Using pattern recognition to search LIDAR data for archeological sites

ARJAN DE BOER

ADC ArcheoProjecten, Nijverheidsweg-Noord 114, a.de.boer@archeologie.nl

ABSTRACT

Since the introduction of LIDAR-data researchers used them to search for archaeological sites. However, standard methods to efficiently explore these data are missing. As pattern recognition techniques have proven their value in other research areas, a pilot study was carried out to determine the potential for archaeological applications. After a brief overview of archaeological use of LIDAR-data, pattern recognition is briefly discussed. Next, the methods and results of the pilot study are described in which a template based technique is deployed. It is shown that with this technique registered burial mounds are traced effectively, their co-ordinates can be determined more accurately and that many 'potential burial mound' locations are found. The paper concludes that archeological pattern recognition techniques offer a valuable potential for tracing archaeological sites and gives directions for further research.

1. INTRODUCTION

In the Netherlands, the upcoming implementation of the European treaty of Valletta (Council of Europe, 1992) has increased the research volume of archaeological inventories. Because the research period for individual projects is often limited, it has become important to develop strategies to explore regions in an accurate and efficient way. Since the introduction of Light Detection And Ranging (LIDAR) several years ago, large amounts of high quality data became available, and ever since, various researchers used the data to search for archaeological sites. However, standard methods to explore these data are missing: results depend heavily on the quality and experience of the individual researchers.

Recently, in the framework of a larger project of the Dutch National Service for Archaeological Heritage (ROB) a study was made to explore the possibilities of pattern recognition techniques in relation to LIDAR-data. The science of pattern recognition studies the design and operation of systems that are able to recognize patterns in data (Theodoridis and Koutroumbas, 1999). Applications can be found in various research areas, such as, fingerprint matching, iris recognition, handwriting recognition and industrial inspection. Considering the degree of success in those areas, pattern recognition algorithms offer a promising opportunity in archaeology to explore large amounts of data in a structural and cost-effective way. Archaeological researchers have developed computerized techniques to identify patterns in post hole distributions (e.g., Bradley and Small, 1985; Fletcher and Lock, 1984). However, the latter applications use fundamentally different data sources, and, therefore, are not easily adapted to the use of LIDAR data. Currently, LIDAR applications in archaeology published so far have shown valuable results, but are usually restricted to visualizing the data and subsequently searching the data for interesting phenomena (Schmidt *et al.*, 2005; Sittler and Daeffler, 2005; Van Zijverden and Laan, 2004; Herzog, 2001; Fletcher and Spicer, 1992).

The aim of this paper is to show the potential of archaeological pattern recognition using the results of a pilot study. The outline of this paper is as follows. Firstly, a brief overview of pattern recognition is given. Next, a pilot-study is described in which a pattern recognition technique is applied to search for burial mounds in LIDAR data. Finally, the conclusions are presented with directions for further research.

2. PATTERN RECOGNITION

Pattern recognition uses techniques such as feature extraction, discriminant analysis, principal component analysis, cluster analysis, neural networks and image processing to search for data with a set of predefined characteristics. Algorithms can be categorized into two classes (e.g. Brunelli and Poggio, 1993): template-based and feature-based.

Template-based algorithms calculate the correlation between one or more templates and (a selection of) the data. Basically, this method isolates a small part of the data representing the phenomenon in question. Next, using a correlation technique, the whole dataset is analyzed for similar patterns, which will yield a high correlation factor.

Feature-based algorithms analyze local features in the data and their spatial relationships. For example, using surface altitude data, features that can be used include the absolute or relative altitude, the slope or the change of slope, the roughness and the smoothness of the surface and the exposition. Using a known relationship between the occurrence of these features and the occurrence of the phenomenon in question, the latter can be traced.

3. PILOT STUDY

3.1 STUDY AREAS

In the pilot study the use of a template-based technique was explored to search LIDAR-data for burial mounds in the central part of the Netherlands. Results were described earlier in a technical report in Dutch (Laan and De Boer, 2005). The hemispherical burial mounds mostly date from the Neolithic until the Iron Age and typically have a diameter of 10 m and a height of 1 m, though their size may vary considerably (see e.g. Theunissen, 1999). As a result of the reclamation of previously uncultivated land – especially in the beginning of the 19th century – many burial mounds were destroyed. Today, the best-preserved mounds occur in forested areas that were considered unsuitable for cultivation during the 19th and 20th centuries. It is likely that more burial mounds exist than known so far because they are relatively difficult to detect in the forest using field surveys.

Three forested study areas were selected for the pilot study (Fig. 1).

The size of the areas is approximately 15 km^2 (Garderen), 12 km^2 (Putterbos) and 52 km^2 (Rhenen). All three areas are located on an ice pushed ridge from the Weichselien glacial period and have altitudes ranging from 10 to 80 m above mean sea level. The terrain is undulating as a result of a variety of glacial morphological features such as dead-ice pits, large scale eolian dunes and glacial meltwater valleys. In addition, in some parts small scale drift sands occur due to local medieval reclamation activities. The currently registered number of burial mounds in the Garderen, Putterbos and Rhenen areas is 32, 69 and 110, respectively[1]. The accuracy of the registered X and Y co-ordinates is either 1, 10 or 1000 m. Most of the mounds in the study area have their coordinates registered with an accuracy of 10 m. A subset of 18 burial mounds in the Rhenen area has been surveyed recently and their co-ordinates are known with centimeter accuracy. It should be stressed that, as these burial mounds are registered, they are – albeit with minimal efforts – monitored, restored when damaged, and periodically cleared of vegetation.

3.2 METHODS

The raw LIDAR data was obtained from the Ministry of Public Works of the Netherlands. Generally, the data has an average point density of one point every 16 m^2. In forested areas the point density may drop to one point every 36 m^2. The standard deviation is reported to be 0.15 m in vertical direction (Ministry of Public Works, 2000). The accuracy of the point measurements in horizontal direction is estimates it to be less than 0.3 m^2. Before data is released, reflections other than those from the surface are removed using a filter[3].

Of each area a regular grid (Digital Elevation Model) was constructed based on raw laser altimetry data using kriging (Burrough and McDonnell, 1998). The DEM's have gridcell sizes of 2 m (Garderen and Putterbos), and 4 m (Rhenen). The regular grids were analyzed using a template technique. A template is a small DEM of the phenomenon under study. The template constructed for the pilot study had to resemble a burial mound. Because of the exploring character of the study, a basic simple shape was chosen. The template is sinusoidal (0 – ¶) in cross-section and circular seen from above (Figs. 2 and 3).

Because the template and the real world DEM were normalized before comparing, the height of the template mound did not matter. To account for varying diameters, the size of the template was varied roughly according to the range of the mounds which was expected and could be detected in the LIDAR data (diameters: 12, 16, 20, 24 and 28 m).

A measure (r) analogue to the Pearson product-moment correlation coefficient was used to determine the correlation between the template and a sample of the DEM location as follows:

$$r = \frac{\sum_{j=1}^{d}\sum_{i=1}^{d}(T_{i,j}-\overline{T})(S_{i,j}-\overline{S})}{\sigma_T\,\sigma_S\,d^2}$$

were i and j are row and column indexes of the template (T) and a sample (S), the over line ($\overline{\ }$) denotes mean grid values, d is the size of the template expressed in cell units and σ is the standard deviation. The correlation ranges from -1 (negative correlation) to +1 (positive correlation). The correlation was calculated for all locations in a study area using the range of template sizes mentioned before, resulting in five correlation maps (one for each template diameter) for each of the three

1 These data were obtained from the National Service for Archaeological Heritage (Dutch acronym: ROB).

2 Public Works agency estimatioin.

3 The contractors who obtain the LIDAR data apply this filter. The precise filter is unknown as the contractor companies consider the algorithms to be company secrets (Ministry of Public Works 2000).

study areas. At locations with a high correlation, the surface resembles a burial mound. Whether burial mounds exist at these locations could only be verified in the field – a time consuming task which was not done due to time limitations. These locations are called 'potential burial mounds', abbreviated as PBM's.
The results of the calculations were assessed as follows. The locations of registered burial mounds were plotted on the correlation maps. Next, the correlation maxima in relation to the location of registered burial mounds were inspected visually on these maps. In addition, the correlation at the location of the mounds in the Rhenen subset was determined on each map so that the 'scores' of the several templates could be compared.

3.3 RESULTS

In Figs. 4, 5 and 6, the DEM's of all study areas are shown. It appears that burial mounds can be traced visually on these images when zoomed in on various parts of the DEM's.
In Figs. 10, 11 and 12, the result the correlation of the DEM's with a 16m template is shown for the enlargements. In these figures, the co-ordinates of the registered burial mounds are indicated with triangles
These figures illustrate the following:

- Most registered burial mounds have a high correlation. It shows that the shape of the template corresponds to 'real world' mounds and that the applied algorithm functions as expected.
- Apart from those at the position of registered mounds, many more PBM locations are found. It suggests that more than the currently known mounds exist. This suggestion is considered plausible by researchers in the area. Clearly, field surveys are necessary to determine the archeological value of these PBM's.
- The co-ordinates of the registered burial mounds are often very inaccurate, see e.g. Fig. 11. It may lead to erroneous protection of archeological sites. This currently presented technique may assist in finding more precise co-ordinates.
- There is missing data right at the location of several burial mound, see e.g. Fig. 12. Probably, the applied filter technique considered the laser measurements of elevated surface at the burial mounds as 'outliers' and removed these. Clearly, 'standard' filter techniques should be used with care, as burial mounds are anomalies from the 'normal' surface. This problem is also illustrated by Waldus and Van der Velde (2005).

The correlation of the templates with various diameters at the location of the Rhenen subset is presented in Fig. 13. Four out of 18 burial mounds were not detected in the LIDAR data as a result of missing data at these particular locations. The figure shows that the correlation is always positive, typically between 0.3 and 0.9. Two burial mounds (no's 5 and 9) have a very low correlation, the reason of which should be given further attention in future research. It is striking that the 16m template usually performs the best. The diameter of the mounds in the Rhenen subset is indeed of order 15m.

4. CONCLUSIONS AND RECOMMENDATIONS

The pilot study showed that registered burial mounds can be traced with a relatively simple technique. Many new locations were found which could be selected for field survey. By plotting the PBM's on a map (Fig. 14) and or using GPS, such a field survey could be done efficiently.
The main application for archeological pattern recognition is in the process of making archeological inventories on a regional scale. In the Valetta framework, the Dutch government presses regional authorities to have these inventories available to be able to protect known or potential cultural heritage. The author is convinced that archeological pattern recognition offers an opportunity improve and extent the existing practice of making these inventories. However, the technique should be developed by:

- using more and different shaped templates,
- using and combining with feature based techniques, and by
- combining the results with earth science resources, such as, geology and pedology.

It is useful to stress that in other landscape types, archeological sites with a shape similar to burial mounds exist, such as, house or village-mounds and coversand-dunes buried by Holocene deposits. These features can probably be traced using techniques similar as those used in the pilot. In general, all phenomena visible in the local surface morphology can potentially be traced using pattern recognition techniques in combination with LIDAR-data, e.g., historical routes, Celtic fields, sand pits, medieval parcel structures, high potential meander belts and crevasse splays. Well-preserved historical landscapes have most potential with this technique, and, as such, it is also outside the Netherlands that this technique may be profitable.

ACKNOWLEDGEMENTS

I am grateful to Dr. Spek for this financial and substantive support of the pilot project. I thank Dr. Kamermans and Drs. Van Zijverden for their enthusiasm and offering the opportunity to present this work in Portugal. Finally, I am in-debted to Drs. Dekker and Drs. Laan for their support with data processing and exchange of ideas during the research.

REFERENCES

BRADLEY, R.; SMALL, C. (1985) – Looking for circular structures in post hole distributions: Quantitative analysis of two settlements from bronze age England. *Journal of Archaeological Science*, 12, p. 285-297.

BRUNELLI, R.; POGGIO, T. (1993) – Face recognition: features versus templates. *IEEE Transactions on PAMI*, 15 (10), p. 1042-1052.

BURROUGH, P. A.; MCDONNELL, R. A. (1998) – *Principles of Geographical Information Systems.* Oxford.

COUNCIL OF EUROPE (1992) – European convention on the protection of the archaeological heritage (revised). *Council of Europe Treaty Series*, 143.

FLETCHER, M.; LOCK, G. (1984) – Post built structures at Danebury hillfort: an analytical search method with statistical discussion. *Oxford Journal of Archaeology*, 3 (2), p. 175-196.

FLETCHER, M.; SPICER, D. (1992) – The Display and Analysis of Ridge-and Furrow from Topographically Surveyed Data. In REILLY, P.; RAHTZ, S., eds. – *Archaeology and the information age: a global perspective.* London: Routledge.

HERZOG, I. (2001) – Ehemalige Materialentnahmegruben erkennen – Auswertung von Höhendaten. *Archäologische Informationen*, 24, 1, p. 39-43.

LAAN, W.; A. DE BOER (2005) – *AHN onderzoek West-Veluwe.* ADC rapport (in press), Amersfoort.

MINISTRY OF PUBLIC WORKS (RIJKSWATERSTAAT) (2000) – *Product Specification AHN 2000 (Productspecificatie AHN 2000)*, in Dutch.

SCHMIDT, S.; BOFINGER, J.; KURZ, S. (2005) – LIDAR – *High resolution raster data as a survey tool*, this publication.

SITTLER, S.; DAEFFLER, M. (2005) – *Assessing ancient landscapes fossilized under forests by using laser scanning: a pilot study to generate 3-D models of ridge and furrow in the upper Rhine Valley,* this publication.

THEODORIDIS, S.; KOUTROUMBAS, K. (1999) – *Pattern Recognition*. London: Academic Press.

THEUNISSEN, L. (1999) – *Midden-bronstijdsamenlevingen in het zuiden van de Lage Landen*. Ph.D. dissertation, Leiden University.

VAN ZIJVERDEN, W.; LAAN, W. (2004) – Landscape reconstructions and predictive modeling in archaeological research, using a LIDAR based DEM and digital boring databases. In: *Archäologie und Computer*. Workshop 7 (Vienna 2004).

WALDUS W.; VAN DER VELDE, H. (red.) (2005) – Archeologie in vogelvlucht: een onderzoek naar de toepassingsmogelijkheden van het AHN voor de archeologie. ADC rapport (in press), Amersfoort.

FIGURES

Fig. 1 – Location of the study areas. All maps in this paper are oriented with North to the top.

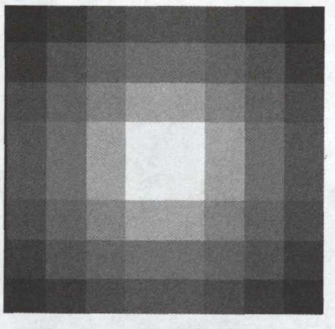

Fig. 2 – Template in 2D view from above. The shades indicate the altitude – a lighter shade is higher.

This is a 16m template in a 2 x 2m grid (8 x 8 gridcells).

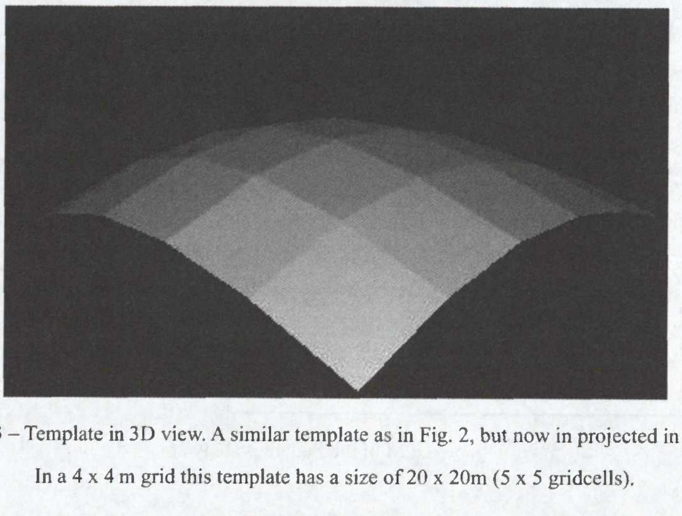

Fig. 3 – Template in 3D view. A similar template as in Fig. 2, but now in projected in 3D.

In a 4 x 4 m grid this template has a size of 20 x 20m (5 x 5 gridcells).

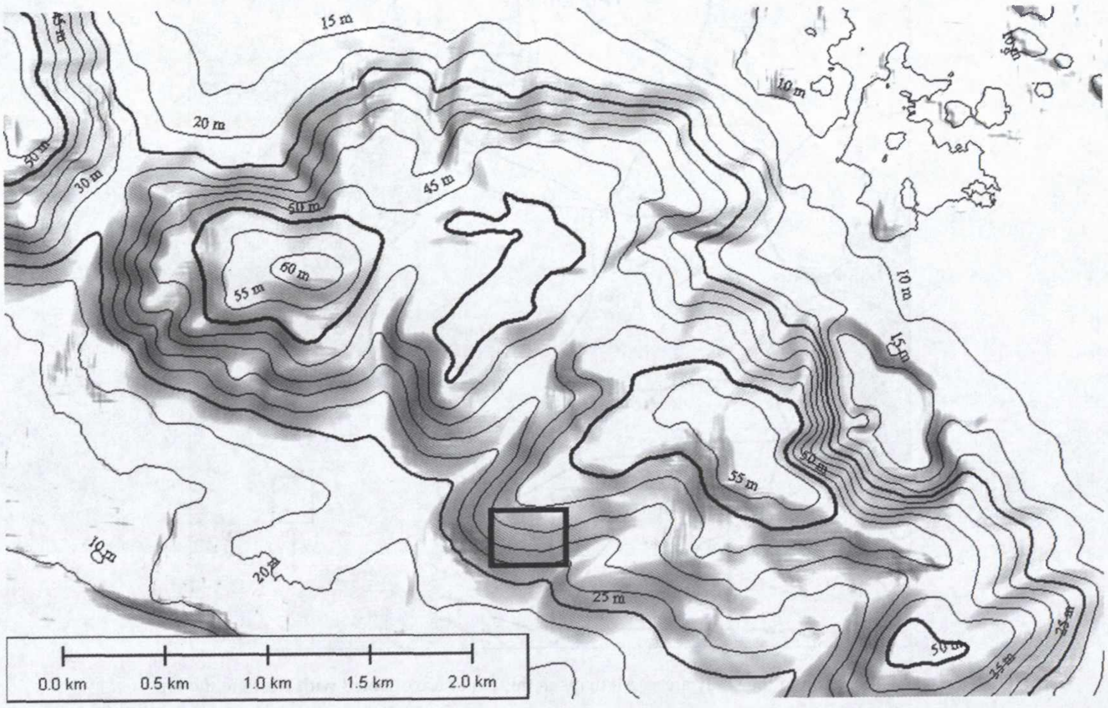

Fig. 4 – DEM of the Garderen study area.

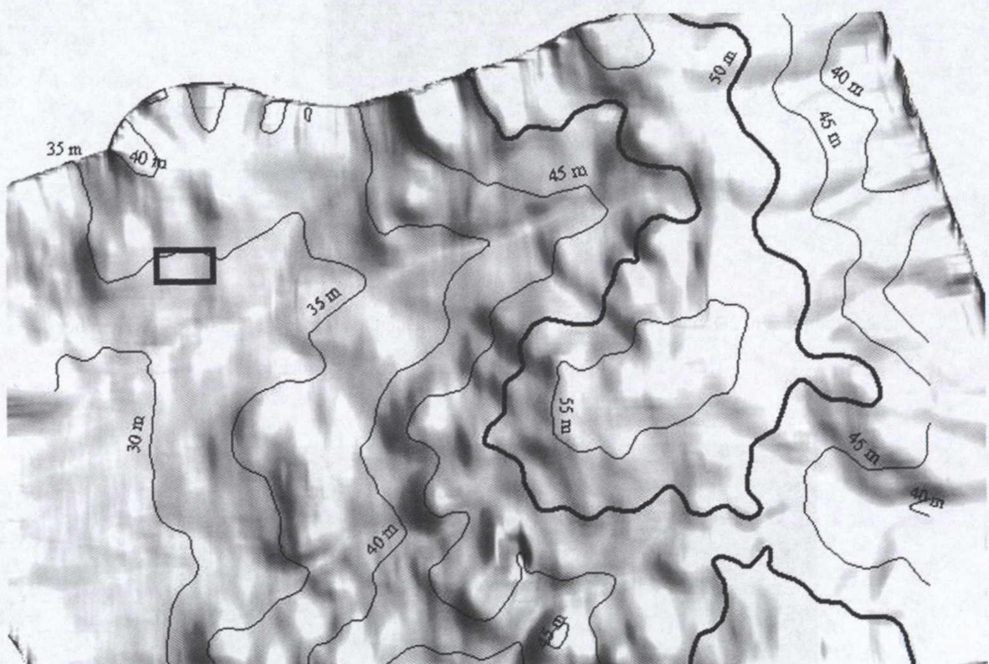

Fig. 5 – DEM of the Putterbos study area.

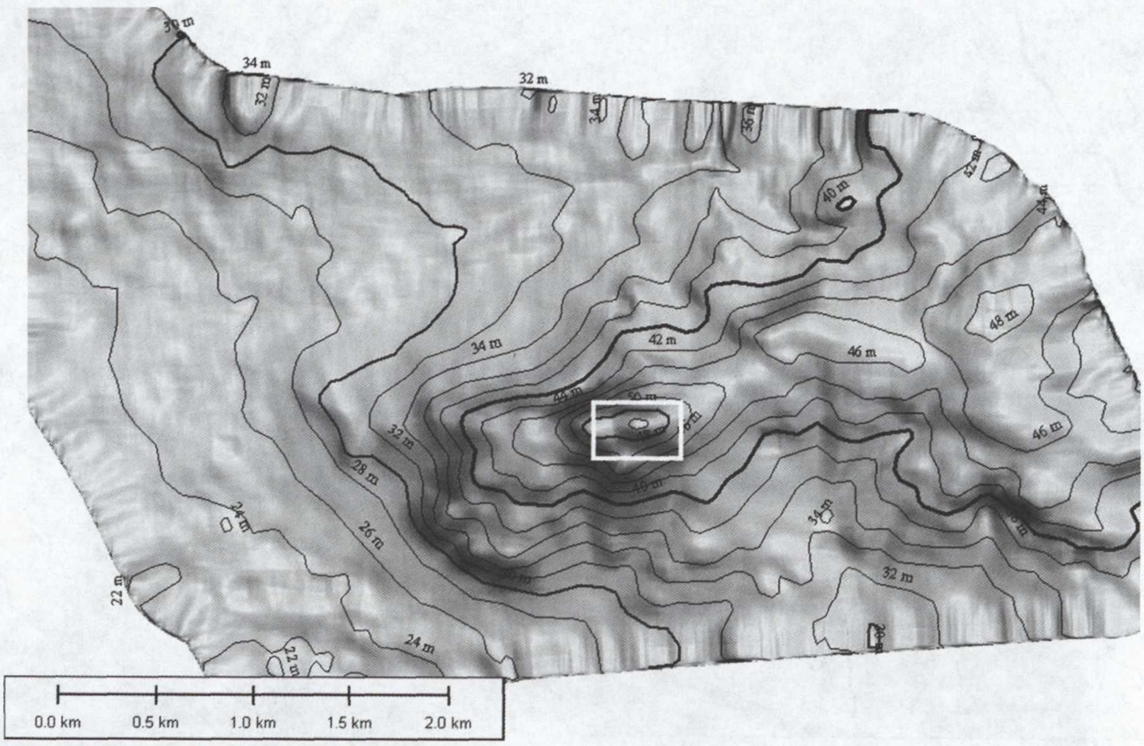

Fig. 6 – DEM of the Rhenen study area.

In Figs. 7, 8 and 9, examples of enlargements in which burial mounds occur are shown.

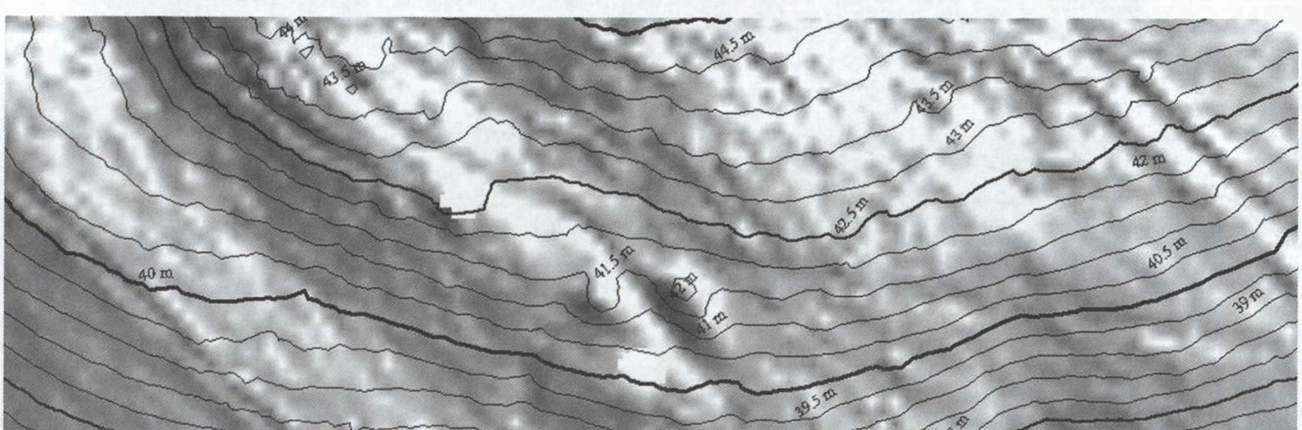

Fig. 7 – Enlarged section of the Garderen DEM.

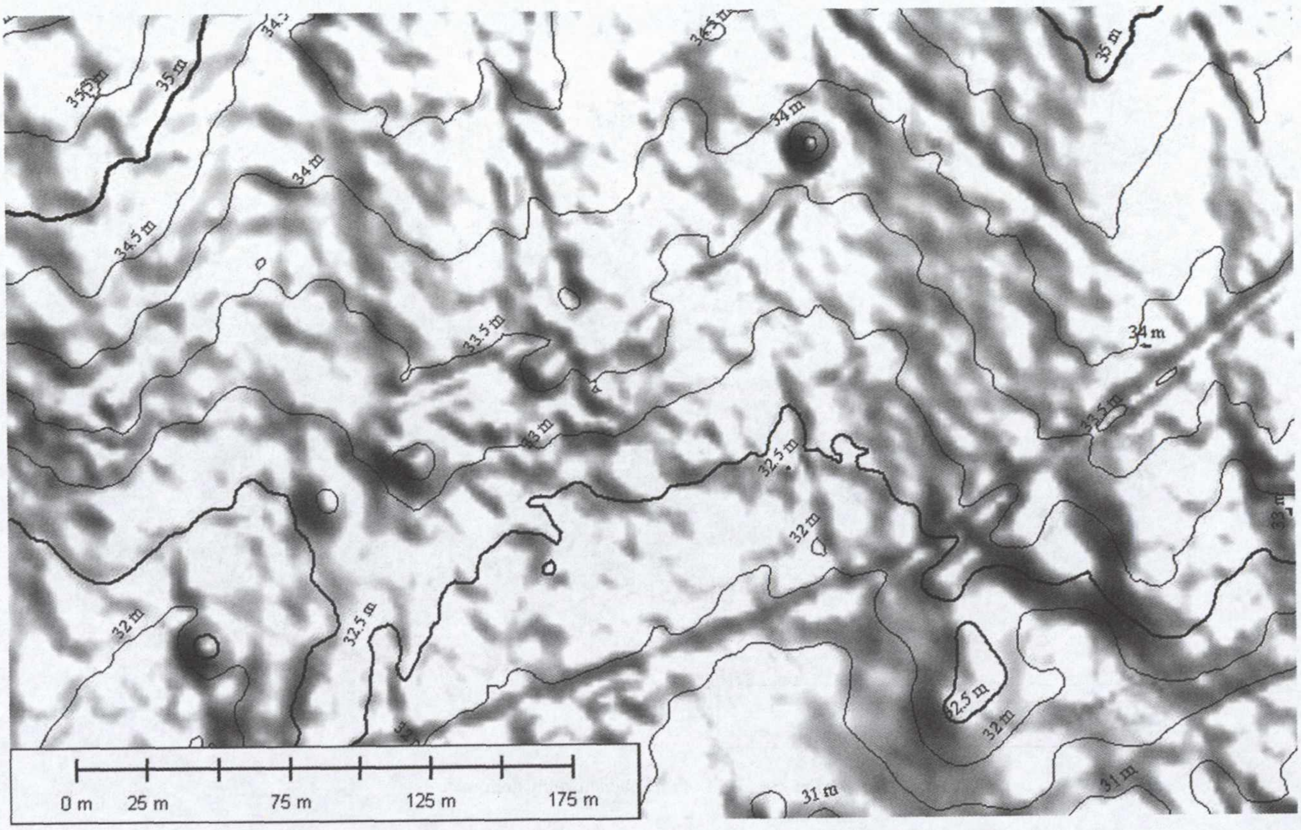

Fig. 8 – Enlarged section of the Putterbos DEM.

Fig. 9 – Enlarged section of the Rhenen DEM.

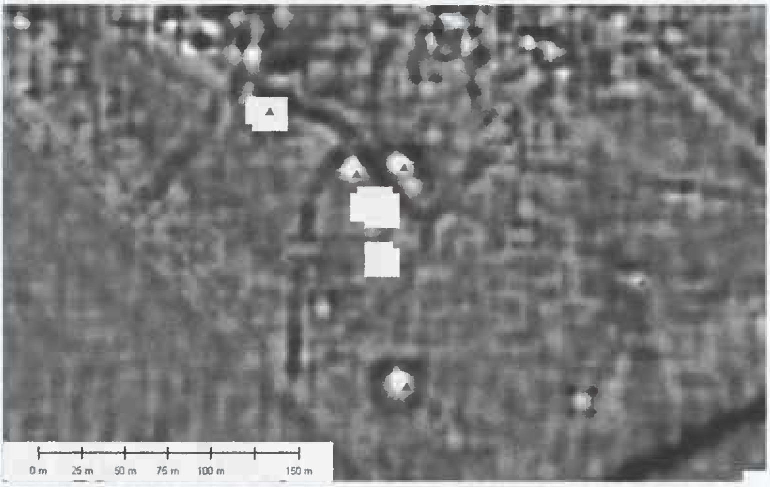

Fig. 10 – Results of the correlation with a 16m template in the Garderen area (compare Fig. 7).

The triangles are plotted on the locations were burial mounds are registered.

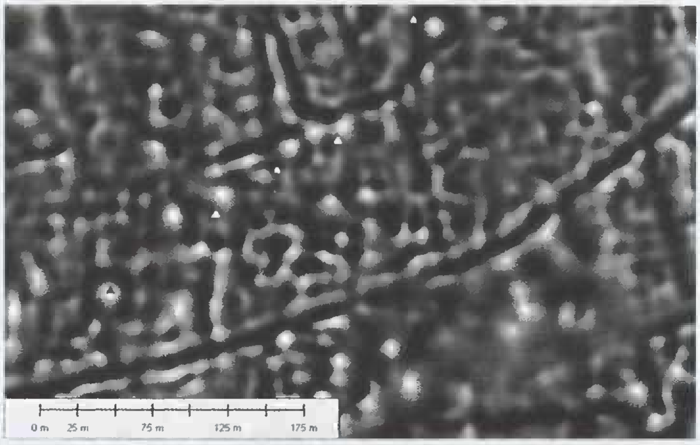

Fig. 11 – Results of the correlation with a 16m template in the Putterbos area (compare Fig. 8).

The triangles are plotted on the locations were burial mounds are registered.

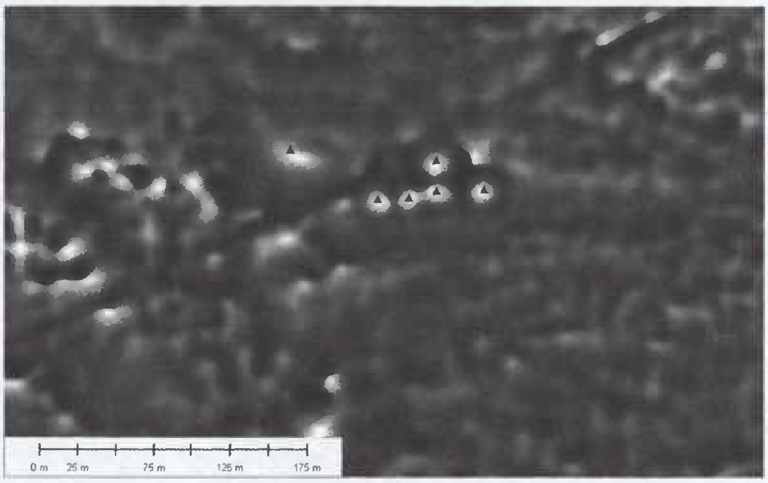

Fig. 12 – Results of the correlation with a 16m template in the Rhenen area (compare Fig. 9).

The triangles are plotted on the locations were burial mounds are registered.

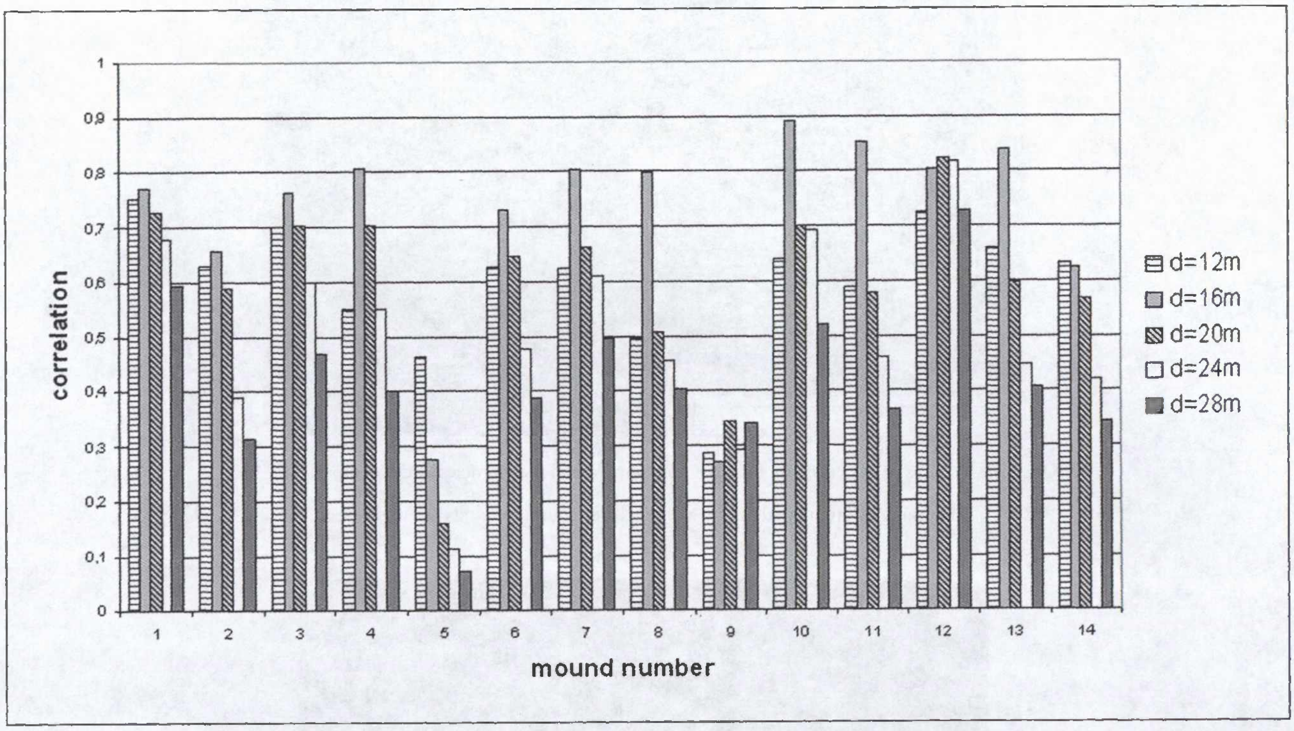

Fig. 13 – Correlation of the LIDAR-data with the 16m template at the location of the mounds in the Rhenen subset area.

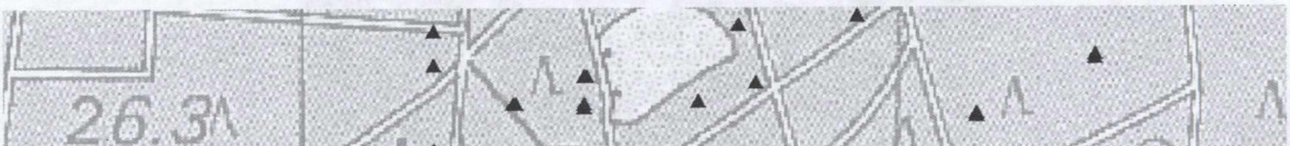

Fig. 14 – Example of a potential burial mound (PBM) map.