

Love your polygons: powerful processing of vectorised geophysics data

Armin Schmidt ^{1*}

¹ Dr Armin Schmidt - GeodataWIZ, Remagen, Germany

* Corresponding author: E-mail: A.Schmidt@GeodataWIZ.com

Abstract

By converting raster data to vectors several spatial attributes become readily available. For non-gridded data the vectorisation can happen after creating a raster, from a Triangulated Irregular Network (TIN), or directly from the xyz data using an efficient spatial index. Classification, clustering and envelop creation is simple from vectorised data. Examples of aggregation and analysis of results are provided using different data sources.

Keywords

anomaly; classification; data envelope; processing; vector

Introduction

While most processing algorithms for archaeological geophysics data rely on raster representations it was shown that the use of vectorised forms opens up new avenues for spatial processing and classification through powerful vector manipulation and filtering engines (Schmidt and Tsetskhladze 2012, 2013; Schmidt 2016, 2019; Schmidt et al. 2019). The popular vector processing database PostGIS, for example, uses the OpenGIS Consortium (OGC) compatible GEOS library and the examples provided in this contribution refer to this specific implementation. The main advantage of vector-based methods is the immediate and fast availability of spatial characteristics for data areas (e.g. size, shape, elongation) and their mutual spatial relationships. The availability of powerful Machine Learning/Deep Learning (ML/DL) algorithms (Verdonck 2019; Küçükdemirci and Sarris 2022) has made some of these spatial classification schemes available to raster data (Sevara et al. 2016; Cornett and Ernenwein 2020). It hence seems timely to provide an overview of the different processing tasks that can be undertaken with vectorised data.

Overview

Vectorisation

To convert archaeological geophysical data into a vectorised form there are three main approaches available: (a) a data raster can be created first (either by combining individual data grids or by interpolation) from which vectors are then derived; (b) an intermediate Triangulated Irregular Network (TIN) is calculated; or (c) a spatial database is used with a spatial index that allows for very fast filtering to support the creation of vector representations. The latter two methods are well suited for data that were collected irregularly and tagged with their individual GPS/GNSS location, requiring interpolation if a raster representation were used.

The simplest vectorisation is the creation of polygons that enclose data points above a threshold value (e.g. based on a statistical analysis, using ST_Reclass and ST_Polygon (Schmidt and Tsetskhladze 2013)). For ground penetrating radar (GPR) data with strong contrast it was found that the resulting polygons only depended weakly on the exact choice of such a threshold.

In another approach the range of data values is represented by contour polygons, which can be created with various algorithms (ST_Contour or gdal_contour from the Geospatial Data Abstraction Library (GDAL)). Most contouring algorithms create contour lines and if their start and end points coincide they can easily be closed to form polygons.

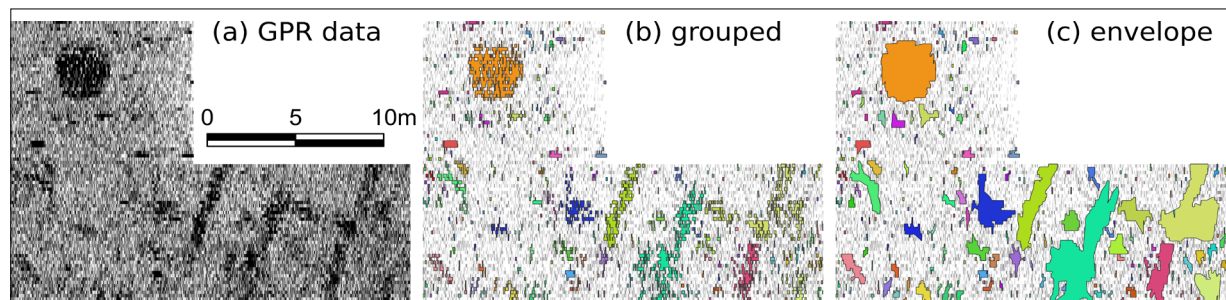


Fig. 1: GPR data from the Dutch Fort in Jaffna, Sri Lanka. (a) GPR data, strong reflections are black; (b) extracted anomalies grouped based on spatial vicinity; (c) envelopes of concave hulls around the grouped anomalies.

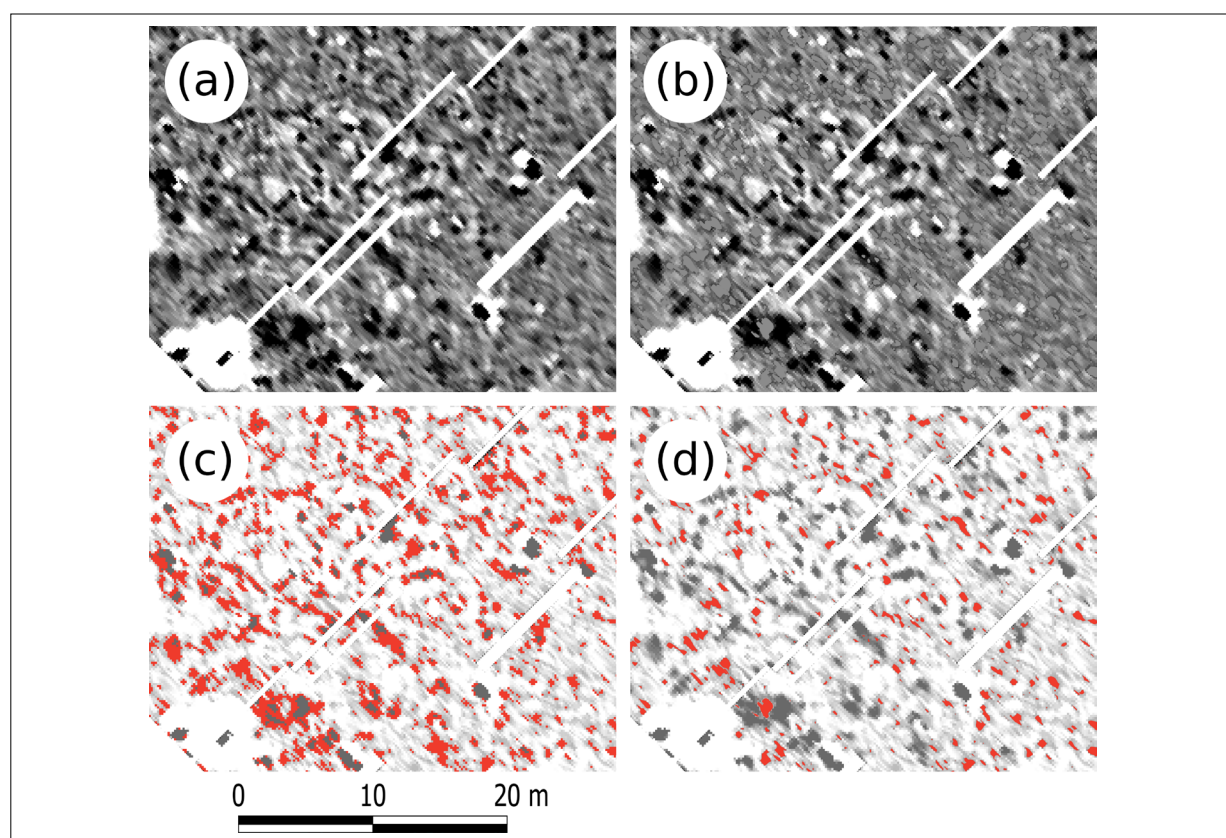


Fig. 2: Magnetometer data from the Iron Age enclosure in Eisenach, Rhineland-Palatinate, Germany. (a) magnetometer data with a greyscale range of -2.5 to +3.0 nT, white to black; (b) data with those anomalies masked that had a maximum between 1 and 3 nT; (c) all raster cells marked in red that have a data range between 1 and 3 nT; (d) only anomalies marked in red whose maximum is between 1 and 3 nT.

Aggregation

Especially in GPR time slices that were created from single-channel surveys, buried features sometimes manifest as disjoint small anomalies. While a human interpreter may be able to join them together, this is more difficult for computer-based classification schemes (Sevara et al. 2016). Using threshold derived polygons for all individual

parts of such a ‘dispersed anomaly’ it is possible for vector processing engines to create automatically combined anomalies based on spatial attributes of the constituent parts (Schmidt and Tsetschladze 2013). For example, anomalies with a very small extent can first be discarded and then all remaining anomalies that are a certain distance apart are combined into a single larger anomaly (Fig. 1).

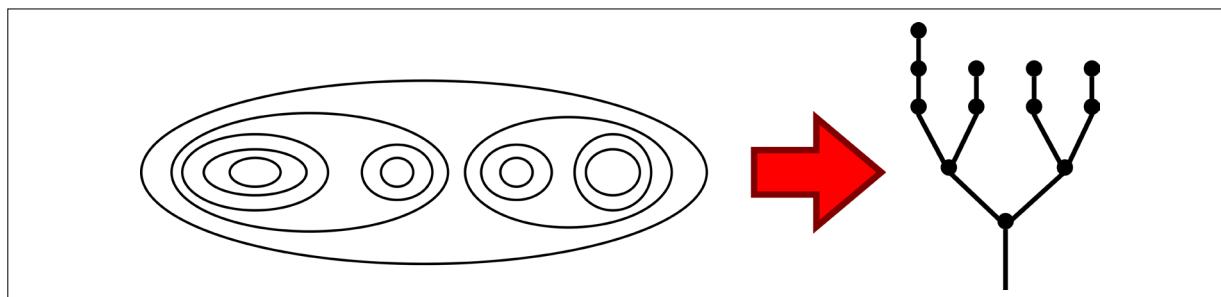


Fig. 3: The topological nesting structure of contour lines (left) can be represented by a tree structure (right).

Analysis

A useful method for the analysis and subsequent interpretation of magnetometer data is the identification of those anomalies that have a certain range of magnitudes, for example from 1-3 nT. A magnetometer survey inside an Iron Age enclosure ('Keltenschanze'), located in a wooded area in Germany, showed many small-size and small-value anomalies, which obliterated the majority of the site (Fig. 2a). They were due to ferrous remnants of shooting exercises after the end of the Second World War. Masking these small anomalies allowed a far easier evaluation of the site overall (Fig. 2b). This task cannot be accomplished easily with raster data alone, as masking a middle-range of data also affects anomalies from stronger features (Fig. 2c). If instead a topological classification of data-contours is undertaken it is possible to relate each contour to the maximum value of its anomaly and then select only those that belong to chosen range (Fig. 2d). Hinterleitner et al. (2019) achieved similar results with a raster method that descended from each raster peak to its base.

A topological tree structure can be built to represent how data-contours nest inside each other (Fig. 3) and by traversing this tree the minimum/maximum of each anomaly can easily be found.

This method was used to identify the contours that represents the magnetic anomalies' "full width at half maximum" (FWHM), which are reasonable proxies for the shape of buried features that are causing these anomalies (Schmidt 2019).

Using such classifications of anomalies it is also possible to identify positive and negative peaks that are within a given proximity to each other and that may jointly be the result of a bipolar anomaly. The data can subsequently be investigated further by fitting analytical bipole models to them (Norton and Witten 1998; Hinterleitner et al. 2019).

Conclusion

Vectorised geophysics data provide natively a variety of spatial attributes that would be far more difficult to derive from raster data. There are hence several vector-based classification schemes that allow for advanced analysis and visualisation of geophysics data using deterministic processing that can be engineered in accordance with the specific characteristics of the geophysical data types (e.g. magnetometer, GPR). The computing power provided by free vector processing engines like PostGIS makes such processing available to many researchers and practitioners, and the examples provided show what can be achieved with this approach. —

References

- Cornett RL, Ernenwein EG. Object-Based Image Analysis of Ground Penetrating Radar Data for Archaic Hearths. *Remote Sensing* 2020;12(16):2539. doi: [10.3390/rs12162539](https://doi.org/10.3390/rs12162539)
- Hinterleitner A, Totschnig R, Löcker K, Schiel H, Tobias B. Modelling the layer between topsoil and subsoil using magnetic prospection data. In Bonsall, J (ed.) *New Global Perspectives on Archaeological Prospection 2019*;282-285. Oxford: Archaeopress Publishing Ltd.
- Küçükdemirci M, Sarris A. GPR Data Processing and Interpretation Based on Artificial Intelligence Approaches: Future Perspectives for Archaeological Prospection. *Remote Sensing* 2022;14(14):3377. doi: [10.3390/rs14143377](https://doi.org/10.3390/rs14143377)
- Norton SJ, Witten AJ. Target Parameter Estimation and Survey Design in Magnetometry. *Journal of Environmental and Engineering Geophysics* 1998;3:147-155. doi: [10.4133/JEEG3.3.147](https://doi.org/10.4133/JEEG3.3.147)
- Schmidt A. Vector Processing and Visualisation of Archaeological GPR Data. In Hadjimitsis, DG, Themistocleous, K, Michaelides, S and Papadavid, G (eds) *Fourth International Conference on Remote Sensing and Geoinformation of Environment: 2016*; 27. Paphos, Cyprus: Cyprus Remote Sensing Society.
- Schmidt A. The limits of a blob: geophysically informed automatic extraction of magnetometer anomalies. In Bonsall, J (ed.) *New*

Global Perspectives on Archaeological Prospection 2019;272-273. Oxford: Archaeopress Publishing Ltd.

Schmidt A, Gunawardhana P, Davis C, Hale D, Coningham R, Pushparatnam P, Villis R, Woolston-Houshold M, Manuel M. Interpreting GPR data from Jaffna Fort, Northern Sri Lanka, using historic maps and new excavations. In Bonsall, J (ed.) *New Global Perspectives on Archaeological Prospection 2019*; 224-227. Oxford: Archaeopress Publishing Ltd.

Schmidt A, Tsetskhladze G. They may all be dead but they still ain't equal ... Recent Work in Archaeological Geophysics 2012;16-17. London, UK: Near-surface Geophysics Group.

Schmidt A, Tsetskhladze G. Raster was Yesterday: Using Vector Engines to Process Geophysical Data. *Archaeological Prospection* 2013;20(1):59-65. doi: [10.1002/arp.1443](https://doi.org/10.1002/arp.1443)

Sevara C, Pregebauer M, Doneus M, Verhoeven G, Trinks I. Pixel versus object - A comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data. *Journal of Archaeological Science: Reports* 2016;5:485-498. doi: [10.1016/j.jasrep.2015.12.023](https://doi.org/10.1016/j.jasrep.2015.12.023)

Verdonck L. Automated detection and analysis of diffraction hyperbolas in ground penetrating radar data. In Bonsall, J (ed.) *New Global Perspectives on Archaeological Prospection 2019*;300-308. Oxford: Archaeopress Publishing Ltd.

Open Access

This paper is published under the Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0/deed.en>). Please note that individual, appropriately marked parts of the paper may be excluded from the license mentioned or may be subject to other copyright conditions. If such third party material is not under the Creative Commons license, any copying, editing or public reproduction is only permitted with the prior consent of the respective copyright owner or on the basis of relevant legal authorization regulations.