Towards an automated analysis of geophysical archaeological prospection data

Immo TRINKS[1,2], Alois HINTERLEITNER[2,3], Mario WALLNER[2], Klaus LÖCKER[2,3], Alexander BORNIK[2], Johannes HÖLLER[2], Matthias KUCERA[2], Roland FILZWIESER[2], Hannes SCHIEL[2,3], Tanja TRAUSMUTH[2], Alexandra VONKILCH[2], Jona SCHLEGEL[2], David RUSS[2], Wolfgang NEUBAUER[2,1]

[1] Vienna Institute for Archaeological Science, University of Vienna, Franz-Klein Gasse 1, 1180 Vienna, Austria
[2] Ludwig Boltzmann Institute for Archaeological Prospection & Virtual Archaeology, Hohe Warte 38, 1180 Vienna, Austria
[3] Zentralanstalt für Meteorologie und Geodynamik, Hohe Warte 38, 1180 Vienna, Austria

Keywords: *semi-automatic data analysis — archaeological prospection — magnetometry — ground-penetrating radar — extensive data coverage — high-resolution data.*

CHNT Reference: Trinks, I., Hinterleitner, A., Wallner, M., Löcker, K., Bornik, A., Höller, J., Kucera, M., Filzwieser, R., Schiel, H., Trausmuth, T., Vonkilch, A., Schlegel, J., Russ, D., Neubauer, W.: 2020. Towards an automated analysis of near-surface geophysical archaeological prospection data

Editorial team. Proceedings of the 25th Conference on Cultural Heritage and New Technologies. DOI:xxxxxxx.

Background

The past decade has seen enormous increases in efficiency regarding the acquisition of extensive high-resolution geophysical archaeological prospection data, particularly in case of magnetic and GPR surveys. These developments have resulted in substantially increased geophysical data sets, calling for a similar increase in the effectiveness of data processing and data analysis capabilities. However, the development of appropriate algorithms and tools is lagging, particularly when it comes to the analysis of 3D GPR data volumes.

Automating magnetic data analysis

Several automatic data processing and analysis steps can considerably enhance the interpretation efficiency of extensive high-resolution magnetic prospection data (Hinterleitner et al. 2015). It is essential to differentiate between the steps of unsupervised data segmentation leading to data classification, and the subsequent step of data interpretation conducted by an experienced interpreter, leading to feature extraction. Data segmentation divides the data into regions of similar magnetisation. Ideally, the segmentation process should leave no part of the data unsegmented. Subsequently, several attributes can be computed and assigned to the detected anomalies, such as the orientation of dipole anomalies derived from the relative location of the positive and negative maxima, the distance between the dipole minimum and maximum, the area covered by the anomaly. Based on the orientation of dipole anomalies, it is possible to automatically extract anomalies that could be caused by thermoremanent magnetisation. In case of anomalies that lack a dipole character and that could be caused by the magnetic infill of pit-like structures, the following parameters can be extracted semi-automatically: maximum magnetisation, area of an anomaly (based on a defined threshold value), the maximum and minimum diameter of the encompassing polygon, length of the describing polygon shape, the width to length ratio, it's circularity. The results of this semi-automatic segmentation process can be classified as the basis of subsequent interpretation and generation of interpretative maps. Semi-automatic algorithms that require only a few input parameters permit the automatic detection, outlining and classification of magnetic anomalies (e.g. iron objects, likely thermoremanently magnetised structures, possible pits or postholes). Large magnetic data sets covering square kilometres can contain many thousands of dipole anomalies and vast numbers of pit-like structures (Fig. 1). In one case study that comprised 968 ha of magnetic data, 119,842 dipole anomalies have been detected semi-automatically, as well as 666,507 archaeologically relevant non-dipole anomalies.

Precondition for the application of semi-automatic anomaly detection and classification algorithms is data of high quality, in which possible errors, systematically introduced through the measurement process, have been corrected. The entire process is optimised for the automatic generation of georeferenced greyscale data images with different black and white clip-off ranges, respectively float GeoTIFFs that are loaded into the Geographic Information System ArcGIS for Analysis with the custom, in-house developed *ArchaeoAnalyst* Toolbox. Dipole and non-dipole anomalies are extracted semi-automatically, requiring only the input of threshold values. In a manual step conducted by an expert in the interpretation of magnetic prospection data, the automatically generated shapes are checked and where necessary adjusted.

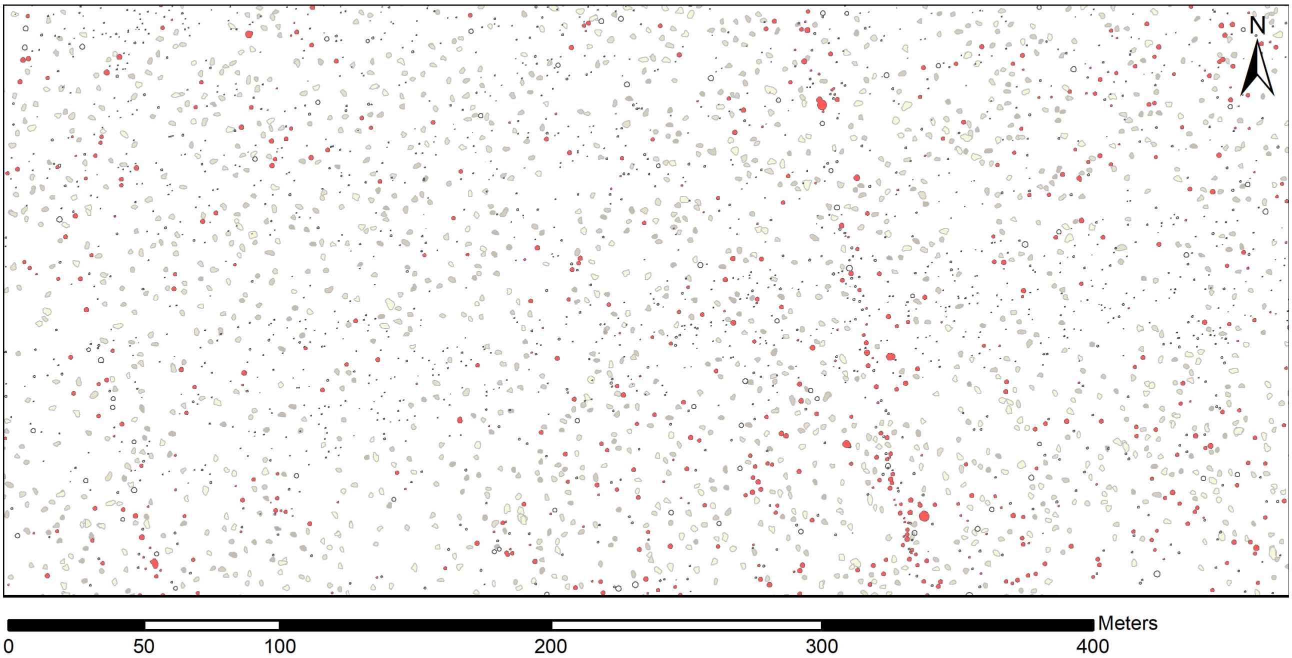

Fig. 1. Semi-automatic classification in numerous pits of different magnetisation (grey polygons) and ferromagnetic dipole anomalies (red) from a small subset of the gradiometer data collected for the Stonehenge Hidden Landscape Project.

Automating GPR data analysis

Due to the two- and three-dimensional nature of GPR data, GPR data analysis involves considerably more parameters than magnetometry. The multitude of different data processing options and visualisations possible for GPR data permits the preparation of various data sets that can be used as input for manual or automatized GPR data analysis. Depending on the quality of the data and the local geology, surface and subsurface conditions, different data visualisation can be favourable. It could be shown that in case of some sites, multi-trace coherence mapping is resulting in better visibility of apparently manmade subsurface structures than traditional amplitude mapping (Trinks and Hinterleitner 2020). Similar to magnetic data analysis, any automated GPR data analysis requires appropriately spatially sampled data with of high quality. Currently, the analysis of extensive high-resolution GPR data sets (Trinks et al. 2018) is rendered more efficient by automatically generating sets of georeferenced GPR depth-slice images of different thicknesses. These images are loaded into an ArcGIS geodatabase using the ArcGIS toolbox *ArchaeoAnalyst*, which subsequently permits the picking of reflection anomalies by semi-automatically drawing closed polygons around high amplitude anomalies. Using a hybrid volume and surface data rendering algorithm and different transfer functions inspired from medical imaging applications, after application of high-frequency de-noising algorithms that preserve the edges of larger structures, it is possible to extract the mayor features from GPR data (Fig. 2) and to generate high-fidelity volumetric interpretation representations (Neubauer et al. 2019).

Discussion

Currently, many geophysical archaeological prospection surveys still hardly go beyond the generation of data images, or basic line or polygon markings of the most prominent anomalies detected. Magnetometry surveys often are still conducted with too coarse sample spacing and sub-optimal data and image processing applied. In case of GPR prospection, there is still ample room for improvement with regard to spatial sampling resolution and frequency spectrum covered in order to achieve optimum image quality. The great potential of utilising GPR single- or multi-trace attributes aside of simple reflection amplitudes is only emerging. Advanced data visualisations derived from medical applications permit the imaging of structural information in the data that is difficult or impossible to derive from 2D slice representations alone. Data and image fusion permits the enhanced visualisation of information contained in different data sets or differently processed data, leading to an improved understanding of the imaged structures. The above described approach based on semi-automatic data segmentation and classification does not yet involve “artificial intelligence” (AI) or “machine learning” (ML). However, in our view it presents the state-of-the-art regarding the interpretation of extensive high-resolution geophysical prospection data sets and suggest that any future use of AI or ML for the enhanced analysis of geophysical archaeological prospection data should be measured against this approach.

We argue that any possible meaningful future exploration of AI or ML for enhanced processing and analysis of geophysical archaeological prospection data requires data of highest imaging resolution and quality. Particularly with regard to GPR data, we often still lack the basic understanding of the actual physical processes that cause anomalies observed in the prospection data. The careful acquisition of extensive high-resolution data sets in different geological and archaeological settings as well as the increased use of data modelling are highly likely to enhance our understanding of the underlying processes. Prior to utilising artificial intelligence and machine learning for automated data analyses, we should ensure that natural intelligence, human learning and the quality of the data subjected to such approaches are up to the task.

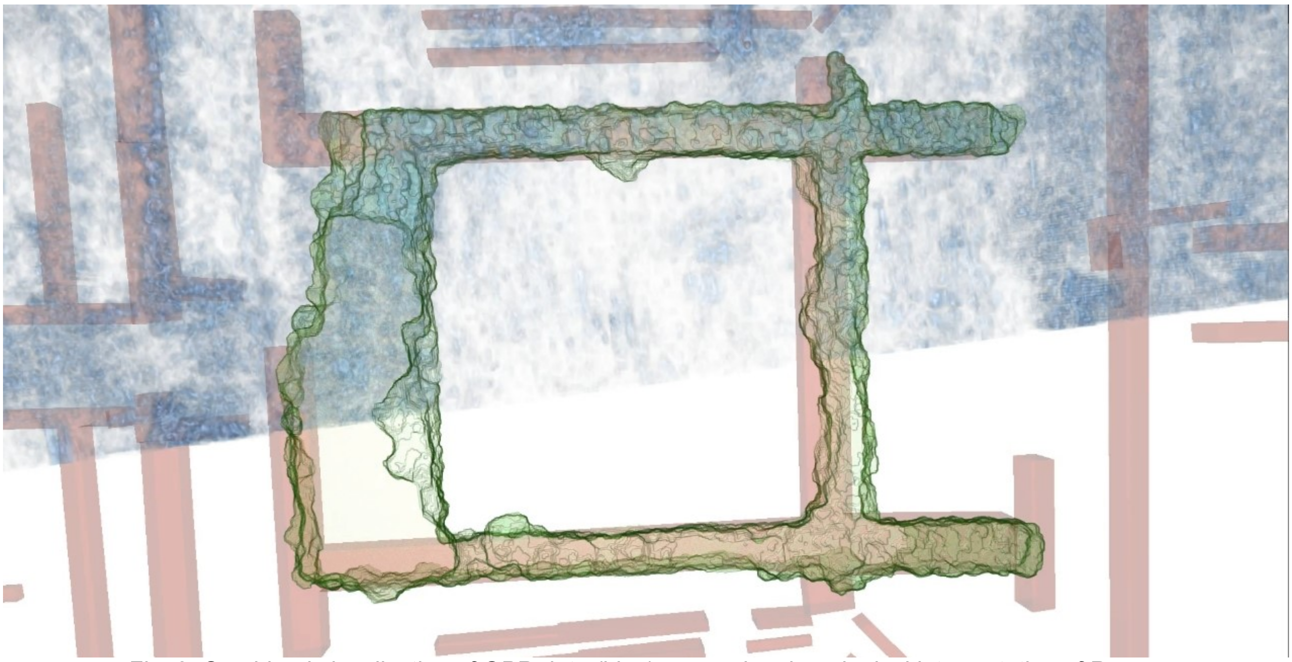

Fig. 2: Combined visualisation of GPR data (blue), manual archaeological interpretation of Roman foundation walls (red) and the detailed semi-automatic segmentation of one building.

References

Hinterleitner, A. Kastowsky-Priglinger, K., Löcker, K. Neubauer, W. Pregesbauer, M., Sandici, V. Trinks, I., Wallner, M. (2015) Automatic detection, outlining and classification of magnetic anomalies in large-scale archaeological magnetic prospection data. Archaeol. Polona 53, pp. 296-299.

Neubauer, W. Bornik, A., Wallner, M., Verhoeven, G. (2019) Novel volume visualisation of GPR data inspired by medical applications. New global perspectives on archaeological prospection. Proceedings of the 13th International Conference on Archaeological Prospection, 28 August – 1 September 2019, Sligo – Ireland

Trinks, I., Hinterleitner, A., Neubauer, W., et al. (2018) Large-area high-resolution ground-penetrating radar measurements for archaeological prospection. Archaeological Prospection 25, pp. 171– 195. https://doi.org/10.1002/arp.1599

Trinks, I., Hinterleitner, A. (2020) Beyond Amplitudes: Multi-Trace Coherence Analysis for Ground-Penetrating Radar Data Imaging. Remote Sens. 12, 1583.