Improving the accuracy of local frequency estimation for interferometric synthetic aperture radar interferogram noise filtering considering large coregistration errors

Jiao Guo\textsuperscript{1,2}, Weitao Zhang\textsuperscript{2}, Yanyang Liu\textsuperscript{2}, Longsheng Fu\textsuperscript{1}

\textsuperscript{1}College of Mechanical and Electronic Engineering, Northwest Agriculture and Forestry University, Yangling, Shanxi 712100, People’s Republic of China
\textsuperscript{2}National Key Lab of Radar Signal Processing, Xidian University, Xi’an 710071, People’s Republic of China

E-mail: jiao.g@163.com

Abstract: This study deals with the problem of estimating the local frequencies for interferometric synthetic aperture radar (InSAR) phase images (i.e., interferograms) noise filtering considering large coregistration errors. The estimation of local frequencies is frequently applied to achieving high performance of interferometric phase noise filtering, which is a key step in InSAR processing procedures. Unfortunately, the generated interferograms suffer seriously from coregistration errors especially for complicated topographies, thus imposing strong limits on the accuracy of local frequencies estimation by the existing methods. Taking large coregistration errors into account, the proposed method makes full use of the neighbouring pixels to construct joint pixel vector and take the separation extent of the signal and noise subspaces as the criterion to accurately estimate the local frequencies so that the effects of the envelope misalignment can be mitigated. Theoretical analysis and the simulated data as well as the real ERS-1/2 data show that the presented method has the ability to provide accurate estimation of local frequencies even if the coregistration error reaches one pixel.

1 Introduction

Synthetic aperture radar interferometry (InSAR) is one of the most important remote-sensing techniques to acquire terrain digital elevation models (DEMs) \cite{1, 2}. InSAR illuminates the same ground scene to obtain two or more complex SAR images from slightly different incident angles. Also, the phases of the SAR image pair are processed in interferometric way to determine the height of each scattering element in the observed scene (i.e., each pixel of SAR image). This advanced remote-sensing technique, which can work either day or night and through cloudy cover, has numerous applications in the fields of topography, geomorphology, seismology, and so on \cite{1–3}.

As is well known, interferometric phase noise suppression is a key step in InSAR processing procedures. To achieve fine performance of phase noise suppression, it is of extreme importance to obtain more filtering samples that are independent and identically distributed (i.i.d.) from the adjacent pixels. The i.i.d. condition requires that the adjacent pixels must have an identical terrain height. In other words, the height difference of the adjacent pixels inevitably results in poor performance of phase noise suppression, consequently deteriorating the DEM accuracy. Therefore, the effects of neighbouring terrain changes must be removed especially for complicated topography, which is usually called topography compensation. Of course, if a prior coarse DEM of the observed scene is available, the effect of terrain changes on the interferograms can be removed conveniently. However, in practice the processor is not always provided with a coarse DEM available. Consequently, the parameter estimation techniques are usually employed to estimate local frequencies of interferograms for topography compensation. On the one hand, the estimation of local frequencies can be completed by the existing methods such as the maximum likelihood (ML) method \cite{4} or the multiple-signal classification (MUSIC) method \cite{5–7}. However, the effectiveness of these methods strongly depends on the coregistration accuracy of each pixel in the SAR image pair, that is, these methods may fail in the presence of large coregistration errors. Unfortunately, the accuracy of coregistration methods is always finite, especially in the areas with complicated topography, since the coregistration offset is a function of terrain changes. For example, if the coregistration error reaches one pixel, the interferograms are decorrelated completely and seem to be purely noisy, and the methods mentioned above inevitably fail to correctly acquire local frequency estimates. On the other hand, the proposed method in the referee \cite{8} uses joint subspace projection method by constructing joint pixel vector. However, it is still necessary to select the neighbouring
pixels to estimate the joint covariance matrix under the assumption that the neighbouring pixels have an identical terrain, that is, topography compensation is completed based on a prior coarse DEM or accurate frequencies estimated. Consequently, if a prior DEM is not available, it is a vital problem to obtain accurate frequencies on the condition of coregistration errors. The work in this paper focuses on frequencies estimation under coregistration errors. Therefore, this manuscript can be looked upon as an extension to the referee [8].

The contribution of our work is to propose a new method to estimate the local frequencies of interferograms in the presence of large coregistration errors that are up to one pixel. The proposed method takes full advantages of the coherence information of neighbouring pixels to obtain robust estimation of local frequencies. Consequently, the effectiveness of local frequency estimation does not depend on the coregistration accuracy any more. Furthermore, combined with the joint subspace projection method [8–10], we can simultaneously estimate local frequencies and improve the phase estimation accuracy, even if there exist large coregistration errors.

The remainder of this paper is arranged as follows: Section 2 presents the signal model of the joint pixel vector and the optimisation functions for local frequency estimation considering large coregistration errors. In Section 3, the block diagram and the processing procedures of the proposed method are described in detail. The performance of the proposed method is investigated by the simulated data as well as real ERS-1/2 data in Section 4. In the end, Section 5 concludes the paper.

## 2 Signal model

Under the assumption that two complex SAR images are precisely coregistered, after phase flattening with a reference height the observed data vector of the pixel pair \((i, j)\) can be expressed as

\[
s(i, j) = a(\varphi_{ij}) \circ x(i, j) + n(i, j)
\]  

where \((i, j)\) represents the range and azimuth position, respectively, \(s(i, j) = [S_1(i, j) \ S_2(i, j)]^T\) is the spatial data vector with \(S_1\) and \(S_2\) the SAR image pair, \(a(\varphi_{ij}) = [1, e^{j\varphi_{ij}}]^T\), is the spatial steering vector corresponding to \((i, j)\), \(x(i, j)\) is the complex reflectivity vector, \(\circ\) is the Hadamard product, \(n(i, j)\) represents the additive noise, superscript \(T\) denotes matrix transpose, and \(\varphi_{ij}\) is the real interferometric phase.

Unfortunately, in any case, the errors of coregistration always exist, especially for the areas with complicated topographies, since the accurate amount of misalignment is always depending on the topography and InSAR system parameters. Therefore, considering large coregistration errors, we take full advantages of the information of neighbouring pixels to construct the joint pixel vector as

\[
js(i, j) = [s(i - 1, j - 1)^T, s(i - 1, j)^T, \ldots, s(i + 1, j + 1)^T]^T
\]  

where a \(3 \times 3\) pixel window centred at \((i, j)\) is adopted under the assumption that the coregistration error is less than or equal to one pixel and its orientation is uncertainty. Correspondingly, the array steering vector can exactly be formulated as

\[
ja(\varphi_{i-1,j-1}, \varphi_{i-1,j}, \ldots, \varphi_{i+1,j+1}) = [a^T(\varphi_{i-1,j-1}), a^T(\varphi_{i-1,j}), \ldots, a^T(\varphi_{i+1,j+1})]^T
\]  

If the condition holds that the neighbouring pixels have an identical terrain height (i.e. the InSAR phases of neighbouring pixels are identical after phase flattening), the joint array steering vector can be approximated by

\[
ja(\varphi_{i-1,j-1}, \varphi_{i-1,j}, \ldots, \varphi_{i+1,j+1}) = [a^T(\varphi_{ij}), a^T(\varphi_{ij}), \ldots, a^T(\varphi_{ij})]^T
\]  

However, the assuming condition is not always satisfied due to the real terrain changes, thus degrading the performance of phase estimation. In practical InSAR processing procedures, the effects of the terrain changes must be mitigated or eliminated in order to improve the accuracy of InSAR phase estimation [11, 12].

If a prior DEM of the illuminated scene is available, we can compute the one-to-one correspondence between SAR images and the coarse DEM [13], removing the effects of topography and achieving more accurate InSAR phase estimation. However, in many cases, we are not provided with a coarse DEM or the height accuracy of the coarse DEM is not sufficient. Therefore, the parameter estimation techniques are compelled to acquire the local frequencies of the interferogram. In this paper, we take the assumption that there exists only one frequency for each dimension in the local window, that is, the local slant plane assumption.

Correspondingly, the joint array steering vector denoted by (4) must be theoretically modified. Likewise, we adopt another equivalent strategy to modify the observed data vector with invariability of the approximated, that is, the effect of terrain change is removed from the data vector with the array steering vector by (4). Based on the local slant plane assumption, the interferometric phase in the local window can be expressed as

\[
\varphi(i + u, j + v) = \tau_a \cdot u + \tau_r \cdot v + \varphi(i, j)
\]  

where \(\tau_a\) and \(\tau_r\) represent the local azimuth and range frequencies, respectively, and the parameters to be determined, \(u\) and \(v\), denote the azimuth and range distances of the pixels in the local window relative to the central pixel \((i, j)\). Therefore, to avoid the effect caused by the topography, the compensation phase is formulated as

\[
\Delta\varphi(i + u, j + v) = \tau_a \cdot u + \tau_r \cdot v
\]  

After the topography phase compensation, the spatial data vector is modified as

\[
s'(i + u, j + v) = [S_1(i, j) \cdot e^{-j2\pi(i+u+j+v)} \ S_2(i, j)]^T
\]  

Correspondingly, the joint pixel vector is reconstructed and can be written as

\[
js'(i, j) = [s'(i - 1, j - 1)^T, s'(i - 1, j)^T, \ldots, s'(i + 1, j + 1)^T]^T
\]
The covariance matrix of the updated joint pixel vector $\mathbf{j}(i, j)$ is expressed as

$$ C_{\mathbf{j}}(i, j) = E\{\mathbf{j}(i, j)\mathbf{j}^H(i, j)\} $$

(9)

which is named joint covariance matrix and the superscript $H$ denotes vector conjugate-transpose. In practice, we compute the covariance matrix from the neighbouring samples, so it should be noted that the neighbouring pixels adopted for estimating the covariance matrix must be included in the local window compensated. By eigendecomposing the joint covariance matrix, we can obtain the joint noise and signal subspaces combined with responding eigenvalues. The joint covariance matrix is assumed to be with the dimensions of $M \times M$, and it can be eigendecomposed to

$$ C_{\mathbf{j}}(i, j) = \sum_{k=1}^{K} \lambda_k \mathbf{b}_{\mathbf{j}c}^k \mathbf{b}_{\mathbf{j}c}^k + \sum_{k=K+1}^{M} \lambda_k \mathbf{b}_{\mathbf{n}c}^k \mathbf{b}_{\mathbf{n}c}^k $$

(10)

where $\lambda_1 > \lambda_2 > \cdots > \lambda_K \gg \lambda_{K+1} > \cdots > \lambda_M$ are the eigenvalues, $\mathbf{b}_{\mathbf{j}c}$ and $\mathbf{b}_{\mathbf{n}c}$ are the eigenvectors corresponding to the smaller and larger eigenvalues, respectively. From (10), we can obtain the signal and noise subspaces according to [8]. The proposed method takes the separation extent of the signal and noise subspaces as the criterion to optimise the estimates of the local frequencies. The more accurate the local frequencies $\tau_s$ and $\tau_r$ are contained in the compensation phase, the higher is the criterion. The separation of the eigenvalues are equivalent

![Fig. 1 Block diagram of estimating the accurate local frequency and phase filtering](image-url)
to that of noise and signal subspaces, which can be measured by the proportion of $\lambda_K$ to $\lambda_{K+1}$. To maximise the proportion $\lambda_K/\lambda_{K+1}$, we can obtain the estimation of local frequencies. Consequently, we define the cost function as

$$\hat{\tau}_a, \hat{\tau}_r = \max_{\tau_a, \tau_r} \left\{ \frac{\lambda_{K}(\tau_a, \tau_r)}{\lambda_{K+1}(\tau_a, \tau_r)} \right\}$$

which is a two-dimensional optimisation problem. To reduce the computational load, in practical applications the azimuth local frequency $\tau_a$ and range local frequency $\tau_r$ are estimated separately. Equation (11) is decomposed into two one-dimensional optimisations as

$$\hat{\tau}_a = \max_{\tau_a} \left\{ \frac{\lambda_{K}(\tau_a)}{\lambda_{K+1}(\tau_a)} \right\}$$

(12)

$$\hat{\tau}_r = \max_{\tau_r} \left\{ \frac{\lambda_{K}(\tau_r)}{\lambda_{K+1}(\tau_r)} \right\}$$

(13)

Solving (12) and (13), the local azimuth and range frequencies are determined.

With the estimated local frequency, the effect of terrain changes could be mitigated and the assumption of i.i.d. condition is more satisfied. Combined with the method proposed by [8], the real interferometric phase is figured out and also its accuracy is improved.

3 Processing procedures

In this section, we will describe the detailed processing procedures of the proposed method mentioned previously. The complete block diagram of the proposed method is shown in Fig. 1, where it is divided into two main parts, the first part (indicated by dot rectangle in Fig. 1) of which is the focus in this paper and mainly comprises the following five steps.

Step one: Coarse coregistration: The SAR images are coarsely registered using the existing methods [14–16]. The allowance errors of coregistration can be up to one pixel, thus greatly decreasing the complexity of coregistration.

Step two: Terrain phase compensation: For the pixel pair $(i, j)$ of the images and each searched value pair of $\tau_a$ and $\tau_r$, the local frequencies $\tau_a$ and $\tau_r$ are utilised to compute the compensation phase according to (7). And then the phases of the neighbouring pixels are compensated by subtracting the computed phase, which is equivalent to remove the interferometric phase caused by the terrain changes.

Step three: Joint covariance matrix computation: The compensated neighbouring pixels are selected to estimate the joint covariance matrix.

$$\hat{\mathbf{C}}_{js}(i, j) = \frac{1}{(2p+1)(2q+1)} \sum_{u=-p}^{p} \sum_{v=-q}^{q} j_s(i+u, j+v) \times j_s^{H}(i+u, j+v)$$

(14)

where $(2p+1)$ and $(2q+1)$ are the number of azimuth and range samples of the local window.

Step four: Joint covariance matrix eigendecomposition: The joint covariance matrix $\hat{\mathbf{C}}_{js}(i, j)$ assumed with the
dimensions of $M \times M$ can be eigendecomposed into

$$\hat{C}_{\mu}(i, j) = \sum_{k=1}^{K} \lambda_k \beta_{\mu i}^{(k)} \beta_{\mu j}^{(k)H} + \sum_{k=K+1}^{M} \lambda_k \beta_{\mu i}^{(k)} \beta_{\mu j}^{(k)H} \quad (15)$$

from which, we can obtain $M$ eigenvalues as $\lambda_1 > \lambda_2 > \cdots > \lambda_K \gg \lambda_{K+1} > \cdots > \lambda_M$.

Step five: The proportion optimisation: Maximising the proportion of $\lambda_k$ to $\lambda_{k+1}$, the corresponding values of $\tau_a$ and $\tau_r$ are the accurate estimation of local frequencies.

The estimation of local frequencies can be obtained by performing the five steps above for each pixel pair of the interferogram regardless of the coregistration error, which will be verified in the following section. With the joint covariance matrix eigendecomposition in the optimal case,
the joint subspaces (i.e. the joint noise subspace and the joint signal subspace) can be derived. Finally, the interferometric phase can be determined by projecting the joint signal subspace onto the joint noise subspace [8].

4 Performance investigation

In this section, we investigate the performance of the proposed method with the simulated data as well as real ERS-1/2 data. The local frequencies estimated by different methods are applied to compensate the terrain phase and the filtered interferograms generated by the same joint pixel projection method are adopted to evaluate the performance of frequency estimation. For comparison, the results by other methods including ML method and MUSIC method are included.

In the first experiment, we use the simulated data to verify the performance of the proposed method. It is assumed that the InSAR system is composed of two formation-flying satellites with effective cross-track baseline length (two-way) 543.14 m. The effective cross-track baseline length is the projection of the cross-track baseline onto the direction perpendicular to the line of sight. The height of the orbit is 600 km, and the incident angle of the antenna beam is 40°. The terrain is simulated by a two-dimensional hamming window (i.e. the hamming window function in MATLAB environment), which is shown in Fig. 2. The statistical model is used to generate the SAR image pair and the signal-to-noise ratio (SNR) is set to be 23 dB.

Fig. 3 shows the original interferograms directly generated by the SAR image pair for the coregistration error of 0, 0.7 and 1 pixel, respectively. The original interferograms for different coregistration errors are used to investigate the performances of the proposed method. Comparing Figs. 3a–c, it can be clearly seen that the noise level in the image increases with respect to the increase of coregistration errors. For the worst case that the coregistration error reaches one pixel, the interferogram seems to be purely noisy.

To illustrate the proportion value in the case of different coregistration errors, first a small data set is used before the proposed method is applied to the whole images. The size of the selected data set depends on the neighbouring pixels that are selected to estimate the joint covariance matrix and is fixed to be 7 × 7, which is indicated by the black rectangle shown in Fig. 4. Figs. 4a–c show the normalised proportion value of $\lambda_K/\lambda_{K+1}$ with respect to local frequencies under three conditions of coregistration errors, which are 0, 0.7 and 1 pixel, respectively. From the results, it is concluded that the true estimates of local frequencies can be acquired by identifying the peak of the cost function $\lambda_K/\lambda_{K+1}$, even though the noise level of the curve are rising with respect to

---

Fig. 6 Interferograms for the coregistration error of 0.7 pixel

a The filtered interferogram without topography phase compensation
b, c and d The filtered interferograms with topography phase compensation using local frequencies estimated by ML, MUSIC and our method, respectively
the coregistration error, that is, there is still information available even if the interferometric phase seems to be purely noisy with the coregistration errors of one pixel.

Figs. 5–7 are the results of the whole images. Fig. 5 shows the filtered interferograms in the case of accurate coregistration. The same joint pixel projection method is applied to generate the filtered interferogram by local frequency estimation of different methods. In other words, the local frequencies are obtained by different methods which include ML, MUSIC and the proposed method, and the local frequencies estimated are applied to joint pixel projection method to generate corresponding interferograms. Fig. 5a is the filtered interferograms without topography compensation. Figs. 5b and c show the filtered interferograms by ML method and MUSIC method, respectively. Fig. 5d is the interferogram generated by the proposed method. From Fig. 5a, it can be seen that the result without topography compensation is satisfied in the area with sparse fringes (i.e. the terrain with low slopes), however, the filtered interferogram in the area with dense fringes (i.e. the terrain with high slopes) deteriorates sharply, indicated by a white ellipse as poor area in Fig. 5a. Figs. 5b–d show the interferograms filtered are of high quality, which is because the local frequencies can be estimated accurately by different methods in the case of accurate coregistration.

Figs. 6 and 7 show the filtered interferograms for the coregistration errors of 0.7 and 1 pixel, respectively. The filtered interferograms are generated through the same
processing procedures as Fig. 5. Figs. 6a and 7a are the filtered interferograms without topography compensation. Figs. 6b and 7b exhibit the filtered interferograms by ML method. Also, the results by MUSIC method are given by Figs. 6c and 7c. Correspondingly, Figs. 6d and 7d are the filtered interferograms by our method. It can be seen that Figs. 7b and c seem to be noisy due to the wrong estimation of local frequencies by the ML and MUSIC methods for the coregistration error of one pixel.

Comparing Figs. 5–7, the joint subspace projection method can generate fine interferograms for large coregistration errors. From Fig. 5a, it can be seen that the result without topography compensation is satisfied in the area with sparse fringes (i.e. the terrain with low slopes), however, the filtered interferogram in the area with dense fringes (i.e. the terrain with high slopes) deteriorates sharply, indicated by a white ellipse as poor area in Fig. 5a. Figs. 5b–d show the interferograms filtered are of high quality, which is because the local frequencies can be estimated accurately by different methods in the case of accurate coregistration. The proposed method in this paper can acquire accurate estimation of local frequencies in the presence of large coregistration errors (up to 1 pixel), thus greatly increasing the performance of interferometric phase noise filtering.

Additionally, the root-mean-square error (RMSE) is employed to evaluate the accuracy of local frequency estimation for different methods. Fig. 8 shows the relationship that the local frequency estimation errors vary with respect to the image coregistration error for different methods.

In the following experiment, the ERS-1/2 real data (ERS-1 orbit = 32 585, ERS-2 orbit = 12 912, frame = 781, 1997–10-08/09, Shangyi, China) is used to confirm the effectiveness of the proposed method. Fig. 9a is the
interferogram obtained from the image pair of accurate coregistration, and Fig. 9b is the corresponding result by the proposed method. Furthermore, Fig. 9c is the interferogram for the coregistration error of 0.8 pixel, and Fig. 9d is the corresponding output result by our method.

5 Conclusion

In this paper, a new method to estimate the local frequencies of interferogram considering large coregistration errors has been presented. Based on the construction of the joint pixel vector and the covariance matrix, the proposed method adopts the separation extent of the signal and noise subspaces as the criteria, optimised to determine the estimates of local frequencies. In the experiments, the simulated data as well as the real ERS-1/2 data has been employed to investigate the performance of the method. The two experiments carried out demonstrate that the proposed method has the ability to acquire accurate estimation of local frequencies even if the coregistration error reaches one pixel. For future InSAR systems with higher spatial resolutions, for example, 1 m × 1 m or even higher [1, 17], it is more difficult to increase the accuracy of image coregistration, so the proposed method would achieve more satisfactory results.

6 Acknowledgements

The authors would like to thank the European Space Agency (ESA) for providing the real SAR data. Many thanks are also given to the anonymous reviewers for their valuable suggestions. This work is supported by the National Natural Science Foundation of China (NSFC) under Grant 41301450, the Fundamental Research Funds for the Central Universities (QN2013066) and Exclusive Talent Funds of Northwest A&F University (Z111021302).

7 References