

Robust Reconstruction of Building Facades for Large Areas Using Spaceborne TomoSAR Point Clouds

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

Abstract—With data provided by modern meter-resolution synthetic aperture radar (SAR) sensors and advanced multipass interferometric techniques such as tomographic SAR inversion (TomoSAR), it is now possible to reconstruct the shape and monitor the undergoing motion of urban infrastructures on the scale of centimeters or even millimeters from space in very high level of details. The retrieval of rich information allows us to take a step further toward generation of 4-D (or even higher dimensional) dynamic city models, i.e., city models that can incorporate temporal (motion) behavior along with the 3-D information. Motivated by these opportunities, the authors proposed an approach that first attempts to reconstruct facades from this class of data. The approach works well for small areas containing only a couple of buildings. However, towards automatic reconstruction for the whole city area, a more robust and fully automatic approach is needed. In this paper, we present a complete extended approach for automatic (parametric) reconstruction of building facades from 4-D TomoSAR point cloud data and put particular focus on robust reconstruction of large areas. The proposed approach is illustrated and validated by examples using TomoSAR point clouds generated from a stack of TerraSAR-X high-resolution spotlight images from ascending orbit covering an approximately 2-km² high-rise area in the city of Las Vegas.

Index Terms—Clustering, facade reconstruction, point density, TerraSAR-X, tomographic SAR (TomoSAR) inversion, 4-D point cloud.

I. INTRODUCTION

AUTOMATIC detection and reconstruction of buildings has been an active research area for at least two decades. Despite an extensive research effort, the topic is still of great interest due to ever increasing growth of urban population which gives rise to a wide range of potential applications in various fields. For instance, building footprints (i.e., 2-D building outline/shape) can be used for urban landscape development [1], urban planning [2], damage assessment [3], disaster management [4], navigation purposes [5], etc. Additionally,

Manuscript received December 13, 2013; revised April 7, 2014; accepted May 7, 2014. This work was supported in part by the Helmholtz Association under the framework of the Young Investigators Group “SIPEO.” This work was also part of the project “6.08 4-D City” funded by the IGSS of Technische Universität München and the German Research Foundation (DFG, Förderkennzeichen BA2033/3-1).

M. Shahzad is with the Lehrstuhl für Methodik der Fernerkundung, Technische Universität München, 80333 Munich, Germany (e-mail: muhammad.shahzad@bv.tum.de).

X. Zhu is with the Lehrstuhl für Methodik der Fernerkundung, Technische Universität München, 80333 Munich, Germany and also with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany (e-mail: xiao.zhu@dlr.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2014.2327391

2-D footprint data in combination with height information can generate 3-D building models. Such models are essential for virtual city modeling [6] and 3-D GIS applications (e.g., commercial software such as Google Earth, Apple Maps, etc.). Other possible usages may include analyzing solar potential over building roofs [7], placing and installing telecommunication antenna towers [8], Web-based mapping [9], tourism [6], architecture [10], augmented reality applications [5], [11], and many more.

Spaceborne synthetic aperture radar (SAR) sensors are able to provide day/night global coverage in virtually all weather conditions. Moreover, due to coherent imaging nature, by using acquisitions taken at different times, it is also uniquely capable of imaging the dynamics of the illuminated area in the scale of centimeters or even millimeters. These benefits have motivated many researchers, and therefore, several methods have been developed, which use very high resolution (VHR) spaceborne SAR imagery for detection and reconstruction of man-made structures in particular buildings. For instance, single-aspect SAR-image-based approaches for building reconstruction are proposed in [12]–[14]. Due to the fact that only single SAR images are used, these approaches predominantly perform well mostly only for isolated buildings but not for dense urban areas where the buildings are densely packed and smaller buildings are often occluded by the higher ones [15]. To resolve this, interferometric SAR (InSAR) data, SAR image pairs taken from slightly different viewing angles, are used, e.g., a modified machine vision approach is proposed in [16] to detect and extract buildings. The algorithm is based on local approximation of best fitting planes in the segmented regions of interest. Similarly, Thiele *et al.* [17] also developed a model-based approach to detect and reconstruct building footprints using orthogonal InSAR images. Another automatic approach based on modeling building objects as cuboids using multiaspect polarimetric SAR images is presented in [18]. In data fusion context, the use of optical imagery has also been exploited along with SAR [19] and InSAR [15] datasets, respectively. Despite of the active ongoing research in the area, the problem of building reconstruction still remains challenging due to inherent characteristics with SAR images such as geometrical projection caused by the side-looking geometry [20]. Moreover, complex building structures and high variability of objects appearing in the images make automatic building detection and reconstruction a challenging task. For example, problems posed by occlusion of smaller buildings/objects from the higher ones render difficulties in large area extension. Therefore, prior knowledge is often incorporated with certain regularization (geometric) constraints for realistic and automatic reconstruction. For instance,

facades are often assumed to be vertical [19], building footprint as regular parallelepipeds [14], roofs as polyhedral structures [21], etc.

Modern spaceborne SAR sensors such as TerraSAR-X/TanDEM-X [22] and COSMO-SkyMed [23] can deliver VHR resolution data that fit well to the inherent spatial scales of buildings. Hence, these VHR data are particularly suited for detailed urban mapping [24]–[33]. However, 2-D SAR imaging projects the 3-D scene onto a 2-D image, making it “noninjective” in urban scenarios due to the presence of vertical structures (e.g., building facades or other man-made objects) [25]. The unwelcoming effects such as layover and foreshortening seriously handicap the interpretation of SAR images. Advanced interferometric techniques, such as persistent scatterer interferometry (PSI) and SAR tomography (TomoSAR), aim at SAR imaging in 3-D or even higher dimensions. Among them, PSI exploits the coherent pixels, i.e., the bright long-term stable objects (persistent scatterers) [30]. However, it assumes single scatterers in one azimuth-range pixel and therefore does not resolve the layover problem. TomoSAR, on the other hand, aims at real and unambiguous 3-D SAR imaging [25], [34]–[36]. By exploiting multipass SAR images taken from slightly different positions, like PSI does, it builds up a synthetic aperture in the elevation direction that enables retrieval of precise 3-D position of dominant scatterers via spectral analysis within one azimuth-range SAR image pixel [25]. Multiple layovered objects in any pixel are therefore separated from the reconstructed reflectivity profile in elevation direction [36]. Moreover, exploiting the fact that different acquisitions are taken at different times, the synthetic aperture can also be extended in the temporal domain to enable 4-D (space-time) focusing of SAR images. The technique is referred to as D-TomoSAR, which combines the strengths of both TomoSAR and PSI [26], [27], [37]–[40]. It is capable of retrieving elevation and deformation information (linear, seasonal, etc.) even of multiple scatterers inside a single SAR image pixel [25], [29]. Retrieval of rich scatterer information from VHR D-TomoSAR enables generation of 4-D (space-time) point cloud of the illuminated area with point (scatterer) density that is comparable to LiDAR, e.g., experiments using TerraSAR-X high-resolution spotlight data stacks show that the scatterer density retrieved using TomoSAR is on the order of 1 million points/km² [28].

Object reconstruction from these high-quality TomoSAR point clouds can greatly support the reconstruction of dynamic city models that could potentially be used to monitor and visualize the *dynamics* of urban infrastructure in very high level of details. Such models would be immensely helpful to ensure safety/security of growing urban population by monitoring of urban infrastructures against potential threats of damage and structural degradation caused by various factors, e.g., ground subsidence or uplift, bad construction, natural disaster, etc. Motivated by this, the very first results of building facade reconstruction from single-view (ascending stack) and multi-view (fused ascending and descending stacks) perspectives over a small test building area (Bellagio hotel, Las Vegas) were presented in [41] and [42], respectively.

In this paper, we present an approach that allows automatic reconstruction of 3-D building facades using these unstruc-

tured TomoSAR point clouds only. The approach proposes new methods as well as modifications to the previously introduced idea in [42] aimed at finding a more general solution toward automatic reconstruction of the whole city area. The basic idea is to reconstruct 3-D building models via independent modeling of each individual facade to build the overall 2-D shape of the building footprint, followed by its representation in 3-D. The following are the innovative contributions that are specific to the approach proposed in this paper.

- 1) A robust M-estimator-based directional SD estimation method is proposed, which provides much better estimates of facade regions compared to the grid-based SD estimation proposed in [42] by incorporating the facade geometry. Moreover, instead of rejecting non-facade points via 2-D morphological operations used in [42], robust 3-D surface normal information is utilized. The use of additional-dimensional with the vertical facade assumption helps in better rejecting nonfacade points.
- 2) K-means clustering with a criterion for prior guessing the number of clusters is used in previous works [41], [42]. This technique provides good results for single buildings, but when it comes to larger areas, there are two major concerns: 1) guessing the number of clusters is not always trivial, and 2) a certain shape of clusters is not very well recognized. For this reason, a three-step automatic (unsupervised) clustering approach that combines both the density-based clustering [43] and mean shift algorithm [44], [45] is proposed in this paper. The proposed segmentation approach is able to work directly on bigger areas without requiring any prior knowledge about the number of clusters.
- 3) Facades are modeled using general (first and second order) polynomial equations to cater for a wide variety of building footprint. A detailed methodological description of the modeling procedure is explained, which is able to cater arbitrary (rotated) orientations of building facades. The coefficients of the model are estimated using weighted total least squares (WTLS) method to cope for localization errors of TomoSAR points in both xy directions.
- 4) During the reconstruction procedure, the presence of smaller clustered segments occurring at facade transition regions handicaps accurate determination of vertex points from the adjacent facade pair and may cause the reconstruction procedure to fail. To deal with this problem, smaller “conflicting segments” are automatically identified and removed for accurate and robust reconstruction of the adjacent facades.
- 5) Side-looking SAR geometry and complexity of the scene in dense high-rise area of interest can cause occlusions of lower height facades scattered around higher building facade structures. As a consequence, a few or sometimes no data are available for the occluded region, rendering incomplete reconstruction or breaking an individual facade into two or more segments. A partial solution is also presented in this paper, which refines the

reconstructed facade footprints via insertion (of additional segments) and extension operations.

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

The aforementioned contributions allow completely *automatic* (but parametric) reconstruction of building facades from TomoSAR point clouds in larger areas.

The remainder of this paper is structured as follows. Section II presents a brief procedural overview of the existing techniques that use 3-D LiDAR point cloud for building reconstruction. Section III presents in detail the proposed approach. In Section IV, the experimental results, obtained from the TomoSAR point cloud generated from a TerraSAR-X high-resolution spotlight data stack (ascending orbit only), are presented and discussed. Finally, in Section V, a conclusion about the proposed approach is drawn, and future perspectives are discussed.

II. RELATED WORK

Most approaches employ airborne LiDAR data for automatic 3-D reconstruction of buildings. Methodologically, the problem is tackled by subdividing the task into two sequential steps, i.e., detection/classification of building points followed by their 3-D modeling and reconstruction.

Detection is generally carried out by first computing the digital terrain model (DTM) by filtering techniques, e.g., morphological filtering [46], gradient analysis [47], or iterative densification of triangular irregular network structures [48]. Nadir-looking LiDAR points essentially give a digital surface model (DSM). Subtracting DSM from the computed DTM provides a normalized DSM (nDSM) which represents the height variation of nonground points. Subsequently, building points are extracted out from nDSM by exploiting geometrical features such as deviations from the surface model, local height measures, roughness, and slope variations. Methods based on building boundary tracing from nDSM [49] or directly from point clouds [50], [51] have also been employed for building detection. With them, finer building boundaries are determined by regularization of the coarsely traced boundaries. All points that lie inside the boundary regions are considered as building points. Building points can also be extracted out by explicitly labeling every point in the data set. For labeling purposes, local neighborhood features such as height, eigenvalue, and plane features have been used in conjunction with supervised [52], semisupervised [21], and unsupervised [53] classification techniques.

Detected building regions or points are, in turn, used for 3-D modeling and reconstruction. Most methods make use of the fact that man-made structures such as buildings usually have either parametric shapes (model driven) or composed of

polyhedral structures only (data driven). The latter is, however, more common in the literature, where local sets of coplanar points are first determined using 3-D Hough transform or RANSAC algorithms and then reconstruction is carried out by surface fitting in the segmented building regions followed by region growing procedure [53] or by building up an adjacency graph [21], [54].

The aforementioned methods and the majority of other techniques present in the literature that make use of 3-D LiDAR data cannot be directly applied to TomoSAR point clouds due to side-looking SAR geometry and different microwave scattering properties of the objects appearing in the scene reflecting different geometrical and material features. Compared to LiDAR, TomoSAR point clouds possess the following peculiarities that should be taken into consideration.

Accuracy and Errors:

- TomoSAR point clouds deliver moderate 3-D positioning accuracy on the order of 1 m [15] as compared to (airborne) LiDAR systems having an accuracy on the order of 0.1 m [21].
- Ghost scatterers [55] may be generated due to multiple scattering that appears as outliers far away from a realistic 3-D position.
- A small number of images and limited orbit speed render the location error of TomoSAR points highly anisotropic, with an elevation error typically one or two orders of magnitude higher than in range and azimuth [25].

Coherent Imaging:

- Due to coherent imaging nature, temporally incoherent objects such as trees cannot be reconstructed from multipass spaceborne SAR image stacks.

Side-Looking Geometry:

- Separation of layover on vertical structures renders geocoded TomoSAR point clouds to possess higher density of points on building facades.
- In order to obtain a full structure of individual buildings from space, multiple views are required [42].

Higher Dimensional Imaging:

- In addition to 3-D spatial information, TomoSAR point clouds also possess the fourth-dimensional information, i.e., temporal or seasonal deformation estimates, making them very attractive for dynamic city modeling.

III. PROCESSING STEPS FOR BUILDING FACADE RECONSTRUCTION

Due to the side-looking SAR geometry, when projected, the TomoSAR point clouds onto ground plane vertical facade regions exhibit higher scatterer (point) density (SD) as compared to nonfacade regions. It is mostly true due to the existence of strong corner reflectors, e.g., window frames on the building facades. Taking this fact into account, in [42], we proposed to extract facade points by projecting all of the points onto the xy grid for estimating SD (rastered image), followed by thresholding and applying morphological dilation operation. This approach works well for high-rise buildings having a much higher point density but limits the extraction of facade points

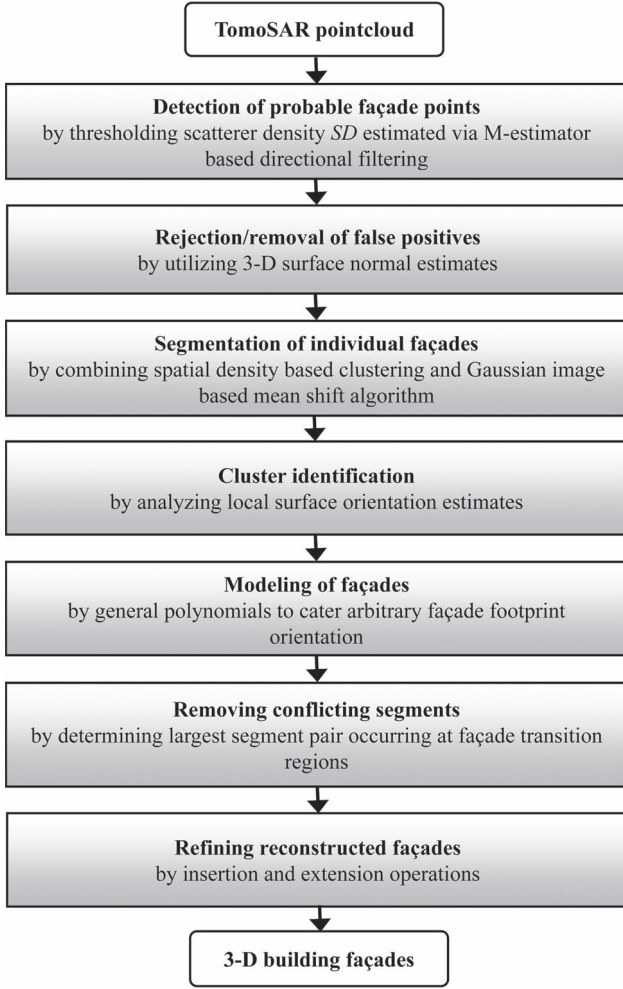


Fig. 1. Block diagram of the proposed method.

from relatively lower buildings. The selection of a particular threshold thus becomes crucial.

To resolve this issue, in this work, a more robust facade extraction approach is proposed, which is based on the directional SD estimation procedure to locally estimate the SD for each point while incorporating the facade geometry [56]. Later, robust 3-D surface normal information is utilized to extract facade points. Automatic segmentation of extracted facade points is obtained by first performing coarse clustering to cluster points belonging to individual buildings. Then, each cluster is further fine-clustered using Gaussian-map-based mean shift clustering algorithm. After that, clusters within clusters are spatially separated. Segmented facades are then classified as flat or curved, and their model parameters are estimated. Subsequently, the geometric primitives such as vertex points are determined from the intersection of adjacent facade pair after removing smaller conflicting segments occurring at transitional regions. Finally, a refinement operation is carried out on the reconstructed facades that remain either incomplete or broken into two or more segments to complete the reconstruction process.

Fig. 1 shows the block diagram of the processing steps involved in the complete methodology. Next, we explain in detail the procedures of the proposed approach in dedicated sections.

A. SD Estimation

For each 3-D TomoSAR point p , points within its local neighborhood v_c are used for SD estimation. v_c includes all of those points that lie inside a vertical cylinder centered at p . To emphasize the building facades, we incorporate facade geometry in estimating SD , i.e., we estimate the direction of the local neighborhood via line fitting using robust M-estimator. The method iteratively reweights the points according to the residuals and computes the so-called M-estimates as follows [57].

- 1) The initial estimates of the line parameters β (e.g., $\beta_1 =$ slope and $\beta_2 =$ offset) are derived from ordinary least squares.
- 2) The weights w_{p_i} of each point $p_i \in v_c$ are then computed using a bisquare function [57], [58]

$$w_{p_i} = \begin{cases} (1 - u^2)^2 & \text{for } |u| < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{where } u = \frac{|y_{p_i} - x_{p_i}\beta_1 - \beta_2|}{4.685\hat{\sigma}\sqrt{1-t}} \quad (1)$$

where t is the leverage computed using parameter estimates β of the fitted line and $\hat{\sigma}$ is the estimate for the scale of the error term computed by $\hat{\sigma} = 1.483 * \text{MAD}$, where MAD is the median absolute deviation of the residuals from their median. The term 1.483 is used to make the estimate $\hat{\sigma}$ consistent to the standard deviation at Gaussian distribution [58], [59].

- 3) Updating β using weights w_{p_i} by applying weighted least squares to solve the following objective function:

$$\arg \min_{\beta} \sum_{p_i \in v_c} w_{p_i}(\beta) |y_{p_i} - x_{p_i}\beta_1 - \beta_2|^2 \quad (2)$$

where x_{p_i} and y_{p_i} represent the abscissa and ordinate (i.e., ground coordinates) of the points within v_c .

- 4) Iterating steps 2 and 3 until a fixed number of iterations.

The estimated line describes the main principal axis of the cylindrical footprint of the local neighborhood. The orthogonal distance for every point in v_c is then calculated from the principal axis (shifted to the point p), and the points having distances less than d are taken as “inliers” and used in SD estimation.

SD for each point is thus defined as the number of points within a directional (cylindrical) neighborhood window divided by the area of the window

$$SD = \frac{\text{number of points in } v_d}{\text{Area of } v_d} \quad (3)$$

where $v_d \subseteq v_c$ but includes only those points that lie close to the principal axis of points in v_c .

Fig. 2 shows the graphical representation of the SD estimation procedure.

B. Facade Extraction

Based on the estimated SD , facade points can be extracted. For a large area, both high and low buildings are present. A *hard* threshold, i.e., removing points below a rather high SD value,

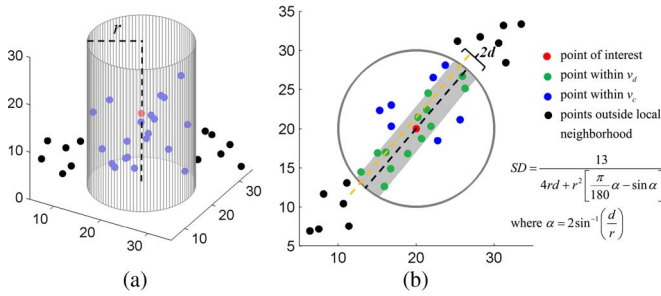


Fig. 2. Illustration of SD estimation. (a) Three-dimensional view of the local cylindrical neighborhood around the point of interest. (b) Top view of (a). The coefficients of the dotted yellow regression line are estimated via M-estimation. The black dotted line shows the shift of the yellow line to the point of interest. The shaded area shows the region of v_d within v_c .

as used in [42], would lead to misdetection of facade points of lower buildings whose SD estimates would be relatively low with respect to high-rise buildings. To avoid such misdetection, we extract facade points in a sequential way. First, we have lowered the SD threshold to detect not only higher buildings but also lower ones. However, a softer threshold would also introduce false positives, i.e., roof points or ground points with a local point density comparable to those of lower buildings. Therefore, we introduce a second step which utilizes 3-D surface normal information by incorporating prior knowledge (i.e., facades are assumed to be vertical surfaces which separate them from nonvertical ground plane and roofs) to eliminate those false positives.

The key issue is then the local surface normal estimation for each 3-D point. Commonly, they are estimated via fitting “best” plane in least squares (LS) sense in the local neighborhood v_c , which is equivalent to performing principal component analysis (PCA) of the points in v_c [60]. This implies that the surface normals can be directly estimated for each 3-D point via eigenvalue/eigenvector analysis of 3-D (i.e., 3×3) covariance matrix Σ_{v_c} . There are two advantages in using eigenvalue/eigenvector analysis of Σ_{v_c} for surface normal estimation.

- 1) First, the eigenvector associated to the smallest eigenvalue of the positive semidefinite matrix Σ_{v_c} is the direct estimate of the local surface normal of the query point [60].
- 2) Second, in addition to giving us the direct solution of estimating local surface normal, it can also help us in determining the dimensionality of each 3-D point [21]. To elaborate, the eigenvectors of Σ_{v_c} essentially give the orthonormal basis for the local neighborhood v_c , with their corresponding eigenvalues representing the magnitude (or variance) of expansion. Analyzing this magnitude implicitly gives us an indication of the beneath surface, e.g., eigenvalue analysis of Σ_{v_c} with all points lying on the plane would ideally return only two nonzero eigenvalues. Similarly, all points lying on a 3-D line would give only one nonzero eigenvalue. Eigenvalue analysis for segmentation and classification of planar points has been exploited in [21], [52], and [53].

Eigenvalue/eigenvector analysis via classical PCA may fail to give precise estimate of the 3-D surface normal using

TomoSAR point cloud due to the presence of outliers and localization errors (see Section II: accuracy and errors). Robust estimation of the covariance matrix Σ_{v_c} is therefore needed. To this end, we estimated Σ_{v_c} using robust minimum covariance determinant (MCD) method [61]. The method finds a subset (fraction) α of the data points $p_i \in v_c$ whose covariance matrix has the lowest determinant. The idea stems from the concept of generalized variance (GV) which is defined to be the determinant of the covariance matrix of any d -D ($d > 1$) random variable [62]. For instance, in case of 2-D (x - y) points, the GV provides a scalar value which measures the overall variability of all points in both x and y dimensions. Points that are clustered tightly together tend to have a smaller GV (i.e., lower determinant of their covariance matrix) as compared to scattered ones. Thus, the subset α of the data points which provides the lowest determinant is taken as the MCD estimate of Σ_{v_c} . If the data points are assumed to have less than 25% outliers, then an appropriate selection of $\alpha = 0.75$ (also used in this work) proposed in [63] provides a good compromise between statistical efficiency and high breakdown value ($\alpha = 0.75$ implies that 75% of the data points has been used in covariance estimation).

The covariance matrix Σ_{v_c} estimated using the MCD method from the local neighboring points $p_i \in v_c$ around (in cylinder) the point of interest $p_o(x_o, y_o, z_o)$ is then used to determine the local 3-D surface normal at p_o . If we denote a plane which robustly fits the neighboring points p_i as $n_x x + n_y y + n_z z + \rho = 0$, with $\rho = -n_x x_o - n_y y_o - n_z z_o$, then $N_o(n_x, n_y, n_z)$ depicts the local 3-D surface normal at p_o . N_o is thus directly estimated from Σ_{v_c} by computing the eigenvector associated to the smallest eigenvalue of Σ_{v_c} (here, v_c includes points in the vicinity of p_o) i.e.,

$$\text{if } \Sigma_{v_c} \cdot v_j = \lambda_j \cdot v_j, \quad j = 1, 2, 3 \text{ (descending order)}$$

then surface normal of the underlying surface at point

$$p_o : N_o(n_x, n_y, n_z) = v_3. \quad (4)$$

From (4), robust 3-D surface normals are computed for each point that is obtained after SD thresholding. Ideally, the direction of surface normal should be parallel to the ground for points on the vertical facades which separate them from non-vertical ground plane and roofs. Taking this fact into account, facade points are extracted out by retaining only those points having normals that are close to the horizontal axis (i.e., parallel to ground for points belonging to a vertical surface). In this manner, the proposed two-step approach allows us to robustly extract facade points over a large area where both high and low buildings are present.

Fig. 3 shows a comparison of the proposed approach with the one presented in [42]. The selected area shown in Fig. 3(a) contains relatively lower height buildings with low and inconsistent density of points on building facades. It can be seen that, in comparison to the SD estimation results from the previous approach depicted in Fig. 3(b), higher and complete density values are obtained for facade regions using the SD estimation method proposed in this paper, shown in Fig. 3(c).

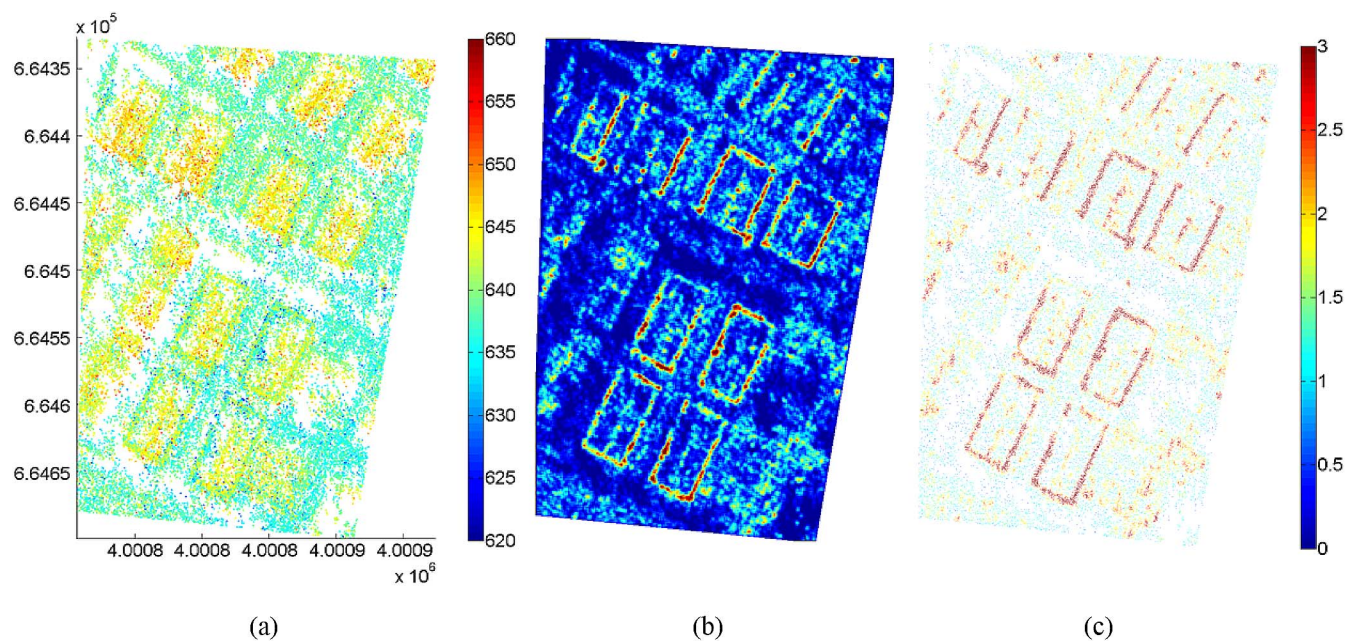


Fig. 3. Comparison of two SD estimation methods using TomoSAR points of an area in Las Vegas containing rectangular shaped buildings. (a) TomoSAR points (top view) generated from ascending stack only. The axes are in UTM coordinates. The height (above sea level) of TomoSAR points is color-coded [unit: meter]. (b) SD image estimated via [42]. (c) SD estimated via M-estimator-based directional filter proposed in this paper. Higher SD regions depict probable facade points. SD is color-coded, with the colorbar representing points/m². (b) and (c) share the same colorbar. Note that the SD estimated in (b) is the rastered image obtained by projecting all points onto the xy grid as compared to (c) where SD is directly computed for each point.

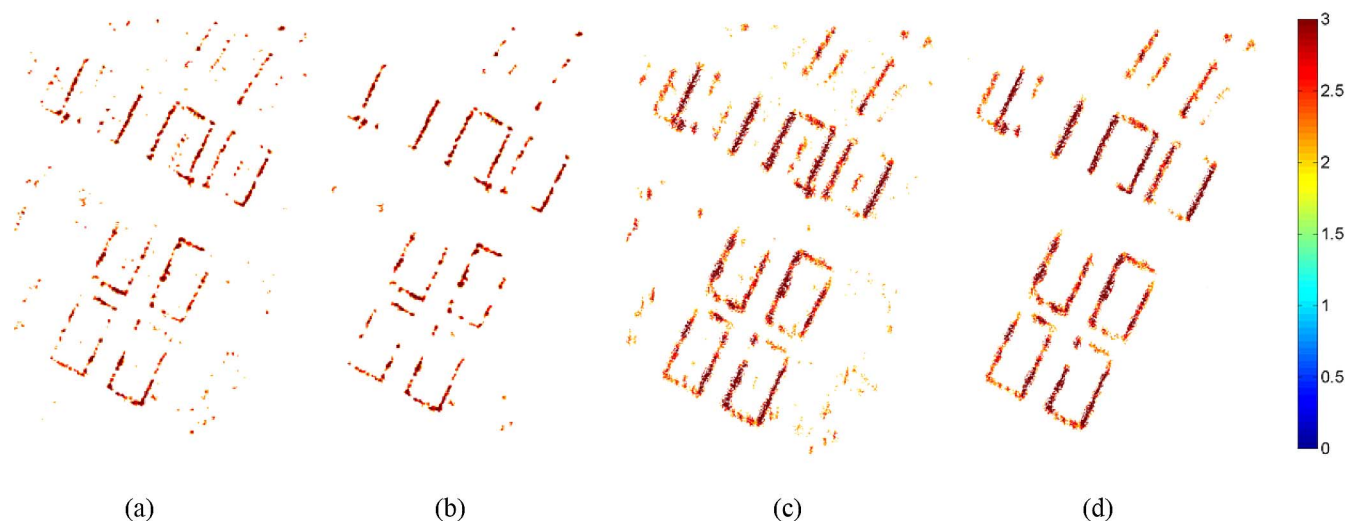


Fig. 4. Extraction of facade regions/points using the SD estimation results from Fig. 3. (a) Building facades obtained by thresholding the rastered SD image of Fig. 3(b) by the TH value. (b) Final extracted facade after 2-D morphological operation on (a) as proposed in [42]. (c) TomoSAR points whose SD estimated in Fig. 3(c) is above TH . (d) Extracted facades from (c) by utilizing robust 3-D surface normal information. The threshold value TH used here is 2 points/m² (empirically found to be optimum in [42]).

Moreover, later use of the third dimension in robust 3-D surface normal estimation provides much better results of extracting facades by rejecting nonfacade points.

Fig. 4 presents the comparison of the facade extraction results obtained using the SD estimates in Fig. 3. Taking the facade point candidates extracted by thresholding SD [Fig. 3(a) and (c)] as inputs, 2-D morphological operations (area opening and dilation) as proposed in [42] and robust 3-D surface normal information as suggested in this paper are performed to reject false positives. The final extracted facades are shown in Fig. 4(b) and (d), respectively.

C. Segmentation of Individual Facades

The extracted facade points belong to different facades. Clustering of points belonging to the same facade is therefore needed. First, a coarse clustering is performed using density-based clustering algorithm [43]. It involves the notion of density connectivity between the points. For example, two points are directly density connected to each other if one 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 all are directly density connected

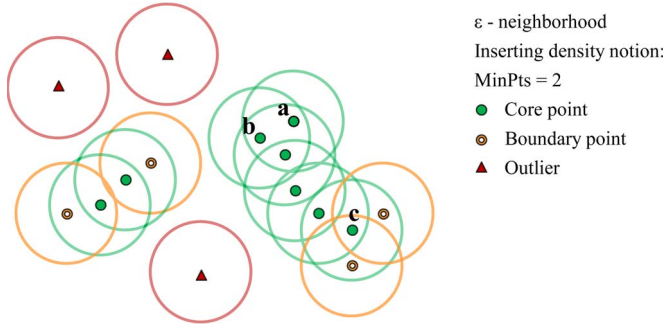


Fig. 5. Density-based clustering. Points a and b are directly density connected to each other, whereas points a and c are density connected to each other since there is a chain of points between them such that they all are directly density connected to each other.

to each other. Two parameters that control the clustering process include the neighborhood parameter ε , i.e., the radius in case of sphere or cylindrical neighborhood, and the minimum number of points MinPts in the ε -neighborhood for any particular point. The resulting clusters K_i thus contain points such that all of the points in any particular cluster are density connected to each other but are not density connected to any other point belonging to another cluster. Moreover, each point inside any particular cluster K_i belongs to one of the three categories (Fig. 5) [43].

- 1) Core points: A point is labeled core point if it contains, within its ε -neighborhood, MinPts number of points.
- 2) Border points: A point is considered border point if it is within ε -neighborhood of any core point but itself is not a core point and does not have MinPts neighbors.
- 3) Outliers: A point neither core point nor border point is termed as an outlier, i.e., any point which do not have density (points) greater than MinPts within its ε -neighborhood and also is not the neighbor of any other point.

Density-connected clusters containing only core and boundary points are used for further processing.

The aforementioned process, however, may merge points of two or more adjacent facade segments into a single cluster. To reconstruct individual facade segments, separation of these segments is therefore necessary. It is done by mapping the facade points in Gaussian image and then employing mean shift clustering.

Let us assume that a coarsely clustered segment K_i consist of one or more vertical adjacent facades F_j , $j = 1, \dots, J$. An image of a map $M: F \rightarrow F^2$ that assigns each point in F to its respective unit surface normal is known as Gaussian image GI of F [64]. Flat F (i.e., planar surface) should ideally be represented by a point in GI . Fig. 6 conceptually illustrates this in an ideal scenario. In practice, surface normals are estimated locally and may fluctuate from one point to another as practical data often contain errors in 3-D positions. However, if the estimation of normals is robust enough, a surface mapped to GI should be represented as a dense cluster of points in GI . The shape of clusters in GI corresponds to the geometry of connected surfaces [44]. The number of clusters in GI tells the number of surfaces in the spatial domain.

If we assume $p_r = 1, \dots, m$ to be 3-D points and n_r as their corresponding 3-D unit normal vectors belonging to one of the

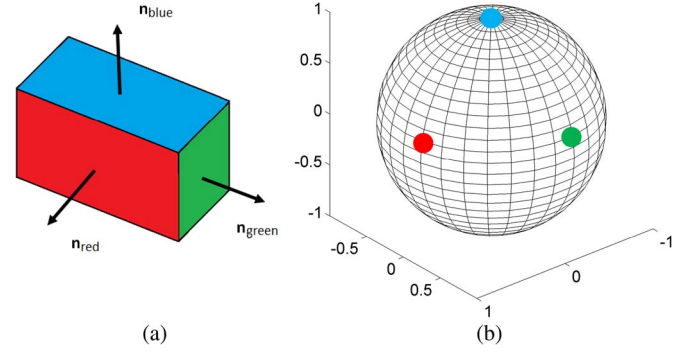


Fig. 6. Gaussian image of three connected planar surfaces. (a) Arrows indicate surface normal vectors (n_{red} , n_{green} , n_{blue}) to the respective surfaces. (b) All points belonging to one particular surface are mapped to the same identical point in GI (ideal scenario).

coarsely clustered segments K_i , then the density at any normal point $n_q (q \in r)$ in GI (feature space) is defined as [44]

$$D_{n_q} = \frac{c}{mb^3} \sum_{r=1}^m g \left(\left\| \frac{n_q - n_r}{b} \right\|^2 \right) \quad (5)$$

where b is the bandwidth parameter and $g(x)$ is a nonnegative, nonincreasing, and piecewise continuous function with definite integral, i.e., $\int_0^\infty g(x) dx < \infty$. From the concept of kernels [45], the function $g(x)$ is defined as the profile of the radially symmetric kernel $G(x)$ satisfying $G(x) = cg(\|x\|^2)$, where c is a normalization constant ensuring that $G(x)$ integrates to 1. Different kernels, such as the unit flat kernel and the Gaussian kernel, can be used to define the density D_{n_q} . However, the latter with the profile function $\exp(-\|(n_q - n_r)/b\|^2)$ has been used in this work.

Density D_{n_q} is higher for points that belong to planar or parabolic surfaces and lower for points that lie at the transition edges between the surfaces [44]. These higher density points in the GI are identified and clustered using mean shift (MS) clustering procedure. MS is a mode-seeking procedure and works iteratively by shifting every data point toward the weighted mean of points within its neighborhood. The shift vector $m(n_q)$ always points toward the direction of the maximum increase in the density D_{n_q} [65] and is computed as

$$m(n_q) = \frac{\sum_{r=1}^m n_r \exp \left(-\frac{\|n_q - n_r\|^2}{b^2} \right)}{\sum_{r=1}^m \exp \left(-\frac{\|n_q - n_r\|^2}{b^2} \right)} - n_q. \quad (6)$$

Applying MS in GI produces clusters whose corresponding points in spatial domain represent different facades. However, it is also possible that spatial points corresponding to any one particular normal cluster in GI may belong to two or more different facades. This can happen if points of two or more facades that are “nearly” parallel to each other (i.e., having close normal directions) are present in K_i . To resolve this, density-based clustering is again performed in the resulting clusters for spatial separation of parallel facade points clustered into one group. Finally, clusters with very few points are removed from further processing for robust reconstruction.

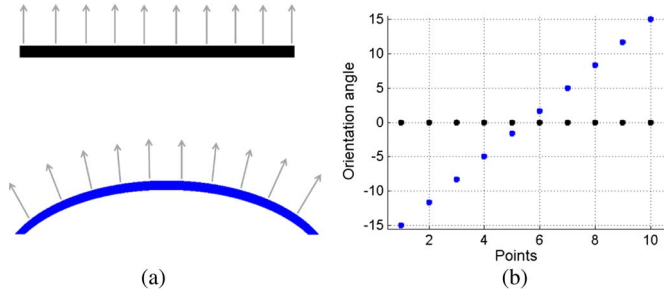


Fig. 7. Illustration of orientation angle for flat and curved vertical footprints (top view). (a) Arrows indicate pattern of change in orientation (azimuthal) angles of ten points on each vertical surface. (b) Plots their respective orientation angles.

D. Cluster Identification

Each cluster is further classified into flat or curved surface by analyzing derivatives of the local orientation angle θ . θ for each 3-D point is equal to the azimuthal angle of the corresponding computed surface normal

$$\theta = \arctan\left(\frac{\lambda_{3y}}{\lambda_{3x}}\right) \quad (7)$$

where λ_{3x} and λ_{3y} represent the x and y components of the surface normal λ_3 of any 3-D point. Ideally, the flat surfaces should have constant orientations, i.e., zero derivatives compared to the curved surfaces that have gradually changing orientations (see Fig. 7). We exploit this fact and compute the first derivative θ' of the orientation angle for each facade footprint. Since the original orientation derivatives θ' are usually noisy, all of the points are first projected to the first principal axis, and polynomial fitting is later applied for denoising. Based on the behavior of θ' , facade footprints are classified as flat or curved.

E. Modeling of Facades

The identified facade clusters in xy plane are then modeled using the following general polynomial equation [42]:

$$f_p(x, y) = \sum_{q=1}^p a_q x^i y^j \quad i + j \leq q \quad (8)$$

where i and j are permuted accordingly, p is the order of the polynomial, and the number of terms in the aforementioned polynomial is equal to $(p+1)(p+2)/2$. Cross terms are introduced in the model in case of the rotated local coordinate system. To solve (8), we restrict ourselves to the first and second orders (i.e., *flat* with $\max(i, j) = 1$ and *curved* with $\max(i, j) = 2$). The coefficients a_q are estimated using the WTLS method, where the total least squares is utilized to cope for localization errors of TomoSAR points in both xy directions and the weight of each point is assigned equal to its corresponding SD . The weighted polynomial fitting (residual) error f_{err} is minimum for the case where we have the unrotated local coordinate system reducing the right-hand side of (8) to $\sum_{i=0}^p a_i x^i$ (i.e., with no cross terms). In case of the rotated local coordinate system (which is often the case), we perform

the following steps to obtain consistent parameter estimates of all facades in a global coordinate system.

- 1) Rotate the points by rotation angle ω , and compute the polynomial fitting error f_{err} by applying the WTLS method.
- 2) Consider coefficients computed with ω_{min} that gives the minimum polynomial fitting error f_{err} as polynomial terms depicting *unrotated* points in the global coordinate system. ω_{min} is computed by using an unconstrained nonlinear optimization procedure to find the minimum of the error function f_{err} by varying ω over $0-360^\circ$ range via the Nelder–Mead simplex algorithm [66].
- 3) Rotate the computed polynomial by replacing the unrotated $(x-, y-)$ axis terms by their rotation counterparts $(x \cos \omega + y \sin \omega, -x \sin \omega + y \cos \omega)$ to yield polynomial terms a_q in global coordinates.

F. Removing Conflicting Segments

After estimation of model parameters, the next step is to describe the overall shape of the building footprint by further identifying adjacent facade pairs and determining the intersection of the facade surfaces. The adjacency of facades is usually described by an adjacency matrix AM that is built up via connectivity analysis [21], [42]. Identified adjacent facade segments are used to determine the vertex points (i.e., facade intersection lines in 3-D) by computing the intersection points between any adjacent facade pair.

Determination of these intersection points can sometimes become difficult if the transition points are segmented as isolated small clusters (also referred to as conflicting segments) rather than part of the corresponding adjacent facade segments. As a consequence, it gets complicated to find a legitimate adjacent facade pair from which intersection points should be computed. To resolve this issue, conflicting segments must be removed prior to vertex point computation. To illustrate how they are removed in an automatic manner, an example is shown in Fig. 8. The labeled line segments indicate the reconstructed facade segments of two different buildings A and B. The endpoints of each segment are denoted as “ Δ ” and “ \bullet .” AM represents the built adjacency matrix, where “1” and “0” denote the adjacent and nonadjacent conditions, respectively. Among the labeled segments, segments [7 4 1 3 6] are “valid” facades, while segments [2 5 8] are the conflicting segments.

The following steps are performed for automatic removal of these conflicting segments.

- 1) The connected series matrix ConnSeg is determined from AM such that rows of ConnSeg represent a set of distinct series of adjacently connected segments, e.g., the i th row of ConnSeg $Seg_i = \{s_j \mid j = 1, \dots, n\}$ represents n segments (i.e., $s_1 \sim s_n$) that are adjacently connected. In Fig. 8, since there exist only two series of adjacently connected facade segments (i.e., belonging to two buildings), ConnSeg therefore consists of two rows only, where the first row contains facade segments $Seg_1 = [2 4 5 7]$ while the second row comprises of segments $Seg_2 = [1 3 6 8]$.
- 2) For each segment, the largest segment that is connected to each endpoint can be identified. Their indices are

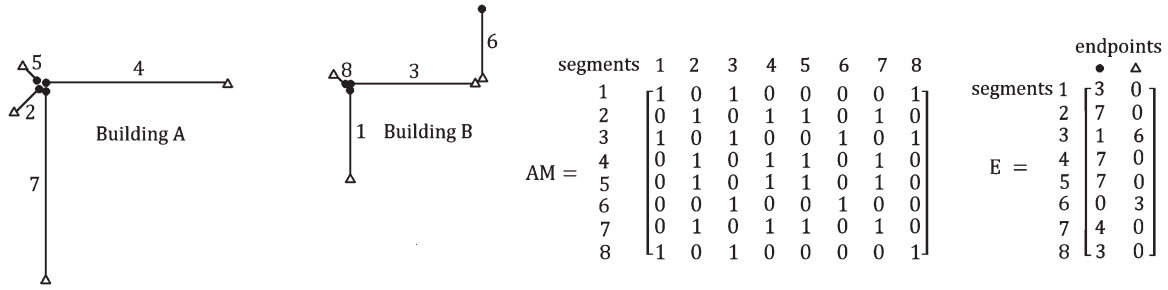


Fig. 8. Example illustrating the removal of conflicting segments.

TABLE I
PROCEDURE TO REMOVE CONFLICTING SEGMENTS

Given: Endpoints matrix EndPts containing endpoints of the reconstructed façade segments & Adjacency matrix $AM = \{a_{vw} | v, w = 1, \dots, k\}$ such that $a_{v=v} = 1$ & $a_{v \neq w} = 1$ only for conditions where façade Segments v and w are adjacent to each other, otherwise $a_{v \neq w} = 0$. k is the total number of façades reconstructed.

1. Initialize: RemoveSeg := \emptyset
2. **for** $i = 1$ to number of rows in ConnSeg
3. Select the adjacently connected series of segments $Seg_i = ConnSeg(i, :)$
4. **if** number of elements in $Seg_i > 2$
5. **for** $j = 1$ to n (number of segments in Seg_i)
6. Extract adjacent segments of s_j in Seg_i from AM.
7. Extract endpoints of all the segments in Seg_i from EndPts.
8. Compute the Euclidean distance of both endpoints in s_j from end points of all other segments $s_k (j \neq k)$ in Seg_i .
9. Find out the largest adjacent segment l_1 connected to one endpoint. Similarly find the largest adjacent segment l_2 connected to the other endpoint of s_j .
10. Store l_1 and l_2 in a two column matrix $E(j, 1:2)$. In case of no adjacent segment to any of the endpoints, insert 0 in the corresponding position in E.
11. **end for**
12. Separately apply union operation to both column of E and store the unique result after excluding 0s in the row matrix RetainSeg.
13. Compare Seg_i and RetainSeg and insert segments of Seg_i that are not members of RetainSeg in the matrix RemoveSeg($i, :$) for removal.
14. **end if**
15. **end for**

recorded in a two column matrix E that captures such an “endpoint”–“largest segment” relationship, e.g., the “•” endpoint of segment 1 shown in Fig. 8 is connected to two segments 8 and 3. Since segment 3 has a larger length than segment 8, therefore 3 is assigned to this endpoint of segment 1 in E . The endpoint matrix E for both buildings is depicted on the right side of Fig. 8. Zeros in E represent the condition of no adjacent facade at that endpoint.

- 3) Applying union operation to all elements in E results in a matrix RetainSeg whose elements contain all building facades that should be retained. Conflicting facades, i.e., the ones that are not part of RetainSeg, are removed. For the example shown in Fig. 8, the union of elements in E gives the RetainSeg [1 3 4 6 7] (zeros are not considered). Subsequently, the segments that are not part of RetainSeg, namely, [2 5 8], are removed.

The pseudocode for the aforementioned procedure is given in Table I.

After removing the conflicting segments, the vertex points are computed from the intersection of valid adjacent segments to complete the reconstruction process.

G. Refining Reconstructed Facades

Sometimes the reconstructed facades remain either incomplete or are broken into more than one segment due to the following reasons: 1) Higher building structures present nearby can partly (or fully) occlude the facades of lower buildings, and 2) due to the geometrical shape, only very few points are available at some parts of building facades. In order to overcome this issue, in this section, we propose a procedure that tries to refine the reconstructed facades by inserting additional

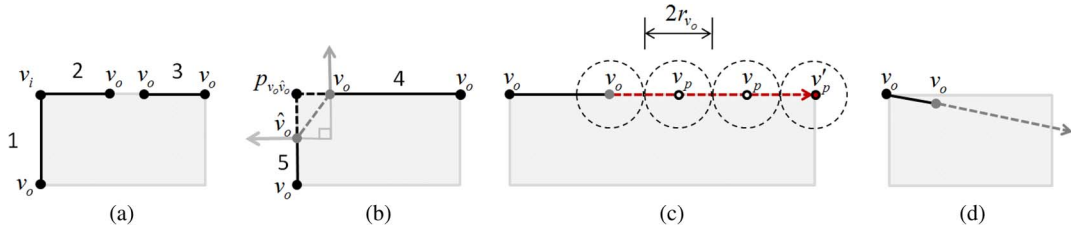


Fig. 9. Vertices for refinement. Gray rectangles depict the 2-D building footprint from the top. (a) Shows the total of five vertices, out of which four are open endpoint vertices and one is an intersection vertex computed from the intersection of segments 1 and 2. (b) Depicts the situation where ignoring condition $C3$ would yield false segment addition shown as a gray dotted line. The gray arrows indicate the local orientation angle θ at open vertices v_o and \hat{v}_o . Two black dotted lines represent the two correct inserted segments between v_o and \hat{v}_o . (c) Depicts the recursive extension procedure of the open vertex v_o . v_p represents the intermediate extension points, whereas v'_p denotes the final point. (d) Illustrates the direction of extension of the gray open vertex over the roof region. This can happen if we only consider h_{\max} and ignore the local standard deviation of height h_σ .

segments between the broken regions and extend those facades that remain incomplete.

Vertex points computed from the previous section are separated into two types: The first type consists of vertices that are computed from the intersection of two adjacent facades, while the second type consists of the other vertices representing “open” endpoints. For each series Seg_i made up of n segments (i.e., the i th updated row of the ConnSeg matrix after removing conflicting segments), there exist two open vertices and $(n - 1)$ intersection vertices. Refinement operations including insertion of additional segment to connect broken facades and extension of incomplete facades are carried out only on the second type, i.e., open endpoint vertices.

If we denote an open end vertex as v_o and an intersection vertex as v_i [see Fig. 9(a)], then the refinement procedure for any one particular adjacently connected series Seg_i having two open vertices is described in Table II.

In Table II, steps 2–4 tries to cope with the broken facades, while steps 5–7 deals with the incomplete facades. Conditions $C1$ and $C2$ in step 4 imply that the two segments are considered part of the same (broken) building facade if both segments are not far enough from each other and at the same time possess data points in between that have close maximum height values. h_{\max} is taken as the mean of at least ten maximum height values (i.e., if there are less than ten points available, then h_{\max} is taken as the mean of all of those point). If conditions $C1$ and $C2$ in step 4 are met, the algorithm then checks condition $C3$. If the two segments belong to the same facade, a segment with vertices (v_o, \hat{v}_o) is inserted, which fills the empty (i.e., broken) regions of the facade. On the contrary, if the open vertex pair $v_o \hat{v}_o$ is not part of the same facade but rather belongs to two different facade segments (determined via the difference in the local orientation angle $> 45^\circ$), then point of intersection $p_{v_o \hat{v}_o}$ is computed, and instead of inserting one segment, two segments with vertices $(v_o, p_{v_o \hat{v}_o})$ and $(\hat{v}_o, p_{v_o \hat{v}_o})$ are inserted. Fig. 9(b) graphically depicts such a situation where the gray open vertices of segments 4 and 5 are (assumed to be) within 2ε distance but have a difference in the local orientation angle of 90° . The gray dotted line shows the addition of new segment without checking condition $C3$. When $C3$ is taken into account, two segments shown in black dotted line are inserted.

In contrast, if any of the conditions $C1$ or $C2$ fail, then the algorithm tries to extend the open vertex point v_o by imposing constraints $C4$ and $C5$ present in step 7. Similar to $C2$, the condition $C4$ ensures that the extended point has the closer

maximum height value. The condition $C5$ ensures that the local 3-D points have certain standard deviation. It is necessary to make sure that the extension is not carried out in the direction that deviates from the facade footprint, i.e., it avoids the extension if the local 3-D points around v_p belong to other nonfacade objects, e.g., roofs, etc. The problem is illustrated in Fig. 9(d), where the gray open vertex can potentially extend over the roof region if the condition $C5$ in step 7 is ignored. Thus, imposing this constraint helps in limiting this false extension.

Finally, the computed vertex points (i.e., the intersection vertices and the open vertices before and after refinement) along with their estimated model parameters are used to reconstruct the 3-D model of the building facades.

IV. EXPERIMENTAL RESULTS AND VALIDATION

A. Data Set

To validate our approach, we tested the algorithm on TomoSAR point clouds generated from a stack of 25 TerraSAR-X high spotlight images from ascending orbit only using the Tomo-GENESIS software developed at the German Aerospace Center [67]. The test area covers approximately 2 km^2 in the high-rise part of the city of Las Vegas. The number of TomoSAR points in the area of interest is about 1.2 million. Fig. 10(a) shows the optical image of our test area, while Fig. 10(b) shows the corresponding TomoSAR point cloud in Universal Transverse Mercator (UTM) coordinates.

B. Results—Extraction of Facade Points

The result of applying the SD estimation procedure is illustrated in Fig. 11(a). The two parameters r (radius of the neighborhood cylinder) and d are empirically set to 5 and 0.9 m, respectively, according to the point density of the data set. One can observe that the TH value influences the number of extracted facade points. A lower TH value results in higher completeness but lower correctness. In [42], we showed the results of estimating SD with varying area sizes and found that a kernel window of size $3 \times 3 \text{ m}^2$ and threshold TH value of about 2 points/ m^2 results in the best tradeoff in terms of completeness and correctness with this class of data. Here, 2 points/ m^2 works well for high-rise buildings but might ignore relatively smaller facades. Therefore, to extract lower facades (and also to automate the procedure), we set the TH

TABLE II
REFINEMENT PROCEDURE

Given: Any one particular adjacently connected series Seg_i

1. Select one of the two open vertex point v_o .
2. Find the nearest open vertex point \hat{v}_o from v_o that belongs to another adjacently connected series $Seg_j (j \neq i)$ and compute their midpoint mid_{v_o} .
3. Locally compute maximum height value h_{\max} and height standard deviation h_{σ} of 3-D points within neighborhood vicinity of v_o , \hat{v}_o and mid_{v_o} . Additionally, compute local orientation θ at v_o and \hat{v}_o .
4. Check following three conditions (C1, C2, C3):

$$C1: \|\hat{v}_o - v_o\|_2 < 2\varepsilon$$

$$C2: abs(h_{\max \text{ at } v_o} - h_{\max \text{ at } mid_{v_o}}) < T_h \ \& \ abs(h_{\max \text{ at } \hat{v}_o} - h_{\max \text{ at } mid_{v_o}}) < T_h$$

$$C3: abs(\theta_{v_o} - \theta_{\hat{v}_o}) > 45^\circ$$

If conditions C1 & C2 are true, then check condition C3.

If C3 is also true

Insert two additional segments with vertices $(v_o, P_{v_o\hat{v}_o})$ and $(\hat{v}_o, P_{v_o\hat{v}_o})$ and then go to step 1 ($P_{v_o\hat{v}_o}$ is the point of intersection between two segments).

else

Insert an additional segment with vertices v_o and \hat{v}_o and go to step 1

end if

else

proceed to the next step 5.

end if
5. Determine point v_p at a distance $2r_{v_o}$ away from v_o using the estimated model parameters (r_{v_o} is the radius used for defining local neighborhood around v_o)
6. Similarly as step 3, locally compute h_{\max} and h_{σ} of 3-D points within neighborhood vicinity of v_p .
7. Check following two conditions (C4, C5)

$$C4: abs(h_{\max \text{ at } v_p} - h_{\max \text{ at } v_o}) < T_h$$

$$C5: abs(h_{\sigma \text{ at } v_p} - h_{\sigma \text{ at } v_o}) < T_{\sigma}$$

if both C4 and C5 are true, then recursively extend the vertex v_o towards v_p (see Figure 9(c)) and finally add a new segment with vertices (v_o, v_p')
8. Continue steps 1 to 7 for the second open vertex point of the series Seg_i

to the maximum of SD histogram value. This, as described in Section III, includes not only the facade points but additionally also some nonfacade points with relatively high SD , e.g., roof points. To reject these points from the set of extracted points after SD thresholding, surface normal information is utilized. Fig. 11(b) shows the extracted facade points by retaining only those points having normals between $\pm 15^\circ$ from the horizontal axis (or equivalently $\pm 90^\circ$ from the vertical axis).

C. Results—Automatic Clustering of Extracted Facade Points

Once the facade points are extracted out, the next step is to cluster them into segments, where each segment corresponds to an individual facade. For this, we apply the clustering procedure using the cylindrical neighborhood definition and cluster all of

the points with parameter settings: $\varepsilon = r = 5$ m and $MinPts = 2$. Here, an important point to notice is that two buildings are considered distinct only in a case when points belonging to facades of two different buildings are separated by ε . Setting ε too small can cause points belonging to a single cluster (i.e., corresponding to an individual facade) to break into more than one cluster. On the other hand, larger values of ε tend to merge points of the nearby facades into one cluster. The value of ε is therefore empirically chosen according to the length and distance among the buildings in the area of interest and implicitly indicates the assumption that two individual facades that belong to different buildings are farther apart than the 5-m radius.

Setting parameter $MinPts$ equal to 2 implies that points are connected to one cluster even if there is a single neighboring point among them. This parameter helps in removing outliers

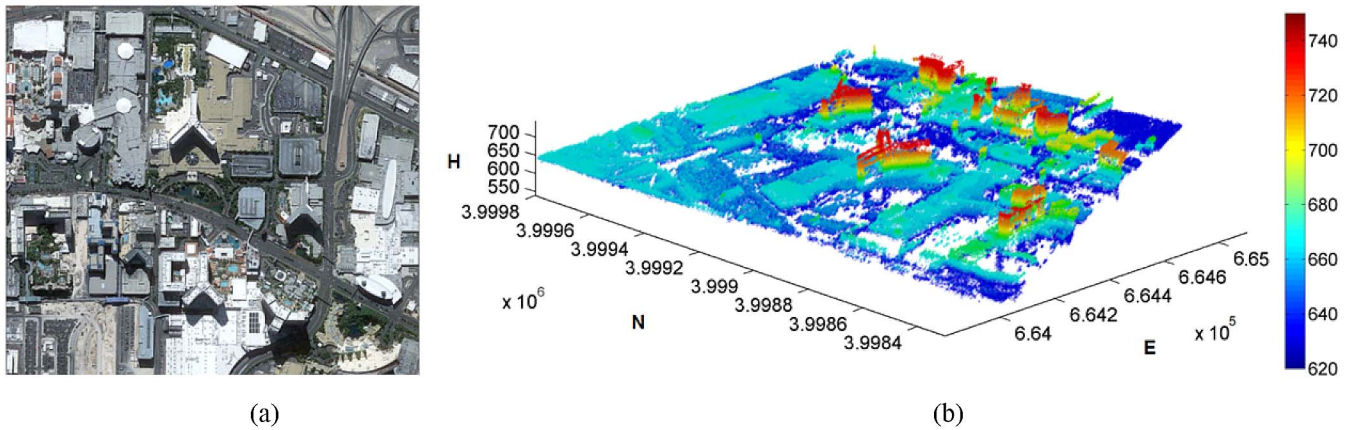


Fig. 10. Dataset. (a) Optical image of the test area in Las Vegas. Copyright Google. (b) TomoSAR points in UTM coordinates of the corresponding test image. The height of TomoSAR points is color-coded [unit: meter].

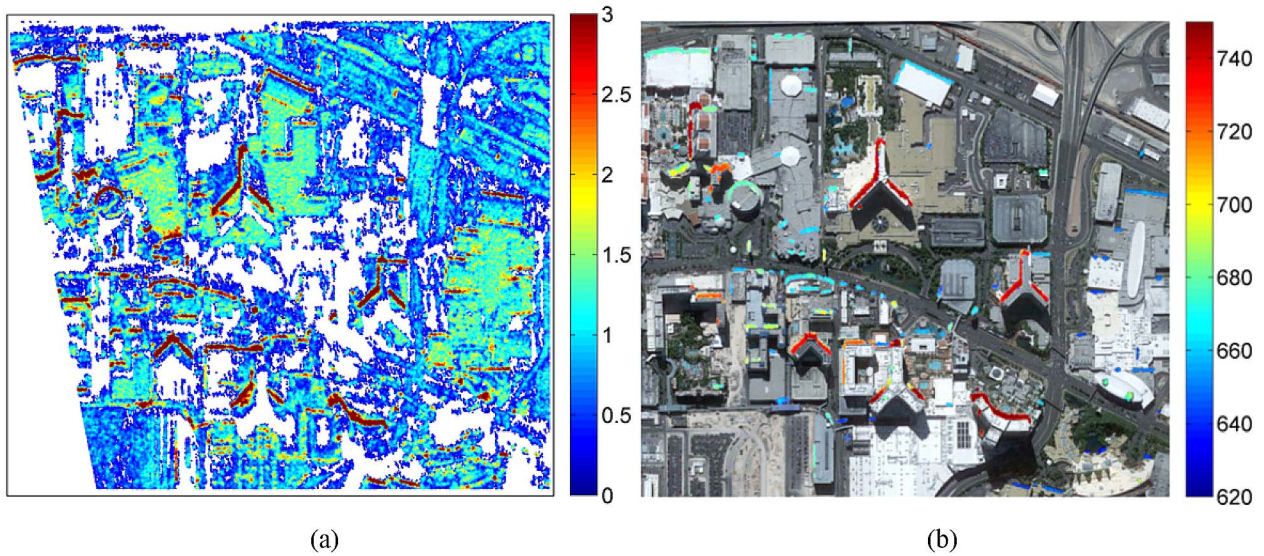


Fig. 11. Facade point extraction. (a) Scatterer (point) density with radius $r = 5$ m and inliers $d = 0.9$ m. (b) Extracted building facade points. Colobar indicates SD and height in meters in (a) and (b), respectively.

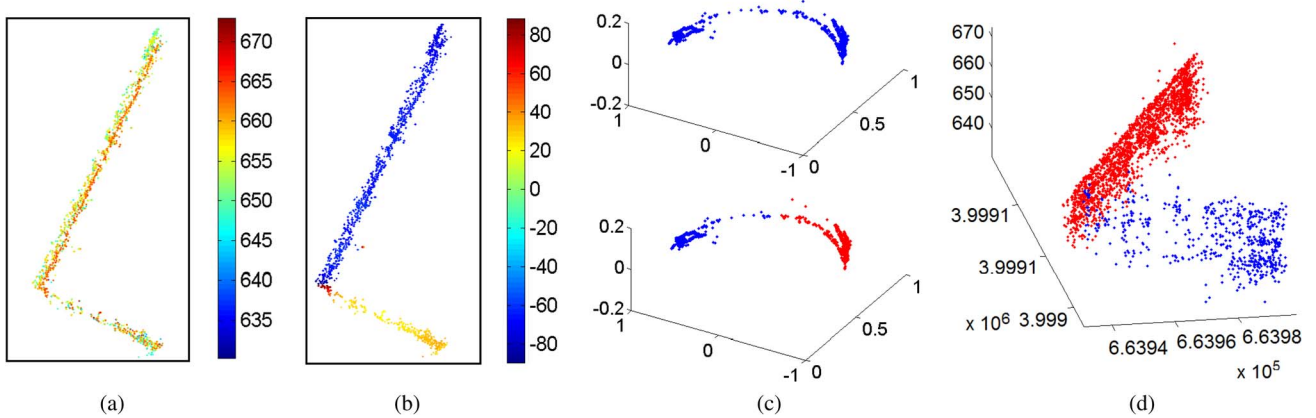


Fig. 12. Fine clustering results after applying mean shift clustering using Gaussian kernel with bandwidth $b = 0.4$ to the coarsely clustered segments in their normal feature space (in GI domain). (a) TomoSAR points of one particular density-connected cluster (top view). The colorbar indicates height in meters. (b) Corresponding orientation angle in degrees. (c) Nonclustered (top) and clustered (bottom) points in the Gaussian image of points in (a). (d) Resulting clustered points in 3-D.

that do not have any neighboring point and produce clusters similar to the clusters obtained from the dendrogram cut at ϵ in case of hierarchical clustering using single link metric [43].

Increasing MinPts can help in retaining more stable core points but, on the other hand, can also break the clusters into two or more clusters. This property is sometimes useful in cases when

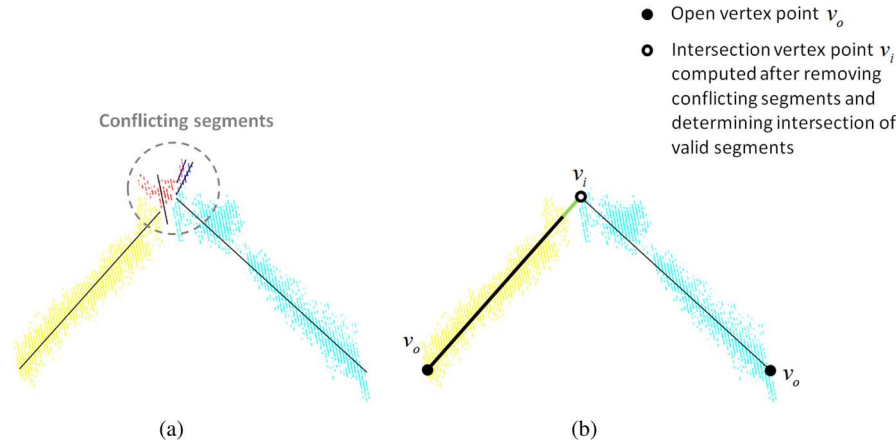


Fig. 13. Removal of conflicting segments and vertex point computation. (a) Shows the case of five clustered point segments belonging to one particular building. At the transitional regions, there are conflicting (smaller) clusters, shown in gray dotted circle, that behave as noisy segments and should be removed before computing the vertex points. (b) Shows the retained segments and their corresponding vertices after removing conflicting segments.

different clusters are merged together by a thin line of points. Estimating the exact value of MinPts is, however, very much dependent on the data set, and certain heuristics based on the “thinnest” cluster in the data set, e.g., k -distance graph, can be employed [43].

In order to reconstruct individual facades, these density-based coarse clusters need to be further clustered. To this end, mean shift clustering has been applied to the coarsely clustered segments in their normal feature space (in GI domain). Fig. 12(b) shows the estimated orientation angle θ for extracted facade points from a single building shown in Fig. 12(a). The variation in orientation angle is quite evident and allows mean shift to cluster points having similar orientations together. Further separation of points in the spatial domain is also required in some cases where the spatially separated points are clustered into one segment. This happens when these points belonging to different facades have similar normals and are spatially closer. Density-based clustering is therefore again applied for spatial separation of the clusters within clusters.

D. Results—Reconstructing Facades

Prior to reconstruction, the segmented facades are first classified to flat and curved surfaces by analyzing the derivatives of the local orientation angle θ . A slope value θ' of 0.3 ($\approx 17^\circ$) is set by empirically testing the computed orientation angles of all of the buildings in the area of interest to distinguish flat and curved surfaces.

After identification, appropriate model parameters are estimated from the core points of the individual clusters. Vertex points are then determined by computing intersections of the adjacent segment pairs. However, in doing so, smaller clusters occurring at facade transition regions behave as noisy segments in the reconstruction procedure. A practical example of these so-called conflicting segments is shown in Fig. 13(a). Following the procedure explained in Table I (see Section III-F), the conflicting (reconstructed) segments occurring at the transitional regions of individual buildings are removed prior to the vertex point computation as exemplified in Fig. 13(b).

Once these transitional clusters are removed, the intersection vertices are determined by computing the intersection point of the two adjacent facades. Refinement operation is then carried out on the open vertices to insert additional segments between the broken facade regions followed by extension of incomplete reconstructed facades.

Fig. 14(a) and (b) depicts the reconstructed facade models of the area of interest before and after refinement, respectively. Green lines show reconstructed facade footprint before refinement. The blue lines indicate additional segments that are added between the vertices of those broken facades that meet the conditions present in step 4 in Section III-G, while the red lines are subsequent extensions of the open vertices after filling the break regions.

E. Results—Validation

The actual ground truth data are missing for exact qualitative evaluation of the approach. In order to provide some quantitative measures of the algorithm performance, we manually counted the actual number of facades that were to be reconstructed. A total of 141 facades are present in the data set, out of which 7 are curved facades and the remaining 134 are flat. Prior to the refinement operation, the algorithm reconstructed a total of 176 facades, i.e., higher than the actual facades present in the data set. As already stated in Section III-G, this is because some individual facades have been broken down into two or more segments due to discontinuity in the number of points available in the data set. After refinement, 29 insertion segments (27 single and 2 double based on the condition $C3$ in Table II) are added between the broken facade regions, whereas 43 facades have been extended. In the final reconstruction, we obtain 147 reconstructed facades, i.e., all 141 facades are successfully reconstructed; among them, five facades remain broken (counted as additional five facades), and there is one case of false alarm which will be explained later. Besides the five cases, we also find seven facades that are not extended and therefore remain incomplete. This is, however, due to the inadequate number of points available in the data.



(a)



- Reconstructed facades without refinement
- Additional facades added between the broken facade regions
- Extension of facades after adding additional segments

(b)

Fig. 14. Reconstructed facades. (a) Two-dimensional view of the facade footprints overlaid onto the optical image prior to refinement. (b) Two-dimensional view of the facade footprints overlaid onto the optical image after refining with parameter settings $T_h = 5$ m and $T_\sigma = 2.5$ m.

As mentioned earlier, there is also one case, shown in Fig. 15, which is considered as false positive (i.e., a facade not actually present but reconstructed by the algorithm). As can

be seen in Fig. 15(c), the reconstructed segment is actually a bridge for pedestrian crossing. Higher number of scatterers is retrieved over the bridge due to its apparently metallic structure.

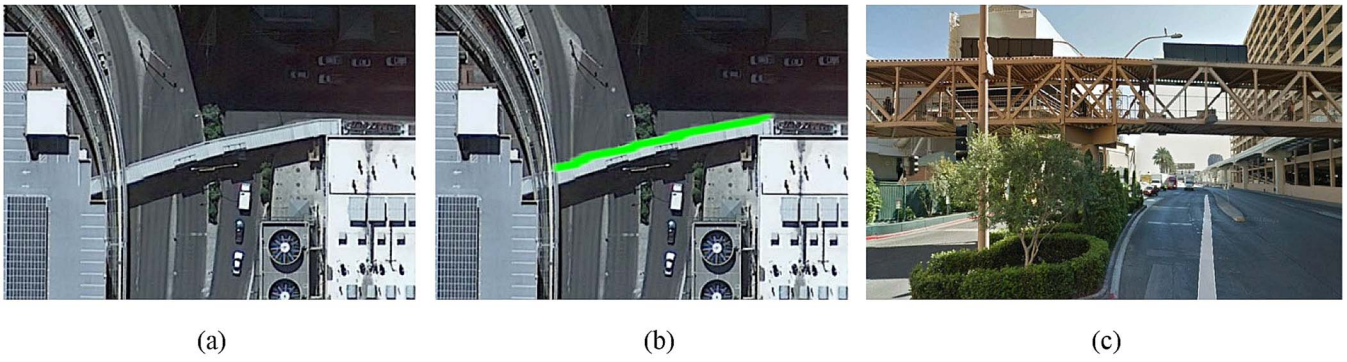


Fig. 15. Case depicting wrong reconstruction of a pedestrian bridge. (a) Optical image of the bridge. (b) Overlays the reconstructed segment in green onto the optical image of (a). (c) Side view of the bridge (copyright Google Street View).

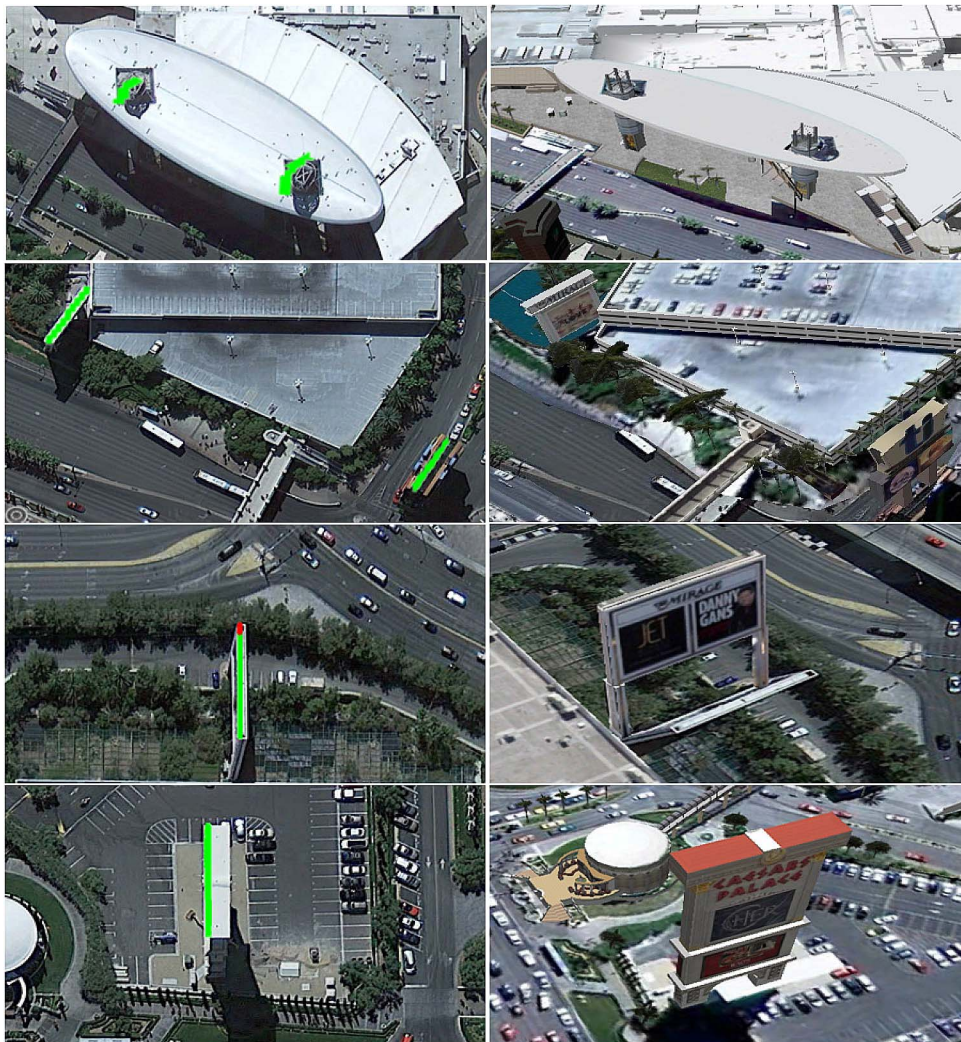


Fig. 16. Reconstructed facades on the left. Their 3-D view on the right (copyright Google Earth).

Moreover, the bridge is also covered from the top, and therefore, scatterers are obtained at the top and bottom and on the metallic rods connecting the upper and lower surfaces of the bridge. The estimated surface normal of these scatterers thus gives a higher horizontal component, and as a consequent, these scatterers are wrongly classified as facade points by satisfying

both extraction constraints: higher SD and higher horizontal component of the surface normals.

It is also interesting here to mention that, in Fig. 14, some small vertical structures on roofs of the buildings or on ground are very well reconstructed. Fig. 16 shows some examples of such objects that might visually appear (or interpreted) as false

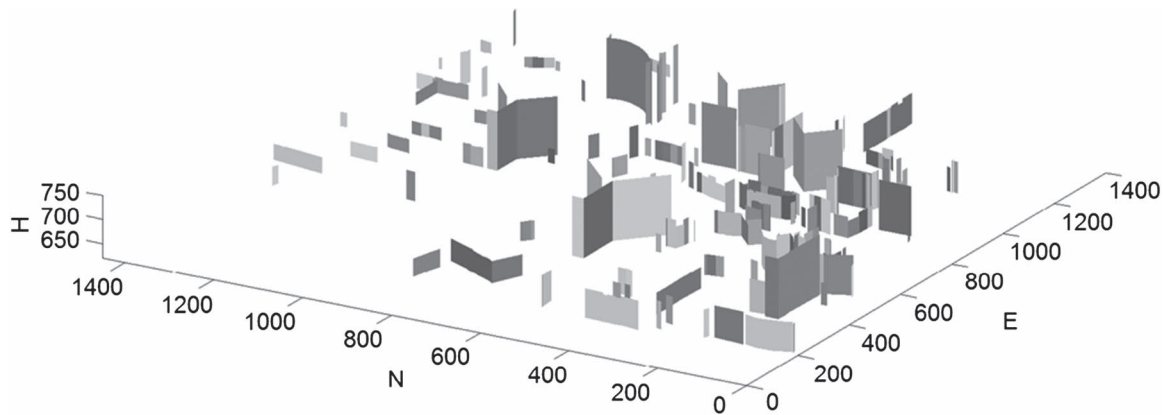


Fig. 17. Three-dimensional view of the final facade reconstruction. The axis is in meter range and has been translated to the origin for better metric clarity by subtracting the UTM easting and northing values by their respective minimum values present in the reconstructed vertices.

reconstructions in Fig. 14 but are actually vertical structures (e.g., advertisement boards, monuments, etc.).

Finally, in Fig. 17, we present the final reconstructed facades in 3-D. As depicted in [42], the shown reconstructed facade 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 that could potentially be used to monitor and visualize the dynamics of urban infrastructure in very high level of details.

V. OUTLOOK AND CONCLUSION

In this paper, we have presented an automatic (parametric) approach for robust facade reconstruction for large areas using TomoSAR point clouds. The approach is modular and works directly on unstructured 3-D points. It allows for a robust reconstruction of both higher facades and lower height structures, and hence, it is well suited for urban monitoring of larger areas from space. A few points, however, need to be addressed.

- 1) During SD estimation, the continuity of an individual facade can be broken due to a limited number of available points. This may result into two or more segments of the same facade. We attempted to cope with this problem by refining the reconstructed facade footprints via insertion and extension operations. Still, the lack of measurements prevents the complete resolution of this problem. The use of 2-D ground plans or cadastral maps can be helpful in this case.
- 2) Since the satellite orbits are bound to pass close to the poles of Earth, we may fail to reconstruct building facades facing North or South due to the missing of measurements. One way to rectify this is by using fused point clouds (i.e., both ascending and descending) and/or inserting new segments by simply connecting the endpoints of the missing facades if they match a certain criteria to get the complete shape of the building footprint.
- 3) The presented approach is a much better option to detect the shape of the building when dense points on the facades are available. However, in cases (usually for lower height buildings) when no or few facade points are avail-

able, one can try to extract roof points and reconstruct the 2-D footprint. This could help in resolving the problems related to the visibility of facades mainly pointing toward the azimuth direction.

In the future, we will work over these considerations and will extend the algorithm toward object-based TomoSAR point cloud fusion and automatic building roof reconstruction.

REFERENCES

- [1] J. Hu, S. You, and U. Neumann, "Approaches to large-scale urban modeling," *IEEE Comput. Graph. Appl.*, vol. 23, no. 6, pp. 62–69, Nov./Dec. 2003.
- [2] J. Benner, A. Geiger, and K. Leinemann, "Flexible generation of semantic 3-D building models," in *Proc. Int. ISPRS/EuroSDR/DGPF-Workshop Next Gen. 3-D City Models*, K. Gröger, Ed., Bonn, Germany, 2005, pp. 18–22, EuroSDR Publication No. 49.
- [3] J. Lee and S. Zlatanova, "A 3-D data model and topological analyses for emergency response in urban areas," in *Geospatial Information Technology for Emergency Response*, L. Zlatanova, Ed. New York, NY, USA: Taylor & Francis, 2008.
- [4] S. Kemeç, S. Duzgun, S. Zlatanova, D. I. Dilmen, and A. C. Yalciner, "Selecting 3-D urban visualization models for disaster management: Fethiye tsunami inundation case," in *Proc. 3rd Int. Conf. Cartogr. GIS*, Nessebar, Bulgaria, Jun. 15–20, 2010, pp. 1–9.
- [5] K. Lee, Y. Kim, S. I. Cho, and K. Choi, "Building detection in augmented reality based navigation system," in *Proc. 13th Int. Conf. Multimedia Model.—Volume Part II*, 2006, pp. 544–551.
- [6] J. Döllner, T. H. Kolbe, F. Liecke, T. Sgouros, and K. Teichmann, "The virtual 3-D city model of Berlin—Managing, integrating, and communicating complex urban information," in *Proc. 25th UDMS*, Aalborg, Denmark, May 15–17, 2006, pp. 1–12.
- [7] A. Jochem, B. Hoefle, M. Rutzinger, and N. Pfeifer, "Automatic roof plane detection and analysis in airborne LiDAR point clouds for solar potential assessment," *Sensors*, vol. 9, no. 7, pp. 5241–5262, 2009.
- [8] J. Kirtner, "Using LiDAR data in wireless communication system design," in *Proc. Am. Soc. Photogramm. Remote Sens. Conf.*, Washington, DC, USA, 2000, pp. 1–8.
- [9] M. Over, A. Schilling, S. Neubauer, and A. Zipf, "Generating Web-based 3-D city models from OpenStreetMap: The current situation in Germany," *Comput., Environ. Urban Syst.*, vol. 34, no. 6, pp. 496–507, Nov. 2010.
- [10] X. Yin, P. Wonka, and A. Razdan, "Generating 3-D building models from architectural drawings: A survey," *IEEE Comput. Graph. Appl.*, vol. 29, no. 1, pp. 20–30, Jan./Feb. 2009.
- [11] First App for LiDAR Data Visualisation on iPad, Accessed Mar. 21, 2014. [Online]. Available: <http://geoawesomeness.com/first-app-lidar-data-visualisation-ipad/>
- [12] M. Quartulli and M. Datcu, "Stochastic geometrical modeling for built-up area understanding from a single SAR intensity image with meter resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1996–2003, Sep. 2004.

- [13] A. Ferro, D. Brunner, and L. Bruzzone, "An advanced technique for building detection in VHR SAR images," in *Proc. SPIE Eur. Remote Sens.*, Berlin, Germany, 2009, vol. 7477, pp. 74770V-1–74770V-12.
- [14] 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.
- [15] J. D. Wegner, U. Soergel, and A. Thiele, "Building extraction in urban scenes from high resolution InSAR data and optical imagery," in *Proc. URBAN*, 2009, pp. 1–6.
- [16] P. Gamba, B. Houshmand, and M. Saccani, "Detection and extraction of buildings from interferometric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 1, pp. 611–617, Jan. 2000.
- [17] A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, 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.
- [18] F. Xu and Y.-Q. Jin, "Automatic reconstruction of building objects from multi-aspect meter-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2336–2353, Jul. 2007.
- [19] H. Sportouche, F. Tupin, and L. Denise, "Extraction and three-dimensional reconstruction of isolated buildings in urban scenes from high-resolution optical and SAR spaceborne images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3932–3946, Oct. 2011.
- [20] U. Stilla, U. Soergel, and U. Thoennessen, "Potential and limits of InSAR data for building reconstruction in built-up areas," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 1/2, pp. 113–123, Jun. 2003.
- [21] A. Sampath and J. Shan, "Segmentation and reconstruction of polyhedral building roofs from aerial LiDAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1554–1567, Mar. 2010.
- [22] R. Bamler, M. Eineder, N. Adam, S. Gernhardt, and X. Zhu, "Interferometric potential of high resolution spaceborne SAR," in *Proc. PGF*, 2009, pp. 407–419.
- [23] S. Mezzasoma *et al.*, "COSMO-SkyMed system commissioning: End-to-end system performance verification," in *Proc. IEEE Radar Conf.*, Rome, Italy, May 2008, pp. 1–5.
- [24] N. Adam, X. Zhu, and R. Bamler, "Coherent stacking with TerraSAR-X imagery in urban areas," in *Proc. JURSE*, Shanghai, China, 2009, pp. 1–6.
- [25] G. Fornaro, D. Reale, A. Pauciuillo, 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. Newslett.*, Dec. 2012.
- [26] 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.
- [27] X. Zhu and R. Bamler, "Let's do the time warp: Multi-component non-linear motion estimation in differential SAR tomography," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 735–739, Jul. 2011.
- [28] 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.
- [29] D. Reale, G. Fornaro, A. Pauciuillo, 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.
- [30] 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.
- [31] Y. Wang, X. Zhu, and R. Bamler, "An efficient tomographic inversion approach for urban mapping using meter resolution SAR image stacks," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, p. 1, Jul. 2012.
- [32] D. Reale, G. Fornaro, and A. Pauciuillo, "Extension of 4-D SAR imaging to the monitoring of thermally dilating scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 12, pp. 5296–5306, Dec. 2013.
- [33] X. Zhu and R. Bamler, "Super-resolving SAR tomography for multi-dimensional imaging of urban areas," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 51–58, Jul. 2014.
- [34] 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.
- [35] P. Pasquali, C. Prati, F. Rocca, and M. Seymour, "A 3-D SAR experiment with EMSL data," in *Proc. IGARSS*, 1995, pp. 784–786.
- [36] G. Fornaro, F. Serafino, and F. Lombardini, "Three-dimensional multipass SAR focusing: Experiments with long-term spaceborne data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 702–714, Apr. 2005.
- [37] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 34–44, Jan. 2005.
- [38] 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.
- [39] G. Fornaro, F. Serafino, and D. Reale, "4-D SAR imaging: The case study of Rome," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 236–240, Apr. 2010.
- [40] F. Lombardini, *Diff-Tomo opening of the urban SAR pixel: Single-look 4-D and nonuniform motion 5D extensions*. Presented at FRINGE Workshop, 2011. [Online]. Available: http://earth.eo.esa.int/workshops/fringe2011/files/Lombardini_FRINGE2011.pdf
- [41] M. Shahzad, X. Zhu, and R. Bamler, "Façade structure reconstruction using spaceborne TomoSAR point clouds," in *Proc. IEEE IGARSS*, Munich, Germany, 2012, pp. 467–470.
- [42] X. Zhu and M. Shahzad, "Façade reconstruction using multi-view spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3541–3552, Jun. 2014.
- [43] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. KDD*, Portland, OR, USA, 1996, pp. 226–231.
- [44] Y. Liu and Y. Xiong, "Automatic segmentation of unorganized noisy point clouds based on the Gaussian map," *Comput.-Aided Design*, vol. 40, no. 5, pp. 576–594, May 2008.
- [45] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 8, pp. 790–799, Aug. 1995.
- [46] G. Sithole and G. Vosselman, "Experimental comparison of filtering algorithms for bare-earth extraction from airborne laser scanning point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 59, no. 1/2, pp. 85–101, Aug. 2004.
- [47] G. Vosselman, "Slope based filtering of laser altimetry data," *Int. Archives Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 33, pt. B3, pp. 935–942, 2000.
- [48] G. Sohn and I. Dowman, "Terrain surface reconstruction by the use of tetrahedron model with the MDL criterion," *Int. Archives Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 34, pt. 3A, pp. 336–344, 2002.
- [49] H. Gross, U. Thoennessen, and W. Hansen, "3-D-modeling of urban structures," in *Proc. Int. Archives Photogramm. Remote Sens.*, Vienna, Austria, 2005, vol. XXXVI/3-W24, pp. 137–142.
- [50] F. Rottensteiner and C. Briese, "A new method for building extraction in urban areas from high-resolution LiDAR data," in *Proc. Int. Archives Photogramm., Remote Sens. Spatial Inf. Sci.*, 2002, vol. 34, pp. 295–301, Part 3A.
- [51] A. Sampath and J. Shan, "Building boundary tracing and regularization from airborne LiDAR point clouds," *Photogramm. Eng. Remote Sens.*, vol. 73, no. 7, pp. 805–812, Jul. 2007.
- [52] C. Mallet, F. Bretar, M. Roux, U. Soergel, and C. Heipke, "Relevance assessment of full-waveform LiDAR data for urban area classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 6, pp. S71–S84, Dec. 2011.
- [53] P. Dorninger and N. Pfeifer, "A comprehensive automated 3-D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds," *Sensors*, vol. 8, no. 11, pp. 7323–7343, 2008.
- [54] G. Forlani, C. Nardinocchi, M. Scaioni, and P. Zingaretti, "Complete classification of raw LiDAR data and 3-D reconstruction of buildings," *Pattern Anal. Appl.*, vol. 8, no. 4, pp. 357–374, Feb. 2006.
- [55] 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. 2006.
- [56] Y. Wang, X. Zhu, R. Bamler, and S. Gernhardt, "Towards TerraSAR-X street view: Creating city point cloud from multi-aspect data stacks," in *Proc. JURSE*, Sao Paulo, Brasil, 2013, pp. 198–201.
- [57] P. Holland and R. Welsch, "Robust regression using iteratively reweighted least-squares," *Commun. Statist., Theory Methods*, vol. 6, no. 9, pp. 813–827, 1977.
- [58] J. O. Street, R. J. Carroll, and D. Ruppert, "A note on computing robust regression estimates via iteratively reweighted least squares," *Amer. Statist.*, vol. 42, no. 9, pp. 152–154, 1988.
- [59] P. J. Huber, *Robust Statistics*. Hoboken, NJ, USA: Wiley, 1981.
- [60] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, "Surface reconstruction from unorganized points," *Comput. Graph.*, vol. 26, no. 2, pp. 71–78, Jul. 1992.
- [61] M. Hubert, P. Rousseeuw, and K. V. Branden, "ROBPCA: A new approach to robust principal components analysis," *Technometrics*, vol. 47, no. 1, pp. 64–79, Feb. 2005.

- [62] R. R. Wilcox, *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*, 2nd ed. New York, NY, USA: Springer-Verlag, 2010.
- [63] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.
- [64] M. P. D. Carmo, *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1976.
- [65] D. Comaniciu and P. Meer, "Mean shift: A robust approach towards feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [66] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965.
- [67] X. Zhu, Y. Wang, S. Gernhardt, and R. Bamler, "Tomo-GENESIS: DLR's tomographic SAR processing system," in *Proc. JURSE*, Sao Paulo, Brazil, 2013, pp. 308–313.



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 has been working toward the Ph.D. degree in the Chair of Remote Sensing Technology (LMF), Technische Universität München (TUM), München, Germany, since December 2011.

His research topic is 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) received the Bachelor's degree in space engineering from the National University of Defense Technology, Changsha, China, in 2006 and the M.Sc. degree, the Dr.-Ing. degree, and the "Habilitation" in the field of signal processing from Technische Universität München (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

In October/November 2009, she was a Guest Scientist with the Italian National Research Council (CNR)—Institute for Electromagnetic Sensing of the

Environment (IREA), Naples, Italy. Since May 2011, she has been a Scientist with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Wessling, Germany, where she is the Head of the Team Signal Analysis and with the Chair of Remote Sensing Technology, TUM. Since September 2013, she has also been a Helmholtz Young Investigator Group Leader and appointed as a TUM junior fellow. Her main research interests are advanced InSAR techniques such as high-dimensional tomographic SAR imaging and SqueeSAR, computer vision in remote sensing including object reconstruction and multidimensional 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 multi-/hyperspectral image analysis.