AUTOMATED AND PRECISE IMAGE REGISTRATION PROCEDURES

U. WEGMÜLLER, C. WERNER, T. STROZZI, AND A. WIESMANN

Gamma Remote Sensing, Thunstrasse 130, CH-3074 Muri BE, Switzerland
Tel: +41 31 9517005, Fax: +41 31 9517008, e-mail: wegmuller@gamma-rs.ch

In this contribution image registration methods are presented and their performance is discussed based on examples. Automated and precise image registration is feasible in many cases. Key elements of the presented registration techniques are the determination of a large number of registration offsets throughout the images, the calculation of a quality measure for each of these estimates, the derivation of a registration function and its quality, and the resampling of the images to the common reference geometry. The registration accuracies achieved clearly meets the "registration error < 0.2 pixel" requirement of SAR interferometry. Application examples include multi-temporal SAR imagery, multi-sensor SAR imagery, optical imagery, registration between SAR and optical data, and between satellite imagery and other spatial map type data, and between SAR imagery and a simulated data set calculated based on a DEM.

1   Introduction

Multi-temporal analysis is one of the most important techniques of remote sensing. Change detection and monitoring applications are directly depending on multi-temporal observations - for other applications as thematic mapping or the retrieval of bio- and geophysical parameters, multi-temporal data allows to enhance the performance by reducing signal noise through multiple independent observations. The use of multi-sensor data, for example SAR and optical data, is another way to increase the information content. Whatever the motivation for the multi-temporal or multi-sensor analysis is, a very important processing step is the registration of the data to a common geometry.

In this contribution two image registration methods are presented. Both methods were originally developed in the context of SAR interferometry. One of the two methods is much more generally applicable. Interferometry requires SAR image registration at sub-pixel accuracy (< 0.2 pixel) in order not to reduce scene coherence. As control point selection is particularly tedious in SAR images because of the image speckle, there was a clear demand for automated and precise registration algorithms. Key elements of the presented registration techniques are the automated determination of a large number of registration offsets throughout the image, the calculation of a quality measure for each of these estimates, the derivation of a registration function and its quality, and the resampling of the images to the common reference geometry. The two methods differ in the registration offset estimation method used. The "coherence optimization method" calculates small interferograms and optimizes its coherence, the "intensity cross-correlation method"

estimates the registration offsets based on the cross-correlation function of corresponding sections of real valued intensity images. The first method requires coherence between at least small sections of the images and is therefore limited to multi-temporal SAR images with a minimum level of coherence. The second method depends on scene contrast and features and is therefore much more widely applicable. The intensity cross-correlation method was found to be very robust for the registration of multi-temporal SAR data of a single sensor. It was also successfully applied to register SAR data from different sensors, optical data, and SAR images with optical images and with landuse inventory data. SAR images were also successfully registered with simulated images calculated based on a digital elevation model, as used to automate terrain corrected SAR geocoding.

2   Concept

The basic concept of the presented image registration procedures consists of 3 main steps. In a first step the geometry used for the registration is selected and all images are brought into this geometry in an initial registration step. In this step it is important to consider the imaging geometry as good as possible; appropriate geometric models for the imaging process are preferred to parametric approaches based on tie points. Repeat-pass SAR images, for example, can be registered in slant range, ground range or map geometry, but a SAR image in slant range geometry should never be registered to an image in ground-range or map geometry. To register images of different sensors a map geometry is usually selected. This means that each image will be geocoded prior to the fine-registration step.

The second step is the derivation of the fine registration function using an automated technique. The steps used in the derivation of this fine registration function are the selection of image chips, the determination of registration offsets for these chips, and a regression analysis on the registration offsets to derive the fine registration function. This step will be discussed in more detail in Section 3.

In a third step the fine registration function is used to resample the data to the common reference geometry. In order to avoid multiple resampling of the data the fine registration function can be used to "refine" the initial registration function, so that the two transformations are conducted in a single resampling step, which allows to minimize problems introduced by the resampling.

3   Registration offset estimation

The key step in the derivation of the fine registration function is the automated estimation of reliable registration offsets. To meet the reliability requirement, quality control is conducted for each single offset estimate. Lower quality estimates have to be rejected and, in addition, a relatively large number of reliable estimates is

required to minimize statistical errors. Fine registration functions with few free parameters to be determined in the regression analysis are preferred.

To meet these requirements a large number of image locations are automatically selected for the registration offset estimation. In our implementation the selected m x n locations are on a regular grid. For each selected location a registration offset is estimated and for each estimate the quality is determined. The two main methods used for the registration offset determination are the intensity cross-correlation method and the coherence optimization method which are both described in more detail below. The estimated offsets and the corresponding quality measures are the input to the regression analysis. Estimates with a quality measure above a specified threshold, are used to determine the parametric fine registration function using a least squares technique. In the case of images which are already in the same image geometry bi-linear functions are usually sufficient, but in principle more complicated functions could also be used. In addition to the registration function the standard deviation of the estimates from this function is calculated and used for quality control and to iteratively reject estimates which deviate more than a specified multiple (for example 3 times) of the standard deviation from the registration function.

Very often more than one hundred estimates with a standard deviation below half a pixel are used in the final determination of the registration function, resulting in a very small statistical error of the registration function of less than a tenth of a pixel.

3.1 Intensity cross-correlation method

For the determination of an individual registration offset a small section of one real valued image is selected and cross-correlated with the other real valued image. The maximum correlation found, or more precisely the maximum of the two dimensional correlation function determined from the correlation values in the proximity of the correlation maximum, determines the offset estimate. The width of the correlation function determines the quality, with narrow correlation functions corresponding to high quality and wide correlation functions to low quality. The intensity cross-correlation method depends on sufficient image contrast. The size of the image sections used for the intensity cross-correlation method should be larger than the typical dimension of features in the image. In practice we typically use image chips of 32 x 32 pixels to 256 x 256 pixels.

3.2 Coherence optimization method

The coherence optimization method is limited to "interferometric SAR data". With interferometric SAR data we mean SAR acquisitions with a significant level of coherence. As in the intensity cross-correlation methods small section of one image

is selected. But instead of optimizing the intensity cross-correlation with the other image the degree of coherence of small interferograms is optimized. The maximum coherence found, or more precisely the maximum of the two dimensional coherence function determined from the coherence values in the proximity of the coherence maximum, determines the offset estimate. The width of this function determines the quality of the estimate, with narrow functions corresponding to high quality and wide functions to low quality. The coherence optimization method depends on sufficient coherence at least for parts of the images. The size of the image sections used for the coherence optimization method can be small. In practice, we use image chips of 4 x 4 pixels to 16 x 16 pixels. In order to be able to calculate the small interferograms and determine its average coherence level single look complex (SLC) data are required.

3.3   Optimization

The presented registration methods offer some flexibility for optimization for a specific example. Several parameters, as for example the size of the image chips used to determine the individual offset estimates, quality thresholds to reject estimates as unreliable, the number of locations included in the analysis, may be adapted to the specific case. To achieve a high quality is only one criteria considered in such an optimization. The robustness of the methodology and its efficiency are also very important. Typically, the objective is not to achieve the highest possible quality but to reliably achieve a high enough quality in a short time.

The size of the image chips used has a strong effect on the quality of the individual estimate as well as on the computational efficiency. The selected chip size needs to be a minimal size to achieve the required quality. This minimal size depends on the data type and the specific example. A good measure of the accuracy of the individual estimates is found in the standard deviation of these estimates from the parametric registration function. Using larger chips also increases the chance for reliable estimates and therefore the robustness of the method. The number of chips selected needs to be high enough to achieve a sufficiently high number of accepted estimates. The spatial distribution of these estimates is another important factor for the accuracy of the registration.

The computational effort depends not only on the size and number of the image chips used but also on the size of the search region in the second image. This search region needs to be large enough that it contains the section corresponding to the search chip, but it should be as small as possible to reduce the computational effort. A good solution is a guided search. An initial registration function is used to pre-determine an expected location. This initial registration functions is typically determined in a first run with fewer, but larger image chips and search regions. The main run is then conducted for a larger number of locations but with smaller chips and search regions.

With our implementation (Werner et al., [1]), a registration is conducted for both methods without operator interaction in a reasonably short time (~ few minutes) on normal PCs or workstations.

Oversampling of the image chips is a possibility to further increase the accuracy of the individual offset estimates. This is important mainly if the use of the individual estimates goes beyond the determination of a parametric function. An example for such use is the determination of surface deformations based on the deviation of the individual offset from an average registration function (Werner et al., [2], Strozzi et al. [3]).

4   Examples

4.1   SAR interferometry

The registration requirements in SAR interferometry are high. Registration errors larger than about 0.2 pixel cause significant additional phase noise and reduce the coherence level. For this case both presented methods are applicable, allowing to validate their accuracy by cross-comparison of the results. Both methods were successfully applied to thousands of image pairs, very often as part of fully automated processing sequences. Typically, more than 100 individual estimates are determined with a standard deviation from the parametric registration function between 0.05 and 0.3 pixels.

The coherence optimization method is applied with small estimator window sizes. Therefore, the spatial resolution of the estimates is higher. This is only a relevant advantage, though, if there are inhomogeneities with higher spatial frequencies, as caused for example in the case of uncompensated terrain effects or surface displacement as in the case of moving glaciers or earthquakes. The coherence optimization method does not depend on image contrast, so it also works in the case of very homogeneous surfaces. On the other hand it depends on a sufficiently high level of coherence, which strongly reduces the number of accepted estimates for longer acquisition intervals and data with larger spatial baselines. In general we prefer the intensity cross-correlation method for its higher robustness.

4.2   Multi-temporal SAR

The main difference between the multi-temporal SAR case and SAR interferometry is that coherence cannot be expected, in general. Therefore, mainly the intensity cross-correlation method is used, which works again very well for repeat-pass ERS, JERS, and Radarsat data. The accuracy achieved is only slightly below the one in interferometric pairs. Reasons for some reduction of the accuracy are longer acquisition intervals, which lead to more temporal change in the scene) and larger differences in the initial geometries (as the data selection has not to be restricted to

small baselines as in the interferometry case), which results in a higher spatial dependence of the registration offsets. Again, many individual estimates with a small standard deviation from the parametric registration function (0.05 and 0.5 pixels) are typically found.

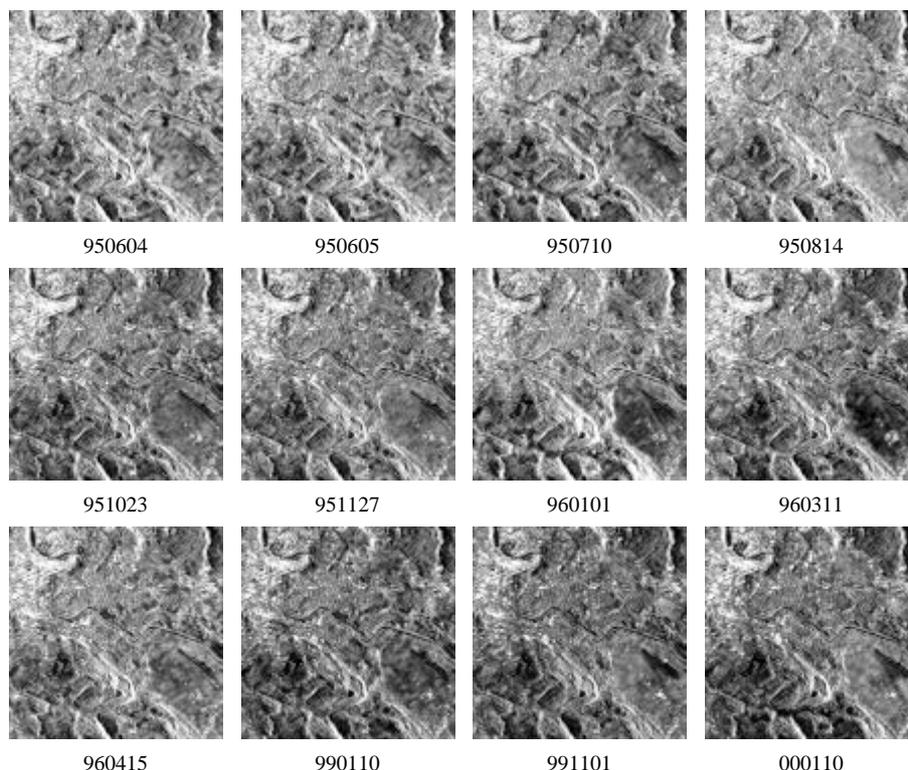|			
950604	950605	950710	950814
951023	951127	960101	960311
960415	990110	991101	000110

Figure 1 Section (320x320 pixels; approx. 6.4 km x 6.4 km) of the twelve repeat-orbit multi-temporal ERS SAR images over Berne (upper left corner with its "old town" within Aare river loop), which were used for the registration experiment.

To investigate the robustness of the registration of multi-temporal SAR images a small registration experiment was conducted on twelve ERS SAR backscatter images data (5-look intensity images) acquired in repeat-orbits over Berne, Switzerland (see Figure 1). This time series includes pairs with very short (Tandem pair with 1 day difference) to very long (4.5 years) acquisition intervals. The selected images cover different seasons and include one image (960101) with wet snow coverage which significantly changes the backscatter levels and image contrast. The intensity cross-correlation method with an oversampling factor of 2 was used for the image registration.

With a correlator size of 64x64 pixels a consistent high registration accuracy was achieved. From the 1024 cross-correlation analyses resulted 741 to 960 accepted estimates with standard deviations of the range and azimuth offset estimates from the regression functions between 0.041 and 0.123 pixel. This means that each image pair was successfully registered at high accuracy. As expected the standard deviations are the smallest for the Tandem pair, and larger for intervals of 35 days to several years. There is not a uniform time dependence, scene parameters as for example the wet snow cover on 960101 are very important factors. The detailed results are shown in Table 1.

Table 1: Registration of multi-temporal ERS. For 1024 evenly distributed locations the intensity cross-correlation approach with a 64 x 64 pixel window and an oversampling factor of 2 was used to estimate individual registration offsets. The number of valid estimates (first row), the standard deviation of the range offset in pixels (second row) and the standard deviation of the azimuth offset in pixels (third row), both relative to the determined regression function, are indicated. The image pairs are defined by its acquisition dates (indicated in the format yymmdd).

	950605	950710	950814	951023	951127	960101	960311	960415	990510	991101	000110
950604	960 0.041 0.043	950 0.075 0.069	874 0.082 0.074	917 0.085 0.076	923 0.089 0.075	832 0.114 0.099	844 0.084 0.078	914 0.098 0.086	854 0.092 0.084	824 0.097 0.088	784 0.117 0.088
950605		955 0.076 0.074	863 0.085 0.073	911 0.084 0.079	924 0.087 0.078	850 0.121 0.108	830 0.089 0.083	903 0.091 0.087	864 0.092 0.087	808 0.098 0.092	790 0.123 0.086
950710			902 0.085 0.078	897 0.070 0.067	852 0.107 0.091	813 0.091 0.081	792 0.082 0.077	887 0.094 0.080	897 0.104 0.088	773 0.121 0.087	741 0.082 0.074
950814				905 0.078 0.079	891 0.081 0.077	831 0.103 0.091	860 0.080 0.087	913 0.088 0.076	872 0.088 0.082	864 0.086 0.090	822 0.106 0.088
951023					922 0.083 0.076	877 0.100 0.088	895 0.077 0.073	907 0.079 0.072	918 0.088 0.082	859 0.097 0.085	801 0.111 0.090
951127						861 0.096 0.098	897 0.073 0.078	918 0.075 0.073	841 0.097 0.085	864 0.091 0.082	773 0.085 0.082
960101							810 0.091 0.102	869 0.092 0.092	845 0.121 0.098	792 0.121 0.105	782 0.086 0.097
960311								898 0.070 0.075	825 0.092 0.084	831 0.090 0.094	752 0.092 0.098
960415									875 0.094 0.083	844 0.098 0.083	789 0.099 0.085
990510										917 0.079 0.072	880 0.107 0.084
991101											884 0.087 0.078

To evaluate the influence of the correlator window size the experiment was repeated with window sizes of 16x16, 32x32, 64x64, 128x128, and 256x256 pixels. During the registration we kept track of the number of valid offset estimates found and of the standard deviation of the estimated range and azimuth offsets from the regression functions. It was found that the registration always succeeded for this data set, even with the smallest window size. For the small window sizes the acceptance rates were significantly lower. In spite of this or more precisely because of this the standard deviations remained quite low, below 0.2 pixel. For the small window sizes significantly more estimates were accepted for the Tandem data. The reasons are that there is very little temporal change for the Tandem pair and that the speckle are very similar for the Tandem data due to the high degree of coherence. Figure 2 shows the values for the Tandem pair, the average of two 35 day interval pairs, and the average of the 3 pairs with acquisition intervals of several years. The corresponding azimuth offset estimation standard deviations are very similar to the values in range direction. For the largest windows the processing efficiency gets very low. As a good compromise for efficient and robust registration a window size of 64x64 or 128x128 can be recommended.
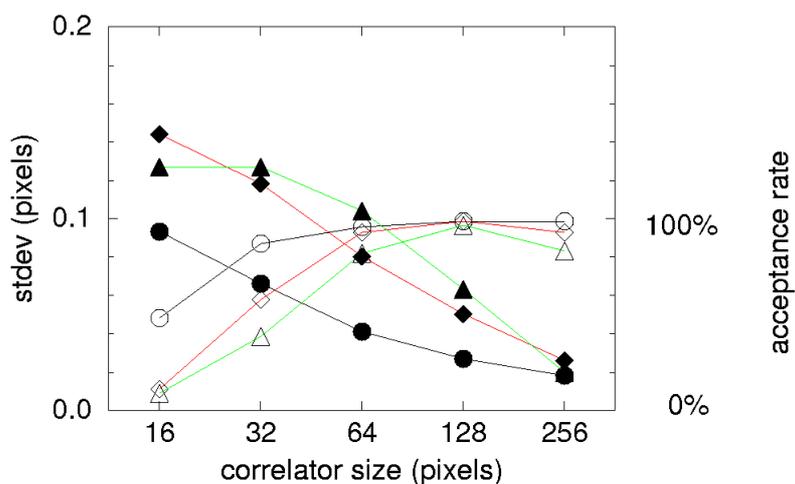

Figure 2 Offset estimate acceptance rate (opaque symbols) and range offset estimation standard deviation (filled symbols) as a function of the estimator window size, for an ERS Tandem pair (circle), 35 day pairs (diamond), and multi-year pairs (triangle) for intensity cross-correlation method applied to 5-look ERS backscatter images.

4.3 Multiple SAR sensor

The registration of SAR images from different sensors differs from the registration of multi-temporal images acquired with just one sensor in the sense that imaging parameters as the frequency, incidence angle, track orientation, and spatial resolution, differ. As mentioned in the concept section the images need to be brought in the same geometry. Typically, a map geometry is selected in this case. Each of the images is first geocoded to the map geometry. Terrain corrected geocoding is, of course, more accurate than just ellipsoid corrected geocoding and is therefore preferred if a digital elevation model (DEM) is available. Once in the same geometry the automated fine registration can be applied.

In such a way a large number of JERS and ERS scenes were successfully registered (Wiesmann et al., [4]). The accuracy achieved depends on the accuracy of the geocoding step as well as on the correspondence of specific features within the images of the two sensors. For the example JERS / ERS everything worked fine, in spite of the very difficult appearance of the images. The standard deviations of the fine registration offsets were usually between 0.5 and 1.0 pixel.

The same technique is also recommended for the registration of images acquired by the same SAR sensor but at different incidence angles, as necessary when working with Radarsat or ENVISAT ASAR images of different modes or in ascending and descending orbits.

4.4 SAR geocoding

The intensity cross-correlation method is also used in the frame of the automated SAR geocoding methodology described by Wegmüller et al., [5]. In this case a simulated SAR backscattering image, calculated based on a DEM and an assumed incidence angle dependence of the backscattering coefficient, is used as the image in the reference geometry. This simulated image does not contain any dependence on the surface type. In spite of this limitation automated fine registration with a real SAR image is feasible as soon as there is some topography in the scene. In this case somewhat larger chips (128 x 128 to 1024 x 1024) are used for the estimation of the individual offsets (Wegmüller et al., [5]).

4.5 Optical and SAR data

The number of optical data registration examples investigated so far with our implementation is much smaller than for the SAR data, simply because of the focus of our work on SAR data. Examples with SPOT and Landsat TM data confirmed nevertheless, that the intensity cross-correlation method is also applicable to the registration of optical data.

To assess the potential more quantitatively we used a data set over Switzerland consisting of a Landsat TM image, channels 2, 3, and 4, an ERS SAR backscatter

image, an ERS Tandem coherence image, and a landuse inventory. All the data sets were geocoded to the same map projection with a pixel spacing of 25m and a total size of 2000x2000 pixels. The landuse inventory data is based on the Swiss "Areal Statistik" which has an original pixel spacing of 100m. The data were oversampled to 25m. In addition, the inventory was reduced to the three classes forest and water, urban, and others. In the optical and in the SAR data forests and water appear in dark gray levels. Therefore, a dark gray level was attributed to the forest and water class, an intermediate level to the others class and a bright level to the urban class. To the SAR backscattering logarithmic scaling was applied. To the optical channels and the coherence image linear scaling was applied. The resulting images which were used for the SAR – optical registration analysis are shown in Figure 3.
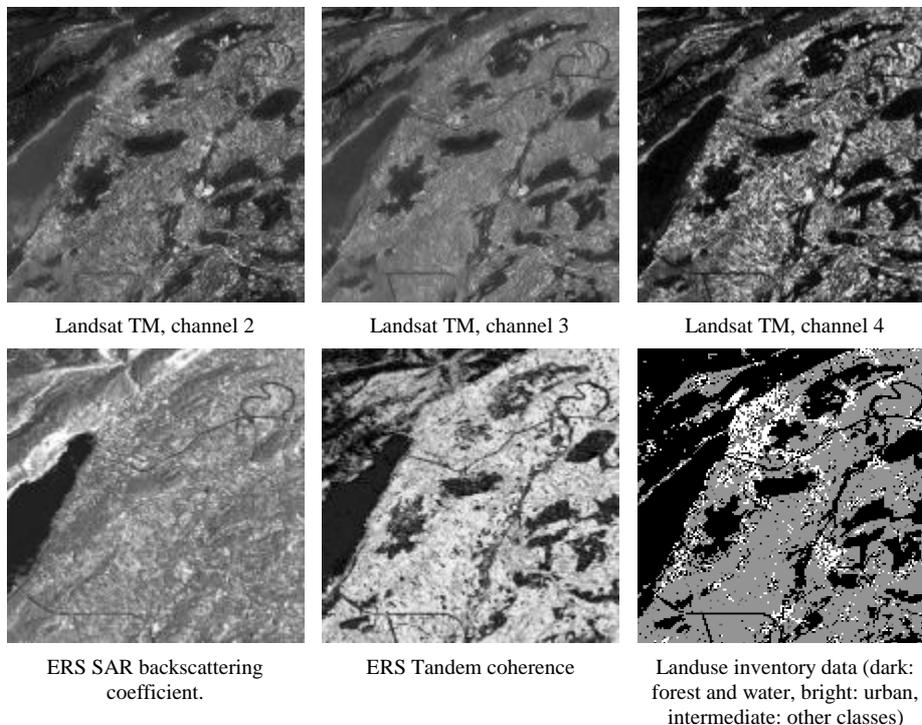

| Landsat TM, channel 2 | Landsat TM, channel 3 | Landsat TM, channel 4 |

| ERS SAR backscattering coefficient. | ERS Tandem coherence | Landuse inventory data (dark: forest and water, bright: urban, intermediate: other classes) |

Figure 3 Section (640 x 640pixels, $16km$ x $16km$) of geocoded images used for the SAR – optical registration analysis.

Table 2: SAR – optical data registration experiment. Easting (diamonds) and Northing (triangles) offset estimation standard deviations versus estimator window size, for Landsat TM image, channels 2, 3, and 4, an ERS SAR backscatter image, an ERS Tandem coherence image, and a landuse inventory data set.

	ERS σ⁰	ERS coherence	Inventory
TM.2			
TM.3			
TM.4			
ERS σ⁰			
ERS Coher-ence			

Table 3: SAR – optical data registration experiment, as Table 2, but for Landsat TM image, channels 2, 3, and 4.
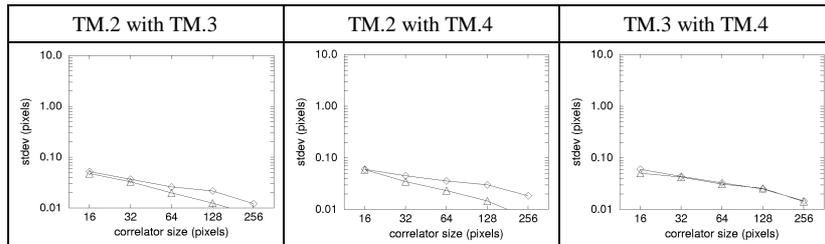
TM.2 with TM.3	TM.2 with TM.4	TM.3 with TM.4

The intensity cross-correlation method with an oversampling factor of 2 was used for the image registration. To evaluate the influence of the correlator window size the registration was repeated with window sizes of 16x16, 32x32, 64x64, 128x128, and 256x256 pixels. The standard deviation of the estimated registration offsets in Easting and Northing from the regression functions are shown in Tables 2 and 3 as a function of the correlator window size. For unsuccessful registration attempts, these are attempts with too few accepted offset estimates and attempts with standard deviations much larger than 1 pixel, no value was plotted for the corresponding correlator window size.

The registration between the different optical channels (Table 3) was successful with very low standard deviations even for the smallest window sizes considered. This is not astonishing as the three optical images are just different channels of the same acquisition. Consequently, this result is not representative for a registration between multi-temporal optical data.

The registration between the Landsat TM channels (Table 2) and the SAR backscattering image was only successful with larger correlator windows. This is not surprising when considering how different the images look. The standard deviations achieved were of the order or 1 pixel in Easting direction and below 0.5 pixel in Northing direction. Our explanation for the higher variability in the Easting direction are related to the special SAR imaging effects at the near and far range forest boundaries (corresponding closely to the eastern and western boundaries). The registration between the Landsat TM channels and the Tandem coherence map was found to be more robust than the registration with the backscatter image. The explanation of this is the quite high contrast between forest and non-forest in the coherence image. The standard deviations were similar as for the backscattering images but based on more accepted estimates. The standard deviation was again higher in Easting direction, for the same reason.

In addition, to satellite imagery the modified landuse inventory data set was considered. Both the optical and the SAR data registered successfully but more robustly with larger estimator windows, as can be expected for this data set based on 100m resolution data. This type of registration can be used to automate the

refinement of the geocoding of SAR or optical imagery with map type data sets used as reference.

5 Conclusions

Automated and precise image registration is feasible in many cases. Two specific algorithms were presented. The coherence optimization method is predominantly for application in the context of SAR interferometry. The registration accuracy achieved clearly meets the "registration error < 0.2 pixel" requirement of SAR interferometry. The intensity cross-correlation method results in similarly accurate results but it is more robust and has a very wide applicability. Application examples include multi-temporal SAR imagery, multi-sensor SAR imagery, optical imagery, as well as registration between SAR and optical data, between satellite imagery and spatial data sets as a landuse inventory, and between SAR imagery and a simulated data set calculated based on a DEM.

6 Acknowledgment

7 References

1. Werner C., U. Wegmüller, T. Strozzi, and A. Wiesmann, "*Gamma* SAR and Interferometric Processing Software", Proceedings of ERS-ENVISAT Symposium, Gothenburg, Sweden, 16-20 Oct. 2000.
2. Werner C., T. Strozzi, A. Wiesmann, U. Wegmüller, T. Murry, H. Pritchard, and A. Luckman, "Complementary measurement of geophysical deformation using repeat-pass SAR", Proceedings of IGARSS 2001, Seattle, Australia, 9-13 July 2001.
3. Strozzi T., A. Luckman, T. Murray, U. Wegmüller, and C. Werner, Glacier motion estimation using SAR offset tracking procedures, submitted to IEEE TGRS in 2001.
4. Wiesmann A., T. Strozzi, and U. Wegmüller, JERS SAR and Interferometric Processing for the boreal forest mapping project SIBERIA, Proceedings of IGARSS'99, Hamburg, 28 June - 2 July 1999 (DD01_03).
5. Wegmüller U., Automated terrain corrected SAR geocoding, Proceedings of IGARSS'99, Hamburg, 28 June - 2 July 1999 (CC03_02).