Digital Processing of Remote Sensing Data

Ivana Čapková, ivana.capkova@fsv.cvut.cz

August 29, 2004

Acknowledgements

Our research is supported by ESA which provided the image data. Precise orbits of the satellite were provided by the DEOS group, Delft University of Technology, the Netherlands.

DORIS is a free open-source batch program for InSAR data processing, being developed also at TU Delft. It is designed to process the SLC data and generate the interferogram with the coherence image, or a geocoded digital elevation model or deformation map. This time, phase unwrapping is not implemented and the external SNAPHU program is used instead. For details, see [7].

1 Introduction

In remote sensing, photographic data are usually not used anymore; all the other data are digital. This paper deals with processing of the imagery data, especially radar data from ERS-1/2 or ENVISAT. This paper will not focus on image acquisition, we will deal with the already acquired data processing.

The purpose of remote sensing imagery data processing is usually to map something: we may want to create a topographic map of the terrain, we may do some thematic mapping, we may look for Earth deformations (landslides, earthquakes etc.) or we may try to observe certain atmospheric features. For geodetic purposes, creating topographic or thematic maps is the most common task, mapping the deformations is also a similar and maybe the most important task.

For thematic mapping, just one image may be sufficient. When processing, the amplitude of the received signal is the most important and corresponds to the reflectivity that is often looked for.

Looking for atmospheric parametres is not a geodetic task so we will not deal with it here.

For the other tasks, we need to introduce the interferometry method.

2 Interferometry

Synthetic aperture radar (SAR) interferometry is a method of processing two SAR images which can result in a topographic (DEM) or deformation map. For this kind of processing, the radar needs to be coherent (that means that the phase of transmitted signal must be stored and later compared with the phase of the received signal, and just their difference is processed). Because the transmitted signal is frequency-modulated, we receive two types of information with the signal: its magnitude (i.e. the reflectivity of the corresponding ground area), and the phase, i.e. the distance.

It looks very simple but the distance is "coded". The phase is naturally defined only in the $(-\pi, \pi)$ interval, so the distance is determined only in the $(-\lambda/2, \lambda/2)$ interval, with the unknown number of $\lambda/2$ laps. For ERS-1/2 and ENVISAT, $\lambda = 5.6cm$. So, the distance from the transmitter/receiver to the reflector is $\frac{k\lambda + \frac{\varphi}{2\pi}\lambda}{2}$ where $\varphi \in (0, 2\pi)$ is the measured phase, λ is the radar wavelength and k is the so-called (range) ambiguity. Let's add that there is no way to determine the ambiguity.

The principle of the interferometry is to compute the phase difference between the two acquired signals, i.e. the distance difference (of course, we can computed just the phase difference between the two signals, not the ambiguity difference; the processing and the consequences will be discussed later). From this difference, we can find out the "height" of the reflector (above some reference surface), or a movement of the reflector between the two acquisitions.

The radar emits C-band microwaves and receives the backscattered radiation. The magnitude (power) of the received radiation is (in contrast to the optical data) independent of the atmosphere and is only influenced

by the dielectric nature of the reflectors and their orientation. For example, a water surface reflects all the energy away from the radar and that's why the water bodies appear black in the image. As for interferometry, studying large water bodies is unadvisable because the phase of the received signal (if any) is very unreliable. Fortunately, "most nature object are diffuse reflectors" [9] which send a part of the reflected radiation back to the satellite. The best reflectors are corner reflectors where the first reflector is usually a water body (trees on a river bank, bridges over river).

By nature of radar, the phase quality is much better in stable areas. In urban areas, where the reflectors don't move for years or move a distance which is much smaller than the radar wavelength, interferometry gives good results. On the other hand, forests cannot be mapped by a radar, and images of fields can be only acquired in spring and in autumn, when there is neither vegetation nor snow there. Also, large ice areas which don't melt can be mapped because they don't move.

In interferometry, as in photogrammetry, baseline between the two radars is considered. The actual baseline can be divided into two parts: the perpendicular (effective) baseline, perpendicular to the beam direction, and the parallel baseline, parallel to the beam. The effective baseline is important for the interferogram quality: if it is too small, the radar images are the same (except for atmospheric effects and eventually some temporal changes (deformation, snow etc.)) and their interferogram can't be used for DEM generation. If the effective baseline is too large, finding the coregistration between the images is more difficult because of the change in the look angle and thus the scattering features of the area. There is a maximum effective baseline (about 1000 km) beyond which the interferometry is not possible. The parallel baseline just determines the phase difference between the two images.

As we already noted, interferometry allows us to process only the phase differences, not the distance differences. The distance may be different but the phase difference varies quite slowly throughout the image. Of course, the varying speed depends on the slope of the terrain and its orientation to the satellite. The phase difference between the neighboring pixels larger than π (i.e. the distance difference larger than $\frac{\lambda}{4} \approx 1.4cm$ is considered a mistake (phase unwrapping will be discussed later). Let's denote that the resolution of the original image is approximately 4m in azimuth and 20m in (slant) range. But, to improve the image features (to reduce the speckle), the image is often spatially averaged (and approximately square pixels are achieved).

Let's emphasize here that radar only measures the distance: this means that deformations in the plane perpendicular to the radar beam can't be measured. Also, the deformations are measured only in this direction, the perpendicular part can never be acquired by radar interferometry.

Interferometry generally consists of three steps: coregistration of the two images, interferogram creation and interferogram interpretation.

To be able to comply with precision requirements, we need to know the orbits precisely — they are needed for the final phases: subtraction of the reference phase, slant-to-height conversion and geocoding. Unfortunately, the orbits obtained from the DEOS precise orbits procedure are usually not precise enough.

3 Coregistration

The first step to be done with the two SAR images is the coregistration. The coregistration is neccessary in order to be able to compare the phases of the two received signals. Coregistration means looking for a shift vector between the two images, and this vector needs to be determined very precisely to get meaningful results later.

In DORIS, coregistration is done in three steps: First, a coarse shift vector is computed from the orbit data of the two satellites. This shift vector is influenced by the orbit errors and sometimes may have a very bad value — so bad that the following step fails. Second, a "coarse coregistration" is done. That means that there are many windows in the image, created with a defined size, and correlation of the images within the window is computed. This is done for each window, and then the shifts from the most correlated windows are averaged. The correlation is given as

$$C(dx, dy) = \sum_x \sum_y (I_1(x, y) I_2(x - dx, y - dy))$$

and the shift vector as

$$(\delta x, \delta y) = max_{dx,dy} C(dx, dy),$$

where I_1 and I_2 are the intensities of the two images respectively (these intensities must be adjusted, e.g. the half of the intensity scope subtracted, so as to achieve both negative and positive values), C is the correlation and x, y are the image coordinates.

These windows should be quite large (hundreds of pixels), especially if the orbit data are not precise. It is important to have the windows much larger than the expected error from the previous step. The obtained shift vector should now be determined within 2 or 3 pixels. Third, the "fine coregistration" is performed. The windows are smaller (e.g. 8 pixels) and the data are oversampled (e.g. 16 times) to get the coregistration precision about 1/10 of a pixel (the precision of finding the convolution maximum is always larger than the smaller element).

Usually, there are many computing windows in an image and the shift vector is different in different areas (because of the nature of affine transformation). That's why a 2D polynomial is computed to better fit the measurements. Also, a correlation C is related to each shift vector (one for each window), and when looking for the polynomial parameters, the vector which are less probable (have smaller correlation) are excluded. After that, one of the images (the slave one) is resampled so as to have the "center of the pixels" at the same place as in the master image.

To compute the convolution faster, fast Fourier transform (FFT) is often used. Anyway, the results obtained are the same as those obtained by computing it in the space domain, which is simpler to explain.

4 Interferogram creation

Before creating the interferogram itself, both images need to be filtered both in the range and azimuth direction, in order to eliminate those parts of spectrum which are not present in the other image. The difference of spectrum in the azimuth direction between the two images is given by SAR processing — if the SAR processing yields the same Doppler centroid for both images, this filtering is not necessary. Doppler centroid is the Doppler frequency measured in the beam center, for more details see [**?**]. The different spectrum in the range direction is a result of different viewing angles — the larger the perpendicular baseline, the smaller part of the spectrum is overlapped (see [6]).

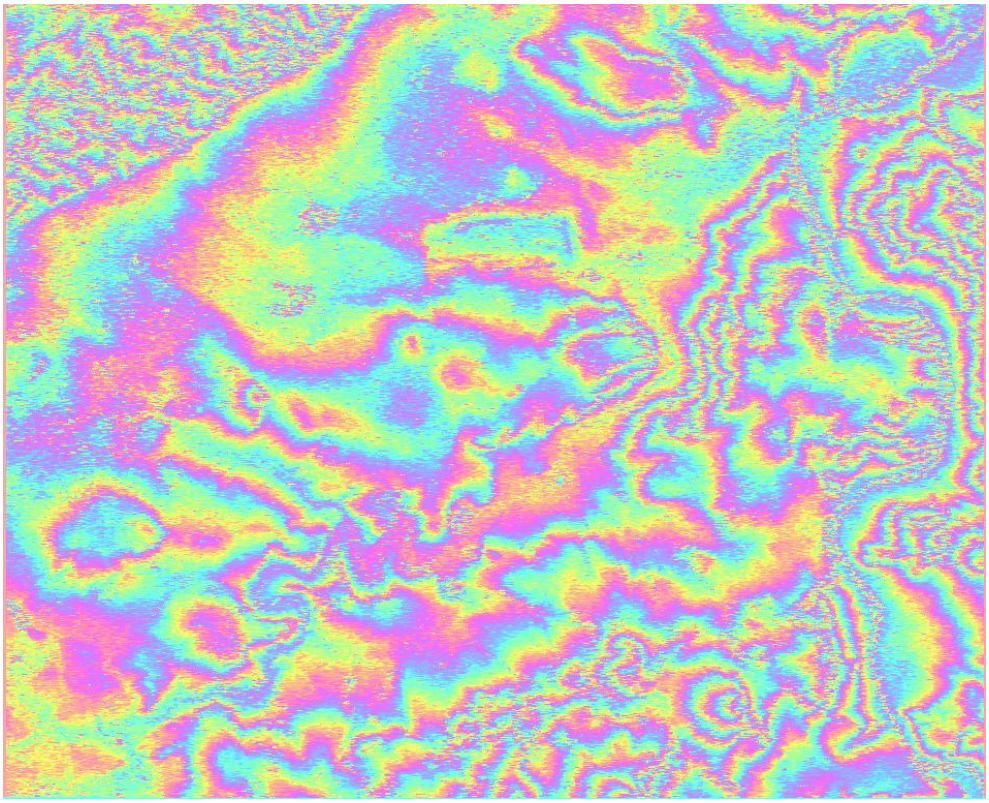

Figure 1: Topo pair interferogram with the reference phase corrected, filtered

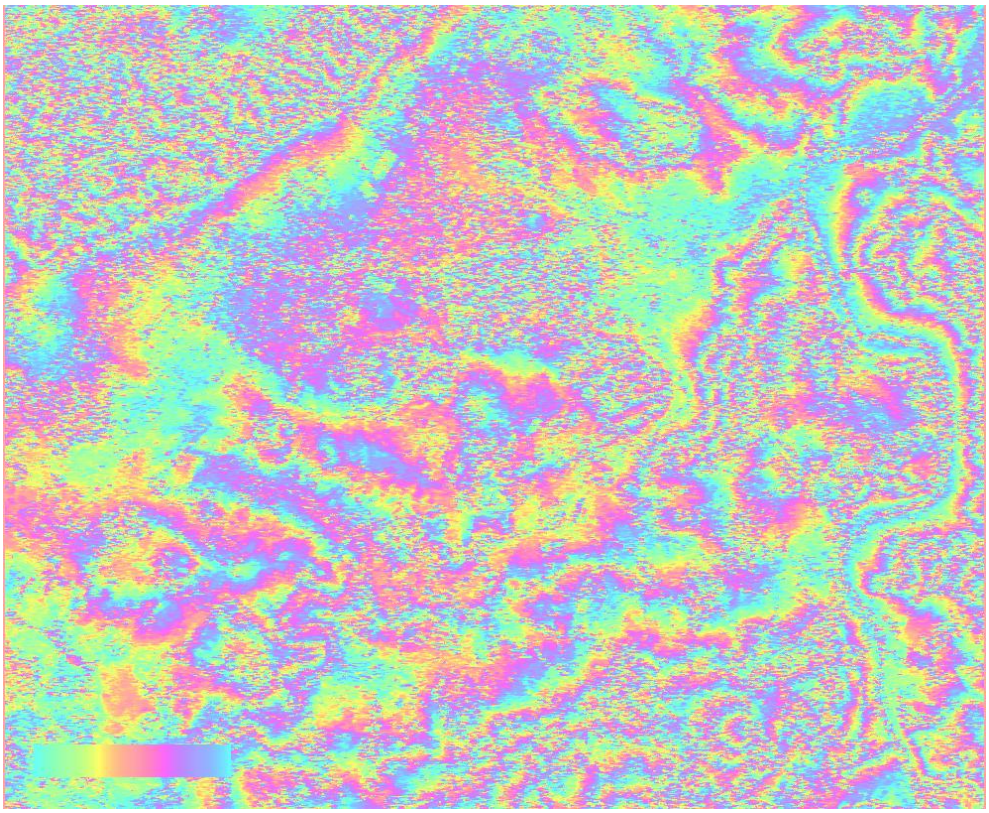

Figure 2: Defo pair interferogram with the reference phase corrected, filtered

The interferogram is defined as the (pixelwise) complex cross-multiplication of the two images:

$$I = M \cdot S^*,$$

so that

$$A_I = A_M * A_S,$$

$$\varphi_I = \varphi_M - \varphi_S,$$

where I is a pixel in the interferogram, M master, S slave image, A is the magnitude, φ is the phase. The magnitude is not interesting for us, but the phase is. The interferogram consist of "fringes" (i.e. 2π phase cycles) if done well.

After that, the phase difference of the reference surface (ellipsoid) must be subtracted. The subtraction is also done by complex multiplication with a signal with unity magnitude. We may either compute the reference phase as the distance between the given pixel and the satellite position (obtained from the precise orbits procedure), but this usually gives bad results due to the orbit errors, or compute the "exact" polynomial that matches the frequency of the fringes in the interferogram. If these "regular" (frequent) fringes (in the range direction) are eliminated (the polynomial must be of a low order), only the "residual" fringes are left, containing the topographic, deformation and atmospheric signals.

To be able to distinguish where the interferogram is "right" and where it is "wrong", the coherence is also computed. The complex coherence is defined (according to [6]) as

$$\gamma_c = \frac{\mathcal{E}\{M \cdot S^*\}}{\mathcal{E}\{M \cdot M^*\} \cdot \mathcal{E}\{S \cdot S^*\}},$$

where the $\mathcal{E}$ operator is the mean value along a predefined area (e.g. 30 (azimuth) by 6 (range) pixels). The real coherence is given by $\hat{\gamma} = |\gamma_c|$ and its value determines the quality of interferogram in that area.

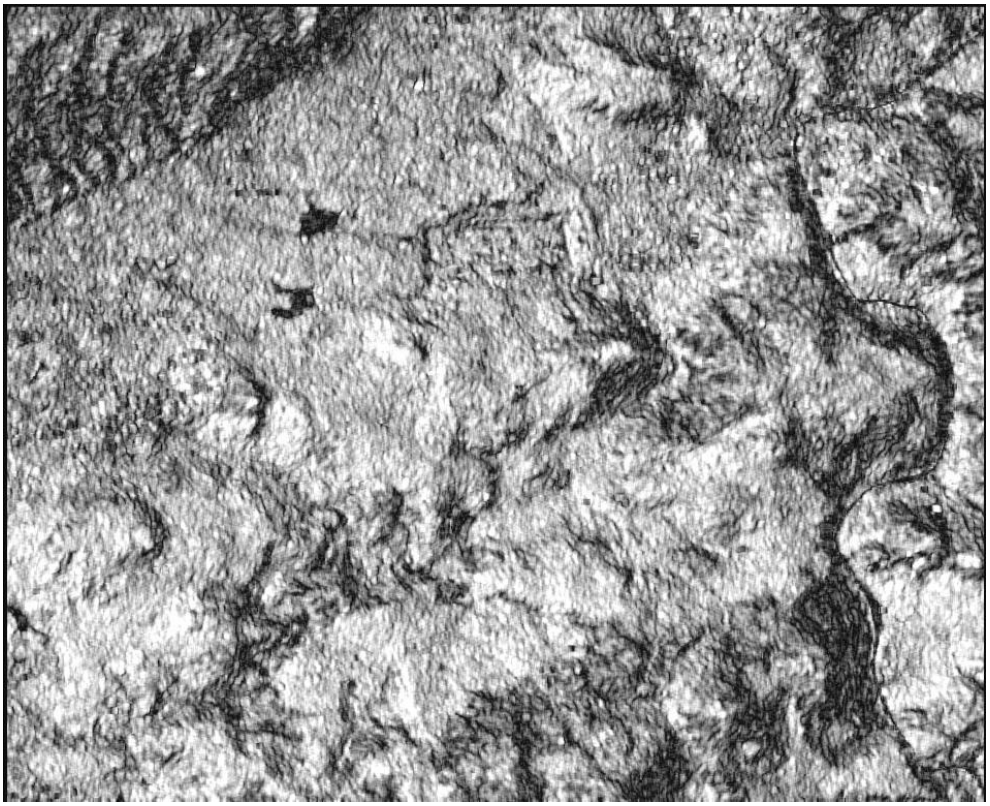

Figure 3: The coherence of the topo pair

As a characteristic feature of an interferogram, altitude of ambiguity is often used. It is the altitude corresponding to one fringe in the interferogram. The value of the altitude of ambiguity depends primarily on the perpendicular baseline (the exact relationship is given e.g. in [9]) and is important for selecting images: a small value is better for extracting the topo signal — the larger the altitude of ambiguity, the less of the topo signal is contained in the interferogram, the better for extracting the defo signal. That is also the reason why a relatively unprecise DEM may be used for reducing the topo signal in an interferogram used for deformation mapping — if the altitude of ambiguity is much larger than the DEM errors. An ideal case for deformation mapping is having two images acquired from the same position — corresponding to the infinite altitude of ambiguity.

Let's add here that the altitude of ambiguity varies slightly across the range and azimuth but for image selection, just one value may be used.

Also, more interferograms may be used to obtain the desired altitude of ambiguity. This procedure is called integer interferogram combination ([1, 9]). The interferograms are "stacked" like

$$I_s = \sum_i q_i I_i$$

where I_s is the stacked interferogram, q_i are integers and I_i are the interferograms to be stacked. The values of q_i cannot be too large because the noise is also added up (the same reason why it is not possible to stack many interferograms). The equivalent altitude of ambiguity of the stacked interferogram is then (for derivation, see [9])

$$\frac{1}{h_{ae}} = \sum_i \frac{q_i}{h_{ai}}$$

where h_a is altitude of ambiguity.

Before unwrapping, the interferogram needs to be (low-pass) filtered in order to achieve a smoother image with less noise.

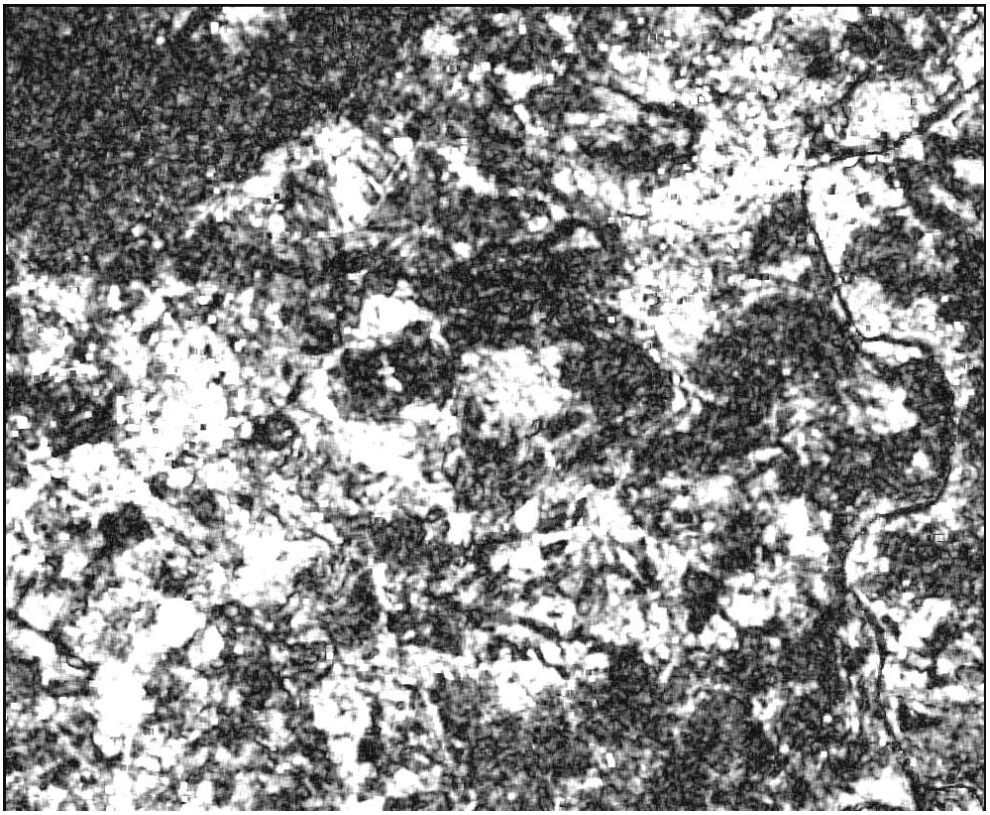

Figure 4: The coherence of the defo pair

5 Phase unwrapping

Now, the problem of unwrapping the phase arises, i.e. extracting it from the $(-\pi, \pi)$ interval to general values. Of course, we need a reference point with a known value, but that is not the main problem. This problem is easy only when some conditions are filled, and this is not the practical case. These conditions are:

- the absolute difference between two neighbouring (original) phase differences is less than π,
- the signal is not too noisy.

The problem of unwrapping the phase is ambiguous by principle and there is no analytical method to solve it. There are two types of numerical methods to solve it: branch-cutting methods and L^P norm methods. Branch-cutting methods are path-following; with the use of the image and coherence map, they create a graph and look for the most quality paths — and the phase is integrated along the paths. A noise may introduce a big problem here. The L^P norm methods are a little bit more reliable, computing the unwrapped phase at the same time for the entire image; but the computations take a much longer time. For more details about phase unwrapping methods, see [2].

6 Differential interferometry

To obtain a deformation map, two interferometric pairs need to be processed. The first one, which can be substituted by the external DEM, is a topo pair, best acquired with a long perpendicular baseline and a small temporal baseline. Before further processing, interferogram of these two images must be unwrapped (after that, it's a DEM, although in the slant range system).

Another way to obtain a deformation map is to use the defo pair and an external DEM.

From the other pair (called the defo pair, usually having a long temporal baseline (so as to allow the deformations to be present) and a small perpendicular baseline (so as to reduce the topographic signal)) we also

create the interferogram. These two image pairs can have a common master but it is not a neccessity. But, in both cases, the two interferograms must be coregistered (and eventually resampled). This is especially problem if an external DEM is used — the orbit errors cause an "geocoding" error which is too large to be ignored when using well geocoded DEM (e.g. SRTM).

When using an external DEM, this DEM must be first converted to the slant range (similar to the unwrapped interferogram).

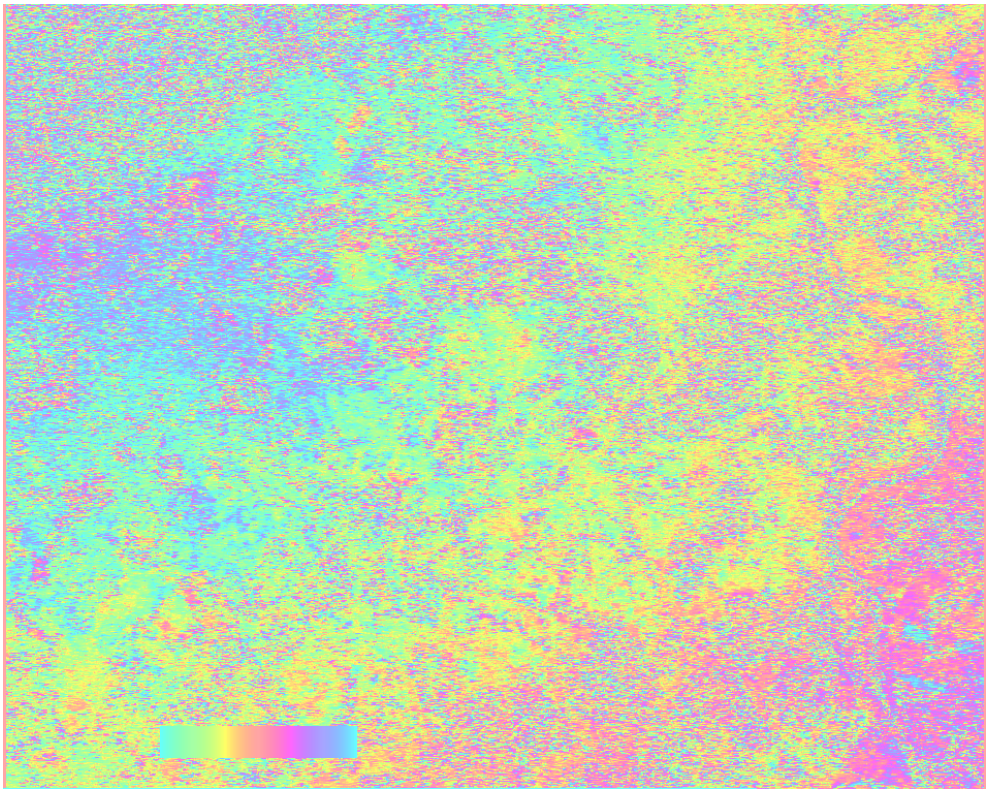

Figure 5: The differential interferogram

For each pixel, individual baselines are computed for both the topo and defo pair, and the topographic signal is scaled according to them.

The phase of the defo interferogram (with the use of the topo interferogram phase) can be written as (according to [6])

$$\varphi' = \varphi \frac{B'_{\parallel}}{B_{\parallel}} + \frac{4\pi}{\lambda}\Delta r,$$

where $B_{\parallel}$ and $B'_{\parallel}$ are the parallel baselines of the topo and defo pairs respectively, λ is the radar wavelength, φ is the phase of the topo interferogram (corrected for the phase of the reference body) and Δr is the deformation in the line of sight occured during the acquisitions. Because of the impossibility to determine the parallel baselines precisely the formula with perpendicular baselines is used for computing the differential phase (for the derivation, see [6]):

$$\varphi' = \varphi \frac{B'_{\perp}}{B_{\perp}} - \frac{4\pi}{\lambda}\Delta r,$$

where $B_{\perp}$ and $B'_{\perp}$ are the perpendicular baselines of the topo and defo pairs, respectively. Thus the phase of the differential interferogram is given as (see [6])

$$\varphi_{\Delta r} = \varphi' - \frac{B'_{\perp}}{B_{\perp}}\varphi.$$

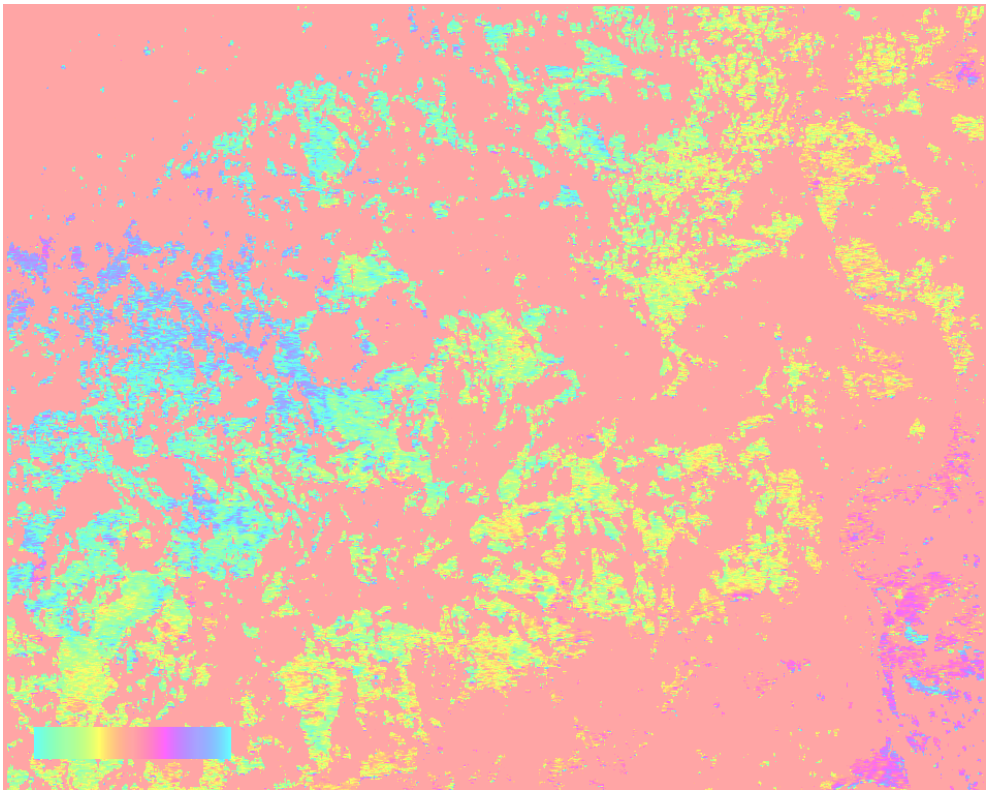

Figure 6: The differential interferogram (only pixels with joint coherence larger than 0.1)

Let's denote here that the look angle is different for each pixel and so are the parallel and perpendicular baseline values.

7 Errors and accuracy

Besides processing errors, there are two main sources of errors introduced during the data acquisition: imprecise orbits and atmospheric delay.

The satellite orbits are usually not very precise although a "precise orbits" procedure was used for computations. The orbit errors are randomly distributed and their values cannot be found. According to [4], "the accuracy of the orbits is approximately 30cm along and cross-track with a slightly higher radial accuracy (8cm)". According to [1], the orbit rms errors of 5 to 10cm correspond to rms erros of 80 to 160 m in the resulting digital elevation model. According to both [4] and [5], precise orbits are not precise enough to allow correct geocoding, and tie points need to be used. Identification of the tie points is not easy in the radar images and their accuracy is not very good — because of the spatial resolution of the image.

An interesting procedure for correcting for the orbit errors is suggested in [9], adjusting the "orbit separation" (perpendicular baseline) according to the number of fringes in the range direction in the image after subtracting the flat-earth phase. First, let's compute the fringes in the range direction on an (azimuthal) border of the interferogram. For each fringe, the distance from the last point of the range profile should be lengthened by $\frac{\lambda}{2}$. The distance from the first point of the profile must be preserved — and the position of the satellite is computed as a 2-dimensional problem (for more details, see [9]). The same procedure shall be performed at the other (azimuthal) end of the interferogram. Although not mentioned, this procedure may be applied only in a flat terrain (relatively to the height ambiguity).

In [1], the authors describe the attempt to "approximate the residual reference phase by a linear and quadratic polynomial". For the quadratic case, the maximum residual phase was about 100 times lower than for the linear case.

The other source of errors is the atmosphere which can generate errors up to 30-35m in digital elevation model (according to [5]) or even up to 100m vertically (according to [1]). For a more detailed description and

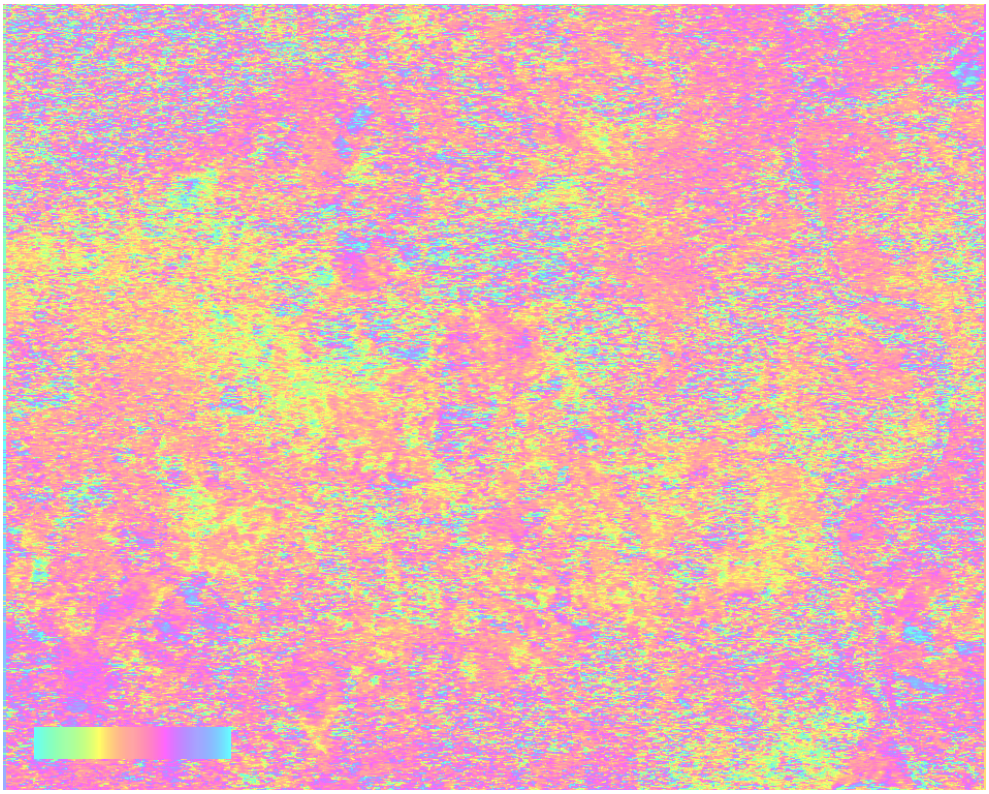

Figure 7: The phase residuals after approximation with the linear polynomial

atmospheric models, see [1]. According to [1], there is no way to correct these errors.

Residuals caused by both orbit errors and atmospheric delay have a long wavelength and for a small area, they can be corrected together. [5] recommends tie points both for atmosphere and orbit errors, for a small region a linear or quadratic polynomial may also be applied.

8 Results

A differential interferogram was generated for the area of Northern Bohemia, topo pair acquired on March 7-8, 1999, defo pair acquired on December 28, 1998, and March 8, 1999 (three-pass differential interferometry). Satellites used: ERS-1 (March 7, 1999), ERS-2 (December 28, 1998, March 8, 1999). Data were provided by ESA. For precise orbit computations, satellite position data from DEOS [8] were used.

Still, precise orbits are quite imprecise, the difference between the shift vector computed from the precise orbits and coarse coregistration was for the defo pair 16 pixels in the azimuth direction and 0 pixels in the range direction, for the topo pair it was 207 pixels in the azimuth direction and 0 pixels in the range direction. Please note that the tandem pair is combined from both satellites; also note that the cross-track accuracy of the orbits is much better than the along-track accuracy.

The interferograms corrected for the reference phase and filtered are shown in figures 1 (topo pair) and 2 (defo pair). The reference phase was computed "exactly", i.e. independently of the computed orbits. The coherence of both interferograms is displayed in figures 3 (topo pair) and 4 (defo pair). Please note that the area of interest is correlated in the topo interferogram but decorrelated in the defo interferogram. The topo interferogram was then unwrapped (using the coherence image) and the differential interferogram (shown in figure 5) was computed, using the wrong baselines.

Also, the joint coherence of the two interferograms was computed (product of the two coherence images). Masked differential interferogram (only pixels with joint coherence larger than 0.1, the other pixels are shown in pink) is shown in figure 6.

Unwrapping was performed OK, so the differential interferogram only contains the deformation residuals, atmospheric delays and orbit errors (no fringes are present). The entire area is "tilted" which may be caused

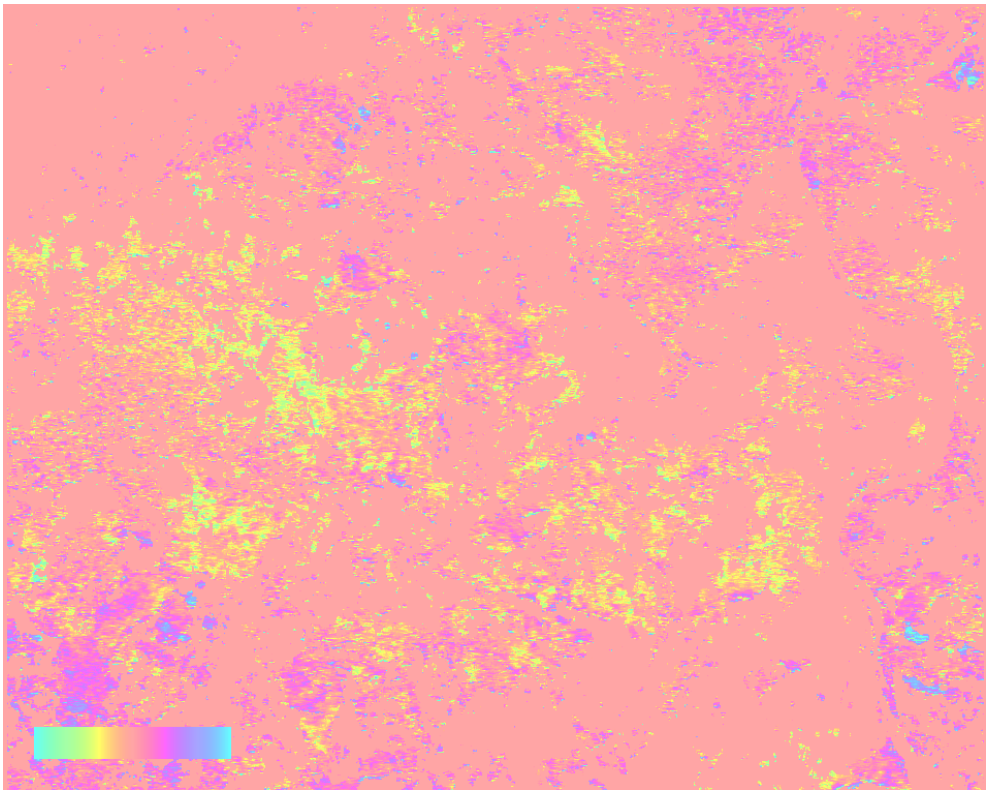

Figure 8: The phase residuals after approximation with the linear polynomial (masked with the 0.08 joint coherence threshold)

by orbit errors (the baseline error for the topo pair is really large) or by atmospheric delay (or both).

The differential interferogram was unwrapped (using the joint coherence image) and an adjustment was performed (using no tie points), the residual phase was first approximated by a plane (with the exclusion of distant observations) — the residual phase is shown in figure 7, or masked in figure 8. During the adjustment, 17,507 out of 722,760 observations were excluded (the criterion was $3\sigma_{apost.}$), $\sigma_{apost.} = 1.23$. Please note that in this image, some areas suspected of landslides can be detected (usually blue color).

Second, phase residuals were approximated by a quadratic polynomial (as recommended in [1]). 17,505 out of 722,760 observations were excluded (the criterion was $3\sigma_{apost.}$), $\sigma_{apost.} = 1.21$. The residual phase is shown in figure 9, masked residual phase in figure 10.

References

[1] Ramon F. Hanssen: Radar Interferometry (Data Interpretation and Error Analysis), Kluwer Academic Publishers, 2001

[2] Dennis C. Ghiglia, Mark D. Pritt: Two-Dimensional Phase Unwrapping (Theory, Algorithms, and Software), John Wiley & Sons, Inc., 1998

[3] Brief Introduction to SAR Interferometry, http://www.images.alaska.edu/documents/pdf/insar_introduction.pdf

[4] D. Small, D. Nüesch: Validation of Height Models from ERS Interferometry, http://www.geo.unizh.ch/rsl/fringe96/papers/small-nuesch/

[5] M. Crosetto: Calibration and Validation of SAR Interferometry for DEM generation, ISPRS Journal of Photogrammetry & Remote Sensing, 57 (2002):213-227

[6] DORIS Manual, http://www.geo.tudelft.nl/fmr/research/insar/sw/doris/Usersmanual/index.html

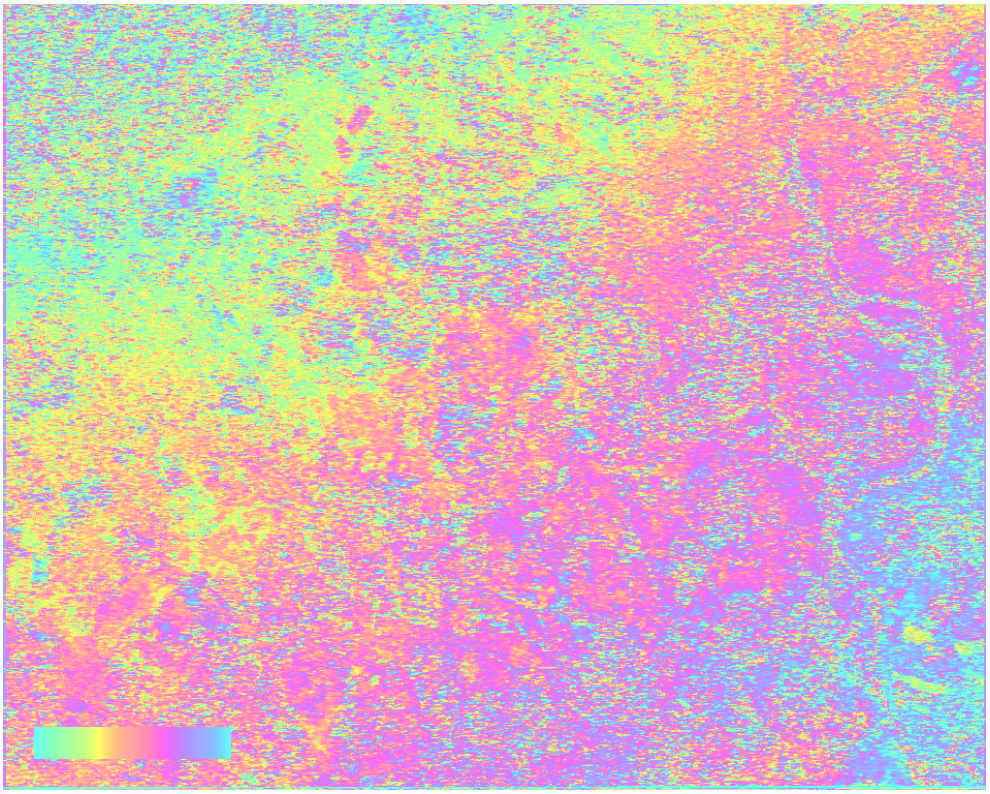

Figure 9: The phase residuals after approximation with the quadratic polynomial

[7] DORIS Home Page, http://www.geo.tudelft.nl/fmr/research/insar/sw/doris/index.html.cgi

[8] Department of Earth Observation and Space Systems Home Page, http://www.deos.tudelft.nl/

[9] D. Massonnet, K. L. Feigl: Radar Interferometry and Its Application to Changes in the Earth's Surface, Reviews of Geophysics, 36(4):441-500

[10] R. Bammler, B. Schättler: SAR Data Acquisition and Image Formation, in G. Schreier: SAR Geocoding: data and systems, pp. 53 – 102, Wichman Verlag, Karlsruhe, 1993

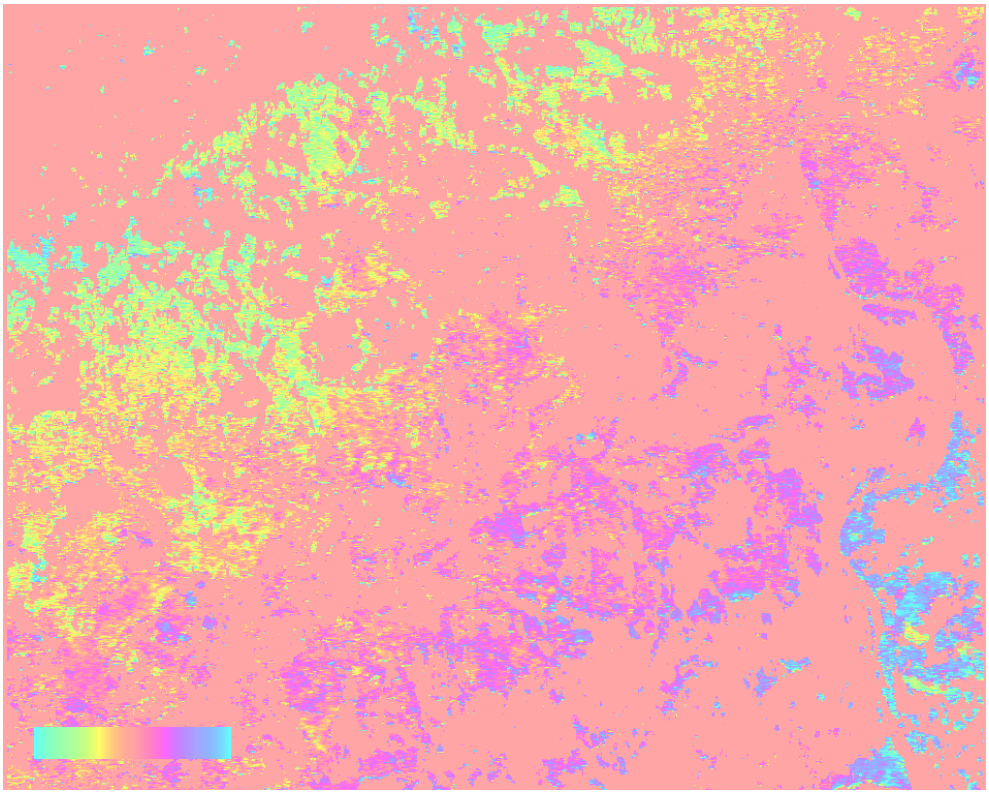

Figure 10: The phase residuals after approximation with the quadratic polynomial (masked with the 0.08 joint coherence threshold)