SEMI-AUTOMATIC DETECTION OF CULTURAL HERITAGE IN LIDAR DATA

Ø. D. Trier [a] and M. Zortea [a]

[a] Norwegian Computing Center, P. O. Box 114 Blindern, NO-0314 Oslo, Norway – trier@nr.no, maciel@nr.no.

ABSTRACT:

This paper presents new methods for the semi-automatic detection of some kinds of cultural heritage in forested areas in Norway, and reports on a work in progress. Some areas have a large number of old pitfall traps that were used for deer hunting 2000-500 years ago. Other areas have a large number of iron production sites that were in use 1400-700 years ago. These two kinds of cultural heritage manifest themselves as pits in the terrain. We have developed methods for the automatic detection of such pits in lidar data with at least 5 emitted pulses per m^2. We are now extending the automatic detection methods to locate grave mounds, stone fences, and old roads in the lidar data.

Experience from on-going archaeological field work clearly demonstrates the benefits of combining automatic detection methods with visual inspection of the lidar data to achieve a map of possible cultural heritage remains, before the actual field survey. The field work can be performed much more efficiently. Since the archaeological feature candidates have already been geo-referenced and measured in the lidar data, the field work is reduced to accepting or rejecting the candidates. Thus, a much larger number of archaeological features can be mapped per day.
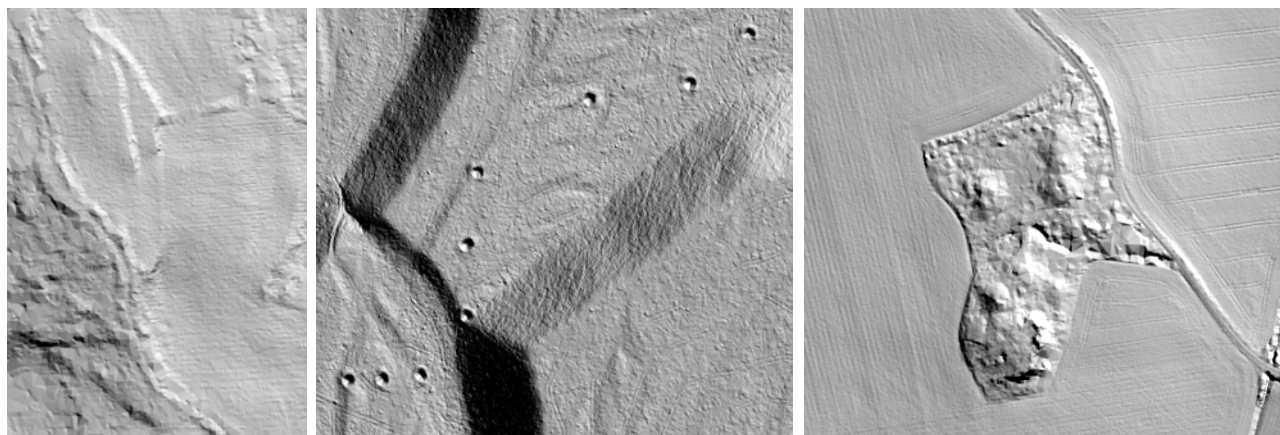
Figure 1. Lidar data from some Norwegian municipalities. Left: Kongsberg, with stone fences. Middle: Nord-Fron, with pitfall traps for moose hunting, which appear as pits. Right: Larvik, with grave mounds, which are seen as heaps in the terrain.

1. INTRODUCTION

Several Norwegian municipalities are experiencing growing pressure on forested land for development, being it new residential areas, new mountain cabins and hotels, or new highways. The traditional mapping of cultural heritage, mainly based on chance discovery and inaccurate positioning, has proven inadequate for land use planning. Therefore, the Norwegian Directorate for Cultural Heritage, in cooperation with some counties and municipalities, are investing in the development of new methods, using new technology, for a more systematic mapping of cultural heritage.

A project was started in 2002 by the Norwegian Directorate for Cultural Heritage, aiming at developing cost-effective methods for surveying and monitoring cultural heritage on a regional and national scale. During the first years, the focus was on the automatic detection of crop marks and soil marks in cereal fields in satellite and aerial images (Aurdal et al., 2006; Trier et

al., 2009). Several of these detections have been confirmed to be levelled grave mounds, dating to 1500-2500 years ago.

However, methods based on optical images are of limited value in forested areas, since the archaeology tends to be obscured by the tree canopies. By using airborne laser scanning data, also called airborne lidar data, and by only keeping the ground returns and not the returns from trees and buildings, the forest vegetation can be removed from the data, and a very detailed digital elevation model (DEM) of the ground surface can be constructed (Devereux et al., 2005). This makes it possible to detect archaeology in a semi-automatic fashion, provided the archaeology manifests itself as features in the digital elevation model of the lidar ground returns, and that these features may be described using some appropriate kind of pattern.

2. DATA AND METHODS

2.1 Airborne lidar height measurements

For an area surrounding the lake Olstappen in Nord-Fron municipality, Oppland County, data was acquired by helicopter, with a minimum of 10 emitted laser pulses per m^2. The data set covered a total area of 29.3 km^2, with 7.3 ground hits per m^2 on average. This terrain is dominated by open pine forest, allowing a large proportion of hits from the ground. This area is known to contain several systems of pitfall traps that were used in moose hunting 500-2000 years ago, and some iron extraction sites with charcoal burning pits dating to 700-1400 years ago. The data set is split geographically in two halves, one western training set, and one eastern test set.
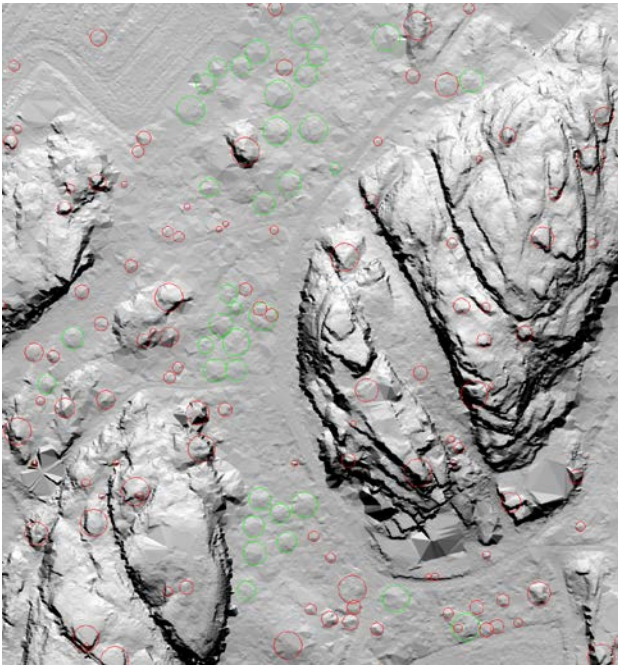

Figure 2. A 210 m × 225 m part of the Kaupang, Larvik training data set for heap detection. True (green) and false (red) grave mounds have been labelled manually.

Larvik municipality in Vestfold County is known to contain a large number of grave mounds in forested areas. From a lidar data set that covers about 150 km^2 of the southern part of Larvik municipality, 12 small portions containing known grave mounds were extracted. Four of these are used as a training set: Kaupang (Figure 2), Store Sandnes, Tanum, and Ødelund. The remaining eight comprise a test set: Berg, Bommestad, Bøkeskogen (Figure 3), Hvatumskjeet, Kjerneberget, Lunde, Valby, and Valbysteinene.

2.2 Automatic detection of circular features

For the detection of circular features, the following general method is applied (Trier and Pilø, 2012):

1. Convert the input LAS files (*LAS Specification*, 2010), containing individual (x, y, z) point measurements to a regular grid of interpolated height measurements, i.e., a digital elevation model (DEM). Only the (x, y, z) points labelled as 'ground' are used to construct the DEM.
2. Convolve the image with templates of varying sizes. Threshold each convolution result to obtain detections.
3. Merge detections that are overlapping, keeping the strongest detections

4. For each detection, compute various attributes that measure the deviation from an ideal model, using different measures than the convolution in step 2
5. Remove detections that have attributes in step 4 outside prescribed intervals.
6. Assign confidence values to the remaining detections.

The above method is applied for each class of archaeological feature. The same templates (Figure 2) may be used to detect pitfall traps, charcoal burning pits, and grave mounds. However, for grave mounds, the range of template sizes is different than for pitfall traps and charcoal pits. Further, a pit template will give negative convolution values for heaps (e.g., grave mounds).
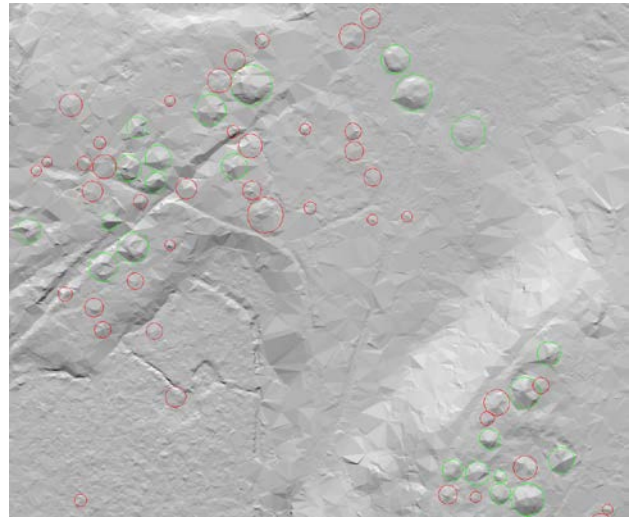

Figure 3. A 245 m × 200 m part of the Bøkeskogen, Larvik test data set for heap detection. True (green) and false (red) grave mounds have been labelled manually.
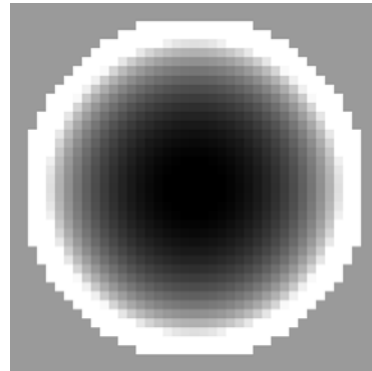
Figure 4. Pit template, shaped as a half-dome circumscribed by a flat ring. White pixels are +1, black pixels are -1, and grey pixels in between. The medium grey pixels outside the white ring edge are exactly zero, thus not contributing to the convolution value. This particular pit template has 3.4 m radius.

2.2.1 Computation of attributes

In step 3 above, the following attributes are computed:

1. Correlation value, obtained from the convolution step.
2. Radius, also obtained from the convolution step.
3. Normalized correlation value, that is, the correlation value divided by the radius.
4. Average pit depth, measured as the height difference between the lowest point inside the pit and the average height on the ring edge outside the pit.

5. Minimum pit depth, measured as the height difference between the lowest point inside the pit and the lowest point on the ring edge.
6. Standard deviation of height values on the ring edge.
7. Root mean square (RMS) deviation from a perfect hemisphere, i.e., a perfect U-shaped pit.
8. RMS deviation from a perfect V-shaped pit.
9. For each pit, a threshold is defined as the value that separates the pixels inside the pit into two groups, the 25% of the pixels that are darker than the threshold, and the 75% that are brighter. Use this threshold to extract a dark blob segment from a square image centred on the pit, with sides equal to six times the radius. This is called the 25%-segment. If this results in a compact, central segment inside the pit, connected to a larger segment outside the pit, with only a few connecting pixels on a ring just outside the pit, then the central segment is separated from the outside segment. From the extracted segment, the following measures are computed:
   a. Offset: distance from pit centre to the segment's centre.
   b. Major axis length, for a definition, see e.g., Prokop and Reeves (1992).
   c. Elongation, defined as major axis divided by radius.
10. Similarly to above, extract the 50%-segment and compute offset, major axis and elongation from that segment as well

2.3 Initial screening

Thresholds are set on some of the attributes to remove detections that are very unlikely to be archaeology, while at the same time keeping all true archaeological features. By sorting a training set of labelled detections on one attribute at a time, one can manually identify attributes that can be thresholded so that all detections labelled as 'true' or 'possible' archaeological feature be kept, keeping several 'unlikely' and 'false' detections as well, but at the same time removing many 'unlikely' and 'false' detections. These thresholds should not be set too tight, to allow for slightly more variation in the attribute values for the 'true' and 'possible' archaeological features than was observed in the training data.

2.4 Statistical classification versus decision tree

For step 6 in the circular feature detection method above, a manually designed decision tree was originally used to assign confidence values 1-6, with 1 meaning 'very low' and 6 meaning 'very high' (Trier and Pilø, 2012). However, this requires that a number of fixed thresholds be set manually, based on training examples. If a large number of training examples are available, an alternative is to use a statistical classifier. We will compare the two approaches below for automatic pit detection in the context of semi-automatic detection of pitfall traps and charcoal burning pits.

2.5 Automatic pit detection method: common steps

The first five steps in the general circular feature detection method are common for both the manually designed decision tree and the statistical classifier approach. These five steps were applied on the Olstappen data set. A number of parameters had to be selected in this process. A DEM grid size of 0.2 m was used to preserve the accuracy of the lidar height measurements,

thus converting the, on average, 7.3 ground hits per m^2 to 25 interpolated height values per m^2. In the convolution step, pit templates corresponding to pit radii from 1.2 to 3.4 m were used, corresponding to the expected pit sizes; each template having 0.2 m larger radius than the next smaller. The initial screening exercise resulted in the following subset of attributes to be used for thresholding as follows:
1. Normalized correlation > 2.0
2. Average pit depth > 0.5 m
3. Minimum pit depth > 0.1 m
4. RMS u-shape < 0.2
5. RMS v-shape < 0.2
6. 25% segment elongation < 4

When applied on the entire Olstappen data set, the initial screening resulted in 2018 detections, which were then labelled manually, resulting in 258 verified archaeological pits. All these were first verified visually by archaeologists. 67 of these were also verified by archaeological field survey. In addition to these 258 confirmed pits, four detected pits were found to be modern by field survey, and 10 archaeological pits were detected visually in the lidar data by archaeologists. Of these 10 pits that were missed by the automatic method, six have been confirmed by field survey. Field survey to verify the remaining 191 + 4 pits is pending additional funding.

The data set was split in two; a training set containing 129 confirmed archaeological pits and 1000 false detections, and a test set containing 128 confirmed pits and 866 false detections. This split was done by listing pit detections ordered alphabetically on tile names, resulting, roughly, in a western training set and an eastern test set.

2.6 Automatic pit detection using manually designed decision tree

Table 1. Thresholds for assigning confidence values for pitfall trap detection.

feature	confidence					
	very low	low	medium	med. high	high	very high*
normalized correlation	≥2	≥2.5	≥2.5	≥3	≥3.5	
minimum depth	≥0.1	≥0.1	≥0.23	≥0.4	≥0.5	≥1.0
average depth	≥0.5	≥0.5	≥0.5	≥0.55	≥0.75	
RMS u-shape	≤0.2	≤0.1	≤0.07	≤0.05	≤0.04	≤0.02
RMS v-shape	≤0.2	≤0.085	≤0.07	≤0.05	≤0.03	≤0.015
25% segment offset	≤40	≤6	≤6	≤6	≤5	
25% segment elongation	≤4	≤2	≤1.5	≤1.3	≤1.2	
assigned tag	1	2	3	4	5	6

Table 2. The effect of running the confidence assignment on the Olstappen training set for pitfall trap detection.

score value	1	2	3	4	5	6		
confidence	very low	low	medium	medium high	high	very high	not detected	sum
pit confirmed in field			2	2	5	16	1	26
modern/other visually								0
pit visually in image		7	27	32	17	21	4	108
not pit visually	329	517	136	15	3			1000
sum	329	524	165	49	25	37	5	1134

By using the thresholds in Table 1 in a decision tree, confidence values were assigned to all the detections in the training set (Table 2). For confidence values from 'very low' to 'high', all the tests have to be fulfilled, and each detection is assigned the best possible confidence according to the rules. For a detection with 'high' confidence, it is upgraded to 'very high' if at least one of the tests for 'very high' are fulfilled. Obviously, by adjusting the thresholds in Table 1, different number of 'true' and 'false' detections will get the various confidence values

(Table 2), and the goal is to achieve a meaningful balance of 'true' and 'false' detections within each confidence level.

2.7 Automatic pit detection using statistical classifier

In this variety of the pit detection method, the decision tree classifier in the confidence assignment step is replaced by a statistical classifier. The following six different classifiers were evaluated (Hastie et al., 2009):
1. Decision tree (CART algorithm)
2. Nearest neighbour
3. Naïve Bayes (assuming independent attributes)
4. Mahalanobis distance
5. Linear discriminant analysis
6. Quadratic discriminant analysis

For each classifier, the best subset of the 13 attributes computed in Section 2.2.1 is determined using the sequential forward attribute selection algorithm (Pudil et al., 1994). The subset of attributes that maximize the 10-fold cross-validation of average accuracy in the training set is retained. The best classifier turned out to be the Mahalanobis distance classifier (Figure 5), with the following seven attributes, in order of importance:
1. Minimum depth
2. RMS V-shape
3. Standard deviation on edge
4. Offset of 25% segment
5. Normalized correlation
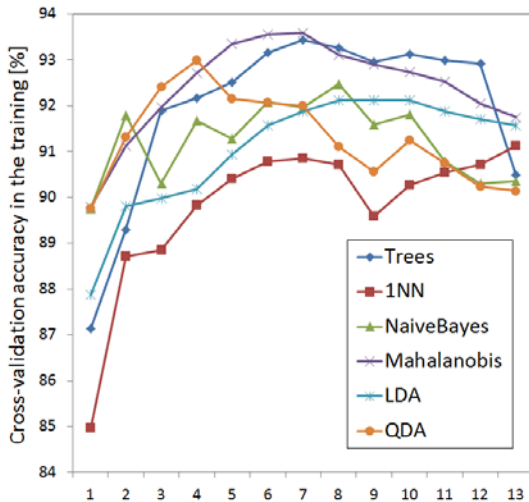6. Average depth
7. Elongation of 25% segment

Figure 5. Performance of the six different classifiers on the Olstappen training set, as a function of the number of attributes.

We will now use the estimated posterior probability, that is, the probability that the detected pit is archaeology, to assign a confidence level to each detection. With six confidence levels, we need to determine five thresholds. As initial threshold values, we use the values corresponding to the 10th percentile, 25th, 50th, 75th and 90th percentile. Then we can count the number of pits and non-pits in each confidence level, multiply with penalty weights (Table 3) and accumulate to obtain a total penalty for the particular choice of thresholds. By adjusting the threshold values, they can be optimized to minimize the total penalty. By doing this on the training set, the thresholds in Table 4 are obtained, which assign 'medium high' or better confidence to most of the true archaeological pits, and 'medium' confidence or lower to most false pits (Table 5).

Table 3. Penalty weights used for optimizing confidence level thresholds.

score value	1	2	3	4	5	6
confidence	very low	low	medium	medium high	high	very high
pit	1024	256	64	16	4	1
non-pit	1	4	16	64	256	1024

Table 4. Optimized threshold values for pit detection.

1	2	3	4	5
0.07374048	0.12168859	0.32615019	0.58907545	0.80731382

Table 5. The result of using the Mahalanobis distance classifier on the Olstappen training set.

score value	1	2	3	4	5	6		
confidence	very low	low	medium	medium high	high	very high	not detected	sum
pit confirmed in field				2	1	22	1	26
pit visually in image		3	7	20	29	45	4	108
not pit visually	27	380	528	62	7			1004
sum	27	383	535	84	37	67	5	1138

2.8 Automatic heap detection using statistical classifier

The method for pit detection can be modified to detect heaps that could be grave mounds. As for pit detection, 0.2 m grid size was used for the DEM. By reversing the sign of the pit templates, heap templates are obtained. Heap templates with radii in the range 1.0-10 m are used, again with each radius being 0.2 m larger than the next smaller. The same attributes as for pits are used, with the obvious exception that average and minimum heap heights are computed instead of the respective pit depths. In addition, two more attributes were computed, by dividing the heap height measurements by the heap radius:
1. Normalized average heap height
2. Normalized minimum heap height

As this is a first attempt, to avoid overlooking true grave mounds, the initial screening uses very relaxed thresholds on a subset of the attributes as follows:
1. Normalized correlation > 1.0
2. Average heap height > 0.2 m
3. Minimum heap height > 0.0 m
4. RMS u-shape < 0.2
5. RMS v-shape < 0.2
6. 25% segment elongation < 5

The result of the initial screening was a training set with 785 heap detections, of which 96 were labelled 'true' and the remaining 'false'; and a test set of 905 heap detections, of which, 96 were labelled 'true' and the remaining labelled 'false'. The labelling was done by a non-archaeologist.

Again, we evaluate six different classifiers and different attribute combinations. The Mahalanobis distance classifier performs best on the Larvik training set, with the following seven features, in order of importance:
1. RMS U-shape
2. Correlation
3. Elongation of 25% segment
4. Offset of 25% segment
5. Standard deviation on edge
6. Major axis of 50% segment
7. Offset of 25% segment

As for pit detection in Section 2.6, thresholds on the posterior probability are used to assign a confidence level to each heap

detection. The thresholds were optimized on the Larvik training set, using the same penalty weights as before (Table 3). The resulting thresholds (Table 6) assigns 'medium' confidence to the majority of the false detections, and 'medium high' or better confidence to heaps that we think are grave mounds (Table 7).

Table 6. Thresholds for confidence assignment for heap detection.

1	2	3	4	5
0.05072984	0.05121662	0.47666119	0.67167690	0.76737689

Table 7. The result of using the Mahalanobis distance classifier for confidence estimation on the Larvik training set.

score value	1	2	3	4	5	6	
confidence	very low	low	medium	medium high	high	very high	sum
grave mounds			16	47	21	12	96
not grave mounds	55	1	554	72	7		689
sum	55	1	570	119	28	12	785

3. RESULTS

3.1 Automatic pit detection using manually designed decision tree

Table 8. The result of running the decision tree confidence assignment on the Olstappen test set for pit detection.

score value	1	2	3	4	5	6		
confidence	very low	low	medium	medium high	high	very high	not detected	sum
pit confirmed in field		2	13	10	7	8	5	45
pit visually in image	1	11	28	24	19	5		88
sum true pits	1	13	41	34	26	13	5	133
modern/other in field	1	1	1	1				4
not pit visually	384	375	90	11	2			862
sum	386	389	132	46	28	13	5	999

Table 9. Accumulated pit detection counts for different confidence levels on the Olstappen test set.

score value	≥1	≥2	≥3	≥4	≥5	≥6		
confidence	very low or better	low or better	medium or better	medium high or better	high or better	very high	not detected	sum
pit confirmed in field	40	40	38	25	15	8	5	45
pit visually in image	88	87	76	48	24	5		88
sum true pits	128	127	114	73	39	13	5	133
modern/other in field	4	3	2	1				
not pit visually	862	478	103	13	2			862
sum	994	608	219	87	41	13	5	999
pits detected	96.24%	95.49%	85.71%	54.89%	29.32%	9.77%		
pits missed	3.76%	4.51%	14.29%	45.11%	70.68%	90.23%		

By running the confidence assignment decision tree with the thresholds in Table 1 on the Olstappen test set, slightly worse results were obtained (Table 8) compared with the training set (Table 2). Fewer true detections obtained very high confidence, and more true detections obtained low or medium confidence. Still, the confidence levels reflect the number of true versus false detections in a meaningful way. All detections with 'very high' confidence are confirmed by archaeologists, either by field survey or by visual inspection of the lidar data. By accumulating the detection counts (Table 9), the trade-off between detecting as many pits as possible while at the same time limiting the number of false detections is more evident. E.g., 114 of 133 pits of archaeological interest were detected with medium confidence or better (Table 9); this is 85.7% of the pits of archaeological interest. At the same time, 103 of the detections with medium or better confidence were false. Alternatively, one may want to accept a higher number of false detections to obtain more true detections. 127 of 133 'true' pits

were detected with low or better confidence, which is 95.5%. This is achieved by accepting 478 'false' pits. Of the remaining 6 pits of archaeological interest that were not detected, 5 were lost due to missing ground returns in the lidar data due to vegetation. Many of the false detections could easily be removed by using digital map overlays, or were otherwise obvious misclassifications due to the context of the terrain.

3.2 Automatic pit detection using statistical classifier

By running the Mahalanobis distance classifier as described in Section 2.6 on the Olstappen test set, more archaeological pits were assigned 'high' and 'very high' confidence (Table 10-Table 11) than when using the manually designed decision tree (Table 8-Table 9). On the other hand, very few false pits were assigned 'very low' confidence by the Mahalanobis distance classifier. It seems like most of the true and false pits that the decision tree classifier labelled with 'very low' or 'low' confidence, were labelled with 'low' or 'medium' confidence by the Mahalanobis distance classifier.

Table 10. The result of confidence level assignment using the Mahalanobis distance classifier on the Olstappen test set.

score value	1	2	3	4	5	6		
confidence	very low	low	medium	medium high	high	very high	not detected	sum
pit confirmed in field	1	1	4	13	6	15	5	45
pit visually in image		5	9	13	27	34		88
not pit visually	22	406	391	41	6			866
sum	23	412	404	67	39	49	5	999

Table 11. Accumulated pit detection counts for the Mahalanobis distance classifier

score value	≥1	≥2	≥3	≥4	≥5	≥6		
confidence	very low or better	low or better	medium or better	medium high or better	high or better	very high	not detected	sum
pit confirmed in field	40	39	38	34	21	15	5	45
pit visually in image	88	88	83	74	61	34		88
sum true pits	**128**	**127**	**121**	**108**	**82**	**49**	**5**	**133**
not pit visually	866	844	438	47	6			866
sum	994	971	559	155	88	49	5	999
pits detected	96,24 %	95,49 %	90,98 %	81,20 %	61,65 %	36,84 %		
pits missed	3,76 %	4,51 %	9,02 %	18,80 %	38,35 %	63,16 %		

For pit detection the best statistical classifier is better than the manually constructed decision tree for high confidence detections. On the Olstappen test set, the Mahalanobis distance classifier assigns 'high' confidence or better to 82 confirmed pits, which is 62% of the confirmed pits, with only six additional false detections. The manually designed decision tree assigns 'medium high' confidence or better to 73 confirmed pits (55%), with 14 false detections. However, for the 'low' confidence detections, the manually designed decision tree seems to work better. The Mahalanobis distance classifier assigns 'very low' confidence to only 22 false pits, while the manually constructed decision tree assigns 'very low' confidence to 385 false pits. Both methods assign 'very low' confidence to only one confirmed pit.

3.3 Automatic heap detection using statistical classifier

By running the Mahalanobis distance classifier on the Larvik test set (Table 12), almost none of the false detections get 'low' or 'very low' confidence. So, in an operational setting, to successfully verify the 14 grave mounds with 'medium' confidence, 647 false detections have to be checked as well. However, for the 'medium high' or better confidence levels, the number of false detections is reasonable (Table 13).

Table 12. Result of running the Mahalanobis distance classifier for confidence estimation on the Larvik test set.

score value	1	2	3	4	5	6	
confidence	very low	low	medium	medium high	high	very high	sum
grave mound			14	39	25	18	96
not grave mound	4		647	144	13	1	809
sum	4	0	661	183	38	19	905

Table 13. Accumulated heap detection counts for the Mahalanobis distance classifier on the Larvik test set.

score value	≥1	≥2	≥3	≥4	≥5	≥6	
confidence	very low or better	low or better	medium or better	medium high or better	high or better	very high	sum
grave mound			14	39	25	18	96
not grave mound	4		647	144	13	1	809
grave mound	**96**	**96**	**96**	**82**	**43**	**18**	**96**
not grave mound	809	805	805	158	14	1	866
sum	994	608	219	87	41	13	999
heaps detected	100,00 %	100,00 %	100,00 %	85,42 %	44,79 %	18,75 %	
heaps missed	0,00 %	0,00 %	0,00 %	14,58 %	55,21 %	81,25 %	

4. DISCUSSION AND CONCLUSIONS

The pit detection results on the Olstappen test set indicate that the manually designed decision tree method is capable of detecting 95% of the pits of archaeological interest that were visible in the terrain, while at the same time producing four times as many false detections as true detections. Experience from field work indicates that this is an acceptable trade-off. Further, the automatic method was able to detect several small pits that were overlooked by visual inspection of the lidar data. The combined use of automatic detection and visual inspection prior to field survey is now being used by archaeologists in Oppland County, Norway, for the mapping of ancient hunting systems and iron production sites.

We have seen that the best statistical classifier for pit detection in the Olstappen data set, i.e., the Mahalanobis distance classifier, seems to work better than the manually constructed decision tree for the detection of pits with high confidence. However, the decision tree seems to be better for low confidence pit detections. Perhaps better performance could be obtained for the Mahalanobis distance classifier by reducing the penalty weights for confirmed pits with low confidence values. Alternatively, one could use a separate method to re-estimate confidence values for pit detections with very low to medium confidence, or to detections with posterior pit probability < 0.5.

It should also be noted that the heap detections have not been verified by archaeologists yet, as this part of our work is in an early stage. We plan a field work campaign in Larvik municipality this summer, after which we plan to redo the evaluation of automatic heap detection methods.

The lidar data could also be used to extract other features of archaeological interest than circular features, like pits and heaps. For linear features, like stone fences and old roads, we propose to use the following general method:
1. Apply multi-resolution edge- and/or ridge-detectors
2. Use digital map overlays to remove detections due to modern features such as roads.
3. For each resolution, threshold the detection result and connect neighbouring detections having approximately the same orientation
4. Close small gaps
5. Compute measures such as length, height above terrain, etc.
6. Classify the linear features either by using fixed thresholds or by using a statistical classifier.

In conclusion, semi-automatic detection of cultural heritage in lidar data is a valuable tool in combination with visual inspection of the lidar data, prior to field survey. Provided that the point density of the lidar data is high enough, the experience from the automatic pit detection method suggests that field survey can be accomplished ten times faster compared to the traditional approach without lidar data. Obviously, this is mainly due to the use of lidar data in itself, but automatic detection contributes both by reducing the time required for visual interpretation and by detecting pits that are missed during visual inspection.

REFERENCES

Aurdal, L., Eikvil, L., Koren, H., Loska, A., 2006. Semi-automatic search for cultural heritage sites in satellite images. In: *From Space to Place, Proceedings of the 2nd International Conference on Remote Sensing in Archaeology*, Rome, Italy, December 4-7, 2006, pp. 1-6.

Devereux, B. J., Amable, G. S., Crow, P., Cliff, A. D., 2005. The potential of airborne lidar for detection of archaeological features under woodland canopies. *Antiquity* 79, pp. 648-660.

Hastie, T., Tibshirani, R., Friedman, J., 2009. *The elements of statistical learning. Data mining, inference and prediction*. Second edition. Springer, New York.

LAS specification, version 1.3 – R11, October 24, 2010. The American Society for Photogrammetry & Remote Sensing, 18 pp. [online 2012-02-07] URL: http://www.asprs.org/a/society/committees/standards/LAS_1_3_r11.pdf

Prokop, R. J., Reeves, A. P. 1992. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing* 54(5), pp. 438–460.

Pudil, P., Novovičova, J., Kittler, J., 1994. Floating search methods in feature selection. *Pattern Recognition Letters* 15, pp. 1119-1125.

Trier, Ø. D., Larsen, S. Ø., Solberg, R., 2009. Automatic detection of circular structures in high-resolution satellite images of agricultural land. *Archaeological Prospection* 16(1), pp. 1-15. DOI: 10.1002/arp.339.

Trier, Ø. D., Pilø, L. H., 2012. Automatic detection of pit structures in airborne laser scanning data. *Archaeological Prospection*, to appear.