

FrFT-Based Scene Classification of Phase-Gradient InSAR Images and Effective Baseline Dependence

Nazli Deniz Cagatay, *Student Member, IEEE*, and Mihai Datcu, *Fellow, IEEE*

Abstract—In the literature, scene recognition from interferometric synthetic aperture radar (InSAR) images has been mainly focused on the joint use of the backscatter intensity and the coherence between interferometric image pairs. However, the terrain height information residing in the interferometric phase requires further exploration for classification purposes. In this letter, taking the interferometric phase information into account together with the backscatter intensity, the whole complex-valued InSAR image is exploited for feature extraction. In addition, a new complex-valued phase-gradient InSAR (PGInSAR) image is defined. A fractional-Fourier-transform-based feature extraction, which was proposed for the classification of single-look complex (SLC) SAR images, is adopted for InSAR and PGInSAR images. For patch-based classification, an image database is generated from bistatic pairs acquired from the same terrain with three different effective baselines. The supervised K -nearest neighbor classification results show that InSAR outperforms SLC by 15%, whereas PGInSAR introduces further 10% improvement over InSAR or a total improvement of 27% over SLC. Moreover, PGInSAR is found to be more robust to effective baseline changes than InSAR, which makes PGInSAR a better candidate for feature extraction.

Index Terms—Effective baseline, feature extraction, fractional Fourier transform (FrFT), interferometric synthetic aperture radar (InSAR), phase gradient.

I. INTRODUCTION

AUTOMATED recognition and classification of remote sensing images is of great importance due to the huge volume of available data. Among various imaging technologies, the weather- and daylight-independent sensing capability of synthetic aperture radar (SAR) makes it a special candidate for scene recognition. The inclusion of a second SAR image acquired from a slightly different position and/or at a different time, as in SAR interferometry, provides some complementary information about the scene, such as the interferometric coherence, the terrain height, etc. Such information can be expected to enhance the recognition performance. Thus, in addition to being in use for the derivation of the surface topography and the observation of temporal surface change, SAR interferometry

has recently found a new application area in scene recognition and classification.

In literature, the research on the classification of InSAR images has mostly focused on the use of interferometric coherence. In particular, the coherence of interferometric pairs, together with the backscatter intensity, was found to increase the accuracy of land-cover classification [1]–[4]. Furthermore, multitemporal InSAR images have been widely utilized for classification based on the temporal variation of the coherence and the backscatter intensity [1], [5].

On the other hand, the terrain height information residing in the interferometric phase requires further exploration [6]. There are several studies that take the interferometric phase information into account for boundary detection [6]–[8] and for damage detection based on height changes [9]. Moreover, the phase-gradient approach is studied for boundary detection [6] and for the motion detection of glaciers in differential InSAR [10].

However, although some amount of research related to the interferometric phase has been reported, not much has been done to explore its potential in classification. Considering the fact that man-made and/or natural land-cover scatterers may differ in height (the interferometric phase) and backscattering properties (the amplitude), the whole complex-valued InSAR image was employed in our previous work to extract feature descriptors for classification purposes [11]. In addition, based on the fact that the gradient of the interferometric phase follows how the surface topography changes throughout the scene [12], a new complex-valued phase-gradient InSAR (PGInSAR) image was defined and employed in feature extraction.

As a feature descriptor, first-order statistics were extracted from the fractional Fourier transform (FrFT) of the images. This method had been already proposed and found to be quite successful in the classification of single-look complex (SLC) SAR images [13]. In [11], by slightly modifying this method for InSAR and PGInSAR images, a comparative study was performed for SLC, InSAR, and PGInSAR, and the superiority of InSAR and PGInSAR over SLC was presented.

In this letter, in order to further investigate the performance of the proposed method in patch-based image classification for different InSAR geometries, a database is generated with three different InSAR acquisitions from the same site, particularly with different effective baselines.

The rest of this letter is organized as follows. In Section II, the relation between the terrain height and the interferometric phase is explained through the InSAR acquisition geometry. Section III introduces the proposed phase-gradient approach. The FrFT-based feature extraction is presented in Section IV. The image database and the classification results are given in Sections V and VI, respectively. Finally, the results are concluded in Section VII.

Manuscript received June 1, 2014; revised October 25, 2014; accepted December 3, 2014. The work of N. D. Cagatay was supported by the German Academic Exchange Service (DAAD).

The authors are with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany (e-mail: Nazli.Kahyaoglu@dlr.de; Mihai.Datcu@dlr.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2014.2385771

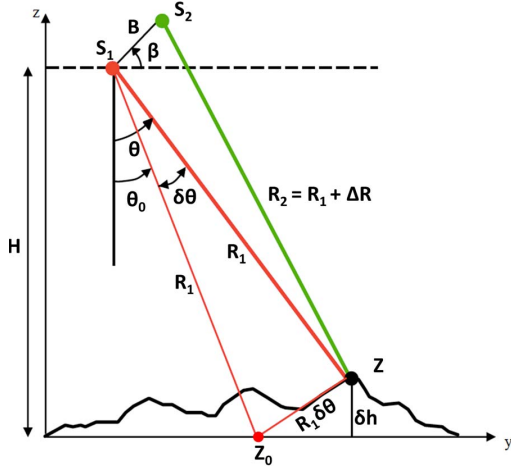


Fig. 1. InSAR geometry.

II. INSAR IMAGING GEOMETRY

A 2-D SLC SAR image is represented as $z = |z| \cdot e^{j\psi}$, with its amplitude $|z|$ and phase ψ . The amplitude depends on the transmitter power, the distances of the scatterer to the transmitter and the receiver, and the scene reflectivity. On the other hand, the phase depends on the total distance that the signal travels before being received by the receiver antenna. Hence, the phase is mainly composed of two components: 1) the geometrical phase due to the total distances between transmitter–scatterer and scatterer–receiver (a deterministic component); and 2) the object’s phase due to the coherent sum of many individual scatterers, i.e., the scattering phenomenon on the ground (a random component).

In SAR interferometry, the backscattered echo from the same region is received from slightly different positions S_1 and S_2 , as in Fig. 1, and an interferometric image pair (z_1, z_2) is obtained. For the case where S_1 transmits and both S_1 and S_2 receive, the phases of the two SAR images will be

$$\begin{aligned}\psi_1 &= -\frac{2\pi}{\lambda}(2R_1) - \psi_{\text{obj}_1} \\ \psi_2 &= -\frac{2\pi}{\lambda}(R_1 + R_2) - \psi_{\text{obj}_2}.\end{aligned}\quad (1)$$

Since the random object’s phase components may be assumed to be identical for these two images, i.e., $\psi_{\text{obj}_1} \approx \psi_{\text{obj}_2}$, interferometric phase ψ_{int} , which is defined to be the phase difference between the two images, reflects the difference in the sensor–scatterer distances of the two images, i.e., $\Delta R = R_2 - R_1$. Unlike the random phase of one SLC image, the interferometric phase carries particular information about the acquisition geometry as follows:

$$\psi_{\text{int}} = \psi_1 - \psi_2 = \frac{2\pi}{\lambda}\Delta R \approx -\frac{2\pi}{\lambda}B \sin(\theta - \beta). \quad (2)$$

Due to a nonzero baseline, even a point on the flat Earth yields a nonzero interferometric phase. Then, a point at height δh has an interferometric phase of

$$\psi_{\text{int}} = -\frac{2\pi}{\lambda}B \sin(\theta_0 - \beta) - \frac{2\pi}{\lambda}B \cos(\theta_0 - \beta)\delta\theta \quad (3)$$

where the first term corresponds to the interferometric phase measured in the absence of any topography, and the second term

corresponds to the flattened interferometric phase that reflects the height variations in the scene relative to the flat Earth. Once an interferogram is generated, it is flattened by removing the first term, and ψ_{flat} is obtained.

Then, together with the knowledge of InSAR parameters, such as baseline B , look angle θ_0 , and signal wavelength λ , a change in the flattened interferometric phase corresponds to a change in the terrain height according to the following relation:

$$\delta h = -\delta\psi_{\text{flat}} \frac{\lambda R_1 \sin \theta_0}{2\pi B_{\perp}} \quad (4)$$

where the *effective baseline* $B_{\perp} = B \cos(\theta_0 - \beta)$ is the projection of the interferometric baseline onto the direction perpendicular to the line of sight. The height change that leads to a 2π change in the flattened interferometric phase is called the *ambiguity height*, i.e.,

$$h_{\text{amb}} = \frac{\lambda R_1 \sin \theta_0}{B_{\perp}} \quad (5)$$

and it can be considered a measure of sensitivity to the topography. From (5), a longer effective baseline yields a smaller ambiguity height, which means better height sensitivity. However, for long baselines, the spatial decorrelation between the two images should be also considered. As a measure of the similarity between the two images, the *complex correlation* is defined as follows:

$$\rho = \frac{E[z_1 \cdot z_2^*]}{\sqrt{E[|z_1|^2] E[|z_2|^2]}} = |\rho_c| e^{j\phi_c} \quad (6)$$

and its magnitude $|\rho_c|$ is called the *coherence*.

III. PGINSAR IMAGE

In this letter, the complex-valued flattened InSAR image is used for feature extraction, as previously proposed in [11].

The amplitude of the InSAR image is slightly modified such that the geometric mean of the amplitudes of the two images is used instead of their multiplication as

$$I = \sqrt{|z_1| \cdot |z_2|} \cdot e^{j\psi_{\text{flat}}}. \quad (7)$$

On the other hand, the interferometric phase suffers from being wrapped over principal interval $[-\pi, \pi)$. Hence, the observed phase contains 2π phase jumps and has to be unwrapped to retrieve the original phase. However, since the whole complex-valued images are used in this letter for feature extraction, not their phases, the challenging phase unwrapping operation is avoided. In addition, the idea of an image gradient is adopted for the interferometric phase. Taking the gradient magnitude of the interferometric phase, the number of phase discontinuities is reduced.

Considering the modifications in the InSAR amplitude and phase, a new complex-valued *PGInSAR* is defined as

$$\text{PG} \equiv \sqrt{|z_1| \cdot |z_2|} \cdot e^{j|\nabla\psi_{\text{flat}}|} \quad (8)$$

where $|\nabla\psi_{\text{flat}}|$ is the magnitude of the interferometric phase gradient in the range and azimuth directions. As it is, the PGInSAR image phase can be considered to express how fast the terrain height changes throughout a scene, whereas the InSAR image phase is related to the absolute terrain height [11].

As in [14], $\nabla\psi_{\text{flat}}$ is computed from the real (I_{real}) and imaginary (I_{imag}) parts of the flattened interferogram as

$$\nabla\psi_{\text{flat}} = \frac{I_{\text{real}}\nabla I_{\text{imag}} - I_{\text{imag}}\nabla I_{\text{real}}}{I_{\text{real}}^2 + I_{\text{imag}}^2} \quad (9)$$

by means of the relation $\psi_{\text{flat}} = \tan^{-1}\left(\frac{I_{\text{imag}}}{I_{\text{real}}}\right)$ and the chain rule in the differentiation. Unlike the computation of the gradient directly from the wrapped phase ψ_{flat} , the real and imaginary parts, which are usually continuous, yield a phase gradient with less jumps [11].

IV. FEATURE EXTRACTION BASED ON FrFT

In this letter, a nonparametric feature extraction method previously proposed in [13] for the classification of SLC images is adopted to InSAR and PGInSAR images [11]. In that method, the feature descriptor is based on the simple statistics extracted from the FrFT of the SLC image. The FrFT of a function $f(\xi)$ is defined as

$$\begin{aligned} \text{FrFT}_\alpha(\xi) &= A_\alpha \cdot \exp(j\pi\xi^2 \cot \alpha) \\ &\quad \cdot \int \exp[j\pi(-2x\xi \csc \alpha + x^2 \cot \alpha)] f(x) dx \\ A_\alpha &= \frac{\exp[-j(\pi \operatorname{sgn}(\sin \alpha)/4 - \alpha/2)]}{|\sin \alpha|^{1/2}} \end{aligned} \quad (10)$$

where $\alpha = p\pi/2$ is the transform angle, and $0 < |p| < 2$ is the transform order of the FrFT. If $\alpha = 0$ or a multiple of 2π , the FrFT corresponds to the identity operator, and for $\alpha = \pi/2$, the FrFT takes the standard Fourier transform [15]. For 2-D images, the transform can be first applied to columns and then to rows, or vice versa, due to the linearity of the FrFT.

Due to the chirplike nature of SLC and InSAR images, a chirplet-based FrFT is expected to be well suited for transform-domain representation. Controlled by the FrFT order, the scaling performed in the phase captures the backscattering and topographic behaviors in SLC and InSAR, respectively [13], [16].

For a given image patch, 17 different FrFT operations are performed with 17 equally spaced transform angles between 0 and π (i.e., $\Delta p = 0.125$), and for each of the transformed images, three log cumulants (*log mean*, *log variance*, and *log skewness*) are computed from its amplitude, yielding a feature descriptor of length 51 [13]. However, for the InSAR and PGInSAR cases, the exploitation of the whole complex data is important in order not to ignore any contribution of the interferometric phase information. Hence, the log cumulants are computed for both the real and imaginary parts of the transformed image and then appended to obtain a feature descriptor of length 102 [11].

V. EXPERIMENTAL BISTATIC SAR IMAGE DATABASE

For patch-oriented classification, an image database of eight different classes is generated. The classes in the database are: **C1**—agricultural fields; **C2**—forest; **C3**—industrial area; **C4**—mixed vegetation (forest and agricultural); **C5**—riverside; **C6**—urban-1 (medium-density residential area); **C7**—urban-2 (high-density buildings); and **C8**—water body (lakes and ponds).

Each class consists of 50 scenes of size 200×200 pixels. For each scene, there exist three TerraSAR-X add-on for Digital Elevation Measurement (TanDEM-X) image pairs acquired in the

TABLE I
INSAR ACQUISITION PARAMETERS OF THE DATABASE

Parameter	Set-1	Set-2	Set-3
Acquisition date	03.02.2012	16.10.2011	06.07.2012
Orbit direction	Ascending	Ascending	Ascending
Look direction	Right	Right	Right
λ	0.031 m	0.031 m	0.031 m
θ	33.8°	33.8°	33.8°
B_\perp	88.2 m	128.4 m	229.6 m
h_{amb}	60.4 m	41.1 m	22.9 m

bistatic operation mode with three different effective baselines. That is, the whole database consists of 3 sets of 400 patches, with each set corresponding to a different effective baseline. The orbit direction, the look direction, the look angle, and the wavelength are identical for these acquisitions. These parameters and the corresponding ambiguity heights are tabulated in Table I.

The image patches are from the test site in Toulouse, France. The spatial resolution is about 2 m, i.e., each scene corresponds to an area of $400 \text{ m} \times 400 \text{ m}$ on the ground. Google Earth has been used as the *ground truth* during the generation of the database. In Fig. 2, each class in the database is visualized by a sample patch with its quick-look view of the SLC image, together with the corresponding optical Google Earth image.

Due to the bistatic imaging (S_1 transmits; S_1 and S_2 receive within the same pass), the temporal baseline is quite short (on the order of seconds), resulting in a low temporal decorrelation. Fig. 2(c) illustrates the spatial decorrelation due to the nonzero baseline with the probability density function (pdf) of the interferometric coherence values estimated for each sample patch and baseline. It can be seen that a longer baseline yields lower coherence as expected, except for C1 likely due to the seasonal agricultural activities. In addition, for Set 3, the sharp decrease in the C2 coherence may be not only due to the long baseline but also due to the presence of leaves in summer. For a given baseline, C3, C6, and C7 have similar coherence values due to the similar scattering mechanisms, and C5 and C8 have the lowest coherence values due to the specular reflection on water.

VI. CLASSIFICATION RESULTS

In this letter, a supervised K -nearest neighbor classifier, with a Euclidean distance and $k = 1$, is used. Four percent of each class, i.e., 2 out of 50 patches, is randomly selected as the training samples for the classifier. Each classification experiment is repeated 100 times with different training sets, and the average of the classification accuracy is reported.

A. Performance of SLC SAR, InSAR, and PGInSAR for Set 2

First, the classification performance of the SLC images, that of the InSAR images, and that of the PGInSAR images are compared for Set 2 (a medium baseline), and the results are given in Fig. 3. It can be seen that InSAR is better than SLC with a 15% enhancement in the mean accuracy. PGInSAR introduces further improvement of 10% over InSAR or a total improvement of 27% over SLC.

The F-measures show that InSAR considerably improves all class accuracy values compared with SLC, except for C5 and C8, which is likely due to the specular reflection in these classes. PGInSAR performs the best among all, and the biggest

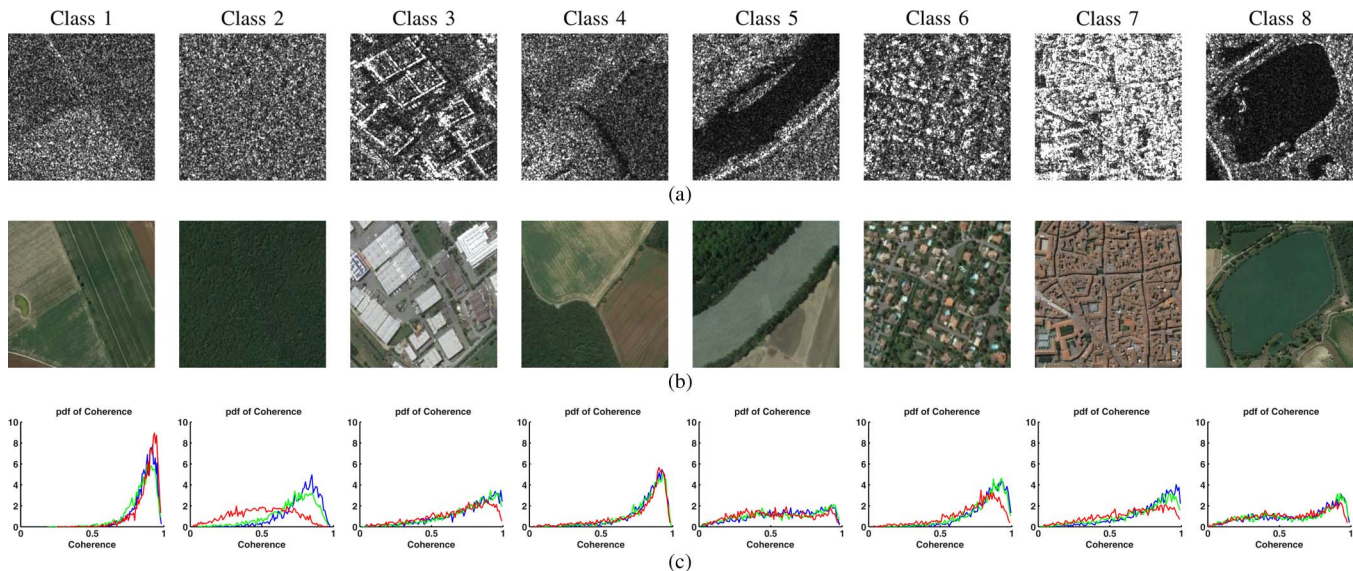


Fig. 2. Classes in the bistatic SAR image database: C1—agricultural fields; C2—forest; C3—industrial area; C4—mixed vegetation; C5—riverside; C6—urban-1 (medium density); C7—urban-2 (high density); C8—water body. (a) Quick-look views obtained from the amplitude of the master SLC data. (b) Optical Google Earth images as the ground truth. (c) PDFs of the interferometric coherence values computed with a 4×4 averaging window for three different baselines. Blue curve: Set 1 (a small baseline). Green curve: Set 2 (a medium baseline). Red curve: Set 3 (a long baseline).

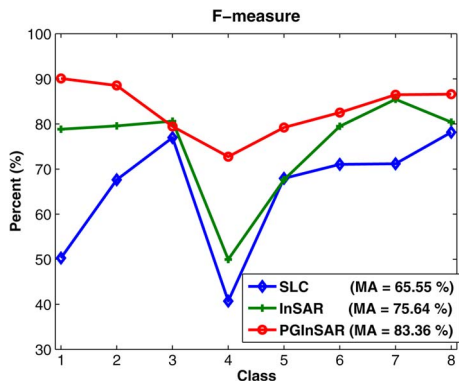


Fig. 3. F-measures and mean accuracy of SLC, InSAR, and PGIN SAR (Set 2).

TABLE II
CONFUSION MATRIX FOR PGIN SAR (SET 2) IN PERCENTAGE

Class index	Predicted classes							
	C1	C2	C3	C4	C5	C6	C7	C8
C1	90.22	4.04	0	5.74	0	0	0	0
C2	0.04	97.58	0	2.38	0	0	0	0
C3	0	0	78.62	0	0.42	14.92	6.04	0
C4	9.92	18.86	0	66.02	5.18	0	0	0.02
C5	0	0	1.70	6.90	78.62	1.10	0.06	11.62
C6	0	0	9.50	0.08	0.12	88.10	2.20	0
C7	0	0	8.12	0	0	9.38	82.50	0
C8	0.12	0	0	0.42	14.20	0.02	0	85.24

improvement is achieved for scenes with natural structures such as C1, C2, and C4. The improvements for C3, C6, and C7 are faint due to the strong phase discontinuities present, even in PGIN SAR.

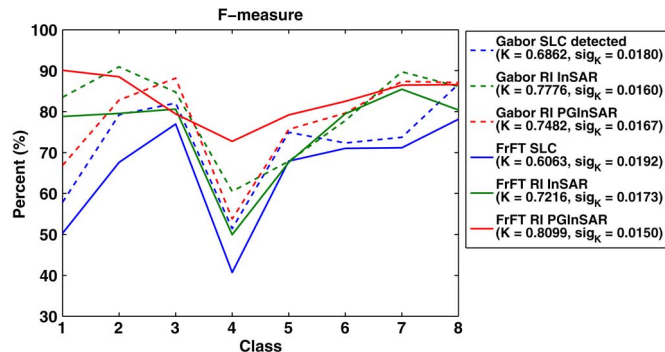


Fig. 4. Comparison of the Gabor- and FrFT-based feature descriptors (Set 2).

For a detailed assessment of individual class accuracy values, the confusion matrix for PGIN SAR is tabulated in Table II. It can be seen that the misclassifications are mostly observed within the following groups: 1) C1–C2–C4; 2) C3–C6–C7; and 3) C5–C8, which seems to be quite reasonable.

In addition, the FrFT-based feature descriptor is compared with a Gabor-based descriptor (with three scales and four orientations), which is a well-known multiscale approach capturing the textural information. The F-measures are given, together with Kappa coefficients κ and standard deviations σ_{κ} , in Fig. 4. The FrFT-based PGIN SAR feature descriptor yields the best performance with the highest κ and the smallest σ_{κ} . Moreover, it seems to be more uniform over all classes than other feature descriptors.

B. Performance of InSAR and PGIN SAR for Sets 1, 2, and 3

As the second experiment, the impact of the effective baseline is studied for the InSAR and PGIN SAR images. As shown in Fig. 5(a)–(c), for each effective baseline, PGIN SAR is better than InSAR. However, with the long baseline, for which the

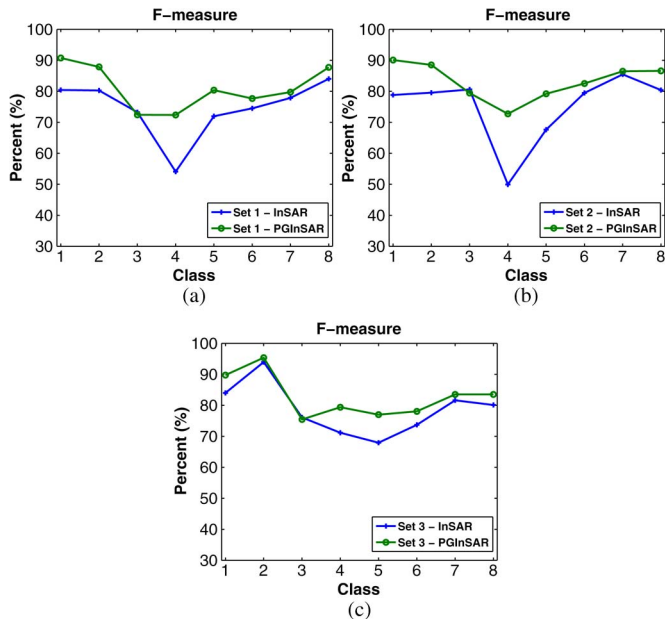


Fig. 5. F-measures for InSAR and PGIN SAR for (a) Set 1 (a small baseline), (b) Set 2 (a medium baseline), and (c) Set 3 (a long baseline).

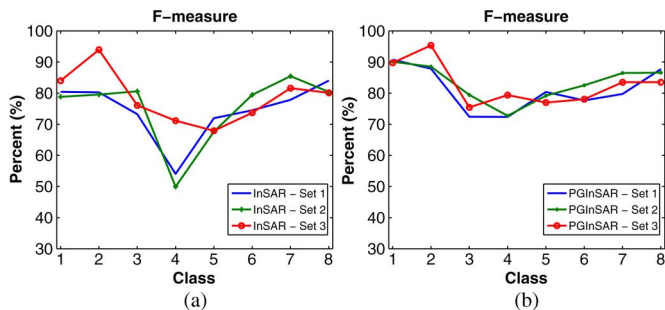


Fig. 6. F-measures for (a) InSAR and (b) PGIN SAR for the three different baselines.

number of fringes and phase discontinuities increase, PGIN SAR can no longer perform better than InSAR for C2, and similarly, the improvement is less for C1 and C4.

Furthermore, Fig. 6(a) and (b) shows how InSAR and PGIN SAR react to different effective baselines in terms of classification accuracy, respectively. In this figure, it can be seen that, for C2 and C4, the long baseline yields better accuracy despite the low coherence in C2. For C3, C6, and C7, the medium baseline is found to be the best among all. For C5 and C8, the small baseline results in better accuracy. That is, the behavior of a class differs under different effective baselines, as expected. Additionally, the effect of the seasonal changes between acquisitions should be considered. Nevertheless, it is important to note that the PGIN SAR feature descriptor is less affected by the effective baseline than the InSAR feature descriptor since the phase discontinuities are reduced compared with InSAR.

VII. CONCLUSION

In this letter, complex-valued InSAR and newly defined PGIN SAR images have been used for image classification, where the InSAR phase is related to the absolute height of the

scene and where the PGIN SAR phase rather reflects how fast the height changes throughout the image. As the feature descriptor, log cumulants are extracted from the real and imaginary parts of the FrFT coefficients of the complex-valued InSAR and PGIN SAR. Classification results show that InSAR and PGIN SAR improve the mean accuracy by 15% and 27% compared with SLC, respectively.

In addition, the impact of the effective baseline on the classification performance of the InSAR and PGIN SAR images is investigated through three interferometric pairs acquired from the same site with different effective baselines. For each baseline, PGIN SAR achieves better classification performance than InSAR since the phase discontinuities are reduced. Moreover, it has been observed that, for scenes containing natural structures with smooth phase changes, PGIN SAR considerably improves the classification accuracy compared with InSAR. However, the gain of PGIN SAR may remain faint for scenes containing man-made objects with abrupt phase changes and even for natural scenes with long baseline acquisitions. The behavior of a class may differ under different baselines, but yet PGIN SAR is more robust to the changes in the effective baseline than InSAR.

REFERENCES

- [1] U. Wegmuller and C. Werner, "Retrieval of vegetation parameters with SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 18–24, Jan. 1997.
- [2] A. Elmzoughi, R. Abdelfattah, and Z. Belhadj, "SAR image classification using the InSAR coherence for soil degradation cartography in the South of Tunisia," in *Proc. IEEE ICIP*, Nov. 2009, pp. 1677–1680.
- [3] V. Nizalapur, R. Madugundu, and C. S. Jha, "Coherence-based land cover classification in forested areas of Chattisgarh, Central India, using environmental satellite-advanced synthetic aperture radar data," *J. Appl. Remote Sens.*, vol. 5, no. 1, Mar. 2011, Art. ID. 059501.
- [4] V. Vijaya and G. J. Niveditha, "Classification of COSMO SkyMed SAR data based on coherence and backscattering coefficient," *Int. J. Comput. Sci. Inf.*, vol. 1, no. 4, pp. 60–63, 2012.
- [5] M. E. Engdahl and J. M. Hyyppa, "Land-cover classification using multitemporal ERS-1/2 InSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 7, pp. 1620–1628, Jul. 2003.
- [6] A. B. Suksmo, A. Handayani, and A. Hirose, "Snake in phase domain: A method for boundary detection of objects in phase images," in *Proc. IJCNN*, 2006, pp. 481–485.
- [7] F. Baselice and G. Ferraioli, "Statistical edge detection in urban areas exploiting SAR complex data," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 185–189, Mar. 2012.
- [8] F. Baselice, G. Ferraioli, and D. Reale, "Edge detection using real and imaginary decomposition of SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 3833–3842, Jul. 2014.
- [9] H. Wang, K. Ouchi, and Y. Q. Jin, "Classification of typhoon-destroyed forests based on tree height change detection using InSAR technology," in *Proc. IEEE IGARSS*, Jul. 2011, pp. 1247–1250.
- [10] A. I. Sharov, K. Gutjahr, and P. E. Pellikka, "Phase gradient approach to the evaluation and mapping of glacier rheology from multi-pass SAR interferograms," in *Proc. MultiTemp*, Jul. 2004, pp. 154–165.
- [11] N. D. Cagatay and M. Datcu, "Scene recognition based on phase gradient InSAR images," in *Proc. IEEE ICIP*, Paris, France, 2014.
- [12] P. A. Rosen *et al.*, "Synthetic aperture radar interferometry," *Proc. IEEE*, vol. 88, no. 3, pp. 333–382, Mar. 2000.
- [13] J. Singh and M. Datcu, "SAR image categorization with log cumulants of the fractional Fourier transform coefficients," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 12, pp. 5273–5282, Dec. 2013.
- [14] D. T. Sandwell and E. J. Price, "Phase gradient approach to stacking interferograms," *J. Geophys. Res.*, vol. 103, no. B12, pp. 30 183–30 204, Dec. 1998.
- [15] H. M. Ozaktas, O. Arikan, M. A. Kutay, and G. Bozdagi, "Digital computation of the fractional Fourier transform," *IEEE Trans. Signal Process.*, vol. 44, no. 9, pp. 2141–2150, Sep. 1996.
- [16] M. Datcu and J. Singh, "Phase-scale analysis of complex valued SAR images," in *Proc. EUSAR*, Jun. 2014, pp. 1121–1124.