DETECTION OF RING SHAPED STRUCTURES IN AGRICULTURAL LAND USING HIGH RESOLUTION SATELLITE IMAGES

S. Ø. Larsen [a, *], Ø. D. Trier[a], R. Solberg[a]

[a] Norwegian Computing Center, P.O. Box 114 Blindern, NO-0314 Oslo, Norway
- siri.oyen.larsen@nr.no, +47 2285 2615

KEY WORDS: Ring graves, Quickbird, template matching, contrast enhancement, archaeology, remote sensing

ABSTRACT:

The use of satellite images in archaeology surveys has become popular the recent years. Ancient remains, with no above-ground structure, often leave crop or soil marks, which under certain ground conditions may be detected with a bird's view. There is often increasing pressure on areas associated with cultural heritage. Advanced detection and mapping of such sites has the potential of saving and protecting much more of our cultural heritage. This paper reports on experiments and results from a Norwegian project for development of remote sensing methodology, including pattern recognition algorithms, to be applied in a software tool assisting archaeologists to search for interesting sites.

We have developed approaches for detection and classification of ring-shaped structures in high resolution satellite images. The ring structures are in many cases ring graves. The ring structures appear in numerous different disguises with varying diameter, thickness and contrast. In several cases the surrounding area contains image noise, the ring is blurred, or only fragments of a circle are left to leave a mark that is visible from space. The main idea behind the approach is to search the image for locations where the correlation with predefined templates representing rings, is high. Various types of filters have been tested. We then attempted object-wise feature extraction followed by decision-tree classification, in order to rule out as many false positives as possible. The features tested are Hu-moment invariants, real weighted Fourier moments, projection means, and an estimate of how complete the ring is. This resulted in too many false positives, and better results were obtained by accepting fewer rings in the template matching step.

The approach has been tested on Quickbird images from south-east Norway. The test areas are rich in known cultural heritage sites and are also expected to contain a large number of unknown sites. The search was confined to agricultural land defined by land-cover maps. Our results indicate that automated processing of satellite images for detection of candidate cultural heritage sites, in particular ring-grave sites, is feasible. Of the rings that were clearly visible in the images, 73% was detected, and of the ones that were fairly visible, 50% was detected. In addition, seven times as many false negatives as false positives were detected.

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.) exerts growing pressure on cultural heritage in the landscape.

In recognition of this, 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 2002 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*.

CultSearcher is currently analyzing soil-marked and crop-marked patterns. Soil-marked sites are typically the remains of a ditch or pit, buried walls, etc. A ditch or a pit would disturb the local soil profile, and refilled material usually has different characteristics, like density and composition. The refilled material is in most cases not so compact, and it might contain more humus components, making it look darker. The refilled material may also affect the soil texture with a grain-size distribution that differs from the undisturbed soil (usually larger number of smaller grain sizes). This results in improved water-storage capacity, so the soil would look darker under certain conditions.

Crop marks are an indirect effect of buried archaeological features. Their visibility depends on the soil, climate and vegetation. So-called positive marks are due to more available water, which makes plants grow higher and ripen later than the plants around. A color-tonal contrast may be created because the vegetation stays green for a longer period and/or that the vegetation is darker green. Crop marks may also be due to a vegetation relief. Plants grow higher, enough to throw a shadow in slanting sunlight. So-called negative marks appear when plants grow over buried stones (e.g. walls) and run out of water sooner, ripen earlier and stay shorter. Almost any crop can develop marks, if conditions are well. Cereals react fast on a Soil Moisture Deficit (SMD) and are growing very close, making contrasts clearer.

* Corresponding author.

Figure 1. The Quickbird images. Left: “Laagen”, right: “Gardermoen”.

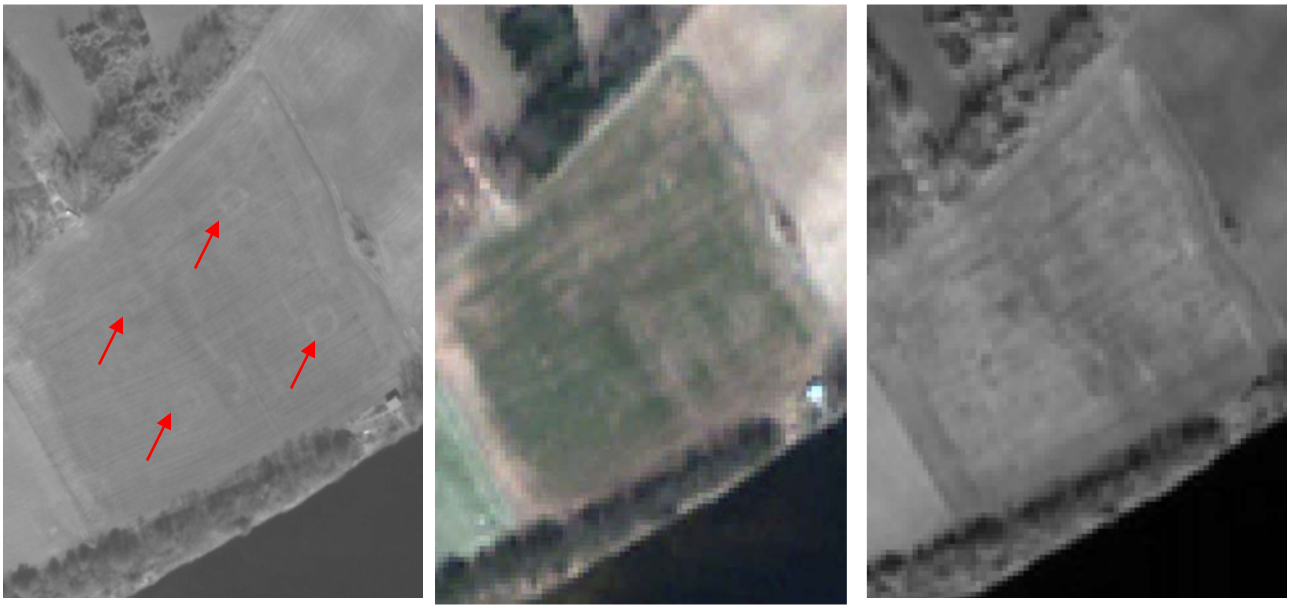
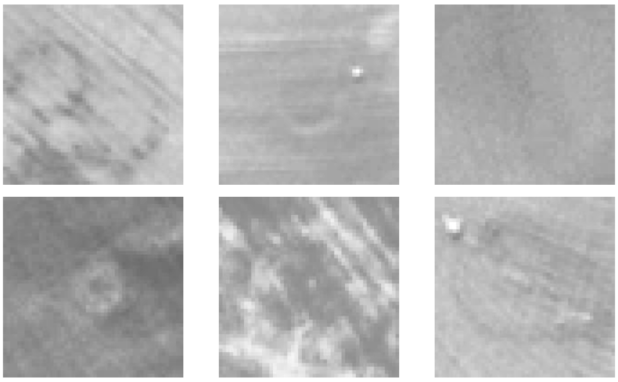


Figure 2. Detail of the image “Laagen”. Left: the panchromatic band, with four rings pointed out; middle: The red, green and blue bands; and right: the infrared band.

Various types of remote sensing sensors, airborne and spaceborne, are useful for detecting remains or patterns due to cultural heritage sites. Soil- and crop-marked sites can be measured with high-resolution optical (visible and infrared) sensors. With the optimal selection of observation wavelengths, high contrast can be obtained (in particular appearing from reflectance contrasts due to soil moisture or vegetation density). The spatial resolution of these sensors should be of 1 m or better to be really useful. Therefore, the project has so far focused on images from Ikonos and QuickBird.

The aim of the software prototype described in this report is to provide computerized assistance to the operator in the analysis of satellite images. In particular, the software identifies circular structures, i.e., potential sites for ring graves, for further inspection by an archaeologist. This means that the archaeologist may concentrate on analyzing the identified sites rather than the entire image.



(a) (b) (c)

Figure 3. Example rings: (a) strong, (b) fair, and (c) weak.

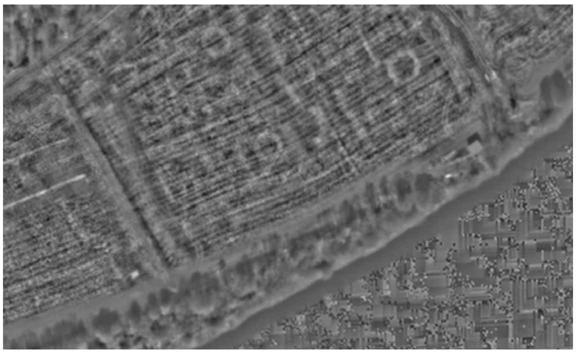


Figure 4. Result of local contrast enhancement.

The rest of the paper is organized as follows. Section 2 describes the experimental data set. Section 3 describes the various methods we have attempted to use, and summarizes the chosen algorithm. Experimental results are described in Section 4, and discussed in Section 5. Finally, conclusions are given in Section 6.

2. EXPERIMENTAL DATA

The data set consists of two Quickbird images (
Figure 1). Image "Laagen" was taken on April 27, 2005 at 10:45AM, from the valley Lågendalen between Kongsberg and Larvik. Image "Gardermoen" was taken 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 multi-spectral image and a panchromatic (gray scale) image. The panchromatic image has 0.6 m wide pixels, and the single band covers the 450-900 nm wavelengths. The multi-spectral image has 2.4 m wide pixels, and the bands are: blue (450-520 nm), green (520-600 nm), red (630-690 nm) and near-infrared (760-900 nm).

Many circular patterns are clearly visible in the panchromatic images, but can hardly be seen in the multi-spectral images (Figure 2). Recently, other research groups have used multi-spectral Quickbird (Lasaponara, 2007) or Ikonos (De Laet, 2007) images, but the objects they were looking for were much larger than the circular patterns in the present work. Since the circular patterns are difficult to spot visually in the multi-spectral images, we chose to use only the panchromatic images.

In the two images, archaeologists have identified 35 rings 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). 11 subimages of 4096×4096 pixels were extracted for the experiments. These subimages included all 35 rings.

3. METHODS

In order to detect as many rings 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. Local contrast enhancement
2. Template matching
3. Feature extraction
4. Decision tree-based classification

Each of these will be discussed below.

3.1 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 $N \times N$ neighborhood centered on the pixel. The pixel value $p_{CE}(x,y)$ in the contrast enhanced image is computed as

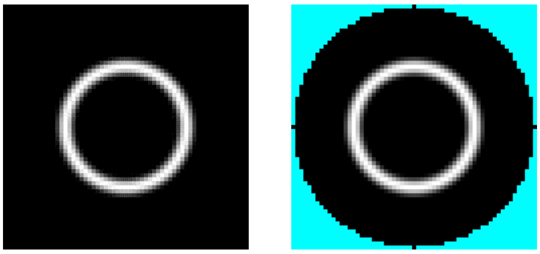
$$p_{CE}(x,y) = \frac{p(x,y) - \mu(x,y,N)}{\sigma(x,y,N)} \tag{1}$$

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 $N \times N$ neighborhood centered on (x,y), and $\sigma(x,y,N)$ is the standard deviation of the gray level values in the same neighborhood.

The choice of the neighborhood size, N, doesn't 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 too much 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. One undesirable effect that is present when using N=21 is that the local contrast is suppressed when the central 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 is a band of almost homogeneous gray values along its bank. The river is of little concern to us, but the band along tree rows may make it difficult to detect rings near the borders of fields.

3.2 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.



(a) (b)

Figure 5. Different template boundaries: (a) square, (b) circular. The cyan pixels are outside the template boundary.

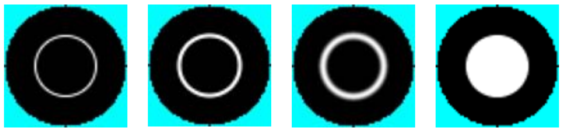
We used ring-shaped templates with radii in the range from 4.5 m to 9.0 m, with 0.5 m increment in radius, giving ten different radii in total. We tried different ring thicknesses, a disk instead of a ring, and also whether the filter boundary should be a square or a circle.

Technically speaking, each ring filter was convolved with the image, producing a new correlation image, where the value at each pixel indicated how well the ring filter, when centered on that pixel location, agreed with the image. A high positive value then indicated 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 extracted. 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. What value to use depends on the following factors:

- If good features can be extracted, resulting in good classification performance, a fairly low value of *T* may be used, since most of the false detections will be removed in the classification step.
- If, on the other hand, the classification performance is poor, then a higher value of *T* might be necessary.
- If the goal is to locate some rings in a large number of images, then a high *T* value is desired.
- If the goal is to locate as many true rings as possible in a limited number of images, and thorough manual inspection is acceptable, then a low value of *T* is desired.
- If the number of false detections is too high, then the usefulness of the automatic recognition is questionable, since it would be just as time-consuming to inspect a large number of ring candidates as to inspect the images manually without the aid of CultSearcher.



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

Figure 6. Different ring shapes: (a) one pixel wide ring, (b) two pixels wide ring, (c) ring from Envi function, (d) disk.

We have experimented with various ring template shapes. A common principle for all these templates is that they have a mean value of zero.

We have used the Envi image software (http://www.ittvis.com/envi/), which has a function to create band-pass filters for use in the frequency domain. This can be used to create ring templates, for use in the image domain, as well. By setting the two cutoff radii to the same value, a fairly thick ring is obtained (Figure 5a). This template has a square-shaped boundary, meaning that some pixels quite far from the center will contribute, while some pixels at the same distance will not. To eliminate this problem, the template boundary was made circular (Figure 5b), and the remaining pixel values adjusted so that the average value was still zero.

Experiments on some of the test data indicated the following:

- A circular template boundary was much better than a square-shaped boundary.
- Two different circular boundary radii were tested: 1.5 times the ring radius, and 2 times the ring radius. 2 times the ring radius was better.
- Three different ring thicknesses were tested: 1 pixel wide (Figure 6a), 2 pixels wide (Figure 6b), and the ring obtained by using the band pass filter function in Envi (Figure 6c).
- A disk template (Figure 6d) was not better than the best ring template for detecting circular patterns that looked more like a disk than a ring. However, with only three disks in the test data, this is hardly a conclusive result.

3.3 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. It is common to use the term *feature vector* to mean the collection of features that were extracted from one object.

Prior to feature extraction, each segmented ring candidate carries the following information:

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

To be able to extract feature vectors, sub-images of each ring were extracted as follows. For each ring candidate, a $4r \times 4r$ sub-image was extracted, centered on the ring candidate's centre, where *r* is the radius found in the template matching. The corresponding sub-image 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 sub-image centre were set to black.

This resulted in four sub-images for each ring candidate – two gray level sub-images and two binary sub-images

A number of feature extraction methods have been suggested in the literature (e.g., Reiss, 1993; Trier et al, 1995). We have tried out the following on the extracted sub images. The two first were used only on the binary sub images.

- Ring cover, that is, the amount of overlap between a binary sub image and a binary version of the ring filter. This is measured as the number of pixels in the intersection between the two binary images.
- 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.
- Hu-moment invariants – these were computed for both gray-level and binary sub-images. The seven moment invariants of order up to three were used (Maitra, 1979).
- Real weighted Fourier moments (Reiss, 1993, p.18) – as for Hu-moment invariants, these were computed for both gray-level and binary sub-images.

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. Even for the most promising features, they still seemed to be drawn from the same population.

3.4 Classification based on a decision tree

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.

correlation threshold		strong rings	fair rings	weak rings	total rings
0.3	true rings	11	5	0	16
	false rings				450
0.33	true rings	11	5	0	16
	false rings				109
0.35	true rings	10	2	0	12
	false rings				39
0.4	true rings	8	0	0	8
	false rings				3
Identified rings		15	10	10	35

Table 1. Detection results.

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 feature that was used, 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 and, possibly, giving too optimistic estimates for the classification results.

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. The hope 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.

Even these classification results were not too promising. Depending on the threshold value on the correlation in the template matching in the segmentation step, 10 to 100 times as many false positives as true positives were detected. Based on this, we simply decided to skip the decision tree in the current version of CultSearcher, meaning that all feature vectors will be classified as rings.

3.5 Algorithm

The algorithm for ring structure detection can be summarized as follows.

1. Define masks of agricultural fields. This can be done by interactively drawing regions of interest, or by importing a GIS vector file.
2. Apply local contrast enhancement
3. Search for rings
    a. Construct ring templates of increasing sizes
    b. Convolve image with a ring template
    c. Threshold result of b to find bright rings
    d. Threshold result of b to find dark rings
    e. Repeat b-d for all ring template sizes

4. EXPERIMENTAL RESULTS

The algorithm, as described above in Section 3, was applied to the entire data set, that is, a collection of 11 sub-images of sizes 4096×4096 pixels, which together covered all the identified true rings.

The number of detected false rings varies dramatically with the correlation threshold (Table 1). A reasonable compromise between not detecting too many false rings and at the same time detect as many true rings as possible, might be when the number of false detections is approximately seven times the number of true detections (yellow line in Table 1). 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 7 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. Further, none of the weak rings are detected.

5. DISCUSSION

One may argue that it is rather disappointing that feature extraction followed by classification 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 on improving feature extraction and classification. So, in order to gain anything, one probably has to lower the threshold on correlation used in the template matching to give more false positives at that stage. Hopefully, a few more true positives may appear as well. Then, the feature extraction and classification steps must be so good that the final recognition results are improved.

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. Features that could be useful might describe acceptable deviations from the circular shape.

One approach we haven't tried is to include multi-spectral information as features, either from the four individual bands, the normalized difference vegetation index, or a pan-sharpened image.

At the moment, statistical classification is not being used, 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. Perhaps the multi-spectral bands or pan-sharpened images could provide suitable features. If a large training set cannot be produced, we could still attempt to use a decision tree on the new features extracted from the multi-spectral 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.

6. CONCLUSION

We have presented a method for detection of circular patterns in the panchromatic band of Quickbird images of agricultural land. These circular patterns are potentially ring graves 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 experiments demonstrate that the proposed algorithm is able to detect many circular patterns. Still, many are also missed by the algorithm. 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 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 this work, we have only used two satellite images, containing in total 35 identified rings. Many more satellite images and identified rings are needed to evaluate the current version of the system, spot weaknesses and experiment with possible enhancements.

All in all, the proposed computer software is a helpful tool for archaeologists in locating circular patterns in high resolution satellite images of agricultural land.

ACKNOWLEDGEMENTS

We thank Ragnar Bang Huseby, Hans Koren, Marit Holden and Jostein Amlien for fruitful discussions. The work presented in this paper was funded by The Norwegian Directorate for Cultural Heritage and The Norwegian Space Centre.

REFERENCES

De Laet, V., Paulissen, E. and 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*, vol. 34(5), pp. 830-841.

Lasaponara, R. and Masini, N., 2007. Detection of archaeological crop marks by using satellite QuickBird multispectral imagery. *Journal of Archaeological Science*, 34(2), pp. 214-221.

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

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

Trier, Ø. D., Jain, A. K. and Taxt, T., 1996. Feature extraction methods for character recognition – a survey. *Pattern Recognition*, 29(4), pp. 641–662.