A STRATEGY FOR AUTOMATIC IMAGE TO MAP REGISTRATION

Heiner HILD, Norbert HAALA and Dieter FRITSCH
Institute for Photogrammetry (ifp), Stuttgart University, Germany
{heiner.hild, norbert.haala, dieter.fritsch}@ifp.uni-stuttgart.de

KEY WORDS: Geo-referencing, Automation, Feature Extraction, Image registration.

ABSTRACT

This paper introduces a fully automatic approach for the registration of satellite imagery by matching extracted image segments against the corresponding objects of an existing a GIS data base. Within the study SPOT imagery is exemplarily applied, whereas the topographic data base is provided by the German ATKIS. Since neither lines nor points allow for the determination of the complete transformation for each pair of corresponding objects, the image-to-map-transformation is based on corresponding regions, which are defined by polygons. The selected GIS objects are matched to the extracted image objects in a three phase process. In the first phase, a coarse global affine transformation between satellite image and vector map is determined. The second phase refines this global transformation by applying a supervised segmentation using seed regions, which are generated based on the results of the first phase. In the third phase individual transformations between corresponding objects are computed. After this step, ground control points at the boundaries of the corresponding polygons are available, which can be used as input for a standard aerial triangulation process.

1 INTRODUCTION

The automatic registration of image data is a key task in photogrammetry. Although remarkable progress has been achieved in integrating GPS and inertial data to directly georeference airborne or spaceborne imagery, the use of corresponding primitives in object and image space in the framework of an aerial triangulation will not become obsolete. Even though high accuracy can be achieved by direct georeferencing, the use of ground control will always be mandatory for system calibration or accuracy control. In contrast to tasks like DEM generation or relative orientation, which can be solved based on image-to-image matching, a general solution for the fully automatic absolute orientation of airborne and spaceborne imagery is not available. The main problem is the required matching of corresponding primitives from reference data, which is usually provided by a GIS to the image. A number of approaches to automatically reconstruct the absolute orientation of aerial images are based on the detection of man-made features like man-holes (Drewniok and Rohr 1997), building roofs (Schickler 1992) or linear topographic features like road and river patterns (Haala and Vosselman 1992). Within our work, similar to the approach described in (Dowman 1998), polygonal objects are applied.

The presented approach is designed to be completely independent from any a priori information on the exterior orientation, it is solely based on the image and topographic reference data itself. The key idea is to automatically match image objects extracted by region growing against corresponding objects provided by the GIS data base. In contrast to approaches aiming at the automatic update of GIS data by classification of (multispectral) images, the complete interpretation of the image is not obligatory for this purpose since only a limited number of control points, i.e. corresponding objects is required. Since the main emphasis is set more on reliability than on completeness of the object extraction process, the algorithm can be restricted to GIS objects with unique spectral properties in order to simplify the detection of the corresponding image segments. For the examples presented in this paper, the required GIS objects are provided by the German ATKIS data base, the image data to be georeferenced is a panchromatic SPOT scene. The polygons to be matched are representing forest regions, since this type of landuse shows homogeneous and prominent spectral properties, thus simplifying the segmentation and extraction of these regions. Additionally, forest regions appear quite frequently and are well distributed in most scenes, which is a prerequisite for their application as ground control primitives. The main disadvantage of using forest regions for that purpose is the limited accuracy of their outlines in the GIS data, nevertheless for the registration of SPOT imagery with a ground pixel size of 10 m this accuracy is sufficient. Of course, the proposed approach is not limited to forest regions, the application of other landuse classes like water bodies is also feasible.

2 THE PROPOSED SYSTEM

In order to provide the ground control point information by the system, the main goal is to extract polygons from the image which will have corresponding shapes to objects in the reference data. Thus, in the first step the relevant polygons have to be extracted from vector and image space. This object generation process will be described in section 3. In order to find corresponding objects, the relevant features have to be extracted (section 4), and a matching and verification step has to be performed (section 5 and 6). The final results of the algorithm, including the assessment of quality, are discussed in section 7.

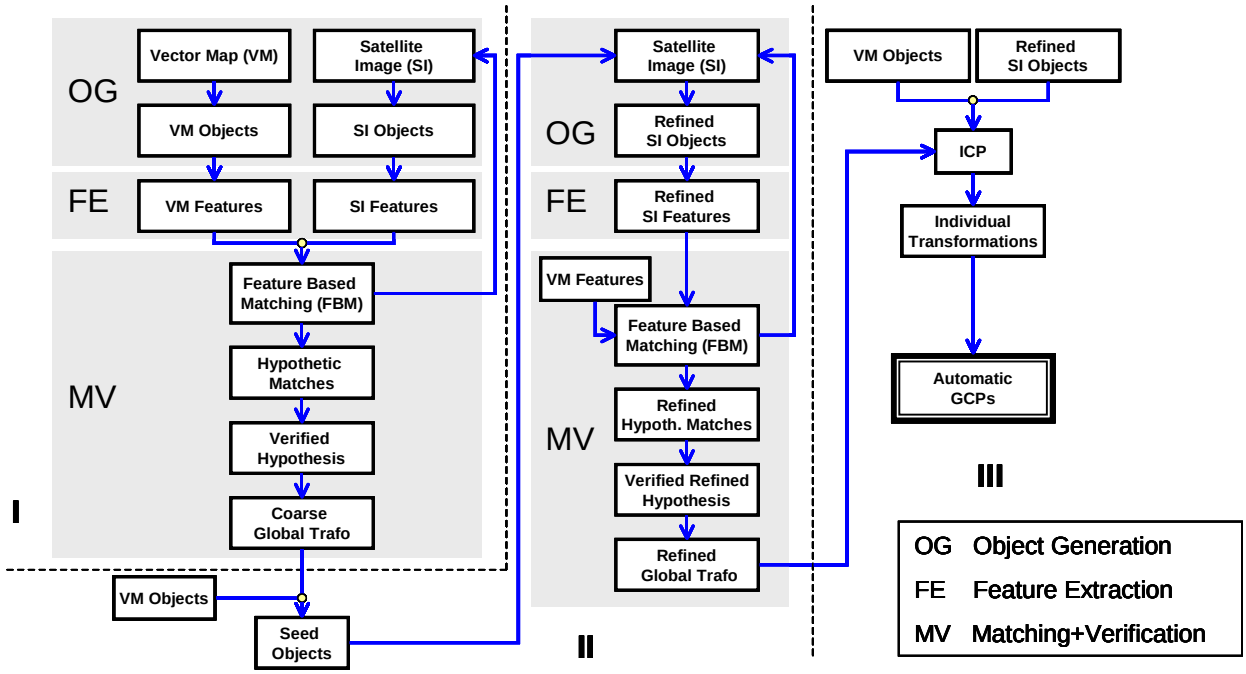


Figure 1. Dataflow of the proposed system

An overview of the proposed system is depicted in figure 1. As it is visible, the algorithm can be divided in three main parts. The phases I and II are separated into the parts *object generation* (OG), *feature extraction* (FE), and *matching and verification* (MV). Phase I aims at the determination of a global transformation between the SPOT imagery and the ATKIS data in order to provide a first approximation on image scale and orientation. All corresponding image and GIS objects are used to determine these parameters. For aerial photography the main problem during the application of global affine transformation between object and image space is relief, whereas for satellite imagery like applied in this study, problems will mainly occur for large tilt angles. Nevertheless, since the global affine transformation is only applied to provide approximate values for the subsequent steps, it was found to be sufficient for our goals.

Since the great influence of the selected thresholds on the results of the image segmentation processes is well known, the basic philosophy of our algorithm is to replace the search for an 'optimal' set of parameters by an iterative process in order to apply a number of different thresholds, automatically. This of course results in a large number of segmentation results. Hence, due to the large number of potential matches between GIS and image objects, the computational effort for the feature extraction and the subsequent matching process has to be optimized. This is realized by a feature based matching procedure. The generated matching hypotheses are verified by a robust estimation of a global affine transformation between the corresponding objects.

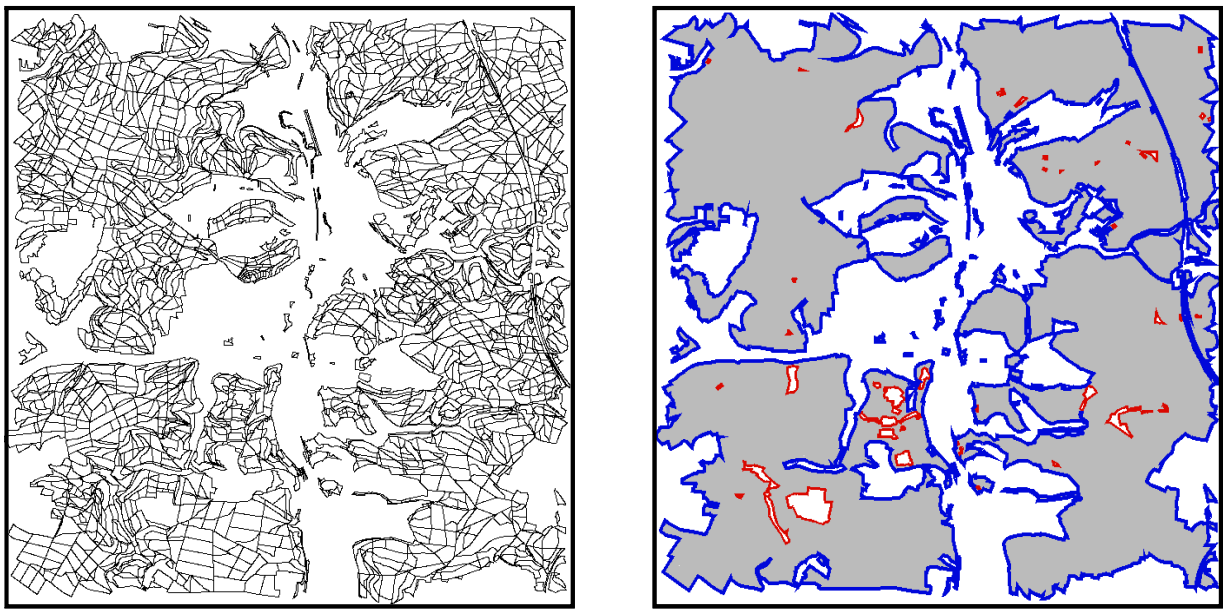
In phase II this global transformation is refined using the parameters of phase I to provide approximate values for a re-segmentation of image objects. For this purpose, the GIS objects are mapped into the image applying the coarse affine transformation of phase I. The GIS objects then provide seed regions, which are applied in phase II for a refined segmentation. Based on this segmentation the matching procedure is repeated, subsequently. This results in an improved, but still global affine transformation. In phase III, individual affine transformations between corresponding objects are computed. This is realized by an iterative closest point algorithm based on the boundaries, which results in single ground control points to be used for a subsequent aerial triangulation.

3 GENERATION OF VECTOR AND RASTER OBJECTS

Before the matching procedure can be applied, significant primitives have to be created and numerical features based on the outline of the selected objects have to be computed. Due to the different nature of object and raster space, different techniques have to be applied for each space to reach this goal.

3.1 Vector space

The reference data in vector space is provided by the German ATKIS data base (AdV 1988). This data base consists of point, line, polygon and complex geometric objects which are assigned to a certain object class. In the first step of the vector object generation, polygonal objects of the landuse class "forest" are extracted. In the second step, adjacent polygons are merged since the individual objects mostly do not represent homogeneous areas in the satellite image but only parts of them. In order to merge neighboring objects, a polygon growing algorithm was developed. Without the computational expensive generation of the topology of the complete data set, this algorithm subsequently assigns potential neighbors by simply checking the bounding boxes of the objects, which is fairly fast. If the bounding boxes are overlapping, common lines are detected and the objects are merged. In this way, a start polygon is grown until no additional neighboring polygons can be found in the data set. The design of the algorithm also allows for the extraction of holes in the objects. Figure 2 shows the result of formerly 2743 forest polygons of the data set (left part) merged into 133 polygons including the detected holes.



Original ATKIS forest polygons Polygon merging result

Figure 2. Hole preserving merging of adjacent vector map polygons

3.2 Raster space

A huge number of segmentation procedures has been developed so far for the creation of polygonal objects of homogeneous grayvalue distribution. In this paper, we use an adaptive threshold region growing algorithm as described by (Levine and Shaheen 1981) in two versions. Version 1 uses start points beginning from the lower left image corner and grows a homogeneous region until no further growth is possible. This version is used in phase I of the overall system. Version 2 uses seed regions, i.e. image regions, which are expanded while taking into account the grayvalue statistics of the seed region. This version is used for the refinement of the global transformation in phase II.

Let N be the actual number of pixels in the actual region R_i, the grayvalue of a pixel in row r and column c may be $g(r,c,)$. Then the relevant statistical values of the region R_i compute to:

mean grayvalue: $M_{R_i} = \frac{1}{N} \sum_{(r,c) \in R_i} g(r,c)$

standard deviation: $\sigma_{R_i} = \frac{1}{N} \left[\sum_{(r,c) \in R_i} \left[g(r,c) - M_{R_i} \right]^2 \right]^{1/2}$

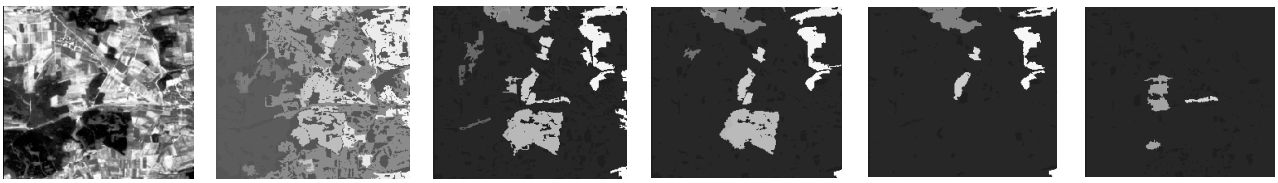
Since the interesting objects are located in the foreground, we apply a 4-neighborhood relation to grow a region R_i. A pixel in the 4-neighborhood of the actual pixel is assigned to the region if $|g_n(r_n,c_n) - g(r,c)| < adaptive\,threshold\ AT$, where (r_n,c_n) are the 4-neighbors of (r,c) and $g_n(r_n,c_n)$ is the corresponding grayvalue.

The adaptive threshold *AT* computes from the region statistics and the user defined fixed threshold *T* to:

$$AT = \left[1 - \min(0.8, \frac{\sigma'_{R_i}}{M'_{R_i}})\right] \times T$$

where σ'_{R_i} and M'_{R_i} include the actual pixel.

The dynamically modified adaptive threshold *AT* can not exceed the fixed threshold *T* but can become much smaller. Adaptive thresholding is useful e.g. in preventing bleeding across smooth image gradients. In Figure 3, the effect of various fixed thresholds on the adaptive threshold segmentation is illustrated with a section of a SPOT PAN scene. For each of these segmentations the FBM tries to find a match in the GIS data set. Hence the hypothetic match list consists of raster objects from different segmentation runs.

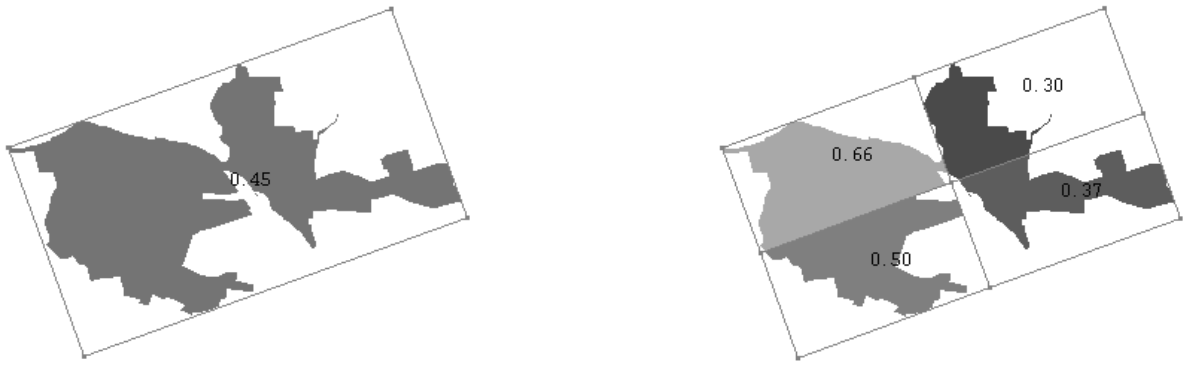


histogr. stretched image section | AT=3 | AT=5 | AT=7 | AT=10 | AT=15

Figure 3. Influence of fixed threshold values on adaptive threshold region growing segmentation of SPOT PAN data

4 FEATURE EXTRACTION

The aim of the presented approach is the fully automatic detection and localization of ground control points (GCP) in a multi-phase procedure. The global transformation from image coordinates to global reference coordinates is chosen to be of affine nature. Hence, the features to be extracted for the matching have to be invariant or almost invariant against affine transformation. Many solutions were already found to extract invariant features from polygonal objects, like invariants based on areal moments, (Reiss 1993), Fourier descriptors, (Arbter and Snyder and Hirzinger and Burkhardt 1990), or boundary chain code frequency, (Abbazi-Dezfouli and Freeman 1994), just to name few. They all have in common, that their computation is rather time consuming. For reasons of simplification, in this approach we present a set of features that are easy to compute, completely invariant under similarity transformation and almost invariant under affine transformation. Since raster and vector data is used simultaneously, the computation of the features should be possible in both domains. A conversion in either raster or vector space is not desirable due to the possible loss of accuracy that results from a conversion.



MBR and F_1 value | MBR quadrants and quadrant fillings F_2, F_2, F_2, F_4

Figure 4. Calculated features for an example GIS object

In our approach the features of the polygonal primitives are based on the so called minimum bounding rectangle (MBR). The MBR of an object is the rectangle that completely contains the object and has minimum size. In the presented system, the MBR is computed iteratively for each raster and vector object. Then the object feature value F_1 is defined by the MBR filling ratio, i.e. the ratio of the object area A_{Obj} and the MBR area A_{MBR}:

$$F_1 = \frac{A_{Obj}}{A_{MBR}}.$$

If the MBR is separated into four equally sized sectors, the filling ratios of each sector provide four additional feature values. If $A_{Obj,i}$ is the area of the object in the i^{th} MBR quadrant, these four numerical values compute to:

$$F_2 = \frac{4A_{Obj,1}}{A_{MBR}} \qquad F_3 = \frac{4A_{Obj,2}}{A_{MBR}} \qquad F_4 = \frac{4A_{Obj,3}}{A_{MBR}} \qquad F_2 = \frac{4A_{Obj,4}}{A_{MBR}}$$

The same features are computed for the raster objects in order to be fed into the matching process. Figure 4 exemplary shows the calculated five features for an example GIS object. The grayvalues of the single areas represent the relative area filling.

The features F_1 to F_5 are obviously invariant under rotation, scaling and translation since the object shape is not affected for these transformations. Both the object (i.e. the nominator) and the MBR (i.e. the denominator) transform similarly under these transformations and the effect of the transformation on the size of the respective areas is eliminated by the application of the ratio. This holds not true for shearing. For shearing, the effect on the MBR is different for each object. Due to the iterative process of the MBR generation, a closed solution can not be given. Therefore, the dependence on the feature values with respect to a rotation of the x axis was exemplary studied. The results are depicted in Figure 5. As expected, the computed features are not fully invariant under shearing. Nevertheless, for most satellite scenes, the change of feature values can be accepted for reasonable shearing values. A relative difference below 20% seems to be sufficient in order to still enable a matching between image and GIS objects with respect to this investigation. One problem is the potential change of the MBR definition in relation to the object border for large shearing values. In that case, the MBR parameters are defined by a different set of points of the boundary polygon. Effects like this can not be avoided due to the iterative nature of MBR computation and therefore have to be eliminated by a robust matching process.

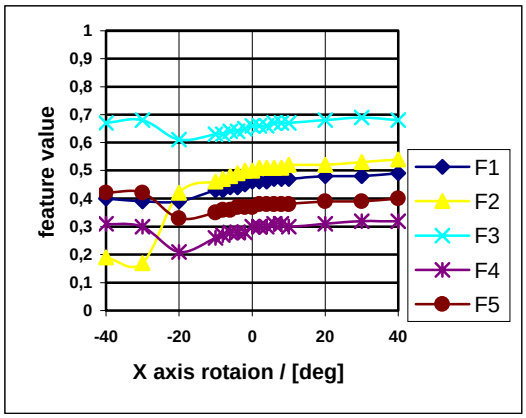


Figure 5. Influence of shear on computed features

5 MATCHING AND VERIFICATION

In order to detect object correspondences a feature based matching (FBM) is applied. An object of set A matches an object of set B if a certain criterion is fulfilled for each pair of corresponding numerical feature values. The situation at the beginning of the matching is illustrated in Figure 6. Let set A consist of M objects and set B consist of L objects. The single features are labeled by F_i. The value of the n^{th} feature of the m^{th} object in set A is then represented by A_{mn}.

ID	F_1	F_2	…	F_N
1	A_{11}	A_{12}	...	A_{1N}
2	A_{21}	A_{22}	...	A_{2N}
…	...	...	...	...
M	A_{M1}	A_{M2}	...	A_{MN}

Set A

ID	F_1	F_2	…	F_N
1	B_{11}	B_{12}	...	B_{1N}
2	B_{21}	B_{22}	...	B_{2N}
…	...	...	...	...
L	B_{L1}	B_{L2}	...	B_{LN}

Set B

Figure 6. Two sets of numerical features to be matched

The matching criterion for a pair of corresponding feature values A_{ij} and B_{kl} is the relative difference (RD) which computes to:

$$RD(A_{ij}, B_{kl}) = \begin{cases} \dfrac{|A_{ij} - B_{kl}|}{Max(A_{ij}, B_{kl})} & \text{for } A_{ij} \neq 0 \text{ or } B_{kl} \neq 0 \\ 0 & A_{ij} = B_{kl} = 0 \end{cases}$$

A match between a pair of values is assigned, if their relative difference falls bellow the threshold for this feature, i.e. 10%.

The implemented procedure first determines the maximum and minimum values for each feature of both input sets. Based on these values, the maximum relative difference $RD_{max,i}$ for each feature is selected. Before the objects are checked for potential correspondences, a presorting algorithm is applied in order to reduce the search space without omitting any potential match. Therefore, the objects of each input set are assigned to a certain category for each feature, separately. A category in this context is represented by a value range. In our implementation, the category width is equal or lower than the RD threshold for this feature. The complete value range for each feature is split into categories of equal size. After presorting, the matching partner of object *n* in category *i* in set A has to be found in one of the categories {*i-1, i, i+1*} of set B and the elements of all other categories can be rejected without any further testing. The search space minimization strategy additionally applies different features based on their RD threshold. Features with smaller RD threshold are considered earlier for matching and only the object pairs that passed the check for one feature are fed into the next feature check. By considering the most rigorous features first, a lot of potential matching candidates can be rejected very early and thus the search space is reduced to a minimum for the applied criterion. The percentage of reduction with this strategy directly depends on the RD thresholds set by the user. Since the orientation of GIS objects with respect to the image is not available at the beginning, the sector filling features can be cyclically permutated. Therefore, the matching procedure must also be able to regard cyclic permutations of certain features. In the implemented program, this option can be activated individually for each feature.

If the matching is terminated, a global affine transformation is computed from the MBR corners of the hypothetic matches. If cyclic feature permutation led to the hypothesis as may be the case, the MBR is rotated first. Based on all MBR corner point pairs, the global transformation is determined by a robust estimation procedure. The necessary condition for a hypothesis verification is fulfilled, if the projected MBR center of one partner is inside the MBR of the other. In this step, erroneous matches are eliminated resulting in the most probable affine transformation between both data sets.

6 REFINED MATCHING

The refinement of the global affine transformation is obtained in phase II of the proposed system by applying a seed image region growing algorithm. This step can be regarded as supervised segmentation. Based on the results of phase I the GIS objects are projected into the satellite image and reduced afterwards in order to avoid the effect of any transformation error. The reduction of the GIS objects is based on a distance transformation. For this purpose, the GIS objects are converted into raster objects and a value is assigned to each object pixel, which represents the shortest distance to the object boundary. In this way, only the most central areas of the objects are remaining and can be used as seed regions for the refined segmentation. Results of the seed image generation are given in Figure 7. Figure 8 shows the used distance image of the GIS objects, which are reduced to 20% of their original size. As it is visible, almost all seed regions are located in the correct image object. The result of the refined segmentation of phase II, which is based on these seed images is illustrated in Figure 9. The improvement of this refined segmentation (right image) with respect to the original segmentation (left image) is clearly visible.

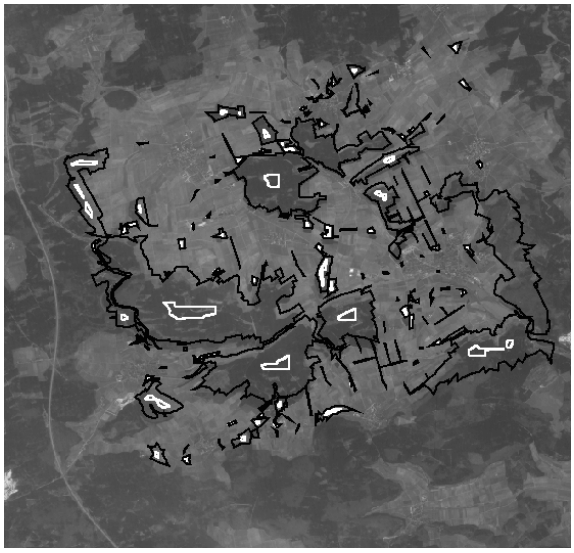


Figure 7. SPOT PAN section with projected GIS objects (black) and seed regions (white)

The individual transformation for objects that have been matched in phase II is performed in phase III by the application of an iterative closest point algorithm (ICP) as proposed by Besl and McKay (1992). This algorithm transforms the GIS object into the image based on the refined global inverse transformation for each matching polygon pair. Afterwards the minimum distance to the raster object boundary is determined for a set of points on the transformed boundary. From this point pair list an affine transformation is directly

estimated for each polygon, seperately. The GIS object is then transformed again with the refined transformation and the procedure is iteratively repeated until a certain threshold in the improvement is undergone or a maximum number of iterations is reached. The convergence of this algorithm has been theoretically proven by Besl and McKay (1992). The ICP algorithm principally can be applied if sufficient approximate values are present – as it is the case after phases I + II of our system.

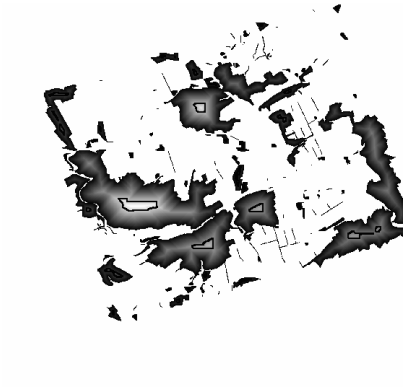
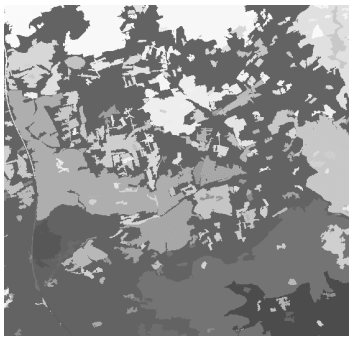
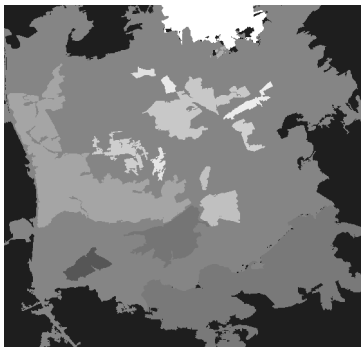


Figure 8. Distance image from projected GIS objects and seed regions from reduced GIS objects



Unsupervised segmentation, AT=6



Supervised segmentation, AT=6

Figure 9. Improvement on SPOT PAN segmentation when seed objects are used in phase II refinement

7 QUALITIY ASSESSMENT

For the quality assessment of the proposed system a set of GCPs was measured manually in a section of a SPOT scene with an extension of approximately 12km x 12km. These manually measured points were used as reference measurement for the evaluation of the accuracy of points provided by the fully automatic process. For each phase of the algorithm, the GCPs of the GIS data were transformed with either the global (phase I + II) or the individual (phase III) affine transformation and the mean, minimum and maximum point error between the automatically generated GIS points and the GCPs measured manually was determined. For phase I + II all points were inspected simultaneously. For phase III the points for each single object were considered separately. The configuration of the GCPs in the satellite image is shown in Figure 10.

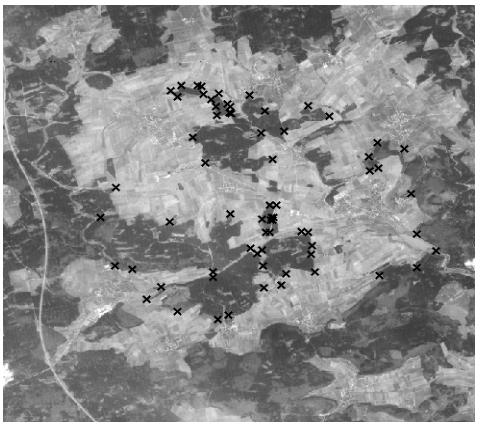


Figure 10. GCP constellation in the SPOT scene for quality assessment

As it is visible in Table 1, the mean point errors as well as the minimum and maximum point error are decreasing from phase to phase. This demonstrates the improvement of the image-map-transformation. The final result of the matching process for two individual objects is depicted in Figure 11. For these examples the corresponding GIS polygons are projected to the relevant image section. The black numbers represent the mean point errors after passing phase III, which is below one pixel in our examples.

Phase	Mean	Min	Max
I	19.87	0.87	40.11
II	4.37	0.42	14.202
III	1.181[1]	0.88[1]	1.66[1]

Table 1. Point errors of various processing phases of the proposed system in pixels

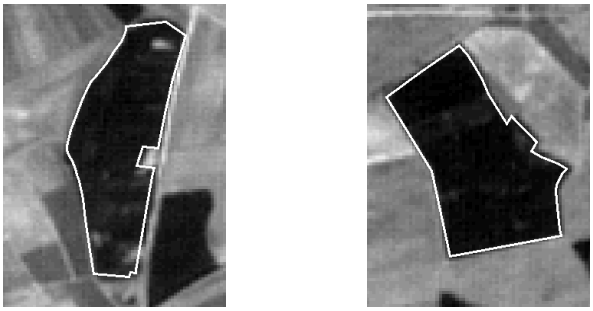


Figure 11. Selected objects with individually transformed GIS objects

[1] selected individual result

CONCLUSION

With this paper a system for fully automatic matching of a SPOT scene section with a vector map (german ATKIS data base) was introduced. With the presented three phase approach, a refinement for the global and individual affine image-to-map transformation could be obtained with each step, resulting in sub-pixel accuracy for the final GCP's. The presented procedure does not need any manual interaction and therefore has the potential to be integrated into an automatic workflow. It is pointed out, that neither pixel size nor orientation nor any other information on the satellite image was used for the demonstrated results. Potential applications of the approach are manifold, like automatic GCP generation, automatic georeferencing of satellite scenes, but also – with some minor modifications – the determination of an approximate transformation for a image-to-image registration in the context of relative orientation or DEM generation.

Although the system works quite well so far, further improvement is still possible. The calculated features are for example not completely invariant under affine transformation and hence could be substituted by invariants based on geometric moments. Furthermore, the verification step can be improved for phases I + II by applying different robust estimation techniques for the determination of the most probable global transformation. In the ICP algorithm of phase III, robust techniques could also be exploited to allow for GCP generation at partially matching objects. A further improvement is possible when using the obtained transformation(s) for detecting and matching point features, like road crossings, thus improving the number as well as the quality of the GCPs.

REFERENCES

Abbazi-Dezfouli, M. and Freeman, T. (1994): Patch Matching in Stereo-Images Based on Shape. IAPRS Vol. 30 Part 3, pp. 1-8

Arbeitsgemeinschaft der Vermessungeverwaltungen der Länder der Bundesrepublik Deutschland (AdV) (1988)

Amtlich Topographisches-Kartographisches Informationssystem (ATKIS)

Arbter, K., Snyder, W.E., Hirzinger, H. and Burkhardt, G. (1990): Application of affine-invariant Fourier descriptors to recognition of 3-D objects. IEEE Trans. Pattern Anal. & Machine Intell. vol. 12, 7, pp. 640-647.

Besl, P.J. and McKay, P.J. (1992): A method for Registration of 3-D Shapes. IEEE Trans. on Pattern Anal. & Machine Intell. vol. 14, 2, pp. 239-256.

Dowman, I. (1998): Automating Image Registration and Absolute Orientation: Solutions and Problems. Photogrammetric Record vol. 16, 91, pp. 5-18.

Drewniok, C. and Rohr, K (1997) Exterior orientation - an automatic approach based on fitting analytic landmark models, Isprs Journal Of Photogrammetry And Remote Sensing (52)3 (1997) pp. 132-145

Haala, N. and Vosselman, G. (1992): Recognition of Road and River Patterns by Relational Matching. IAPRS Vol. 29 Part B3, pp. 969-975

Hild, H. and Fritsch, D. (1998): Integration of Vector Data and Satellite Imagery for Geocoding. IAPRS, Vol. 32 Part4, pp. 246-251

Levine, M.D. and Shaheen, S.I. (1981): A Modular Computer Vision System for Picture Segmentation and Interpretation. IEEE Trans. on Pattern Anal. & Machine Intell. vol. 3, 5, pp. 540-556.
Reiss TH (1993) Recognizing Planar Objects Using Invariant Image Features (Springer)

Schickler, W. (1992): Feature Matching for Outer Orientation of Single Images Using 3-D Wireframe Controlpoints. Proc. ISPRS Congress Comm. III, pp. 591-598