Geophysical Research Abstracts
Vol. 17, EGU2015-2793, 2015
EGU General Assembly 2015

Detecting Buried Archaeological Remains by the Use of Geophysical Data Processing with 'Diffusion Maps' Methodology

Lev Eppelbaum
Tel Aviv University, Faculty of Exact Sciences, Dept. of Geosciences, Tel Aviv, Israel (levap@post.tau.ac.il)

Geophysical methods are prompt, non-invasive and low-cost tool for quantitative delineation of buried archaeological targets. However, taking into account the complexity of geological-archaeological media, some unfavourable environments and known ambiguity of geophysical data analysis, a single geophysical method examination might be insufficient (Khesin and Eppelbaum, 1997). Besides this, it is well-known that the majority of inverse-problem solutions in geophysics are ill-posed (e.g., Zhdanov, 2002), which means, according to Hadamard (1902), that the solution does not exist, or is not unique, or is not a continuous function of observed geophysical data (when small perturbations in the observations will cause arbitrary mistakes in the solution). This fact has a wide application for informational, probabilistic and wavelet methodologies in archaeological geophysics (Eppelbaum, 2014a).

The goal of the modern geophysical data examination is to detect the geophysical signatures of buried targets at noisy areas via the analysis of some physical parameters with a minimal number of false alarms and miss-detections (Eppelbaum et al., 2011; Eppelbaum, 2014b). The proposed wavelet approach to recognition of archaeological targets (**AT**) by the examination of geophysical method integration consists of advanced processing of each geophysical method and nonconventional integration of different geophysical methods between themselves. The recently developed technique of diffusion clustering combined with the abovementioned wavelet methods was utilized to integrate the geophysical data and detect existing irregularities. The approach is based on the wavelet packet techniques applied as to the geophysical images (or graphs) versus coordinates. For such an analysis may be utilized practically all geophysical methods (magnetic, gravity, seismic, GPR, ERT, self-potential, etc.).

On the first stage of the proposed investigation a few tens of typical physical-archaeological models (*PAM*) (e.g., Eppelbaum et al., 2010; Eppelbaum, 2011) of the targets under study for the concrete area (region) are developed. These *PAM* are composed on the basis of the known archaeological and geological data, results of previous archaeogeophysical investigations and 3D modeling of geophysical data. It should be underlined that the *PAM*s must differ (by depth, size, shape and physical properties of **AT** as well as peculiarities of the host archaeological-geological media). The *PAM*s must include also noise components of different orders (corresponding to the archaeogeophysical conditions of the area under study). The same models are computed and without the **AT**.

Introducing complex *PAM*s (for example, situated in the vicinity of electric power lines, some objects of infrastructure, etc. (Eppelbaum et al., 2001)) will reflect some real class of **AT** occurring in such unfavorable for geophysical searching conditions. Anomalous effects from such complex *PAM*s will significantly disturb the geophysical anomalies from **AT** and impede the wavelet methodology employment. At the same time, the "self-learning" procedure laid in this methodology will help further to recognize the AT even in the cases of unfavorable S/N ratio.

Modern developments in the wavelet theory and data mining are utilized for the analysis of the integrated data. Wavelet approach is applied for derivation of enhanced (e.g., coherence portraits) and combined images of geophysical fields. The modern methodologies based on the matching pursuit with wavelet packet dictionaries enables to extract desired signals even from strongly noised data (Averbuch et al., 2014).

Researchers usually met the problem of extraction of essential features from available data contaminated by a random noise and by a non-relevant background (Averbuch et al., 2014). If the essential structure of a signal consists of several sine waves then we may represent it via trigonometric basis (Fourier analysis). In this case one can compare the signal with a set of sinusoids and extract consistent ones. An indicator of presence a wave in a signal $f(t)$ is the Fourier coefficient $\int f(t)\ \sin wt\ dt$. Wavelet analysis provides a rich library of waveforms available and fast, computationally efficient procedures of representation of signals and of selection of relevant waveforms. The basic assumption justifying an application of wavelet analysis is that the essential structure of a signal analyzed consists of not a large number of various waveforms. The best way to reveal this structure is representation of the signal by a set of basic elements containing waveforms coherent to the signal. For structures of the signal coherent to the basis, large coefficients are attributed to a few basic waveforms, whereas we expect small coefficients for the noise and structures incoherent to all basic waveforms.

Wavelets are a family of functions ranging from functions of arbitrary smoothness to fractal ones. Wavelet procedure involves two aspects. The first one is a *decomposition*, i.e. breaking up a signal to obtain the wavelet coefficients and the 2^{nd} one is a *reconstruction*, which consists of a reassembling the signal from coefficients

There are many modifications of the WA. Note, first of all, so-called *Continuous WA* in which$s$ignal $f(t)$ is tested for presence of waveforms $\psi\left(\frac{t-b}{a}\right)$. Here, a is scaling parameter (dilation), $b$determines location of the wavelet $\psi\left(\frac{t-b}{a}\right)$ in a signal $f(t)$. The integral

$$(W_\psi f)\ (b,a) = \int f(t)\,\psi\left(\frac{t-b}{a}\right)\,dt$$

is the Continuous Wavelet Transform.When parameters a, b in $\psi\left(\frac{t-b}{a}\right)$ take some discrete values, we have the Discrete Wavelet Transform. A general scheme of the Wavelet Decomposition Tree is shown, for instance, in (Averbuch et al., 2014; Eppelbaum et al., 2014).

The signal is compared with the testing signal on each scale. It is estimated wavelet coefficients which enable to reconstruct the 1^{st} approximation of the signal and details. On the next level, wavelet transform is applied to the approximation. Then, we can present A_1 as $A_2 + D_2$, etc. So, if S – *Signal*, A – *Approximation*, D – *Details*, then

$$S = A_1 + D_1 = A_2 + D_2 + D_1 = A_3 + D_3 + D_2 + D_1.$$

Wavelet packet transform is applied to both low pass results (approximations) and high pass results (*Details*).

For analyzing the geophysical data, we used a technique based on the algorithm to characterize a geophysical image by a limited number of parameters (Eppelbaum et al., 2012). This set of parameters serves as a signature of the image and is utilized for discrimination of images (a) containing **AT** from the images (b) non-containing **AT** (let will designate these images as **N**). The constructed algorithm consists of the following main phases: (a) collection of the database, (b) characterization of geophysical images, (c) and dimensionality reduction. Then, each image is characterized by the histogram of the coherency directions (Alperovich et al., 2013). As a result of the previous steps we obtain two sets: containing **AT** and **N** of the signatures vectors for geophysical images. The obtained 3D set of the data representatives can be used as a reference set for the classification of newly arriving geophysical data.

The obtained data sets are reduced by embedding features vectors into the 3D Euclidean space using the so-called diffusion map. This map enables to reveal the internal structure of the datasets **AT** and **N** and to distinctly separate them. For this, a matrix of the diffusion distances for the combined feature matrix $\mathrm{F} = \mathrm{F}_N \cup \mathrm{F}_C$ of size 60 x C is constructed (Coifman and Lafon, 2006; Averbuch et al., 2010). Then, each row of the matrices F_N and F_C is projected onto three first eigenvectors of the matrix $\mathrm{D}(\mathrm{F})$. As a result, each data curve is represented by a 3D point in the Euclidean space formed by eigenvectors of $\mathrm{D}(\mathrm{F})$. The Euclidean distances between these 3D points reflect the similarity of the data curves. The scattered projections of the data curves onto the diffusion eigenvectors will be composed. Finally we observe that as a result of the above operations we embedded the original data into 3-dimensional space where data related to the **AT** subsurface are well separated from the **N** data. This 3D set of the data representatives can be used as a reference set for the classification of newly arriving data. Geophysically it means a reliable division of the studied areas for the **AT**-containing and not containing (**N**) these objects.

Testing this methodology for delineation of archaeological cavities by magnetic and gravity data analysis displayed an effectiveness of this approach.

References

Alperovich, L., Eppelbaum, L., Zheludev, V., Dumoulin, J., Soldovieri, F., Proto, M., Bavusi, M. and Loperte, A., 2013. A new combined wavelet methodology applied to GPR and ERT data in the Montagnole experiment (French Alps). Journal of Geophysics and Engineering, **10**, No. 2, 025017, 1-17.

Averbuch, A., Hochman, K., Rabin, N., Schclar, A. and Zheludev, V., 2010. A diffusion frame-work for detection of moving vehicles. *Digital Signal Processing*, **20**, No.1, 111-122.

Averbuch A.Z., Neittaanmäki, P., and Zheludev, V.A., 2014. *Spline and Spline Wavelet Methods with Applications to Signal and Image Processing*. Volume I: Periodic Splines. Springer.

Coifman, R.R. and Lafon, S., 2006. Diffusion maps, Applied and Computational Harmonic Analysis. *Special issue on Diffusion Maps and Wavelets*, **21**, No. 7, 5-30.

Eppelbaum, L.V., 2011. Study of magnetic anomalies over archaeological targets in urban conditions. *Physics and Chemistry of the Earth*, **36**, No. 16, 1318-1330.

Eppelbaum, L.V., 2014a. Geophysical observations at archaeological sites: Estimating informational content. *Archaeological Prospection*, **21**, No. 2, 25-38.

Eppelbaum, L.V. 2014b. Four Color Theorem and Applied Geophysics. *Applied Mathematics*, **5**, 358-366.

Eppelbaum, L.V., Alperovich, L., Zheludev, V. and Pechersky, A., 2011. Application of informational and wavelet approaches for integrated processing of geophysical data in complex environments. Proceed. of the 2011 SAGEEP Conference, Charleston, South Carolina, USA, **24**, 24-60.

Eppelbaum, L.V., Khesin, B.E. and Itkis, S.E., 2001. Prompt magnetic investigations of archaeological remains in areas of infrastructure development: Israeli experience. *Archaeological Prospection*, **8**, No.3, 163-185.

Eppelbaum, L.V., Khesin, B.E. and Itkis, S.E., 2010. Archaeological geophysics in arid environments: Examples from Israel. *Journal of Arid Environments*, **74**, No. 7, 849-860.

Eppelbaum, L.V., Zheludev, V. and Averbuch, A., 2014. Diffusion maps as a powerful tool for integrated geophysical field analysis to detecting hidden karst terranes. *Izv. Acad. Sci. Azerb. Rep., Ser.: Earth Sciences*, No. 1-2, 36-46.

Hadamard, J., 1902. Sur les problèmes aux dérivées partielles et leur signification physique. Princeton University Bulletin, **13**, 49-52.

Khesin, B.E. and Eppelbaum, L.V., 1997. The number of geophysical methods required for target classification: quantitative estimation. Geoinformatics, **8**, No.1, 31-39.

Zhdanov, M.S., 2002. *Geophysical Inverse Theory and Regularization Problems*. Methods in Geochemistry and Geophysics, Vol. 36. Elsevier, Amsterdam.