FUSION OF MULTI-SENSOR IMAGES AND DIGITAL MAP DATA FOR THE RECONSTRUCTION AND INTERPRETATION OF AGRICULTURAL LAND-USE UNITS

Thomas Löcherbach *
Institut für Photogrammetrie, Universität Bonn, Germany
e-mail: thomas@ipb.uni-bonn.de

KEY WORDS: image interpretation, least-squares reconstruction, multi-sensor and map fusion, model-based segmentation, mixed-pixel problem

ABSTRACT:

The paper describes an approach to reconstructing agricultural land-use areas from remotely sensed images using digital polygon maps as prior information. Our goal is to update the geometry and class label of agricultural parcels. The approach integrates an estimation of the vector polygons and the derivation of object (i.e. land-use unit) related features for the determination of the object classes.
Both a feature edge model for the transition between two land-use units and the assumption of homogeneity are used to reconstruct the land-use units in a least squares approach. The theoretical concept and its technical realization are described, first results are presented and a critical evaluation of the results with a discussion of possible extensions is given.

KURZFASSUNG:

Der Aufsatz beschreibt einen Ansatz zur Rekonstruktion landwirtschaftlich genutzter Gebiete aus Fernerkundungsdaten mit digitalen Polygonkarten als Vorinformation. Unser Ziel ist es, die Geometrie und Klasse landwirtschaftlicher Parzellen zu aktualisieren. Der Ansatz integriert eine Schätzung der Vektorpolygone mit der Herleitung objektbezogener (d.h. auf die Landnutzungseinheit bezogener) Merkamle zur Bestimmung der Objektklassen.
Sowohl ein Merkmalskantenmodell für den Übergang zwischen zwei Landnutzungseinheiten und die Annahme der Homogenität werden benutzt um die Landnutzungseinheiten in einem kleinste-Quadrate-Ansatz zu rekonstruieren. Das theoretische Konzept und seine technische Realisierung werden beschrieben, erste Ergebnisse werden vorgestellt und eine kritische Bewertung der Ergebnisse mit einer Diskussion möglicher Erweiterungen gegeben.

1 INTRODUCTION

Automation in photogrammetry has led to an increasing integration of geometric and thematic aspects. Thus automatic determination of object geometry and automation of interpretation have to be linked conceptually as well as technically. Our aim is to classify images of agricultural areas. Many approaches perform tasks like surface reconstruction and point determination in object space. Simultaneously they derive the reflectance properties of the object surface, e.g. to produce orthophotos (Ebner *et al.* 1987, Fua and Leclerc 1993, Heipke 1992, Helava 1988, Wrobel 1987). Also in classification pixel-based approaches are more and more being replaced by object-based approaches in order to link geometric and thematic aspects and to overcome some of the deficiencies of pixel based approaches (Janssen and van Amsterdam 1991, Johnsson and Kanonier 1991, Lemmens 1988, Lemmens 1992, Mason *et al.* 1988):

- During classification multitemporal and multisensoral data as well as data from other sources (e. g. maps, knowledge) have to be fused for several reasons, e. g. the problem with cloud coverage or the phenology of vegetation. The separability of classes can be increased by making use of domain knowledge which is related to objects, not to pixels etc. The pixel is not a temporarely stable reference for the fusion of these data. This is especially true for the phenology of vegetation.
- Purely pixelbased approaches cannot cope with the mixed pixel problem. Object boundaries need to be known or estimated to solve this problem.
- Objects rather than pixels are classified, leading to a direct link to geographic information systems (GIS) (cf. Lemmens 1988).

As a prerequisite for an object-based classification object geometry and location must be known. Janssen *et al.* 1990 take the object geometry from a GIS. In most cases, however, actual land-use boundaries are not available. Due to the high complexity of natural scenes, all information sources must be used for the recovery of the geometry, including images, map data and knowledge. Schneider 1992 takes advantage of the straightness of field-boundaries to improve on the mixed-pixel problem. Pan and Förstner 1994 and Janssen *et al.* 1992 incorporate knowledge about the field structure. Lemmens 1988 states ways to check initial boundaries from a vector-format-GIS. He simultaneously incorporates thematic information from the GIS,

*Supported by Bundesministerium für Forschung und Technologie, "Erfassung des biophysikalischen Zustandes landwirtschaftlicher Kulturpflanzenbestände während des Wachstums mit Hilfe multitemporaler ERS-1 SAR-Daten", Förderungskennzeichen 01QS90015

both to predict the type of the boundary and to search for possible new boundaries.

Our approach uses vector polygons, e.g. legal boundaries, out of date land-use boundaries or the result of a segmentation (cf. Pan and Förstner 1994) as prior information.

In the next section the underlying model is described. Its technical realization is outlined in section 3 (cf. also Löcherbach 1992). Details of the implementation are not described here. First results are presented in section 4. Finally in section 5 the advantages, drawbacks and possible extensions of the procedure are discussed.

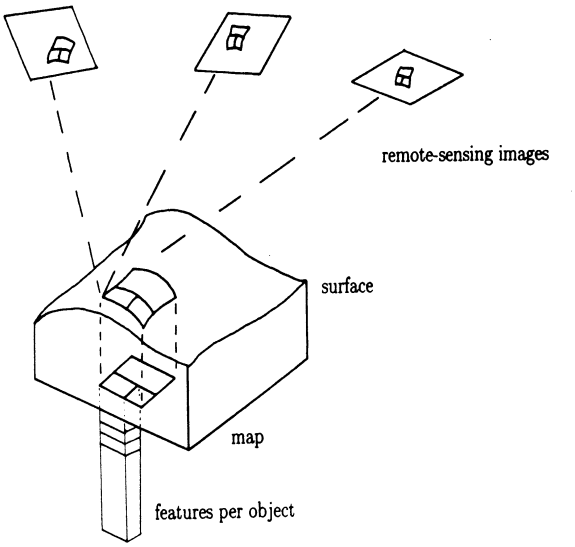

Figure 1: Map-generated hypotheses of land-use units (from Förstner 1991)

2 A MODEL FOR FUSING RASTER AND VECTOR-DATA

In this section the assumptions made by the underlying model are described (cf. fig. 1).

- The object-model contains 3D polygons, representing the geometry of the land-use units. (The terms land-use unit, object, field and area here are taken as synonyms.)

  The radiometric part consists of one feature vector per object, e.g. the field mean, within-field variance, or a field histogram (e.g. in an ERS-1 scene). This represents the object class.

- The geometry of the image model is represented by 2D polygons. The radiometric part contains the assumption of homogeneous features within a given object and a feature edge model describing the transition between two neighboring fields along the land-use boundary.

The feature edge model is matched to the observed edge in the images by least-squares matching. In other words a global optimization is applied to simultaneously estimate the unknown parameters from a set of multitemporal and multisensoral and hence multiresolution images. Unknown parameters describe object geometry and radiometry, global transformation and type and form of the feature edge.

Due to discrepancies between the model and reality, a pre-segmentation and a robust estimation are necessary to cope with isolated objects or inhomogeneous regions.

Generalizations of this model are possible and will be discussed later.

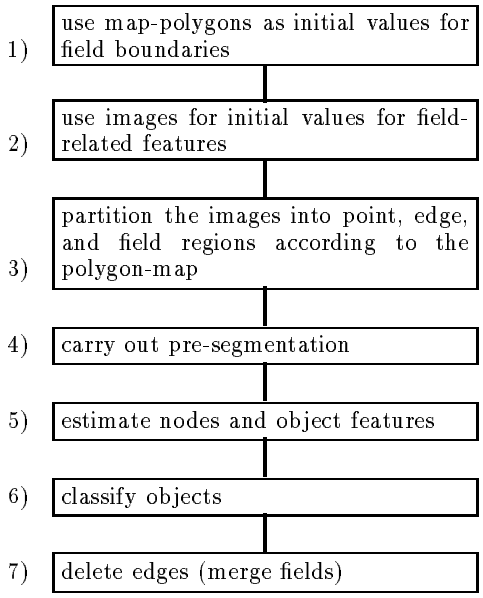

Figure 2: flow-chart

3 TECHNICAL REALIZATION

In this section the technical realization is described. A general description of the algorithm (cf. fig 2) is followed by a more detailed description of the observation equations.

3.1 Overall Algorithm

1. Polygons are given as initial values for the geometry. Either map polygons or polygons obtained from a segmentation (Pan and Förstner 1994) can be used.
2. Images yield initial values for object features.
3. A partitioning of the images into point, boundary and field regions is done by using a distance transformation of the boundaries and a connected component algorithm for the fields (cf. also Fig. 4).

   Thus each pixel can be identified as belonging to a certain field or to a point or to a boundary surrounding a field. This information is needed in the following two steps.
4. A local pre-segmentation is carried out within each field region. This step serves to determine initial values for the features, to determine a weight for each

observation and eventually to split a field. After this step the topology of the polygons remains fixed.

5. In the adjustment process geometry and radiometry of the objects are estimated. Different types of observation equations are used:
   We use a feature edge model for pixels within boundary-regions to estimate geometry and radiometry. We estimate radiometry from pixels within a field region and we use observation equations to include prior information.

6. Objects can now be classified (Janssen and van Amsterdam 1991).

7. In a final step objects can be merged, if the same class for two neighboring objects was estimated in the classification step or a low probability was found for an edge in the estimation step.

This paper deals with the first 5 steps, i.e. the reconstruction of the geometry and the derivation of field-related features. Due to the highly nonlinear observation equations the estimation process must be done iteratively. The partitioning has to be repeated in each step of the iteration. The observation equations for the estimation are now described.

3.2 Observation Equations

The following notation is used:

$$\begin{array}{lcl} \text{set of images} & \mathcal{I} & = [1,\ldots,i,\ldots,I] \\ \text{set of channels} & \mathcal{K} & = [1,\ldots,k,\ldots,K] \\ \text{set of pixels} & \mathcal{J} & = [1,\ldots,j,\ldots,J] \\ \text{set of nodes} & \mathcal{N} & = [1,\ldots,n,\ldots,N] \\ \text{set of edges} & \mathcal{E} & = [1,\ldots,e,\ldots,E] \\ \text{set of areas} & \mathcal{A} & = [1,\ldots,a,\ldots,A] \end{array} \quad (1)$$

We thus have the observed intensity values g_{ijk} to estimate the unknown parameters, i. e.

- the object coordinates of the nodes $P_n = (X, Y, Z)_n$ of the polygons
- the global transformation parameters p_i of image i
- the object-related feature vector m_{iak} of area a
- the parameters q_{eik} describing type and form of the edge e

We use different observation equations

- for pixels g_{ij_ek} close to a boundary e
- for pixels g_{ij_ak} within an area a
- for prior information.

3.2.1 Observation Equation for Boundaries

The boundary $\overline{n_\alpha, n_\omega}$ has been transformed to the pixel system (r, c), cf. Fig. 3. A y-axis perpendicular to the boundary then is positioned through each observed pixel g_j. The edge-model $f = f(y)$ is matched to the signal edge $g = g_j$ by a radiometric and a geometric transformation T_R, T_G (cf. Förstner 1992, Pertl 1984).

$$\boxed{g_{ij_ek} = T_R(m_{ia_rk}, m_{ia_lk}; f(T_G(P_\alpha, P_\omega, p_i, q_{eik}))} \quad (2)$$

The radiometric transformation consists of a linear shift and scale described by the features m_{ia_rk} and m_{ia_lk}. a_r and a_l denote the areas on the right- and left-hand sides of the boundary.

The geometric transformation is obtained by writing the y-coordinate as a function of the object coordinates P_α and P_ω, the global transformation paramerters p_i and the parameters q_e.

The feature edge currently is modelled by a normalized and smooth function

$$f(y) = \frac{1}{1 + e^{-y}} \quad (3)$$

This model has proved to be useful. However, it will be shown also in analyzing the examples that also other types of models should be included, e. g. a line model (cf. Malik and Perona 1990). The type of model will then have to be estimated in advance.

The observation equation 2 has to be linearized.

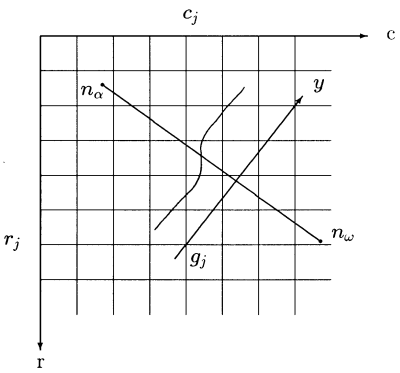

Figure 3: For every observed intensity value g_j a y-axis is positioned perpendicular to the line $\overline{n_\alpha n_\omega}$

3.2.2 Observation Equation for Homogeneous Areas

The derivation of equation (2) is based on the presence of an edge. However, only pixels within a certain distance of the edge are influenced by the reflectance of the neighbouring field and may thus be used for estimating the position

of the edge using equation (2). For pixels within an area a we use as observation equation

$$\boxed{E(g_{ij_ak}) = m^0_{aik} + dm_{aik}} \tag{4}$$

since according to our model the area a is supposed to have a homogeneous feature m_{aik} in channel k of image i.

3.2.3 Prior Information

Prior information can be used in addition to the observed intensity values of the images. Singularities may easily occur if the edges in the map are not visible in the image, the geometry of the polygons is unfavourable, points are not well defined by intersecting edges, etc.
The following prior information may be used if available:
For nodes:

$$\boxed{E(P_n) = P^0_n + dP_n} \tag{5}$$

For transformation parameters:

$$\boxed{E(p_i) = p^0_i + dp_i} \tag{6}$$

The strength of this prior information may be varied by means of a priori weights $w_n \to 0$ to $w_n \to \infty$ respectively $w_i \to 0$ to $w_i \to \infty$, representing weak or strong prior information (cf. Mikhail and Ackermann 1976).

4 EXAMPLES

This section gives some first results of the approach together with with an evaluation of the examples. Conclusions drawn for future work are given in the next section. The examples are organized as follows: Example 1 again shows the principle of the algorithm. Examples 2 through 7 are simulated examples. The purpose of these synthetic data is to allow the analysis of an isolated and exaggerated effect under well-defined conditions. Finally a test on real data is given where we have to cope with a mixture of these and other effects.

- Example 1 (cf. Fig. 4) shows the principle of the algorithm. Vector polygons are given as initial values for the geometry (upper left image). They are superimposed on each image (lower left image). The image is partitioned into regions around nodes, boundaries and inside a field (upper right image). The normal equations are calculated according to the partitioning. Partitioning and solving the normal equations is done iteratively, so that finally the new polygon together with transformation parameters and feature vectors is estimated (lower right images).

- Example 2 and 3 (cf. Fig. 5 and 6) relate to the topology of the initial polygon. In the left-hand image of figure 5 the initial polygon is superimposed on the image. From the partitioning (middle image) it can be seen that the boundary in the middle contains three internal points. As a consequence the boundary can be fitted to the image data compared to a joint (right image). In figure 6 the initial boundary from the map

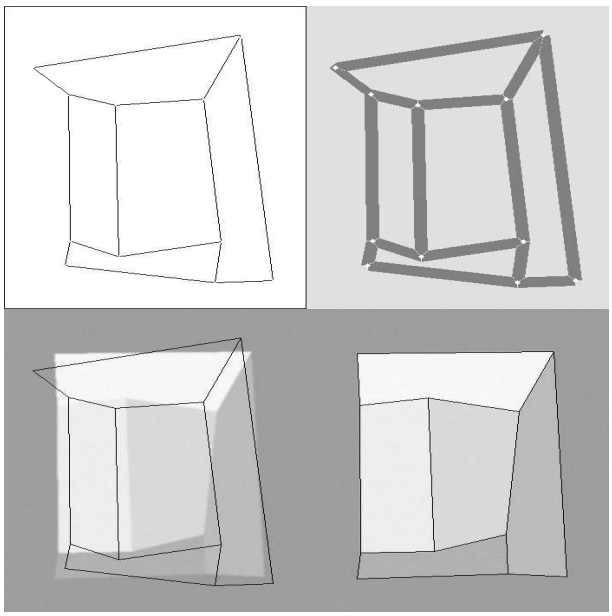

Figure 4: Principle of the algorithm: the initial polygon (upper left) is transformed to each image (lower left); the image is partitioned into regions (upper right); different observation equations are used for pixels within a boundary- or a field-region. After an iteration the adjusted polygon fits the image data (lower right)

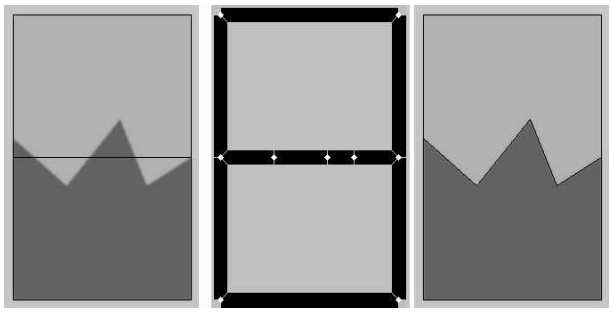

Figure 5: Provided that the topology of the initial polygon is complete, the initial boundary is fitted to the image edge compared to a joint

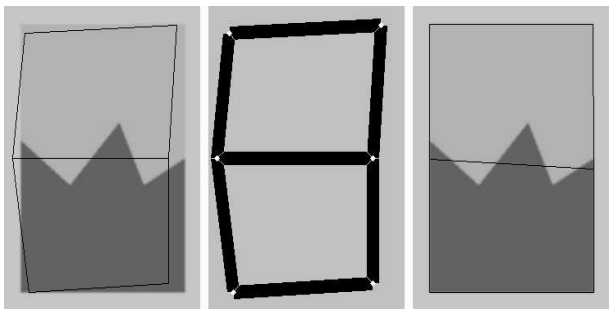

Figure 6: In the case of incomplete topology, the initial boundary cannot be merged with the image edge. The remaining boundaries, however, are adjusted correctly

does not contain these internal points, hence it cannot be fitted to the image. The initial boundary from the map then is retained. The remaining boundaries, however, can be adjusted correctly.

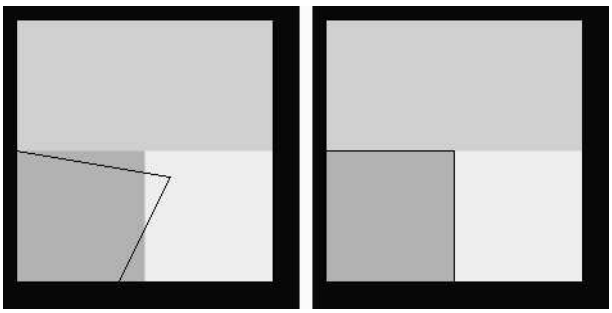

Figure 7: The topology of the initial polygon is incomplete. Nevertheless the existing part of the polygon can be matched to the image

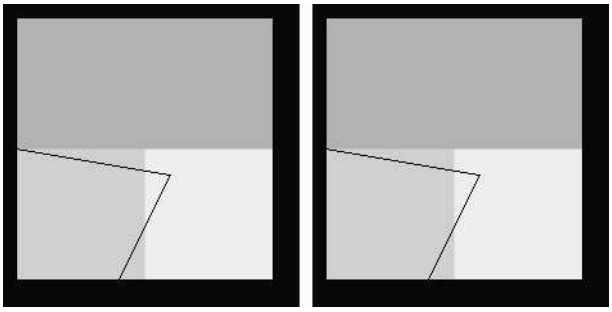

Figure 8: The topology of the initial polygon again is incomplete. However, the differences of the intensity values are more extreme. The algorithm cannot cope with these conditions without a local pre-segmentation

- In example 4 and 5 (cf. Fig. 7 and 8) a boundary that can be seen in the image is not contained in the initial map. Nevertheless the existing part of the polygon in figure 7 can be matched to the image. In the example of figure 8 this is not the case. The reason is that the intensity value of the lower left area lies between the two different intensity values of the larger area. In this case the model assumption of homogeneity within each area is too severely violated and a local pre-segmentation is necessary.

- Example 6 (Fig. 9) shows two fields of the same intensity value that are separated by a feature *line*. However, our algorithm currently contains only a feature *edge* model and cannot cope with the appearance of a line. The nodes of the respective boundary are undefined in row direction. The remaining boundaries are merged correctly.

- In example 7 (cf. Fig. 10) the 3D-shape was reconstructed from three images simulating a flight with an aerial camera, thus using perspective projection.

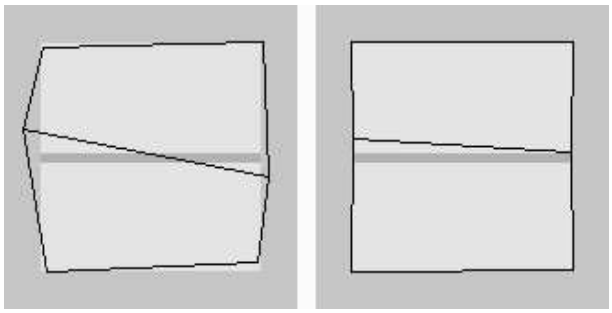

Figure 9: The edge model does not fit the line in the image

The imaging-geometry is given in Table 1. The rotation angles were set to 0, the flight direction was in Y-direction. Table 2 shows the convergency of the iteration. Observe the fast convergency. The low empirical accuracy results from the large pixelsize.

No. of images: 3
c = 150mm
Pixelsize: 100 μm
Fleight-height: 1500 m
Baseline: 1000 m
Extension of the terrain: 500 m
No. of polygon-nodes: 76

Table 1: Imaging geometry for three images.

Iteration	μ_X [m]	μ_Y [m]	μ_Z [m]	$\hat{\sigma}_0^2$
0	0.00	0.00	2.35	
1	0.18	0.08	1.25	32.6
2	0.17	0.09	0.67	17.6
3	0.16	0.07	0.39	10.9
4	0.16	0.07	0.29	7.6
5	0.16	0.07	0.29	6.7
6	0.16	0.07	0.28	6.5

Table 2: Accuracy of the result in each iteration with three images.

- Example 8 (cf. Fig. 11 and 12) shows a test on real image data. The "map data" are still simulated data. Only the area inside the surrounding polygon was adjusted. Figure 11 shows the start and figure 12 shows the result of the adjustment which took eight iterations. Boundaries that are rotated and even boundaries with a relatively strong parallel shift are adjusted correctly. Problems occur especially for the paths in the map. In the upper left area there are two crossing paths. In the image they appear as a line or even as a composition of two lines. Here the algorithm yields unsatisfactory results, destroying the parallelism of the two boundaries. In the lower left area the map contains some paths that are not visible in the image. Here the initial geometry from the map is retained, due to prior information that is stronger than the observations. This is in accordance to our intention of preserving the old map information if no information from the observed image data is available. This also holds for two long within-field-boundaries, one in the light field in the upper part of the image, the other in the dark field in the middle right area. Due to missing image information the initial boundary is retained. Some problems still occur for short edges where the intensity profile along the boundary varies.

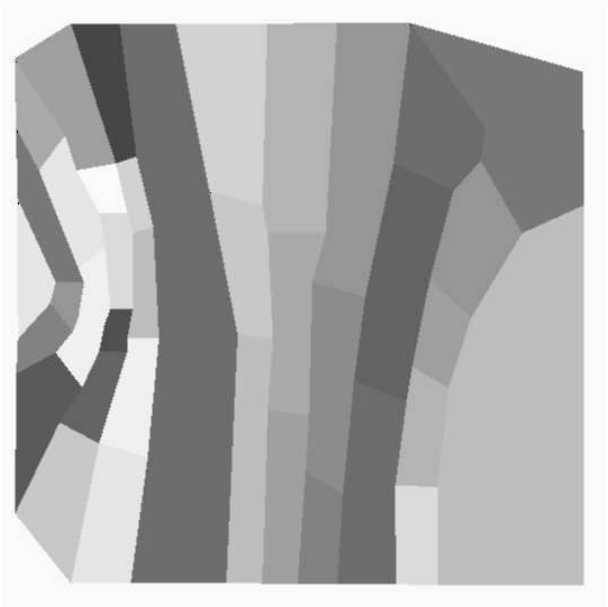

Figure 10: Three images of a valley were simulted to test the reconstruction of the 3D-shape

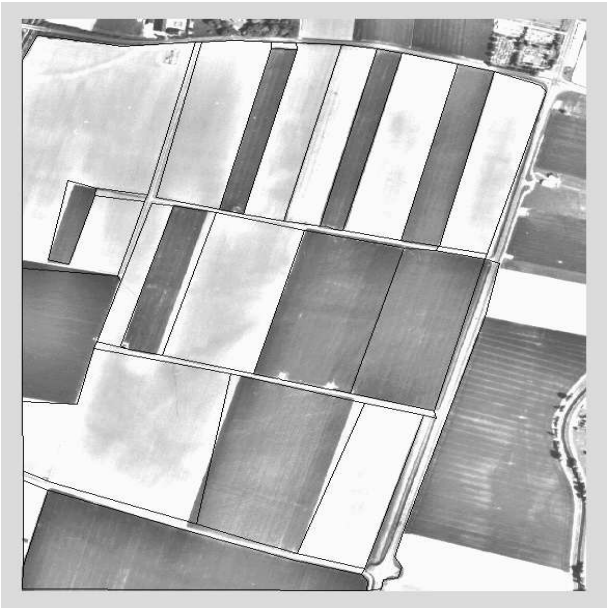

Figure 11: shows a test on real data

5 CONCLUSIONS AND FUTURE WORK

Many authors have stated the necessity for an interaction between GIS and image data. The GIS serves as additional information source for analyzing the remote sensing data. On the other hand the extracted information can be fed back into the GIS. Here a possible approach was shown to update the geometry of a vector-format GIS. Conceptually the approach links geometric and thematic aspects as it simultaneously estimates object-related feature vectors to classify the objects. The objects are used as references for the fusion of map and multitemporal and multisensoral image data. The approach can be incorporated in a broader context for image analysis (cf. Fuchs *et al.* 1993). There are several advantages of the approach:

Figure 12: shows the estimated polygons

- Closed polygons are derived.
- The approach helps to handle the mixed-pixel problem in remote sensing. A distinction of pure and mixed (=boundary) pixels is possible. The transition between two classes is modeled. By projecting the map into the images a resampling is avoided, which is useful for classification.
- Map data are used as a reliable information source in image analysis.
- The approach is both area- and edge-based.
- If the image data meet the assumptions from the model, the convergence behaves favourable.
- The unknown parameters, object coordinates and transformation parameters are also treated as observations with variable a priori weights. This unified approach for the adjustment has turned out to be useful: In case of missing or bad image information the geometry from the map is preserved and singularities are avoided.
  By varying e.g. the a priori weights either X, Y- or Z-coordinates or orientation parameters or a combination can be estimated.

From analysing the examples several conclusions can be drawn for future work:

- A *line* model must be integrated or possibly a combination of different types of *lines* and *edges*, e.g. to handle ditches and paths in images of agricultural areas.
- Constraints can be included such as "paths have parallel boundaries".

- The type of feature edge should be predicted e.g. by investigating the feature transition for each boundary in advance. Possibly a change of the kind of transition along the boundary (e.g. a shading) must be taken into account. The type of edge also can be predicted from the thematic information from the map (cf. Lemmens 1988).

- A shading of the features within a field can also be modeled, e.g. by means of a triangulation.

- A robust estimation should be included by varying weights for the observations depending on the local pre-segmentation. The pre-segmentation should be one of the main issues to be investigated.

- Depending on the quality of the initial values the algorithm should start at a higher level of a pyramid.

- A digital terrain model can be included. In this case the number of internal points in the boundaries must be increased, depending on the undulation of the terrain surface. A problem of matching the nodes in the different images does not exist, because the nodes are given in object space.

- The approach could be applied to other problems than analysing natural scenes provided that the objects are composed of homogeneous areas.

References

Ebner, H.; D., Fritsch; W., Gillesen; Ch., Heipke (1987): Integration von Bildzuordnung und Objektrekonstruktion innerhalb der digitalen Photogrammetrie. *BuL*, 55(5):194 – 203, 1987.

Förstner, Wolfgang (1991): Object Extraction from Digital Images. *Schriftenreihe des Instituts für Photogrammetrie der Universität Stuttgart*, 15:269 – 279, 1991.

Förstner, Wolfgang (1992): Least-Squares Matching. In: Haralick; Shapiro (Eds.), *Robot and Computer Vision, Vol. II.* Addison Wesley, 1992.

Fua, P.; Leclerc, Y. G. (1993): Object Centered Surface Reconstruction: Combining Multi-Image Stereo and Shading. *Technical Report, Workshop on Understanding Aerial Images*, 1993.

Fuchs, C.; Löcherbach, Th.; Pan, H.-P.; Förstner, W. (1993): Land-Use Mapping from Remotely Sensed Images. *Colloquium on Advances in Urban Spatial Information and Analysis, 19-22 Oct. 93, Wuhan, China*, 1993.

Heipke, Ch. (1992): A global approach for least squares image matching and surface reconstruction in object space. *Photogrammetric Engineering and Remote Sensing*, 58(3):317 – 323, 1992.

Helava, U.V. (1988): Object-Space Least-Squares Correlation. *PE&RS*, 54(3):711 – 714, 1988.

Janssen, L. L. F.; van Amsterdam, J. (1991): An Object-Based Approach to the Classification of Remotely Sensed Images. *Proceedings, IGARSS-Symposium, Espoo Finland*, 1991.

Janssen, L. L. F.; Jaarsma, M. N.; van der Linden, E. T. M. (nov 1990): Integrating Topographic Data with Remote Sensing for Land-Cover Classification. *PE&RS*, 56(11):1503 – 1506, nov 1990.

Janssen, L. L. F.; Schoenmakers, R. P. H. M.; Verwaal, R. G. (1992): Integrated Segmentation and Classification of High Resolution Satellite Images. In: Molenaar, Martien; Janssen, Lucas; van Leeuwen, Hans (Eds.), *Proc. of Int. Workshop on Multisource Data Integration in Remote Sensing for Land Inventory Applications*, pages 65–84, Delft, 1992. International Association of Pattern Recognition TC7.

Johnsson, K.; Kanonier, J. (1991): Knowledge Based Land-Use Classification. In: *Proceedings, IGARSS-Symposium, Espoo, Finland*, 1991.

Lemmens, M. J. P. M. (1988): GIS - Digital Image Interaction. In: *ISPRS-Congress, Kyoto*, 1988.

Lemmens, M. J. P. M. (1992): Integration Levels of Topo Data Bases and Geo Imagery. *Proc. of the ISPRS and OEEPE joint workshop on Updating Digital Data by Photogrammetric Methods*, pages 141–149, 1992.

Löcherbach, Th. (1992): Reconstruction of Land-Use Units for the Integration of GIS and Remote Sensing Data. In: Molenaar, Martien; Janssen, Lucas; van Leeuwen, Hans (Eds.), *Proc. of Int. Workshop on Multisource Data Integration in Remote Sensing for Land Inventory Applications*, pages 95–112, Delft, 1992. International Association of Pattern Recognition TC7.

Malik, J.; Perona, P. (1990): Detecting and Localizing Edges Composed of Steps, Peaks and Roofs. *Proceedings 3rd ICCV, Osaka, Japan*, pages 52 – 57, 1990.

Mason, D. C.; Corr, D. G.; Cross, A.; Hoggs, D. C.; Lawrence, D. H.; Petrou, M.; Tailor, A. M. (1988): The Use of Digital Map Data in the Segmentation and Classification of Remotely-Sensed Images. *Int. Journal of Geographical Information Systems*, 2(3):195–215, 1988.

Mikhail, E. M.; Ackermann, F. (Contributions) (1976): *Observations and Least-Squares.* University Press of America, 1976.

Pan, H.-P.; Förstner, W. (1994): Segmentation of Remotely Sensed Images by MDL-principled Polygon Map Grammer. *Proceedings ISPRS Comm. III Symposium, Munich*, 1994.

Pertl, A. (1984): Digital Image Correlation with the Analytical Plotter Planicomp C100. In: *Int. Archives of Photogrammetry*, 1984.

Schneider, W. (1992): Sub-Pixel Analysis of Optical Satellite Image Data. *Poster presented at the ISPRS-Congress, Washington D. C.*, 1992.

Wrobel, B. (1987): Digitale Bildzuordnung durch Facetten mit Hilfe von Objektraummodellen. *BuL*, 55(3):93 – 101, 1987.