Combination of Maximin and Kriging Prediction Methods for Eddy-Current Testing Database Generation

Sándor BILICZ[1,2], Marc LAMBERT[1], Emmanuel VAZQUEZ[3] and Szabolcs GYIMÓTHY[2]

[1] Département de Recherche en Électromagnétisme, Laboratoire des Signaux et Systèmes UMR8506 (CNRS-SUPELEC-Univ Paris-Sud), Gif-sur-Yvette, France

[2] Budapest University of Technology and Economics, Hungary

[3] École Supérieure d'Électricité (SUPELEC), Gif-sur-Yvette, France

Abstract

The numerical simulation of eddy-current non-destructive testing methods involves high complexity and expensive computational load. However, one needs to reach reliable solutions for these problems in order to be able, in particular, to solve the related inverse problem. A way to overcome such a difficulty is to propose a “surrogate modelization” of ECT experiments through the generation of a problem-specific database (for a given ECT application), containing the calculated probe responses (“data”) for well-chosen defect configurations (at certain values of “defect parameters”). Surrogate models attempt to imitate the behaviour of the rigorous model but remain much simpler (thus faster to run). In other words, the expensive ”simulator” is replaced by a cheap “emulator” in the inversion process. The emulator is always based on some “true” data (computed by the simulator). The choice of these samples strongly influences the performance of the emulation and then requires some sophisticated design-of-experiment (DOE) methods. In our context, DOE attempts to achieve an “optimal sampling” of the given ECT experiment.

However, the construction of such databases is not a simple task. The database should be “optimal” in some sense. Optimality is a rather general property. In our context, it means that the database achieves a prescribed precision when one applies a given interpolation method (based on the stored data points), while the size of the database remains as small as possible. Nowadays, the generation of optimal databases is quite challenging and an intensively studied field in non-destructive evaluation. To the best of our knowledge, the most preferred approach of database generation so far is the application of mesh generation methods (mostly originated from finite-element meshing techniques), see for instance [1, 2].

The present contribution provides a new methodology of optimal database generation. Let us introduce the following notation: the n^{th} entry of the database consists of a D-dimensional parameter vector $\mathbf{p}_n$ (the defect is characterized by D parameters), and an M-dimensional data vector $\mathbf{z}_n$ ($\mathbf{z}_n$ can be complex, e.g., the impedance variation of a probe coil). The parameter vectors, which relate to all possible defects that one expects to find in the given experiment, span the parameter space P. The data vectors, which are related to the parameter vectors in the parameter space, span the data space Z. The basic idea of our method is to construct a database, whose data points $\mathbf{z}_n$ (n=1, 2, …, N) are as spread as possible over Z. In other words, the data space Z is uniformly sampled by the database.

The proposed methodology is a combination of a *maximin* method (a well-known tool in decision theory, game theory, statistics) and a *kriging* prediction (a stochastic method for interpolating functions, based on some observed function values, using a Gaussian process to model the functions). A close idea appears in a quite different context in [3] but the

methodology presented therein does not work in the case of high-dimensional data spaces, like in our ECT application (note that M might be several hundred).

The sought uniform output sample distribution is attempted to achieve by using the *maximin* method. In our context, *maximin* means maximizing the minimal distance between any pair of output samples. In other words, a sample distribution is aimed where the distance of all output samples from their nearest neighbour is as large as possible. We propose an incremental algorithm to generate databases of this *maximin* property. Once the distance function (usually depending on the direct solver) has been choosen the method consists in mainly two stages. At the beginning, some initial samples are chosen by hand. Then, samples are inserted one-by-one according to a *maximin* criterion applied onto the choosen distance function. The latter has to be computed for a large number of input space's samples which is numerically expensive, then a *kriging* approach is used to predict the distance functions at unobserved locations.

To illustrate the method numerical simulations will be discussed and comparison of the performances of the databases generated by such a method and a "classical" method will be carried out. Preliminary results show that the proposed algorithm spreads the data vectors on the data space better than some conventional sampling methods.

Acknowledgments

This research is partially supported by the French ANR through the "INDIAC"project and by the DIGITEO cluster's project N° 2008-15D. Laboratoire des Signaux et Systèmes et SUPELEC are founding members of the DIGITEO cluster.

References

1. J. Pávó, Sz. Gyimóthy, "Design method for optimal database used for inverse problems in ECT," in T. Sollier, D. Prémel, D. Lesselier (Eds.), Electromagnetic Nondestructive Evaluation (VIII), IOS Press, Amsterdam, 2004, pp. 69-76.
2. J. Pávó, Sz. Gyimóthy, "Adaptive inversion database for electromagnetic nondestructive evaluation," NDT&E International, vol. 40, no. 3, 2007, pp. 192-202.
3. R. Bettinger, P. Duchêne, L. Pronzato, E. Thierry, "Design of experiments for response diversity," J. Phys.: Conf. Ser. 135, 012017, 2008.