

OBJECT-BASED INSAR DEFORMATION RECONSTRUCTION WITH APPLICATION TO BRIDGE MONITORING

Jian Kang⁽¹⁾, Yuanyuan Wang⁽¹⁾, Marco Körner⁽²⁾, Xiao Xiang Zhu^(1,3)

(1) Signal Processing in Earth Observation, Technical University of Munich

(2) Chair of Remote Sensing Technology, Technical University of Munich

(3) Remote Sensing Technology Institute, Germany Aerospace Center

ABSTRACT

Deformation monitoring by multi-baseline 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 optimal deformation parameters retrieval on a pixel-basis. Only until recently, the first demonstration of object-based urban infrastructures monitoring by fusing SAR interferometry (InSAR) and the semantic classification labels derived from optical images was presented in [1]–[3]. This paper proposes an algorithm for object-based joint InSAR deformation reconstruction using these classification labels. We derive an object-based multi-baseline InSAR reconstruction model, and propose an efficient algorithm for bridge detection in optical images.

INTRODUCTION

Long-term millimeter-level deformation monitoring over large areas is so far only achievable through multi-baseline SAR interferometry (InSAR) techniques such as persistent scatterer interferometry (PSI) [4]–[7] and differential SAR tomography (TomoSAR) [8]–[10]. Through modelling the interferometric phase of the scatterers in the SAR image, we are able to reconstruct the 3-D position and the deformation history of each scatterer.

Recent research on multi-baseline InSAR is mainly focused on the optimum retrieval of the phase history parameters of individual pixel. Although distributed scatterer (DS)-based methods do employ information of multiple pixels, they exploit purely statistical information. SAR images, as images themselves, contain much geometrical information, e.g. the semantics in urban area. Such information can be integrated into the inversion process, and retrieve more reliable parameter estimates.

The aim of this paper is to derive an object-based multi-baseline InSAR inversion model given a spatial classification label. In this paper, the classification label is derived from optical images. As we primarily targeted at bridges monitoring, a bridge detection algorithm is also developed.

The overall workflow contains the following three main steps:

- Semantic classification in optical image
- Classification label projection to SAR image geometry
- Object-based joint deformation parameters reconstruction of the bridge areas

OBJECT-BASED JOINT INVERSION MODEL

Without losing generality, we derive the object-based joint inversion model based on the persistent scatterer (PS) model. However, this can be extended to other multi-baseline InSAR algorithms.

Given a PS with the spatial coordinate (x, y) , its time series observations can be modeled as follows:

$$\bar{\mathbf{g}}(x, y) = A \exp \left\{ -j \left(\frac{4\pi}{\lambda r} s(x, y) \mathbf{b} + \frac{4\pi}{\lambda} d(x, y) \mathbf{p} \right) \right\} \quad (1.1)$$

where $\bar{\mathbf{g}}(x, y)$ denotes the observation data given (x, y) , \mathbf{b} is the spatial baseline vector, $s(x, y)$ denotes the corresponding elevation as a function of the spatial coordinate, and $d(x, y)$ is the spatial deformation map under a certain motion model, e.g. $\mathbf{p} = \mathbf{t}$ for linear motion, $\mathbf{p} = \sin(2\pi(\mathbf{t} - t_0))$ for the seasonal motion model with the time baseline \mathbf{t} .

Given a set of PS pixels of a same cluster of semantic label, e.g. road or bridge, we can exploit the prior knowledge of their spatial deformation function, e.g. smoothness. In order to jointly estimate them, we solve the following optimization problem:

$$\begin{aligned} & \{ \hat{\mathbf{S}}(\mathbf{x}, \mathbf{y}), \hat{\mathbf{D}}(\mathbf{x}, \mathbf{y}) \} \\ & = \arg \min_{\mathbf{S}(\mathbf{x}, \mathbf{y}), \mathbf{D}(\mathbf{x}, \mathbf{y})} \frac{1}{2} \sum_i \left\| \mathbf{g}(x_i, y_i) - \bar{\mathbf{g}}(\mathbf{S}(x_i, y_i), \mathbf{D}(x_i, y_i)) \right\|_2^2 \\ & \quad + \lambda TV(\mathbf{D}(\mathbf{x}, \mathbf{y})) \end{aligned} \quad (1.2)$$

where i is the spatial index, $\mathbf{g}(x_i, y_i)$ is the PS observations, and $\mathbf{S}(\mathbf{x}, \mathbf{y}), \mathbf{D}(\mathbf{x}, \mathbf{y})$ are the 2-D images of elevation and deformation parameters, respectively. The first term in the minimization is the joint log-likelihood of the phase history parameters w.r.t. the observations, and second term is the anisotropic total variation (TV) of the 2-D image of the deformation parameters, with a regularization parameter λ .

Due to the complex exponential function in the likelihood, the objective function is nonconvex. The minimization method is achieved by the well-known bound constrained limited memory quasi Newton method – Broyden–Fletcher–Goldfarb–Shanno (LBFGSB).

BRIDGES DETECTION

As our targets to be monitored are bridges, this part describes a bridge extraction algorithm in optical images. One could also directly classify the SAR images. However, due to the speckle noise and layover phenomenon in urban areas, optical images are more suitable for this task. Moreover, other semantic classification labels, such as roads, façade can also be determined and the corresponding object-based deformation reconstruction can be obtained.

The proposed bridge extraction algorithm is based on the following prior knowledge of bridges:

1. Bridges are above rivers.
2. One bridge should completely cross and divide the river.

According to this, the proposed algorithm is developed as follows:

A. River segments extraction

As the first step, considering the homogenous texture of the rivers, we classify the river and the rest pixels using Local Binary Pattern (LBP) [11] as the feature. LBP descriptor is grayscale invariant and rotation invariant, which has recently been used in hierarchical segmentation of remote sensing images [12]. It is defined as follows:

$$\text{LBP}_{P,R} = \sum_{p=0}^{P-1} 2^p s_p, \text{ where } s_p = \begin{cases} 0, & i_p < i_c \\ 1, & i_p \geq i_c \end{cases} \quad (1.3)$$

Here, P is the total number of pixels in the circle with radius R centered at the target pixel. and i_p, i_c are the intensities of the respective pixels. Classification is done in a supervised manner, with some training samples and Supported Vector Machine (SVM) with linear kernel.

B. Precise retrieval of river segments contours

From the above method, a coarse detection of the river segments can be obtained. But it is not sufficient to precisely determine the discontinuity regions of the rivers (which are bridges). Given the initial river mask, we employed the Chan-Vese method [13] – an active contour algorithm – to accurately extract the river segments.

By using the level set formulation $C = \{(x, y) | \phi(x, y) = 0\}$ where C represents the closed curve, $\phi(x, y) = 0$ is the zero-level set function. Chan-Vese segmentation basically solves the following minimization function:

$$F(c_1, c_2, \phi) = \mu \int_{\Omega} \|\nabla H(\phi)\| + \int_{\Omega} (I(x, y) - c_1)^2 H(\phi) dx dy + \int_{\Omega} (I(x, y) - c_2)^2 (1 - H(\phi)) dx dy \quad (1.4)$$

where $H(\cdot)$ is the Heaviside function, c_1, c_2 are the two unknown parameters, Ω denotes the domain to be segmented.

C. Bridge extraction

According to the accurate river segments obtained by Chan-Vese algorithm, bridge area detection can be achieved by extracting discontinuities between the river segments. We apply morphological closing to accomplish this task [1].

EXPERIMENT

A. Bridge detection in optical image

The supervised classification of the rivers requires some training patches. We manually selected a dozen of training patches. The parameters in the LBP descriptor are set to be $P=8$ and $R=1$. Due to the time consuming work of manually labeling new training samples in different dataset every time, the classification can also be achieved by applying domain adaptation with some training samples from online land-use dataset. The rivers are classified by an SVM in a sliding-window manner. Outliers appeared as small clusters are later removed by spatial connectivity analysis.

For the refinement of the classified river segments, the iteration number of Chan-Vese segmentation is set to be 150. This parameter determines the final area of segmentation, larger number induces overflow of segmentation, otherwise, insufficient segmentation. As illustrated in Figure 1, blue area represents the river and bridges are covered by green. Due to the shadow phenomenon, some parts of river banks are also regarded as rivers. However, five main bridges in the scene are efficiently detected.

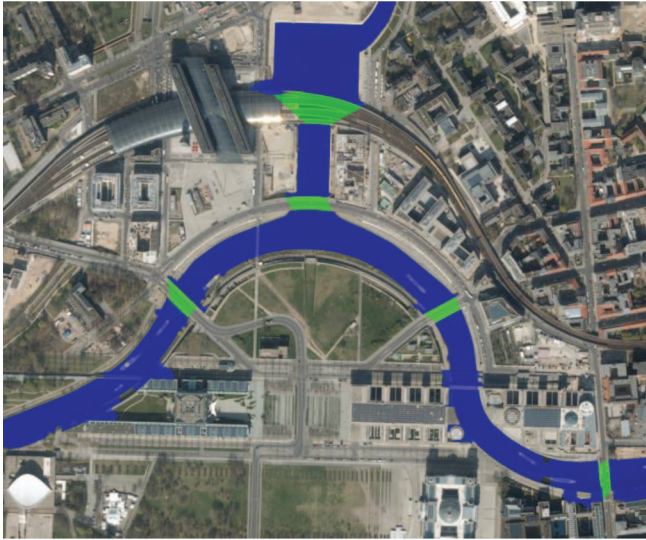


Figure 1 Bridge and River classification in optical image (Green area represents bridges and River areas are covered by blue)

B. Joint deformation reconstruction

Simulations

We simulated a multibaseline SAR image stack with the spatial deformation pattern as shown in Figure 2. The linear deformation rate ranges from 1 mm/year to 2.5 mm/year. We choose a spatial baseline comparable to TerraSAR-X and a time baseline with regular spacing. The number of images is 20, with $20 \times 20 = 400$ pixels in each image. Uncorrelated complex circular Gaussian noise was added to the simulated SAR image stacks with a signal-to-noise ratio (SNR) of 5 dB. The penalty parameter λ was set to be 350.

Figure 2 compares the linear deformation rate estimated pixel-wisely and jointly, where the left subfigure is the simulated ground truth. It shows that, without considering the smoothness along the spatial direction, the single-pixel periodogram result is much noisier comparing to the result of joint reconstruction. For a quantitative comparison, the standard deviation (SD) of the result of from the two methods with SNR from 0dB to 10dB are plotted in Figure 3. The proposed method is particularly effective for low SNR areas, where improves the accuracy by more than 10-25 times than single-pixel reconstruction.

Real data

We utilize TerraSAR-X dataset of Berlin area to be our test data. The whole data stack composes 109 images. The bridge labels projected to the SAR image is shown in Figure 4(a). The area used for joint deformation reconstruction is determined by the largest inscribed rectangle in the bridge label. The result of joint reconstruction with 1/3 number (37) of SAR images is illustrated in the right side of Figure 4.(b),

together with the results of single-pixel with the whole 109 images (Left) and 37 images (Middle) for comparison. A zoom in to the results cropped by black rectangular areas are shown in (c)

As shown from the zoom in areas, with the limited number (37) of images, single-pixel based Periodogram is too noisy to demonstrate the real pattern of spatial deformation. However, the proposed approach can get rid of wrongly estimated pixels in the deformation map induced by single pixel based Periodogram and obtain more reliable result even reconstructed by limited images.

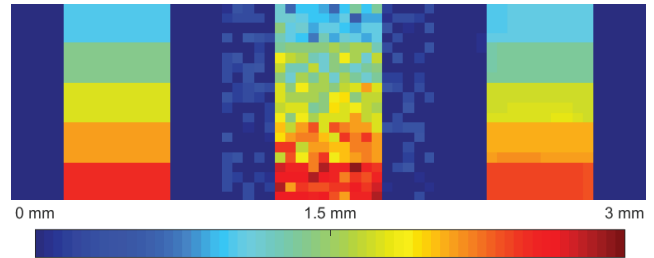


Figure 2 (Left) Ground truth linear deformation rate map, (Middle) Deformation map estimated by single-pixel based Periodogram, and (Right) Deformation map obtained by Joint Deformation Reconstruction

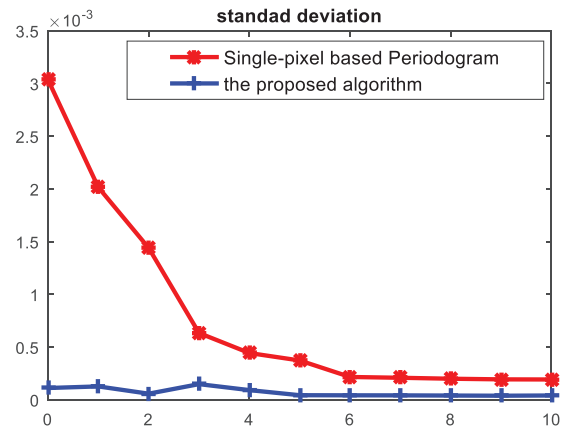
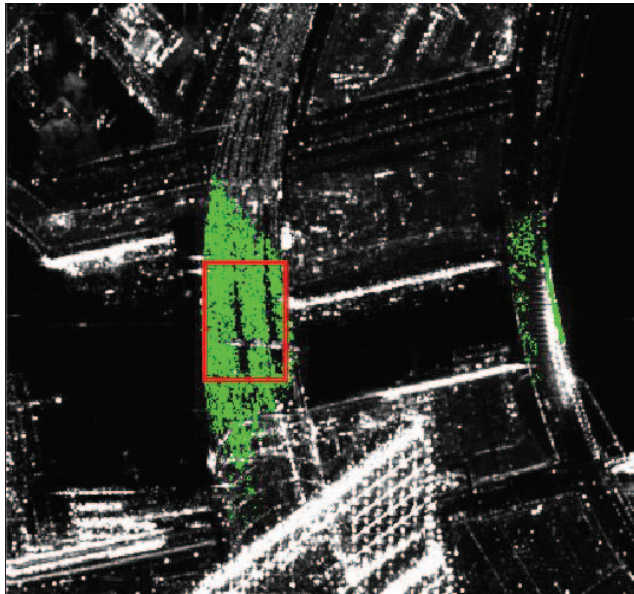
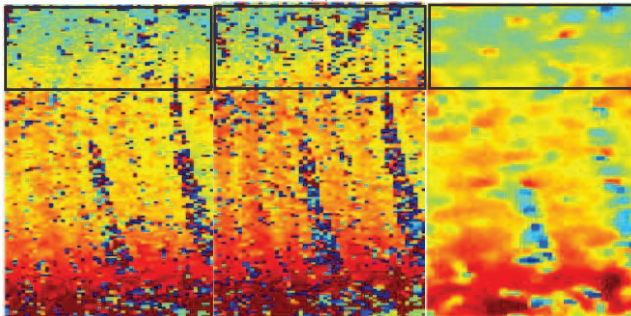


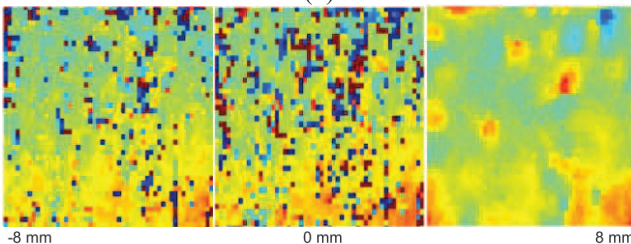
Figure 3 Standard Deviation value of two methods with respect to SNR ranges from 0dB to 10dB



(a)



(b)



(c)

Figure 4 (a) Bridge semantic label in SAR image (b) Deformation result of Single pixel based Periodogram with the whole data stack (109 images)(Middle), Deformation result of Single pixel based Periodogram with the half data stack (37 images)and (Right) Joint Deformation Reconstruction with the half data stack (37 images)

CONCLUSION

This paper proposed a new deformation retrieval framework – object-level joint deformation reconstruction based on the work of InSAR semantic interpretation. We demonstrated that the smoothness prior of the spatial deformation signal can be exploited by adding a total variation penalty term in

the joint phase parameter inversion. Real data experiments have illustrated the effectiveness of the proposed method, especially the improvement with the limited number of SAR images.

In the future, we seek to improve other InSAR techniques guided by the semantic labeling as well as different prior knowledge according to the objects.

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] 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.
- [4] N. Adam, B. Kampes, M. Eineder, J. Worawattanamateekul, and M. Kircher, "The development of a scientific permanent scatterer system," in *ISPRS Workshop High Resolution Mapping from Space, Hannover, Germany, 2003*, vol. 2003, p. 6.
- [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] B. M. Kampes, *Radar Interferometry - Persistent Scatterer Technique*, vol. 12. Dordrecht, The Netherlands: Springer, 2006.
- [8] G. Fornaro, S. Verde, D. Reale, and A. Pauciuolo, "CAESAR: An Approach Based on Covariance Matrix Decomposition to Improve Multibaseline-Multitemporal Interferometric SAR Processing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2050–2065, Apr. 2015.
- [9] 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.
- [10] 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.
- [11] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Anal. Mach. Intell. IEEE Trans. On*, vol. 24, no. 7, pp. 971–987, 2002.
- [12] Z. Hu, Z. Wu, Q. Zhang, Q. Fan, and J. Xu, "A spatially-constrained color-texture model for hierarchical VHR image segmentation," *Geosci. Remote Sens. Lett. IEEE*, vol. 10, no. 1, pp. 120–124, 2013.
- [13] T. F. Chan, L. Vese, and others, "Active contours without edges," *Image Process. IEEE Trans. On*, vol. 10, no. 2, pp. 266–277, 2001.