Potential of the "SARptical" System

Yuanyuan Wang[1], Xiao Xiang Zhu[1,2], Sina Montazeri[2], Jian Kang[1], Lichao Mou[1], Michael Schmitt[1]

(1) Signal Processing in Earth Observation, Technische Universität München, Arcisstraße 21, 80333 Munich, Germany
(2) Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Weßling, Germany

1. INTRODUCTION

Very high resolution SAR images in dense urban area are not trivial to interpret due to the inevitable layover caused by the side-looking imaging geometry. With the growing attention on very high resolution SAR data, the fusion of optical and SAR images in dense urban area has become an emerging and timely topic, because the complementary of these two data types can lead us to unprecedented insights and findings, such as the unique scattering mechanisms of different urban infrastructures. Lying at the basis of such fusion topic is the challenging task of the co-registration of SAR and optical images. Such two images are acquired with intrinsically different imaging geometries, and thus are nearly impossible to be co-registered without a precise 3-D model of the imaged scene. Only until recently, the "*SARptical*" [1], [2] system proposed a promising solution to tackle this challenging task. SARptical can trace individual SAR scatterers in corresponding high resolution optical images where we can analyze the geometry, material, and other properties of the imaged object. Vice versa, the similar study can also be done in the SAR image coordinate. This paper demonstrate the capability of the SARptical system, and its potential in various different applications and research directions.

2. THE SARPTICAL SYSTEM

The general framework of the proposed approach is shown in Figure 1. The framework applies to a stack of SAR images and a pair of (or more) optical images. The focus of SARptical is put on linking the attributes from optical image to the SAR image by 3-D matching and projection. The basic idea is to match the 3-D models derived from SAR and optical images respectively. As a result, the 2-D SAR and optical images will also be matched. Based on the matched images, subsequent tasks such as semantic label texturing and joint deformation analysis can be conducted. The detailed procedures of SARptical are as follows.

- 3-D reconstructions

a) Retrieve the 3-D positions and deformation parameters of the scatterers from the SAR image stacks. Since urban area is of our main interest, tomographic SAR inversion (TomoSAR), including SAR tomography and differential SAR tomography, is employed in order to resolve a substantial amount of layovered scatterers.

TomoSAR is the most computationally expensive step in the framework. In addition, TomoSAR and other multipass SAR interferometry (InSAR) algorithms typically requires a fairly large SAR image stack (>20 images). The computational and image resource are the main limitation for this step.

b) Retrieve the 3-D positions of points from the optical images using stereo matching with structure from motion (SfM) if necessary. For covering large urban area, aerial or spaceborne images are preferred. This step also calibrates the camera parameters.

Stereo matching and SfM are well studied topics. Many matured algorithms and software are readily available.

- 3-D matching: Co-register the TomoSAR point cloud and the optical point cloud.

The main challenges present in this step are the different modalities of optical and TomoSAR point clouds, i.e. nadir-looking and side-looking, as well as the relatively large anisotropic noise in the TomoSAR point cloud. However, considering the large amount of points compared to the few co-registration parameters to be estimated, the co-registration accuracy is expected to be high enough for the following steps.

- Optical image classification: applying semantic classification to the optical images.

This part is not the focus of SARptical. Depending on the application, different classification algorithms can be applied.

- Semantic texturing: Texture the InSAR point cloud with the attributes derived from optical images, e.g. RGB color, semantic classification label, object bounding box, etc.

The main challenge of this step is to project the optical image to TomoSAR point cloud without explicit 3-D surface reconstruction in the TomoSAR point cloud. Therefore, we choose point-based rendering technique.

The main limitation of this step is the relatively poor positioning accuracy (1 to 10m) of spaceborne TomoSAR point cloud. This error will directly translate to the projection accuracy of the TomoSAR points in optical image.

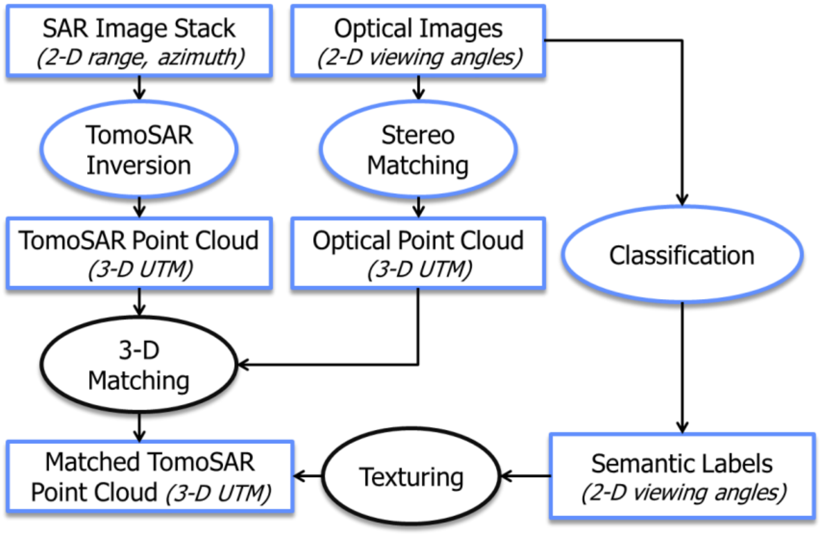

Figure 1. Flowchart of the proposed framework. It matches the 3-D TomoSAR and optical point clouds, and then transfers the attributes from optical images to the TomoSAR point cloud via a point-based texturing.

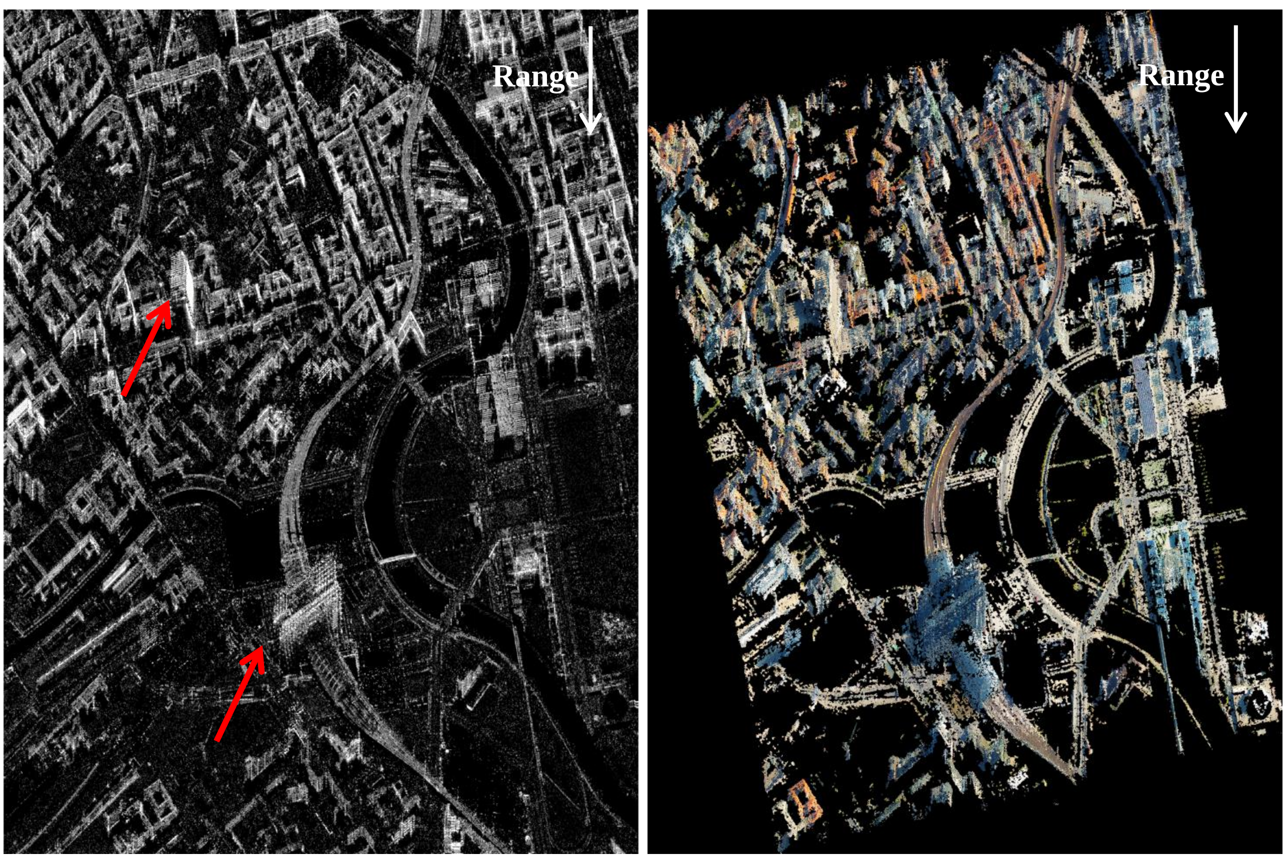

Figure 2. Left: the SAR amplitude image, and right: the optical image in SAR image geometry. Both images are in SAR azimuth and range coordinate. The red arrows indicate the locations of Berlin central station and the high-rise building of Universitätsmedizin Berlin.

Figure 2 is a direct demonstration of the SARptical system. It shows the SAR amplitude image and the optical image projected in SAR range-azimuth geometry. This is probably the first high resolution urban optical image projected in SAR geometry. The Berlin central station (complex structure) and the university hospital (high-rise building) are indicated by the red arrows. On the other hand, it is also possible to project the 2-D optical image to the 2-D SAR image geometry, which is not shown here due to space limit.

3. DEMONSTRATION OF APPLICATIONS

3.1 Automatic lamp poles detection for geodetic InSAR

Only until recently, it has been demonstrated that absolute localization with centimeter accuracy can be achieved for manually matched persistent scatterer (PS)s from TerraSAR-X images acquired from cross-heading geometries [3]. For automatically localizing a large network of such PSs for geodetic applications like the “Geodetic SAR Tomography” [4], we found that cylindrically vertical structures like lamp poles on the street are, most probably, the only natural ones visible in SAR images acquired from both ascending and descending orbits [5]. Detecting these PSs in SAR images can be extremely challenging, while it is much promising to achieve in optical images.

Thus, the demonstrated methodology includes the identification of lamp posts from high resolution optical data and project them into the cross-heading SAR images using the SARptical system. The precise absolute 3-D coordinates of the points are retrieved from the corrected TerraSAR-X timing measurements using the stereo SAR method [3]. Results for a test site in the city of Berlin acquired from TerraSAR-X high resolution spotlight mode will be demonstrated in the full paper.

3.2 Object-based multibaseline InSAR

Deformation monitoring by multi-baseline repeat-pass synthetic aperture radar (SAR) interferometry is so far the only imaging-based method to assess millimeter-level deformation over large areas from space. Past research mostly focused on the optimal deformation parameters retrieval on a pixel-basis. Only until recently, the first demonstration of object-based urban infrastructures monitoring by fusing InSAR and the semantic classification labels derived from optical images was presented in the SARptical system [1], [2], [6].

In this paper, we proposed a general framework, given such classification label in the SAR image, for object-based InSAR parameters retrieval where the estimation of the parameters is achieved in an object-level instead of pixel-wisely. Another new development presented in this paper is to introduce a robust phase recovery step in prior to the parameters inversion, in order to handle outliers in real data. The demonstrated method outperforms the current pixel-wised estimators, e.g. periodogram, by a factor of as much as dozens in the accuracy of the linear deformation estimates, at various situations such as signal-to-noise ratio, and outlier percentage. For practical demonstration, we presented a full workflow of long-term bridge monitoring using the proposed approach in the final paper.

3.3 SAR and optical images matching

The identification of similar image patches certainly is a frequently demanded task in remote sensing-related image analysis, especially in the framework of stereo applications. While many established feature-based approaches, specifically designed for the matching of optical images e.g. SIFT [7], already exist. To this date, the matching of images acquired by different sensors still remains an open challenge. This particularly holds for a joint exploitation of SAR and optical imagery. The challenge is caused by two completely different sensing modalities: optical imagery is acquired passively in a perspective projection at visible to inferred band (hundreds of THz), whereas SAR imagery is acquired actively in a cylindrical projection at microwave frequency (several GHz). Thus, particularly structures elevated above the ground level, such as buildings in urban areas, show strongly different appearances in both image types.

Thanks to the SARptical system, we have collected tens of thousands of matched SAR and optical images patches which can be used for exploitation the SAR and optical patch similarity. We demonstrate a convolutional neural network (CNN)-based approach, which allows to identify similar patches of very high resolution (VHR) optical and SAR imagery of complex urban scenes. The underlying similarity function is learnt directly from automatically generated training data and does not resort to any hand-crafted features. First evaluations show that the network provides an overall accuracy of more than 93% with a false alarm rate of 0%, thus indicating great potential for further development to a generalized multi-sensor matching procedure. We will show the example using real data in the final paper.

4. SUMMARY

SARptical is a novel concept that allows a pixel-level matching between high resolution SAR and optical images of dense urban areas. This is probably the first time that we see an optical image in dense urban area in a SAR geometry, and vice versa. We demonstrated the potentials of SARptical, including difficult target detection, object-based InSAR, and SAR optical image matching which were otherwise very challenging without the aid of SARptical.

REFERENCE

[1] Y. Wang and X. X. Zhu, “Fusing Meter-Resolution 4-D InSAR Point Clouds and Optical Images for Semantic Urban Infrastructure Monitoring,” *IEEE Trans. Geosci. Remote Sens.*, 2016.

[2] Y. Wang and X. X. Zhu, “InSAR Forensics: Tracing InSAR Scatterers in High Resolution Optical Image,” presented at the Fringe 2015, 2015.

[3] C. Gisinger *et al.*, “Precise Three-Dimensional Stereo Localization of Corner Reflectors and Persistent Scatterers With TerraSAR-X,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1782–1802, Apr. 2015.

[4] X. X. Zhu, S. Montazeri, C. Gisinger, R. F. Hanssen, and R. Bamler, “Geodetic SAR Tomography,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 18–35, 2015.

[5] S. Montazeri *et al.*, “SAR Ground Control Point Identification with the Aid of High Resolution Optical Data,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2016, Beijing, China*, 2016.

[6] Y. Wang and X. X. Zhu, “Semantic Fusion of SAR Interferometry and Optical Image with Application to Urban Infrastructure Monitoring,” presented at the La Grande Motte, France, La Grande Motte, France, 2015.

[7] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.