Texture Classification applied on aerial imagery in Forestry

Dietmar P.F. Moeller, Christian Koerber
Department Computer Science, AB TIS
University of Hamburg
Vogt-Koelln-Straße 30, 22527 Hamburg, Germany

Christoph Kaetsch
Department of Forest and Wood Science
University of Stellenbosch
Privaatsak X1, 7602 Matieland, South Africa

E-Mail: dietmar.moeller@informatik.uni-hamburg.de

KEYWORDS
multi-channel Image Analysis, Digital Signal Processing, Texture Classification, Object Extraction, Clustering

ABSTRACT

Feature extraction from aerial images is an important research topic with a wide area of applications, like traffic or agriculture monitoring, natural disaster early warning system, etc. Because, to deal successfully with, information gained by remote sensing is by several factors much more cost effective compared to manually accomplished measurements. Certain Information is actually only available by remote sensing since, as the area under investigation is as huge, that any other means would be infeasible. The extraction of certain objects from aerial images has proven to be a very difficult problem especially if the investigated objects do not have sharply bounded lines. This is common the regular case when dense forests are concerned during image analysis. Here, trees usually occlude each other and are hard to differentiate from epiginous vegetation, that make geometrical approaches of object identification, as well as direct representations, hard to apply. On the other hand, counting's performed in the frequency domain offer the advantage of transformation invariance and suffer lesser from diffuse object boundaries. The determination of a clear signature is difficult, if the objects in question are quite similar, though. Hence the paper suggests an approach derived from texture classification that achieves better results in the forestry.

INTRODUCTION

Monitoring of forests for several purposes has proven itself as being a demanding and costly task. Gathering data on this diverse natural resource providing information for sustainable forest management means identifying single trees as well as their most important properties like species, age, health-status, height, crown and trunk diameter. In a normal forest inventory covering 8000 ha to 10.000 ha the number of trees to be measured easily exceeds 10.000 or even 12.000. To minimize costs, current forest inventory systems are generally based on ground measurements taken on sample plots distributed regularly over the whole forest. In addition, terrestrial measurements are supported by using aerial photographs. The synoptic overview provided by the images can help to stratify the forest and allows to optimize the terrestrial work. In fact, remote sensing is used for nearly 100 years (Hildebrandt 1996) in the forestry domain resulting in a widely used set of tools for analogue image analysis in forest inventory.
Despite all efforts towards optimized forest inventory concepts measurements in forests remain costly. Expensive equipment required for image interpretation and time consuming manual Photogrammetry have not met the high expectations.

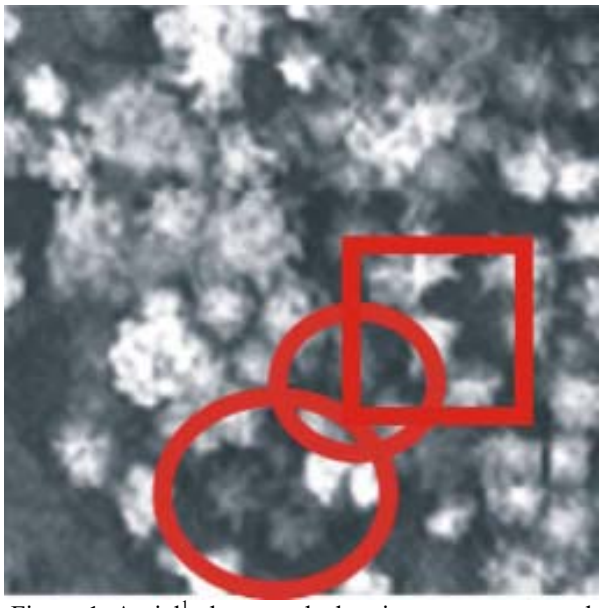

Figure 1. Aerial[1] photograph showing an uneven-aged mixed forest stand formed by Norwegian Spruce, White Oak and European Beech. Spruces with low intensity are marked by circles, the square corresponds to Fig. 3.

Growing and improved availability of digital images may change this situation dramatically. While digital images are easy to handle and allow sophisticated image processing, reliable methods capable of addressing a

[1] IRC chromatic aerial image courtesy of Institut für Forsteinrichtung und Ertragskunde, University of Göttingen. Scale approximately 1:6000, scanned with a photogrammetric scanner, resolution 12 µm, GSD approximately 8 cm.

Proceedings 19th European Conference on Modelling and Simulation
Yuri Merkuryev, Richard Zobel, Eugène Kerckhoffs © ECMS, 2005
ISBN 1-84233-112-4 (Set) / ISBN 1-84233-113-2 (CD)

broader range of stock enumeration issues may pave the road towards automatic forest inventory systems.

Since the early 1990's several efforts to extract data on trees and forest stands from digital images in a widely automated manner have been published (a.o. Pinz, 1988, 1992; POLLOCK, 1994, 1996, STRAUB & HEIPKE 2001, KÄTSCH 2002). Despite very promising results presented from different authors none of the methods proposed so far can be seen as an universal tool ready to meet all challenges related to different stand structures, tree species or age-classes found in forests around the globe.

This paper presents first results of study carried out in uneven aged Central European forest stands using multi-channel texture classification, which in combination with existing methods may lead to more general approaches of automated stock taking in forestry.

Fig. 1 shows a cutout of a grayscale version of a chromatic aerial image, displaying the Hills area in south western Lower Saxony, Germany. The (1) mixed forest shown is characterized by a tree population of (2) very varying age (40 to 120 years), consisting mainly of oaks, beeches and spruces. The (3) differing height structure of the trees leading to shadowing effects is clearly visible. All three mentioned factors result in an increased complexity regarding automated image analysis. In this paper we restrict ourselves to issues (1) and (2), which are, among others, characterized by the difference between high intensity broad-leaved trees and low intensity conifers, the latter marked by a circle in Fig. 1, a characteristic which classical image processing algorithms tend to miss, as exemplified in Fig. 2, showing the results of an edge detection filter using a Sobel kernel as in (Paulus and Hornegger 2003).

THE MULTI-CHANEL APPROACH

This paper suggests an approach, based on the two pass filter algorithm, introduced by (Zell 1994), which itself was derived from (Hatzigeorgiou 1993). The first pass compromises a windowed multi-band signature analysis, the second pass consists of an artificial neural network classifier. Our contribution introduces a multi-resolution[2] representation of the first pass and replaces the neural network classifier by *k*-means clustering, supersede the automatic algorithm by a semi-automatically one, i.e. one, that needs currently at least some manual interaction, and in such manner gaining increased accuracy for an extremely narrow feature set.

[2] (Zell 1994) refers to the multi-band approach as multi-resolution which is appropriate in the sense that lower frequencies contribute more to the signal than higher frequencies and that thus a low pass filtered image has a different resolution than a high pass filtered one. We refer to the term multi-resolution in the sense of different window sizes.

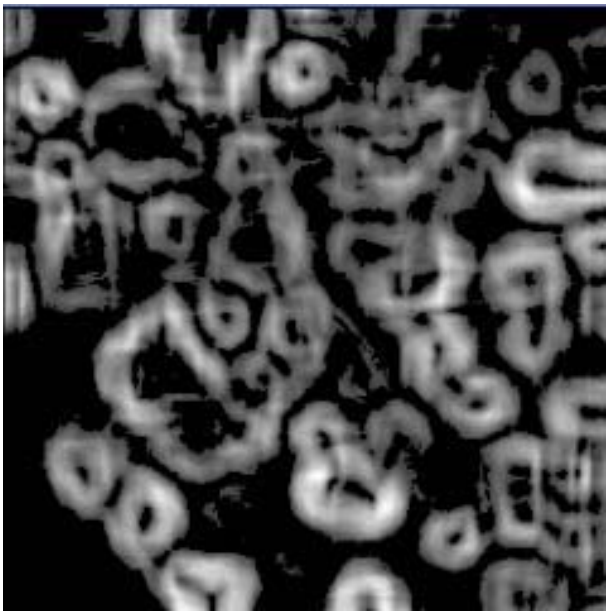

Figure 2. Edge detection applied to Fig. 1 after low pass filtering. The conifers, marked by circles in Fig. 1, tend to vanish

The **first pass** consists of five steps:

1. Window extraction
   A square window of size $N \times N$ is extracted for each pixel of the image, centred at the pixel, where w is an odd number, see Fig. 3 (left). Windows not fitting completely into the image are disregarded so that, if the image has $M \times M$ pixel, only the inner $(M-N+1) \times (M-N+1)$ pixel can be considered. Typical window sizes which had been proven to be the most effective are rather smaller than those reported by (Zell 1996) and range from 7 to 25 pixel. The example shown in Fig. 3 has a size of 64 x 64 pixel only for demonstration purposes.

2. Window smoothing
   The extracted sub image becomes smoothed to zero at the edges during the next step in order to avoid undesired edge effects. Smoothing is performed by multiplying the extraction window pixel wise with a filter that shows good stop-band attenuation if transformed to frequency space, as shown in Fig.3 (middle and right).

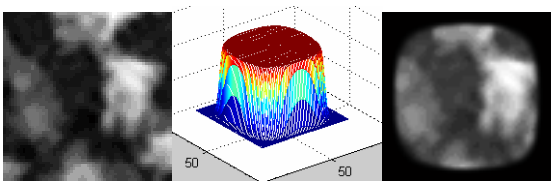

Figure 3. Extraction window corresponding to the square region of Fig. 1, multiplication filter, resulting extraction window (respective extracted sub-image, from left to right)

3. 2D-FFT
   Window smoothing is followed by calculating the magnitude of the two-dimensional discrete Fourier transform[3],

$$X[k_1,k_2]=\frac{1}{N^2}\sum_{n_1}^{N-1}\sum_{n_2}^{N-1}x[n_1,n_2]e^{-i2\pi\left(\frac{k_1n_1+k_2n_2}{N}\right)}$$

   whose locality in pixel space obviously depends on the size of the extraction window. Too large windows result in failing to detect texture boundaries, while smaller windows cannot detect larger patterns. Therefore the window size depends on the type of application and on the scale of the aerial photographs or satellite images. Images, window sizes between 7 and 25 rendered, show the best results. The extracted sub-image had to be zero-padded to achieve at least a magnitude spectrum of *128 x 128* pixel in order to avoid numerical complications in subsequent steps of the algorithm.

4. Multi-band pass filtering
   Step 4 consists of 8 successive band pass filters completely partitioning the spectrum. The radius of the succeeding filter is $\sqrt{2}$ times the radius of its predecessor, which is the reason for the greater size of the spectrum than the original image's size. Hence the first filter, the low pass, must thus be relatively small, and would be smaller than a single pixel if the extracted image wouldn't be zero padded. The adopted filter was a perfect band pass, obtained by setting all values of the spectrum to zero, which don't fall into the circlet of the respective band. Since no inverse transform is intended no disadvantages are expected from avoiding a more sophisticated filter as e.g. the gradients of a Gauss function.

5. Signature generation
   Finally, the signature of the pixel which lies at the centre of the extraction window, is generated by computing the arithmetic mean value per band pass. This results in eight scalar values, as eight band passes are used, representing an eight-dimensional vector, that is referred to as signature of the pixel.

The **second pass** completely deviates from (Zell 1996) and consists of three steps:

1. *k*-means clustering
   k-means clustering is a statistical method of vector quantization where a "codebook" *K* of *k* vectors is build up by the algorithm with the objective to heuristically determine its *k* vectors so, that the average minimum Euclidean distance for the *m* vectors of a set *M* (here the signatures) to one of the *k* codebook vectors becomes minimized (with $k < m$). Applying this kind of clustering results in good outcomes. Fig. 4. gives an indication though it refers already to the improved version of the algorithm, described in the next section. It has to be noticed that the clustering was not able to distinct noise from low intensity conifers remarkably better than other methods. It has also been observed that increasing the number of clusters didn't increase the quality if the distinction, an indication that the band which separates the spectrum of the conifers from noise, must be vary narrow. An improved version of clustering, which lead to significantly better results, is described in the following section.

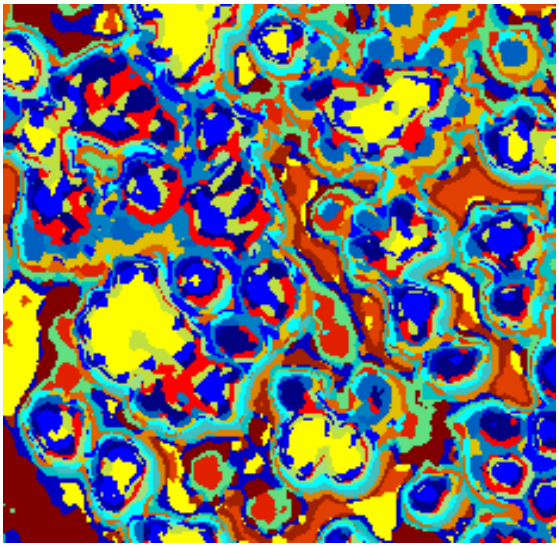

Figure 4. Merging the signatures into *k* cluster obtained from a 16-digit signature

2. Merging clusters to classes
   Once the clusters are obtained, they need to be merged down to a more concise number of 3 or 4 clusters which are referred to as classes, see Fig. 5. This step currently requires manual user interaction and this is the sense where some kind of "intelligence" needs to be applied. Since this is limited to merging about 32 color regions to larger ones, we expect this to be achievable in about 5 to 10 minutes, based on our observations on manually merging clusters within an imaging application and assuming an appropriate user interface is provided, which we consider tolerable. In the conclusion a thorough discussion of this subject will be given.

3. Morphological operations
   Morphological operations are not necessarily an integral part of the algorithm but are required by the application in order to utilize the gained information. These operations include dilation and erosion in order to get a better object separation without losing too much object area, binarization,

[3] Following the notation for the complex DFT of (Smith 1999) and using the FFT for the implementation

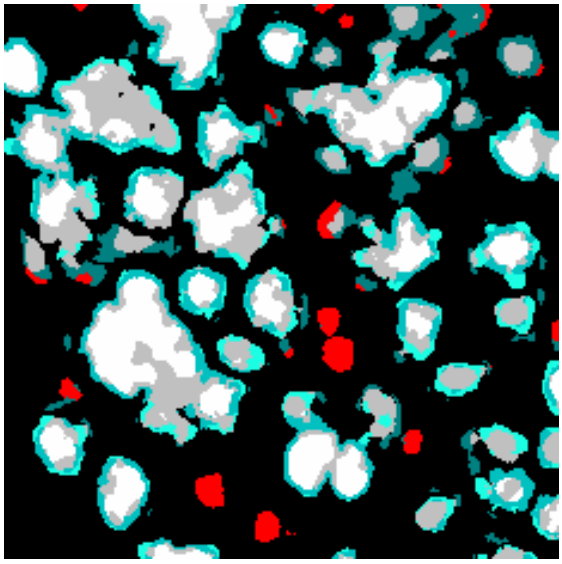

Figure 5. Merging clusters (as in Fig. 6) into classes, red regions refer to regions identified as conifers

required for object identification and ellipse embedding (unaccounted by us yet), used for further formalization and subsequent automated processing. Fig. 6 shows the identified objects after dilation, erosion, binari zation and contouring overlayed on the original image.

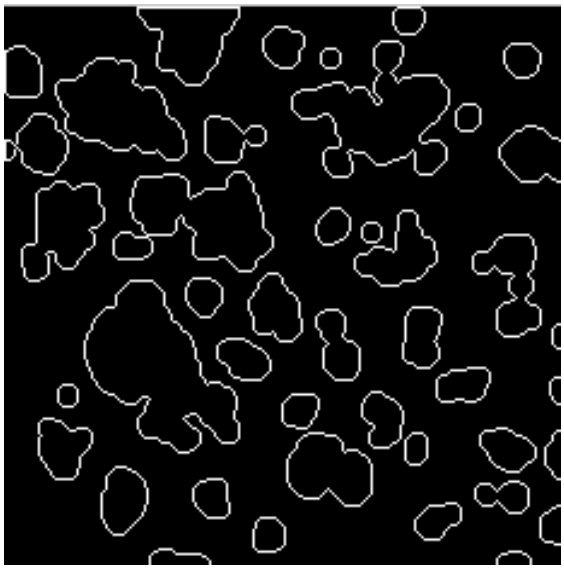

Figure 6. Identified objects after dilation, erosion, binarization and contouring

IMPROVEMENTS

The originally obtained results showed similar disadvantages as filter-based algorithms. The algorithm had difficulties to distinguish low-intensity conifers from noisy, non vegetable parts of the signal as it can be seen in the lower left corner of the investigation area in Fig. 1, that represents a forest track. This could not be remedied by increasing the number of clusters, which had only led to the same finer subdivision in both areas, the one containing the conifers and the one containing the noise.

One obvious solution was to use information gathered by varying image sizes, especially the tendency of low intensity objects to vanish with increasing window size, because noise remains noise. We applied two different algorithms, both utilized a 7 pixel wide and a 25 pixel wide extraction window. The first algorithm combined two clustered images by generating *k1 * k2* (basically by appending *k2 to k1*) new clusters from the *k1* respective *k2* clusters, computed by the method suggested above. This not only generates numerous meaningless clusters respective needlessly separated well bounded clusters and thus hampered the successive classification step, it also introduced unwanted artefacts, though it indeed result in a better object distinction. Much better results are gained by appending one signature to the other, thus gaining a 16-digit[4] instead of a 8-digit signature vector. and performing the k-means clustering in a 16-dimensional instead of a 8-dimensional vector space, see Figs. 4 to 6.

CONCLUSION AND PERSPECTIVE

- Our work so far emphasized the detection of low intensity objects (conifers) which obviously had the greatest prospect of success for our application. More work is required with respect to object separation, see Fig. 6. We expect to gain better results in this area either with improved morphological operations (i.e. testing with successive erosion operations whether objects are separable) or with geometrical operations (ellipse embedding.

- We anticipate further improvements through utilizing other spectra, especially the infrared channel. Hyper-spectral Images which had been introduced to remote sensing in the context of forestry may give this approach a hole new direction as the up to 220 channels could be utilized for direct clustering.

- The elimination of the semiautomatic classification step could probably be reached by e.g. artificial neural networks (perceptron, similar to character recognition) But since the requirements for manual interaction are already condensed to a very small effort, which can easily be further streamlined by providing a simple interface, we would rather improve object separation which will improve the overall usefulness of the algorithm to a much higher degree.

- Some investigations may be appropriate whether at least partial results of this algorithm could be ported to operations in the spatial domain. The solution to work with different window sizes for the 2D-FFT corresponds to working with different

[4] with each digit being a floating point number

kernel sizes for convolution (low pass filtering, edge detection).

REFERENCES

Hatzigeorgiou, A. "Texturerkennung mit neuronalen Netzen" In Zell, A (Ed.): *Workshop SNNS-93: Simulation neuronaler Netze mit SNNS, Universität Stuttgart, Fakultät Informatik, Bericht Nr. 10/93*. pp. 50-58, 1993

Hildebrandt,G. "Fernerkundung und Luftbildmessung für Forstwirtschaft, Landschaftsökologie und Vegetationskartierung. Wichmann Verlag. 1996

Kätsch, Chr. Towards an automatic forest inventory system based on remotely sensed data and digital image processing. Proceedings of the 29th ISRE-Conference, Argentina 2002

Paulus, D.W.R., Hornegger, J. *"Applied Pattern Recognition"*, Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiebaden, pp. 146-156, 4th Ed. 2003

Zell, A. *"Simulation Neuronaler Netze"*. Addison-Wesley, Bonn;Paris; Reading, Mass., pp. 519-531, 1994

Pollock, Richard J., The Automatic recognition of Individual trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model, Vancouver, Canada, June 1996: The University of British Columbia, 1996.

Pinz, A.: Ein Bildverstehendes Expertensystem zur Erkennung von Bäumen auf Farb-Infrarot-Luftbildern. PhD thesis, Technische Universität Wien, 1988

Smith, S.W., *"The Scientist and Engineers's Guide to Digital Signal Processing"*, California Technical Publishing, San Diego, California, p. 570, 2nd. Ed. 1999

Straub, Bernd-M. and Heipke, , "Automatic Extraction of Trees for 3D-City Models from Images and Height Data," (Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), Band 3 hg. von A. G. L. v. G. E. Baltsavias) A.A.Balkema Publishers: Lisse/Abingdon/Exton(PA)/Tokio, 2001, 267-277.

AUTHORS

DIETMAR P. F. MÖLLER was born in Preetz, Germany. He enrolled at the Universities of Lübeck, Bremen Mainz and Bonn, where he studied Electrical Engineering and Human Medicine. He obtained his doctoral degree in 1980. From 1985 to 1991 Dr. Möller lead the anaesthesia division of Dräger AG in Lübeck, Germany. 1991 he has been elected as Full Professor for Computer Engineering (TI) at the Technical University of Clausthal. In 1998 he has been elected as Full Professor for Computer Engineering Systems (AB TIS) at the University of Hamburg where he holds the chair for Computer Engineering. Since 1998 he also is Adjunct Professor at the California State University Chico. His E-Mail address is dmoeller@informatik.uni-hamburg.de and his Web-page can be found at http://www.informatik.uni-hamburg.de/TIS/

CHRISTIAN KÖRBER was born in Hamburg, Germany. He enrolled at the University of Hamburg, where he studied Computer Science and obtained his diploma degree in 2002. He joined the Computer Engineering Department (AB TIS) as a research assistant where he lead research work on Geo Information Systems, http://www.verve-gis.com with specific emphasis on computer graphics and geometric modeling. E-mail address/web page is koerber@informatik.uni-hamburg.de; `http://www.informatik.uni-hamburg.de/TIS/`

CHRISTOPH KÄTSCH was born in Göttingen, Germany. He studied Mechanical Engineering at the Technical University of Clausthal and Forestry Sciences at the University of Göttingen. He finished his Phd in 1990 and received the venia legendi (Dr. habil.) in 1997. From 1982 to 1987 he was a Forest Officer at the State Forest Service of Lower Saxony where he became head of the department of Informatics and Accounting in 1988. 1992 he was appointed as Professor for Geoinformatics & Remote Sensing at the University for applied Science & Art in Hildesheim. Since 2004 he is holding a chair on Geoinformatics & Remote Sensing at the Faculty of Agriculture and Forestry Sciences, Stellenbosch University, South Africa.
Email: christop@sun.ac.za,
Web: http://www.academic.sun.ac.za/forestry