Applied vegetation mapping of large-scale areas based on high resolution aerial photographs – a combined method of remote sensing, GIS and near comprehensive field verification.

Jörg Petersen, Otto Dassau, Hans-Peter Dauck and Nicole Janinhoff, nature-consult, Hildesheim, Germany, http://www.nature-consult.de, office@nature-consult.de

Abstract

Nature-consult has developed and established a method that combines classical vegetation mapping with modern methods of multi-spectral image interpretation. The technique combines specific knowledge of the disciplines vegetation ecology, GIS and remote sensing to produce maps of a very high quality, geometric accuracy and coincident efficiency.

This methodology is described and important notes for implementation in similar vegetation mapping projects are given. Additionally the method is documented as the basis of four reference projects carried out for the Wadden Sea National Park Administrations of Lower Saxony and Schleswig-Holstein, the German Federal Institution for Hydrology (BfG) and the Federal Administration of Waterways and Shipping (WSA-Bremerhaven). Within these projects, different high resolution camera systems - matrix cameras (UltraCam-D, DMC) and line scanner (HRSC-AX) - were used.

1. Introduction

Today, as in the past, mapping in the field is the most important and most exact way to document current vegetation conditions. These data provide insight into existing conditions and arising changes. For example, they provide an important resource needed to fulfil nature conservation, coastal conservation and general planning requirements. In this context, typical tasks and reporting obligations include Natura 2000, the EU Water Framework Directive, the Trilateral Monitoring and Assessment Program (TMAP), Major Nature Reserves, and Environmental Impact Assessment (EIA).

For vegetation mapping, the evaluation and analysis of digital (multi-spectral) imagery plays an important role. This is particularly true if the area under investigation and to be mapped is very large, resulting in time and budgetary constraints. Recent methods concentrate either on classical vegetation mapping in the field, where aerial images are used exclusively as background maps, or on pure digital image analysis by means of vegetation classification with little verification in the field. The nature-consult approach combines the advantages of classical vegetation mapping, modern remote sensing and Geographical Information Systems (GIS) analysis. This 'direct' co-operation of vegetation, GIS, and remote sensing experts is a requirement for the successful implementation of this method.

2. Applied method (6 working steps)

In the course of several projects this future-oriented approach implementing remote sensing to support vegetation mapping was designed, used and developed further (Figure 1).

Data pre-processing (2.1)

Prior to each project, selection of the vegetation units which have to be surveyed takes place. For priority areas, a finely differentiated classification is used. For thematic boundary regions, a somewhat rougher subdivision of vegetation types turned out to be optimal (BfN, 2001; Drachenfels. 2004). Generally speaking, 'less is more' in terms of a vegetation classification.

If, at the beginning of a project, the aerial flight has not yet been undertaken, an additional consultation takes place. During this process a data requirements analysis is carried out to determine which data (aerial photographs, elevation data) are useful and necessary for the survey. At the

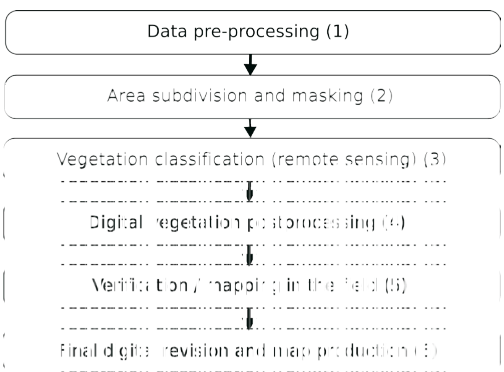

Figure 1: General overview of the applied method (working steps).

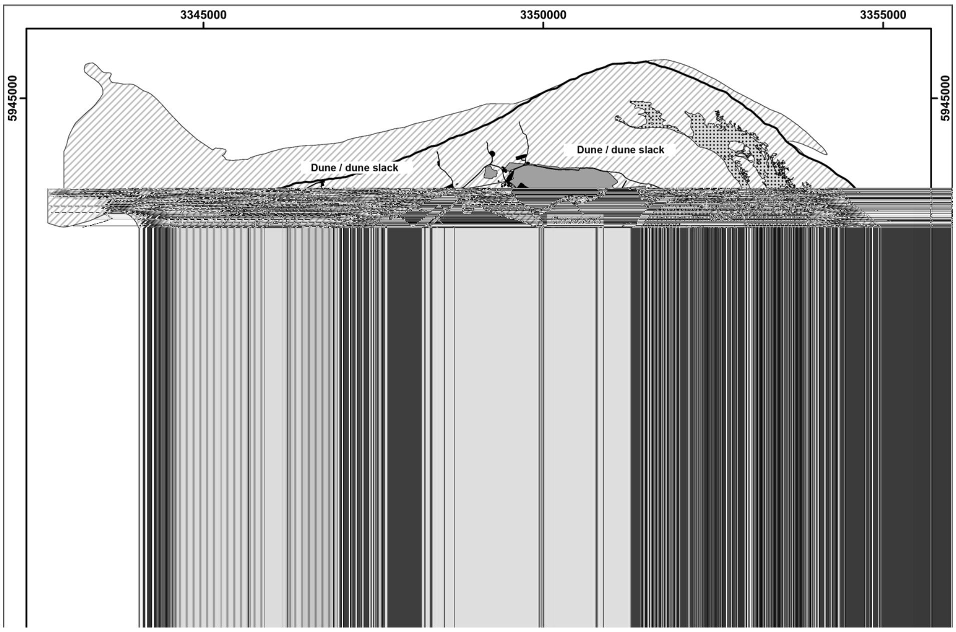

Figure 2: Categorization and masking example for the East Frisian Island Borkum (Advantage: Separate classification of dune/dune valleys, salt meadows and grassland; for a project summary see Table 1).

same time, decisions are made as to how the data processing should be carried out in order to obtain optimal analysis results for the classification. The differentiation of vegetation units places special requirements on the aerial photographs to be used for the project.

For pre-processing of aerial imagery, image enhancements are used to reduce radiometric and geometric distortions which occur during data acquisition. Typical analyses are contrast enhancement, digital filtering, fourier transforms and principal component analysis (Schowengerdt, 2007).

The collection and integration of additional vegetation related data plays an increasingly important role for vegetation classification of multi-spectral aerial photographs. These 'artificial' channels can provide additional information about elevation, geology, texture, slope, or soil characteristics for the area under investigation (Albertz, 2007).

Area subdivision and masking (2.2)

Before classification, areas which do not have to be analysed are masked to define and minimize the area size and the amount of information to be processed. The goal is to reduce the number of vegetation units in order to optimize the classification result. If possible, mask elements such as roads or settlements are taken from official data (ALK, ATKIS or DBWK2). This allows the customer later to exactly overlay the project data with topographic and other maps for visualization and further analyses.

Carrying out an unsupervised classification of the aerial imagery data first is often helpful. For example, it can be used to identify and integrate water surfaces into the mask. However, in practice it is often necessary to digitize mask elements manually during mask creation.

Classification and filtering (2.3)

Training areas are the basis of the applied classification method. These assign specific spectral characteristics to the different vegetation units and are used for their separation into different classes. During the definition of vegetation units, it is important to make sure that the spectral characteristics of the training data are differentiated as clearly as possible (Jähne, 2005). For interpretation purposes, frequency distributions of the training data can be derived and visualized as one or multi-dimensional spectral plots (Lillesand *et al.*, 2003, Figure 3).

Training areas are defined over the entire area under investigation using high resolution GPS equipment (*e.g.* Trimble: GeoExplorer with GeoBeacon). To provide an exhaustive set of training data, a subsequent definition of training areas is carried out in a GIS application. This is done by the vegetation experts who performed the collection of training data in the field. According to comparative investigations, this two step process is necessary to achieve high quality classification results and to process independent classifications with distinct training data for smaller sub-regions (McCauley and Engel 1995).

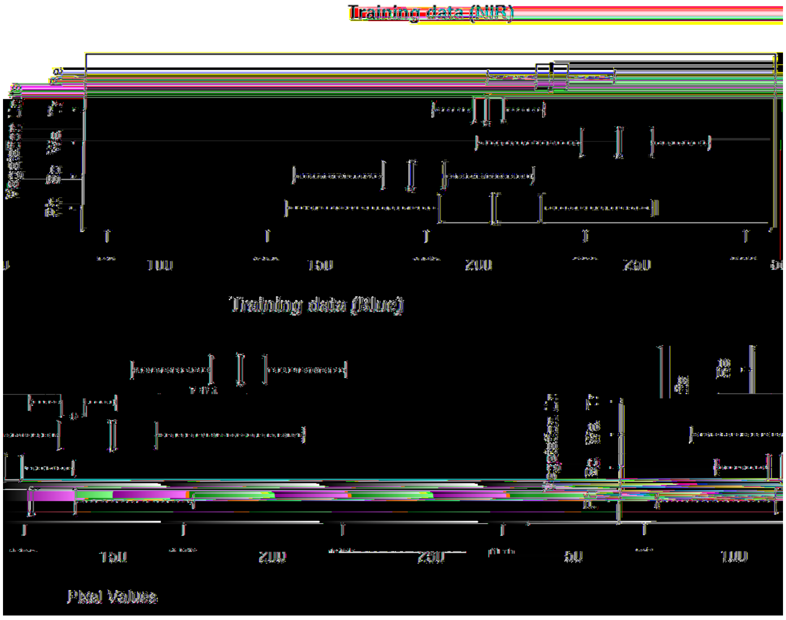

Figure 3:
Spectral plots for training data obtained in 2 bands (NIR, B) for 5 vegetation units.

The supervised combined geometric and radiometric classification is carried out using the open source software GRASS (http://grass.osgeo.org). At the beginning of the classification, a cluster analysis of the image data takes place based on the training areas. In practice the spectral and 'artificial' data pre-processed for the classification (*e.g.* G, R, NIR, NDVI, elevation and texture information) are combined in a map group. On the basis of the spectral data characteristics within the training areas, several statistic analyses take place. The results are stored as training signatures in the form of Gaussian mixture distribution and are later used to classify the vegetation units. For parameter estimation, the EM-algorithm is used (Dempster *et al.*, 1977). The number of normal distributions is either statistically specified during the evaluation of the training data using the Rissanen's Minimum Description Length (MDL) criterion (Rissanen, 1983), or can be specified manually in the GIS. Weighting of the normal distributions is carried out next, using *a posteriori* probabilities of the subclasses for creation of the training signatures. For the classification, the 'Sequential Maximum A Posteriori' (SMAP) segmentation algorithm combined with the 'Multiscale Random Field' (MSRF) model is used (Bouman and Shapiro, 1996). The group of layers, previously defined for the cluster analysis, is analyzed using the provided training signatures and assigned to the different vegetation classes (Neteler and Mitasova, 2008). Contrary to pixel-based procedures, a scale-dependant approach is used which uses radiometric and geometric information in the form of neighbourhood relations between pixel values during the class assignment (Ripley, 1996). The MSRF works as an image pyramid with different layers segmented at different image resolutions. Starting from a coarse scaled layer, the probability of affiliation to a class is determined for each pixel within a pre-defined window and passed on to the next, finer scaled, layer (Bouman and Shapiro, 1996).

After the classification, several filter processes take place. A multilevel, raster-based neighbourhood filtering of the results is carried out next. This process minimizes small areas which occur in each classification due to the mixing of pixel and spectral heterogeneity, without manipulating the classification result.

Following this first filter step, the data are converted into a vector shapefile and further processed for the production of the field maps. This entails:

- Reducing the data volume by decreasing vector vertices, without degrading the geometry of the classification substantially.
- Additional filtering of the vegetation geometries on the basis of a project-specific minimum area size. Generally a minimum area of 200 to 500 m^2 proved to be useful, because smaller areas cannot be represented at 'usual and desired' mapping scales (1:3 000 to 1:5 000). Furthermore, it must be possible to almost exhaustively verify the classification within a justifiable period

of time in the field. For selected vegetation units (*e.g.* types with high protection status, important for nature conservation) the definitions deviate from the general minimum area size specified in the project.

Digital vegetation post-processing (2.4)

After classification and filtering, a complete digital revision of the vector data is essential. This is done by the vegetation experts who were already responsible for working on the training areas in the field.

The steps that follow are: the definition of an optimal mapping scale; the preparation of the field map grids; and the printing of the field maps for verification / mapping.

Verification/Mapping in the field (2.5)

After the classification is filtered and digitally revised, an exhaustive survey of the area under investigation takes place. Under optimal conditions this leads to verification. The focus lies on areas which proved difficult or uncertain during the digital revision. Thus, the effort in the field is strongly reduced and is efficient when compared to conventional mapping methods.

Final digital review and map production (2.6)

After verification/mapping in the field, the results are revised again and finally integrated into the digital field maps (GIS project). The better the quality of the preceding working steps, the smaller the effort required for the final creation of the project maps.

3. Project examples

The method introduced here has been successfully used since 2003. This is documented in four reference projects, which were undertaken on behalf of the Wadden Sea National Park Administrations of Lower Saxony and Schleswig-Holstein as well as the German Federal Institute for Hydrology and Federal Administration of Waterways and Shipping (Table 1).

For the entire terrestrial area of the Wadden Sea National Park of Lower Saxony, a vegetation map of habitat types, supplemented by TMAP units (Trilateral Monitoring and Assessment Program) and biotope types, was established using HRSC-AX and DMC aerial photographs. This project was carried out in order to report on the implementation of the EC Habitats Directive and on the basis of the results of the TMAP working groups (Bakker

Table 1:
Key data of presented reference projects.

TMAP types, Natura. 2000 habitat types	TMAP types, plant communities	Plant communities, TMAP and biotop types	Biotop and TMAP types, plant communities	Vegetation typolo
HRSC-AX (GSD 32 cm), DMC (24)	UltraCam-D (GSD 10 cm)	DMC (GSD 25 cm)	DMC (GSD 25 cm)	Camera system
B, G, R, NIR, PAN (+ Texture)	G, R, NIR (+ Texture)	B, G, R, NIR (+ Texture, laser scan data)	B, G, R, NIR, PAN (+ Texture, laser scan data)	Spectral channels
1:3,000	1:4,400	1:4,300	1:4,300	Mapping scale
ca. 1,300	ca. 650	ca. 340	ca. 329	Field maps (DIN A
12 weeks (3-4 people)	11 weeks (3 people)	2 weeks (3 people)	6 weeks (1 person)	Field work duratio
8 people	8 people	6 people	4 people	Team
Natura 2000 conservation status	Agricultural exploitation	Photo documentation	Interpretation	Additional service

et al,. 2005; Petersen and Lammerts, 2005). The entire area of the National Park (about 33 000 hectares) was mapped using image classification by four vegetation experts. A maximum map scale of 1:3 000 was used (Petersen and Pott, 2005; nature-consult 2006).

Following this project, the entire terrestrial area of the Wadden Sea National Park Schleswig-Holstein was mapped using aerial imagery taken with an UltraCam-D camera. In only 11 weeks of field work, about 15,000 ha of salt marshes, dunes, dune slacks, grassland vegetation, as well as land use and scarps, were mapped (nature-consult ,2008).

On behalf of the German Federal Institute for Hydrology (BfG), reed and adjacent vegetation on both banks of the River Elbe from Otterndorf to Hamburg was mapped. The evaluation used digital DMC aerial imagery and laser scan data together with exhaustive field work. With the described procedure, about 10,000 ha were classified, digitally revised, mapped and verified in the field over a period of only two and a half months (nature-consult, 2007). A similar project (approximately 8,000 ha) was conducted for the Weser River (also the Lesum, Hamme and Wümme rivers) from the Wadden Sea to Bremen for the Federal Administration of Waterways and Shipping (WSA-Bremerhaven, Table 1).

4. Conclusion

The method developed by nature-consult has proven to be successful for large-scale vegetation mapping. With this approach it is possible to efficiently provide high quality, optimized field maps. For example the East Frisian Island Borkum, with a total area of about 3,830 ha, was completely mapped and verified by four field workers in only four days. For this complete survey, 126 analogue field maps (DIN A4) at a scale of 1:3 000 were produced.

As a result, this approach enables the provision of vegetation maps with a very high accurateness of geometrical location and vegetation assignment. Although remote sensing cannot replace field surveys, it can make vegetation mapping more efficient and thus more economical.

References

Albertz, J., 2007. Einführung in die Fernerkundung. Grundlagen der Interpretation von Luft- und Satellitenbilddaten. Wissenschaftliche Buchgesellschaft, Darmstadt.

Bakker, J.P., Bunje, J., Dijkema, K., Frikke, J., Hecker, N., Kers, B., Körber, P., Kohlus, J., Stock, M., 2005. Salt Marshes - Chapter 7, Wadden Sea Quality Status Report No 19. Common Wadden Sea Secretariat, Wilhelmshaven.

BfN, 2000 (Hrsg). Verzeichnis und Rote Liste der Pflanzengesellschaften Deutschlands (Bearb. E. Rennwald). Schriftenreihe für Vegetationskunde: 35, 800 S., Bonn.

Bouman, C. and Shapiro, S., 1996. A Multiscale Random Field Model for Bayesian Image Segmentation. In: IEEE Trans. On Image Processing, 3(2):162-177.

Dempster, A. P.; Laird N. M. and Rubin, D. B., 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. In: Journal of the Royal Statistical Soc./ B, 39 (1), 1- 38.

Drachenfels, O. v., 2004. Kartierschlüssel für Biotoptypen in Niedersachsen. Naturschutz und Landschaftspflege in Niedersachsen A/4, Hildesheim.

Jähne, B., 2005. Digitale Bildverarbeitung. 6. Auflage, Springer-Verlag, Heidelberg.

Lillesand, T.; Kiefer, R. and Chipman, J., 2003. Remote Sensing and Image Interpretation. 5th Edition, Wiley, New York.

McCauley, J. and Engel, B. 1995. Comparison of Scene Segmentation: SMAP, ECHO and Maximum Likelihood. IEEE Trans. On Geoscience. & Remote Sensing 33 (6): 1313-1316.

nature-consult, 2006. Kartierung der gesamten terrestrischen Bereiche des Nationalparks Niedersächsisches Wattenmeer im Jahr 2004. Im Auftrag der Nationalparkverwaltung Niedersächsisches Wattenmeer. Unveröffentlichter Abschlussbericht, 125 S.

nature-consult, 2007. Erfassung der Röhrichtbestände und angrenzender ufernaher Vegetationseinheiten der Unter-/Außenelbe aus den Daten der Luftbildbefliegung 2006. Im Auftrag der Bundesanstalt für Gewässerkunde. Unveröffentlichter Abschlussbericht, 58 S.

nature-consult, 2008. Vegetationskartierung der Salzwiesen an der Westküste von Schleswig-Holstein 2006/2007. Im Auftrag der LKN-SH Nationalparkverwaltung. Unveröffentlichter Abschlussbericht, 57 S.

Neteler, M. and Mitasova, H., 2008. Open Source GIS: A GRASS GIS Approach. 3rd Edition, Springer, New York.

Petersen, J. and Lammerts, E.-J., 2005. Dunes. Chapter 9.2. In Essink *et al.*, (eds) Wadden Sea Quality Status Report 2004. Wadden Sea Ecosystem No. 19, Common Wadden Sea Secretariat, Wilhelmshaven.

Petersen, J. and Pott, R., 2005. Ostfriesische Inseln. Landschaft und Vegetation im Wandel. 160 S. u. Kartenband (12 Vegetationskarten auf DIN A1), Hrsg: Nds. Heimatbund.

Ripley, B.D., 1996. Pattern recognition and neural networks. Cambridge.

Rissanen, J., 1983. A Universal Prior for Integers and Estimation by Minimum Description Length. In: Annals of Statisitics, 11 (2), 417- 431.

Schowengerdt, R., 2007. Remote sensing: Models and methods for image processing. 3rd ed., Tucson, Arizona.