Knowledge Based Analysis of Satellite Images

Klaus-Jürgen Schilling, Thomas Vögtle, Peter Müßig, Dept. of Photogrammetry & Remote Sensing, Karlsruhe University, D-76128 Karlsruhe, Germany
Tel.: ++49721/6084131, Fax: ++49721/694568
schi@ipf.bau-verm.uni-karlsruhe.de, voegtle@ipf.bau-verm.uni-karlsruhe.de, muessig@ipf.bau-verm.uni-karlsruhe.de
WWW-Server: http://ipfr.bau-verm.uni-karlsruhe.de/

KEYWORDS: image analysis, classification, ATKIS, rectification, feature extraction, semantic modelling

ABSTRACT:
This paper shows several working steps for updating the "Digitale Landschaftsmodell 200 (DLM 200)" using satellite images. It is based on a two-step approach: *verification* and *classification*. First the existing semantic model (DLM 200) is used for the knowlegde based object oriented analysis of the satellite images. At the second stage the information gained from the first step serves to prove and update the DLM 200.
Since the DLM 200 is produced by digitizing the map layers of the "Topographische Übersichtskarte (TÜK 200)", typical cartographic aspects (e.g. generalisation, displacement) have to be considered. Some examples illustrating these effects on a representative class (e.g. forest) of the DLM 200 will be shown. After the determination of these geometric relations between the DLM 200 and the images (i.e. a rectification of the satellite images) the "knowlegde", based on the DLM 200, will back up the object based analysis of the satellite images. Image areas which do not fit the DLM 200 will be examined at the second stage.
The classification - a minor topic in this presentation - has to assign the changes detected in the course of the verification to appropriate classes of the DLM 200. This process will use the parameters of the image analysis as additional information.

KURZFASSUNG:
Der vorliegende Beitrag behandelt einzelne Arbeitsschritte auf dem Weg zur Fortführung des DLM 200 unter Verwendung von Satellitenbildern. Das Verfahren beruht auf einem zweistufigen Ansatz: Verifikation und Klassifizierung. Das vorhandene semantische Modell (DLM 200) unterstützt zuerst die wissenbasierte objektbezogene Klassifizierung der Satellitenbilder. Die aus der Klassifizierung gewonnenen Ergebnisse werden dann zu Überprüfung und ggf. Fortführung des DLM 200 genutzt.
Da das DLM 200 aus praktischen Gründen durch Digitalsieren der Topographischen Übersichtskarte (TÜK 200) erstellt wird, müssen typische kartographische Gesichtspunkte (z.B. Generalisierung, Verdrängung) berücksichtigt werden. Anhand von Beispielen einer repräsentativen Klasse des DLM 200 (hier: Wald) werden diese Effekte verdeutlicht. Sind nun die geometrischen Beziehungen zwischen DLM 200 und Satellitenbild bekannt (Entzerrung des Satellitenbildes), kann das DLM 200 als zusätzliche Informationsquelle für eine objektorientierte Analyses des Satellitenbildes herangezogen werden. Bildobjekte, die nicht mit dem DLM 200 übereinstimmen, werden in einem zweiten Schritt untersucht.
In einem Klassifizierungsprozeß werden nun die in der Verifikationsphase festgestellten Veränderungen den entsprechenden Klassen des DLM 200 zugewiesen. Dieser Prozeß kann die Parameter der Bildanalyse als zusätzliche Informationsquelle nutzen.

1 Introduction

The project presented in this paper is part of a German research activity called *semantic modelling* where different photogrammetric and surveying institutions are involved. The aim of the IPF (Institut for Photogrammetry and Remote Sensing, University of Karlsruhe) is to investigate the reciprocal effects and interactions of satellite image analysis and the topographic information system ATKIS (<u>A</u>mtliches <u>T</u>opographisch-<u>K</u>artographisches <u>I</u>nformationssystem) of the German surveying administration, which is built up currently. Its topographic part consists of 3 data layers with different abstraction levels of information with regard to the corresponding scales: DLM 25 (<u>D</u>igitales <u>L</u>andschafts-<u>M</u>odell 1 : 25 000), DLM 200 (1 : 200 000) and DLM 1000 (1 : 1 000 000). In a first step the DLM information will be built up by digitizing analog maps, so the DLM objects contain all cartographical influences like generalisation and local or regional displacements.

The first topic of our investigations is an improvement of satellite image analysis by using additional non-spectral information of the DLM 200. On the base of a improved image analysis the possibilities of updating the contents of DLM 200 will be the second topic (fig. ??).

Therefore only those objects can be taken into account, which can be extracted from image information, i.e. mainly those with an areal characteristic (e.g. water, forest, settlements, agricultural structures etc.).

2 Conception of an extended satellite image analysis

As additional non-spectral information for an improved image analysis the contents of DLM 200 will be used. In the DLM 200 data base the main permanent object-classes, which can be represented in a scale of 1 : 200 000, are stored digitally. In this case the image information of commonly available satellite systems such as Landsat/TM, SPOT or ERS-1 contains normally more details in information than maps in such scale, so remotely sensed data are on principle suitable for updating a DLM 200. But this image information is not abstracted or generalized like a map would be, so we often have to expect a mixture of more than one object inside one pixel. Another general problem in usual satellite image analysis is the pixel-oriented classification (e.g. like *Maximum Likelihood Classification* or *Cluster Analysis*), which takes only the spectral characteristic of only a single pixel into consideration to classify its semantic meaning. The experiences in the last years have shown, that the results which can be achieved with this methods, are not in every case satisfying, especially for specific tasks like environmental monitoring or updating landuse databases.

Figure 1: Interaction between image analysis and topographic database (DLM 200)

Therefore in the approach described in this paper an extended pixel-oriented classification will be only a small part of a more general object-oriented classification. The result of this image analysis should be a higher/optimal differencing between the relevant 'reflection classes'. Afterwards a knowledge-based sophisticated decission process will be used for updating DLM 200 information. So the whole procedure of image analysis is devided into 3 main steps:

1. local, error-tolerant matching process of image- and DLM 200-data
2. verification process using DLM 200 information
3. object-oriented classification process

The matching process in this case can not be carried out as usual, i.e. one general transformation for the whole scene. Because of the (necessary) local, unsystematic displacements in maps – especially in those of a small scales – a local object matching procedure has to be developed. Additionally this procedure must have an error-tolerant characteristic with regard to the (above mentioned) generalisation of the object contour lines in the database.

After matching the relevant objects a verification process can be started based on the a priori knowledge given by the DLM 200 about the semantic meaning of these objects. Therefore several descriptions for areas of the same class/meaning can be created, which are not only based on spectral features (e.g. mean value, variance, texture etc.) but also on structural, geometrical and other attributes (neighborhood relations, dimensions, form, topography etc.)

These descriptions can now be used in a object-oriented classification process to determine the classes which are pre-defined by the DLM 200. Other classes – like different agricultural areas – have to be related to a semantic meaning by a general knowledge-based analysis (e.g. semantic network, expert system), where all available a priori information has to be used.

3 Matching process

In order to use the DLM 200 as additional information for classification or respectively updating the DLM 200 by means of the classification, it is necessary to determine the geometric relations between the DLM 200 and the satellite images. Because of the generalisation caused by the basis (TÜK 200) of the DLM 200 the usual methods of rectification (e.g. polynoms) cannot be applied. Therefore the method to be used has to be able to take care of the local effects of the generalisation. This method should be automated in practical operation due to the huge number of ground control points necessary for that purpose. The following procedure is planned for this reason:

- A „coarse“ rectification based on the parameters of the satellite images (e.g. inclination, coordinates of the center) will limit the area of interest for the next steps.
- Searching significant features in the DLM 200 and assigning of these objects to the correspondent objects of the satellite image, which have to be transformed from raster format to vector format (feature extraction). This leads to an improved global rectification of the satellite image (Haala, 1993).
- Local and robust matching of all relevant objects of the DLM 200.

First investigations of the size and the influence of the generalisation of the DLM 200 have been done by examination of forest boundaries. For that purpose typical boundaries were digitized from maps of scale 1:25000, 1:50000, 1:100000 and 1:200000 and compared to each other and to the boundaries extracted from satellite images. In the sequel of the work these examinations will be extended to further classes of the DLM 200.

4 Verification process

The aim of the verification process is to find valid and effective descriptions for the objects in satellite images based on the a priori knowledge / semantic model of the DLM 200. With this information we have a prediction of the semantic meaning for every segment in the whole image. Even if some areas in the images have changed since the last update of the DLM 200 data (e.g. increasing of settlement areas, clearing areas in forest etc.), it can be supposed, that by far most information of this database is true. Therefore several descriptions for areas of the same class/meaning – out of a huge number of samples – can be created. Beside of some well known spectral features, which still are the main attributes, also structural, geometrical and other features are used to build up descriptors, because in some cases reflection classes and semantic classes of DLM 200 are not identical. Here it has to be taken into account, that in DLM 200 data only the permanent landuse classes can be stored, not the periodic changing areas like for instance agricultural field, therefore in that areas great differences will appear between only one semantic class 'agriculture' in DLM 200 and many reflection classes in the satellite image.

Those object descriptions can be defined:

- statistical parameters (e.g. mean value / variance of reflection values)
- texture parameters
- structure parameters
- size parameters
- form parameters
- topographical parameters (e.g. slope-, exposition-, elevation-dependence)

etc.

For each class defined in DLM 200 a specific suitable combination of the parameters listed above is used in the verification and classification process, i.e. for homogeneous objects (e.g. water, agricultural fields) another combination will be necessary to describe the class than for inhomogeneous objects (e.g. settlement areas, forest).

Because of the huge number of training areas the true class will be found by majority. Especially for spectral parameters, single outlayers, i.e. objects which have changed compared with the DLM semantic, will have only a small effect on the descriptors or can be seperated statistically by dividing the real distribution into normal distributions which represent sub-classes. Therefore valid and robust descriptors can be extracted without disturbing influences. A similar procedure can be carried out with the other non-spectral parameters.

By means of this descriptions an object-oriented classification is performed where an extended Maximum-Likelihood classification is only a small part of a more general approach. As result all objects are detected which have not changed and are equivalent to the DLM semantic, and on the other hand those areas which are different. It can be expected, that most objects can be verified and excluded from the further process.

5 Classification process

After verification only the remaining areas have to be submitted to a classification process (fig. **??**).

Because of the lack of prediction for this areas, in a first step the same descriptions as in the verification process are used to try to relate this areas to another predifined DLM classes. As result some objects can be classified (e.g. forest which is now a clearing area, agricultural field which is now settlement area etc.).

The rest has to be submitted to a general image analysis procedure by means of external knowledge base (segmentation process, semantic network, context information, relations between objects, achived data etc.). The result again is a number of areas which can be assigned to one of the semantic classes.

The further treatment of the last few areas which are still unclassified can either be the assignement to the most probable class or the rejection of that area.

6 Conclusions

In this approach the pixel-oriented classification is extended to a object-oriented procedure by means of additional non-spectral information of a topographic data base (DLM 200). Not only spectral parameters are used to get an extended image analysis, but also structural, geometrical and other features.

The presented approach is suited for all regions with (future) digital Land Information Systems (LIS) / Geo-Information Systems (GIS). Especially in Central Europe and North America the administrations will establish local and global data bases of this kind (e.g. ATKIS in Germany).

But also this method of an extended image analysis can be applied in many cases in developing countries by digitizing analog maps, if only specific tasks have to be solved and therefore only a few layers have to be created. Every additional information leads to an improvement of this object-oriented classification process.

Figure 2: Procedure of an extended satellite image analysis

References

[AdV89] AdV: Amtliches Topographisch-Kartographisches Informationssystem (ATKIS) – Gesamtdokumentation. Hannover 1989

[Argialas90] D. P. Argialas, Ch.A. Harlow: Computational Image Interpretation Models: An Overview and a Perspective. Photogrammetric Engineering and Remote Sensing 56, Nr. 6, June 1990

[Brüggemann90] H. Brüggemann: Das ATKIS-Datenmodell. Nachrichten aus dem Karten- und Vermessungswesen, Reihe I, Heft 105, 1990

[Busch90] A. Busch, K.-R. Koch: Reconstruction of Digital Images Using Bayesian Estimates. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 6, 1990

[Bähr84] H.-P. Bähr: Abschätzung einiger geometrischer Fehlerkomponenten bei der multispektralen Klassifizierung. Bildmessung und Luftbildwesen 52, Heft 1, 1984

[Bähr91a] H.-P. Bähr, Th. Vögtle (Hrsg.): Digitale Bildverarbeitung - Anwendung in Photogrammetrie, Kartographie und Fernerkundung. Wichmann, Karlsruhe, 1991

[Bähr91c] H.-P. Bähr, F.-J. Behr, H. Gossmann, H. Saurer: Aufbau eines Geographischen Informationssystems zur Ermittlung von Waldschäden und ihrer Veränderung. Forschungsbericht KfK-PEF 83, Kernforschungszentrum Karlsruhe, 1991

[Baumgart90] J. Baumgart: Digitale Erfassung von Veränderungen der Landnutzung unter Verwendung von gebietsbezogenen Gelände- und Fernerkundungsdaten. DFG-Forschungsbericht, Institut für Photogrammetrie und Fernerkundung, Universität Karlsruhe, 1990

[Baumgart91] J. Baumgart: Satellitenfernerkundungsdaten und ihre Integration in ein Geo-Informationssystem zur Nutzung in der Raum- und Landschaftsplanung, Dissertation, Institut für Photogrammetrie und Fernerkundung, Universität Karlsruhe, 1991

[Förstner91] W. Förstner: Object Extraction from Digital Images. Schriftenreihe des Instituts für Photogrammetrie, Universität Stuttgart, Heft 15, 1991

[Lichtner76] W. Lichtner: Ein Ansatz zur Durchführung der Verdrängung bei der EDV-unterstützten Generalisierung in topographischen Karten. Wiss. ArbUH, Nr. 66, Hannover 1976

[McKeown92] St. J. Ford, D. M. McKeown: Information Fusion of Multispectral Imagery for Cartographic Feature Extraction. ISPRS Congress, Washington, 1992

[Palub92] G. Palubinskas: Development of Advanced Information Extraction Techniques for Analysis of Remote Sensing Images, ISPRS Congress, Washington, 1992

[Pfeiffer91] B. Pfeiffer: Satellitenbildverarbeitung beim Aufbau Digitaler Landschaftsmodelle. Zeitschrift für Photogrammetrie und Fernerkundung 5/1991

[Schoedl91] S. Schoedl: Untersuchungen zur Verbesserung von Klassifizierungergebnissen multispektraler Fernerkundungsdaten mit Hilfe topographischer Zusatzinformationen in einem Geo-Informationssystem, Diplomarbeit am Institut für Photogrammetrie und Fernerkundung, Universität Karlsruhe, 1991

[Segl88] K. Segl: Neuentwicklung eines Programmpakets zur überwachten Klassifizierung für mehrkanalige Bilddaten, Diplomarbeit am Institut für Photogrammetrie und Fernerkundung, Universität Karlsruhe, 1988

[Vosselmann92] G. Vosselmann, N. Haala: Erkennung topographischer Paßpunkte durch relationale Zuordnung, Zeitschrift für Photogrammetrie und Fernerkundung, 1992, Heft 6

[Weber83] W. Weber: Konzept zur automationsgestützten Fortführung der Topographischen Übersichtskarte 1 : 200 000 im Rastermodus. Nachrichten aus dem Karten- und Vermessungswesen, Reihe I, Heft 91, S. 55-75, Frankfurt a. M., 1983