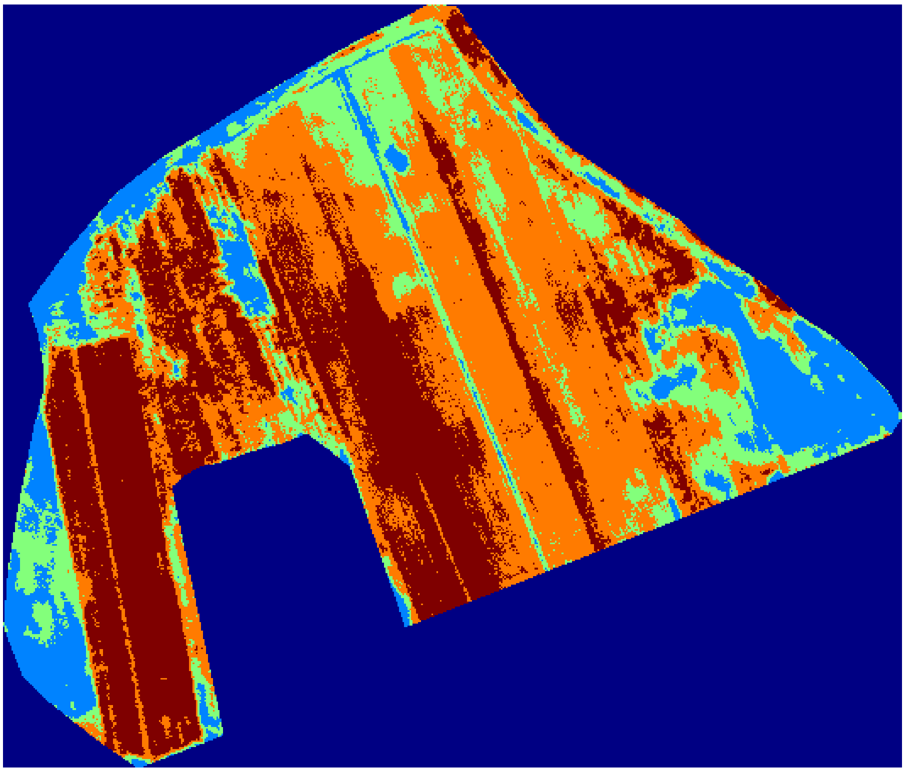
Development of methods for satellite monitoring of cultural heritage sites: A preliminary study

NR Norwegian Computing Center
APPLIED RESEARCH AND DEVELOPMENT

NOTE



SAMBA/25/03

Lars Aurdal

December 3, 2003

Norwegian Computing Center
APPLIED RESEARCH AND DEVELOPMENT

NR Note

Title: Development of methods for satellite monitoring of cultural heritage sites: A preliminary study

Date: December 3, 2003
Year: 2003
Note no.: SAMBA/25/03

Author: Lars Aurdal <lars.aurdal@nr.no>

Abstract: The increasingly intensive use and modification of the landscape as a result of modern demands for efficient infrastructure and land use (agricultural production, mining, energy sources, leisure/tourism facilities) exerts growing pressure on areas and sites associated with our cultural heritage. The use of modern support technologies is imperative if such rapid changes are to be balanced against the sustainable management of this resource.

This pilot project directly addresses these issues by initiating the development of a basis for a sustainable, up-to-date and cost-efficient decision-support methodology which relies upon satellite remote-sensing, mapping and monitoring of cultural heritage sites.

This report concerns the automatic processing of satellite images in order to detect candidate cultural heritage sites. Using publicly available land use maps, the treatment can be confined to agricultural land. We then describe how images of these regions can be preprocessed in order to remove confusing ground structures. Segmentation (clustering) techniques are then used to group pixels into possible candidate sites, a supervised classification step is finally used to separate between interesting and uninteresting sites.

Our principal conclusion is that automated processing of satellite images for detection of candidate sites is feasible, although with clear limitations. Based on our findings, it seems reasonable to investigate the use of satellite data for this purpose further.

Keywords: cultural heritage satellite monitoring

Target group: Public

Availability: Open

Project: Kulturminner

Project no.: 220060

Research field: Remote sensing

No. of pages: 24

Norwegian Computing Center
Gaustadalléen 23, P.O. Box 114 Blindern, NO-0314 Oslo, Norway
Telephone: +47 2285 2500, telefax: +47 2269 7660, http://www.nr.no

Contents

1 Introduction

It is generally recognized that the increasingly intensive use and modification of the landscape resulting from modern demands for efficient infrastructure and land use (agricultural production, mining, energy sources, leisure/tourism facilities) exerts growing pressure on cultural heritage in the landscape.

In order to match the political intentions of updated and sustainable cultural heritage management, a necessary first step is to create a representative picture of the resource that has to be managed. In Norway, where extensive white areas are still to be found on cultural heritage maps; where the registered cultural heritage sites display an unrepresentative concentration in areas with high human activity; and where the registered positions of the sites can easily be 30-40 meters from their true location, it is obvious that something has to be done in order to achieve even this basic goal.

In recognition that a) it will never be realistic to obtain funding for thorough survey and monitoring of the enormous tracts in question using traditional field-survey methods, and b) there is a demand for access to representative and comprehensive cultural heritage data to create a basis for the development of a flexible and up-to-date cultural heritage management, the Norwegian Space Centre (NRS) and the Norwegian Directorate for Cultural Heritage (RA) decided to support a pilot project designed to prepare the ground for the development of survey and monitoring methodology involving multispectral satellite data.

The project's aim is to develop a cost-effective method for surveying and monitoring cultural heritage sites. The costs of systematically surveying areas of the scale involved here by means of conventional fieldwork provides the incentive for the development of alternatives. Depending on which field methods are employed, and the type of landscape surveyed, costs for conventional fieldwork will normally be around 250,000 Norwegian Crowns (NOK) per square kilometer. In comparison, high-resolution satellite data cost less than NOK 1,000 per square kilometer, a fraction of conventional fieldwork costs.

Even though the costs connected with the processing of the satellite data will not be insignificant, and fieldwork can probably never be entirely replaced by high-technological methods, it seems plausible that an essentially cheaper, and possibly even qualitatively better, method for the surveying and monitoring of cultural heritage sites can be developed by using multispectral satellite data to target the fieldwork to a degree not possible today.

It must be anticipated that certain types of cultural heritage sites will be discriminated against by a method based on satellite data, although probably not to the extent to which current systematic fieldwork seems to overlook some categories, such as settlements, for instance. For example, in large areas, burial mounds comprise the predominant registered type of monument to such a degree that one wonders where the people buried in them lived. A systematic and standardized method for analysis of the satellite images should make it possible to make statistical corrections for such bias in the data.

Because the multispectral satellite images are normally delivered geo-corrected and related to a positioning system, the data contains a built-in registration of positions that is superior to that which can be obtained with a GPS in the field.

The ability of high-resolution multispectral satellite images (previously the term high-resolution was applied to the Landsat images; it now seems logical to restrict it to IKONOS and Quickbird images) to distinguish minor variations in the vegetation and in the soil is superior to what is possible using aerial photos. Furthermore, the development of multispectral techniques in the direction of more bands and finer resolution suggests that the potential of these techniques will increase considerably in coming years.

The most exciting aspect of multispectral techniques is to see how far they can be developed in the direction of distinguishing cultural heritage sites that have no visible physical manifestations preserved above ground level. This is the central critical issue for cultural heritage management at present, as demonstrated in the case of cultivated land by this report.

Accessibility is essential for the development and application of sustainable, up-to-date, and cost-effective decision-support methodologies based on high-resolution satellite data. Further development of the recently formed Satellite Data Archive within the framework of the Norwegian Mapping Authority is an important means of making the relevant types of satellite data accessible for the various sectors which can benefit from them (Grön and Loska, 2002).

1.1 Goals

The aims of the overall project are (Grön and Loska, 2002):

- To clarify what the current state-of-the-art is in terms of the location and monitoring of cultural heritage by satellite data.
- To evaluate the practical potential of multispectral satellite data with different resolutions.
- To look for relationships between anomalies visible in multispectral satellite data and ground features that can be distinguished by soil chemistry, ground spectrometry, and vegetation analysis.
- To suggest a strategy for further national initiatives in this field.

A central part of this project is the use of computers for processing satellite data from the study areas in order to find sites that could possibly be cultural heritage sites. Although the satellite data are readily available at a modest price, the automatic extraction of interesting sites from these massive data sources is by no means trivial. As part of the effort undertaken in this project, the Norwegian Computing Centre (NR) was given the task of suggesting and evaluating interesting methods and algorithms to be used in this processing. This document reports the findings and conclusion of the work performed by NR.

1.2 Organization and funding

The pilot project is funded by The Norwegian Space Centre (NRS), [1] and also by contributions from the Norwegian Directorate for Cultural Heritage (RA) [2] and the Norwegian Institute for Cultural Heritage Research (NIKU) [3].

The project leader is Ole Grøn (NIKU). The Steering Committee consists of Guro Dahle-Strøm (NRS), Anna Lena Eriksson, Anke Loska, Joel Boaz (RA) and Ole Grøn (NIKU).

[1] http://www.spacecentre.no/
[2] http://www.riksantikvaren.no/
[3] http://www.ninaniku.no/

The project's participants are Ole Grøn (NIKU), May-Britt Håbjørg and Joel Boaz (RA), Hans Tømmervik and Lars Erikstad (Norwegian Institute for Nature Research, NINA), Rune Solberg, Hans Koren and Lars Aurdal (NR), Finn Christensen (GeoKem, GK), and Nancy Child (University Museum of Cultural Heritage, UKM).

Apart from the officially-defined project elements, an important factor in the fieldwork has been the collaboration and helpfulness of the local landowners, Rygge municipal administration, as well as the Cultural Heritage department in Østfold County Council (Grön and Loska, 2002).

2 Study area and data coverage

2.1 Data coverage

The chosen study area was defined by the borders of an IKONOS image data set of the Rygge Municipality, Østfold County, taken in August 2001. This image was made accessible to the project by NRS and the Satellite Data Archive of the Norwegian Mapping Authority. In WGS-84 UTM coordinates the study area is delimited by 592626-603618 E, 6575139-6586131 N (system 32), an area of 10,992 km E-W and 10,992 km N-S. The image data are geo-corrected. A check against digital ØK maps (1:5000) proved it to be accurate to within ±2m (Grön and Loska, 2002) Geographically, the study area is situated as shown in figure 1.

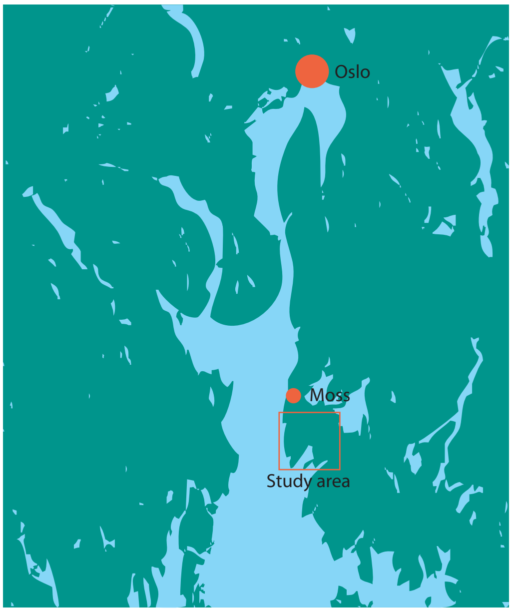


Figure 1: The study area is approximately 11 by 11 kilometers in area and covers a part of Rygge municipality, Østfold County. This figure shows the situation of the study area in relation to major cities in the area.

The image data set comprises a panchromatic image with a $1m$ resolution and a multispectral (four channels, near infrared, red, green and blue) data set with a $4m$ resolution.

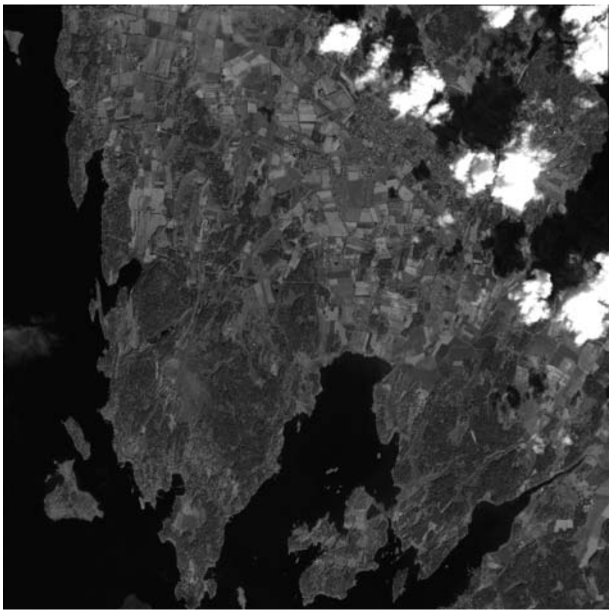


Figure 2: The study area is approximately 11 by 11 kilometers in area and covers a part of Rygge municipality, Østfold County. This figure shows the panchromatic IKONOS data.

A LANDSAT7 ETM+ image (taken in june 2001) partly covering the same study area was also available for this study. The very low resolution of these images ($60m$) renders them useless for this study. Towards the end of the project period, a Quickbird image (taken in July 2003) and also partly covering the same study area became available. Although the resolution of this image is astounding ($0.6m$ for the panchromatic data and $2.4m$ for the multispectral data), it too is less suited for this project since it seemingly is acquired at time of the year when the interesting features are less visible.

The multispectral IKONOS data were analyzed using ENVI (Research Systems, Inc.) using the Maximum Noise Fraction transform to check whether a data dimensionality reduction should be performed. This analysis indicated that the information in the four bands was sufficiently decorrelated so that no dimensionality reduction should be performed. The results of this analysis is shown in figure 3. The result of this analysis is not surprising considering the low number of bands and the high spectral separation between them (for more details concerning the Maximum Noise Fraction transform see (Green et al., 1988) or the ENVI (Research Systems, Inc.) documentation). Analysis was therefore based on the panchromatic and full multispectral data. To allow for easy comparison of the panchromatic and multispectral data, the multispectral data was interpolated to a $1m$ grid using nearest neighbor interpolation. This report is exclusively based on these interpolated IKONOS data.

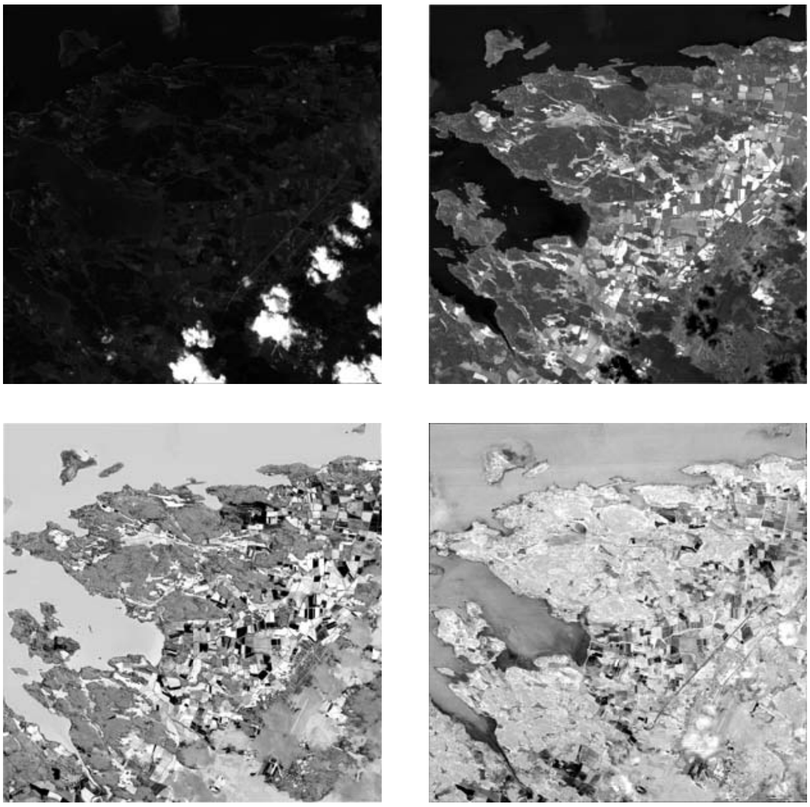


Figure 3: The result of the Maximum Noise Fraction analysis of the multispectral data. The images are ordered lexicographically in order of falling corresponding eigenvalue. It is clear that the spatial correlation in all images is sufficiently high to exclude the use of dimensionality reduction methods.

2.2 Study area

The study area is a typical, intensively exploited, agricultural production area with a quite moderate topography in Norwegian terms. With the exception of a few protruding ridges and rocks, the landscape is flat and hilly. Areas not occupied by fields and not covered by forests form an extremely limited proportion of the area (Grön and Loska, 2002).

The size of the study area is more than 100 square kilometers and the IKONOS data comprise more than 100 million pixels (at a resolution of 1×1 m). For this report we have chosen to concentrate on small parts of the total study area. In particular, it is known from field studies that the areas around the Gipsund farm are rich in cultural heritage sites. To demonstrate the potential of automatic image processing applied to the satellite data we therefore chose to extract subimages of this farm and use these subimages for the image processing.

With the current technology for satellite born imaging sensors it will not be possible to treat forest areas. The only parts of the study area that are of interest are therefore the agricultural fields. The limits of these could possibly be determined automatically, but a much better alternative wold be to use the digital land use maps provided by the Norwegian institute of land inventory (`www.nijos.no`). Such maps were not available for this study and masks for the fields were therefore made by hand. This does not detract from the relevance of this study since it is expected that such masks could easily be made using the digital land use maps.

2.3 Gipsund farm

Gipsund farm is situated in the north-eastern corner of the total study area and the fields surrounding this farm are known to comprise many interesting cultural heritage sites. We therefore extracted a subimage from the total study area shown in figure 2. This subimage comprises the central farm area along with the neighboring fields. Figure 4 shows this subimage along with the delimiting UTM coordinates. Figure 5 shows the same image on which we have overlaid the limits of the three fields of interest. This figure also indicates possible cultural heritage sites that are visible in this subimage.

Figure 4: The Gispund farm and the surrounding fields, panchromatic IKONOS data (1m resolution). The scene comprises 741 rows of 917 pixel for a total of 679497 pixels. The delimiting UTM coordinates are given in the corners of the image.

3 Methods and results

3.1 Ideal result

The ideal result of an application of an algorithm for automatic detection of cultural heritage sites in satellite data might be something like the result shown in figure 6. In this figure we show the subimage of the Gipsund farm area from figure 4 where possibly interesting sites have been labeled.

Obtaining a result like the one shown in figure 6 is not trivial and based on the current project it is obviously impossible to indicate the exact methods and algorithms to be used. However, based on experience, it is clear that the process leading from the raw input data to the kind of labeling shown in figure 6 would comprise a preprocessing step, a segmentation (clustering)[4] step and a

[4]There is much overlap in terminology concerning algorithms for grouping pixels according to their spectral

Figure 5: The Gipsund farm and the surrounding fields, panchromatic IKONOS data (1m resolution). The colored lines indicate masks for the three fields of interest. The figure also indicates possible cultural heritage sites.

classification step. In the following sections we will detail methods that we believe are relevant for the different steps of the overall algorithm.

3.2 Preprocessing

As already stated, the current sensor technology makes it impossible to detect sites of interest in areas covered by forests. It is therefore not interesting to process forest covered areas. The only interesting areas are basically agricultural fields. It would probably be possible to find these in the raw images using statistical texture classification methods, however a much better alternative would be using the digital land use maps as stated previously. We assume that generating masks for agricultural areas (like the ones shown in figure 5) would be a relatively simple task based on these maps. Such a masking would therefore be the first preprocessing step.

The next preprocessing step would have to address noise and artifact removal in the masked images. The noise present in the images is very weak and no countermeasures are actually required. The primary disturbance affecting the processing of the images are plow furrows in the fields. We suggest removing these using standard Fourier analysis techniques [5]. In Fourier analysis, the images being analyzed are projected onto a set of complex exponential functions. These projections basically measure to what degree the images under analysis resemble the different members of the set of complex exponentials. The set of complex exponentials is designed in such a way as to contain both slowly and rapidly varying members. If the image being analyzed comprises slow variations, then the projection onto the part of the set of complex exponentials that themselves vary slowly

properties. Some authors will use the term segmentation for this process, others will refer to this as clustering. We will continue to use both terms throughout this report.

[5] Fourier analysis is a standard technique in image processing, a very good reference is (Jain, 1989)

Figure 6: The Gipsund farm and the surrounding fields, panchromatic IKONOS data (1m resolution). This image shows the conceptual output of an ideal algorithm for the processing of such data.

will be important. Likewise, images with rapid variation will have an important projection onto the rapidly varying complex exponentials. Thus the Fourier analysis will provide a measure of the *frequency content* of the image being analyzed.

If the field being analyzed contains plow furrows, then this will be manifest as a fairly high frequency component in the image. A Fourier transformation of the image will reveal this as illustrated in figure 7. This figure shows what is commonly known as the **spectrum** of the image of field 2. Once the part of the spectrum that corresponds to the unwanted plow furrows has been detected, this part of the spectrum can be masked and the original image reconstructed without the unwanted features. This is illustrated in figure 8. Having masked these furrows, the remaining processing is largely facilitated.

The principal problem with this method is actually to determine exactly what parts of the frequency spectrum to mask, this is not a trivial problem. In order to determine what frequencies to mask, it is essential to detect the plow furrows in such a way that their orientation and spacing can be determined. A very good candidate for this purpose is morphological methods [6]. These methods comprise a set of image processing tools that address the local shape of the image when viewed as a surface. The image to be analyzed is actually analyzed using a small image, typically referred to as a structuring element, that captures some local feature of the overall image. In our case, we choose the structuring element to consist of a small linear image that we can rotate to search for local fits with the image being analyzed. Since the plow furrows are linear structures, the structuring element will give a good local match when oriented along the plow furrows. This

[6] Again, this is a standard image processing technique. A clear and detailed reference is (Soille, 2003)

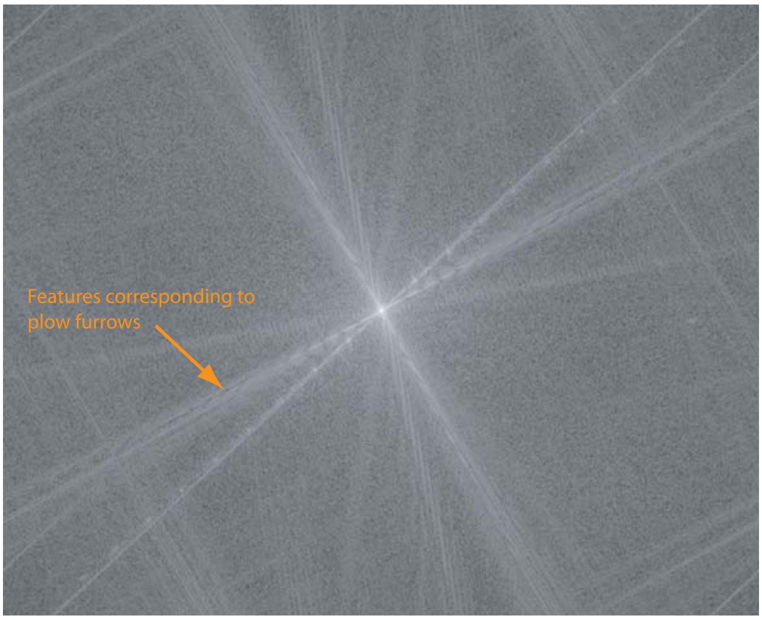


Figure 7: The Fourier spectrum of an image of the field labeled as field 2 in figure 5. The indicated line corresponds to the unwanted plow furrows.

can be used to make an overall map of plow furrow orientation.

More formally, the images to be analyzed, represented by their subgraphs, are successively eroded by small linear structuring elements and the volume of the resulting subgraph is calculated. The erosion resulting in the largest subgraph volume is assumed to correspond to an erosion were the structuring elements are aligned with the plow furrows. This method seems very robust and could be made even more robust exchanging the erosion with a parametric erosion. In figure 9 we show the subimage of field 2 along with a polar plot of the eroded subgraph volumes. We use a polar plot in order to illustrate the dependence on structuring element orientation. Notice the clear peak occurring at roughly 135 degrees, indicating the orientation of the plow furrows.

3.3 Segmentation (clustering) methods

Once the images are preprocessed, a mask should exist for every field, and in every field periodic artifacts such as plow furrows should have been heavily suppressed. Our aim at this stage is to group pixels in the fields into clusters of pixels sharing certain spectral properties. There are many ways of doing this, but the principal question to be answered before doing so is whether the fields should be treated all together or one at a time.

Before answering this question, consider the different alternative segmentation (or clustering algorithms). These can roughly be divided into two types, supervised and nonsupervised segmentation algorithms. In supervised segmentation, the user *trains* the algorithm by indicating what pixels should be grouped together. Thus the user would choose some (not all, the point is obviously to manually indicate only a few training regions) groups of pixels representative of cultural heritage sites. The algorithm will then search through all the data looking for pixels with sufficiently similar

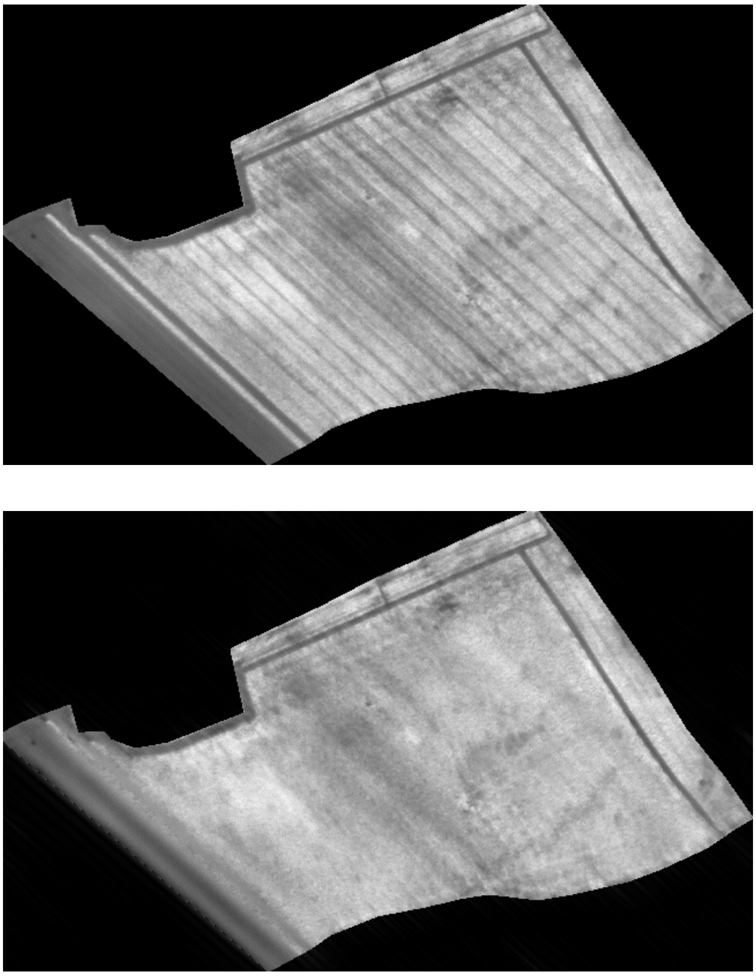


Figure 8: Subimage of field 2 before and after masking of plow furrows.

spectral characteristics and mark these.

In unsupervised segmentation, the algorithm groups pixels autonomously based on some user defined criterion. These algorithms will search through all the data looking for pixels with sufficiently similar spectral characteristics and mark these, without user interaction. In general these methods will be outperformed by the supervised segmentation methods for the simple reason that the supervised algorithms actually use extensive operator input to guide its operation. As we shall see shortly however, making good training data for the problem at hand is not easy. This problem is due to the simple fact that pixels belonging or not belonging to interesting sites may **still share spectral properties to a point where it is impossible to distinguish between them based purely on spectral properties.** The consequence of this is that training the algorithm in one field or in a small number of fields will in no way allow for overall segmentation of pixels into interesting/uninteresting groups. In sections 3.3.1 and 3.3.2 this will be made clearer by example.

3.3.1 Supervised segmentation (clustering) methods

The attractiveness of this approach is twofold. First, the algorithm can be trained according to the users desire. By including, or excluding, certain regions in the images, the user can partly control the performance of these algorithms. Secondly, these algorithms are designed, once trained, to process all the fields in one large operation. This is unfortunately also the weak point of these algorithms. The aim is obviously to keep the training set as small as possible while obtaining the desired segmentation (clustering). But small training sets will typically represent only a few very local characteristics of the cultural heritage sites, in fields far from the training site, the truth represented by the training data might no longer be valid. The alternative to supervised segmentation (clustering), unsupervised segmentation, is the subject of section 3.3.2.

The segmentation (clustering) problem is basically that of assigning labels or *classes* to all pixels in the image. In the case of segmentation (clustering) based on spectral properties, pixels sharing the same label or class will have similar spectral properties. The supervised segmentation (clustering) scheme that we evaluated in this project is the so called **maximum likelihood** segmentation (clustering). Here we will only give a very brief introduction to this algorithm, the reader is referred to the bibliography for details. Omitting the next sections will not detract from the overall understanding of this document.

In maximum likelihood segmentation (clustering) we seek to assign a class ω_i, $i = 1, \ldots, M$ to each pixel, there are thus M classes. When trying to determine what class to assign to a given pixel, we will only consider the conditional probability:

$$p(\omega_i|\mathbf{x}), \quad i = 1, \ldots, M$$

Here, the vector $\mathbf{x}$ is a column vector of spectral values for the pixel being considered. $p(\omega_i|\mathbf{x})$ is the *likelihood* that the correct class for the pixel considered is ω_i given that the pixel has spectral values $\mathbf{x}$. The classification is based on the following simple rule:

$$\mathbf{x} \in \omega_i \quad \text{if} \quad p(\omega_i|\mathbf{x}) > p(\omega_j|\mathbf{x}) \quad \text{for all } j \neq i$$

That is, the pixel with spectral values $\mathbf{x}$ belongs to class ω_i if $p(\omega_i|\mathbf{x})$ is the largest.

The problem with this seemingly simple decision rule is that the probabilities $p(\omega_i|\mathbf{x})$ are unknown. But what if training data were available, that is, what if we had marked regions of pixels to which the user had **assigned** a class. We could use this to estimate the probability $p(\mathbf{x}|\omega_i)$, that is, the probability of certain spectral signatures **given its class membership**. For each class we would then know the probability that a given spectral signature belonged to the class. The beauty of this approach is that the two types of probabilities are related by the following simple formula (Bayes formula):

$$p(\omega_i|\mathbf{x}) = p(\mathbf{x}|\omega_i)p(\omega_i)/p(\mathbf{x})$$

where $p(\omega_i)$ is the probability that the class ω_i occurs in the image, $p(\mathbf{x})$ is simply the probability of a given spectral signature. Based on this formula, we can now make the following decision rule:

$$\mathbf{x} \in \omega_i \quad \text{if} \quad p(\mathbf{x}|\omega_i)p(\omega_i) > p(\mathbf{x}|\omega_j)p(\omega_j) \quad \text{for all } j \neq i$$

where the $p(\mathbf{x})$ have been removed as a common factor. This simple decision rule is at the basis of the so called **maximum likelihood classification**. An introduction to the vast subject of supervised clustering can be found in (Richards and Jia, 1999). An excellent in-depth treatment is provided in (Duda et al., 2000).

A simple test of this algorithm on our images reveals its weaknesses in the present situation. Starting by defining a training area for a single class, we mark a region on the Gipsund fields as belonging to the class of cultural heritage sites (the particular site marked is identified by the collaborating archaeologist as an interesting site). Given this training region, we use ENVI to calculate the corresponding *rule image*. This is simply a new image in which every pixel is labeled with its probability (its likelihood) of belonging to the class defined by the training class. Both the image showing the region used as a training region and the resulting rule image is shown in figure 10.

Notice in figure 10 the high degree of confusion, very many pixels **outside** the training region share the spectral properties of the training region.

From this simple test it is possible to draw very important conclusions. In particular, we can conclude that it will **not** be possible to distinguish between interesting sites and features of no interest using this simple scheme. As is clearly seen from this example, interesting and non-interesting sites share spectral properties to a degree where there is no hope of distinguishing between them in a automatic manner. This is made worse by the fact that the training region and application region were very close in this case, actually the whole test is applied to only a few of the Gipsund fields. If the training and application regions were more separated geographically, it is likely that results would become even worse.

Another simple way of observing this is by plotting the histograms of different regions around Gipsund. We show this in figures 11 and 12. The upper part of these figures show (in red and orange) the two different regions whos histograms we compare. Figure 11 shows the raw histograms, in figure 12 we show the same histograms, but normalized and offset for clarity. Notice the large degree of overlap between the histograms of the two regions, this explains why it is hard to distinguish between the regions based on purely spectral data.

Based on these observations it is reasonable to assume that supervised segmentation (clustering) will play a very limited role in this project. As we have seen, interesting and uninteresting sites share spectral properties to a large degree. Secondly, the fairly low resolution of the images relative to the features of interest also makes it difficult to obtain good statistical information to separate between interesting and uninteresting sites. As we shall see in section 3.4, supervised segmentation (clustering) will play a role in processing **shape and size features** of the spectral clusters. For this application, it will be more appropriate to refer to this method as a classification method.

3.3.2 Unsupervised clustering methods

For a segmentation problem like the one at hand, we have seen that the operator input is of little value. What we would actually like to have is an algorithm that could treat every field separately (or possibly in conjunction with its neighbors) and make **local** decisions as to whether the fields observed contain regions with desirable spectral parameters.

One much used method for unsupervised segmentation (clustering) is the so called k-means algorithm. The name of this conceptually very simple algorithm reflects its basic operation, it tries, in an automatic fashion, to assign the pixels to one of k classes by assigning the pixel to the class whose mean spectral value is the closest. This algorithm therefore tends to to assign pixels to classes in such a manner that the mean squared distances from pixels to class centra are minimized. Here we will only give a very brief introduction to this algorithm, the reader is referred to the bibliography for details. Omitting the next sections will not detract from the overall understanding of this document. The algorithm works as follows:

1. Initialize the procedure by selecting k points in multispectral space to serve as candidate cluster centers. Let these be:
$$\hat{\mathbf{m}}_{\mathbf{i}}, \quad i = 1, \ldots, k$$
The cluster centers can be chosen arbitrarily, but no two should be the same. A much used approach is simply to space the class centra uniformly in spectral space.

2. The location of each pixel in spectral space is examined and the pixels are assigned to the nearest cluster centre. 'Nearness' would typically be measured using a Euclidean distance measure.

3. Based on the new grouping, recalculate cluster centers. Let these be:
$$\mathbf{m}_{\mathbf{i}}, \quad i = 1, \ldots, k$$

4. If $\mathbf{m}_{\mathbf{i}} = \hat{\mathbf{m}}_{\mathbf{i}}$ for all i the procedure is terminated. Otherwise, $\hat{\mathbf{m}}_{\mathbf{i}}$ is redefined as the current value of $\mathbf{m}_{\mathbf{i}}$ and the procedure is reiterated from step 2.

This very simple algorithm, once initialized, is capable of assigning pixels to classes in an autonomous fashion. The **number** of classes must be given by the user, as we shall see shortly this is not necessarily a large problem. Our intended use of this algorithm is to apply it **within each field**. As previously stated the fields can be delimited using land use masks so the fieldwise application should be no problem. Another problem that must be dealt with in the future is that of cultural heritage sites straddling several fields, this could be solved by treating small groups of neighboring fields instead of each field on its own.

An initial attempt at using the algorithm in the Gipsund fields is illustrated in figure 13. We use three clusters ($k = 3$) in this example, one is intentionally introduced to capture the background, the two others are used to cluster actual field data. This initial attempt is clearly seen to result in an underclustering of the data. Increasing the total number of classes to five ($k = 5$), we get the result shown in figure 14.

3.4 Classification methods

An important question to be answered is obviously what clusters among those resulting from the above procedures correspond to interesting sites. Notice that k-means will only cluster the pixels according to a distance measure applied in spectral space. The **meaning** of these clusters is not considered during the clustering so **any one of the clusters** could represent an interesting cluster.

Although supervised segmentation (clustering) was found to perform poorly in grouping pixels according to *spectral properties*, we expect these methods to perform better when trying to *separate between clusters based on their shapes etc.* The reason for this is simple: As we have seen, candidate sites may have a number of spectral signatures, not in any way distinct from those of non-candidate regions. This is why we suggest a *fieldwise* treatment of the data, processing each field with a locally adaptive algorithm.

We expect the clusters corresponding to interesting regions to share shape (and other characteristics) to a much larger degree than was the case for spectral properties. This is so *simply because remains of prehistoric human activity must have certain shape and size restrictions*. In particular, very large clusters can be ruled out from consideration based on the simple fact that man-made structures from the periods in question were of limited size. Furthermore, it is reasonable to assume that interesting regions will have certain compactness and shape properties that can be used to distinguish them from non-interesting regions

We have not gone far in the investigation of this problem, but believe that the following properties might be apropriate for separating between interesting and uninteresting clusters (see also examples in figure 15.

1. **Upper limit on size:** Remains of prehistoric human activity must be upper limited in size.

2. **Lower limit on size:** Considering clusters that are too small will make the algorithms much to sensitive to noise.

3. **Rectangularity and polygonality:** Very rectangular or polygonal regions are probably results of current-day activity.

4. **Compactness:** Interesting clusters are probably fairly compact. Thus clusters surrounding others are probably not interesting

5. **Aspect ratio:** This measure closely related to the previous measures the ratio between the minimal and maximal length of lines inscribable in the considered cluster. If this is too small the object will be a very elongated structure.

6. **Grouping:** Human activity is often concentrated to small areas, the presence of one interesting cluster should increase the probability of nearby clusters being considered as interesting.

7. **Others:** A very large number of shape properties have been devised. Testing all these for their discriminatory power will be necessary

4 Conclusion

Our principal conclusion based on this initial study is that computerized processing of satellite data in search for cultural heritage sites is feasible. Furthermore, the processing will most probably be a three-step process consisting of the following steps:

1. Preprocessing: We are convinced that a field by field treatment is necessary. In preparation for further processing, the field masks derived from land use maps must be applied so that each field can be treated apart. Furthermore, every field should be preprocessed in order to suppress artifacts that could interfere with the clustering. In particular, plow furrows must be removed. Fourier analysis in combination with mathematical morphology will probably be a good candidate.

2. Spectral segmentation (clustering): Again, this must be performed on a field by field basis. Due to the large variability in the data, an unsupervised clustering (k-means) must be applied, possibly in a multistage process, to each field. This will cluster the pixels in the field according to their spectral properties.

3. Object parameter clustering (classification): Having clustered the pixels according to spectral properties, a final processing step consists in processing the different clusters (objects) looking for clusters having shape and size characteristics that make them candidates for cultural heritage sites. It is our belief that this step will be fairly simple, since little can be said in general of the shape of a cultural heritage site. Most probably, the discrimination will be made on some very simple size an shape criteria.

We have had access to satellite images from three different satellites during this project, in order of increasing resolution these are: LANDSAT7 ETM+ , IKONOS and QuickBird. From this initial study it would seem that a resolution equal to or better than that of the IKONOS data is necessary, that means that the panchromatic data should have a resolution of roughly $1m$ or better and the multispectral data a resolution of roughly $4m$ or better. The multispectral part of the data can then be interpolated to the same scale as the panchromatic data providing a reasonable resolution in the final images. The choice of a minimum of $1m$ resolution also seems reasonable based on a consideration of the structures that are to be resolved on the ground. Data from the IKONOS or QuickBird satellites will probably be well suited for a project like this.

Another issue concerning the satellite data is obviously the time of the acquisition. It is very surprising to see how ground details visible in one set of data might be more or less invisible in others. We strongly suggest that a future project addresses this issue.

Furthermore, based on this initial study, we are convinced that a fully automated system is nearly impossible. This is due to the simple fact that even human specialists frequently reach different conclusions interpreting the images. The decision as to whether a site is or is not interesting depends on many details. The spectral signature and shape of the site is only one of several factors that must be taken into account. Knowledge in archaeology and local history coupled with geography and, obviously, knowledge from field surveys will often be of primary importance in the final interpretation. We believe that if the development of software for processing satellite images looking for cultural heritage sites is undertaken, the aim should be to provide the human specialists with a tool for detecting *potentially interesting sites*, leaving the final interpretation to the human specialists. Such a tool would greatly reduce the burden on the human specialist as it would be able to guide the specialist from site to site in the images. The specialist would then concentrate on the actual interpretation of the different sites that are detected. Furthermore, such a tool could learn from input from different users thus paving the way for more objective decisions.

5 Future work

Assuming that a larger scale project is initiated in order to further develop this work, this is a (partial) list of issues that we recommend addressed in order to move forward in the project:

1. A more thorough investigation of the different sources of satellite data is warranted. We have considered only LANDSAT7 ETM+, IKONOS and QuickBird images. We currently believe

that a panchromatic resolution of $1m$ is required, it is necessary to detail what satellites are capable of producing such data. This is necessary for several reasons. First of all, this is interesting in order to keep the investment in satellite data to a minimum. Satellite data are expensive, and the least costly source should be considered. Secondly, as mentioned in the conclusion, satellite coverage and the time of the acquisition is critical. Having access to several satellite data sources might provide images taken at optimal points in time.

2. A study should be undertaken to detail the causes of variation in the satellite data. It is very surprising to see the large differences between the IKONOS and QuickBird images, in the future it should be possible to explain these differences and to suggest optimal times of the year as well as optimal climatic conditions for the data acquisitions.

3. It would greatly help in the interpretation of the satellite data findings if more was known about the ground truth. The ongoing investigations of soil chemistry at potentially interesting sites are very important and should be used to further increase the quality of the interpretation of the satellite data.

4. The use of masks for agricultural land should be further investigated. These masks are currently available from NIJOS, but have not been tested in this project. The availability and quality of these masks should be assessed. Potentially, methods for automatic detection of agricultural land should be explored, both in order to deal with national inconsistencies and to facilitate the use of these methods in regions were such masks might not be available.

5. The locally adaptive method for clustering regions within each field should be further developed.

6. Criteria concerning the size and shape of regions to be considered as possible cultural heritage sites should be established. This has not been a focus of the current project since it is expected that only relatively simple criteria can be used for this purpose. Nevertheless, such criteria must be established.

7. Extensive testing of the methods developed will require a close connection to the system managing the satellite data. Developing such a system from scratch will require an enormous amount of work, and integration to an already existing system is the only sensible way of doing this. We have used ENVI for this project and found it to be flexible and well adapted. ENVI can easily be extended with user-specified functionality (written in C/C++ and /or IDL).

References

R. Duda, D. Stork, and P. Hart. *Pattern Classification*. Number ISBN 0471056693. John Wiley & Sons Inc., 2nd edition, 2000. 12

A. Green, M. Berman, P. Switzer, and M. Craig. A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1):65–74, jan 1988. 4

Ole Grön and Anke Loska. Development of methods for satellite monitoring of cultural heritage sites. Technical report, NIKU, December 2002. 2, 3, 5

A. Jain. *Fundamentals of Image Processing*. Number ISBN 0133361659. Prentice Hall Inc., 1st edition, 1989. 7

J. Richards and X. Jia. *Remote Sensing Digital Image Analysis: An Introduction*. Springer Verlag, 2nd edition, 1999. 12

P. Soille. *Morphological Image Analysis, Principles and Applications*. Springer Verlag, 2nd edition, 2003. 8

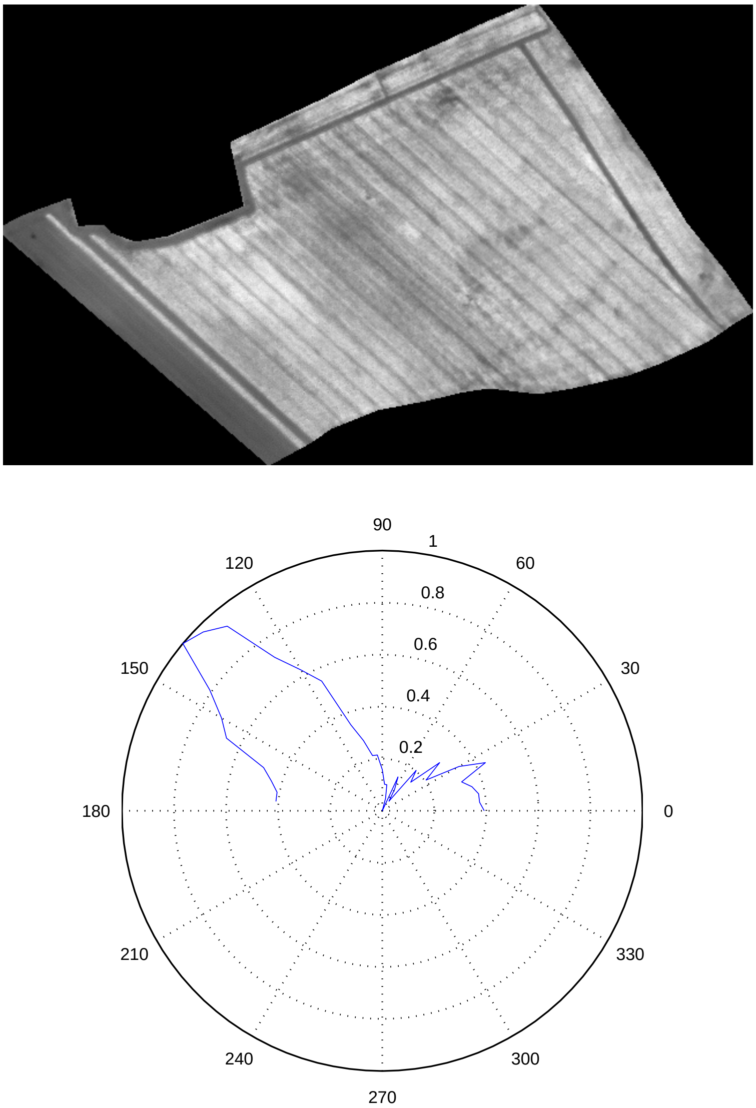


Figure 9: Subimage of field 2 and polar plot of normalized eroded subgraph volume. Notice the distinct peak at roughly 135 degrees indicating the predominant direction of plow furrows.

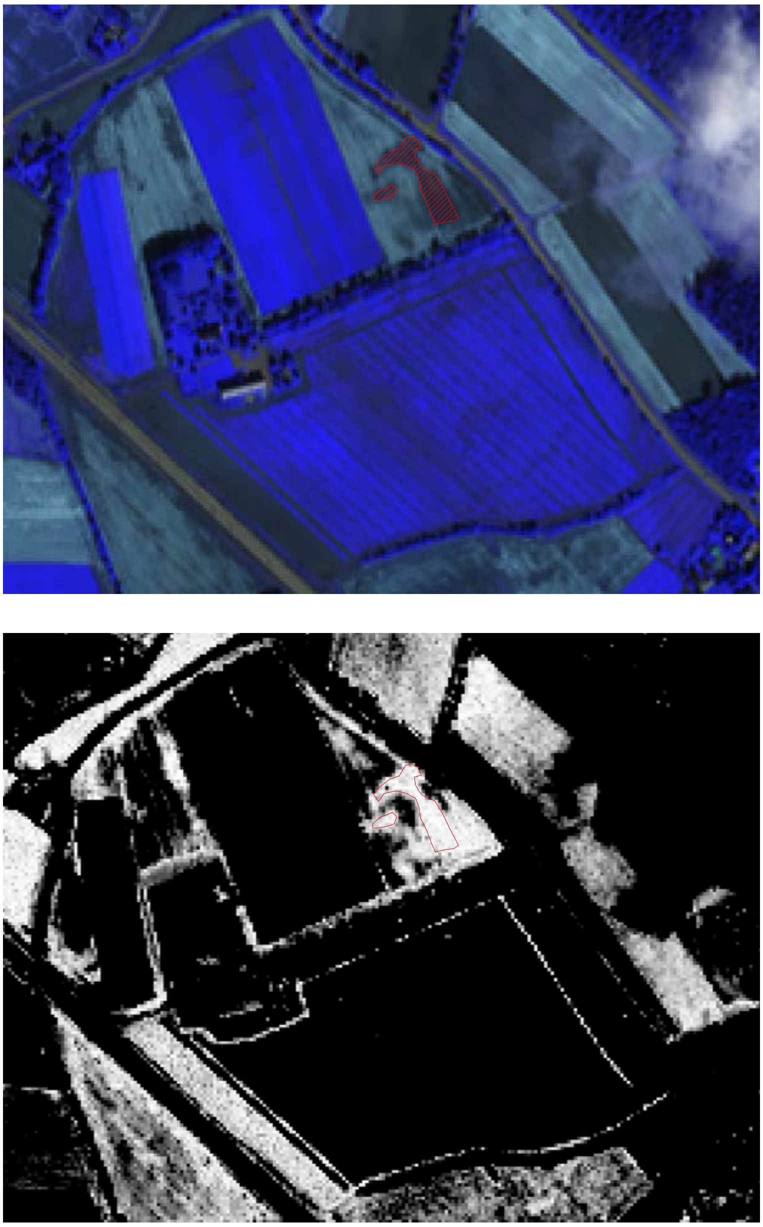


Figure 10: The region used as a training region and the resulting rule image. Probabilities are mapped to gray levels. Dark colors indicate a low probability, light colors a high probability. Notice the high degree of confusion, very many pixels **outside** the training region share the spectral properties of the training region. Among the pixels sharing the spectral properties of the training region are pixels that clearly do not represent interesting sites.

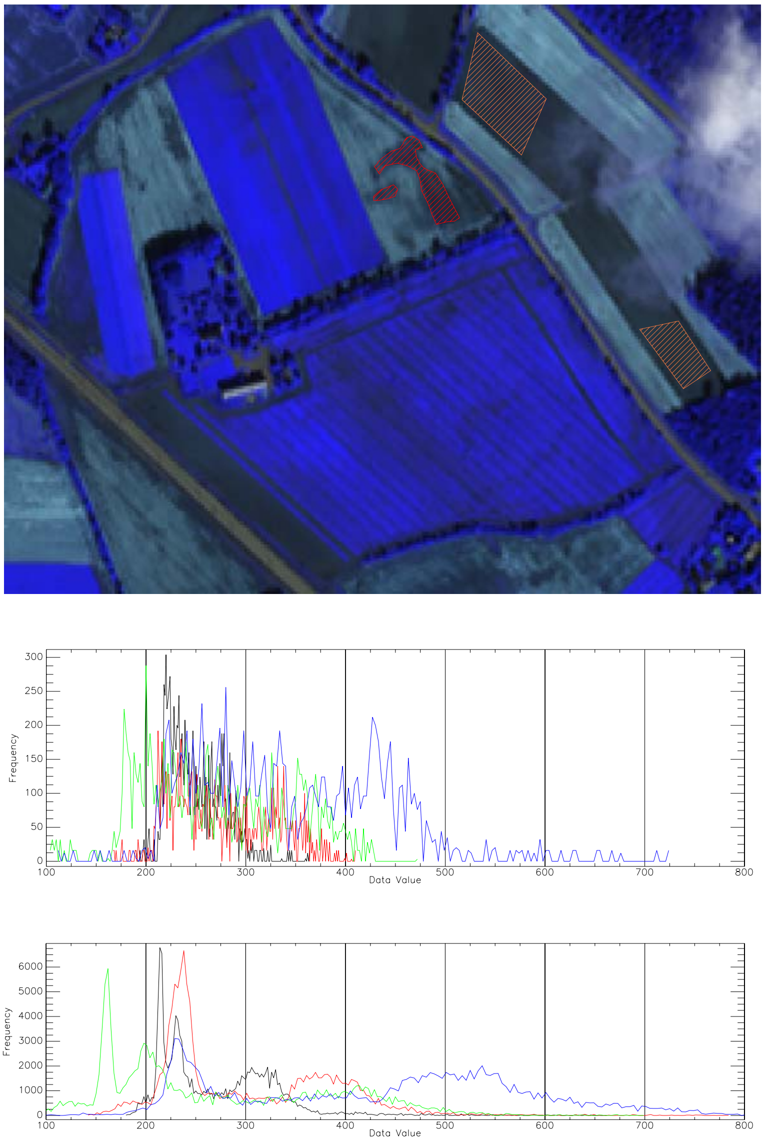


Figure 11: The upper image shows the two different regions used to compare regional histograms. The histogram of the red region appears in the middle plot, the histogram of the orange region in the bottom plot. The four different lines in each plot correspond to the four different spectral ranges in the multispectral data. Notice the large degree of overlap, this explains the difficulties we are having in separating these regions based purely on spectral values.

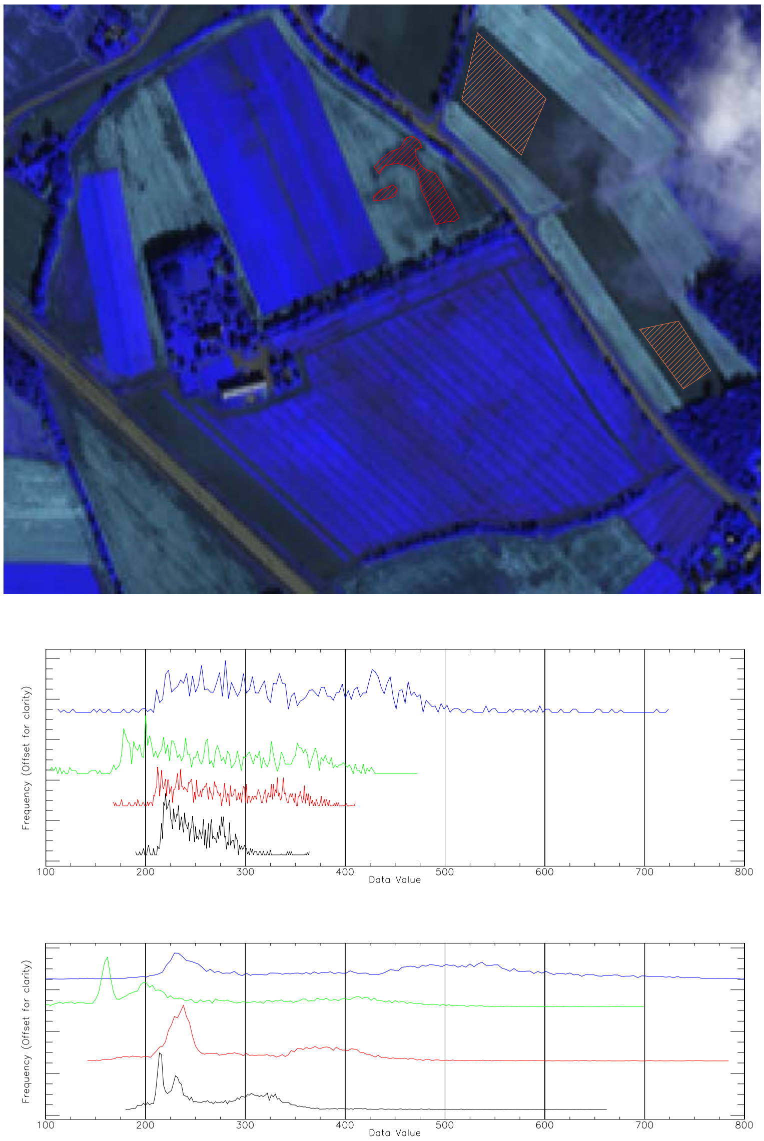


Figure 12: The upper image shows the two different regions used to compare regional histograms. The histogram of the red region appears in the middle plot, the histogram of the orange region in the bottom plot. The four different lines in each plot correspond to the four different spectral ranges in the multispectral data. In this figure the histograms have been normalized and offset for clarity. Notice the large degree of overlap, this explains the difficulties we are having in separating these regions based purely on spectral values.

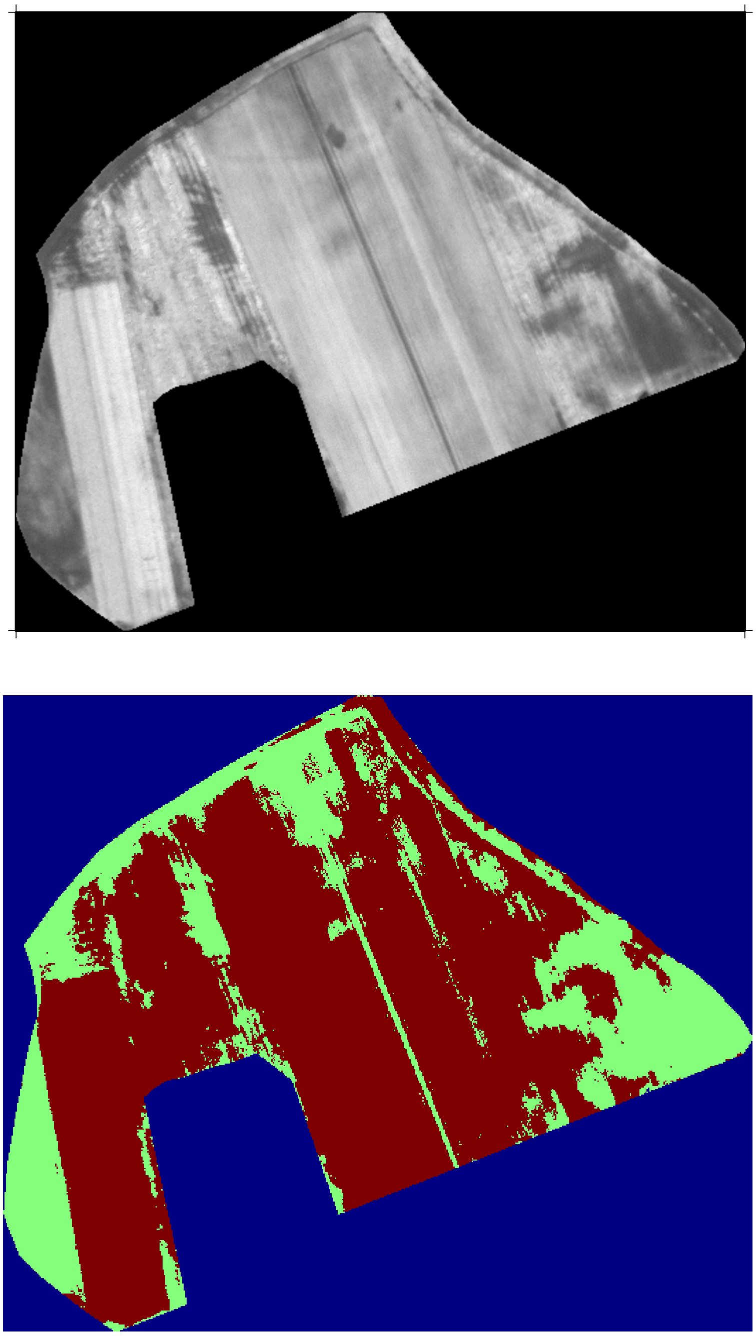


Figure 13: A Gipsund field masked from the remaining image using manually made masks. The bottom part of the figure shows a k-means clustering of this image into three ($k = 3$) clusters. One cluster is intentionally used to capture the background, the other two are used to cluster actual field data.

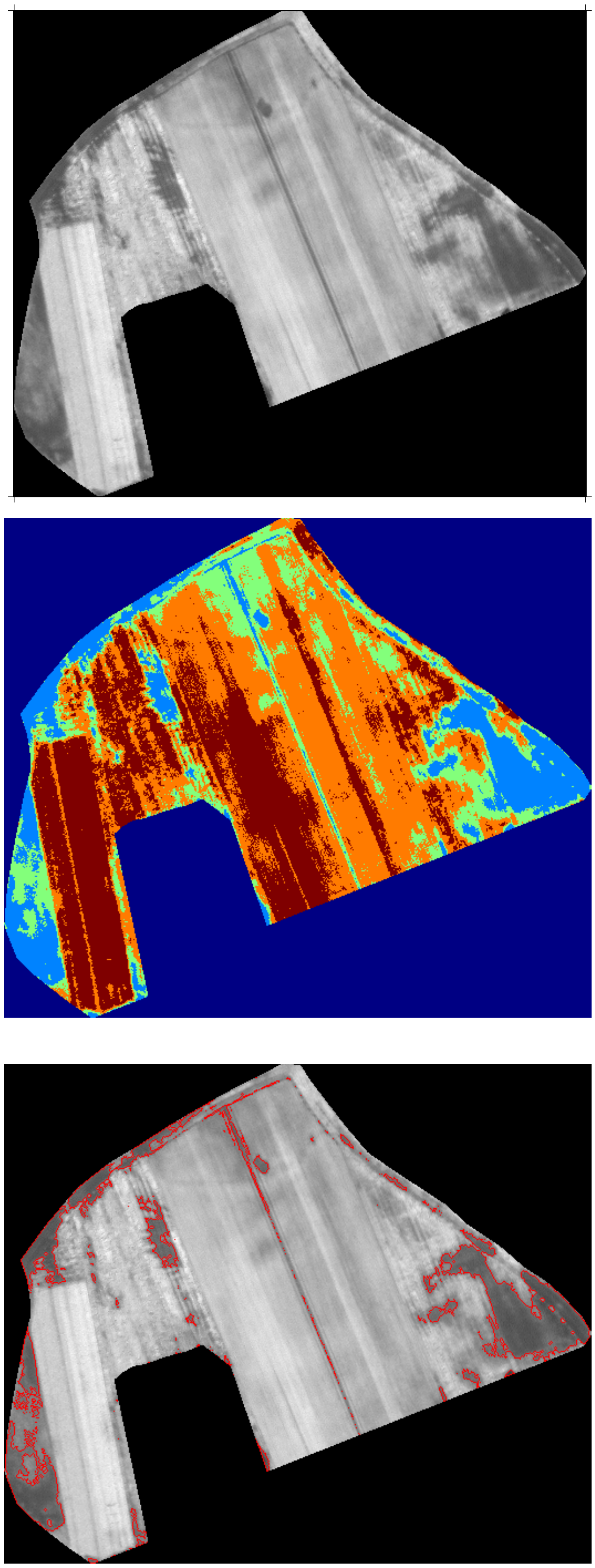


Figure 14: k-means clustering using 5 classes on a Gipsund field. The upper left image shows the initial field, the upper right image shows the clustering result and the lower image shows the contour of the interesting class overlaid on top of the original image.

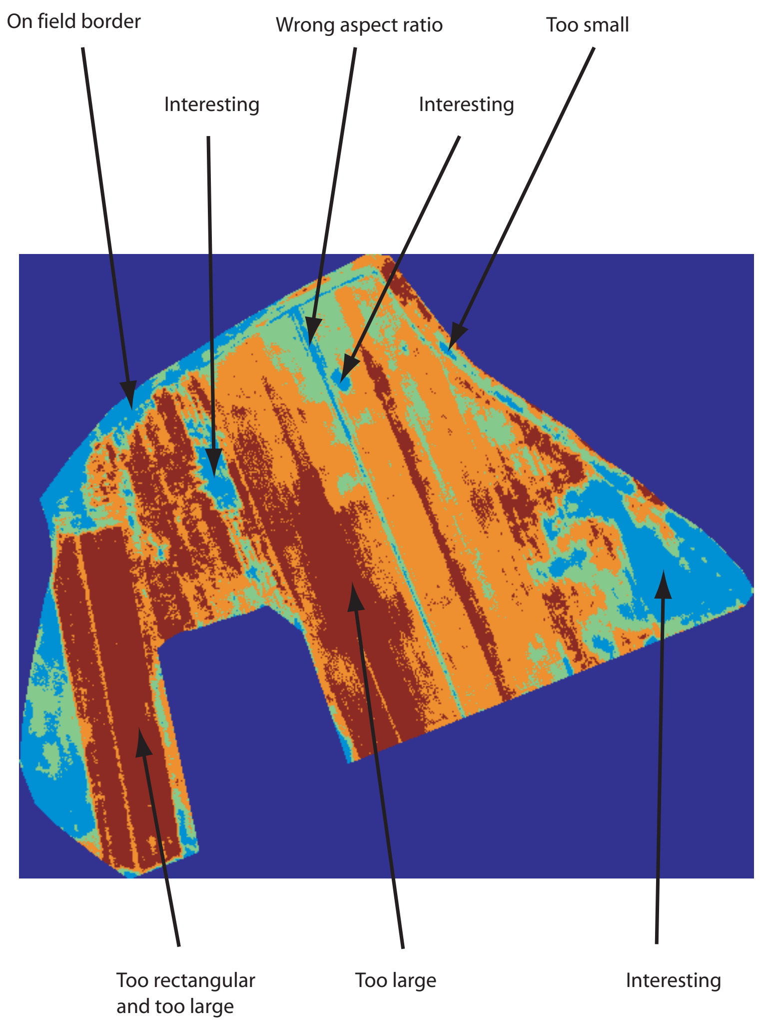


Figure 15: Different properties of the clusters from figure 14.