building extraction algorithm. Completeness tells up to what percentage the algorithm has detected the roof points, whereas correctness provides a measure of correct classification. Quality combines both completeness and correctness metrics to provide an overall measure of the algorithm performance. Results of the evaluation statistics for both data sets are provided in Table III.

In both data sets, all of the buildings having relative heights (with respect to ground) of more than 5 m are detected by the extraction procedure. However, an example of two smaller buildings, for which no points could be extracted, is highlighted by a black circle in Fig. 11(a). These buildings have a footprint
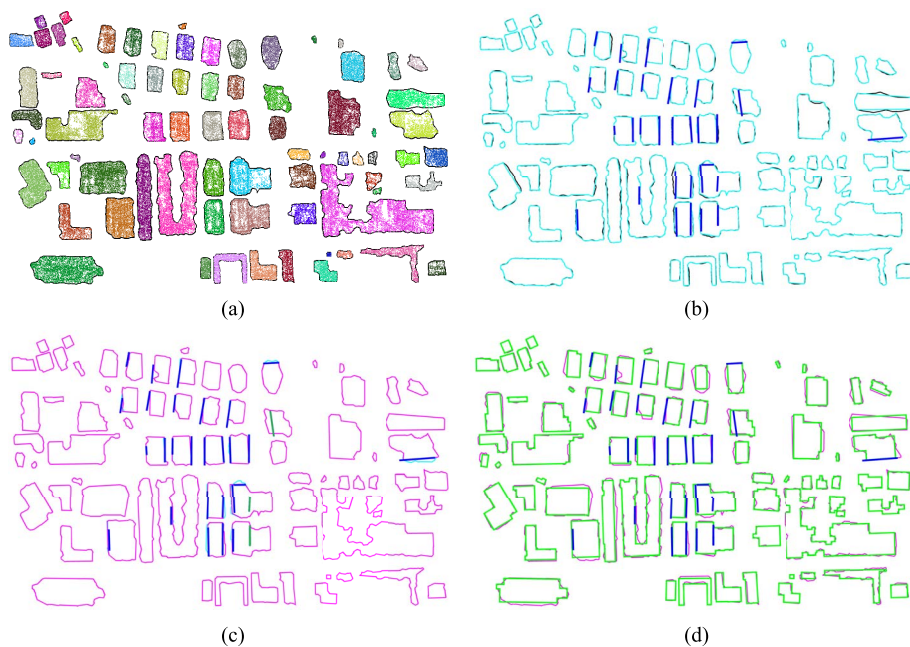
Fig. 13. Reconstruction results: Las Vegas. (a) Clustered (or segmented) building points. Black polygons V_{alpha} surrounding individual segmented building points are the initial coarse boundary/outline obtained using the alpha shape algorithm. (b) Refined (cyan) polygons $V_{refined}$ obtained after applying a recursive angular deviation approach together with 2-D reconstructed facades (depicted in blue, are overlaid onto alpha shape polygons). (c) facades are then incorporated into the refined polygons from (b) depicted in magenta, symbolized as $V_{facadefused}$. facades either identified as P^- or having condition $C1$ in Table II not satisfied are depicted in green and are not utilized during the *facade-Polygon* fusion process. (d) Final rectilinearized polygons V_{final} obtained after adding rectilinear constraints.

area of approximately $(28 \times 5 =)140$ m^2 but possess relative heights of only 2 m. Since during the extraction procedure, seed points are chosen based on local height jumps of 5 m, no seed point could thus be chosen for these buildings. As a consequence, they remain undetected.

In terms of false alarms, it is worth mentioning that it might happen that points belonging to some small vertical structures on ground (e.g., advertisement boards and monuments) also get detected during our extraction procedure. Few examples of such cases are also highlighted by yellow circles in Fig. 11(a). The reason for this occurrence of false positives is our implicit definition (or assumption) pertaining to buildings, i.e., higher points with a higher degree of planarity are detected as buildings by the extraction procedure.

E. Results: Reconstruction of Building Footprints

Extracted building points are then spatially segmented such that each cluster represents an individual building. Figs. 13(a) and 14(a) depict the result of spatially clustering points into individual buildings in the Las Vegas and Berlin data sets, respectively. The initial coarse outline of each cluster is then determined using the alpha shape algorithm. Black polygons in Figs. 13(a) and 14(a) surrounding each individual segmented building cluster depict its corresponding alpha polygon.

Refinement of the initial coarse alpha vertices is then carried out by computing the angular deviation at each vertex point. The threshold value $\theta_{ang} = 20°$ is used, which consequently removes all vertices having angular deviations less than $20°$ from their adjacent neighboring vertices. Refined or smoothed alpha polygons are then fused with the reconstructed facades. Later,

rectilinear constraints are added to the building polygons that are identified to be rectilinear. Figs. 13(b)–(d) and 14(b)–(d) depict the results of building footprint reconstruction on the Las Vegas and Berlin data sets, respectively.

F. Validation: Reconstruction of Building Footprints

To evaluate the reconstruction results, we rasterized both the reconstructed and reference polygonal footprints onto an image with pixel resolution of 1 m (i.e., one pixel corresponds to 1 m^2 spatial area). A difference image created by subtracting the reconstructed footprint image from the reference footprint image is then used to compute the commission and omission errors as follows:

$$\text{Commission error}(\%) = \frac{F_N}{A_{ref}} \times 100$$

$$\text{Omission error}(\%) = \frac{F_P}{A_{ref}} \times 100 \quad (9)$$

where A_{ref} is the area of the reference polygons, whereas F_N and F_p are the number of pixels in the difference image having values of -1 and 1, respectively.

Fig. 15 presents the common (top row) and difference (bottom row) images. The red pixels in the difference images indicate the building regions that are not reconstructed by the proposed algorithm contributing to the omission errors, whereas blue pixels are over reconstructed regions, i.e., pixels not part of the reference footprint image but present in the reconstructed image.

Table IV lists the commission and omission errors obtained for the reconstructed footprints.
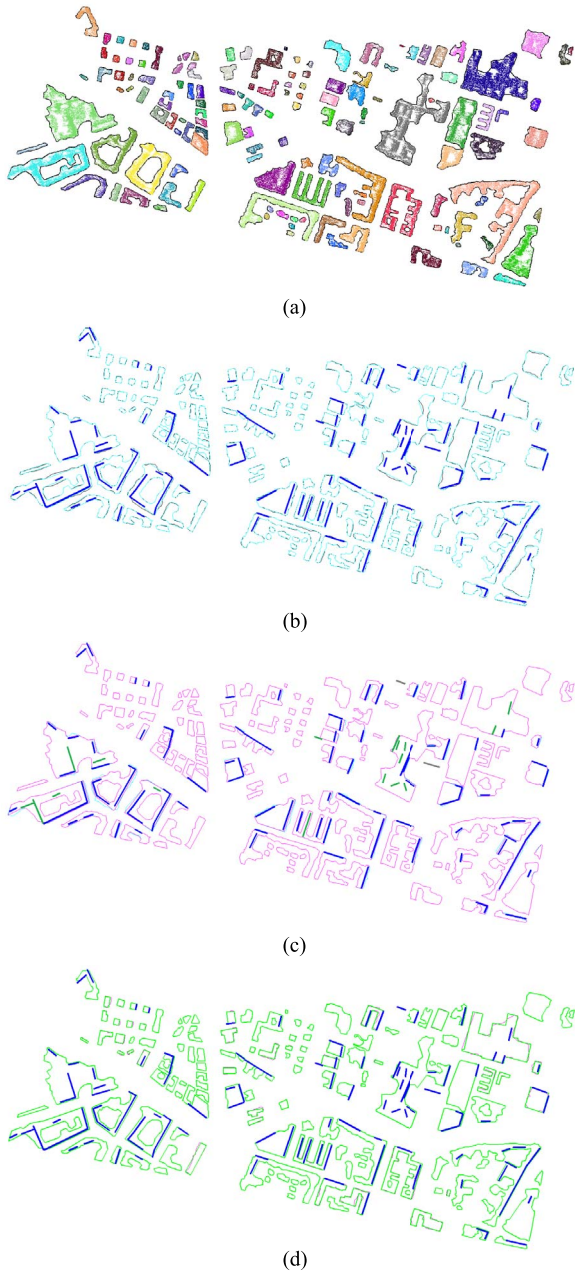
(a)

(b)

(c)

(d)

Fig. 14. Reconstruction results: Berlin. (a) Clustered (or segmented) building points. Black polygons V_{alpha} surrounding individual segmented building points are the initial coarse boundary/outline obtained using the alpha shape algorithm. (b) Refined (cyan) polygons $V_{refined}$ obtained after applying a recursive angular deviation approach together with 2-D reconstructed facades (depicted in blue, are overlaid onto alpha shape polygons). (c) facades are then incorporated into the refined polygons from (b) depicted in magenta, symbolized as $V_{facadefused}$. facades either identified as P^- or having condition $C1$ in Table II not satisfied are depicted in green, whereas facades not associated to any building polygon (i.e., case 1) are depicted in gray. Both red and gray facades are not utilized during the *facade-Polygon* fusion process. (d) Final rectilinearized polygons V_{final} obtained after adding rectilinear constraints.

Hypothetically, the reconstruction results will be improved with higher density of TomoSAR points because more points would be available for parameter estimation. Numerical experiments also demonstrated that reconstruction accuracy is better for buildings with higher density of roof points. For low-density roof regions, the reconstruction accuracy is, however, restricted

by the number of available points, which consequently reduces omission errors. A further improved-model-based approach might be helpful in this regard.

Additionally, the reconstruction errors between the final and coarse 2-D topologies (polygons) are also varying. Thus, there is a tradeoff in achieving geometrically correct footprints while simultaneously keeping the commission and omission errors in control. As evident, the best tradeoff can be obtained by incorporating facades to the coarser building polygons. However, more visually appealing results are produced by introducing rectangular constraints to the rectilinear buildings.

Finally, in Fig. 16, we present the final reconstructed building shapes/footprints in 3-D. As depicted in [25], the shown reconstructed building model can be used to refine the elevation estimates of the raw TomoSAR points. Moreover, with known deformation estimates of the scatterers, such a model can also lead to the reconstruction of dynamic city models [55] that could potentially be used to monitor and visualize the dynamics of urban infrastructure in very high level of details.

G. Discussion on Parameter Selection

Although the parameter values have been tuned according to the investigated scenes, some parameters are not strictly related to the scene under consideration and, therefore, can be easily set *a priori* by using some general rules or constraints. Moreover, most of the parameters involved have a clear physical meaning associated to them, which aids the user in incorporating it as prior knowledge on a particular scene. Based on aforementioned guidelines provided in the previous two sections, in the following, some more insights regarding the selection of parameters and their effects on the results are discussed in detail.

a) Extraction of building points: Building points are extracted by following a two-step procedure: First, the building points are extracted using a region-growing procedure. Later, among the remaining points, an energy minimization formulation is adopted to extract points belonging to lower height buildings for which no facade information was available. Parameters that control the extraction procedure are fac (controlling the minimum height constraint value $h_{\min}$), threshold on angular difference $\theta_{normals}$, the tuning factor adjusting the sensitivity of the height feature ε, the radius size r_N, and the parametric value of local height jumps.

- fac adaptively sets the value of $h_{\min}$ for each seed point during the region-growing procedure. Theoretically, the value of fac should be close to 1. The reason for this is because the height of facades that are part of the building boundary is either lower than or equal to the height of the building roof (i.e., it is lower for polyhedral roofs and equal for flat roofs with the assumption of no overhanging roof structure). Thus, setting fac equal to 1 would set $h_{\min}$ (almost) equal to the height of the facade, and therefore, only those points that are planar and have height values greater than $h_{\min}$ would be clustered for a particular seed point during the region-growing procedure. Lower fac values, on the other hand, may cause flat (or planar) ground points to be added in the growing procedure. As
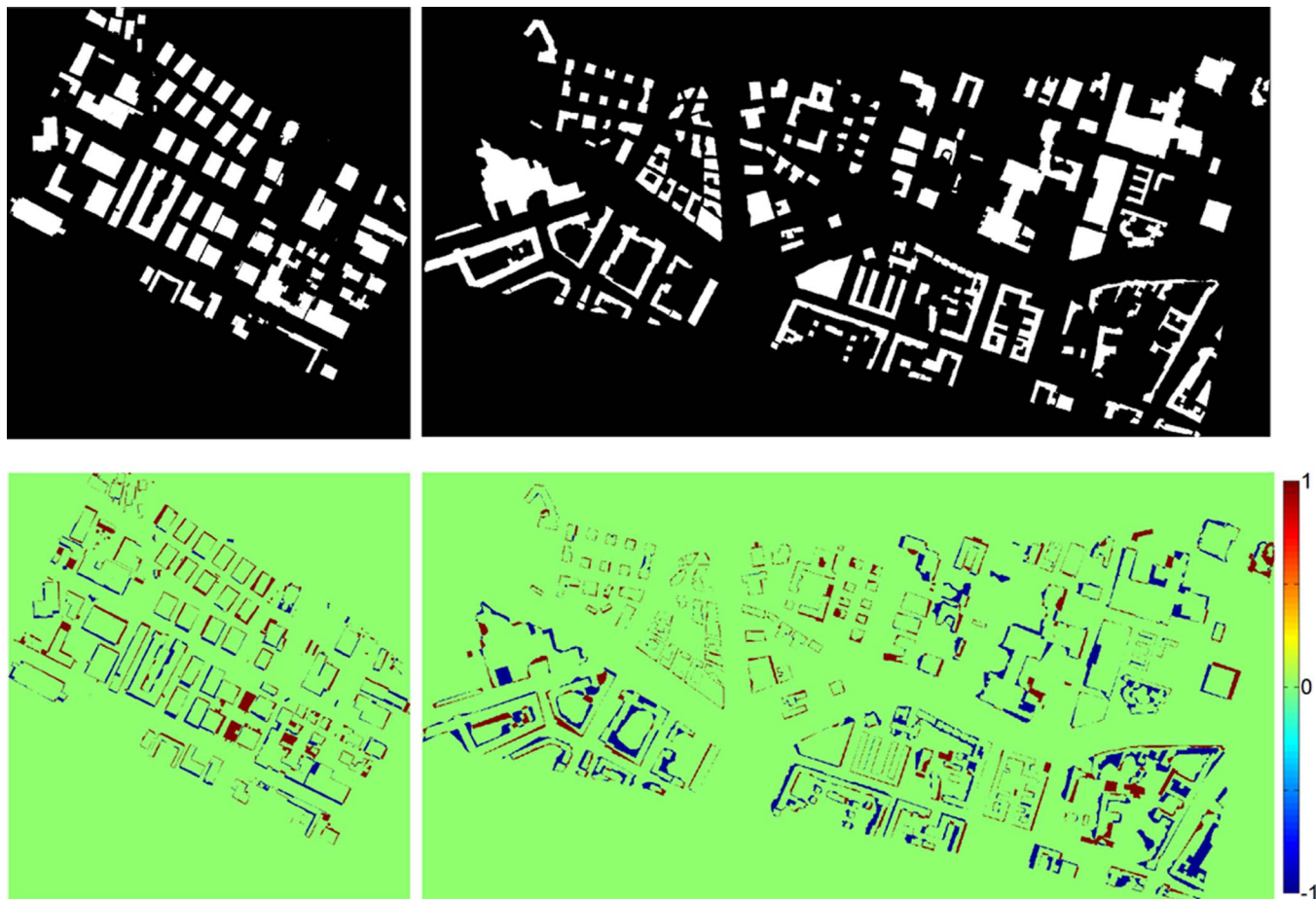
Fig. 15. Common (top row) and difference (bottom) images computed using the reference footprint image and final reconstructed footprints. The first row depicts the common regions, whereas the second row shows the difference image computed by subtracting the final reconstructed footprint image from the reference image. Note that green pixels indicate no difference between reconstructed and reference rastered images.

previously mentioned in Section I, since TomoSAR point clouds have a rather higher elevation error, the suggested range of values for fac based on empirical evaluation for different scenes is between 0.5 and 0.7. However, for other sensors giving better positioning accuracy, a higher fac value may be used.

- The angular difference threshold value $\theta_{normals}$ is mainly dependent on noise in the data. If the normals are robustly estimated, this parameter is more or less independent of both the scene and sensor configurations. The possible range of values for this parameter should be set between $10°$ and $20°$.

- ε can be easily set based on *a priori* knowledge related to the average height of the buildings in the scene. For instance, with $\varepsilon = 10$ m (used in this work), it is ensured that any planar point above 10 m provides a minimum data discrepancy term in (1) and is therefore regarded as belonging to the building structure. ε is a scene- and data-dependent parameter. For data with relatively higher density and positioning accuracy, a lower value for ε may be used. For instance, a reasonable value for ε equal to the height of two floors (i.e., 6 m) has been proposed in [32].

- The role of radius size parameter r_N is twofold: First, it is used to compute local neighbors for feature computation,

TABLE IV
FOOTPRINT RECONSTRUCTION STATISTICS

Reconstructed footprints / Errors(Las Vegas / Berlin)	V_{alpha}	$V_{refined}$	$V_{facadefused}$	V_{final}
Commission error (%)	14.01 / 20.20	13.56 / 20.43	13.13 / 19.13	10.79 / 19.43
Omission error (%)	12.49 / 12.31	13.39 / 12.84	13.14 / 13.24	17.20 / 14.57

and second, it is used to spatially cluster extracted points such that they belong to individual buildings for subsequent reconstruction. Based on experimental evaluation over TomoSAR point clouds, a good choice for radius size between 5 and 10 m providing reasonable results for feature computation have been utilized [24], [56]. The use of a particular radius size, however, puts a constraint on the minimum distance between the buildings. That is, setting $r_N = 5$ m allows the algorithm to separate two buildings only if they are at least farther than 5 m from each other. Otherwise, the algorithm will merge them into one single cluster.

- Apart from these four parameters controlling the whole extraction procedure, the parametric value of local height
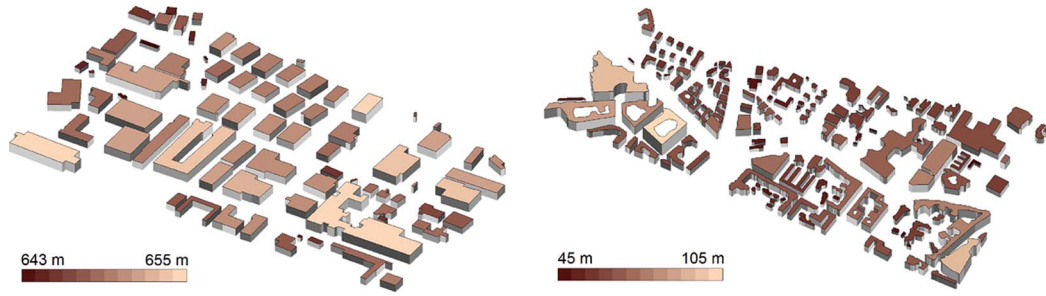
Fig. 16. Final reconstructed 3-D building footprints. (Left) Las Vegas data set. (Right) Berlin data set.

jumps that is used to select seed points while approximating the terrain surface also puts a constraint on minimum building height that could be detected. This parameter is merely data dependent and, therefore, has been empirically fixed in this work to 5 m. With a different sensor of higher resolution and positioning accuracy, this value may be lowered.

b) Reconstruction of building footprints: The whole reconstruction procedure is essentially governed by only two parameters: α and θ_{ang}.

- α provides good initial estimates of building outlines. However, the value of α effects the shape of the initial coarse polygon. Setting a larger α restricts in obtaining concave boundaries, whereas lower values may result in more smaller boundary polygons that are actually present. In addition, with smaller α, it is also possible that the outer and inner polygons share one (or more) common vertex and, hence, leads to improper footprint geometry. The use of one particular value for α may not be feasible for scenes containing arbitrary-sized buildings. Thus, to adaptively select an appropriate value of α, we initialize $\alpha = 5$ m (reasonable tradeoff for our data), which is recursively incremented by 1 m if resulting polygons share common vertices or the minimum area of any resulting polygon is less than 50 m^2.
- θ_{ang} is used to refine the initial (coarse) building boundary obtained using the alpha shape algorithm. $\theta_{ang} = 0$ results in no refinement or regularization, i.e., the original alpha polygons are returned. Setting too high a value for θ_{ang} may, however, result in overrefinement/smoothing.

Both α and θ_{ang} are very stable parameters for a variety of different input scenes with arbitrary building sizes. For α, the point density of the data plays an important role. For lower density data, the value of α should be higher, whereas for high-density data, lower α values may be used. For θ_{ang}, relatively stable values for data from different scenes (or sensors) range from 5° to 25°.

Although all of the remaining subsequent operations, including the *facade-Polygon* fusion, estimating the principal direction, and rectilinearization, are completely parameter free, one fixed threshold value, i.e., $L_f = 0.75$, providing a careful balance for our data, has been used to identify the rectilinear

TABLE V
LIST OF PARAMETERS TOGETHER WITH THEIR VALUES USED AND DEPENDENCY TYPE

Module	Parameters	Dependency type	Value used (stable range)
Building extraction & segmentation	*fac*	Neither scene nor data dependent	0.55 (0.5 ~ 0.7)
	$\theta_{normals}$	Data dependent	15° (10° ~ 20°)
	ε	Scene dependent	10 m (10 ~ 20m)
	r_N	Both scene and data dependent	5 m (5 ~ 10m for features computation)
	local height jump	Both scene and data dependent	5 m
Reconstruction	α	Data dependent	Initial value of 5 (3 ~ 15)
	θ_{ang}	Data dependent	20° (5° ~ 25°)
	L_f	Data dependent	0.75 (0.7 ~ 0.85)

polygons. With lower values of L_f, it is possible to wrongly classify more polygons as rectilinear, whereas higher values of L_f may increase the chance of a miss hit. This parameter also depends on the point density of the input data, and therefore, with higher density point clouds, values of L_f close to 1 may be employed.

To summarize, in total, eight parameters control the complete processing chain. Although the proposed approach is able to work with unstructured 3-D point clouds generated from any other sensor with a similar configuration (i.e., oblique geometry), the point density and positioning accuracy plays an important role in tuning these parameters. For spaceborne TomoSAR point clouds utilized in our experiments, Table V provides the values used in this work together with their possible (stable) range, which has been empirically evaluated for a variety of different input scenes. Additionally, categorization of each parameter according to its dependency type has also be presented.

VI. OUTLOOK AND CONCLUSION

In this paper, we have presented an automatic (parametric) approach that utilized unstructured spaceborne TomoSAR point clouds from one viewing angle only to detect and reconstruct 2-D/3-D building shapes/footprints. The approach is modular and allows for a robust detection of both tall and low buildings and, hence, is well suited for urban monitoring of larger areas from space. The approach is completely data driven and, therefore, imposes no restrictions on the shape of the building, i.e., any arbitrarily shaped footprint could be reconstructed. Moreover, the presented approach utilizes roof points in determining the complete shape of the buildings and, therefore, resolves problems, as mentioned in [24], related to the visibility of facades mainly facing toward the azimuth direction.

In the following, several aspects in further extension/improvement of the proposed approach are mentioned.

- The approach presented in [24] is a much better option for detecting the shape of the building when dense points on the facades are available. However, there are cases occurring particularly for lower height buildings or buildings having a smoothed facade structure where no or only very few scatterers are detected. The approach proposed in this paper presents nice solutions to such cases by exploiting roof points. However, the availability of roof points is somewhat dependent on the structure and geometry (area) of the roofs. For the part of the roof that is smooth or completely invisible to the sensor, no points would be obtained, which may result in the underreconstruction of the footprint. Such situations can be rectified by using points from other viewing angles (e.g., use of TomoSAR points generated from both ascending and descending orbits) and/or by incorporating some prior model knowledge of the shape to obtain complete building footprints.
- Building points are extracted based on two assumptions, namely, planarity and higher relative heights. This assumption also enables the detection/reconstruction of other man-made vertical structures, e.g., advertisement boards and monuments. Imposing dimensional constraints on the reconstructed footprints may help in further distinguishing them from buildings.
- Unknown absolute height values of different reference points, which are chosen independently while processing individual VHR SAR data stacks, are a hindrance to the direct fusion of TomoSAR point clouds from multiple views. To this end, 2-D/3-D footprints are independently produced from each single aspect, and TomoSAR point cloud could be matched/correlated to develop a more precise object-based geometric TomoSAR fusion method compared with two existing ground-point-based [57] and feature-based [30] approaches. An alternative approach is geodetic TomoSAR [58], where the absolute 3-D coordinates of the reference points are retrieved using imaging Geodesy [59] and stereoSAR [60].

In the future, we will also explore the potential of extending the algorithm toward the generation of automatically reconstructed complete watertight prismatic (or polyhedral) 3-D/4-D building models from space.

References

[1] P. Sehrawat and K. Kensek, "Urban energy modeling: GIS as an alternative to BIM," 2014 ASHRAE/IBPSA-USA Building Simulation Conf., Atlanta, GA, USA, 2014.
[2] B. Tomaszewski, *Geographic Information Systems (GIS) for Disaster Management*. Boca Raton, FL, USA: CRC Press, 2014.
[3] T. H. Kolbe, G. Gröger, and L. Plümer, "CityGML-3D city models and their potential for emergency response," in *Geospatial Information Technoly Emergency Response*, vol. 257. London, U.K.: Taylor & Francis, 2008.
[4] J. Döllner, T. H. Kolbe, F. Liecke, T. Sgouros, and K. Teichmann, "The virtual 3D city model of Berlin–Managing, integrating, and communicating complex urban information," in *Proc. 25th UDMS*, 2006, pp. 15–17.
[5] A. Köninger and S. Bartel, "3D-GIS for urban purposes," *Geoinformatica*, vol. 2, no. 1, pp. 79–103, Mar. 1998.
[6] J. Kirtner, "Using LIDAR data in wireless communication system design," in *Proc. Amer. Soc. Photogramm. Remote Sens. Conf.*, Washington, DC, USA, 2000, pp. 1–8.
[7] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, "Interferometric potential of high resolution spaceborne SAR," *Photogramm.-Fernerkund.-Geoinformation*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.
[8] A. Ferro, D. Brunner, and L. Bruzzone, "Automatic detection and reconstruction of building radar footprints from single VHR SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 935–952, Feb. 2013.
[9] A. Thiele, E. Cadario, K. Schulz, U. Thonnessen, and U. Soergel, "Building recognition from multi-aspect high-resolution InSAR data in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3583–3593, Nov. 2007.
[10] Y.-Q. Jin and F. Xu, "Automatic Reconstruction of Building Objects from Multi-Aspect SAR Images," in *Polarimetric Scattering and SAR Information Retrieval*. Hoboken, NJ, USA: Wiley, 2013, pp. 235–274.
[11] L. Zhao, X. Zhou, and G. Kuang, "Building detection from urban SAR image using building characteristics and contextual information," *EURASIP J. Adv. Signal Process.*, vol. 2013, no. 1, p. 56, Mar. 2013.
[12] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.
[13] P. Pasquali *et al.*, "A 3-D SAR experiment with EMSL data," in *Proc. IGARSS*, 1995, vol. 1, pp. 784–786.
[14] 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.
[15] G. Fornaro, D. Reale, A. Pauciullo, X. Zhu, and R. Bamler, "SAR tomography: An advanced tool for spaceborne 4-D radar scanning with application to imaging and monitoring of cities and single buildings," *IEEE Geosci. Remote Sens. Soc. Newslett.*, pp. 9–17, Dec. 2012.
[16] 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.
[17] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 37–44, Jan. 2005.
[18] X. X. Zhu and R. Bamler, "Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 735–739, Jul. 2011.
[19] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.
[20] S. Gernhardt and R. Bamler, "Deformation monitoring of single buildings using meter-resolution SAR data in PSI," *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 68–79, Sep. 2012.
[21] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.
[22] D. Reale, G. Fornaro, A. Pauciullo, X. Zhu, and R. Bamler, "Tomographic imaging and monitoring of buildings with very high resolution SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 661–665, Jul. 2011.
[23] G. Fornaro, D. Reale, and S. Verde, "Bridge thermal dilation monitoring with millimeter sensitivity via multidimensional SAR imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 677–681, Jul. 2013.
[24] M. Shahzad and X. X. Zhu, "Robust reconstruction of building facades for large areas using spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 752–769, Feb. 2015.
[25] 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.
[26] X. X. Zhu and R. Bamler, "Demonstration of super-resolution for tomographic SAR imaging in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3150–3157, Aug. 2012.
[27] G. Fornaro, A. Pauciullo, D. Reale, and S. Verde, "Multilook SAR tomography for 3-D reconstruction and monitoring of single structures applied to COSMO-SKYMED data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 7, pp. 2776–2785, Jul. 2014.
[28] X. X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, Jan. 2012.

[29] S. Auer, S. Gernhardt, and R. Bamler, "Ghost persistent scatterers related to multiple signal reflections," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 919–923, Sep. 2011.

[30] Y. Wang and X. X. Zhu, "Automatic feature-based geometric fusion of multiview TomoSAR point clouds in urban area," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 3, pp. 953–965, Mar. 2015.

[31] M. Hubert, P. J. Rousseeuw, and K. V. Branden, "ROBPCA: A new approach to robust principal component analysis," *Technometrics*, vol. 47, no. 1, pp. 64–79, Feb. 2005.

[32] F. Lafarge and C. Mallet, "Creating large-scale city models from 3D-point clouds: A robust approach with hybrid representation," *Int. J. Comput. Vis.*, vol. 99, no. 1, pp. 69–85, Aug. 2012.

[33] J. Yan, J. Shan, and W. Jiang, "A global optimization approach to roof segmentation from airborne lidar point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 94, pp. 183–193, Aug. 2014.

[34] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.

[35] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[36] M. Shahzad and X. X. Zhu, "Reconstructing 2-D/3-D building shapes from spaceborne tomographic synthetic aperture radar data," in *Proc. ISPRS Spatial Inf. Sci.*, Aug. 2014, vol. XL-3, pp. 313–320.

[37] B. Fulkerson, A. Vedaldi, and S. Soatto, "Class segmentation and object localization with superpixel neighborhoods," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, 2009, pp. 670–677.

[38] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, Feb. 2004.

[39] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," *Data Mining Knowl. Discov.*, vol. 96, no. 2, pp. 226–231, Jun. 1996.

[40] H. Edelsbrunner and E. P. Mücke, "Three-dimensional alpha shapes," *ACM Trans. Graph.*, vol. 13, no. 1, pp. 43–72, Jan. 1994.

[41] 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.

[42] K. Zhang, J. Yan, and S.-C. Chen, "Automatic construction of building footprints from airborne LIDAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2523–2533, Sep. 2006.

[43] X. Zhu, Very High Resolution Tomographic SAR Inversion for Urban Infrastructure Monitoring - A Sparse and Nonlinear Tour, C. Reihe Ed. Munich, Germany: Deutsche Geodätische Kommission, Verlag der Bayerischen Akademie der Wissenschaften, 2011, No. 666, ISBN 978-3-7696-5078-5.

[44] X. X. Zhu, Y. Wang, S. Gernhardt, and R. Bamler, "Tomo-GENESIS: DLR's tomographic SAR processing system," in *Proc. JURSE*, 2013, pp. 159–162.

[45] CyberCity3D, Inc. [Accessed: 07-Apr-2015]. [Online]. Available: http://www.cybercity3d.com/

[46] GEOFABRIK," 2015. [Accessed: 07-Apr-2015]. [Online]. Available: http://www.geofabrik.de/data/download.html

[47] Stats-OpenStreetMap Wiki. [Accessed: 22-Jan-2015]. [Online]. Available: http://wiki.openstreetmap.org/wiki/Statistics

[48] H. Fan, A. Zipf, Q. Fu, and P. Neis, "Quality assessment for building footprints data on OpenStreetMap," *Int. J. Geograph. Inf. Sci.*, vol. 28, no. 4, pp. 700–719, Apr. 2014.

[49] M. Haklay, "How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets," *Environ. Plan. B, Plan. Des.*, vol. 37, no. 4, pp. 682–703, 2010.

[50] M. Helbich, C. Amelunxen, and P. Neis, "Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary geodata," in *Proc. AGILE*, 2011, pp. 24–33.

[51] P. Mooney, P. Corcoran, and A. C. Winstanley, "Towards quality metrics for OpenStreetMap," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, New York, NY, USA, 2010, pp. 514–517.

[52] R. Hecht, C. Kunze, and S. Hahmann, "Measuring completeness of building footprints in OpenStreetMap over space and time," *ISPRS Int. J. Geo-Inf.*, vol. 2, no. 4, pp. 1066–1091, Nov. 2013.

[53] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, "A comparison of evaluation techniques for building extraction from airborne laser scanning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 2, no. 1, pp. 11–20, Mar. 2009.

[54] G. Sohn and I. Dowman, "Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 1, pp. 43–63, May 2007.

[55] L. Ding, X. Zhu, and L. Meng, "Scientific visualization for 4-D building deformation data analysis," presented at the 26th Int. Cartographic Conf., Dresden, Germany, 2013.

[56] O. D'Hondt, S. Guillaso, and O. Hellwich, "Geometric primitive extraction for 3D reconstruction of urban areas from tomographic SAR data," in *Proc. JURSE*, 2013, pp. 206–209.

[57] S. Gernhardt, X. Cong, M. Eineder, S. Hinz, and R. Bamler, "Geometrical fusion of multitrack PS point clouds," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 38–42, Jan. 2012.

[58] X. X. Zhu, S. Montazeri, C. Gisinger, R. Hanssen, and R. Bamler, "Geodetic SAR Tomography," in *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 18–35, Jan. 2016.

[59] M. Eineder, C. Minet, P. Steigenberger, X. Cong, and T. Fritz, "Imaging geodesy—Toward centimeter-level ranging accuracy with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 661–671, Feb. 2011.

[60] C. Gisinger *et al.*, "Precise three-dimensional stereo localization of corner reflectors and persistent scatterers with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1782–1802, Apr. 2015.

Muhammad Shahzad (S'12) received the B.E. degree in electrical engineering from the National University of Sciences and Technology, Islamabad, Pakistan, and the M.Sc. degree in autonomous systems from the Bonn Rhein Sieg University of Applied Sciences, Sankt Augustin, Germany. He is currently working toward the Ph.D. degree at Signal Processing in Earth Observation (SiPEO), Technische Universität München, Munich, Germany.

His research interests include the automatic 3-D reconstruction of objects from point clouds retrieved from spaceborne synthetic aperture radar image stacks.

Xiao Xiang Zhu (S'10–M'12–SM'14) received the Bachelor's degree in space engineering from the National University of Defense Technology, Changsha, China, in 2006 and the M.Sc., Dr.-Ing., and "Habilitation" degrees in signal processing from the Technische Universität München (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

She was a Guest Scientist or a Visiting Professor with the Italian National Research Council (CNR-IREA), Naples, Italy; Fudan University, Shanghai, China; and the University of Tokyo, Tokyo, Japan, in 2009, 2014, and 2015, respectively. Since 2011, she has been a Scientist with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Oberpfaffenhofen, Germany, where she is also the Head of Team Signal Analysis. Since 2013, she has been a Helmholtz Young Investigator Group Leader and appointed as a TUM Junior Fellow. In 2015, she was appointed as a Professor for signal processing in Earth observation with TUM. Her main research interests include advanced interferometric synthetic aperture radar techniques, such as high-dimensional tomographic synthetic aperture radar imaging and SqueeSAR; computer vision in remote sensing, including object reconstruction and multi-dimensional data visualization; and modern signal processing, including innovative algorithms such as compressive sensing and sparse reconstruction, with applications in the field of remote sensing such as multispectral/hyperspectral image analysis.

Dr. Zhu is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.