# AUTOMATIC LARGE AREA RECONSTRUCTION OF BUILDING FAÇADES FROM SPACEBORNE TOMOSAR POINT CLOUDS

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

(1) Chair of Remote Sensing Technology (LMF), Technical University Munich (TUM), Arcisstrasse 21, 80333 Munich, Germany
(muhammad.shahzad@bv.tum.de)
(2) Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Wessling, Germany
(xiao.zhu@dlr.de)

## ABSTRACT

Improved resolution of SAR sensors and advanced multi-pass interferometric techniques, such as tomographic SAR inversion (TomoSAR), opens up new possibilities of 4D (or even higher dimensional) imaging that can be potentially used to reconstruct dynamic models of entire cities, i.e., city models that can incorporate temporal (motion) behaviour along with the 3D information. Motivated by these chances, this paper presents an approach that systematically allows *automatic* reconstruction of building façades from 4D point cloud generated from tomographic SAR processing and put particular focus on *robust* reconstruction of large areas. The approach is modular and is illustrated/validated by examples using TomoSAR point clouds generated from a stack of TerraSAR-X high-resolution spotlight images from ascending orbit covering approx. 2 km$^2$ high rise area in the city of Las Vegas.

***Index Terms***— TomoSAR, point cloud, façade reconstruction, 4D city model, TerraSAR-X

## 1. INTRODUCTION

Modern spaceborne SAR sensors such as TerraSAR-X and COSMO-SkyMed can deliver meter-resolution data that fits well to the inherent spatial scales of buildings. This very high resolution (VHR) data is therefore particularly suited for detailed urban mapping. In particular, using stacked VHR SAR images, advanced multi-pass interferometric techniques such as tomographic SAR inversion (TomoSAR) allow to retrieve not only the 3D geometrical shape but also the undergoing temporal motion of individual buildings and urban infrastructures [1][2]. The resulting 4D point clouds have a point (scatterer) density that is comparable to LiDAR. E.g. experiments using TerraSAR-X high-resolution spotlight data stacks show that the scatterer density retrieved using TomoSAR is on the order of 1 million pts/km$^2$ [3]. Object reconstruction from these high quality TomoSAR point clouds can greatly support the reconstruction of dynamic city models that could potentially be used to monitor and visualize the *dynamics* of urban infrastructure in very high level of details. Motivated by this, we presented very first results of façade reconstruction from single view (ascending stack) and multi-view (fused ascending and descending stacks) perspectives over a small test building area (Bellagio hotel, Las Vegas) in [4] and [5] respectively. Figure 1 shows the TomoSAR points colorcoded according to the amplitude of the seasonal motion overplotted onto the resulting façades model.

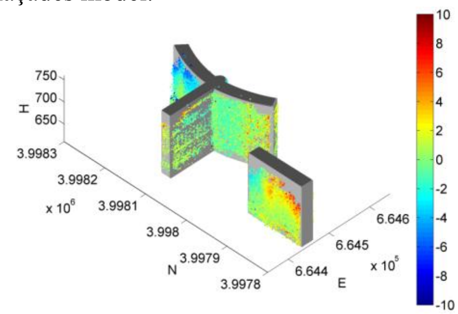



Figure 1: Reconstructed façade model with overplotted TomoSAR points [5]. Colorbar represents the amplitude of seasonal motion (line-of-sight) caused by thermal dilation in millimeters. Axis labels represents 'N' for northing, 'E' for easting in UTM coordinates and 'H' for height in meters.

Yet towards automatic reconstruction for the whole city area, more robust and fully automatic approach is needed. In this paper, we extend our previously introduced approach aiming at finding more general solution towards automatic reconstruction of the whole city area. The four major contributions presented in this work include: 1) Robust façade extraction procedure incorporating the facade geometry; 2) Unsupervised clustering without prior knowledge about the number of clusters; 3) More robust reconstruction of façades by removing conflicting segments occurring at façade transition areas; 4) Partially coping with the problem of façade occlusion which is caused by nearby higher building structures present in the area of interest.

## 2. FAÇADES (DETECTION) EXTRACTION

The proposed approach takes into account the characteristics of TomoSAR point clouds introduced by the side-looking SAR geometry. When we project the TomoSAR point clouds onto ground plane, the scatterer (point) density *SD* of the vertical façade regions depends on the building heights and is in general higher than the non-façade regions. It is mostly true due to the existence of strong corner reflectors, e.g., window frames on the building façades. Taking this fact into account, in [5], we proposed to extract façade points by a simple $SD$ thresholding followed by some morphological operations. The approach works well for high rise buildings having much higher point density but limits the extraction of façade points from lower buildings. The selection of a particular threshold thus becomes crucial. To resolve this issue, we use a more robust façade extraction approach based on directional $SD$ estimation procedure that locally estimate the $SD$ for each point while incorporating the façade geometry [6]. To be more explicit, we estimate direction of the local (cylindrical) neighborhood $v_c$ via line fitting using robust M-estimator. The method computes M-estimates by iteratively applying weighted least squares to the objective function $\arg\min_{\beta} \sum_{p_i \in v_c} w_{p_i}(\beta) \left| y_{p_i} - f_{p_i}(\beta) \right|^2$ where weights $w_{p_i}$ are computed using bisquare function of the form [7]:

$$w_{p_i} = \begin{cases} \left(1-u^2\right)^2 & \text{for } \mathrm{abs}(u) < 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{with} \quad u = \frac{\left| y_{p_i} - f_{p_i}(\beta) \right|^2}{4.685\hat{\sigma}\sqrt{1-t}} \qquad (1)$$

where $\beta$ are estimated line parameters that iteratively updated, $t$ is the leverage computed from least squares fit, $\hat{\sigma}$ is the estimate for the scale of the error term computed by $\hat{\sigma} = 1.483 * \mathrm{MAD}$ where MAD is the median absolute deviation of the residuals from their median. The term 1.483 is used to make the estimator consistent for the estimation of standard deviation at Gaussian distribution and has been practically used for robust initializations for solving problems involving computation of M-estimates [7][8].

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

Based on the estimated $SD$, façade points can be extracted. For large area, both high and low buildings are present. To avoid miss-detection of lower buildings, we extract façade points in a sequential way. Firstly, points having $SD$ value less than a *soft* threshold are removed. Usually, remaining points include not only façade points but also other non-façade points having higher $SD$, e.g., building roof points. These non-façade points are removed prior to further processing via normals estimation. Façade points are thus extracted out by retaining only those points having normals between ±15 degrees from the horizontal axis.

### 3. RECONSTRUCTION VIA SEGMENTATION

Extracted façade points are further segmented to points belonging to the same individual façade as follows:

- Extracted points are coarsely clustered by a density based connectivity approach as proposed by [10];
- Surface normals are computed locally for each point and the mean shift algorithm is used for clustering points having smaller angular difference in feature space (Gaussian image *GI*) into one cluster [11];
- Previous step results in clusters of points that have similar normal directions but may be spatially far from each other. To cope this, spatial connectivity is used for further clustering of points.

Each cluster is further classified into flat or curved surface by analyzing derivatives of the local orientation angle $\theta$ (= azimuthal angle of the surface normal). Identified façade clusters in $xy$ plane are then modeled using the general polynomial equation [5]:

$$f_p(x,y) = \sum_{q=1}^{p} a_q x^i y^j \text{ with } i + j \le q \qquad (2)$$

Here $i$ and $j$ are permuted accordingly, $p$ is the order of polynomial, the number of terms in the above polynomial is equal to $(p + 1)(p + 2)/2$. Cross terms are introduced in the model in case of rotated local coordinate system. To solve (2), we restrict ourselves to 1st and 2nd order (i.e., *flat* with $\max(i,j) = 1$ & *curved* with $\max(i,j) = 2$). The coefficients $a_q$ are estimated using weighted total least squares (WTLS) method where total least squares is utilized to cope for localization errors of TomoSAR points in both $x$ and $y$ directions and the weight of each point is assigned equal to its corresponding $SD$. The weighted polynomial fitting (residual) error $f_{err}$ is minimum for the case where we have unrotated local coordinate system reducing right hand side of (2) to $\sum_{i=0}^{p} a_i x^i$ (i.e. with no cross terms). In case of rotated local coordinate system (which is often the case), we perform following steps to obtain consistent parameter estimates of all façades in a global coordinate system: 1) Rotate the points by rotation angle $\alpha$ and compute polynomial fitting error $f_{err}$ by applying WTLS method; 2) Consider coefficients computed with $\alpha_{\min}$ that gives the minimum polynomial fitting error $f_{err}$ as polynomial terms. $\alpha_{\min}$ is computed by using an unconstrained nonlinear optimization procedure to find the minimum of the error function $f_{err}$ by varying $\alpha$ over 0~360 degree range via Nelder-Mead simplex algorithm [12]; 3) Rotate the computed polynomial by replacing the *unrotated* $(x-, y-)$ axis terms by their rotation counterparts $(x\cos\alpha + y\sin\alpha, -x\sin\alpha + y\cos\alpha)$ to yield polynomial terms $a_q$ in global coordinates.

After estimation of model parameters, the next step is to describe the overall shape of the building footprint by further identifying adjacent façades pairs and determining the intersection of the façade surfaces. The adjacency of façades is usually described by an adjacency matrix AM that is built up via connectivity analysis [5]. Identified adjacent façade segments are used to determine the vertex points (i.e., façade intersection lines in 3D) by computing the intersection points between any adjacent façade pair. Determination of these intersection points can sometimes become difficult if the transition points (i.e., points occurring at the transition region of two adjacent façades) are segmented as isolated small clusters rather than part of the corresponding adjacent façade segments. As a consequence, it gets complicated to find a legitimate adjacent façade pair from which intersection points should be computed. To resolve this issue, such cases are first identified and then the intersection point is computed from the two largest segments only.

Moreover, the reconstructed façades could remain either incomplete or break into more than one segment due to the following reasons: 1) Higher building structures present nearby can partly (or fully) occlude the façades of lower buildings; 2) Due to the geometrical shape, only very few points are available at some parts of building façades. Computed vertex points are therefore first categorized into two types: First type consists of vertices that are computed from the intersection of two adjacent façades, while the second type consists of the other vertices representing "open" endpoints. Reconstructed façades are later refined by inserting additional segments between the broken regions and extend those façades that remain incomplete by statistically analyzing and matching the local height distribution of the nearest open endpoint vertices.

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

## 4. EXPERIMENTAL RESULTS

To validate our approach, we tested the algorithm on TomoSAR point clouds generated from a stack of 25 TerraSAR-X high spotlight images from ascending orbit only using the Tomo-GENESIS software developed at the German Aerospace Center [13]. The test area covers approx. 2 $km^2$ in the high rise part of the city Las Vegas. The number of TomoSAR points in the area of interest is about 1.2 million. Figure 2 shows the TomoSAR point cloud of our test area in universal transverse mercator (UTM) coordinates. The result of applying *SD* estimation procedure is illustrated in Figure 3. The two parameters *r* (radius of the neighborhood cylinder) and *d* are empirically set to 5m and 0.9m respectively according to the point density of the data set. One can observe that *TH* value influences the number of extracted façade points. Lower *TH* value results in higher completeness but lower correctness. In [5], we showed the results of estimating *SD* with varying area sizes and found that a kernel window of size 3x3$m^2$ and threshold *TH* value of about 2pts/$m^2$ results in best trade-off in terms of completeness and correctness with this class of data. 2pts/$m^2$ works well for high rise buildings but might ignore relatively smaller façades. Therefore to extract lower façades (and also to automate the procedure), we set the *TH* to the maximum of *SD* histogram value. This, as described in section 2, includes not only the façade points but additionally also some non-façade points with relative high *SD*, e.g., roof points. To reject these points from the set of extracted points after *SD* thresholding, surface normals information is utilized.

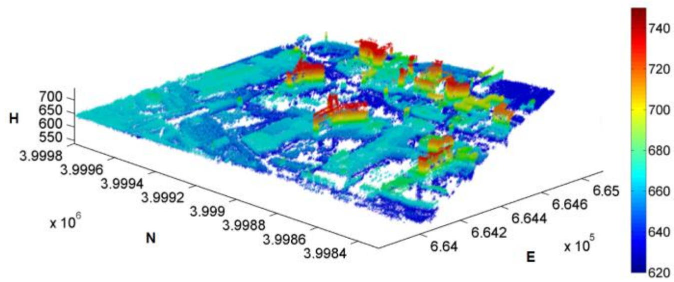



Figure 2: TomoSAR points in UTM coordinates of the area in the city of LasVegas. Height is color-coded.

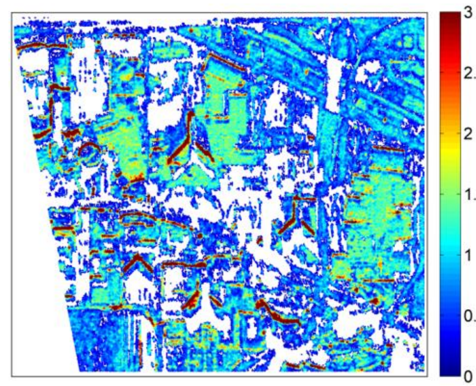



Figure 3: Scatterer (point) density with radius *r* = 5m and inliers *d* = 0.9m

Extracted façade points are then coarsely clustered based on connectivity of points. Subsequently, in order to reconstruct individual façades, mean shift clustering is applied in normals feature space (in *GI* domain) to the obtained density based coarse clusters. Figure 4(b) shows the estimated orientation angle $\theta$ for extracted façade points from single building shown in Figure 4(a). The variation in orientation angle is quite evident and allows mean shift to cluster points having similar orientations together. Further separation of points in the spatial domain is also required in some cases where the spatially separated points are clustered into one segment. This happens when these points belonging to different façades have similar normals and are spatially closer. Density based clustering is therefore again applied for spatial separation of the clusters within clusters.

Finally, Figure 5 depicts the reconstructed façades models of the area of interest. As depicted in [5], the shown reconstructed façade model can be used to refine the elevation estimates of the raw TomoSAR points. Moreover, with known deformation estimates of the scatterers, such a model can also lead to the reconstruction of dynamic city models that could potentially be used to monitor and visualize the dynamics of urban infrastructure in very high level of details.

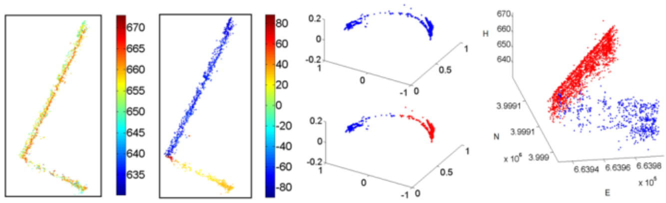
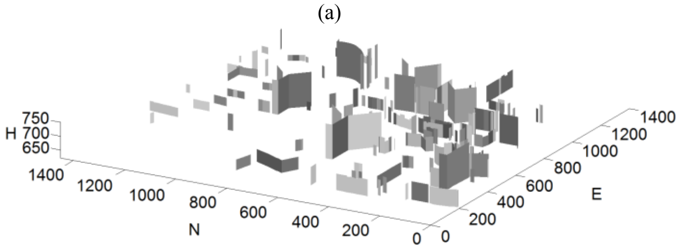

(a) (b) (c) (d)

Figure 4: Fine clustering results after applying mean shift clustering: (a) TomoSAR points of one particular density connected cluster (top view). Colorbar indicates height in meters; (b) Corresponding orientation angle in degrees; (c) Non clustered (top) and clustered (bottom) points in the Gaussian image of points in (a); (d) Resulting clustered points in 3D.


(a)



(b)

Figure 5: Reconstructed façades: (a) 2D view of the façade footprints after refinement overlaid onto the optical image; (b) 3D view. The axis is in meters range and has been translated to the origin for better metric clarity.

The actual ground truth data is missing for exact qualitative evaluation of the approach. In order to provide some quantitative measures of the algorithm performance, we manually counted the actual number of façades that were to be reconstructed. Total of 141 façades are present in the dataset out of which 7 are curved façades and remaining 134 are flat. Prior to refinement operation, the algorithm reconstructed in total of 176 façades, i.e., higher than the actual façades present in the dataset. As already stated in section 2, this is because some individual façades have been broken down into two or more segments due to discontinuity in the number of points available in the dataset. In the final reconstruction after refinement, we obtain 147 reconstructed facades. I.e., all 141 façades are successfully reconstructed; among them 5 façades remain broken (counted as additional 5 façades) and there is one case of false alarm which is actually covered metallic pedestrian bridge. Besides the 5 cases, we also find 7 facades that are not extended and therefore remain incomplete. This is however due to the inadequate number of points available in the data.

## 5. CONCLUDING REMARKS

In this paper we presented an automatic (parametric) approach for robust façade reconstruction for large areas using TomoSAR point clouds. It consists of three main steps: façade points extraction, segmentation of points belonging to individual façades and reconstruction via estimating façade model parameters and computing geometric 3D vertices. The approach allows for robust reconstruction of both higher façades and lower height structures, and hence is well suited for urban monitoring of larger areas from space. In the future, we will extend the algorithm towards automatic roof reconstruction and object based TomoSAR point clouds fusion.

## 6. REFERENCES

[1] R. Bamler, M. Eineder, N. Adam, S. Gernhardt, X. Zhu, "Interferometric Potential of High Resolution SpaceborneSAR", *PGF* 5/2009, pp. 407- 419, 2009.

[2] X. Zhu and R. Bamler, “Very High Resolution Spaceborne SAR Tomography in Urban Environment”, *IEEE TGRS* 48(12), pp. 4296-4308, 2010.

[3] X. Zhu, R. Bamler, “Demonstration of Super-resolution for Tomographic SAR Imaging in Urban Environment”, *IEEE TGRS* 50 (8), pp. 3150-3157, 2012.

[4] M. Shahzad, X. Zhu, R. Bamler, "Façade Structure Reconstruction Using SpaceborneTomoSAR Point Clouds", *IEEE IGARSS*, Munich, Germany, 2012.

[5] X. Zhu and M. Shahzad,"Façade Reconstruction Using Multi-View Spaceborne TomoSAR Point Clouds", *IEEE TGRS*, in press, 2013.

[6] Y. Wang, X. Zhu, R. Bamler, S. Gernhardt, "Towards TerraSAR-X Street View: Creating City Point Cloud from Multi-aspect Data Stacks", *JURSE 2013*, Sao Paolo, Brasil, 2013.

[7] J. O. Street, R. J. Carroll, D. Ruppert, "A Note on Computing Robust Regression Estimates via Iteratively Reweighted Least Squares" *The American Statistician*. Vol. 42, 1988, pp. 152–154.

[8] P. J. Huber, "Robust Statistics", Hoboken, NJ: John Wiley & Sons, Inc., 1981.

[9] M. Hubert, P. Rousseeuw and K. V. Branden, "Robpca: A new approach to robust principal components analysis", *Technometrics* 47, 2005.

[10] M. Ester, H. P. Kriegel , J. Sander , X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise", *KDD-96*, Portland, Oregon, USA, 1996.

[11] Y. Liu and Y. Xiong, "Automatic segmentation of unorganized noisy point clouds based on the gaussian map", *Computer-Aided Design* 40(5), pp. 576–594, 2008.

[12] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization", *The Computer Journal* 7, 308-313, 1965.

[13] X. Zhu, Y. Wang, S. Gernhardt, R. Bamler, "Tomo-GENESIS: DLR's tomographic SAR processing system", *JURSE 2013*, Sao Paulo, Brazil, 2013.