Chapter 8

Image Registration

For many remote sensing applications it is beneficial to combine the information from various types of images. This is only possible if all images can be warped into a common coordinate system. The purpose of image registration is to find and apply the transformation that allows for such a warping. The aim of registration is thus to find a geometrical transformation that maps one image on the other(s).

Many image registration methods are based on the detection of common points or objects in the different images to be registered. These common objects are called matching primitives. Registration methods based on matching primitives consist of the following steps:

- extract matching primitives from each image
- match corresponding primitives from each image
- determine a transformation that maps the primitives from one image to the ones in the other image
- apply the transformation

The development of a registration method involves choices for the following four components [73]:

- a feature space: the primitives extracted from the images that will be used for the matching
- a search space: the type of transformations that will be applied to align the images
- a search strategy: the method used to choose the next transformation from this space
- a similarity metric: a measure of the relative merit of each transformation

The choice of these four components method depends on several factors. The *choice of matching primitives* (the feature space) is determined by the following properties:

- Similarity: the matching primitives should be invariant between the 2 images to be registered

- Geometric Distribution: the primitives should be geometrically distributed in such a way that they cover the whole region of interest
- Selectivity: if the characteristics of the primitives are highly selective (unique) the matching is simplified
- Density: too many matching primitives result in higher calculation time for the matching and increases the probability that two characteristics are confused. On the other hand, if too few primitives are available, the sensitivity to noise and inconsistencies increases.

The *choice of the search space* mainly depends on the difference in image acquisition geometry between the two images as well as on the geometry of the imaged terrain. It is clear that a different type of transformation will need to be used when registering two images taken by the same sensor from a slightly different viewpoint (e.g. in stereoscopy) than when registering a SAR image with an optical image or a map.

The *choice for the search strategy* depends mainly on the choices made for the two previous steps but also on another important factor that is the level and accuracy of available prior knowledge. For a satellite or airborne platform in many cases the platform's navigational parameters and the image acquisition parameters are known within some accuracy. Using these parameters considerably facilitates the registration process as they provide an initial estimate of the transformation parameters. The actual registration is then a fine-tuning of these parameters.

8.1 Evaluation Method for the Registration

For the evaluation of the registration a set of corresponding control points is manually selected on the images to be registered (or the image and the map). The transformation that is found by the automatic registration is then used to warp the control points from the first image into the second and the distance between these warped points and the control points in the second image is used to define the evaluation criterion for the quality of registration.

Let $P^m(i), i = 1..N$ be the control points in the first image and $P^s(i), i = 1..N$ the corresponding points in the second image. Φ is the transformation resulting from the automatic registration. For each point in the set of control points the distance is then calculated between the $P^s(i)$ and the warped point $\Phi(P^m(i))$:

$$\forall i = 1..N \quad d_i = Dist(P^s(i), \Phi(P^m(i))). \tag{8.1}$$

Note that this definition of distance is independent of the type of transformation that is used in the registration. If one is interested in finding out how well the automatic registration method is able to find the parameters of a given type of transformation, this transformation could be calculated from the control points and the result compared with the automatically found transformation [65].

The three measurements used to estimate the quality of the registration are the maximum, the average and the root-mean-square of this distance:

$$D_{Max} = \max_{i=1..N} d_i, \tag{8.2}$$

$$D_{Mean} = \frac{1}{N} \sum_{i=1}^{N} d_i, \tag{8.3}$$

$$D_{RMS} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} d_i^2}. \tag{8.4}$$

8.2 Development of a Strategy for Registering SAR Images

The purpose of the image registration described in this chapter is to warp the different images into a common coordinate system. The idea that is pursued in this chapter is to register each image first with a map, i.e. into a coordinate system that is independent of any single image, and then use this as a starting point for the actual registration between images. In this work we have used topographic maps at a scale 1/25000. Registration with a map has several advantages:

- it automatically results in a geocoded image;
- it is a first step for automatic updating of existing maps;
- the information on the map can be used to enhance the results of the image interpretation;
- the information from the map can be used for selecting the primitives to be used for the registration (e.g. it only makes sense to use the position of forests as a matching primitive if there are forests in the region of interest);
- for any type of image it should always be possible to identify some types of objects that are present on the map and automatically detectable on the image. The types of objects used for registering various types of images with the map do not have to be the same.

However, when registering an image with a map several problems can arise:

- inconsistencies can exist between the information present on the map and on the image;
- the same object can appear very differently on the map and the image;
- the map and image do not in general cover exactly the same area;
- the accuracy of localisation is different in the map and the image;
- some objects are moved on the map for better display.

Furthermore it is possible that a given object on the map is detected in the SAR images as a set of separate objects or vice versa. This means that the registration can not rely on a one-to-one relationship between the primitives on the map and those found in the image.

In fig. 8.1 some of the problems mentioned above are illustrated. A visual image of the area around the airfield of Oberpfaffenhofen was manually registered with a map at a scale 1/25000 of the same area. The image was acquired in 1997 while the map was printed in 1992. An affine transformation was used for the registration.

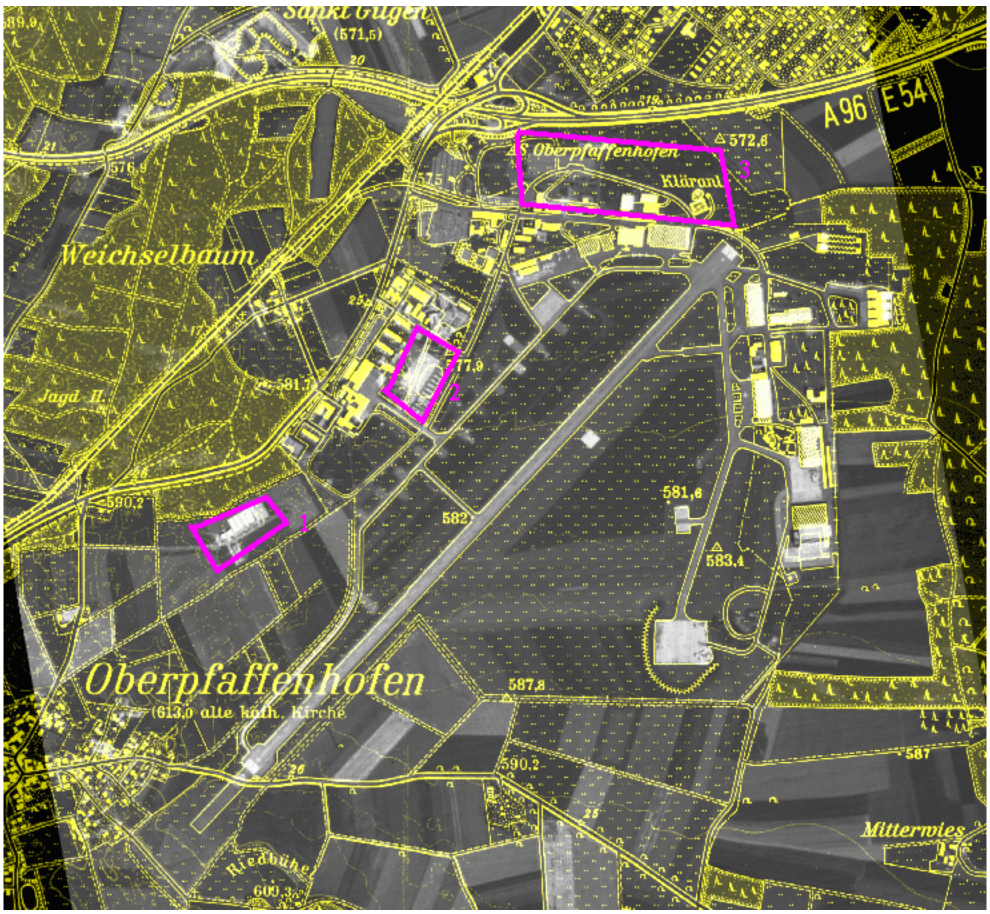

Figure 8.1: Map objects projected on a visual image

Fig. 8.1 shows the symbols from the map overlayed on the visual image. The three indicated regions show inconsistencies between the image and the map. The corresponding part of the map and the image are shown in figs. 8.2 and 8.3 in two separate images.

In region 1 some new buildings and a road have been constructed. In region 2 a road seems to have moved and the same has happened in region 3 where also some new buildings have appeared. The overall accuracy of the map seems satisfactory though, i.e. the roads that are present on both the map and the image seem to overlap quite well. This is not necessarily the case [74]: the roads on a map can be moved (e.g. when too close to a more important feature such as a river or a railway) and they can be simplified (e.g. in mountaneous regions roads with a lot of curves are simplified).

Figure 8.2: Visual image

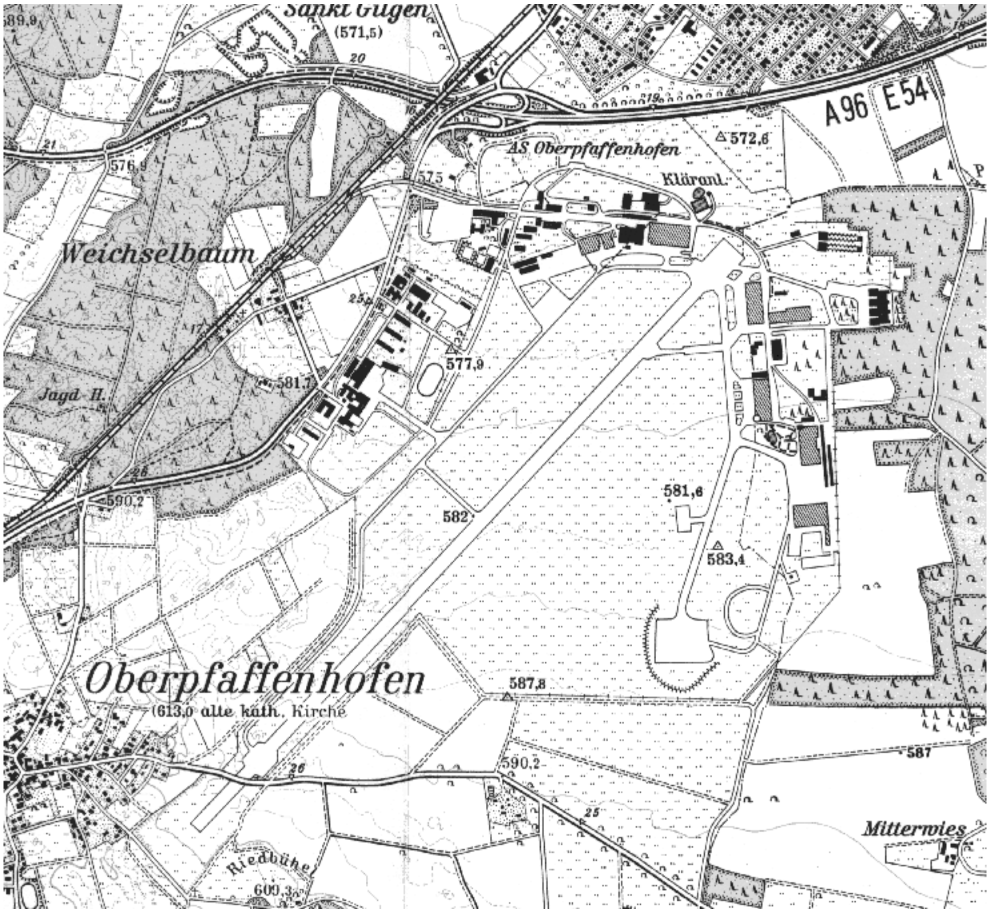

Figure 8.3: Part of map corresponding to visual image

Using a map for registration thus presents several advantages but it is unlikely that a very accurate registration between a map and an image can be achieved in all cases. The map can only be used as a first step and it is worthwhile checking its accuracy. A map, even at a scale 1/25000 has a limited accuracy. Visual images with an appropriate resolution provide a much better localisation of objects and furthermore they contain more information that can be used for registration, such as contours of fields, that is not available on a topographic map. A visual image could thus be used to obtain a better registration.

When using a map in the registration process we should be aware of the kinds of problems mentioned above. The registration should thus be able to cope with inconsistencies and it is very difficult to match the landmarks found on the map with those found on the image a priori.

In this section we will try to find a good combination of the four components of a registration method for registering a SAR image with a map.

8.2.1 The Feature Space

Evidently the appearance of the objects on a map is completely different from their appearance on a SAR image. It is therefore impossible to use pixel-based features. We need to extract similar objects from both the map and the image and use some properties of these objects to construct a feature space for the matching. Features that come into mind are the location (centre) and contours of forested and built-up areas and the network of communication lines (roads, railways and rivers). Using forests and built-up areas only provides a rough registration. For the forests this is primarily due to problems of layover and shadowing. For the built-up areas the main problem is the delimitation of such areas. The communication lines should be used to refine the registration.

Once a registration between each image and the map is found, the resulting transformation is used to determine a rough registration between the different images. This registration will need to be refined using features that

- are likely to be present on different images (even non-SAR images),
- have a limited 3D structure,
- are dense enough to allow, if necessary, to find local transformations.

Edges detected in the different images are good candidates for this refinement step.

The exact form of the features that are used, are chosen in function of the search strategy.

8.2.2 The Search Space

The geometry of a map is completely different from that of a SAR image. However, the information about the SAR image acquisition can be used to convert the SAR image into a geometry that is more similar to that of a map. This is done in two steps. In the first step the slant range image is transformed into ground range. Then the information about spatial resolution is used to transform the image into square pixels. After this first transformation, which is completely controlled by a priori information, a global polynomial transformation should provide a first registration between image and map. For a flat scene, if the SAR image has square pixels, a similarity transform (see below) should be sufficient. Otherwise a semi-affine or affine transformation is more appropriate. In scenes with a more complex 3D structure, an affine transformation will be inadequate and higher order polynomials or local transformations will be necessary.

8.2.3 The Search Strategy

For the type of features and transformations defined above, point-mapping or landmark mapping techniques are most appropriate [73]. They consist generally of three stages:

- Computation of the feature points
- Matching of the feature points
- Determining a spatial mapping between the matched sets of feature points

For landmarks that are determined automatically, in some cases matches can be determined based on the properties of these points, such as curvature, principal axis,...

For obtaining the first rough registration between the SAR image and a map we have selected built-up areas and forests as matching primitives. These types of primitives have the advantage of being relatively randomly distributed in a given region but the disadvantage of changing rapidly. The apparent borders of forests can change rapidly when forests are cut or old trees are replaced by young ones and the borders of built-up areas can also change when more buildings are erected. It is thus very difficult to find properties of this type of regions that are specific and reliable enough to obtain an a priori matching. The size of the regions is relatively stable and we can hope that the same is true for their relative position. However, neither of these two characteristics is sufficiently selective to allow an a priori matching between the objects on a map and on an image.

A way to deal with such a situation where it is not possible to match the primitives a priori is to combine the matching with the determination of the optimal global transformation in a single step. Several methods along this line were presented in literature. They involve clustering, relaxation, iterative methods, or step-wise methods. An example of an iterative method is the hypothesis generation/verification method presented in [65]. We have used a step-wise method called *feature consensus* [75] which is explained in detail in section 8.4.1.

The result of the feature consensus method is a first registration between map and SAR image. This registration can be refined using the contours of forests, which will be explained in section 8.4.2. Because the forest contours can change quickly over time and because they are not well-localised due to geometrical distortions, it is better to use the network of communication lines for refining the registration. This is explained in section 8.4.3.

For registering different images, once an affine transform mapping them to the same map is known, edges, preferably between fields or other objects with a low 3D structure, can be used. It will then only be necessary to find corresponding edges in a small neighbourhood and to refine the already available registration.

The global strategy used for registering SAR images with a map is schematically presented in fig. 8.4. The refinement using contours of forests is optional. The conversion from slant range to ground range and the resampling to approximately square pixels is performed in a pre-processing step that relies only on a priori information about the image acquisition. This pre-processing is explained in the next section. The following sections respectively present the rough registration based on the feature consensus method and the refinements of the results using contours of forests and the network of communication lines.

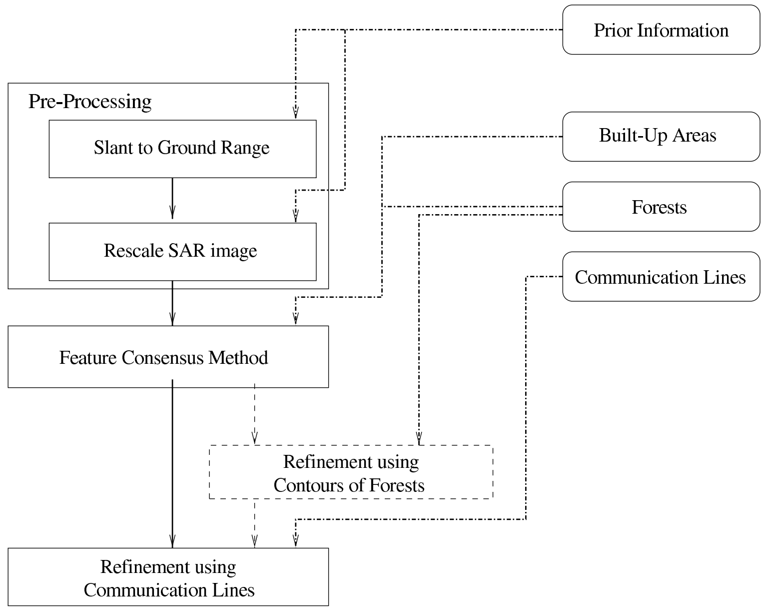

Figure 8.4: Registration of SAR images with a map: Overview of the method

8.3 Pre-Processing based on A Priori Information

The pre-processing includes the conversion of the SAR image from slant range to ground range and its resampling to (almost) square pixels. Both steps are solely based on prior information about the image acquisition. They do not involve any matching process or an iterative search for some optimal transformation. This is the reason that we call it pre-processing.

The conversion to ground range is based on the assumption of a flat terrain. At this point of the registration process it is not possible to incorporate any available terrain elevation data because we have not yet geocoded the image.

In order to facilitate the feature consensus method, the SAR image is also resampled such that the pixel spacing is approximately equal to the pixel spacing in the map.

8.4 Registration of SAR Images with a Map

8.4.1 Registration by Feature Consensus

The method used to find a first registration based on the position of built-up areas and forests is based on a technique called feature consensus that was first described in an article by Shekhar, Govindu and Chellappa [76] and refined by the same authors in [77, 75]. Their approach does not require feature correspondence or area correlation. The geometric transformation to be used is reparametrised into a sequence of elementary stages. At each stage a single transformation parameter is estimated from the available features by a method called feature consensus. The method consists in choosing the parameter that is most consistent with all possible pairs of features, i.e. for each feature pair, the transformation that would map them onto each other is determined. This means that the parameter must be "visible" from the features that are used, e.g. when finding a rotation angle, it is not possible to use single points as features; one should at least use line segments.

For all true feature pairs, the estimated parameters should cluster around the true value of the parameter while for all other pairs, parameters will be randomly distributed. The parameter is then chosen by selecting the maximum of the distribution of the parameters. The distribution for a given parameter is called the *consensus function* for that parameter.

The method allows to find a global transformation only if it can be reparametrised as a sequence of elementary transformations. Even if this limits the possible choices for the transformation, it can be used as a first step, yielding a rough registration and, perhaps more importantly, yielding the correct feature pairs. The resulting feature pairs can then be used to find the parameters of more adequate transformations. While the method is reported to be robust against inconsistencies between features found in both images, it does present problems when too many features are available. In that case, the distribution can have several peaks or even become flat. In order to detect this type of problems the authors [75] defined two measurements of the quality of parameter estimate. A method of progressive feature filtering combined with a tree descent along peaks of progressively smaller importance is proposed to avoid these problems.

Similarity Transform

This transformation is characterised by four parameters (rotation angle β, translation t_x, t_y and scale s). Under this transformation a point x,y in the first image maps to the point X',Y' in the second image according to:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = s \begin{bmatrix} cos\beta & -sin\beta \\ sin\beta & cos\beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \tag{8.5}$$

This transformation can be rewritten as a sequence of 4 stages:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix} = T_{t_x} T_{t_y} T_s T_\beta \begin{pmatrix} x \\ y \end{pmatrix}. \tag{8.6}$$

The first parameter to be estimated is the rotation angle. This can be observed from the slopes of line features in the image. If l and l' are corresponding line features with slopes ϕ and ϕ', we have:

$$\phi' = \phi + \beta. \tag{8.7}$$

The rotation angle can thus be evaluated from the difference in line slopes. Please note that this does not allow to distinguish between rotation angles that differ by 180^o (i.e. β and $\beta + 180^o$). The correct choice has to be made either based on prior knowledge about the sensor platform's flight path or by backtracking from later stages of the transformation. Once the rotation angle has been found the rotation is applied to the objects from the map (or the first image) and then the scale factor is determined. This parameter is visible from the lengths of line segments or the distance between two point features. Only line segments that have a relative orientation that is approximately consistent with the already found rotation angle are considered in the voting. This limits the number of feature pairs to be considered while at the same time increasing the contrast between the peak of the votes and the random background. Limiting the number of votes to be considered at the various stages of the algorithm allows to increase the robustness of the method. Translation parameters are observed trough point features.

The Semi-Affine Transform

As mentioned previously, before performing the registration between map and SAR image, the SAR image was transformed into ground range and resampled to square pixels. In principle it should therefore be possible to register the resulting image with the map using the similarity transform (in flat regions) discussed above. However, in order to transform the slant-range SAR image into a square-pixeled ground-range image, the information from the image header has to be used. This prior information is not necessarily very accurate. The resulting image was indeed noted not to have exactly square pixels. This means that we need to allow a different scale factor along x- and y-coordinates (they will however both be close to 1.0). We thus need to replace the similarity transform by a semi-affine transform, i.e. a rotation combined with a scaling that can be different along x- and y-coordinates and a translation:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} cos\beta & -sin\beta \\ sin\beta & cos\beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \tag{8.8}$$

In this case the rotation angle and scale factors are no longer visible separately. The solution to this problem is to reparametrise the transformation with respect to a new set of parameters. The set of parameters proposed in [75] is $(\beta, t_x, t_y, \Delta, \rho)$ where the first three are defined as before and $\Delta = \sqrt{s_x s_y}$ is the scale factor and $\rho = \sqrt{\frac{s_y}{s_x}}$ is the square root of the scale ratio.

The decomposition can be written as:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = T\begin{pmatrix} x \\ y \end{pmatrix} = T_{t_x} T_{t_y} T_\Delta T_\rho T_\beta \begin{pmatrix} x \\ y \end{pmatrix}, \tag{8.9}$$

where $T_\rho = \begin{bmatrix} 1/\rho & 0 \\ 0 & \rho \end{bmatrix}$ and $T_\Delta = \Delta$.

The parameters β and ρ are observable using pairs of lines (one pair in each image). β can then be found by considering the ratios of line slopes. This can be seen by applying the T_β and T_ρ to a unit vector v at an angle ϕ ($v = \begin{bmatrix} cos\phi \\ sin\phi \end{bmatrix}$).

$$v' = T_\rho T_\beta v = \begin{bmatrix} \frac{1}{\rho} cos(\phi + \beta) \\ \rho sin(\phi + \beta) \end{bmatrix}. \tag{8.10}$$

If ϕ' is the orientation angle of v' we get

$$tan\phi' = \rho^2 tan(\phi + \beta). \tag{8.11}$$

Using two non-collinear line segments β can thus be found by dividing the above expression for the two line segments.

Once β is found, ρ is found by first applying the corresponding rotation. We then have for one (rotated) line segment that $tan\phi' = \rho^2 tan\phi$ from which ρ can be found. The other parameters are estimated in the same way as for the similarity transform.

Practical Issues

The Evaluation Criteria

Ideally the distribution of a given parameter (the consensus function) should have one single clear peak. There are mainly two ways in which the function can deviate from this ideal situation: there can be several peaks of almost equal strength (lack of peak distinctness) or the function can be too flat (lack of parameter visibility). The algorithm should be able to detect these two cases and deal with them in an appropriate fashion. In [75] two criteria to quantify these effects are proposed:

- The first is a measure of *peak distinctness* and is defined as the ratio between the strength of the first peak (h_1) and that of the second peak (h_2) of the consensus function: $Q_{pd} = h2/h1$. In the ideal case Q_{pd} should be zero.

- The second measure is called *parameter visibility* and it is defined in terms of the entropy when the consensus function is viewed as a probability function. First the consensus function for parameter θ, $C(\theta)$ is normalised:

$$P(\theta) = \frac{C(\theta)}{\sum_\theta C(\theta)}. \tag{8.12}$$

  The entropy is then calculated as $E = \sum_\theta -P(\theta)ln(P(\theta))$. The parameter visibility is defined as:

$$Q_{pv} = exp\left[-\sqrt{1 - \frac{E}{ln(B)}}\right]. \tag{8.13}$$

Finding the Maxima

As the number of inconsistencies or differences between the set of objects found on the map and in the image increases, several peaks may occur in the distribution of the different parameters. It is not necessarily the highest peak that is correct. Therefore it is not sufficient to find the highest value of the consensus function for a given parameter, we need to find all local maxima. Starting from the local maximum with highest value, the transformation parameters are determined and the registration is evaluated at different levels. In order to find the local maxima of the parameter's histogram we used an approach based on 1-dimensional morphological operators. The method is schematically represented in figure 8.5. In the figure, "H" is the morphological structuring element.

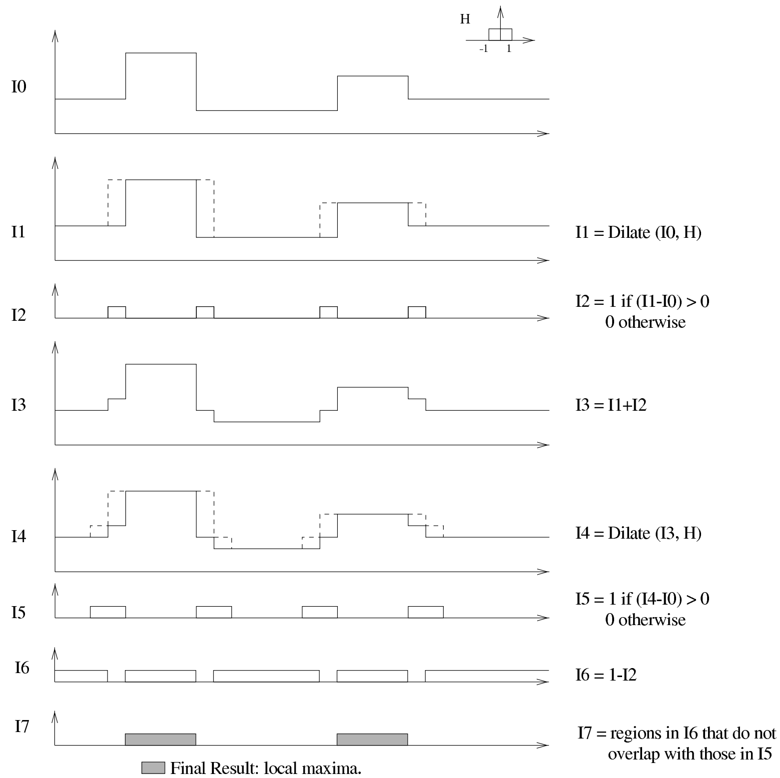

Figure 8.5: Detection of local maxima in a histogram

The Strategy used for the Feature Consensus Method

Our strategy for applying the feature consensus method involves three main changes with the original method as presented in [75]. The first change is to distribute the vote of each pair of objects over neighbouring parameter values in the consensus function. The second is the use of 2D consensus functions for finding the scale factors (for the semi-affine transform) and the translation parameters. The third difference is the evaluation criterion that is used. Below, the three changes are described in detail.

Gaussian Filter of the Voting Functions

The characteristics of the primitives that are used for the feature consensus method are the mass-centres of the built-up areas and the forests. The extent of the built-up areas and thus also their mass-centre depends on the threshold applied to the results of the logistic regression in the detector for built-up areas (see capter 7). Furthermore the size of the built-up areas as well as that of the forests can change very quickly over time. It is therefore unlikely that the physical location of the mass-centres of these objects that is found on the SAR image corresponds precisely to the mass-centres of the same objects on the map. To take this imprecision into account, each primitive's vote in the consensus function is distributed over a range of transformation parameters. This is implemented by applying a Gaussian filter to the calculated parameter value in the consensus function.

2D Consensus Functions

In the original method [75] all parameters of the transform are determined sequentially, e.g. first the translation in the x- direction and then the translation in y-direction. However, it is better to couple some of the parameters. In case of the semi-affine transform for instance, if the scale factors are close to 1, the rotation angle can be determined independent of the scale factors [75]. Thus the rotation angle is determined and applied first and then line segments that are compatible with that rotation are used to first determine the scale factor along the x-direction and then the one along the y-direction. This gives two one-dimensional consensus functions for the scale factors. In fig. 8.6 an example of such a set of one-dimensional consensus functions for estimation the scale factors of a semi-affine transform. An alternative is to couple the two scale factors in order to obtain a two-dimensional consensus function. The 2D consensus function, corresponding to the one-dimensional case of fig. 8.6, is shown in fig. 8.7.

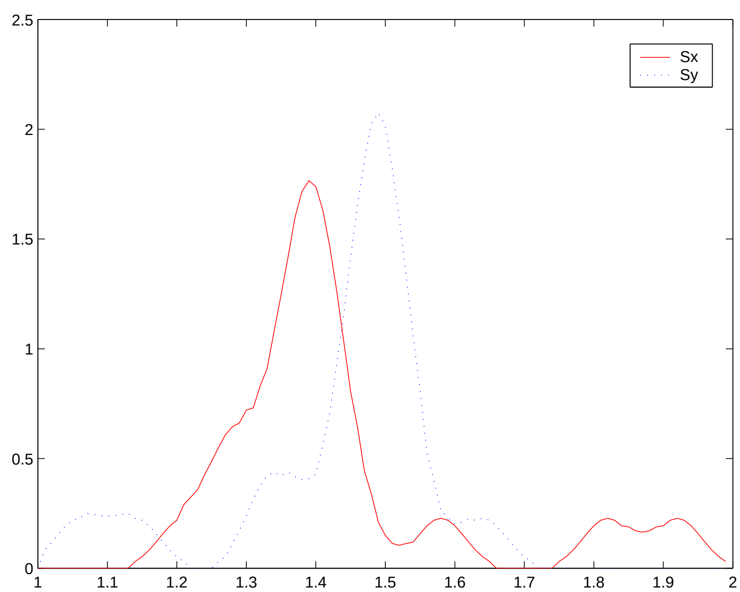

Figure 8.6: 1D consensus functions for the 2 scale factors

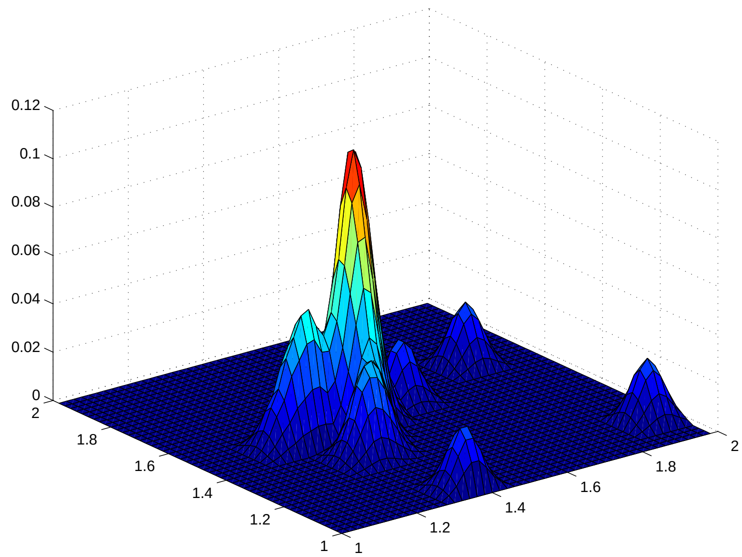

Figure 8.7: 2D consensus functions for the 2 scale factors

The 2D function offers several advantages. The 1D functions corresponds to the projection of the 2D function in the $x = 0$ and $y = 0$ planes respectively. However, this projection will sum the peaks that are masked by each other along the projection direction. This means that two secondary peaks might merge to become the most prominent peak when the 1D consensus functions are used. The second reason to use 2D functions is that the two scale factors are coupled in a natural way, i.e. the best scale factor along the x-direction will only give a good transformation if it is coupled with the corresponding scale factor along the y-axis.

The Evaluation Criterion

We have tried to use the evaluation criteria that were presented above, i.e. the peak distinctness and the parameter visibility, to guide the search trough the tree of possible transformation parameters. However their use does not lead to a robust method. Sometimes it is possible to find a clear peak for some parameter that is not the correct one. The only evaluation criterion that we found to be robust is the overlap between the objects detected on the SAR image and the objects from the map that are warped into the coordinates of the SAR image using the transformation that was found. This means that the complete depth of the search tree has to be explored before the evaluation can take place, i.e. all transformation parameters have to be determined before the overlap can be evaluated.

In order to determine the overlap, the map objects are warped onto the SAR image. Let G_M be the set of pixels corresponding to the interior of a warped object from the map and falling inside the SAR image. G_S is the set of pixels in the interior of the same type of object detected in the SAR image. The overlap quality factor Q_o is then defined as:

$$Q_o = \frac{Card(G_M \bigcap G_S)}{Card(G_M \bigcup G_S)}, \tag{8.14}$$

with $Card$ representing the number of points in a set.

If various types of objects are used in the feature consensus method, for each of them the quality factor is determined and the average quality factor over all objects is used as an evaluation criterion. The feature consensus method selects the set of transformation parameters which yields the highest value of Q_o.

Application of the Feature Consensus Method to the SAR Image

As already mentioned, the feature consensus method is used to find a first rough registration between the SAR image and a topographic map. The matching primitives that are used here are the location of built-up areas and/or forests. In fig. 8.8 these primitives are presented on the complete map of the region that was used. Apparently the geographical distribution of both forests and built-up areas is relatively random. Using their relative position should thus allow for a first registration. In fig. 8.9 the region of the objects detected on the SAR image are shown at the left and the part of the map corresponding to the SAR image on the right.

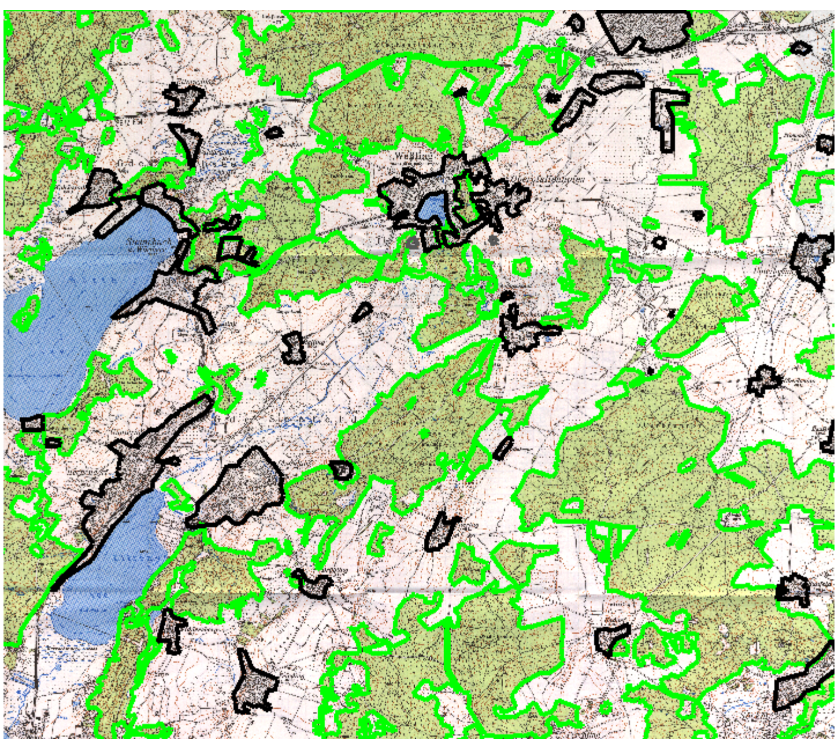

Figure 8.8: Objects from the map used in the feature consensus method

Figure 8.9: Objects detected in SAR image and region on the map corresponding to the SAR image

It is possible to constrain the search in the feature consensus method by limiting the range of possible transformation parameters. This can be done on the basis of knowledge of the SAR image acquisition parameters (i.e. heading, pixel spacing and GPS coordinates).

In order to test the robustness of the method we have set the ranges for the different parameters very wide:

Parameter	Range
Rotation Angle	$0 - 360^o$
Scale factors	$1 - 3$
Translation	-3000-3000

Table 8.1: Range of parameters used for the feature consensus method

Because the pixel spacing of the SAR image is always approximately known, it is reasonable to set the range for the scale factor somewhat narrower than for the other parameters. The feature consensus method was applied to search for a similarity transform. The resulting matched objects are then used to find the parameters of an affine transform using a least square method. In table 8.2 the estimated parameters are presented. The first two columns show the results obtained using only the position of built-up areas. The last two columns show the results obtained when used the position of built-up areas as well as forests.

Parameter	Built-Up Areas		Built-Up Areas and Forests	
	Similarity	Affine	Similarity	Affine
Rotation	47.0	48.64	46.65	48.52
Sx	1.42	1.50	1.41	1.51
Sy	1.42	1.25	1.41	1.37
Tx	-940	-671	-968	-805
Ty	-2692	-2920	-2612	-2916
Skew	0	0.0104	0	-0.00385

Table 8.2: Registration parameters found by the feature consensus method

The built-up areas from the map were transformed to the coordinates of the SAR image using the parameters given in table 8.2. In fig. 8.10 the results of the feature consensus method using only the position of built-up areas is shown. The objects (built-up areas) detected in the SAR image are delimited in yellow. The results of applying a similarity transform found by the feature consensus method are shown in red. The feature consensus method also delivers matched objects. These can be used to find parameters of another transform using a least square method. The result of using an affine transform found in such a way is presented in green in the figure.

In figure 8.11 the results of using both the built-up areas and the forests is shown. On the left a similarity transform found directly by the feature consensus method is applied. On the left the results of applying the affine transform are shown. In both figure the objects warped from the map into the coordinate space of the SAR image are delimited by green and red lines for respectively the forests and the built-up areas. The corresponding objects detected in the SAR image are delimited with orange and yellow lines.

Contrary to what could be expected, the result of the feature consensus method using only the built-up areas is slightly better than when the positions of the forests are used too. This can be seen by comparing the red objects in fig. 8.10 and fig. 8.11(left). This is due to the fact that several of the forests cross the border of the image which results in a localisation of the centres of the forests that is different in the image and the map. These differences widen the peaks in the consensus function which makes the parameter estimation less precise. However, in both cases the objects are matched correctly and using the least square method to find the parameters of an affine transform does result in a comparable registration (compare the green objects in fig. 8.10 with the red ones in fig. 8.11(right)). In fact, the result after the least square method is slightly better when both built-up areas and forests are used.

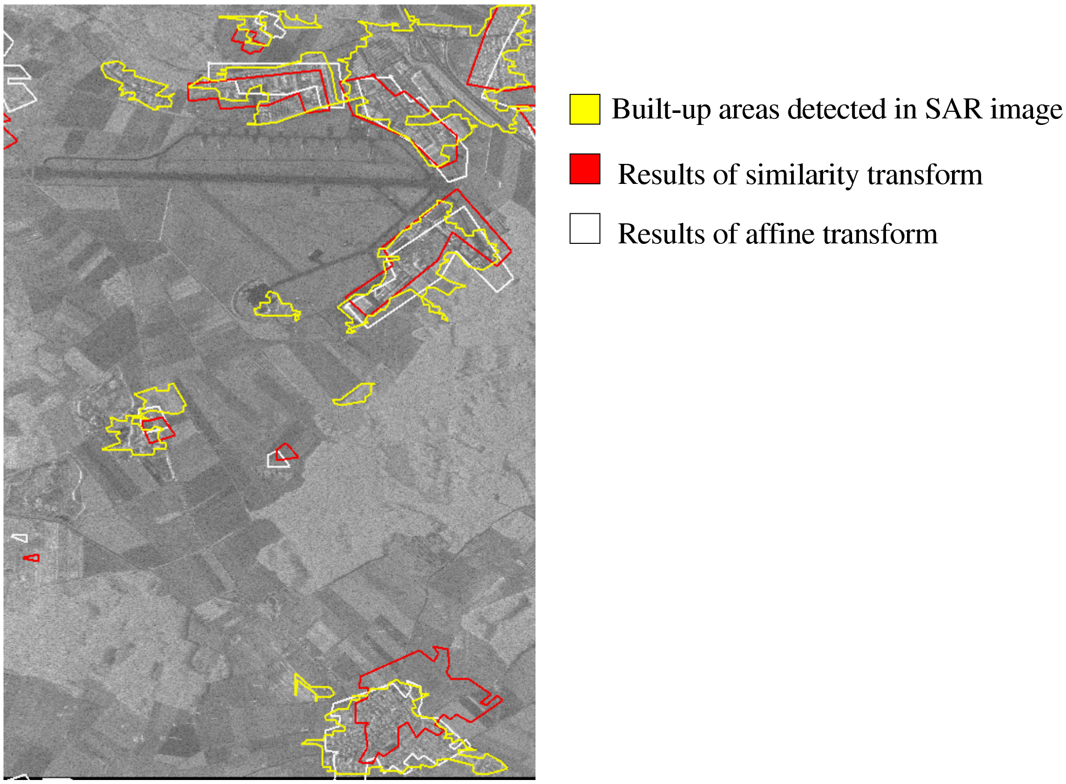

Figure 8.10: Result of the feature consensus method using only built-up areas.

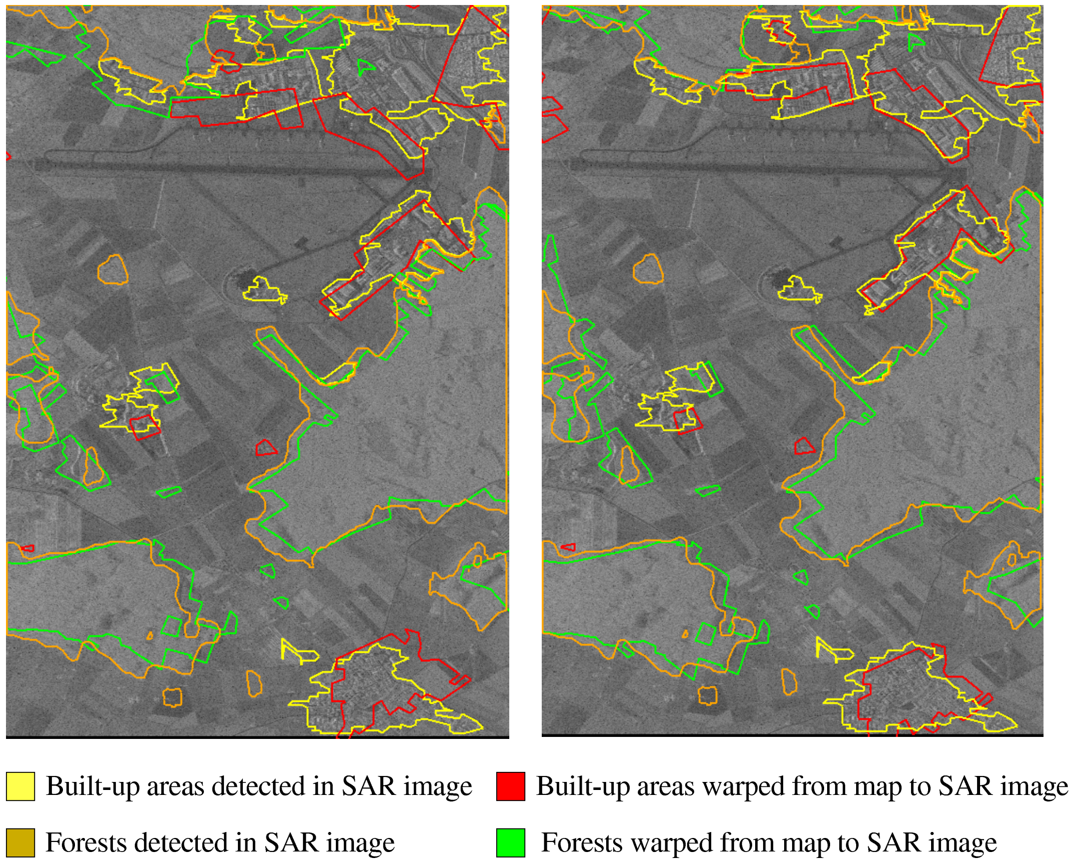

Figure 8.11: Result of feature consensus method using built-up areas and forests (left), LSQ method using matched objects (right)

These results of visual comparison are confirmed by the evaluation method described earlier. Results of this evaluation are shown in table 8.3.

Input	Method	Transform	D_{mean}	D_{RMS}	D_{max}
Built-Up Areas	Feature Consensus	Similarity	31	35	63
	Least Square	Affine	26	33	71
Built-Up Areas and Forests	Feature Consensus	Similarity	49	52	75
	Least Square	Affine	17	18	35

Table 8.3: Evaluation of the registration results

8.4.2 Refinement of Registration using Contours of Forests

Once a first registration between the map and the SAR image is found, the contours of the forests are used to refine this registration. First both contours are approximated by polygons. Then, for each vertex of the polygons of the "map forests", nearby vertices are searched in the "SAR forests". Couples of these matched vertices are used to find a better affine transformation by the method of least squares. The "map forests" are warped using this new transformation. It is likely that after this transformation other vertices are now close together and some incorrect matches used in the first iteration will be further apart as they do not fit the global transformation that was found. The vertices that are nearby after applying the transformation found by the first iteration are now used to improve the parameters of the transformation. The process is iterated until the parameters of the affine transform remain unchanged. The results of this refinement are shown in fig. 8.12. In this case the refinement procedure stabilised after 15 iterations.

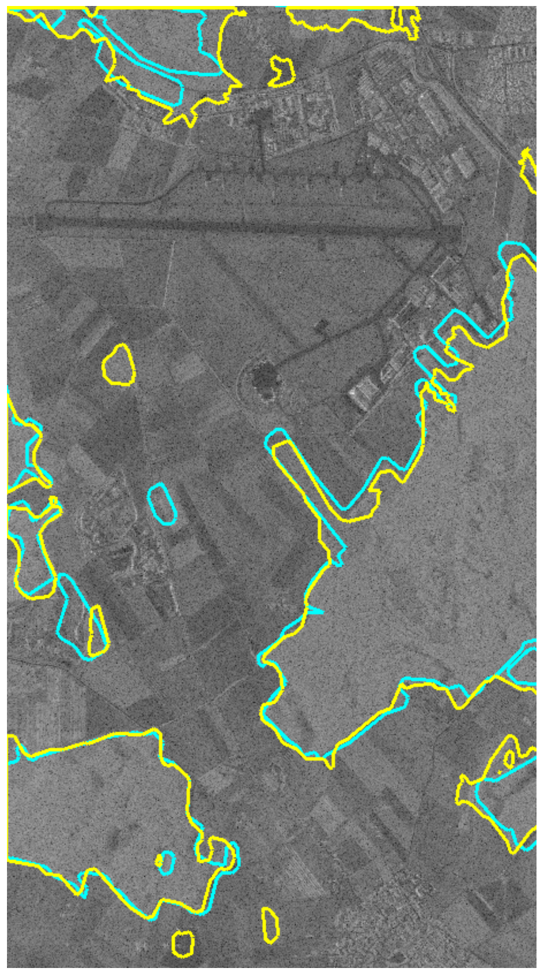

Figure 8.12: Results of first refinement step

The contours of the forests in the SAR image were found by the Freeman decomposition method. However this yields very rough contours because the decomposition is performed on the basis of relatively large averaging windows and afterward a majority voting method is used to find the forests, resulting in even more imprecision of the localisation of forest contours. A way to improve the localisation of the contours is to use the output of the region merging method (section 6.1). In order to do that we need to find out which regions are part of the forest. This was done in two steps.

The first step is a K-means clustering of the regions based on the elements of the covariance matrix. For this clustering 50 classes were selected. For the learning phase only large regions are used and an erosion is applied to them in order to take into account only the interior of the regions.

The second step is to find out which of the 50 classes correspond to the forests. This is done determining the number of pixels (N_{tt}) in each class that fall within the boundaries of either the forest contours from the map or those from the SAR image. As it is hard to distinguish forests from built-up areas based on the covariance matrix, pixels that fall into the regions detected as built-up areas are not counted as false targets. The pixels of a given class that fall outside the already detected forests and built-up areas are counted as false targets for the class (N_{ft}).

For each class the number of correctly detected forest pixels (N_{tt}) and number of false targets (N_{ft}) is calculated and a confidence measure that the class belongs to the forests is defined as:

$$Conf(forest) = \frac{N_{tt}}{N_{tt} + N_{ft}}. \tag{8.15}$$

A threshold at 50 % is applied on the resulting confidence image and the contour of the corresponding region is vectorised. Regions that are too small (< 500 pixels) are rejected in this process. Although the classification is relatively crude because built-up areas and forests are again mixed, the contours that do belong to the forest are better localised than those that were previously found by the decomposition method. Therefore, when the same refinement method is used as above, the resulting registration is closer to the actual forest contours in the image. The confidence map for forests (and built-up areas) as well as the results of the second registration refinement step are shown in fig. 8.13.

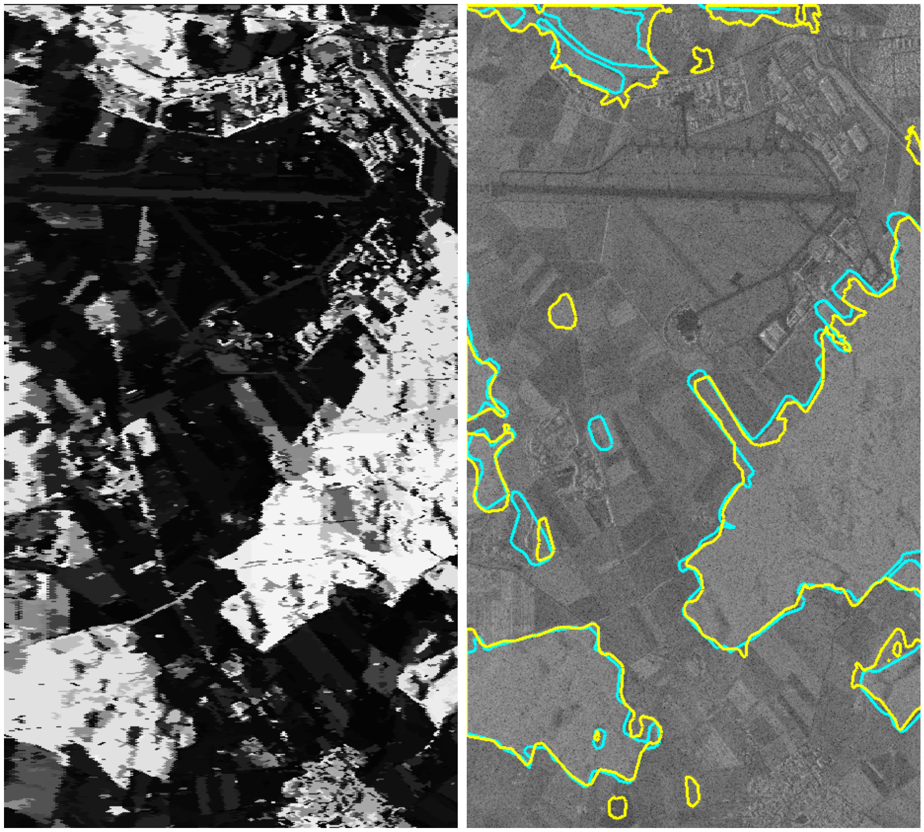

Figure 8.13: Forest confidence map (left) and results of second refinement step (right)

8.4.3 Refinement of Registration using Communication Lines

The forests and built-up areas can only provide a first registration because their 3D structure results in an inaccurate localisation (see sect. 2.3.1). The obtained registration should be further refined using features with a low 3D structure. Good candidates for refining the registration between maps and images are roads, railroads and waterways. This communication network, extracted from the map, is warped onto the SAR image using the affine transformation that was found so far. The result (fig. 8.14) shows that detecting the communication network on the SAR image is indeed likely to provide a possibility to improve the registration further.

Figure 8.14: Communication lines extracted from the map and projected onto the SAR image

The detector for communication lines that was introduced in section 5.8 gives a set of line segments that potentially correspond to roads (or railways or rivers). The result of this detector can thus be used for refining the registration. The proposed refinement procedure relies on the following hypotheses:

- the current registration is already relatively accurate. This means that candidate road segments only need to be searched in a restricted neighbourhood of the roads that were warped from the map onto the SAR image.
- as we only have detected segments of roads in the SAR image, a matching between the roads from the map to these segments will only yield the displacement component perpendicular to the road.
- we search for a global affine transformation

The first hypothesis allows to match most of the line segments found on the SAR image to one or more roads found on the map. The matching can be restricted to a small neighbourhood and to lines that are almost parallel.

The second hypothesis restricts the methods that can be used to find the transformation. Several methods were tried out and compared. They are discussed in the next sections.

Please note that the methods presented below were applied directly to the results of the feature consensus method with least square approximation of the affine transform but using only the positions of built-up areas. The forests where used neither in the feature consensus method nor were the results refined using the contours of forests. The idea is that forests could be considered as features that are not stable over time, i.e. forests can disappear or their surface area and contour can change very quickly.

Least Squares Approximation

The affine transformation can be written as:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} T_x \\ T_y \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}. \tag{8.16}$$

Applying the affine transform with a given set of transformation parameters $\Theta = T_x, T_y, a, b, c, d$ will transform a point (X, Y) into a point (X', Y') and the displacement of the point P : $\overrightarrow{dP} = dX\overrightarrow{1_x} + dY\overrightarrow{1_y}$ is given by:

$$\begin{aligned} dX &= T_x + aX + bY - X \\ dY &= T_y + cX + dY - Y \end{aligned} \tag{8.17}$$

If a matching between the line segment from the SAR image and a road from the map can be found, we know the component, perpendicular to the road on the map, of the displacement that the points of the SAR line segment should undergo, when the optimal transform is applied.

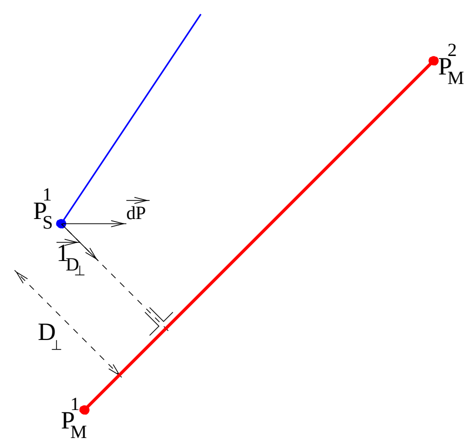

Figure 8.15: Principle of optimisation by least square method

The aim of the optimisation method is thus to find a set of transformation parameters that yields a displacement that minimizes this perpendicular component, i.e.

$$MIN_\Theta \left(\sum_i (\overrightarrow{dP_i}\overrightarrow{1_{D_\perp}} - D_\perp(P_i) \right)^2, \tag{8.18}$$

with i an index over the points of the SAR image for which a match with the map was found. $D_\perp(P)$ is the value of the perpendicular distance and $\overrightarrow{1_{D_\perp}}$ a unit vector pointing along the perpendicular direction from the point of the SAR segment to the road from the map (see fig 8.15).

If P_S is the point of the SAR segment and the corresponding road segment from the map is given by $[P_M^1, P_M^2]$ then $D_\perp(P)$ is given by:

$$D_\perp(P) = \frac{\| \overrightarrow{P_M^1, P_S} \times \overrightarrow{P_M^1, P_M^2} \|}{\| \overrightarrow{P_M^1, P_M^2} \|} \tag{8.19}$$

and the unit vector perpendicular to the road segment from the map can be found from the equation of the line passing through P_M^1 and P_M^2. Let $\overrightarrow{1_{D_\perp}} = (1_{Dx}, 1_{Dy})$ be the unit vector. The minimisation problem can then be written as a system of equations:

$$[1_{Dx}]\ T_x + [X\ 1_{Dx}]\ a + [Y\ 1_{Dx}]\ b + [1_{Dy}]\ T_y + [X\ 1_{Dy}]\ c + [Y\ 1_{Dy}]\ d - [X\ 1_{Dx} + Y\ 1_{Dy} + D_\perp(P)] = 0. \tag{8.20}$$

The unknowns $\Theta = T_x, T_y, a, b, c, d$ can be found by a least square method. Although the least square method should yield the optimal set of parameters in one go, i.e. without iterations, applying the method iteratively improves the results. This is due to the fact that in the first run a number of matches can be false. These pull the solution away from the correct one. It is not possible to identify these incorrect matches a priori. However, as with the feature consensus method, the incorrect matches are hoped to be more or less randomly distributed while the correct matches all induce a preferred displacement corresponding to the correct transformation. The least square method will thus, in the first step, approximate the solution. After applying the transformation that is found, some of the incorrect matches will now become invalid and in the next iteration, a higher percentage of points will vote for the correct transformation.

Table 8.4 shows the distance to the ground control points at successive iterations. The line corresponding to 0 iterations is the result of the feature consensus method using only built-up areas (the second line in table 8.3). The distances were calculated using a set of manually indicated ground control points (GCP's) on the map and the SAR image. Each of the GCP's from the map are transformed using the transformation parameters that were found so far.

Iteration	D_{Mean}	D_{RMS}	D_{Max}
0	26.0	33.5	71.0
1	23.0	28.6	62.7
2	22.8	27.3	59.7
3	21.2	24.2	48.5
4	18.1	20.5	42.4
5	13.8	15.7	33.1
6	8.5	9.5	20.2
7	6.8	7.5	15.1
8	5.6	6.5	13.9

Table 8.4: Distance between GCP's as a function of the number of iterations for the least square method

The Simplex Method

The simplex method [78, 79, 80] is a direct search method that does not use numerical or analytical gradients. If n is the number of parameters of the transformation to be found, a simplex in n-dimensional space is characterised by n+1 distinct vectors that are its vertices. In 2D a simplex is a triangle, in 3D a pyramid. As the affine transform has 6 parameters, the simplex has 7 vertices in a 6D space. Each vertex thus corresponds to a given set of parameters and for each of these a cost function can be determined. At each step of the search one of the vertices is replaced by a new vertex which lies in or near the current simplex. This is done by determining the worst point $W = P_k$ of the current simplex and forming its symmetrical image through the centre of the opposite hyperface of the simplex:

$$P_{k,new} = M + (M - P_{k,old}), \tag{8.21}$$

with $M = \sum_{i \neq k} P_i / N$.
This is called *reflection*. If the new point is the best, an *expansion* is applied:

$$P_{k,new'} = M + a(P_{k,new} - M) \quad a > 1. \tag{8.22}$$

If the result is worse than the second worst, a *1D contraction* is applied:

$$P_{k,new'} = M + b(P_{k,old} - M) \quad 0 < b < 1. \tag{8.23}$$

If the result is still not satisfactory, a *multiple contraction* is made around the best point $B = P_j$:

$$P_{i,new'} = B + m(P_{i,new} - B) \quad 0 < m < 1 \quad i \neq j. \tag{8.24}$$

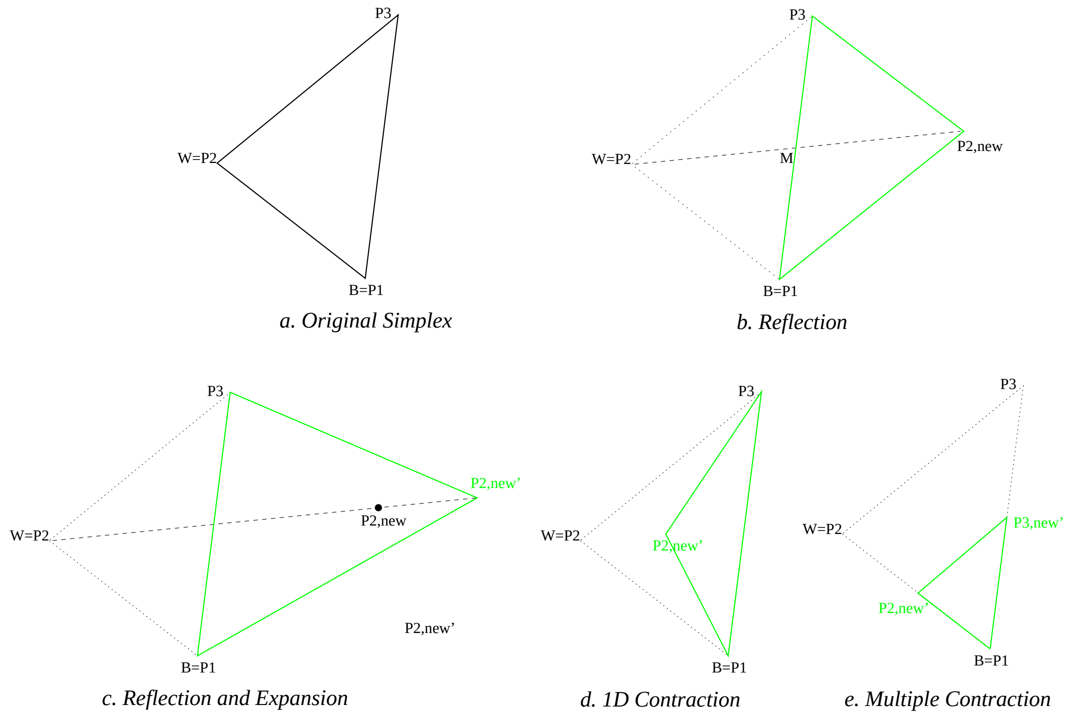

Figure 8.16: Principle of optimisation by Simplex method (for 2D problem)

These steps (see also fig. 8.16) are repeated until the distance between all vertices of the simplex is smaller than a given precision. The simplex method is repeatedly applied. Each time the transformation that is found is applied to the objects that were found on the map. Because at each step the transformation will become more accurate, there will be less incorrect matches and the cost function will be estimated more correctly.

Iteration	D_{Mean}	D_{RMS}	D_{Max}
0	26.0	33.5	71
1.	28.3	35.6	76.9
2	28.2	34.8	76.2
3	27.5	33.8	74.5
4	26.7	32.9	72.7
5	26.7	32.9	73.2
6	25.0	31.1	70.3
7	23.1	29.2	67.6
8	16.9	21.7	50.4
9	10.5	13.2	30.5
10	8.6	10.5	23.6
11	7.6	9.0	19.5
12	7.3	8.5	18.6

Table 8.5: Distance between GCP's as a function of the number of runs for the simplex method

Table 8.5 shows the evaluation of the transformation that is obtained at each iteration of the simplex method. From the table it appears that the simplex method converges slower to an optimal solution than the least square method. Furthermore, the first iteration seems to give a transformation that is worse than the initial transform and only after several iterations an improvement is noticeable.

Results of the Refinement Method

In fig. 8.17 the final results for the refinement using the Simplex and the Least-Square method are shown. As expected from the numerical values presented in tables 8.5 and 8.4, the results are very close.

Figure 8.17: Results of the refinement using the simplex method (green) and the LSQ method (cyan)

In fig. 8.18 the result of applying the feature consensus method for registering image 2 (in fig. 3.6) with the map is shown. The feature consensus method was run using only the location of built-up areas. Although the detected built-up areas from the SAR image were correctly matched with the objects from the map, the registration obtained in fig. 8.18 is still inaccurate. This is due to the fact that most objects used for this registration overlap the image, resulting in a spread of the peaks of the consensus functions. In fig. 8.19 the results of the refinement using the communication lines is shown. Here the least square method is used. Apparently the refinement method does allow to correct the inaccuracies of the feature consensus method.

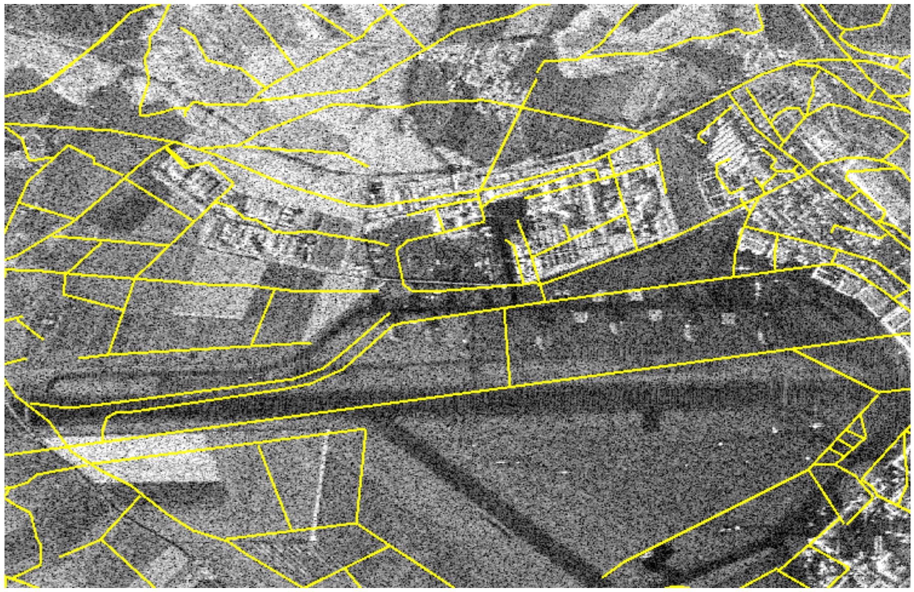

Figure 8.18: Results of the feature consensus method applied on image 2

Figure 8.19: Feature consensus results combined with the refinement based on communication lines for image 2

8.5 Registration of SAR Images

Once each image is registered with a map, the registration between different images becomes easy:

The transformation between the images and the map is used to obtain a first approximation of the registration between the different images. This registration is then refined using the position of edges detected in each image. In fig. 8.20 the method used to register different SAR images is schematically represented. Each SAR image was first transformed to ground range and rescaled to approximately square pixels. Then an affine transformation that warps these resampled images onto a digitised map was determined. Because only a global affine transformation was used for registering each image with the map, the registration transformations can be easily inverted and combined. This allows to find an affine transformation that should warp one image onto the other.

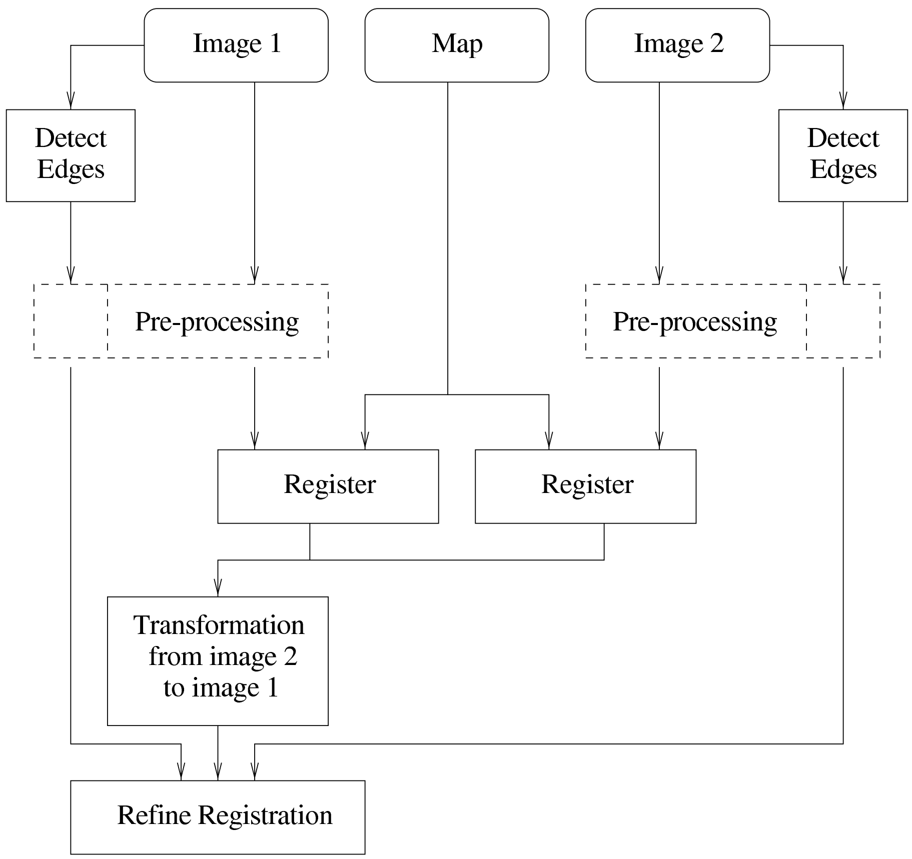

Figure 8.20: Overview of the complete registration method for SAR images

That affine transformation is refined using the position edges detected in the two images. The edge detectors were applied to the original slant range images and therefore the resulting objects need to be first converted to “ground range” and rescaled in the same way as the SAR images.

In fig. 8.21 a part of image 2 is shown. In blue the edges detected in that image are represented. The yellow objects correspond to edges detected in image 1 and warped into the second image using the affine transform that was found so far. The two sets of edges represented in the figure are used to refine the registration. In fig. 8.22 the results of the refinement are shown.

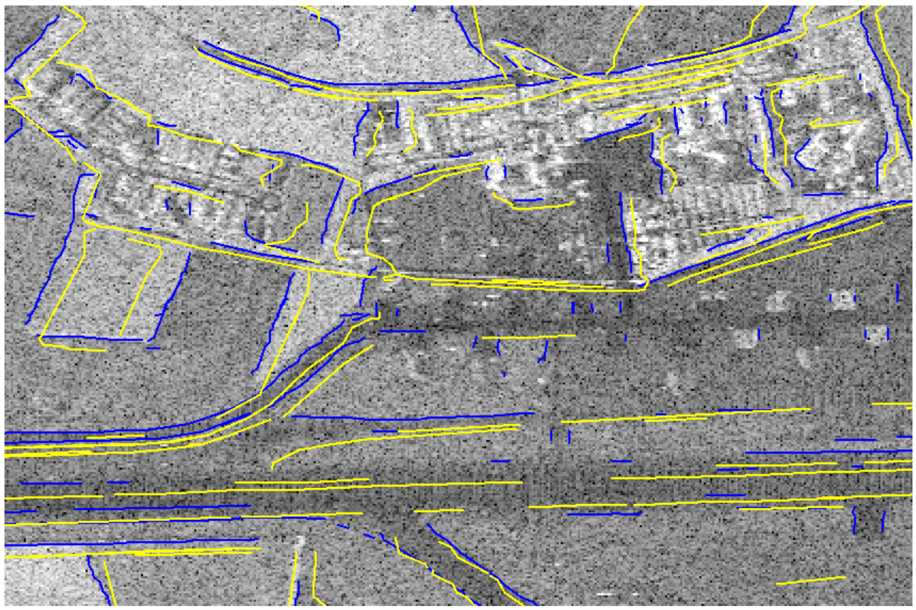

Figure 8.21: Detected and warped edges superimposed on image 2

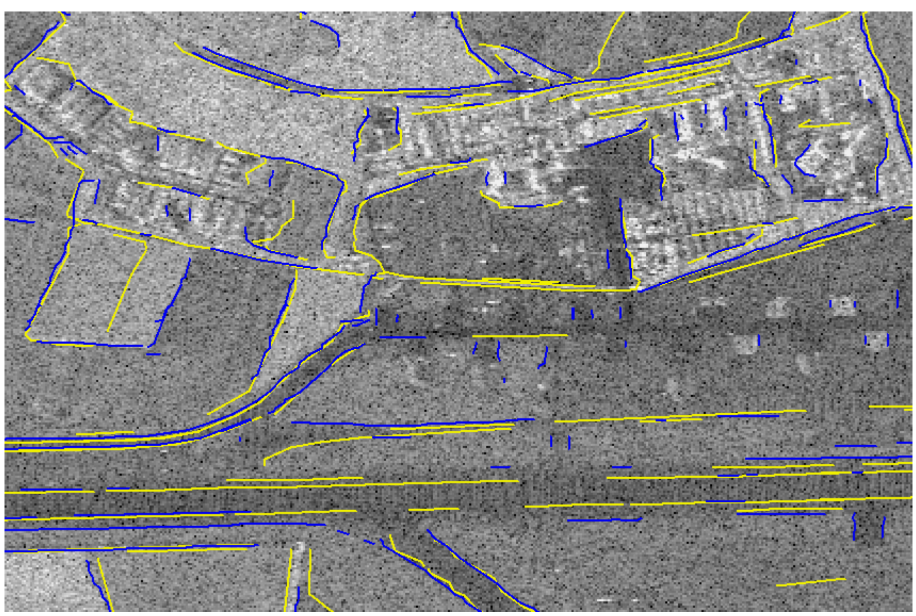

Figure 8.22: Result of registration refinement based on detected edges

After this last refinement step we have the affine transforms that allow any warping between the map and the two images. In fig. 8.23 the map and the first image are warped into the coordinates of the second image. In the figure the first image is represented in green and the second image in blue. For the parts of the map that overlap with either of the images, the dark objects from the map (legend and borders of objects) are superimposed in red.

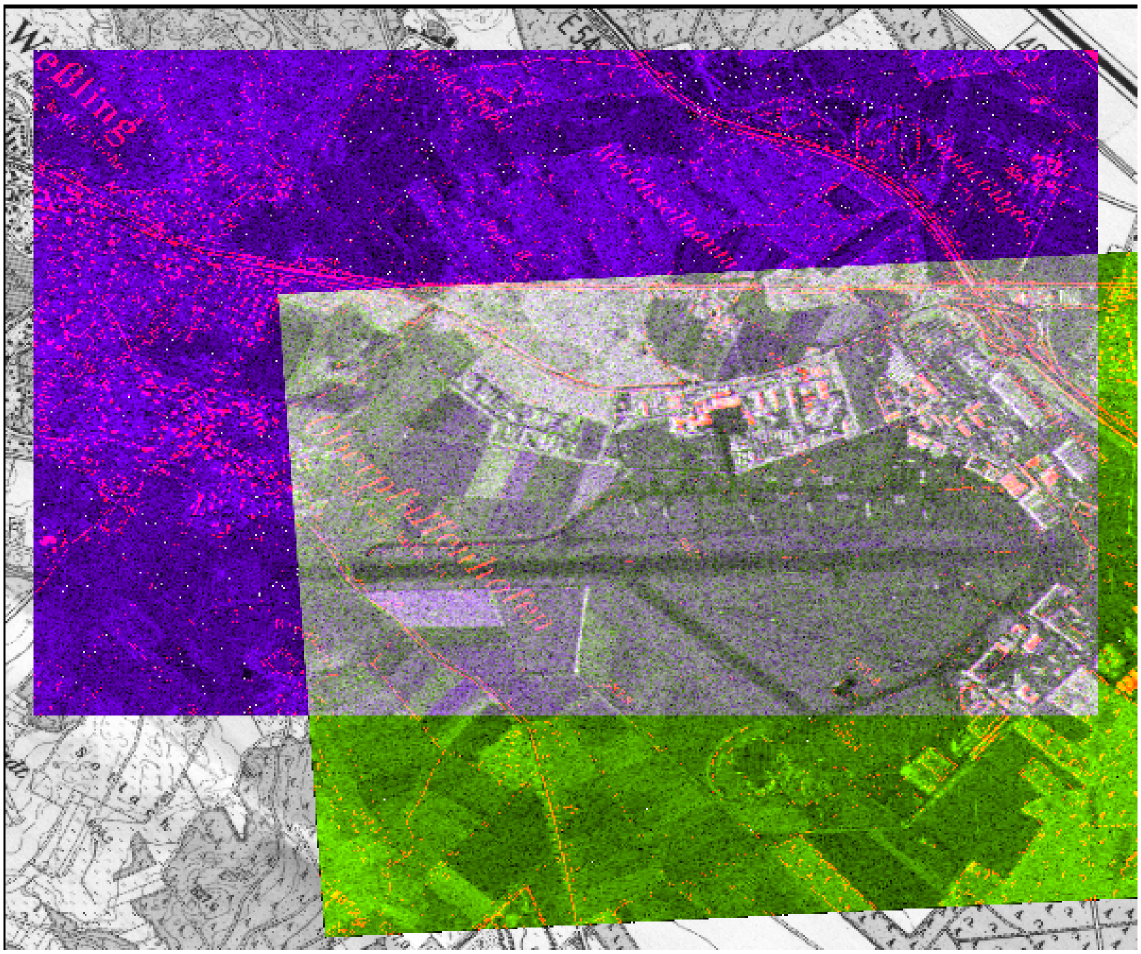

Figure 8.23: Final result of the registration (Green=Part of image 1 warped to image2, Blue=Image2)

In order to get a more quantitative idea of the accuracy of the obtained registration between the two images, the evaluation method described before can be used. However, the evaluation method uses manually determined ground control points. For a rough registration, the error made by manually selecting ground control points such as road crossings can be neglected with respect to the registration error. Now, after the final refinement step, this is no longer the case and we need to find control points that are the same in both images to within an accuracy of one pixel. The only possible candidates are the corner reflectors on the airfield. The control points are the pixels of highest intensity of each of the corner reflectors. The results of the evaluation are shown in table 8.6. The first line in the table is the result of the registration only using the combination of the registration of both images with the map. The second line presents the results after refinement using edges.

Method	D_{mean}	D_{RMS}	D_{max}
Before refinement	2.428	2.523	3.299
After refinement	1.018	1.106	1.726

Table 8.6: Evaluation of the final registration results