Automatic detection of circular structures in high resolution satellite images of agricultural land

Short title: "Automatic detection of circular structures in satellite images"

Øivind Due Trier*, Siri Øyen Larsen, and Rune Solberg†

September 9, 2008

*Corresponding author, telephone +47 22852698, fax +47 22697660

†Ø. D. Trier, S. Ø. Larsen and R. Solberg are with the Norwegian Computing Center, Gaustadalléen 23, P.O. Box 114 Blindern, NO-0314 Oslo, Norway, email: oivind.due.trier@nr.no, siri.oyen.larsen@nr.no, rune.solberg@nr.no.

Abstract

Archaeological sites are sometimes visible in satellite images as soil or crop marks. At best, the marks are distinct, but they tend to have less contrast with the background than many other patterns in the images. Consequently, reliable automated detection based on pattern recognition is very difficult.

Our method detects circle shaped soil and crop marks in the panchromatic band of high-resolution satellite images of agricultural fields. Such circular marks may be caused by burial mounds.

In our approach, local contrast enhancement is applied in order to make weak marks more distinct. The image is then convolved with ring templates of varying sizes, giving high absolute values at candidate circular mark locations. Each candidate mark is presented to an operator, who may reject it.

We tested our method on Quickbird images from south-east Norway. The number of detected candidate marks could be varied by changing a threshold value. A reasonable compromise between not detecting too many false rings and at the same time detecting as many true rings as possible, might be when the number of false detections is approximately seven times the number of true detections. In this case, 11 out of 15, or 73%, of the strong rings were detected, and 5 out of 10, or 50%, of the fairly strong rings were detected. This is 16 out of 25 of the strong and fairly strong rings, or 64%.

Archaeologists state that the software tool we develop will be helpful for locating potential cultural heritage sites. Although it makes many false detections, it will relieve the operators from time consuming manual inspection of entire images.

Keywords Pattern recognition, soil marks, crop marks, Quickbird, contrast enhancement, remote sensing.

1 Introduction

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, etc.) exert growing pressure on cultural heritage in the landscape. In order to match the political intentions of updated and sustainable cultural heritage management, it is necessary to develop a cost effective method for locating and monitoring cultural heritage sites. Given the enormous costs of surveying the areas in question by traditional fieldwork, alternatives must be sought. The use of modern support technologies is imperative, if such rapid changes are to be balanced against the sustainable management of this resource.

Aerial photographs have been used for many years to identify archaeological sites, both visible and buried (Wilson, 1982); for recent research, see, e.g., (Campana et al., 2006; Musson et al., 2006). Many sites manifest themselves as soil marks, crop marks, shadow marks or frost marks (Wilson, 1982). These marks may be visible in the images as contrasts in spectral reflectance to the surroundings.

In addition to airborne multispectral imaging, recent advances in airborne LIDAR are also promising (e.g., see Devereux et al., 2008; Risbøl et al., 2006; Sittler and Schellberg, 2006). To collect airborne imagery and LIDAR is nevertheless both time consuming and costly, thus it imposes severe limitations on the size of the areas that can be investigated.

In this respect, multispectral and panchromatic images from sun-synchronous earth observing satellites provide a complementary source of image data. Panchromatic images have ground pixel sizes varying from 15 m (Landsat 7, since 1999), 10 m (Spot, since 1986) to 1.0 m (Ikonos, since 1999), 0.6 m (Quickbird, since 2001), and 0.5 m (Worldview, since

October 2007). Multispectral images have ground pixel sizes of 30 m (Landsat), 20 m (Spot), 4 m (Ikonos) and 2.4 m (Quickbird) (Cracknell and Hayes, 2007). These have been used for manual and automated detection of archaeological sites by several research groups (see below). Despite coarser resolution than aerial images, digital satellite images are attractive due to global coverage and a simpler viewing geometry because of the long distance between the sensor and the target. Pattern recognition methods have been applied successfully on high resolution satellite images in other research areas, such as automatic detection of vehicles (Larsen et al., 2008a), vehicles and roads (Oostdijk et al., 2008), and buildings (Sumer and Turker, 2008). It should be noted that prior to 1999, 10 m ground resolution was the best available from satellites, with the majority of the research devoted to classification of multispectral pixels, mainly for land use assessment. So, pattern recognition of human made structures in high resolution satellite images is a relatively new research field in general.

Satellite image based location, surveillance and monitoring of cultural heritage sites has been the subject of some recent research. Trelogan *et al.* (1999) and Carter *et al.* (2000) used maximum likelihood classification of two Landsat TM images, acquired in 1988 and 1992, to automatically detect urban development surrounding an archaeological site. Blom *et al.* (1997; 2000) used Landsat images in combination with other remote sensing sources to visually locate sites. Aminzadeh and Samani (2006) used various image enhancement methods on Landsat ETM+ images to assist visual interpretation, including various band ratios, spectral principal component analysis, edge detection, and high-pass filtering. This resulted in different false-color images that could be used for visual inspection and detection. Goossens *et al.* (2006) and Fowler and Fowler (2005) used historic images from the former Corona military reconnaissance satellite (operated from 1960 to

1972; 2 m pixel size) for visual inspection. Lasaponara and Masini (2007) used pan-sharpened multispectral Quickbird images to detect crop marks, and found that the near infrared band was better for crop marks in dry vegetation, and the normalized difference vegetation index (NDVI) was better for green vegetation. Based on this, edge detection in the near infrared band or the NDVI image was used to automatically identify candidate crop marks. This was followed by manual line extraction. De Laet *et al.* (2007) compared three different methods to automatically detect above ground sites in pan-sharpened multispectral Ikonos images: (1) edge enhancement (least successful), (2) maximum likelihood classification of pixels, and (3) segmentation followed by nearest-neighbour classification. All the methods were able to extract some archaeological objects, but also had major limitations, due to large variations in shape and color.

Earlier work at our institute (Aurdal et al., 2006) focused on detecting crop and soil marks of arbitrary shape, but recent experiments on the current data set gave inconclusive results. Further, archaeologists have identified several circular crop and soil marks in the data, which led us to consider methods tailored to detecting circular marks. Lemmens et al (1993) used edge detection to find circular objects in a scanned aerial image from Oxfordshire. From each circle center candidate, a search was made in eight directions (0°, 45°, 90°, ...) to simultaneously locate edges of the correct direction at the same radius. Partly occluded circles were allowed by accepting fewer than eight simultaneous matches. However, in our data, the rings are often weak, and a method based on edge detection would not be successful.

The rest of the paper is organized as follows. Section 2 outlines the background for the *CultSearcher* prototype system. Section 3 describes the experimental data set. Section 4 describes the various methods we have attempted to use, and summarizes the chosen

algorithm. Experimental results are described in Section 5, and discussed in Section 6. Finally, conclusions are given in Section 7. Preliminary results from this work have been presented at international conferences (Larsen et al., 2008b; Trier et al., 2008b).

2 The CultSearcher prototype system

The Norwegian Directorate for Cultural Heritage, in collaboration with the Norwegian Computing Center, the Norwegian Institute for Cultural Heritage Research, the Museum of Cultural History at the University of Oslo, and Vestfold County Administration, started in 2003 a project with the overall aim of developing a cost effective method for surveying and monitoring cultural heritage sites on a regional and national scale. The Norwegian Computing Center has been responsible for developing the automatic detection methodology and implementing this into a prototype software system, *CultSearcher*.

A main objective of the study presented here was to analyze the utility of satellite images for archaeological mapping. It was partly motivated by the fact that mapping on a national scale is cheaper with satellite imagery than with aerial images. It was also motivated by the fact that geometrical effects, like radial distortion, are much smaller in satellite images that aerial images.

CultSearcher is currently analyzing soil marked and crop marked patterns. Soil marked sites are typically the remains of ditches or pits, buried walls, etc. Crop marks are an indirect effect of buried archaeological structures.

The aim of CultSearcher is to provide computerized assistance to the operator in the analysis of satellite images. In particular, the software identifies potential sites, for further inspection by an archaeologist. This means that the archaeologist may concentrate

on analyzing the identified sites rather than the entire image. Note that the system is designed to detect candidate sites and that no claim is made that these candidates are true cultural heritage sites. Even human specialists cannot make such an assertion based on satellite imagery alone. The verification of a potential site always depends on some kind of field inspection.

3 Experimental data set

The data set cosists of two Quickbird images (Figure 1). Image 'Laagen' was acquried on April 27, 2005 at 10:45AM, from the valley Lågendalen between Kongsberg and Larvik. Image 'Gardermoen' was acquried on July 29, 2003 at 10:23 AM, from an area surrounding, but not including, the Oslo Gardermoen airport. Both images consist of a four-band multispectral image and a single-band panchromatic image. The panchromatic image has 0.6 m ground pixel size, and covers the 450-900 nm wavelengths. The multispectral image has 2.4 m pixels, and the four bands are: blue (450-520 nm), green (520-600 nm), red (630-690 nm), and near-infrared (760-900 nm).

There have been numerous archaeological investigations in both test areas, and the areas are known for their relatively high density of archaeological sites.

The Gardermoen area is an undulating glacial landscape consisting mainly of well-drained sandy soils cut by an intricate system of ravines. This fertile area has many agricultural fields in which crop marks have previously been reported. The image acquisition date should be ideal for detection of crop marks, since this is the time of the year when the crops are in the process of ripening and thus turning yellow.

The Lågendalen area is a distinct u-valley in which the river Lågen has eroded its

meandering way through a thick layer of marine clay. The area now consists of agricultural fields and forest. The area includes a known Iron Age grave field at Odberg farm. The date of this image is too early for crop marks to have been fully developed, but soil marks could be expected.

Further understanding of the optimal time and optimal conditions for soil and crop marks is important for large-scale use of satellite images for archaeological mapping. This would require an extensive data set acquired over several years (due to the cloud conditions in Norway), and is intended to be carried out in a future study.

Archaeologists have identified many circular patterns that are clearly visible in the panchromatic images (Figure 2a), but can hardly be seen in the multispectral images (Figure 2b–2c). These are believed to be the remains of burial mounds. The mound itself has been destroyed, but the circular ditch often remains. Recently, other research groups have used multispectral Quickbird (Lasaponara and Masini, 2007) or Ikonos (De Laet et al., 2007) images to locate cultural heritage sites, but the objects they were looking for were much larger than the circular marks in the present work. Since the circular patterns are difficult to spot visually in the multispectral bands, this might indicate that the lower resolution multispectral bands will not contribute significantly. However, it would have been interesting to perform experiments where the multispectral information was used as well, and we plan to do so in a future study.

In the two Quickbird images, archaeologists have identified 35 locations with ring marks that they would like the system to recognize. We have visually classified 15 of these as 'strong', 10 as 'fair' and 10 as 'weak' (Figure 3). At a few of the locations, there are two rings located only a few meters apart; these are counted as one ring only, with the understanding that if one is correctly located, then the archaeologist has been directed

to an interesting site, and whether one or two rings are marked by the program is not important. 11 subimages of 4096×4096 pixels were extracted for the experiments. These subimages included all 35 ring mark locations.

4 Methods for detecting ring marks

As noted in Section 3, the remains of burial mounds are often visible as ring marks in the images.

In order to detect as many ring marks as possible, while at the same time keeping the number of false positives at a minimum, variations of the following sequence of methods have been tried out:

1. low pass-, band pass- or high pass-filtering in the frequency domain;
2. local contrast enhancement;
3. template matching;
4. feature extraction; and
5. decision tree-based classification.

Each of these will be discussed below. Both the various approaches tested and the chosen methods are described.

4.1 Filtering in the frequency domain

Since the ring marks are relatively weak in the images, we wanted to remove image information not related to the ring marks, including plow furrows. We investigated if this

could be done in the frequency domain by use of the fast Fourier transform followed by low pass, band pass or high pass filtering.

The filter that seemed to preserve the rings best and at the same time suppress as much other information as possible, was the band pass filter with cut-off radii 100 and 800. We verified visually that there was no trace of the rings in the information that was removed, by using a band cut filter. This was double-checked by using a low pass and a high pass filter. A lot of low- and high-frequency information was removed. However, the plow furrows were too close in frequency to the rings to be removed. The plow furrows were only slightly reduced in strength.

4.2 Local contrast enhancement

Even the most distinct rings in the test images have relatively low contrast with their surroundings. In order to be able to detect any rings at all, the local contrast has to be more or less constant over the entire image. This can be achieved by, for each pixel, computing the local mean gray level and associated standard deviation in an $N \times N$ square neighbourhood centered on the pixel. The pixel value $p_{CE}(x, y)$ in the contrast enhanced image is computed as

$$p_{CE}(x, y) = \frac{p(x, y) - \mu(x, y, N)}{\sigma(x, y, N)}, \tag{1}$$

where $p(x, y)$ is the gray level value in the input image, $\mu(x, y, N)$ is the mean gray level value in an $N \times N$ neighbourhood centered on (x, y), and $\sigma(x, y, N)$ is the standard deviation of the gray level values in the same neighbourhood.

The choice of the neighbourhood size, N, does not seem to be critical. We have chosen to use $N = 21$. However, using, say, $N = 15$ or $N = 35$ also works quite well. Having a too small value for N may result in exaggeration of even small local variations in gray level.

Having a too large value for N may suppress the local contrast needed to identify the rings. Another undesirable effect is that the local contrast is suppressed when the center pixel is less than $N/2$ pixels from a very dark or very bright object in the image. For example, along a row of trees, the plough furrows have almost been suppressed (Figure 4). Similarly, in the river, there are bands of almost homogeneous gray values along its banks. The river is of little concern to us, but the band along tree rows may make it difficult to detect marks near the borders of fields.

We also tried global contrast enhancement, but this was not useful.

4.3 Template matching

The principle in template matching is to have some predefined 'ideal' images that we slide across the image, and for each template and each location, we compute some similarity measurement. The locations with the highest similarity values are regarded as detections.

We used ring shaped templates with radii in the range from 4.5 m to 9.0 m, with 0.5 m increment in radius, giving ten different radii in total. We tried different ring thicknesses, a disk instead of a ring, and also whether the filter boundary should be square or circular.

Each ring filter was convolved with the image, producing a new correlation image, where the value at each pixel indicate how well the ring filter, when centered on that pixel location, agrees with the image. A high positive value indicates a bright ring, and a high negative value a dark ring.

In order to extract ring candidates, a threshold value T is used twice on the correlation image. First, bright rings are identified at regions with correlation $> T$. Next, dark rings are identified at regions with correlation $< -T$. By selecting a high T, few ring candidates will be extracted. This will reduce the number of false detections, but may also reduce the

number of true detections. By selecting a low T, many ring candidates will be detected. This will increase the number of false detections, but may also increase the number of true detections.

The threshold value T is the single most sensitive parameter that the user may adjust, and several values were used in the experiments.

We have experimented with various ring template shapes. A common principle for all these templates is that they have a mean value of zero. Circular template boundaries performed better than square ones, a template radius of twice the ring radius worked best, and disk templates did not appear to offer improvements over ring templates.

The result of the template matching step is a list of ring candidates, each carrying the following information:

- x- and y-coordinates of the ring center
- ring radius r,
- ring type – dark or bright.

4.4 Feature extraction

The purpose of feature extraction is to measure some qualities of each object. The idea is that each object is described by its features. Good features are features that will make the subsequent classification of the objects easy.

To be able to extract feature vectors, subimages of each ring candidate were extracted as follows. For each ring candidate, a $4r \times 4r$ subimage was extracted, centered on the ring candidate's center, where r is the radius found in the template matching step. The corresponding subimage from the local contrast enhanced image was extracted, and two

thresholded versions of it were created. For bright rings, the two thresholds were 0.5 and 1.0, and each pixel with a higher value than the threshold resulted in a white pixel in the corresponding location, otherwise black. For dark rings, the thresholds were -0.5 and -1.0, and each pixel with a lower value than the threshold resulted in a white pixel, otherwise black. In addition, all pixels outside a $2r$ radius from the subimage center were set to black.

This resulted in four subimages for each ring candidate – two gray level subimages and two binary subimages.

Note that thresholding the local contrast enhanced image is equivalent to binarizing the original image with Niblack's method (Niblack, 1986). The method compares each pixel value with a local threshold value $t(x, y)$ to decide whether the pixel belongs to an object or to the background. The threshold value is

$$t(x, y) = \mu(x, y, N) + k\sigma(x, y, N), \tag{2}$$

where $\mu(x, y, N)$ and $\sigma(x, y, N)$ are as defined in Section 4.2.

A number of feature extraction methods have been suggested in the literature (e.g., see (Reiss, 1993; Trier et al., 1996)). For ring marks, we wanted to have features that could describe deviations from the circular shape. In addition, we wanted to have general shape descriptions that captured the shape differences between ring marks and other locations where the ring templates gave a good match, e.g., plow furrows (Figure 5). We have tried the features below on the extracted subimages.

- **Ring cover** – the amount of overlap between a binary subimage and a binary version of the ring filter. This is measured as the number of pixels in the intersection

between the two binary images. This feature indicates how much the binary image resembles a ring.

- **Mean value** of binary image. The mean x- and y-coordinates of the binary pixels, normalized to a value between 0.0 and 1.0. A symmetric image would give 0.5 for both the mean x- and y-coordinates, and a significant deviation from (0.5, 0.5) indicates a deviation from a ring shape.

- **Hu moment invariants** (Maitra, 1979). Geometric moments can be regarded as a way of describing a shape in statistical terms, and the Hu moment invariants are moment combinations that remain constant even if the shape is rotated, moved or scaled.

- **Real weighted Fourier moments** (page 18 in (Reiss, 1993)). These have quite complicated formulae, based on orthogonal functions, which are used to sample the image, and result in moments that are themselves invariant to rotation, translation and scaling.

Ring cover and mean value are measuring directly how well the segmented rings agree with the ring templates. Hu moment invariants and real weighted Fourier moments are general shape descriptions that are invariant to translation, rotation, and scale changes. Ring cover and mean value were only computed from the binary images, whereas the Hu moments and real weighted Fourier moments were computed for both gray level and binary subimages.

For details on the computation of Hu moments and real weighted Fourier moments, see pages 41–43 in (Trier et al., 2008a).

In order to identify the most promising features, scatter plots were made for all features, two features at a time. In the scatter plots, all 'true' rings from the test images were included, along with 'false' rings from one or two test images. If there had been an individual feature that was able to discriminate between rings and 'no-rings', then the two classes would have manifested themselves as two distinct clusters in the corresponding scatter plot. However, the 'best' features only seemed to have a smaller variance for rings than for 'no-rings', and the dark ring and bright ring classes seemed to be contained within the 'no-ring' class (Figure 8). Since the number of rings was much smaller than the number of no-rings, the apparently smaller variance might just be a coincidence.

In principle, the scatter plots could miss a situation where the clusters were apparent only in a multidimensional feature space. Statistical classification techniques (Duda and Hart, 1973), e.g., based on the multivariate Gaussian distribution, can be used to reveal such a situation and build an optimal classifier. However, this requires an extensive training data set, with a number of ring samples orders of magnitude larger than that we had available.

4.5 Classification based on a decision tree

Since we had a very limited test data set, containing only 35 identified rings, we could not train a statistical classifier and get any reliable estimate for a covariance matrix. Instead, a classifier based on simple if-tests was used. For each selected feature, a lower and an upper bound for the acceptable values were set. These values were determined from the scatter plots, thereby training on the test data, giving too optimistic estimates for the classification results. Even these classification results were not too promising. Depending on the threshold value of the correlation in the template matching in the segmentation

step, 10 to 100 times as many false positives as true positives were detected.

4.6 Combinations of the above methods

In order to create the best possible classification performance, we investigated if any of the methods described in Subsections 4.1–4.5 should be left out. If local contrast enhancement was skipped, then none of the true rings were detected in the subsequent template matching, only false detections were made. If the band pass filtering step was applied, then local contrast enhancement still had to be applied after the band pass filtering step.

We next investigated feature extraction and classification. The idea is that template matching could produce a lot of ring candidates, and that the false rings could be removed by a classifier, based on features with high discriminative power. Although the individual scatter plots did not identify any features that were clearly able to separate the true rings from the false rings, some of them seemed to be able to remove a few false rings. For each such feature, we could set an interval which the feature had to be within. The idea was that by including all these features and intervals in a decision tree, different rings would be excluded by different features. Experiments demonstrated that this was indeed the case, but still many false rings remained. It should be noted that the intervals were determined from the test set, thereby training on the test set, so the less than promising results are in fact too optimistic.

Based on this, we concluded that classification based on a decision tree should be omitted. Consequently, feature extraction is not necessary.

4.7 The best ring mark detection approach

The algorithm can be summarized as follows (Figure 6).

1. Define masks of agricultural fields, either by drawing regions of interest, or by importing a vector file.

2. Perform image processing on subimage, as follows.

   (a) Optionally, do band pass filtering in the frequency domain, as follows.

      i. Apply the fast Fourier transform on the image.

      ii. Apply a band pass filter with cut-off radii 100 and 800, to remove low and high frequencies that do not contribute to rings.

      iii. Apply inverse fast Fourier transform.

   (b) Apply local contrast enhancement.

3. Search for rings, as follows.

   (a) Construct ring templates of increasing sizes.

   (b) Convolve the image with a ring template.

   (c) Threshold the convolution result to find bright rings.

   (d) Threshold the convolution result to find dark rings.

   (e) Remove detections outside masks.

   (f) Merge a new ring with an existing one if they are less than 5 pixels apart, keeping the values of the ring with the higher correlation value.

   (g) Repeat b – f for all ring template sizes.

4. Validate ring detection by letting the operator delete false detections (Figure 7).

5 Experimental results

The algorithm for ring mark detection, as described above in Section 4, was applied to the entire data set, that is, a collection of 4096×4096 subimages, which together covered all the identified true rings. Three parameters were varied:

- The correlation threshold
- Whether band pass filtering in the frequency domain was used or not.
- Whether a normal ring or a thin ring template was used in template matching.

The number of detected false rings varies dramatically with the correlation threshold (Table 1). If band pass filtering in the frequency domain is used ('band pass' = 'yes' in Table 1), then the correlation threshold should be around 0.04 higher to obtain similar results as without band pass filtering. With this adjustment, the number of false detections is of the same order of magnitude regardless of whether band pass filtering is used or not. When the number of detected true rings is similar, we observe that a few rings were detected with band pass filtering, but not without, and also *vice versa*.

If a template with a thinner ring was used, more false detections were made.

A reasonable compromise between not detecting too many false rings and at the same time detecting as many true rings as possible, might be when the number of false detections is approximately seven times the number of true detections. In this case, 11 out of 15, or 73%, of the strong rings were detected, and 5 out of 10, or 50%, of the fairly strong rings were detected. This is 16 out of 25 of the strong and fairly strong rings, or 64%.

The number of false positives can be reduced, at the cost of reducing the number of true positives as well. For example, by reducing the number of false positives from seven

times to less than half the number of true positives, the number of detected strong and fair rings decreased from 64% to 32%. On the other hand, even if the correlation threshold is set so low that almost 30 times as many false rings as true rings are detected, many of the strong and fairly strong rings are not detected. Furthermore, none of the weak rings are detected.

6 Discussion

The experiments demonstrate that the proposed algorithm is able to detect many circular patterns. Still, many are also missed by the algorithm, and many false detections are made. If the goal is to detect each and every circular pattern, then the algorithm needs to be improved to be really useful.

For a thorough search in a limited area, a high number of false positives might be acceptable. On the other hand, for a massive search through a large number of images, the number of false positives might be kept at a minimum, as long as some sites are detected. Some circular patterns may only be visible from time to time. In order to find these, one may have to process images from, say, a ten year period, and, say, 5-10 images per year. In this perspective, our approach can be used to process large volumes of satellite images that would otherwise not be inspected, thus detecting many new sites.

In the experiments, we have only used two satellite images, containing 35 identified ring marks in total. Many more satellite images and identified ring marks are needed to evaluate the current version of the system, spot weaknesses, and experiment with possible enhancements.

Regarding ring marks, one may argue that it is rather disappointing that feature

extraction followed by a decision tree classifier was not able to separate the false rings from the true rings. On the other hand, we are trying to recognize a rather simple geometrical shape, which should be well suited for template matching. The better the template matching result, the less there is to gain from improving feature extraction and classification.

Another issue is what information one can hope to extract from rings. Since the ring shape is used in template matching, one could argue that this shape information is already used, so features describing the ring shape itself would not represent any new information. The features we used were (1) measuring the deviations from the circular shape, and (2) describing the shape. In both cases, the true and false rings appeared to be drawn from the same population. This may be caused by unsuccessful segmentation. In the binary images (Figure 5), the true rings often have spurious, small blobs attached or very close. These are probably adding noise to the feature values. Another kind of feature that could be used to measure deviations from the circular shape is deformable templates (e.g., Yuille et al., 1989; Jain et al., 1996).

At the moment, statistical classification is not being used for detecting circular marks, mainly because we lack a sufficiently large training set. If a large training set could be produced, then we could investigate if a statistical classifier could obtain acceptable recognition rates. However, classification performance is highly dependent on whether features with high discriminatory power have been extracted from the data. One approach we have not tried is to include multispectral information in the features, either from the four individual bands, the normalized difference vegetation index (NDVI), or a pan-sharpened image. If a large training set can not be produced, we could still attempt to use a decision tree on the new features extracted from the multispectral information.

At present, only one parameter is varied for the different ring templates, namely the radius. One could also vary the thickness of the ring, and see if that enables us to use a high correlation threshold, thus eliminating many of the false detections while at the same time detecting more true rings.

The main problem when detecting general cultural heritage sites, or amorphous marks, is their wide range of shapes and sizes. This makes it very hard to select features with high discriminatory power. Further, a very limited number of identified sites makes it difficult to both train the system and test its classification performance. Grøn *et al.* (2004) attempted to use in situ chemical profiling of potential sites. If successful, this could be a way of providing ground truth, but no reliable relationship has yet been established.

7 Conclusion

We have presented a method for detecting circular patterns in the panchromatic band of Quickbird images of agricultural land. These circular patterns are potentially the remains of burial mounds or other circular archaeological sites. The method is based on image processing to enhance the appearance of low contrast rings, followed by template matching. The method has been tested on images from two different areas in Norway, containing a total of 35 manually identified circular patterns.

The system is currently deployed at three user sites, and feedback from the users on performance, precision, and usability will provide important information for the further development of the methods and an efficient graphical user interface. As the interface and underlying software matures we expect to deploy it to a larger number of user sites.

Archaeologists state that the software tool will be helpful for locating potential cultural

heritage sites. Although it makes many false detections, it will relieve the operators from time-consuming manual inspection of entire images.

Acknowledgment

We thank Lars Gustavsen, of the Museum of Cultural History at the University of Oslo, and Christier Tonning, of Vestfold County Administration, for relevant information on the test data and useful comments; our colleagues Jostein Amlien, Line Eikvil, Marit Holden, Ragnar Bang Huseby, and Hans Koren for fruitful discussions; and the anonymous referees for constructive criticism, improving the paper.

This work was funded by The Norwegian Directorate for Cultural Heritage and The Norwegian Space Centre.

References

Aminzadeh B and Samani F. 2006. Identifying the boundaries of the historical site of persepolis using remote sensing. *Remote Sensing of Environment*, 102(1–2):52–62. DOI:10.1016/j.rse.2006.01.018.

Aurdal L, Eikvil L, Koren H, Loska A. 2006. Semi-automatic search for cultural heritage sites in satellite images. In *Proceedings of 'From Space to Place', 2nd International Conference on Remote Sensing in Archaeology*, Rome, Italy, Dec. 4–7. BAR International Series 1568; 1–6.

Blom R, Crippen R, Elachi C, Zarins J. 1997. Space technology and the discovery of the lost city of Ubar. In *Proceedings of the IEEE Aerospace Conference*, volume 1, Feb. 1–8; 19–28. DOI: 10.1109/AERO.1997.574258.

Blom RG, Chapman B, Podest E, Murowchick R. 2000. Applications of remote sensing to archaeological studies of early Shang civilization in northern China. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2000)*, volume 6, July 24–28; 2483–2485. DOI: 10.1109/IGARSS.2000.859614

Campana S, Frankovich R, Corsi M, Pericci F. 2006. Aerial survey project: seven years of flights over Tuscany. In *Proceedings of 'From Space to Place', 2nd International Conference on Remote Sensing in Archaeology*, Rome, Italy, Dec. 4–7. BAR International Series 1568; 497–503.

Carter JC, Crawford M, Lehman P, Nikolaenko G, Trelogan J. 2000. The Chora of Chersonesos in Crimea, Ukraine. *American Journal of Archaeology*, 104(4):707–741.

Cracknell AP, Hayes L. 2007. *Introduction to remote sensing.* CRC Press, Boca Raton, Florida, USA, second edition.

De Laet V, Paulissen E, Waelkens M. 2007. Methods for the extraction of archaeological features from very high-resolution Ikonos-2 remote sensing imagery, Hisar (southwest Turkey). *Journal of Archaeological Science*, 34(5):830–841. DOI:10.1016/j.jas.2006.09.013.

Devereux BJ, Amable GS, Crow P. 2008. Visualisation of LiDAR terrain models for archaeological feature detection. *Antiquity*, 82(316):470–479.

Duda RO, Hart PE. 1973. *Pattern Classification and Scene Analysis.* John Wiley & Sons, New York.

Goossens R, De Wulf A, Bourgeois J, Gheyle W, Willems T. 2006. Satellite imagery and

archaeology: the example of Corona in the Altai mountains. *Journal of Archaeological Science*, 33(6):745–755. DOI: 10.1016/j.jas.2005.10.010.

Grøn O, Aurdal L, Christensen F, Tømmervik H, Loska A. 2004. Locating invisible cultural heritage sites in agricultural fields: Evaluation of methods for satellite monitoring of cultural heritage sites – results 2003. Technical Report ISBN: 82-7574-033-9, The Norwegian Institute for Cultural Heritage Research. http://www.riksantikvaren.no/filestore/satellite-Rap.pdf.

Fowler MJF, Fowler YM. 2005 Detection of archaeological crop marks on declassified Corona KH-4B intelligence satellite photography of southern England. *Archaeological Prospection*, 12(4):257–264. DOI: 10.1002/arp.266.

Jain AK, Yu Zhong, Lakshmanan S. 1996. Object matching using deformable templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(3):267–278. DOI: 10.1109/34.485555

Larsen SØ, Amlien J, Koren H, Solberg R. 2008a. Mapping road traffic conditions using high resolution satellite images. In *Proceedings of GEOBIA 2008 – Pixels, Objects, Intelligence: Geographic Object-Based Image Analysis for the 21st Century*, Calgary, Alberta, Canada, Aug. 5–8; 93–98.

Larsen SØ, Trier ØD, Solberg R. 2008b. Detection of ring shaped structures in agricultural land using high-resolution satellite images. In *Proceedings of GEOBIA 2008 – Pixels, Objects, Intelligence: Geographic Object-Based Image Analysis for the 21st Century*, Calgary, Alberta, Canada, Aug. 5–8; 48–53.

Lasaponara R, Masini, N. 2007. Detection of archaeological crop marks by using satel-

lite quickbird multispectral imagery. *Journal of Archaeological Science*, 34(2):214–221. DOI: 10.1016/j.jas.2006.04.014.

Lemmens JPMM, Stančič Z, Verwaal RG. 1993. Automated archaeological feature extraction from digital aerial photographs. In Andresen J, Madsen T, Scollar I, editors, *Proceedings of the 20th CAA conference 'Computing the Past: Computer Applications and Quantitative Methods in Archaeology'*, Aarhus University Press; 45–52.

Niblack W. 1986. *An Introduction to Digital Image Processing*, Prentice Hall; 115–116.

Maitra S. 1979. Moment invariants. *Proceedings of the IEEE*, 67(4):697–699.

Musson C, Driver T, Pert T. 2006. Air photo applications in Wales, UK. Exploration, landscape analysis, conservation and public presentation. In *Proceedings of 'From Space to Place', 2nd International Conference on Remote Sensing in Archaeology*, Rome, Italy, Dec. 4–7. BAR International Series 1568; 55–60.

Oostdijk A, van Persie M, Noorddbergen HHS, van Rijn JM. 2008. Multi scale object based detection and classification of roads and vehicles in high resolution optical satellite imagery. In *Proceedings of GEOBIA 2008 – Pixels, Objects, Intelligence: Geographic Object-Based Image Analysis for the 21st Century*, Calgary, Alberta, Canada, August; 356–361.

Reiss TH. 1993. *Recognizing planar objects using invariant image features*, volume 676 of *Lecture Notes in Computer Science*. Springer-Verlag.

Risbøl O, Gjertsen AK, and Skare K. 2006. Airborne laser scanning of cultural remains in forests: some preliminary results from a Norwegian project. In *Proceedings of*

'From Space to Place', 2nd International Conference on Remote Sensing in Archaeology, Rome, Italy, Dec. 4–7. BAR International Series 1568; 107–112.

Sittler B, Schellberg S. 2006. The potential of LIDAR in assessing elements of cultural heritage hidden under woodland canopies. Possibilities and limits in detecting microrelief structures for archaeological surveys. In *Proceedings of 'From Space to Place', 2nd International Conference on Remote Sensing in Archaeology*, pages 117–122, Rome, Italy, Dec. 4–7. BAR International Series 1568; 117–122.

Sumer E, Turker M. 2008. Building detection from high-resolution satellite imagery using adaptive fuzzy-genetic approach. In *Proceedings of GEOBIA 2008 – Pixels, Objects, Intelligence: Geographic Object-Based Image Analysis for the 21st Century*, Calgary, Alberta, Canada, Aug. 5–8; 87–92.

Trelogan J, Crawford M, Teng L, Kwon O, Carter J. 1999. Mapping the features of the Chora of Chersonesos via remotely sensed data. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '99)*, volume 5, June 28–July 2; 2569–2571. DOI: 10.1109/IGARSS.1999.771579.

Trier ØD, Jain AK, Taxt T. 1996. Feature extraction methods for character recognition – a survey. *Pattern Recognition*, 29(4):641–662. DOI: 10.1016/0031-3203(95)00118-2.

Trier ØD, Larsen SØ, Solberg R. 2008a. Detection of circular patterns in high-resolution satellite images of agricultural land with CultSearcher. Note SAMBA/16/08, Norwegian Computing Center. http://publ.nr.no/.

Trier ØD, Loska A, Larsen SØ, Solberg R. 2008b. Detection of burial mounds in high-resolution satellite images of agricultural land. In *Proceedings of the First International*

Workshop on Advances in Remote Sensing for Archaeology and Cultural Heritage Management, Rome, Italy, Sep. 30–Oct. 4.

Wilson DR. 1982. *Air Photo Interpretation for Archaeologists.* St. Martin's Press, New York.

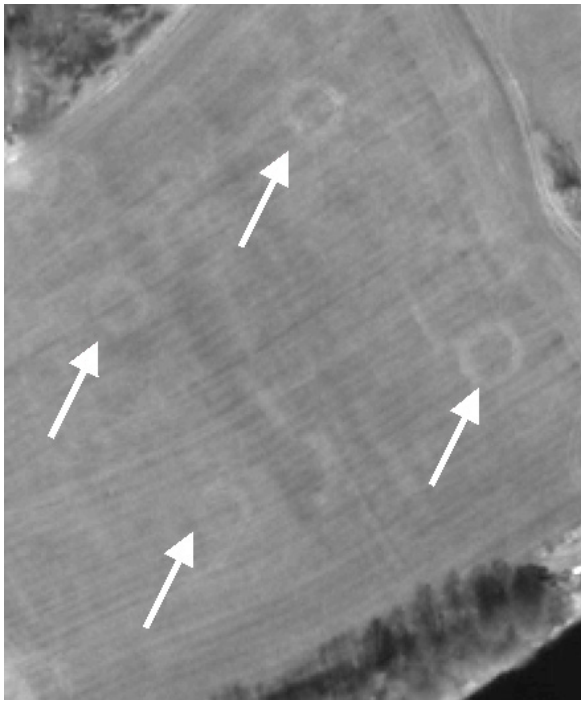
Yuille AL, Cohen DS, Hallinan PW. 1989. Feature extraction from faces using deformable templates. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '89)*, June 4–8; 104–109. DOI: 10.1109/CVPR.1989.37836.

Table 1: Ring mark detection results

band pass	corr. thres.	strong rings	fair rings	weak rings	true rings	false rings
no	0.30	11	5	0	16	450
no	0.33	11	5	0	16	109
no	0.35	10	2	0	12	39
no	0.40	8	0	0	8	3
yes	0.35	12	3	0	15	174
yes	0.38	11	2	0	13	48
yes	0.39	10	2	0	12	31
yes	0.40	9	1	0	10	12
ground truth		15	10	10	35	

(a) (b)

Figure 1: The Quickbird images. (a) “Laagen”, (b) “Gardermoen”

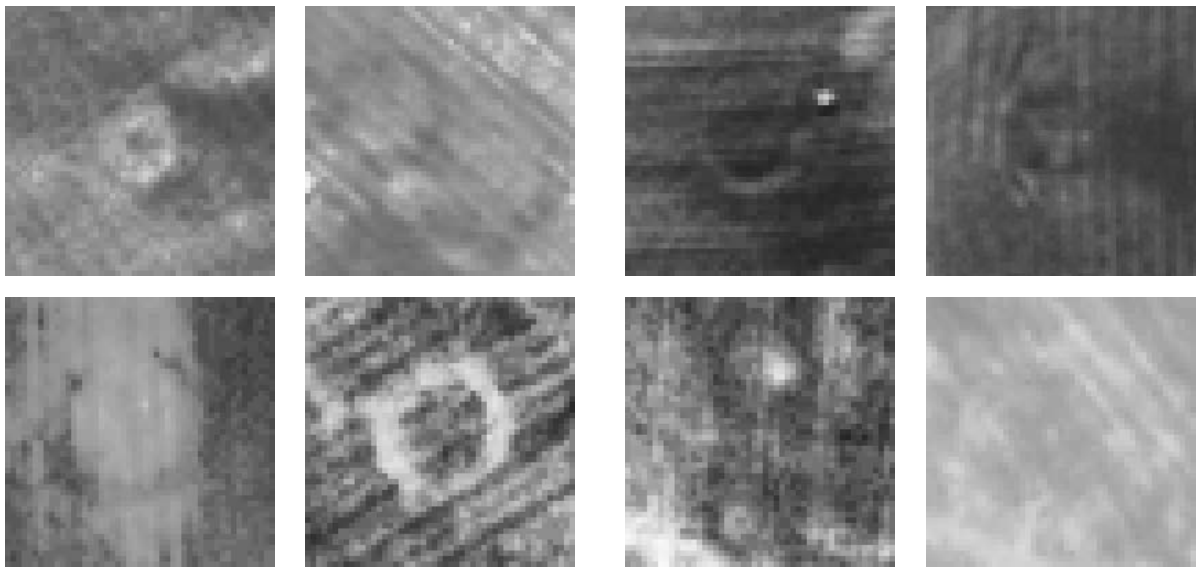


(a)

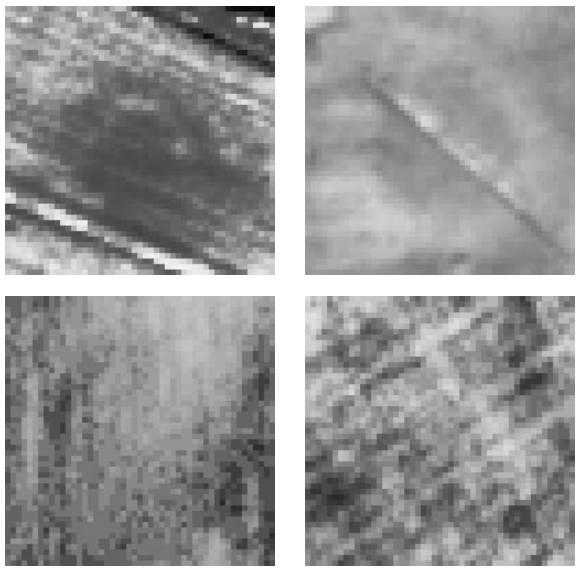
(b)

(c)

Figure 2: Detail of the image 'Laagen'. (a) The panchromatic band, with four rings pointed out; (b) the red, green and blue bands; and (c) the near-infrared band.



(a) (b)



(c)

Figure 3: Example ring marks. The contrast has been adjusted in each case to highlight the rings. (a) strong rings, (b) fairly strong rings, (c) weak rings.

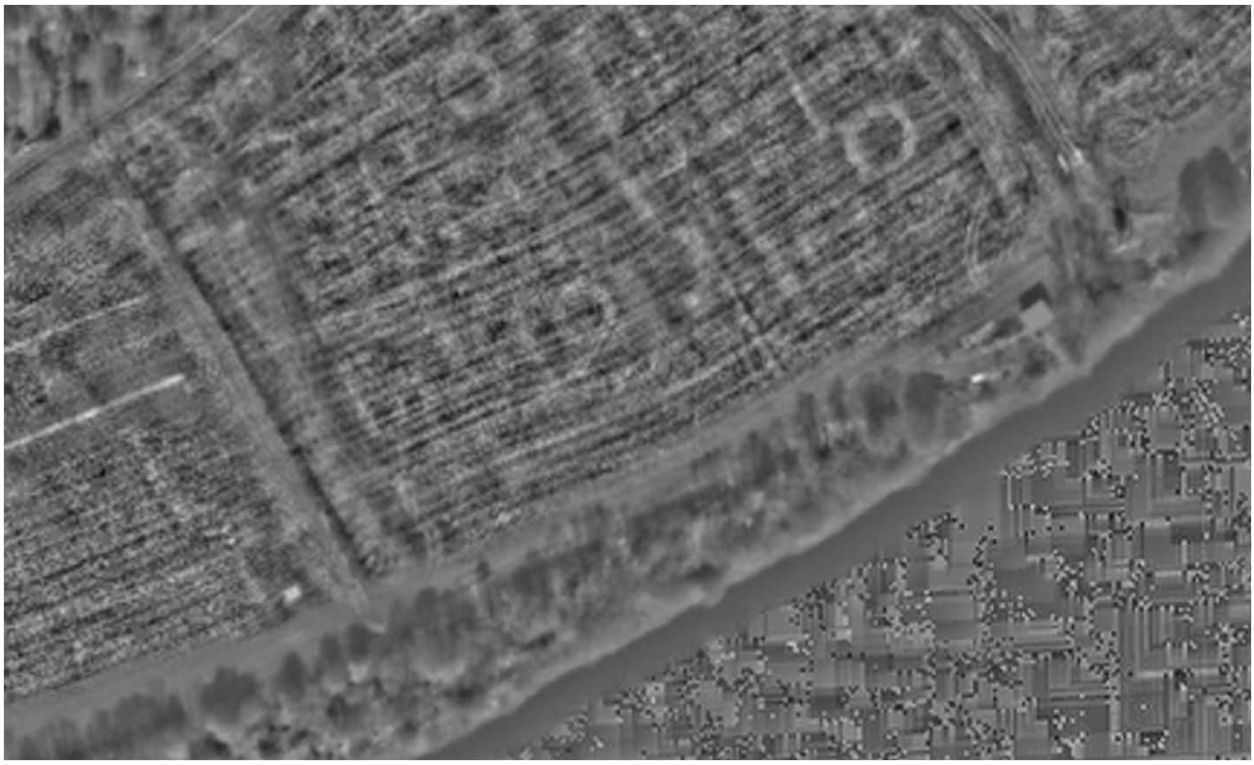
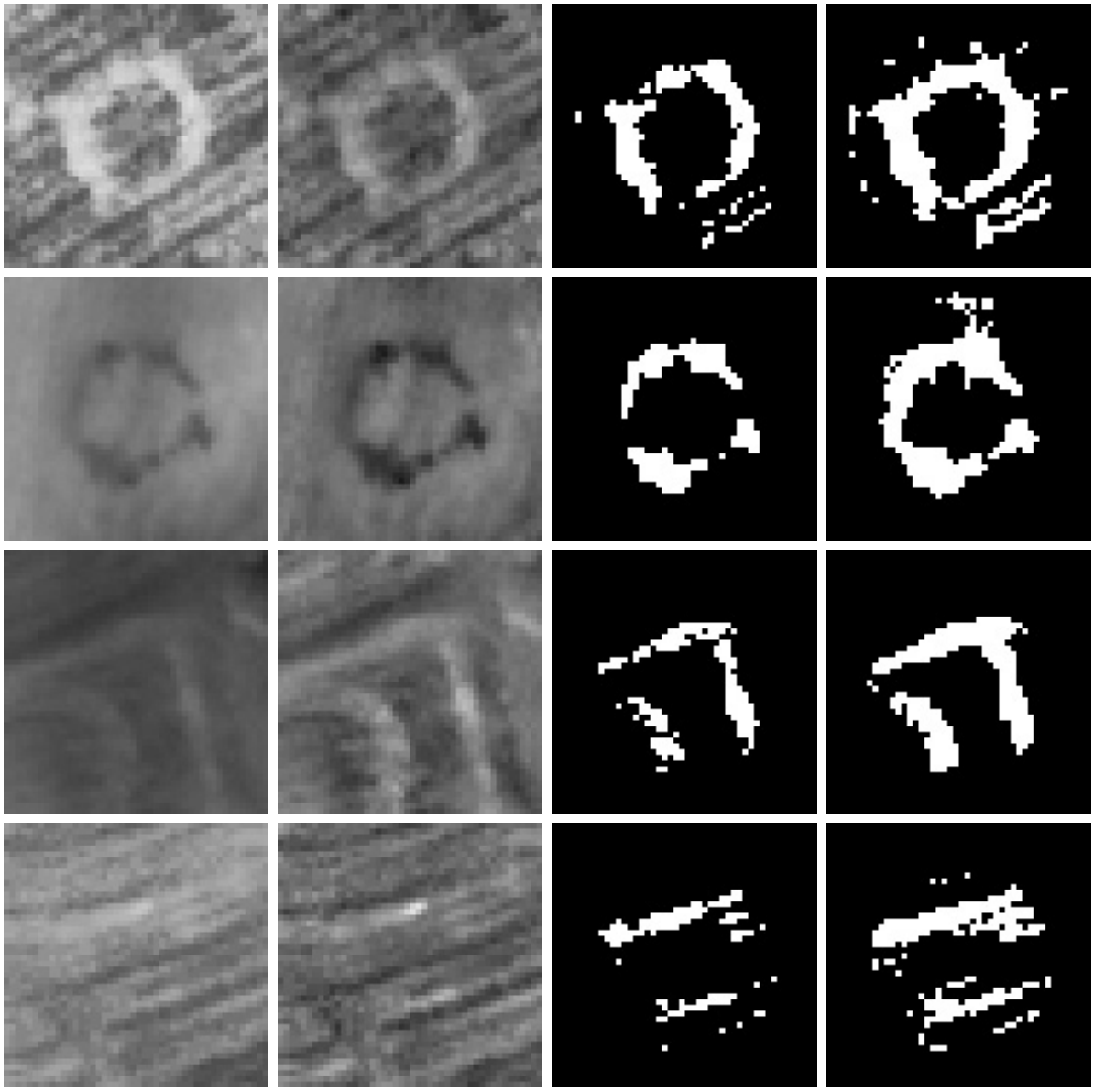


Figure 4: Result of local contrast enhancement, with $N = 21$.



(a) (b) (c) (d)

Figure 5: Subimages used in feature extraction. Top two rows: two of the true rings. Bottom two rows: two of the false detections. (a) Original; (b) after local contrast enhancement; binarized with Niblack's method, with (c) $|k| = 0.5$, and with (d) $|k| = 1.0$.

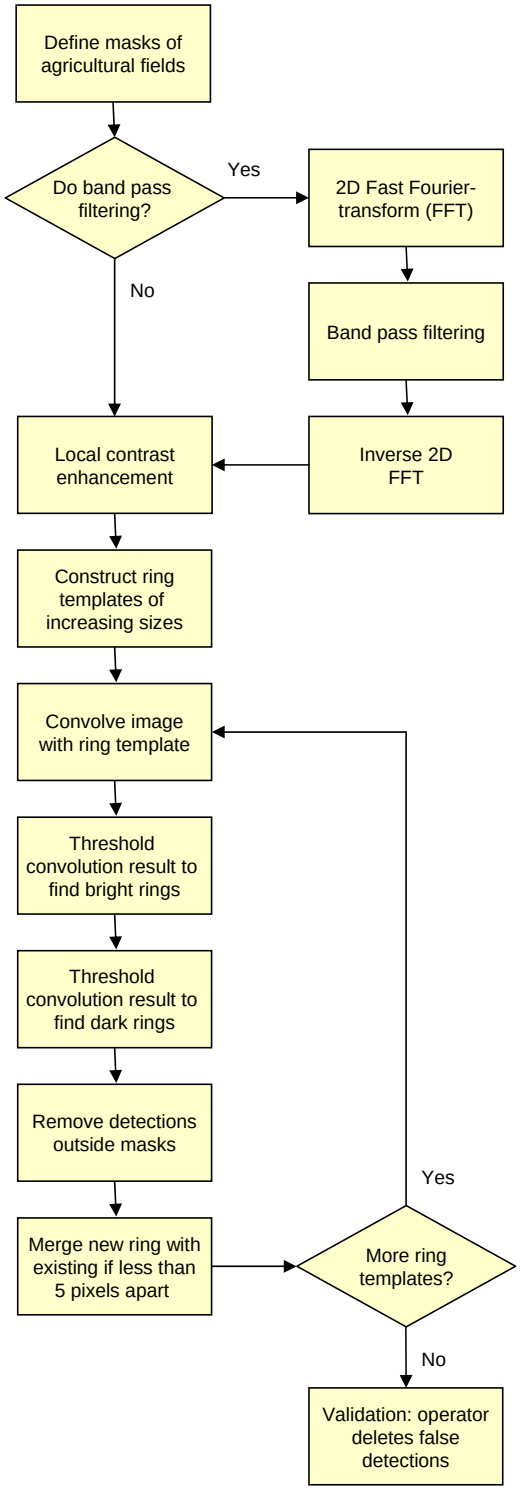
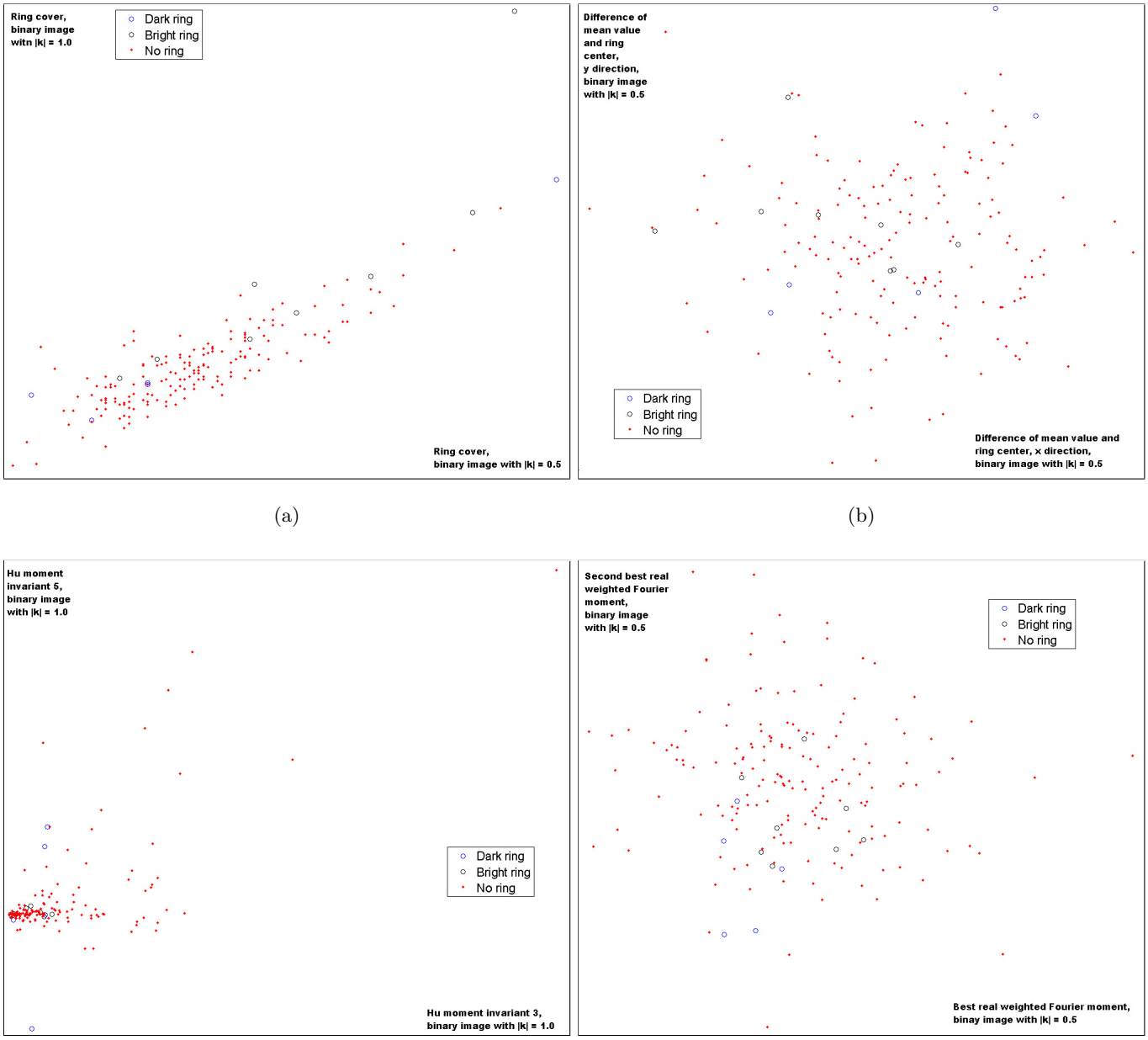


Figure 6: Flowchart of the algorithm for ring mark detection.

(a) (b)

Figure 7: Ring validation. (a) All detected rings, with correlation threshold $T = 0.325$; (b) validated rings.



(a) (b)

(c) (d)

Figure 8: Scatterplots of selcted features extracted from the binary subimages of true and false rings. (a) Ring cover, binarized with Niblack's method and $|k| = 0.5$ versus $|k| = 1.0$; (b) Difference between mean value and ring center, in x direction versus y direction, both with $|k| = 0.5$; (c) the two best Hu moment invariant features, invariant 3 versus invariant 5, both with $|k| = 1.0$, and (d) the two best real weighted Fourier moments, both with $|k| = 0.5$.