Detection of Buildings in Spaceborne TomoSAR Point Clouds via Hybrid Region Growing and Energy Minimization Technique

Muhammad Shahzad(1), Xiao Xiang Zhu(1,2)

(1) Helmholtz Young Instigators Group “SiPEO” ,
Technische Universität München,
Arcisstraße 21, 80333 Munich, Germany
Email: muhammad.shahzad@bv.tum.de

(2) Remote Sensing Technology Institute (IMF),
German Aerospace Center (DLR),
D-82234 Wessling, Germany
Email: xiao.zhu@dlr.de

Abstract—**In this paper, an automatic approach is presented to detect/extract buildings from spaceborne TomoSAR point clouds. The approach is systematic and allows robust detection of both tall and low height buildings and is, therefore, well suited for urban monitoring of larger areas from space. The presented approach is illustrated and validated by examples using TomoSAR point clouds generated from a stack of TerraSAR-X high resolution spotlight images covering an area of approximately 1.5 km^2 containing mostly moderate sized buildings in the city of Berlin, Germany. The depicted results validate the effectiveness of the proposed approach.**

I. INTRODUCTION

Modern spaceborne 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. These VHR data, when processed with advanced interferometric techniques, e.g., SAR tomography (TomoSAR) or persistent scatterer interferometry (PSI), become particularly suited for detailed urban mapping. In urban environments, TomoSAR, in particular, is able to retrieve up to 1 million scatterers/km^2. Geocoding these high density of scatterers into world coordinates enable the generation of high quality TomoSAR point clouds, containing not only the 3D positions of the scatterer location but also estimates of seasonal/temporal deformation, that are very attractive for generating 4-D city models from space. However there are some special considerations associated to these point clouds that are worth to mention [1]: 1) TomoSAR point clouds deliver moderate 3D positioning accuracy on the order of 1 m; 2) few number of images and limited orbit spread 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 [2]; 3) Due to the coherent imaging nature, temporally incoherent objects such as trees cannot be reconstructed from multipass spaceborne SAR image stacks; and 4) TomoSAR point clouds possess much higher density of points on the building façades due to side looking SAR geometry enabling systematic reconstruction of buildings footprint via façade points analysis. As depicted over smaller and larger areas in [1] and [3], façade reconstruction turns out to be an appropriate first step to detect and reconstruct building shape from these point clouds when dense points on the façade are available. Especially, 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 façade points. However, there are cases when no or only few façade points are available. This happens usually for lower height buildings and renders detection of façade points/regions very challenging. Moreover, problems related to the visibility of façades mainly pointing towards the azimuth direction can also cause difficulties in deriving the complete structure of an individual building. These problems motivate us to derive full 2-D building footprint via roof point analysis [4]. In this paper, we propose a hybrid approach based on region growing and energy minimization framework to automatically extract building roof/façade points directly from the 3D TomoSAR points. Two possible cases may exist which are solved sequentially:

- *Enough points on the building façades available*. In this case, the available façade information is incorporated to detect the probable building regions. Seed points around the probable regions are then chosen and a surface normals based region growing procedure is adopted to expand the probable regions.
- *No or very few façade points available*. The problem of extracting roof points from the remaining set of points is resolved by computing local height and planar features and formulating them into a simple energy minimization framework. Graph cuts are later employed to globally extract roof/building points.

II. METHODOLOGY

A. Probable building regions

The main idea of the algorithm is to first identify probable building regions. This is done by including prior knowledge into the processing chain by incorporating information pertaining to façades (i.e., regions corresponding to higher point density indicate probable façade regions). Thus, in this step, building façade points are extracted, segmented to points belonging to individual, and further reconstructed. Detailed processing schemes are described in [1][3].

B. Region growing procedure

The reconstructed façades provides direct hints for the desired building regions. The idea is to search in the nearby vicinity of the reconstructed façade to select seed points and then employ a surface normals based region growing algorithm to extract existing roof points. The seed points are selected by first computing the midpoint of the reconstructed façade segment and then analyzing the local height statistics in the two opposite directions perpendicular to the reconstructed segment. Figure 1 provides the graphical illustration of the seed selection procedure.

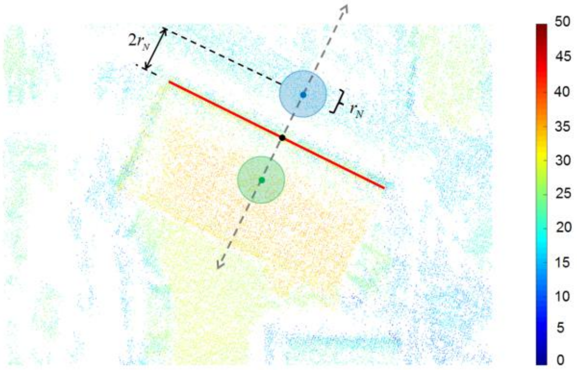

Figure 1. Graphical illustration of the seed point selection procedure. Top view of 3D TomoSAR points of a small region containing one building is depicted. The height in meters of each TomoSAR point is color-coded. The red line shows the reconstructed facade segment with black point as its midpoint. The green and blue points show the selected points in two opposite orthogonal directions with respect to the reconstructed façade (depicted as gray dotted lines). r_N is cylindrical radius used to determine local neighbors (i.e., points within shaded circles in respective colors). Local mean height of the blue and green points are 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 latter is true in the depicted case.

The seed points extracted 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 (i.e., points having angular difference between their surface normals less than θ_{ang} are added to the grown cluster). In addition to using the points based on surface normals only, however, one could also add ground points during growing procedure. An example of such a case will be where both roof and the neighboring ground are flat. A minimum height constraint is therefore needed and is incorporated in the growing procedure to restrict addition of points below a certain height $h_{\min}$. $h_{\min}$ is adaptively computed for each seed point as $m_p + \left(\text{height of } \mathbf{s}_p - m_p\right)\cdot fac$ where $m_p = m_b$ if $m_g > m_b$ or $m_p = m_g$ if $m_b > m_g$ (see Figure 1). The factor *fac* (empirically set to 0.55 in this work) can be adjusted between 0~1 to increase or decrease $h_{\min}$.

The above procedure extracts the existing roof points only for those buildings whose façades have been determined/reconstructed. However, as already mentioned, still there exist cases when no or very few points are available on the building façades rendering detection of façade points/regions very difficult. Moreover, if data is acquired from one aspect/direction e.g., ascending orbit only, the façades of lower buildings could get fully (or partly) occluded due to the presence of nearby higher building structure. As a consequence, existing roof points will be miss-detected. To cope with aforementioned issue, we sequentially formulate the remaining points into a simple energy minimization framework to extract those building/roof points where no information pertaining to façades could be exploited.

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 [5]-[7]:

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

where $\mathbf{P}=\left\{p_i \mid i=1,\dots,n\right\}$ denotes the set of remaining n 3D points, $\mathbf{N}$ denotes the set containing pairs of neighboring points $\left(p_i,p_j\right)$, f_{p_i} represents the label assigned to the point p_i, $\delta(\cdot)$ is an indicator function, and $w_{p_{ij}}$ is the weight of each pair of neighboring points $\left(p_i,p_j\right)$ and is defined as $w_{p_{ij}}=\exp\left(-\left\|\mathbf{p}_i-\mathbf{p}_j\right\|\right)$. The second summation term in (1) ensures the 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 *f* assigned to the set of points $\mathbf{P}$. It is defined as [5]:

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

where r_{p_i} is the plane residual distance feature computed for each point p_i by locally fitting robust RANSAC plane among its local neighbors and h_{p_i} is the differential height of the point p_i obtained after subtracting the height of the underlying terrain. 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}. Value of η less than 1 (e.g., η = 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.

Both features h_{p_i} and r_{p_i} are normalized to the scale of 0~1 by adopting the forms [5]: $h_{p_i} = \min\left(1, h_{p_i}/\varepsilon\right)$ and $r_{p_i} = \min\left(1, r_{p_i}/r_N\right)$ where ε (= 20m) is the tuning factor adjusting the sensitivity of the height feature (i.e., it ensures that all points having relative heights of greater than ε provides minimum data discrepancy term in (1)), and r_N (= 5m) is the radius size used to extract local neighbors.

D. *Approximation of terrain height*

The remaining point cloud $\mathbf{P}$ may contain roof points belonging to buildings for which no or very few façade points are available. To adopt the above energy based formulation to extract these remaining roof points, height of the underlying terrain surface is approximated by fitting a cubic polynomial surface to the non building/ground points via robust least absolute residuals (LAR) method. Ground points are extracted out via successive reduction of non ground 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 5m. These higher jump points are referred to as transition points.
- Cluster these transition points, and for each transition cluster having at least 10 points, begin a region growing procedure (similar to the one explained in section IIB).
- Probable ground points are then extracted out by removing all the grown regions from the set of remaining points $\mathbf{P}$.

It is worth to mention here that the grown regions can also be incorporated into the set of extracted roof points as depicted in [4]. 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 cuts based optimization library using *αβ*-swap move algorithm [7]-[10]. The minimum energy corresponds to the labeling *f* such that higher planar points are detected as building roof points. Incorporating them into the set of roof points extracted via façade information thus completes the extraction procedure.

III. EXPERIMENTAL RESULTS

To validate our approach, we tested the algorithm on TomoSAR point clouds generated from a stack of 102 TerraSAR-X high resolution spotlight images from ascending orbit using the Tomo-GENESIS software developed at the German Aerospace Center (DLR) [11]. The test area covers approximately 1.5 km^2 in the city of Berlin, Germany. The number of TomoSAR points in the area of interest is about 0.52 million.

Figure 2 shows the result of applying façade reconstruction procedure over the test area. Seed points are selected from each reconstructed façade and the region is grown using surface normals based similarity measure with θ_{ang} set to 15°. Later, among remaining points, roof points are extracted by adopting energy minimization procedure explained earlier (section IIC). Figure 3 shows the final extracted building points i.e., both roof and façade points.

The actual ground truth data are missing for exact quantitative evaluation of the approach. In order to provide some qualitative measures of the algorithm performance, we compared our building extraction results to reference polygons downloaded from the OpenStreetMap (OSM) [12]. Figure 3(b) shows the reference polygons overlaid onto the extracted building points. It can be visually seen that the extracted building points fits well to these reference polygons. Moreover, by analyzing the detected buildings from TomoSAR point clouds and validating using optical data, we completed few missing buildings in OSM dataset, polygons depicted as blue polygons in Figure 3(b). The performance of the (detection) extraction procedure is then assessed by employing the following evaluation metrics:

$$\left.\begin{aligned} \text{Completeness }(\%): comp &= 100 \times \left(\frac{TP}{TP+FN}\right) \\ \text{Correctness }(\%):\ corr &= 100 \times \left(\frac{TP}{TP+FP}\right) \\ \text{Quality}(\%): Q &= \frac{comp \times corr}{comp + corr - comp \times corr} = \frac{TP}{TP+FP+FN} \end{aligned}\right\} \tag{3}$$

The abovementioned metrics assess the overall performance of the building extraction algorithm. Completeness tells up to what percentage the algorithm has detected the roof points while 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 are provided in Table 1.

Table 1: Performance evaluation statistics.

Detected building points inside the reference building polygons i.e., True positives *TP*	295367
Detected non building points inside the reference building polygons i.e., False negatives *FN*	16269
Detected building points outside the reference building polygons i.e., False positives *FP*	50834
Detected non building points outside the reference building polygons i.e., True negatives *TN*	154420
Completeness (%) - *comp*	94.779
Correctness (%) - *corr*	85.316
Quality (%) - *Q*	81.487
Total approx. area in square meters of the buildings polygons	298869.78

IV. CONCLUDING REMARKS AND OUTLOOK

We have presented an approach that only utilized unstructured TomoSAR point clouds to detect building structures. The approach 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 proposed approach is automatic but parametric. The free parameters including r_N, η, *fac*, ε, and θ_{ang} are set empirically in this work. A further detailed sensitivity analysis of these parameters is therefore necessary. Moreover, we have compared our results to the OSM data which is openly available and regularly updated but not yet fully complete, therefore a more accurate ground truth would be needed for assessing exact qualitative and quantitative performance of the approach.

REFERENCES

[1] X. Zhu, M. Shahzad, "Façade Reconstruction Using Multi-View Spaceborne TomoSAR Point Clouds", *IEEE TGRS*, 52(6), pp. 3541-3552, 2014.

[2] X. Zhu, R. Bamler, "Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR", *IEEE TGRS*, 50(1), pp. 247-258, 2012.

[3] M. Shahzad, X. Zhu, "Robust Reconstruction of Building Façades for Large Areas Using Spaceborne TomoSAR Point Clouds", *IEEE TGRS*, 53(2), pp. 752-769, February 2015.

[4] M. Shahzad, X. Zhu, "Reconstructing 2-D/3D Building Shapes From Spaceborne Tomographic SAR Point Clouds", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-3, ISPRS Technical Commission III Symposium, Zurich, Switzerland, 2014.

[5] F. Lafarge, C. Mallet, “Creating large-scale city models from 3D-point clouds: A robust approach with hybrid representation”, *IJCV*, pp. 1–17, 2012.

[6] J. Yan, J. Shan, W. Jiang, "A global optimization approach to roof segmentation from airborne lidar point clouds", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 94, pp. 183–193, 2014.

[7] Y. Boykov, V. Kolmogorov, "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision", *IEEE TPAMI*, vol. 26, no.9, pp. 1124-1137, 2004.

[8] B. Fulkerson, A. Vedaldi, S. Soatto, "Class Segmentation and Object Localization with Superpixel Neighborhoods", Proceedings of the International Conference on Computer Vision, 2009.

[9] Y. Boykov, O. Veksler, R. Zabih, "Efficient Approximate Energy Minimization via Graph Cuts", *IEEE TPAMI*, vol. 20, no. 12, pp. 1222-1239, November 2001.

[10] V. Kolmogorov, R. Zabih, "What Energy Functions can be Minimized via Graph Cuts?", *IEEE TPAMI*, vol. 26, no. 2, pp- 147-159, 2004.

[11] X. Zhu, Y.Wang, S. Gernhardt, R. Bamler, “Tomo-GENESIS: DLR’s tomographic SAR processing system,” In Proceedings JURSE, Sao Paulo, Brazil, pp. 308–313, 2013.

[12] GEOFABRIK--Downloads, Accessed Sep. 9th, 2014. [online]. Available: http://download.geofabrik.de/europe/germany/berlin.html.

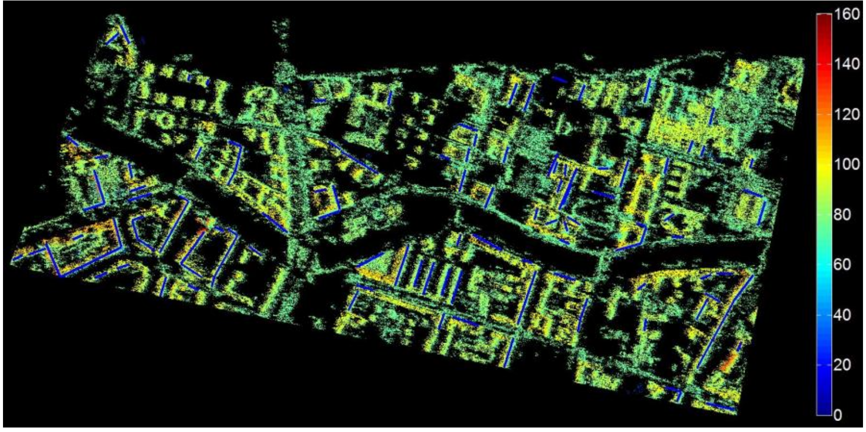

Figure 2. Top view of the three dimensional TomoSAR points in UTM coordinates of the area of interest in Berlin. Blue lines depict the reconstructed façade segments. The height of TomoSAR points is color-coded [unit: m].

(a)

(b)

Figure 3. Results of building extraction: (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 non building points respectively. The overlaid green polygons are reference buildings downloaded from OSM [12]. Blue polygons are manually extracted buildings not present in OSM data. Gray polygons are newly constructed buildings that are not present in our dataset where as magenta colored polygons are buildings that do not actually exist but present in OSM data. Both gray and magenta polygons are not included in the evaluation.