SPATIOTEMPORAL MINING OF ENVISAT SAR INTERFEROGRAM TIME SERIES OVER THE HAIYUAN FAULT IN CHINA

N. Méger[1], *R. Jolivet*[2], *C. Lasserre*[2], *E. Trouvé*[1], *C. Rigotti*[3], *F. Lodge*[2], *M-P. Doin*[4], *S. Guillaso*[5], *A. Julea*[1,6,7] *and Ph. Bolon*[1]

[1]Université de Savoie/Polytech Annecy-Chambéry, LISTIC Laboratory. B.P. 80439, F-74944 Annecy-le-Vieux Cedex, France
{andreea.julea|nicolas.meger|emmanuel.trouve|philippe.bolon}@univ-savoie.fr

[2] Université Joseph Fourier, IsTerre Laboratory, CNRS - UMR 5559. B.P. 53, F-38041 Grenoble Cedex 09, France
{romain.jolivet|felicity.lodge|cecile.lasserre}@ujf-grenoble.fr

[3]Université de Lyon, CNRS, INRIA, INSA-Lyon, LIRIS, UMR5205, F-69621, France
christophe.rigotti@insa-lyon.fr

[4] Ecole Normale Supérieure, Laboratoire de Géologie, CNRS - UMR 8538. 24 rue Lhomond, F-75231 Paris Cedex 05, France
doin@geologie.ens.fr

[5] Technische Universität Berlin, Computer Vision and Remote Sensing, Franklinstraße 28/29, D-10587 Berlin, Germany
stephane.guillaso@tu-berlin.de

[6]Politehnica University of Bucharest/Faculty for Electronics, Telecommunications and Information Technology. Bucharest, Romania.

[7]Institutul de Stiinte Spatiale. Bucharest, Romania

ABSTRACT

In this paper, an original approach for analyzing InSAR time series is presented. The interferograms forming such time series allow ground deformation occurring between acquisition dates to be measured with high precision. Nevertheless, they can be affected by variations in atmospheric conditions. The proposed approach is designed to handle these varying atmospheric conditions. The stratified atmosphere is first removed and the phase evolution is built using a Small BAseline Subsets (SBAS) strategy. Then, *frequent grouped sequential patterns* are extracted. These patterns allow InSAR time series to be described spatially and temporally while discarding atmospheric perturbations. Experimental results on an ENVISAT InSAR time series covering the Haiyuan fault in the north-eastern boundary of the Tibetan plateau are presented.

***Index Terms*—** InSAR, time series, data mining, frequent grouped sequential patterns, SBAS

1. INTRODUCTION

Repeat pass SAR Interferometry (InSAR) measures phase differences between SAR images acquired at different dates. The resulting interferograms allow ground deformation occurring between the acquisition dates to be measured over large areas with a high precision: a fraction of the wavelength (5.6 cm in C band). The two main limitations of this technique are surface changes, which reduce the coherence between SAR images, and variations in atmospheric conditions between the different acquisitions. The loss of phase information is easy to detect by thresholding coherence images whereas the atmospheric perturbations are difficult to discriminate from the displacement signal. The contribution due to the stratified atmosphere can be roughly estimated by using digital elevation models (DEM) and meteorological data, but the effects of the turbulent atmosphere still degrade the interferograms. Different approaches have been developed to reduce these difficulties by using interferogram time series: the permanent scatterer (PS) technique [1] which analyses the temporal signal on specific targets, the Small BAseline Subsets (SBAS) strategy [2] which selects the most reliable pairs according to temporal and spatial baselines or the STAMPS method [3] which incorporates both approaches. They are used by geophysicists to monitor ground surface deformations and their temporal evolutions.

In this paper, the data are first prepared by removing contributions from the variations of the stratified troposphere [4]. An SBAS-based technique is then applied to compute the phase evolution throughout time. This data preparation step is described in Section 2. The application of such a method to large Satellite Image Time Series (SITS) creates huge amounts of data to be analyzed. Furthermore, these data

This work was supported by the French Research Agency (ANR) through the EFIDIR project (ANR-2007-MCDC0-04, http://www.efidir.fr) and the FOSTER project (ANR-2010-COSI-012-04, http://foster.univ-nc.nc/). The authors wish to thank the European Space Agency (ESA) for the ENVISAT SAR data (Dragon project Dragon ID5305) over the Haiyuan fault.

can still be affected by the turbulent atmosphere. Powerful techniques are required to extract the relevant spatiotemporal information carried by InSAR images. A recent data mining technique [5] allowing the characterization of SITS spatially and temporally is used. It extracts so-called *frequent grouped sequential patterns*. As reported in [5], this technique can discard atmospheric perturbations while extracting ground deformation. These patterns are presented in Section 3. Finally, experimental results on a ENVISAT InSAR time series covering the Haiyuan fault in the north-eastern boundary of the Tibetan plateau are described in Section 4.

2. DATA PREPARATION

The data are first prepared using a refined version of the SBAS approach [6]. The goal of this InSAR multitemporal processing stage is to derive displacement measurements by taking data characteristics into account without introducing any specific user knowledge to the studied ground deformation. The processing chain includes the following steps:

1. SAR image coregistration to a single master and interferogram generation with local adaptive range filtering,
2. joint inversion of residual orbital and atmospheric delays,
3. validation of atmospheric corrections using the ERA40 global atmospheric model (ECMWF) and correction of each unwrapped interferogram [7],
4. inversion of the interferogram series on a given track to obtain the increments of LOS radar delays between acquisition dates, adapting the SBAS time series analysis technique [8, 9]. As no smoothing is applied, both ground deformation and atmospheric turbulence contribute to the phase evolution.

This processing chain has been applied to an ENVISAT InSAR time series covering the Haiyuan fault in the north-eastern boundary of the Tibetan plateau. This areas was affected by several major earthquakes in the early 20th century. The InSAR technique is applied to locate and measure possible continuous crustal deformations. A set of 24 raw SAR images from an ascending track, acquired over the 2003-2009 period, has been used to generate 130 interferograms [10]. The 24 Line of Sight (LOS) displacement images resulting from the temporal inversion are illustrated on 3 different dates in Figure 1.

3. FGS-PATTERNS

Once the data have been prepared, they can be mined to uncover the relevant spatiotemporal structures that are present in InSAR time series. This data mining stage is an automatic screening technique which explores the huge amount of measurements which are still affected by noise, processing and atmospheric artefacts. The goal is to extract FGS-patterns and to present them to the end-user to draw attention to possible unknown displacements or displacement evolutions.

This approach relies on evolution and sub-evolution extraction at the pixel level. A pixel evolution or sub-evolution is described using a sequential pattern, denoted $A_1 \to A_2 \to \ldots \to A_n$, where $A_1, A_2, \ldots, A_n$ are symbols representing discrete pixel states at n different dates which are not necessarily consecutive. These pixel evolutions and sub-evolutions are used to find, in an unsupervised way, groups of pixels that could be of interest to end-users. In order to output pixel sets making sense spatially and temporally, sets having at least σ pixels (i.e. a minimum surface) sharing the same temporal evolution α are selected. Pixels sharing α are said to *be covered* by α and are denoted $cov(\alpha)$. The size $|cov(\alpha)|$ is called the *support* of α and is denoted $support(\alpha)$. The pixels covered by a pattern are also required to exceed a minimum connectivity threshold κ. The connectivity measure used is called the *average connectivity*. It gives, for the pixels sharing α, the average number of neighboring pixels also sharing α. The 8 nearest neighbors (8-NN) are taken into consideration. Let us consider a *local connectivity function* $LC((x, y), \alpha)$ that returns, for a pixel (x, y), the number of neighbors covered by α. The average connectivity of α is then defined as follows: $AC(\alpha) = \frac{\sum_{(x,y)\in cov(\alpha)} LC((x,y),\alpha)}{|cov(\alpha)|}$. Formally, an evolution (or sub-evolution) α is thus retained if $|cov(\alpha)| \geq \sigma$ and if $AC(\alpha) \geq \kappa$. In this case, it is called a *Frequent Grouped Sequential Pattern (FGS-pattern)*. These constraints can be actively pushed into data mining algorithms (e.g., [11, 12]) to prune the search space and to make extractions tractable. These operative aspects are detailed in [13]. Our prototype, based on the *PrefixGrowth* [12] algorithm, has been written in C using our own data structures.

4. INSAR TIME SERIES MINING: RESULTS

The 24 Line of Sight (LOS) displacement images (781×501 pixels) resulting from the temporal inversion (see Section 1) have been mined to extract FGS-patterns (see Section 3. Negative values correspond to motion away from the satellite, along the Line-Of-Sight. The whole displacement time series was considered and the input values were strongly quantified using 3 symbols ('1', '2' and '3') by using the 33rd and the 66th centiles. Symbol '1' represents large negative values, symbol '2' corresponds to low negative values and symbol '3' is linked to positive values. The FGS-pattern extraction was run on a standard laptop (Intel Core i5 CPU M 520 @ 2.40 GHz, 4 GB of RAM, Linux 2.6.37 kernel) by setting the minimum surface threshold σ to 100000 pixels and the average connectivity threshold κ to 6. Within less than 1 hour and 8 minutes, using 2 GB of RAM, 3413 FGS-patterns were extracted along with their respective surface and average connectivity measurements. In order to focus on the most specific patterns, the longest FGS-patterns were selected. This leads to 18 patterns with ten symbols and one pattern having eleven symbols. For each longest FGS-pattern, an image was built to observe the pixels that are affected by the evolution given by the pattern. Color coding was used to indicate the date at which the evolution has been fully observed. The dates of the time series are linearly linked to a color palette which can be found in 2. The red colors relate to the early dates while the violet colors correspond to late dates.

The results obtained for 2 different FGS-patterns, 1,1,1,1,-1,1,1,1,1,1 and 3,3,3,3,3,3,3,3,3,3, are illustrated in Figure 2. In the near field, creep is revealed by both patterns, coherent with left-lateral motion. For both patterns, rainbows can be observed: these pixels evolutions thus propagate throughout space and time. The propagation that is observed for pattern 1,1,1,1,1,1,1,1,1,1 is not radial w.r.t. to the creeping zone, and creep migration along the fault could explain such patterns.

5. CONCLUSION

These results illustrate the potential of a combined use of advanced InSAR multitemporal processing for deriving displacement time series and a data mining technique for extracting, in an unsupervised way, spatiotemporal features corresponding to slow ground deformations. The refined SBAS approach is intended to reduce, as far as possible, systemic uncertainty, whereas the data mining technique offers end-users the possibility of exploring huge time series and discovering temporal evolutions which could be hidden by random uncertainty.

6. REFERENCES

[1] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 8–20, 2001.

[2] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, pp. 2375–2382, 2002.

[3] A. Hooper, "A multi-temporal InSAR method incorporating both persistent scatterer and small baseline approaches," *Geophysical Research Letters*, vol. 35, 2008.

[4] O. Cavalié, M.-P. Doin, C. Lasserre, and P. Briole, "Ground motion measurement in the Lake Mead area, Nevada, by differential synthetic aperture radar interferometry time series analysis: Probing the lithosphere rheological structure," *Journal of Geophysical Research*, vol. 112, Mar. 2007.

[5] A. Julea, N. Méger, P. Bolon, C. Rigotti, , M.-P. Doin, C. Lasserre, E. Trouvé, and V. Lăzărescu, "Unsupervised spatiotemporal mining of satellite image time series using grouped frequent sequential patterns," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, pp. 1417–1430, 2011.

[6] P. Lopez-Quiroz, M.-P. Doin, F. Tupin, P. Briole, and J.-M. Nicolas, "Time series analysis of Mexico City subsidence constrained by radar interferometry," *Journal of Applied Geophysics*, 2009.

[7] M.-P. Doin, C. Lasserre, G. Peltzer, O. Cavalié, and C. Doubre, "Corrections of stratified tropospheric delays in SAR interferometry: validation with global atmospheric models," *Journal of Applied Geophysics*, vol. 69, no. 1, pp. 35–50, 2009.

[8] M.-P. Doin S. Guillaso G. Peltzer R. Dailu R. Jolivet, C. Lasserre and J. Sun, "Shallow creep on the haiyuan fault, gansu, china, revealed by sar interferometry," *International Journal in Remote Sensing*, vol. submitted, 2011.

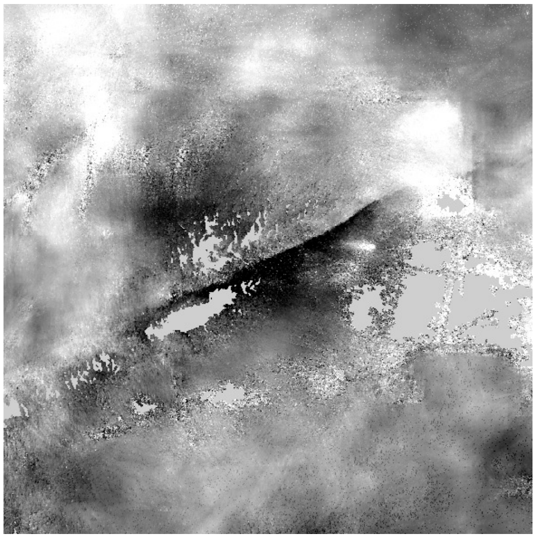
[9] M.-P. Doin G. Peltzer R. Dailu R. Jolivet, C. Lasserre and J. Sun, "Temporal variations of the creep rate along a millenial seismic gap on the haiyuan fault, china,," *Geophysical Journal International*, vol. submitted, 2011.

[10] R. Jolivet, C. Lasserre, N. Lin, M. Simons, M.-D. Doin, E. A. Hetland, P. Muse, G. Peltzer, S. Jianbao, and R. Dailu, "Time series and MinTS analysis of strain accumulation along the Haiyuan fault (Gansu, China) over the 2003-2010 period, from ENVISAT InSAR data," in *American Geophysical Union (AGU) Fall Meeting*, San Francisco, California, USA, Dec. 2010.

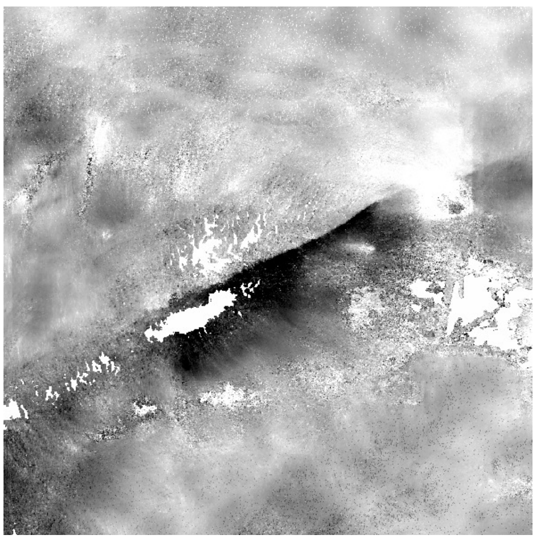
[11] R. Agrawal and R. Srikant, "Mining sequential patterns," in *Proc. of the 11th International Conference on Data Engineering (ICDE'95)*, Philip S. Yu and Arbee S. P. Chen, Eds., Taipei, Taiwan, 1995, pp. 3–14, IEEE Computer Society Press.

[12] J. Pei, J. Han, and W. Wang, "Constraint-based sequential pattern mining: the pattern-growth methods," *Journal of Intelligent Information Systems*, vol. 28, no. 2, pp. 133–160, 2007.

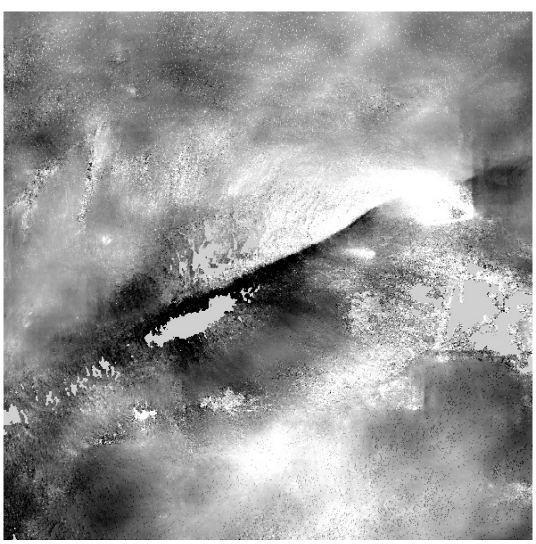
[13] A. Julea, N. Meger, C. Rigotti, M.-P. Doin, C. Lasserre, E. Trouvé, P. Bolon, and V. Lăzărescu, "Extraction of frequent grouped sequential patterns from satellite image time series," in *Proc. of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2010)*, Honolulu, Hawaii, USA, 2010, pp. 3434–3437.



(a) 2006/09/21 (first acquisition date, radar geometry)

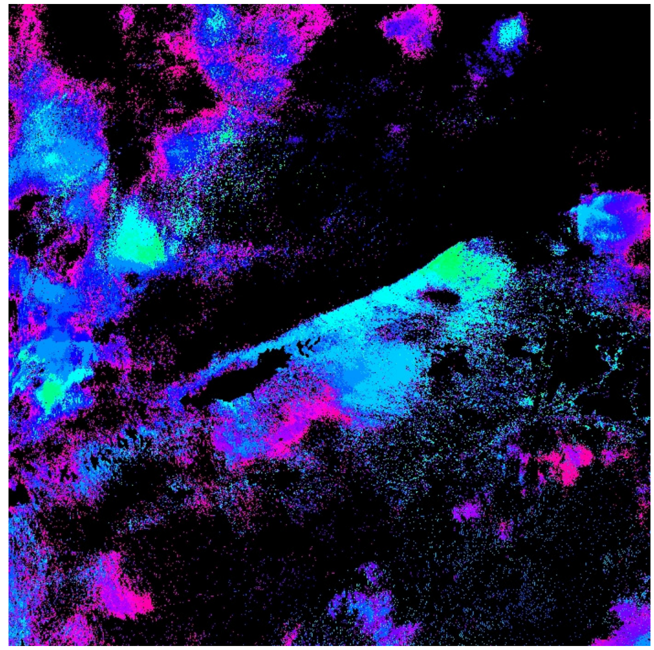


(b) 2008/02/28

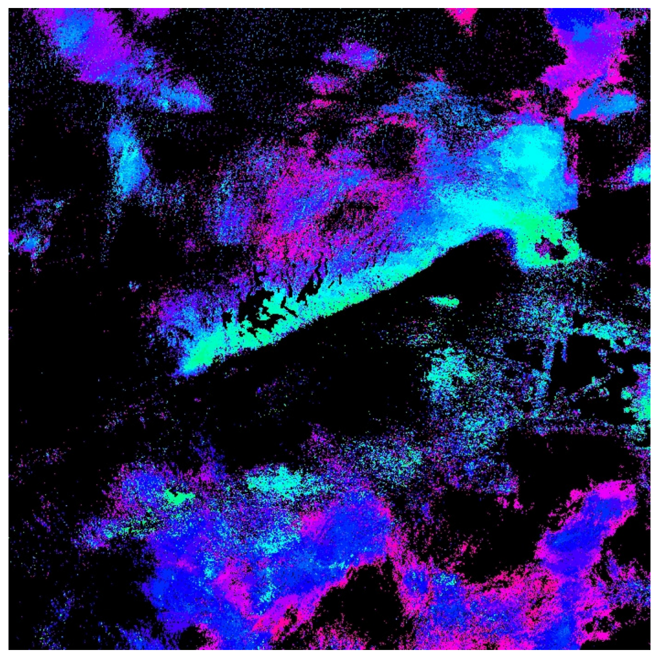


(c) 2009/08/06 (last acquistion date)

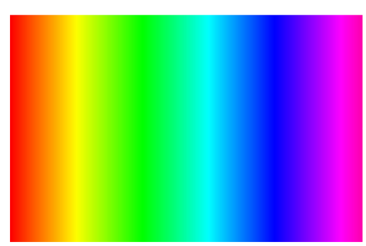
Fig. 1. LOS displacement at 3 different dates.



(a) FGS-pattern 1,1,1,1,1,1,1,1,1,1



(b) FGS-pattern 3,3,3,3,3,3,3,3,3,3



(c) Color palette

Fig. 2. Spatiotemporal localization of 2 FGS-patterns along with the used color palette.