Journal of Computational Science 12 (2016) 28–37

Contents lists available at ScienceDirect

Journal of Computational Science

journal homepage: www.elsevier.com/locate/jocs

A multi-fidelity surrogate-model-assisted evolutionary algorithm for computationally expensive optimization problems

Bo Liu [a,*], Slawomir Koziel [b,c], Qingfu Zhang [d]

[a] Department of Computing, Glyndwr University, UK
[b] Engineering Optimization & Modeling Center, Reykjavik University, Iceland
[c] Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Poland
[d] School of Computer Science & Electronic Engineering, University of Essex, UK

ARTICLE INFO

Article history:
Received 11 June 2015
Received in revised form 31 October 2015
Accepted 15 November 2015
Available online 27 November 2015

Keywords:
Multi-fidelity
Multilevel
Variable fidelity
Surrogate-model-assisted evolutionary algorithm
Expensive optimization

ABSTRACT

Integrating data-driven surrogate models and simulation models of different accuracies (or fidelities) in a single algorithm to address computationally expensive global optimization problems has recently attracted considerable attention. However, handling discrepancies between simulation models with multiple fidelities in global optimization is a major challenge. To address it, the two major contributions of this paper include: (1) development of a new multi-fidelity surrogate-model-based optimization framework, which substantially improves reliability and efficiency of optimization compared to many existing methods, and (2) development of a data mining method to address the discrepancy between the low- and high-fidelity simulation models. A new efficient global optimization method is then proposed, referred to as multi-fidelity Gaussian process and radial basis function-model-assisted memetic differential evolution. Its advantages are verified by mathematical benchmark problems and a real-world antenna design automation problem.

1. Introduction

Solving many real-world engineering design problems requires both global optimization and expensive computer simulation for evaluating their candidate solutions. For example, computational models utilized in photonics and microelectromechanical system optimization require well over 1 h simulation time per design [1–3]. For these problems, successful surrogate-model-assisted local search methods have been developed [2], but in terms of global optimization, many state-of-the-art surrogate-based optimization techniques are still prohibitively expensive. In the context of global optimization, these tasks are considered as very expensive. Addressing such problems is the objective of this paper.

High cost of evaluating real-world simulation models often results from the necessity of solving complex systems of partial differential equations using numerical techniques or Monte-Carlo analysis. Direct handling of such models is often computationally prohibitive and utilization of cheaper representations (surrogates) of the system at hand might be necessary. Two classes of such replacement models are normally considered. The first type is function approximation model (usually, data-driven surrogates constructed by approximating sampled simulation model data, e.g., radial basis function). Optimization methods making use of such models are often referred to as surrogate/metamodel-based optimization (SBO) methods [4]. The other type is low-fidelity simulation model (e.g., coarse-mesh model in finite element analysis), which exhibits relaxed accuracy but shorter evaluation time. Low-fidelity model is typically used with occasional reference to the high-fidelity model. The methods using such models are often referred to as multi-fidelity/multilevel/variable-fidelity optimization (MFO) methods [5]. For simplicity, two-level modeling is considered in this paper: the coarse model is referred to as the low-fidelity model, whereas the fine model is referred to as the high-fidelity model.

Recently, a trend to combine SBO and MFO in a single algorithm for further speed improvement has been observed; successful examples include [6–9]. Refs. [6,9] demonstrated an approach which iteratively updates a co-kriging surrogate model [10] using samples from coarse and fine model evaluations accumulated over the entire optimization process. These techniques are mathematically sound and often feature good reliability. However, the success of these methods relies on a high-quality co-kriging surrogate model constructed by the initial one-shot sampling, which

* Corresponding author.
E-mail addresses: b.liu@glyndwr.ac.uk (B. Liu), koziel@ru.is (S. Koziel), qzhang@essex.ac.uk (Q. Zhang).

http://dx.doi.org/10.1016/j.jocs.2015.11.004

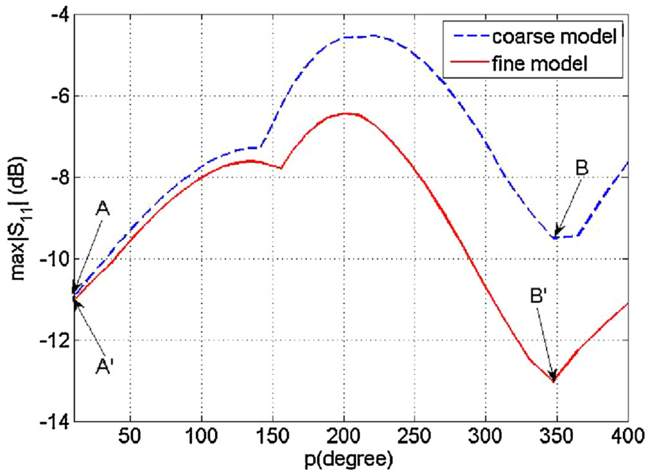

Fig. 1. Curves of coarse and fine models.

determines the effectiveness of the consecutive adaptive sampling. For higher-dimensional design spaces or complex function landscapes, the computational cost of building the initial high-quality co-kriging model may be prohibitive, as the necessary number of training samples grows exponentially with linear increase of the number of design variables [11].

In order to alleviate these difficulties, a new hierarchical algorithm structure has been proposed in [7,8]: It can be considered as an MFO, but SBOs are used for some optimization stages with certain fidelities to replace standard optimization methods without data-driven surrogate models. For example, a coarse model is used for a surrogate model-assisted evolutionary algorithm to explore the space and accurate but expensive fine model evaluations are only used for local search starting from the most promising solutions obtained from space exploration [7,12]. These methods are scalable if proper SBOs are used, but the reliability of the MFO structure becomes a challenge, which is detailed as follows.

Because both the coarse and the fine model describe the same function (or a physical system), it is reasonable to use the cheaper coarse model for filtering out some non-optimal subregions. However, considering discrepancy between the models of various fidelities, there is a lot of uncertainty regarding the "promising" solutions found using the coarse model. Fig. 1 illustrates this issue using an example of a microwave filter. The problem is to minimize the maximum value of the reflection coefficient, i.e., $max(|S_{11}|)$ for a given frequency range of interest. It can be observed by sweeping one of the nine design variables of the device, that although many non-optimal regions for the coarse model are also non-optimal for the fine model, the two critical challenges appear:

- wasting precious fine model evaluations because some promising locations of the coarse model-based landscape (such as *A* and its projection *A'*, which may correspond to multiple design variables) may have substantial distance to the desired optimal regions of the fine model-based landscape (like *B'*);
- making the MFO framework unreliable because the desired optimum such as the point *B'* is difficult to be reached from points such as *A'* by exploitation. Note that only one variable is changed in Fig. 1. When considering multiple design variables, the point *B* may have a low probability to be selected for exploitation because there may be quite a few points with better fitness values according to the coarse model.

Clearly, the lower the fidelity of the coarse model, the higher the efficiency of the space exploration stage of an MFO, but the higher the risk induced by model discrepancy. Ref. [8] investigates the discrepancy problem using practical antenna design cases and indicates that, in many cases, a large number of fine model evaluations may be needed which may result in the same or even higher overall design optimization cost than that of direct optimization of the fine model; also, the MFO may simply fail to find a satisfactory design. There are some methods that do not directly use the promising points from the coarse model optimization, include equalization of the computational effort for models of each fidelity [5], space mapping and model correction, where a correction function is applied to reduce misalignment between the coarse and the fine model [8,13]. However, the above two critical challenges still remain.

To address the above two challenges, a new MFO framework is proposed in this paper. Its goal is to make full use of available expensive fine model evaluations and substantially improve the reliability compared to existing MFO frameworks, and thus, addressing the targeted very expensive design optimization problems. Based on this framework, a data mining method is proposed to address the discrepancy between the coarse and the fine model. A new method, referred to as multi-fidelity Gaussian process and radial basis function-model-assisted memetic differential evolution (MGPMDE), is subsequently proposed. Empirical studies on mathematical benchmark problems with different characteristics as well as a real-world antenna design automation problem verify the advantages of MGPMDE.

The remainder of this paper is organized as follows. Section 2 formulates the optimization problem and introduces the basic techniques. Section 3 describes the new MFO framework, the data mining method and the MGPMDE algorithm. Section 4 presents the experimental results of MGPMDE on test problems. Concluding remarks are provided in Section 5.

2. Problem formulation and basic techniques

2.1. Problem formulation

We consider the following problem:

$$\begin{aligned} &\min \quad f_f(x) \\ &x \in [\bar{a}, \bar{b}]^d \end{aligned} \tag{1}$$

where $f_f(x)$ is the fine model function, which is expensive but accurate. There is a coarse model function, $f_c(x)$, which is much cheaper than $f_f(x)$, but less accurate than $f_f(x)$, and, consequently, with a distorted landscape. Ref. [5] provides an effective method to construct mathematical benchmark problems for MFO, which is as follows.

$$f_f(x) = f_c(s_f \times (x - s_s)) \tag{2}$$

where $f_c(x)$ (also $f_f(x)$) is a periodic function, and there exist minimal and maximal values in each period. s_f is called a frequency shift, which mimics the loss of peaks of $f_c(x)$. For example, when $f_c(x) = cos(x)$, $f_f(x)$ can be $cos(s_f \times (x))$. When s_f is set to 1.3, about 30% of the peaks are not accounted for by $f_c(x)$. s_s is called a spatial shift, which shifts the positions of the optimal points. The frequency shifts and the spatial shifts often happen for expensive evaluations obtained by solving suitable systems of partial differential equations, where the coarse model is a coarse-mesh model and/or with reduced number of solver iterations. This kind of expensive optimization problem is very (if not the most) popular in computationally expensive engineering design optimization, because most physics simulations (e.g., electromagnetic simulation) are based on solving partial differential equations.

It is worth to determine the focused extent of discrepancy before proposing methods to address it. From the point of view of practical industrial problems [2,5,14,15], we focus on reasonably large discrepancy between computational models of various fidelities in