Towards Global 3D/4D Urban Modeling Using TanDEM-X Data

Xiao Xiang Zhu, German Aerospace Center, Oberpfaffenhofen, Germany, and Technical University of Munich, Munich, Germany (xiao.zhu@dlr.de)
Yao Sun†, German Aerospace Center, Oberpfaffenhofen, Germany
Yilei Shi†, Technical University of Munich, Munich, Germany
Yuanyuan Wang, Technical University of Munich, Munich, Germany
Nan Ge, German Aerospace Center, Oberpfaffenhofen, Germany
† Authors contributed equally

Abstract

TanDEM-X mission delivers a global digital elevation model of high quality (12 m posting) close to the HRTI-3 standard. However, when it comes to urban areas, layover effect caused by the side-looking nature of the radar satellites handicaps the use of classical multi-baseline phase unwrapping and TanDEM-X data alone for a precise 3D reconstruction. This calls for advanced SAR image processing methods or complementary data, such as volunteered geographic information. This paper proposes a series of novel algorithms aiming at delivering the world's first global 3D and 4D urban model using primarily the TanDEM-X data.

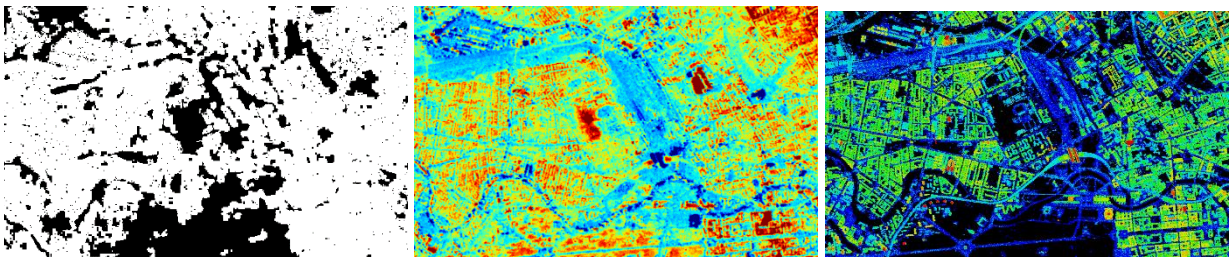


Figure 1. Current knowledge of urban 2D/3D morphology derived from SAR data. From left to right: the global urban footprint (GUF) binary map, 12m posting TanDEM-X raw DEM (image from [1]), and the tomographic 3D reconstruction using approximately 200 images, of the same area in Berlin, respectively. The height in the latter two plots are color-coded (blue-red: ca. 0m-70m). The building height in the TanDEM-X DEM is often underestimated.

1 Introduction

1.1 Our vision

By 2050, around three quarters of the world's population will live in cities. The ongoing new dimension of global migration into the cities poses fundamental challenges to our societies across the globe. Global 3D/4D urban model would play a fundamental role for the stakeholders in understanding such rapid urbanization. Yet our knowledge of global urban morphology so far only exist in a 2D binary classification form, which is known as the Global Urban Footprint (GUF) [2]. The available global DEMs, such as the TanDEM-X DEM, are optimized for nonurban areas. Global 3D *urban* model is non-existent. Not even to mention urban models with temporal updates that is so-called 4D urban model. This calls for high quality products beyond the current knowledge of urban morphology.

Figure 1 is an overview of the current knowledge of 2D and 3D/4D urban morphology derived from SAR data. From left to right, it shows the GUF binary urban-nonurban mask, the TanDEM-X raw DEM with 12m posting, and the SAR tomographic (TomoSAR) reconstruction using over 200 repeat-pass TerraSAR-X images. The first two results are available globally. However, the building height in the standard TanDEM-X DEM is often underestimated. On the other hand, the highly accurate TomoSAR reconstruction requires stack with large number of interferograms, which is only available for very limited number of cities. According to the TanDEM-X data availability map shown in Figure 2, most part of the world is only covered by two to three acquisitions. Only difficult terrains such as mountainous areas are covered by more acquisitions. Therefore, conventional multibaseline InSAR techniques, e.g. TomoSAR [3], [4] and persistent scatterer interferometry (PSI) [5], [6], are not directly applicable to these small TanDEM-X image stacks.

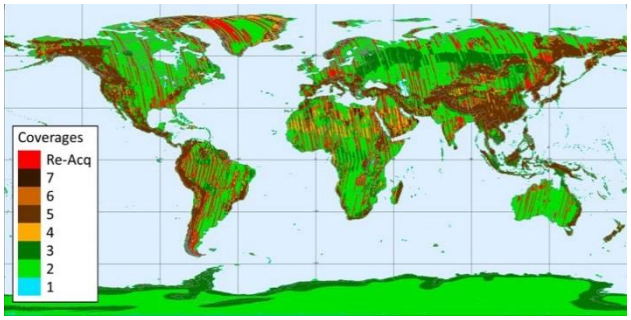


Figure 2. Global number of acquisitions of TanDEM-X data. Image quoted from [7].

Our vision is to reconstruct a global 3D urban model using primarily TanDEM-X data. To overcome the above-mentioned challenges, we propose a series of advanced

SAR and InSAR techniques, as well as complementary data, such as terrestrial optical images or data in geographic information system (GIS) that can bridge the gap of limited number of TanDEM-X images over an area and satisfactory 3D reconstruction accuracy.

1.2 The proposed algorithms

This paper proposed a series of novel algorithms aiming at global urban DEM generation. The proposed algorithm series consider three levels of data availability:

1) Single acquisition: only a pair of single look complex image or one interferogram can be exploited. We propose to use purely geometry-based method on the intensity image or the interferogram, after appropriate adaptive filtering on the images [8], [9].
2) Small TanDEM-X stack: a stack of a few interferograms are available. This requires advanced multi-baseline InSAR techniques, such as spectral estimation based on joint sparsity [10], and InSAR stack filtering based on robust low rank tensor decomposition [11].
3) Normal TanDEM-X stacks on which typical multi-baseline InSAR algorithms can be applied.

Table 1 summarizes the aforementioned algorithms and the required complementary data with respect to the three levels of data availability.

Table 1. Overview of the possible 3D reconstruction algorithms that can be applied w.r.t. an increasing number of TanDEM-X acquisitions.

No. of images	Complementary data required	Methodology
one	GIS building footprint	geometry-based methods
a few	-	Advanced filtering (e.g. nonlocal, and low rank) with spectral estimation
	GIS building footprint	joint sparsity-based spectral estimation
dozens	-	Typical spectral estimation, e.g. PSI, TomoSAR

2 Algorithms

2.1 Single acquisition – GIS-SAR

In very high resolution (VHR) SAR images, bright pixels on building façade often appears as parallel building contours, covering from the bottom of building (far-range) to the roof of building (near range). These signals are usually originated from dihedral or trihedral structures, such as window or balcony corners [12]. Therefore, the number of such parallel contours is often identical to the number of building floors. They define the layover extend of the façade, hence also the building height. This section introduces the GIS-SAR algorithm that makes use of a single VHR SAR image together with GIS building footprints to determine building heights. The workflow of the algorithm is as follows.

- Firstly, the GIS building footprints are projected to the SAR image. A visibility check is performed, so that only the segments that are visible in that SAR image is kept [10].
- These footprint segments are shifted from the bottom of the building to the top towards the near-range direction on the SAR image. The integral of the pixel intensity along the footprint segments at each shift position is recorded. They form as an intensity profiles as a function of the range shifts.
- The value of the periodic signal on the intensity profile is analyzed to determine: (a) if there exists a building, and (b) the layover length if building exists.
- Finally, the layover length is converted to building height according to basic SAR imaging geometry.

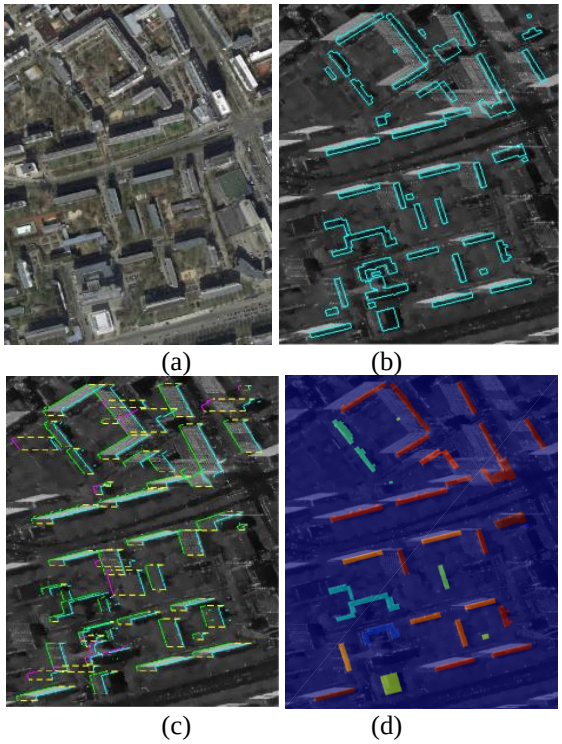


Figure 3. (a) the Google image of the test area, (b) the GIS building footprint in cyan projected to the SAR image, (c) the visible segments of the footprint in cyan and the determined extend of the layover in green, and (d) the final reconstructed building models with color coded height (blue-red: 4-33m). Among the 45 buildings in this area, 33 are detected and the heights are estimated with an overall accuracy of 2.8m and a bias of -2.4m. Twelve buildings were not detected as building by the current algorithm due to the lack of prominent reflection on the building façade or the layover with other buildings. Their building footprints are shown in magenta.

Figure 3 shows an example of the proposed algorithm. From (a) to (d), they are the optical image from Google (a), the building footprints projected to SAR image (b), the visible footprint segments and the determined extend of the layover area (c) and the obtained building models (d). Among the 45 buildings in the test area, 33 are detected and the heights are estimated with an accuracy of

2.8m, as well as a bias of negative 2.4m w.r.t. to the ground truth (a 7cm-resolution optical DEM obtained from aerial images using semi-global matching [13]). The negative bias is expected to systematically occur, because the bright reflection seen as the top of the building in a SAR image is often not originated from the roof of the building. Instead, it is from the bottom window corner of the highest floor. The height difference between such window corner and the roof is about 2.3m, i.e. typical floor height 3m minus typical height of windowsill above the floor 0.7m. Twelve buildings were not detected in the algorithm, because their intensity profiles appear to be homogenous and the intensity values are similar to the intensity of the ground or shadow. This is due to the lack of prominent corner reflectors on the illuminated façades, or the layover of multiple buildings.

2.2 Small stack bistatic InSAR

Conventional multibaseline InSAR techniques usually requires a fairly large stack (>20 images) for a reliable reconstruction. They are not directly applicable to very small (less than 6) InSAR stacks for the following reasons:

- the accuracy of height estimates are linearly proportional to the product of signal-to-noise ratio (SNR) and the number of measurements
- large ambiguities in the height estimates are easier to occur at low number of images as well as at low SNR
- typical spectral estimators for height reconstruction are only asymptotically optimal, i.e. the estimator is suboptimal at low number of images

To improve the estimation accuracy, the algorithm should improve the SNR of the pixels in the image, as the number of images is fixed. The proposed algorithm considers the following two aspects.

- Outlier filtering: outlier exists because the images are sensed from few hundreds of kilometre away. We adapt the state-of-the-art robust low rank tensor decomposition method [11] to detect outlier pixels from a group of neighboring pixels that share a similar or homogeneous structure in intensity or interferometric phase. Such metaphysical property is known as *low rank* in linear algebra.
- SNR improvement: We adapt the nonlocal filter which was recently demonstrated to be successful in SAR interferogram filtering [8], [9]. It groups and averages pixels of similar height to improve their SNR.

After the abovementioned pre-processing, a typical spectral estimation, such as periodogram or TomoSAR, is followed. Lastly, a robust weighted average is performed on the height estimates to mitigate possible outliers in the results.

Figure 4 shows the reconstructed elevation (w.r.t. a reference point) of Munich, Germany using just five TanDEM-X interferograms. The color indicates the estimated elevation. Compared to TomoSAR without the filtering steps, the proposed approach can improve the accuracy of elevation estimates by a factor of three [11].

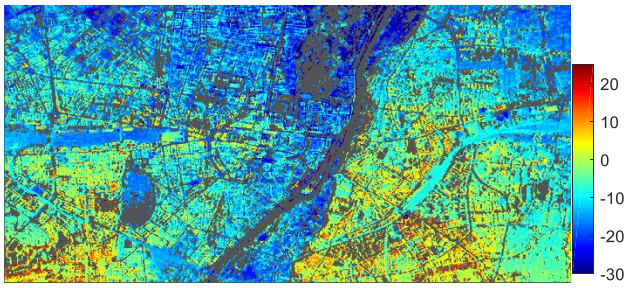
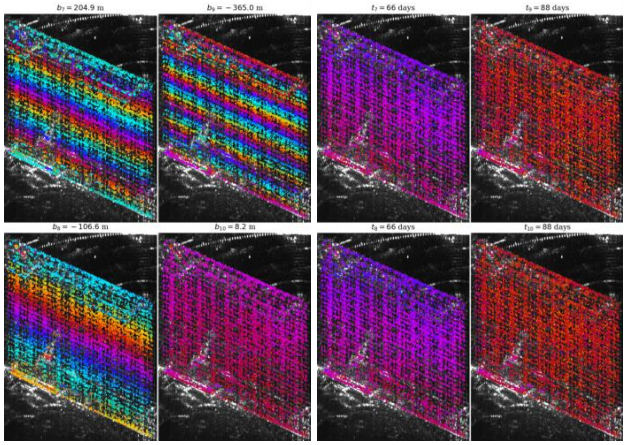


Figure 4. The estimated elevation w.r.t. a reference point reconstructed by the proposed method using five TanDEM-X interferograms over Munich. The elevation is color coded [unit: m].

2.3 Advanced multibaseline InSAR for small stacks- DefoSAR

Above all, advanced multibaseline InSAR techniques can be applied to selected high-rise buildings. According to our previous study, the M-SL1MMER algorithm [10] can retrieve highly accurate height estimates using merely six interferograms (cf. the typically required 20–100) by introducing GIS building footprints as prior knowledge, and jointly reconstructing pixels along iso-height lines. Please refer to [10] for a detailed treatise on M-SL1MMER.

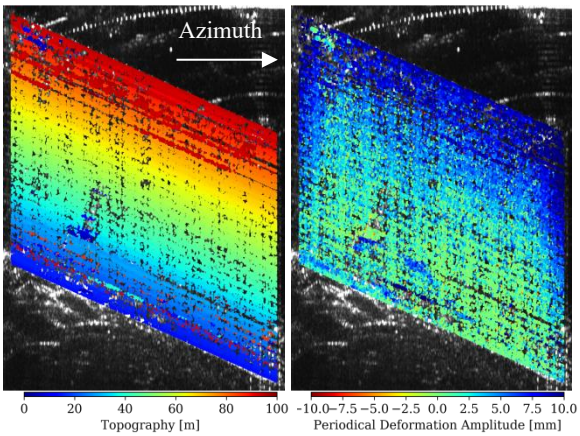
Based on such high-quality 3D reconstruction results, see for example Figure 6(a), accurate deformation signal can also be extracted from repeat-pass interferograms (known as the *DefoSAR*). The TanDEM-X pursuit monostatic acquisitions provide a perfect opportunity for both 3D reconstruction and deformation monitoring in high quality. As a first demonstration, 10 repeat-pass interferograms were generated from 6 pursuit monostatic staring spotlight pairs acquired during the TanDEM-X Science Phase. Figure 5 shows four of these interferograms before and after topography compensation using the height estimates obtained by M-SL1MMER (shown in Figure 6(a)). Each subfigure is annotated with the corresponding spatial (perpendicular) or temporal baseline.



(a) (b)

Figure 5. Time series of four repeat-pass interferograms before (a) and after (b) topography compensation using the high quality 3D height reconstruction from Figure 6(a).

Under the assumption that the motion in the period of study is mainly caused by periodical temperature change, a sinusoidal function was fitted to the phase history after topography compensation. The estimated amplitude of periodical deformation is shown in Figure 6(b). The strong correlation of the amplitude estimate with the height complies with the expected deformation behavior. Yet, due to its requirement for small stacks of TanDEM-X datasets and external GIS building footprints, this algorithm is only applicable to selected buildings.



(a) (b)

Figure 6. Estimated elevation (a) and amplitude of periodical deformation of single point-like scatterers (b).

2.4 Big stack InSAR – Tomo-GENESIS

Typical TomoSAR inversion algorithms can be applied to stacks with sufficient number of images. We will employ the tomographic inversion processor *Tomo-GENESIS* [14] that is based on our previous work [4], [15]–[19]. As these works have been extensively discussed in previous publications, they are not repeated in this paper.

3 Discussion

The three algorithms provides solution to different dataset, hence their performance is not directly comparable. However, as there result will be fused into a single product, the following contents discuss the advantage and disadvantages of these algorithm, and compares their expected accuracy of height estimates.

3.1 Comparison among algorithms

GIS-SAR

Being an incoherent method, exploiting only the intensity provides a complete different view compared to the other two InSAR methods. As mentioned in section 2.1, an accuracy of less than 3m can be achieved in general. Our previous study also shows that the accuracy can be as high as 1m for certain buildings with simply geometry and no layover with other objects [20]. It is also shown in [20] that roofs of different heights in a single building complex can also be estimated with such high accuracy. However, difficulty may arise when considering larger area processing:

- GIS building footprints are usually only available for cities in developed countries. In addition, there accuracy is also limited.
- Buildings without sufficient amount of corners on the facades facing the sensor, or buildings that are occluded by surrounding objects are difficult to be exploited in the algorithm.
- The algorithm relies on a fairly accurate coregistration of the SAR image and the GIS data. This usually requires an accurate terrain model that might not be available for many areas. Otherwise, future algorithm should consider the unknown terrain height.
- Unmixing of layovered building facades has not been addressed in this algorithm.
- Systematic underestimation exists because of the lack of strong reflection from the building roof.

Small stack InSAR

This algorithm performs full tomographic inversion without additional data. Its performance can be characterized mainly by the number of images and the SNR of the pixels. Therefore, it will be applied in a global scale. However, the limited number of acquisitions per area restricts its accuracy. According to the current analysis, we expect to reach a relative accuracy of better than 2m.

M-SL1MMER+DefoSAR

M-SL1MMER provides height estimates with unprecedented accuracy compared to the state-of-the-art multi-baseline InSAR algorithms. The follow-up DefoSAR algorithm retrieves precise deformation signal. However, due to its requirement for TanDEM-X and GIS datasets, it is only used to reconstruct individual buildings in selected cities. Yet, for those selected buildings, M-SL1MMER reaches an accuracy of less than half meter in the height estimation from just 10 images [10], [21].

3.2 Expected relative accuracy

Table 2 summarizes the coverage and the expected relative height accuracy of the proposed three algorithms. The expected final urban model will be a fusion of the results from all the algorithms. The base map will be the result from small stack InSAR, as it will cover the global urban area. We aim at a relative vertical accuracy of better than 2m. To give an impression of the global urban model, an example is given in Figure 7 in which a 3D view of Munich city is shown. For most buildings in developed cities, GIS footprint with SAR can provide urban models with an height accuracy from 1m to 3m. The traditional tomographic inversion will also be applied on some selected cities where TanDEM-X is abundant. We expect a relative height accuracy of around 1m from our previous publications. For certain buildings in the selected cities, a relative accuracy better than half meter can be achieved by applying the M-SL1MMER algorithm. Last but not least, as the accuracy of individual algorithms varies, we also aim to provide a quality map.

4 Conclusion and outlook

This paper proposes a series of algorithm for the reconstruction of a global 3D urban model from TanDEM-X data, and as well as freely available GIS building footprints. The proposed algorithm series provides a complete solution to TanDEM-X data of different availability, from single acquisition, to tens of images. The expected final product will be a fused global 3D urban model from the results of all the algorithms.

Still, challenges remain at several points. Firstly, a reliable tomographic method needs to be developed for small TanDEM-X InSAR stacks, as the current small stack InSAR algorithm is only based on the assumption of single scatterer. Secondly, robust method is required to tackle the less controlled quality of the openly available building footprint, and lastly, method of obtaining urban model changes on a regular basis is non-existent. These will be studies in the future work.

Table 2. The proposed algorithms, their coverage, and the expected relative accuracy of urban models.

Algorithm	Coverage	Rel. Accuracy
GIS-SAR	Most of the buildings in developed cities	1m~3m
Small stack InSAR	Global urban area	<2m
M-SL1MMER	Selected buildings in selected cities	<0.5m
Large stack TomoSAR	selected cities	<1m

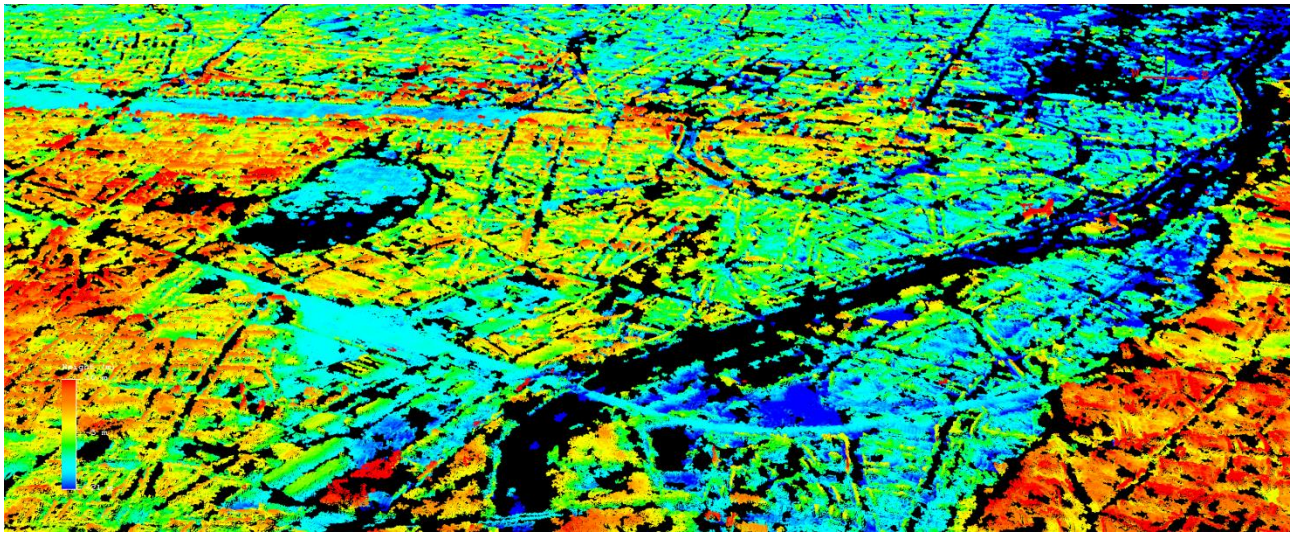


Figure 7. The expected global outcome by small stack InSAR – Munich in 3D, geocoded results of Figure 5. The color represents the height.

5 Acknowledgement

We gratefully acknowledge the support of the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No [ERC-2016-StG-714087], Acronym: *So2Sat*). We also gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (www.lrz.de).

6 Literature

[1] C. Rossi, T. Fritz, and M. Eineder, "TanDEM-X DSM uncertainty measures and demonstrations," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 3830–3833.

[2] T. Esch *et al.*, "Urban Footprint Processor - Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1617–1621, Nov. 2013.

[3] 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.

[4] 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, 2010.

[5] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.

[6] 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.

[7] P. Rizzoli *et al.*, "Generation and performance assessment of the global TanDEM-X digital elevation model," *ISPRS J. Photogramm. Remote Sens.*, vol. 132, pp. 119–139, Oct. 2017.

[8] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jager, "NL-SAR: A Unified Nonlocal Framework for Resolution-Preserving (Pol)(In)SAR Denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2021–2038, Apr. 2015.

[9] X. X. Zhu, M. Lachaise, F. Adam, and R. Bamler, "Towards 6m TanDEM-X DEM: Non-local Methods for In-

SAR Filtering," in *4th TanDEM-X Scientific Meeting Final Program and Abstracts*, Oberpfaffenhofen, Germany, 2013.

[10] X. X. Zhu, N. Ge, and M. Shahzad, "Joint Sparsity in SAR Tomography for Urban Mapping," *IEEE J. Sel. Top. Signal Process.*, vol. 9, no. 8, pp. 1498–1509, Dec. 2015.

[11] J. Kang, Y. Wang, M. Körner, and X. Zhu, "Robust Object-based Multi-pass InSAR Deformation Reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4239–4251, 2017.

[12] S. Auer and S. Gernhardt, "Linear Signatures in Urban SAR Images - Partly Misinterpreted?," *Geosci. Remote Sens. Lett. IEEE*, vol. 11, no. 10, pp. 1762–1766, Oct. 2014.

[13] H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 328–341, Feb. 2008.

[14] X. Zhu, Y. Wang, S. Gernhardt, and R. Bamler, "Tomo-GENESIS: DLR's Tomographic SAR Processing System," in *Urban Remote Sensing Event (JURSE), 2013 Joint*, 2013, pp. 159–162.

[15] X. Zhu and R. Bamler, "Tomographic SAR Inversion by L1-Norm Regularization -- The Compressive Sensing Approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, 2010.

[16] X. X. Zhu and R. Bamler, "Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, 2012.

[17] X. X. Zhu and R. Bamler, "Superresolving SAR Tomography for Multidimensional Imaging of Urban Areas: Compressive sensing-based TomoSAR inversion," *Signal Process. Mag. IEEE*, vol. 31, no. 4, pp. 51–58, Jul. 2014.

[18] 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, pp. 1250–1254, 2014.

[19] Y. Wang and X. Zhu, "Automatic Feature-based Geometric Fusion of Multi-view TomoSAR Point Clouds in Urban Area," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 3, pp. 953–965, 2015.

[20] Y. Sun, M. Shahzad, and X. X. Zhu, "Building height estimation in single SAR image using OSM building footprints," in *2017 Joint Urban Remote Sensing Event (JURSE)*, 2017, pp. 1–4.

[21] N. Ge and X. X. Zhu, "Bistatic-like Differential SAR Tomography," *IEEE Geosci. Remote Sens. Lett.*, 2018.