Abstract—Recently, a new global digital elevation model (DEM) with pixel spacing of 0.4 arcsec and relative height accuracy finer than 2 m for flat areas (slopes < 20%) and better than 4 m for rugged terrain (slopes > 20%) was created through the TanDEM-X mission. One important step of the chain of global DEM generation is to mosaic and fuse multiple raw DEM tiles to reach the target height accuracy. Currently, weighted averaging (WA) is applied as a fast and simple method for TanDEM-X raw DEM fusion, in which the weights are computed from height error maps delivered from the Integrated TanDEM-X Processor (ITP). However, evaluations show that WA is not the perfect DEM fusion method for urban areas, especially in confrontation with edges such as building outlines. The main focus of this paper is to investigate more advanced variational approaches such as TV-L1, L1 norm total variation, and derived from the coherence values and the given geometric configuration (HEM) [14], which are a byproduct of the InSAR process. Form raw SAR data take to the final global DEM, a workflow including different phases such as interferogram generation, phase unwrapping (PU), data calibration, DEM block adjustment, and mosaicking is implemented at DLR [3]. A main step of the DEM generation procedure is carried out in the Integrated TanDEM-X Processor (ITP), which leads to primary raw DEMs for each bistatic acquisition [4]. During the raw DEM generation, some potential error sources are removed by instrument and baseline calibration [5]. After that, the vertical bias, which usually lies between 1 and 5 m, is corrected by a least squares block adjustment [6]. The block adjustment is performed by using ICESat data and connecting points in the overlapping areas of raw DEM tiles. However, dependent on the terrain morphology, some error sources still remain after the block adjustment. The effect of these errors can be decreased through fusion of several DEM coverages within the DEM Mosaicking Processor (DMP) [7]. The TanDEM-X raw DEM coverage over different terrain types is displayed in Fig. 1. As can be seen, the most of the world is covered by at least two nominal acquisitions with height of ambiguities (HoA) between 30 and 55 m. The main objective of TanDEM-X DEM fusion is to improve the final accuracy by employing several coverages over different areas [8].

Diverse methods have been designed for the fusion DEMs with different properties, which can be seen as an application of data fusion in remote sensing [9]. Among them, weighted averaging (WA) is well established as a simple approach with low computational cost [8], [10]–[12]. However, its performance strongly depends on the weights that describe the height error distribution for each pixel [13]. For SAR interferometric-derived DEMs, the weights can be achieved from height error maps (HEM) [14], which are a byproduct of the InSAR process and derived from the coherence values and the given geometrical configuration [15]. However, it should be noted that HEMs
cannot represent all error sources as they do not reflect deterministic effects such as layover and shadow effects. Another way is to compute weight maps by a comparison with ground truth data that are not necessarily available for every arbitrary study area [16].

A current approach for implementing the DEM fusion in DMP is WA. In addition to WA, some logic for clustering consistent heights and upgrading weights regarding the influences of other significant factors such as HoA, PU methodology, and pixel locations relative to the border of the DEM scene is considered to finally reach the target relative accuracy and minimize PU errors remaining from primary steps [8]. While the WA approach can realize the predefined goals in the DMP for global DEM generation, it does not perform optimally in difficult terrains with complex morphology such as urban areas, which contains many high-frequency contents such as edges. After WA-based TanDEM-X DEM fusion, visualization shows that outlines of buildings are not perfectly sharp and still some amount of existing noise spoils the footprints of buildings [for example, see Fig. 9(c)]. As Fig. 1 illustrates, most areas are only covered by two nominal acquisitions (shown in green). Since this holds for many important urban areas as well, this also motivates the development of more sophisticated approaches.

An advanced approach for DEM fusion was proposed by Papasaika et al. [17]. They exploited sparse representations for multisensor DEM fusion. Zach et al. implemented an $L_1$ norm total variational model for range image fusion [18]. In another study, Pock et al. proposed to use total generalized variation (TGV) for fusing DEMs derived from airborne optical imagery [19]. Kuschk et al. evaluated weighted TGV to fuse DEMs derived from spaceborne optical imagery with different resolutions [20]. In another study, weighted TV-$L_1$, in which weights were predicted by neural networks, were applied for the Cartosat-1 and TandEM-X DEM fusion over urban areas [21]. Overall, in spite of the high computational cost of advanced methods for DEM fusion, they perform more efficiently than simple WA.

In this paper, we will investigate the application of more sophisticated DEM fusion approaches that are able to efficiently preserve edges and outlines of buildings while still reducing noise effects. For this purpose, two variational models, namely $L_1$ norm total variation (TV-$L_1$) and Huber model are implemented. Apart from these regularization approaches, we will also investigate the potential of employing raw DEMs with different properties such as different baseline configuration for the TanDEM-X DEM fusion. Therefore, this paper is structured into several sections. In Section II, the methodology of DEM fusion based on regularization methods is explained. Then, the description of the study subsets and experimental results from DEM fusion are provided in Section III. Finally, the performance of the implemented DEM fusion methods for TanDEM-X data over urban areas will be discussed in Section IV.

II. METHODS FOR TANDEM-X RAW DEM FUSION

In this paper, two approaches are implemented for TanDEM-X raw DEM fusion. The characteristics of each model will be explained in the following. Before the fusion, raw DEMs at first are aligned to each other by DEM coregistration approaches such as least squares matching [22], iterative closest point [23], or manual registration. The coregistration of DEMs decreases their translational and rotational differences. For stability reasons, in addition, the height data should be normalized to the interval $[0, 1]$

\[
h_{nk}(x, y) = \frac{h_k(x, y) - h_{\text{min}}}{h_{\text{max}} - h_{\text{min}}} \quad (1)
\]

where $h_k(x, y) > 0$ is the elevation of the study DEM with index $k$ at location $(x, y)$, $h_{\text{max}} > 0$ and $h_{\text{min}} > 0$ ($h_{\text{min}} < h_{\text{max}}$) are the lowest and highest elevations among all input DEMs. The output gives the normalized height in the considered location.

A. Background: Weighted Averaging

The most popular, very fast, and low computational cost method for DEM fusion is WA, which is implemented by

\[
f = \sum_{i=1}^{k} w_i \odot h_i \quad (2)
\]

where $h_i$ are 2-D arrays representing the input DEMs, $w_i$ are the corresponding weight maps, and $\odot$ is a pixelwise product. It is worth to note that other simple methods, such as a pixelwise
median- or mode-based fusion, can also be employed for DEM fusion, especially when multiple DEMs are available [24].

As explained in Section I, the main critical issue for using WA for DEM fusion is to apply appropriate weights that are fairly representative of expected height errors in the source DEMs. For TanDEM-X DEM fusion, generally, these weights are delivered as HEMs from the ITP. For each height of the TanDEM-X DEM, the corresponding HEM value can be estimated by

$$\sigma_j = H_{\text{amb}} \frac{\sigma_{\phi,j}}{2\pi}$$

(3)

where $H_{\text{amb}}$ is the height of ambiguity and $\sigma_{\phi,j}$ is the interferometric phase error that is estimated from the interferometric coherence and the InSAR geometry [2]. Then, from these values, the respective weights can be calculated for each pixel location by

$$w_j = \frac{1}{\sum_{j=1}^{N} \sigma_j^2}.$$  

(4)

**B. Regularization-Based Models**

Variational models were first used for signal and image denoising [25], [26]. Generally, in variational denoising approaches, an energy functional is constituted by fidelity and regularization terms. The fidelity is considered to enforce the output image being similar to the input images while the regularization term (also called penalty term) is embedded to reduce the effect of noise in the final result. The desired output is achieved by minimizing the constructed energy functional. Different energy functionals can be formed according to different functions for defining data and penalty terms [19].

A popular type of variational models is the total-variation-based model (TV) in which the gradient of a desired output image is selected to form the regularization term based on different norms. The main advantage of the TV-based variational model is its convexity that guarantees to find a solution by minimizing the energy functional.

In the problem of TanDEM-X DEM fusion, several input raw DEMs are fused using variational models. The data term makes the fused DEM similar to the input tiles while the TV-based regularization term is defined to provide a sharp output at the end by preserving the edges and reducing the noise. This property is beneficial for fusing TanDEM-X raw DEMs over urban areas where footprints of buildings as edges often appear very noisy because of the inherent SAR imaging properties.

The basic gradient-based variational model for image denoising and data fusion is a quadratic model in which $L_2$ norm is used for both regularization and data terms [27]. However, the quadratic regularization term causes oversmoothing for edges. Therefore, using the $L_1$ norm instead was proposed by Rudin et al. which is called ROF model correspondingly [25]. Since the ROF model still uses the $L_2$ norm for the data term, it does not provide robustness against outliers when applied to DEM fusion. As a solution, the $L_1$ norm can be substituted for the $L_2$ norm [28]. The TV-$L_1$ model consists of the data fidelity and the penalty term

$$\min_{f} \left\{ \sum_{i=1}^{k} \| f - h_i \|_1 + \gamma \| \nabla f \|_1 \right\}$$

(5)

where $h_i$ are noisy input DEMs and $f$ is the desired DEM, which should be achieved by minimizing the functional energy mentioned above. The penalty term is formed based on the gradients of the newly estimated DEM to preserve the edges at the end. The regularization parameter $\gamma$ trades off between penalty and fidelity terms. Increasing $\gamma$ will influence the smoothness and will produce a smoother fused DEM in the end.

While the main advantage of TV-$L_1$ is its robustness against strong outliers as well as edge preservation [19], it suffers from the staircasing effect, a phenomenon that creates artificial discontinuities in the final output and particularly affects high resolution DEM fusion [29]. Moreover, the $L_1$ norm is not necessarily the best choice for all data fusion and denoising cases. As an alternative, the Huber regularization model is proposed to rectify the drawbacks of the TV-$L_1$ model [19]. It applies to the Huber norm instead of the $L_1$ norm in both fidelity and penalty terms [30]

$$\| x \|_H = \begin{cases} \frac{|x|^2}{2\eta} & \text{if } |x| \leq \eta \\ |x| - \frac{\eta}{2} & \text{if } |x| > \eta. \end{cases}$$

(6)

Here, $\eta$ is a parameter that determines a threshold between the $L_1$ and $L_2$ norm in the model. Based on this, the Huber model can be defined as [31]

$$\min_{f} \left\{ \sum_{i=1}^{k} \sum_{\Omega} \| f - h_i \|_\alpha + \gamma \sum_{\Omega} \| \nabla f \|_\beta \right\}$$

(7)

where both data and penalty terms are constituted based on the thresholds $\alpha$ and $\beta$ that are substituted as $\eta$ in the Huber norm relation (6) to form these terms and $\Omega$ denotes the raster DEM space. It should be noted that the Huber norm is a generalized form of the $L_1$ norm. However, in this study, the Huber norm is also used to strictly penalize the outliers.

Using the quadratic norm in the regularization term penalizes high-frequency changes more than $L_1$ norm, and thus, it reduces the noise at the cost of oversmoothing edges. The Huber norm, dependent on $\eta$ values, treats as a norm between $L_1$ and $L_2$ norms. However, for $\eta = 1$, its behavior is nearly similar to $L_1$. In other words, the Huber norm with higher $\eta$ provides DEMs with smoother building footprints but not as much as the quadratic norm. The influence of the parameters of variational models on the quality of fused DEM will be discussed in Section IV-A with more details. In the remainder of this paper, we use $\alpha = 4$ to smooth relative height errors larger than 4 m (considering the relative accuracy of the TanDEM-X DEM), and $\beta = 1$ based on data-driven experiments on different datasets. Consequently, $\gamma$ can be calculated by the L-curve method [32].

**C. Implementation**

It is mathematically proven that the TV-based energy functional based on $L_1$ or the Huber norm is convex. The main
Algorithm 1: Dual primal algorithm.

Input: Primary DEMs to configure primal problem
1: Initialization: \( \tau \sigma \|K\| \leq 1 \), \((\mathbf{u}^0, \mathbf{v}^0) \in \mathbf{U} \times \mathbf{V} \), \( \mathbf{u}^0 = \mathbf{u}^0, \theta \in [0, 1] \).
2: for \( i = 0 \) to stopping criteria do
   \( \mathbf{v}^{i+1} = (I + \sigma \partial F^*)^{-1} (\mathbf{v}^i + \sigma K \mathbf{u}^i) \)
   \( \mathbf{u}^{i+1} = (I + \tau \partial G)\mathbf{u}^i - \tau K^T \mathbf{v}^{i+1} \)
   \( \mathbf{u}^{i+1} = \mathbf{u}^{i+1} + \theta (\mathbf{u}^{i+1} - \mathbf{u}^i) \)
3: end for
Output: \( \mathbf{u} \)

characteristic of a convex problem is that the desired output (i.e., the global minimum) will be certainly found through an optimization process. One popular strategy for finding the minimum of a convex optimization is to reformulate the functional energy as a primal-dual problem [33]. For variational models, the energy functional can be expressed in a general form such as

\[
\min \{ \mathcal{G}(\mathbf{u}) + \mathcal{F}(K\mathbf{u}) \} \tag{8}
\]

where \( \mathcal{G}(\mathbf{u}) \) is the data term, and \( \mathcal{F}(K\mathbf{u}) \) is the regularization term. \( K \) refers to an operator that is used for defining the regularization term (for TV-based variational models, this is the gradient \( \nabla \) of the desired output). In the TanDEM-X raw DEM fusion, \( \mathbf{u} \) (as primal variable) is the final fused DEM (f) and the energy functional (terms \( \mathcal{G}(\mathbf{u}), \mathcal{F}(K\mathbf{u}) \)) is defined by relations (5) and (7). Then, the dual-problem formulation of the energy functional can be written as

\[
\min \max \{ \mathcal{G}(\mathbf{u}) + \langle \mathbf{v}, K\mathbf{u} \rangle - \mathcal{F}^*(\mathbf{v}) \} \tag{9}
\]

where \( \mathcal{F}^*(\mathbf{v}^*) = \sup_{\mathbf{v} \in \mathbf{V}} (\mathbf{v}^*, \mathbf{v}) - \mathcal{F}(\mathbf{v}) \) \( \tag{10} \)

and \( \mathbf{v} \) is the dual variable and \( \mathcal{F}^* \) is defined as the convex conjugate of \( \mathcal{F} \). The dual-problem algorithm for minimizing (9) is presented in Algorithm 1. It should be noted that the median-based fusion of the input DEMs can be used to initialize \( \mathbf{u}^0 \) to speed up the optimization. More details of the algorithm can be found in [33].

III. EXPERIMENTS

In this paper, we investigate TanDEM-X raw DEM fusion over urban areas by using diverse TV-based variational models such as TV-\(L_1\) and Huber models. In addition, the effect of fusing raw DEMs with different baseline configurations will be investigated. The baseline configurations for TanDEM-X data differ by changing orbit direction (ascending or descending) and also changing HoAs. Furthermore, the results of DEM fusion implementation will be evaluated for different land types with an emphasize on urban areas. These land types are as follows.

1) Industrial areas that are characterized as areas with large buildings, often not very high.
2) Inner city areas that include very densely packed buildings, relatively high.
3) Residential areas that are typically specified with low-rise single family homes.

In addition, we also considered some nonurban study areas such as agricultural and forested areas to evaluate the performance of variational models in those areas too.

In TanDEM-X raw DEMs produced with a pixel spacing of 0.2 arcsec (around 6 m), most building footprints in industrial and inner city areas can be visualized but the height accuracy and quality of building shapes suffer from noise and systematic errors such as layover and shadow. The visualization and quality of building become worse in residential areas because of the small sizes and heights of buildings in these areas. However, we will evaluate the performance of variational models for enhancing the quality of buildings appeared in the final fused DEM over different aforementioned land types. After resampling and coregistration, the raw DEMs are fused by the different approaches explained in Section II.

### A. Fusion of TanDEM-X Raw DEMs With Similar Baseline Configuration

Most of the global coverage achieved with the TanDEM-X raw DEMs is generated by two nominal bistatistic acquisition (see Fig. 1), but there are more tiles in overlapping areas at the border of the tiles. The first investigation includes data takes that have similar baseline configurations as well as HoAs. The study subsets are selected from two nominal TanDEM-X raw DEMs over Munich city in Germany. The characteristics of these raw DEMs are presented in Table I.

- **TanDEM-X raws DEMs:** Munich area

| Acquisition Id | 1023491 | 1145180 |
|----------------|---------|---------|
| Acquisition mode | Stripmap | Stripmap |
| Center incidence angle | 38.25° | 37.03° |
| Equator crossing direction | Ascending | Ascending |
| Look direction | Right | Right |
| Polarization | HH | HH |
| Height of ambiguity | 45.81 m | 53.21 m |
| Pixel spacing | 0.2 arcsec | 0.2 arcsec |
| HEM mean | 1.33 m | 1.58 m |
± 20 cm. The final reference DEM is achieved by interpolation in a grid with pixel spacing as same as input TanDEM-X DEMs.

The results of raw DEM fusion using TV-$L_1$ and Huber models for study areas are presented in Table II. The regularization parameter is calculated using the L-curve method. For comparison with the common fusion method, the results of fusion by WA are also provided. The DEM quality after and before fusion was evaluated by statistical metrics, mean, root mean square error (RMSE), mean absolute error (MAE), normal median absolute deviation (NMAD), and standard deviation (STD).

In addition to the statistical analysis, to evaluate the performance of variational models, the residual maps of the input DEMs and the fused DEMs achieved by different methods for the industrial and inner city 1 study areas are displayed in Fig. 4. The results illustrate that using variational models in the fusion process can finally improve the quality of the TanDEM-X DEM over the quality achievable with classic WA. It is explicitly displayed on residual maps that variational modes can finally reduce the noise effects and also makes the footprints of buildings more apparent than WA. Furthermore, fusing ascending and descending DEMs can improve the DEM quality in particular for the shadow- and layover-affected areas in which significant errors occur.

### B. Fusion of TanDEM-X Raw DEMs With Different HoAs

In the first experiment, the study areas were selected from two TanDEM-X raw DEMs that have nearly similar properties. Both DEMs were acquired in the same orbit, same look directions, and also with nearly the same incidence angles and HoAs.

As an additional experiment, we investigate the performance of variational models for fusing TanDEM-X raw DEMs with different HoAs over urban areas. For this purpose, one experimental ITP raw DEM with different HoA over Munich city in Germany is considered. It should be noted that this product has not been used in final global DEM generation but in this study is applied for implementing an experiment of fusing rawDEMs with different HoAs. The specifications of this raw DEM are provided in Table III. Fig. 5 also provides a depiction of the new raw DEM, which is acquired over the same location as tile 1023491 with identical overlap.

The main property that discriminates this tile from those introduced in the previous section is its bigger HoA. Regarding nearly similar incidence angle and slant range, the larger value for HoA means this tile is derived from data takes that were acquired with a shorter baseline is considered helpful in areas where PU errors are dominant [34]. On the other hand, the quality and resolution of this DEM is lower than those with smaller
## TABLE II

| Study area | DEM | Mean | RMSE | MAE | NMAD | STD |
|------------|-----|------|------|-----|------|-----|
| Industrial | Raw DEM id: 1023491 | 0.71 | 4.40 | 3.08 | 2.37 | 4.34 |
|            | id: 1145180 | 0.71 | 4.64 | 3.27 | 3.01 | 4.58 |
| Fused DEM  | TV-L1 | 0.69 | 3.67 | 2.69 | 2.03 | 3.60 |
|            | Huber | 0.71 | 3.74 | 2.84 | 2.40 | 3.67 |
| Inner 1    | Raw DEM id: 1023491 | 0.78 | 7.79 | 5.95 | 6.49 | 7.75 |
|            | id: 1145180 | 0.78 | 8.08 | 6.30 | 7.15 | 8.04 |
|            | WA | 0.84 | 7.51 | 5.83 | 6.49 | 7.46 |
| Fused DEM  | TV-L1 | 0.77 | 6.11 | 5.00 | 5.72 | 6.06 |
|            | Huber | 0.78 | 6.14 | 5.09 | 5.67 | 6.09 |
| Inner 2    | Raw DEM id: 1023491 | 0.18 | 7.90 | 5.44 | 6.36 | 7.00 |
|            | id: 1145180 | 0.18 | 7.16 | 5.57 | 6.51 | 7.16 |
|            | WA | 0.20 | 6.82 | 5.33 | 6.23 | 6.82 |
| Fused DEM  | TV-L1 | 0.12 | 5.83 | 4.78 | 6.16 | 5.83 |
|            | Huber | 0.18 | 5.82 | 4.82 | 6.23 | 5.82 |
| Residential | Raw DEM id: 1023491 | 0.95 | 2.68 | 2.10 | 2.05 | 2.50 |
|            | id: 1145180 | 0.95 | 2.92 | 2.25 | 2.31 | 2.76 |
|            | WA | 0.96 | 2.61 | 2.05 | 1.99 | 2.43 |
| Fused DEM  | TV-L1 | 0.89 | 2.41 | 1.96 | 1.98 | 2.24 |
|            | Huber | 0.95 | 2.44 | 1.98 | 1.98 | 2.24 |
| Agricultural | Raw DEM id: 1023491 | 0.13 | 0.86 | 0.57 | 0.59 | 0.84 |
|            | id: 1145180 | 0.13 | 1.64 | 1.13 | 1.20 | 1.64 |
|            | WA | 0.14 | 0.78 | 0.51 | 0.54 | 0.76 |
| Fused DEM  | TV-L1 | 0.06 | 0.55 | 0.29 | 0.20 | 0.54 |
|            | Huber | 0.13 | 0.72 | 0.48 | 0.47 | 0.71 |
| Forested   | Raw DEM id: 1023491 | 2.25 | 4.84 | 3.54 | 3.46 | 4.28 |
|            | id: 1145180 | 2.25 | 4.58 | 3.36 | 3.24 | 3.99 |
|            | WA | 2.28 | 4.51 | 3.30 | 3.17 | 3.89 |
| Fused DEM  | TV-L1 | 2.25 | 4.34 | 3.18 | 3.09 | 3.71 |
|            | Huber | 2.25 | 4.36 | 3.21 | 3.12 | 3.73 |

The bold values indicate the best results.

---

Fig. 4. Absolute residual maps of the initial input raw DEMs and the fused DEMs obtained by different approaches for the industrial areas. (a) Inner city areas. (b) Study areas over Munich.
TABLE III
PROPERTIES OF THE NONOFFICIAL TANDEM-X RAW DEM TILE FOR MUNICH AREA

| TanDEM-X raws DEMs: Munich area | Acquisition Id 1058842 |
|----------------------------------|-----------------------|
| Acquisition mode                 | Stripmap              |
| Center incidence angle           | 38.33 °               |
| Equator crossing direction       | Ascending             |
| Look direction                   | Right                 |
| Polarization                     | HH                    |
| Height of ambiguity              | 72.02                 |
| Pixel spacing                    | 0.2 arcsec            |
| HEM mean                         | 2.58                  |

HoAs. Comparing the raw tiles displayed in Figs. 2 and 5 confirms a drop of quality of DEM, i.e., with much more noise, with id 1058842, which has larger HoA.

In this experiment, a study subset is extracted from an area that has lots of inconsistent heights due to PU errors. For this aim, a relatively large subset from an urban area, which is covered by trees and also includes a river crossing, is selected. Fig. 6 displays the selected study area suffering from PU errors. The corresponding DEM data are derived from tiles 1023491 and 1058842 with HoAs about 45 m and 72 m, respectively.

The PU errors appearing in this subset originate from the volume decorrelation phenomenon that happens in an area covered by trees (like the selected study subset) and also a coherence change due to transition from dry land to water (river). PU errors typically are at the range of multiples of the HoA value. The inconsistent heights can be determined by

\[ dh_{th} = 0.75 \times \min(|HoAs|) - 4. \]  \hspace{1cm} \text{(11)}

Those height residuals bigger than \( dh_{th} \) are denoted as inconsistent height values emerging because of PU errors.

Table IV collects the results of fusing DEMs with different HoAs in the selected study area. Again, the accuracy was evaluated respective to a reference DSM interpolated from a point cloud with high density (more than eight points per square meters). Moreover, Table V compares the fused DEMs with different approaches and initial DEMs in terms of number of PU errors, maximum and minimum height residuals. The PU threshold for each DEM is computed based on the respective HoA value using (11). It is obvious that the DEM 1058842 has lower number of PU errors because of larger HoA, but for DEM fusion quality analysis, the minimum value of HoAs (here 45.81) is considered to enumerate the number of PU errors. It should be noted that mean values presented in Tables II and IV do not present the real level of canopy penetration of the X-band radar signal. In our previous study [35], we found some amount of vegetation penetration to remain after DEM coregistration.

The results from Tables IV and V demonstrate the efficiency of the Huber model for fusion of two tiles of TanDEM-X raw DEMs in the problematic area. The results show that using the Huber model can significantly improve the RMSE of fused DEMs by up to nearly 2 m while the DEM quality enhancement by means of WA is not remarkable. Apart from this, the Huber model is absolutely more powerful than WA to reduce the PU errors. The maximum and minimum discrepancies also confirm the better performance of the Huber model to deal with PU errors in comparison to the other method. TV-\( L_1 \) also can decrease the noise effect in the final fused DEM and reduce the number of PU errors but the improvement is not as large as for the Huber model.

C. Fusion of TanDEM-X Raw DEMs With Different Baseline Configuration

In the final experiment, we focus on the fusion of DEMs acquired by different baseline configurations including different orbit directions and HoAs. Table VI provides the properties of the tiles used for this experiment. The raw DEMs covering...
TABLE IV
HEIGHT ACCURACY (IN METER) OF THE TANDEM-X DATA WITH DIFFERENT HOAS BEFORE AND AFTER DEM FUSION IN THE PROBLEMATIC STUDY AREA

| DEM         | Mean | RMSE | MAE | NMAD | STD  |
|-------------|------|------|-----|------|------|
| Raw DEM     | -2.35| 10.77| 8.46| 10.10| 10.51|
| id: 1023491 | -2.35| 10.57| 8.27| 9.69 | 10.30|
| id: 1058842 | -2.37| 10.45| 8.23| 9.81 | 10.17|
| WA          | -2.63| 9.24 | 7.13| 8.03 | 8.86 |
| Fused DEM   | -2.35| 8.60 | 6.70| 7.65 | 8.277|
| TV-L1       |      |      |     |      |      |
| Huber       |      |      |     |      |      |

The bold values indicate the best results.

TABLE V
EFFECT OF DEM FUSION TO REDUCE THE NUMBER OF PU ERRORS USING TILES WITH DIFFERENT HOAS IN THE PROBLEMATIC STUDY AREA

| DEM         | HoA  | PU Threshold | No. of PU Errors | Max Discrepancy | Min Discrepancy |
|-------------|------|--------------|------------------|-----------------|-----------------|
| Raw DEM     | 45.81| 30.36        | 2032             | 51.80           | -73.13          |
| id: 1023491 | 72.02| 50.01        | 51               | 58.82           | -54.76          |
| id: 1058842 | 45.81| 30.36        | 1339             | 50.74           | -53.39          |
| WA          |      |              |                  |                 |                 |
| Fused DEM   | 45.81| 30.36        | 102              | 19.16           | -33.76          |
| TV-L1       |      |              |                  |                 |                 |
| Huber       | 45.81| 30.36        | 0                | 16.97           | -28.71          |

The bold values indicate the best results.

TABLE VI
PROPERTIES OF THE NOMINAL ASCENDING AND DESCENDING TANDEM-X RAW DEM TILES OVER TERRASSA AND VACARISSES CITIES

| TanDEM-X raws DEms | Acquisition Id | 1058683 | 1171358 |
|---------------------|----------------|---------|---------|
| Acquisition mode    | Stripmap       | Stripmap|         |
| Center incidence angle | 33.71°        | 34.82°  |         |
| Equator crossing direction | Ascending    | Descending|       |
| Look direction       | Right          | Right   |         |
| Polarization         | HH             | HH      |         |
| Height of ambiguity  | 60.18 m        | 48.58 m |         |
| Pixel spacing        | 0.2 arc-sec    | 0.2 arcsec|       |
| HEM mean             | 1.17 m         | 1.40    |         |

Terrassa and Vacarisses cities located in Spain were produced by ascending and descending acquisitions. In addition to orbit directions, the HoAs of tiles are also not similar to each other. Fig. 7 shows the TanDEM-X raw DEMs used in this study, which mostly covers difficult terrain, the common area is specified by black polygons. Due to morphologically difficult type of terrain, the acquisitions from ascending and descending flight paths have been applied for global DEM generation in this area. However, the study cities are located in the relatively flat part of the area common between tiles. Again, from these tiles, study subsets located in different land types were selected. Fig. 8 display each study subsets from different types extracted from the common area of ascending and descending raw tiles.

The results of fusing ascending and descending raw DEMs in different land types over urban area are provided in Table VII. The accuracy evaluation is performed by comparing each DEM respective to a LiDAR DSM, which was achieved by interpolation of the LiDAR point cloud presented by the ISPRS foundation as a benchmark [36]. On an average, the density of the point cloud is about one point per square meter.

The results of DEM fusion again illustrate that using variational models can increase the accuracy of the initial input raw DEMs. In urban study subsets, the performance of the Huber model is slightly better than TV-L1 according to the statistical metrics, but their differences are not really significant. It can be concluded that both models produce similar results in terms of statistical measurements. In comparison to WA, variational models also give a more accurate DEM in urban areas and the agricultural subsets.

IV. DISCUSSION

In this study, the TV-L1 and Huber variational models were implemented to fuse TanDEM-X raw DEMs over urban areas as well as surroundings. In particular, we investigated these models with respect to the fusion of raw DEMs produced from data takes with different baseline configurations and HoAs. In conclusion, the results demonstrated the efficiency of variational models in comparison to simple WA for the TanDEM-X raw DEM fusion. To clarify the role of smoothness constraint and data term, we...
carried out an experiment regarding the DEM quality improvement to be achieved by just carrying out TV-$L_1$ denoising of a single input DEM. Comparing these results with those achieved by TV-$L_1$ DEM fusion (which employs elevation data of at least two DEM tiles) revealed that fusion is always favorable (cf. Table VIII). Furthermore, it can be seen that in the industrial subset, the main improvement arises from the smoothness term, which is caused by the regular scene structure. However, adding another tile in a fusion manner can still improve the quality of the final DEM. In contrast, for the agricultural subset, the TV-$L_1$ denoising could not change the DEM quality, while DEM fusion could finally produce a DEM with higher accuracy. Using more DEM tiles is furthermore vital for areas suffering from layover and shadowing effects or containing PU errors. More examples of these areas and requirement for employing several tiles can be found in [8]. In addition, regarding the strict quality control policy of the TanDEM-X mission, obtaining lower than 2 m relative height accuracy for slopes lower than 20% and better than 4 m for steeper slopes in each pixel means that the pixelwise TanDEM-X target accuracy can only be realized by DEM fusion. Table VIII provides some statistics relevant to TanDEM-X quality control indicating percentages of pixels with an accuracy better than 2 and 4 m as well as the percentage of pixels with accuracy worse than 4 m. The results confirm that DEM fusion can lead to obtaining more reliable pixels in comparison to just the denoising of single DEMs. Another important problem with using a single tile is the selection of the accurate subsets in different land types. As an example, in experiment 1, the accuracy of DEM with id 1145180 is higher than the accuracy of DEM 1023491 for the forested area while for other study subsets the quality of DEM 1023491 is better than for the other DEMs. As a result, in practice, it is beneficial to carry out DEM fusion in general, as this always improves the quality of the final DEM.

A. Use of TV-Based Variational Models

The main property of TV-based models is to reduce the effect of noise by minimizing the TV term. It should be noted that both data and regularization terms in the energy functional defined for TV-$L_1$ and Huber models are positive terms. Choosing TV as a regularization term leads to preserving the beneficial high-frequency image contents such as footprints of buildings while minimizing its value through the fusion causes to reduce the effects of undesirable noise. Fig. 9 shows the performance of TV-based variational models in comparison to WA in a 3-D view. The displayed patch was selected from an industrial area located in Munich, which was used in the first experiment.

The 3-D display of the fused DEMs clearly shows that the TV-based model can reduce the noise effect and excellently reveal the edges while the WA-based fused DEM still suffers from noise effects. As displayed, the Huber model produces a smoother output in comparison to TV-$L_1$ because of mixing the quadratic norm and the $L_1$ norm to form data and regularization terms. Since the quadratic norm tends to penalize the high-frequency contents more severe than $L_1$, it leads to DEMs with more smoother edges. Apart from the type of norm used to form an energy functional, the amount of smoothing induced by TV-based variational models depends on the regularization parameter, which trades off between the TV term as a regularization term and the $L_1$ norm to form data and regularization terms. Since the quadratic norm tends to penalize the high-frequency contents more severe than $L_1$, it leads to DEMs with more smoother edges. Apart from the type of norm used to form an energy functional, the amount of smoothing induced by TV-based variational models depends on the regularization parameter, which trades off between the TV term as a regularization term and the data fidelity term. While only one regularization parameter is required to be tuned for DEM fusion by TV-$L_1$, using the Huber model for fusion demands to tune three parameters. Selecting different thresholds to form the norms used in the Huber model changes the amount of smoothness that emerges in the final output of DEM fusion. Fig. 10 displays the effect of changing one of the parameters while the others are constant on the final output. Selecting small $\alpha$, which is used for a data term, does not severely penalize discrepancies between the initial DEMs and the desired output strongly, i.e., giving an output fused DEM with more similarity to input data. In contrast, increasing $\alpha$ penalizes the discrepancies intensively and the optimization process tries to lower the total energy that provides a smoother DEM at the end. An identical interpretation can be derived for $\beta$ while this parameter performs in a
reverse manner because it is used to form the regularization term. It should be noted that the regularization parameter $\gamma$ trades off between two terms in functional energy that means by increasing $\gamma$, the effect of TV will become lower such that ultimately smoother DEM is produced. Appropriately tuning the regularization parameter and the Huber model thresholds also influences the accuracy of the final fused DEM. The effect of Huber model parameter values on the final accuracy of DEM fusion is depicted in Fig. 11. Different methods can be used for tuning the regularization parameter. One option is to learn it from data if some training data are available. Another option is to use the L-curve approach [32]. Finally, the parameter can be manually selected based on a visual analysis of different output DEMs.

### B. Fusion Over Different Land Types

In this study, the TanDEM-X DEM fusion by variational models was implemented over different land types that are typically found in urban areas and in their surroundings. Fig. 12 depicts the accuracy improvement (in meters) by means of different fusion algorithm respective to the quality of the initial DEMs for each study land type used in the first experiment (see Section III-A). Similarly, Fig. 13 compares the performance of fusion methods for different land types that were used in the third experiment (see Section III-C). It should be noted that since both variational model have similar performance in terms of RMSE, for each plot in Figs. 12 and 13, just the performance of the best variational model is compared to WA.

---

**TABLE VII**

| Study area       | DEM       | Mean | RMSE | MAE | NMAD | STD |
|------------------|-----------|------|------|-----|------|-----|
| Raw DEM          | 1058683   | -0.19| 3.49 | 2.49| 2.60 | 3.48|
|                  | 1171358   | -0.19| 3.56 | 2.44| 2.34 | 3.55|
| Industrial WA    | -0.26     | 3.06 | 2.13 | 2.07| 3.05 |
|                  | TV-L$_1$  | -0.34| 2.92 | **2.09**| 2.07 | 2.90|
|                  | Huber     | -0.19| 2.89 | 2.10| 2.14 | **2.88**|
| Raw DEM          | 1058683   | -0.78| 5.05 | 3.52| 3.70 | 4.99|
|                  | 1171358   | -0.78| 5.11 | 3.53| 3.62 | 5.05|
| Inner WA         | -0.76     | 4.66 | 3.22 | **3.36**| 4.59 |
|                  | TV-L$_1$  | -0.91| 4.35 | **3.08**| 3.40 | **4.25**|
|                  | Huber     | -0.78| 4.34 | 3.13| 3.52 | 4.27 |
| Raw DEM          | 1058683   | -0.54| 4.24 | 3.11| 3.19 | 4.20|
|                  | 1171358   | -0.54| 4.42 | 3.21| 3.26 | 4.38|
| Residential WA   | -0.62     | 3.94 | 2.87 | 2.83| 3.90 |
|                  | TV-L$_1$  | -0.76| 3.96 | 2.88| 2.77 | 3.88|
|                  | Huber     | -0.54| **3.86**| 2.86| 2.74 | **3.82**|
| Raw DEM          | 1058683   | 0.44 | 2.38 | 1.68| 1.71 | 2.34|
|                  | 1171358   | 0.44 | 1.93 | 1.23| 0.98 | 1.88|
| Agricultural WA  | 0.35      | **1.60**| 1.04| 0.83| 1.57 |
|                  | TV-L$_1$  | **0.27**| **1.60**| **1.04**| **0.78**| 1.59 |
|                  | Huber     | 0.44 | 1.62 | 1.12| 0.91 | **1.56**|

The bold values indicate the best results.

**TABLE VIII**

| Study area | Strategy                  | RMSE  | MAE  | NMAD |
|------------|---------------------------|-------|------|------|
| Residential| TV-L$_1$ DEM denoising    | 3.88  | 2.88 | 2.20 |
|            | TV-L$_1$ DEM fusion       | 3.67  | 2.69 | 2.03 |
| Residential| TV-L$_1$ DEM denoising    | Error $< 2m$ | 49 % | 78 % | 22 % |
|            | TV-L$_1$ DEM fusion       | Error $< 4m$ | 54 % | 81 % | 19 % |
| Study area | Strategy                  | RMSE  | MAE  | NMAD |
| Agricultural| TV-L$_1$ DEM denoising  | 0.86  | 0.57 | 0.59 |
|            | TV-L$_1$ DEM fusion       | 0.55  | 0.29 | 0.20 |
| Agricultural| TV-L$_1$ DEM denoising  | Error $< 2m$ | 95 % | 100 % | 0 % |
|            | TV-L$_1$ DEM fusion       | Error $< 4m$ | 100 % | 100 % | 0 % |

The bold values indicate the best results.
Fig. 9. 3-D display of initial TanDEM-X raw data and the results of DEM fusions using different methods in the industrial area used in the first experiment. (a) TanDEM-X (tile a). (b) TanDEM-X (tile b). (c) WA. (d) Huber model. (e) TV-L1 model. (f) LiDAR.

Fig. 10. Effect of varying Huber models’ parameters on the final DEM. (a) $\gamma = 1$, $\alpha = 0.5$, $\beta = 1$. (b) $\gamma = 1$, $\alpha = 10$, $\beta = 1$. (c) $\gamma = 1$, $\alpha = 1$, $\beta = 0.5$. (d) $\gamma = 1$, $\alpha = 1$, $\beta = 10$.

Fig. 11. Influence of different Huber norm parameters on the RMSE of fused DEM.

Fig. 12. Improvement of the TanDEM-X DEM tiles (a), (b) using variational models (here, TV-L1) in comparison to WA in different study areas located in Munich. The bars indicate the difference between the RMSE of input TanDEM-X DEM and final fused DEM. (a) Refers to tile 1023491. (b) Refers to tile 1145180.

Fig. 13. Improvement of the TanDEM-X DEM tiles (a), (b) using variational models (here, Huber) in comparison to WA in different study areas located in Vacarisses and Terrassa. The bars indicate the difference between the RMSE of input TanDEM-X DEM and final fused DEM. (a) Refers to tile 1058683. (b) Refers to tile 1171358.
The plots demonstrate that variational models exhibit maximum efficiency in inner city land types in both experiments while WA has a nearly similar performance in different land types. The lowest accuracy improvement by variational models is for residential subset and nonurban study areas. The inner city land type includes a lot of building footprints that mostly appear as noisy edges because of inherent properties of SAR sensor imaging. Consequently, using the TV-based variational model can significantly improve the DEM quality in these areas. On the other hand, residential subset areas include single family, small homes usually located in a sparse pattern, and the footprint of buildings, which cannot appear as a strong edge in TanDEM-X raw DEM due to resolution restriction of data takes acquired in the stripmap mode. In nonurban areas, the edginess is usually lower than in urban subsets. Thus, the smoothness term of the variational models has lower performance in those kinds of land types. However, the quality of the final DEM still increases due to the DEM fusion encoded in the data term.

C. Effect of Geometry

While most urban areas covered by global TanDEM-X dataset are generated by two nominal acquisitions that mostly have similar HoAs and geometries, we also investigated the fusion of several TanDEM-X DEMs with different properties to investigate the performance of variational models for these data. In the first experiment, the results identified that the variational models can perfectly fuse the raw DEMs with nearly similar baseline configuration and HoAs acquired over urban areas. The output is a DEM with higher accuracy and more enhanced building footprints. However, the Huber model generates a smoother DEM at the end.

A significant result was yielded for problematic areas where the effects of PU errors are dominant. The selected study subset (see Fig. 6) is mostly affected by noise because of the volume decorrelation due to trees and the low coherence due to river. For these problematic areas fusing one DEM with nominal HoA to another DEM with larger HoA is more useful to reduce the effect of PU errors. In this experiment, fusing two DEMs with different HoAs by using the Huber model could substitute inconsistent heights with logical values and also resulted in a more accurate DEM. This proves, in addition to the DEM fusion methodology, that selecting appropriate raw DEM tiles dependent to problem is significant for a successful fusion. Among variational models, TV-$L_1$ can decrease the number of PU errors and improve the accuracy but more quality enhancement is achieved by the Huber
model. Since, the Huber model also uses the quadratic norm, it produces a smoother fused DEM while TV-L1 tends to save more high-frequency contents that can also be caused by noise.

Fusing ascending and descending DEMs in problematic areas reduces the layover and shadow effects in the final fused DEM. Consequently, in the final experiment, two ascending and descending DEMs with different HoAs were fused. As shown in plots 12 and 13, in comparison to results of fusing DEMs with similar baseline configuration and HoAs, the variational models lead to least significant quality improvement in the final fused DEM in terms of RMSE. However, a display of an exemplary study subset (industrial area) in Fig. 14 demonstrates the efficiency of variational models in comparison to WA for fusing these types of DEMs. For making a correct judgment about the performance of variational models on fusing ascending and descending DEMs versus DEMs with similar flight paths, two DEMs with similar baseline configuration and HoA from the study areas are required. Theoretically, apart from the DEM fusion method, using ascending and descending DEMs instead of using DEMs with similar orbit directions improves the final DEM quality in the difficult terrains and problematic areas such as urban areas that are under the shadow and layover effects. In practice, it is confirmed in [8] that using ascending and descending raw TanDEM-X DEMs can produce highly accurate fused DEM at the end in the shadow- and layover-affected areas.

V. CONCLUSION

In this paper, we proposed to apply TV-based variational models (TV-L1 and Huber models) for TanDEM-X raw DEM fusion at the phase of DEM mosaicking instead of WA. The main focus of this study was to enhance final DEMs in urban areas where the footprints of buildings are influenced by noise effects due to SAR imaging properties. For this purpose, different study subsets were selected from different land types, which mostly are explored over urban areas and surroundings. Apart from this, DEM fusion was investigated for raw DEMs with different geometries. At first, two nominal acquisitions with similar baseline configurations and HoAs were fused over different land types. In the next experiment, two raw DEMs with different HoAs were fused over a problematic terrain that suffers from PU errors. At the end, two DEMs with ascending and descending orbit directions as well as with different HoAs were used. In all experiments, it was demonstrated that using variational models leads to DEMs with higher quality. A great performance of the Huber model was recorded for fusing two raw DEMs with different HoAs over the selected problematic area. Also, in urban areas, variational models with reducing the noise effect and enhancing the outlines of buildings, absolutely performs better than WA. However, the Huber model tends to provide a smoother fused DEM than TV-L1. The results also demonstrated that the variational models, particularly TV-L1, could improve the quality of DEMs significantly in comparison to WA. Using variational models could improve the DEM quality by up to 2 m particularly in inner city subsets. In conclusion, carrying out TanDEM-X raw DEM fusion using variational models with an ability to enhance the building footprints and other useful high-frequency contents along with smoothing the noise, finally produced a DEM with higher quality.

ACKNOWLEDGMENT

The authors would like to thank Dr. Fritz of DLR and Dr. Baier for providing the TanDEM-X raw DEM for the Munich area; and the Bavarian Surveying Administration for providing the LiDAR data of Munich.

REFERENCES

[1] E. Rodriguez, C. S. Morris, and I. E. Belz, “A global assessment of the SRTM performance,” Photogramm. Eng. Remote Sens., vol. 72, no. 3, pp. 249–260, 2006.
[2] G. Krieger et al., “TanDEM-X: A satellite formation for high-resolution SAR interferometry,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 11, pp. 3317–3341, Nov. 2007.
[3] P. F. Noceti et al., “Generation and performance assessment of the global TanDEM-X digital elevation model,” ISPRS J. Photogramm. Remote Sens., vol. 132, no. Suppl C, pp. 119–139, 2017.
[4] T. Fritz, H. Breit, C. Rossi, U. Balss, M. Lachaise, and S. Dueque, “Interferometric processing and products of the TanDEM-X mission,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., Jul. 2010, pp. 1904–1907.
[5] J.-H. Gonzalez et al., “Bistatic system and baseline calibration in TanDEM-X to ensure the global digital elevation model quality,” ISPRS J. Photogramm. Remote Sens., vol. 73, pp. 3–11, 2012.
[6] A. Gruber, B. Wessel, M. Huber, and A. Roth, “Operational TanDEM-X DEM calibration and first validation results,” ISPRS J. Photogramm. Remote Sens., vol. 73, pp. 39–49, 2012.
[7] B. Wessel et al., “Design of the DEM mosaicking and calibration processor for TanDEM-X,” in Proc. 7th Euro. Conf. Synthetic Aperture RADAR, Jun. 2008, pp. 1–4.
[8] A. Gruber, B. Wessel, M. Martone, and A. Roth, “The TanDEM-X DEM mosaicking: Fusion of multiple acquisitions using InSAR quality parameters,” IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens., vol. 9, no. 3, pp. 1047–1057, Mar. 2016.
[9] M. Schmitt and X. X. Zhu, “Data fusion and remote sensing: An ever-growing relationship,” IEEE Geosci. Remote Sens. Mag., vol. 4, no. 4, pp. 6–23, Dec. 2016.
[10] R. Deo, C. Rossi, M. Eineder, T. Fritz, and Y. S. Rao, “Framework for fusion of ascending and descending pass TanDEM-X raw DEMs,” IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens., vol. 8, no. 7, pp. 3347–3355, Jul. 2015.
[11] R. Reinartz, R. Müller, D. Hoja, M. Lehner, and M. Schroeder, “Comparison and fusion of DEM derived from SPOT-5 hrs and SRTM data and estimation of forest heights,” in Proc. Earsoil Workshop 3D-Remote Sens. Earsoil Symp., Porto, Portugal, Jun. 2005, p. (on CD ROM).
[12] D. Hoja and P. d’Angelo, “Analysis of DEM combination methods using high resolution optical stereo imagery and interferometric SAR data,” Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci., vol. 37, 2009, p. (on CD ROM).
[13] A. Roth, W. Knopfle, G. Strunz, M. Lehner, and P. Reinartz, “Towards a global elevation product: combination of multi-source digital elevation models,” Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., vol. 34, no. 4, pp. 675–679, 2004.
[14] W. Knöpfle, G. Strunz, and A. Roth, “Mosaicking of digital elevation models derived by SAR interferometry,” Int. Arch. Photogramm. Remote Sens., vol. 32, pp. 306–313, 1998.
[15] D. Just and R. Bamler, “Phase statistics of interferograms with applications to synthetic aperture RADAR,” Appl. Opt., vol. 33, pp. 4361–4368, Jul. 1994.
[16] H. Bagheri, M. Schmitt, and X. X. Zhu, “Fusion of TanDEM-X and Cartosat-1 elevation data supported by neural network-predicted weight maps,” ISPRS J. Photogramm. Remote Sens., vol. 144, pp. 285–297, 2018.
[17] H. Papasaika, E. Kokkiopoulou, E. Baltasavias, K. Schindler, and D. Kressner, “Fusion of digital elevation models using sparse representations,” in Proc. ISPRS Conf. Photogramm. Image Anal., 2011, pp. 171–184.
[18] C. Zach, T. Pock, and H. Bischof, “A globally optimal algorithm for robust TV-L1 range image integration,” in Proc. IEEE 11th Int. Conf. Comput. Vis., Oct. 2007, pp. 1–8.
[19] T. Pock, L. Zebbed, and H. Bischof, TGV-Fusion. Berlin, Germany: Springer, 2011, pp. 245–258.
Hossein Bagheri received the master’s degree in photogrammetry and remote sensing engineering from Tafresh University, Tafresh, Iran, in 2013. He is currently working toward the Ph.D. degree with the chair of Signal Processing in Earth Observation from Technical University of Munich, Munich, Germany. In 2015, he spent six months with the Photogrammetry Group, Department of Surveying and Geomatics Engineering, University of Tehran, Tehran, Iran. His research interests include multisensor data fusion applied to SAR and optical data, 3-D reconstruction, and digital elevation models.

Michael Schmitt (S’08–M’14–SM’16) received the Dipl.-Ing. degree in geodesy and geoinformation and the Dr.-Ing. degree in remote sensing from the Technical University of Munich (TUM), Munich, Germany, in 2009 and 2014, respectively.

Since 2015, he has been a Senior Researcher and the Deputy Head with the Professorship for signal processing, Earth Observation, TUM. In 2016, he was a Guest Scientist with the University of Massachusetts Amherst, Amherst, MA, USA. His research interests include signal and image processing as well as machine learning for the extraction of information from remote sensing data, data fusion with emphasis on the joint exploitation of optical and radar data, 3-D reconstruction by techniques, such as SAR interferometry, SAR tomography, radargrammetry, or photogrammetry, and millimeter-wave SAR remote sensing.

Dr. Schmitt is a Co-Chair of the International Society for Photogrammetry and Remote Sensing Working Group I/3 on SAR and Microwave Sensing. He frequently serves as a Reviewer for a number of renowned international journals. In 2013 and 2015, he was elected as the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS Best Reviewer, leading to his appointment as an Associate Editor of the journal in 2016.

Xiao Xiang Zhu (S’10–M’12–SM’14) received the master’s (M.Sc.) degree, her Doctor of Engineering (Dr.-Ing.) degree, and her Habilitation in the field of signal processing from the Technical University of Munich (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

She is currently the Professor for signal processing in earth observation (www.sipeo.bgu.tum.de) with the Technical University of Munich (TUM), Munich, Germany and German Aerospace Center (DLR), Cologne, Germany; the Head of the Department “EO Data Science” at DLR’s Earth Observation Center; and the Head of the Helmholtz Young Investigator Group “SiPEO” at DLR and TUM. She was a Guest Scientist or Visiting Professor with the Italian National Research Council (CNR-IREA), Naples, Italy, Fudan University, Shanghai, China, the University of Tokyo, Tokyo, Japan, and the University of California, Los Angeles, CA, USA, in 2009, 2014, 2015, and 2016, respectively. Her main research interests are remote sensing and earth observation, signal processing, machine learning, and data science, with a special application focus on global urban mapping.

Dr. Zhu is a member of young academy (Junge Akademie/Junge Kolleg) at the Berlin-Brandenburg Academy of Sciences and Humanities and the German National Academy of Sciences Leopoldina and the Bavarian Academy of Sciences and Humanities. She is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.