Applied Soft Computing xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Applied Soft Computing

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

Improving surrogate-assisted variable fidelity multi-objective optimization using a clustering algorithm

Yan Liu, Matthew Collette*

Department of Naval Architecture and Marine Engineering, University of Michigan, 2600 Draper Drive, Ann Arbor, MI 48109, USA

A R T I C L E I N F O

Article history:
Received 31 July 2013
Received in revised form 23 June 2014
Accepted 28 July 2014
Available online xxx

Keywords:
Clustering
Evolutionary computation
Surrogate model
Kriging
Optimization
Variable fidelity optimization

A B S T R A C T

Surrogate-assisted evolutionary optimization has proved to be effective in reducing optimization time, as surrogates, or meta-models can approximate expensive fitness functions in the optimization run. While this is a successful strategy to improve optimization efficiency, challenges arise when constructing surrogate models in higher dimensional function space, where the trade space between multiple conflicting objectives is increasingly complex. This complexity makes it difficult to ensure the accuracy of the surrogates. In this article, a new surrogate management strategy is presented to address this problem. A *k*-means clustering algorithm is employed to partition model data into local surrogate models. The variable fidelity optimization scheme proposed in the author's previous work is revised to incorporate this clustering algorithm for surrogate model construction. The applicability of the proposed algorithm is illustrated on six standard test problems. The presented algorithm is also examined in a three-objective stiffened panel optimization design problem to show its superiority in surrogate-assisted multi-objective optimization in higher dimensional objective function space. Performance metrics show that the proposed surrogate handling strategy clearly outperforms the single surrogate strategy as the surrogate size increases.

Introduction

In contemporary engineering design, the ability to rapidly understand trade-offs between multiple conflicting objectives is emphasized. Multi-objective genetic algorithms (MOGA) are able to find the Pareto front in a single optimization run, making them attractive to solve this type of problems. A non-dominated sorting genetic algorithm (NSGA-II) proposed by Deb et al. [1], along with other MOGAs published, have proved to be very robust in converging to the true Pareto optimal sets. In recent years, MOGA have seen wide spread applications to marine structures optimization. Despite displaying robust performance in multi-objective optimization, genetic algorithms are often slow to execute, since genetic algorithms can require a large amount of fitness function evaluations to determine domination status. This is especially true when time-consuming high fidelity models are used for objective functions. Surrogate-assisted MOGA optimization is a potential solution to this problem. In surrogate-assisted optimization, computationally efficient approximation models are constructed to replace the original computationally expensive fitness functions. This article introduces a technique of using *k*-means clustering in variable-fidelity optimization to assist surrogate model construction as surrogate model size increases.

Several successful surrogate model techniques have been reported in literature to date. These include the response surface methodology [2], artificial neural networks [3], and Kriging models [4]. Kriging was originally presented to assist optimization in Sacks's work [4]. Kneijien [5] provides a more recent review paper for Kriging models used in optimization. Owing to its stochastic process framework, Kriging models provide both an estimation of fitness value and a model error estimation. Due to this error estimation property, Kriging models are popularly employed as a surrogate model in optimization.

To replace an objective function, the surrogate model must be accurate and efficient. Various model managements in surrogate-assisted evolutionary algorithms have been developed for this task. In a previous paper, the authors put forward a novel variable fidelity optimization (VFO) method [6], in which a Kriging surrogate model was constructed online to scale a low fidelity version of the fitness function to a high fidelity version of the same function. The VFO method schedules the high-fidelity simulations in given generations so that it uses only a fixed computational budget while converging close to the true Pareto front. This method was demonstrated on the ZDT standard test problems as well as a two-objective structural design problem. However, the previous method becomes inefficient and struggles to converge when the number of objectives is increased beyond two. A key reason for this struggle is the increasingly large size of the Kriging model needed for larger problems.

In the development of the VFO approach, the Kriging model is constructed around the evolving Pareto set in the optimization run. As the optimization problem moves to a higher number of objectives, the location of non-dominated solutions in the independent variable space becomes more diverse. To maintain accuracy, more points are required causing the surrogate model size to consequently increase. The challenge can increase when the number of independent variables increases as well. During the Kriging modeling process, an N^*N matrix (N is the total number of data points) will be inverted. Thus computational cost grows quickly with sample size. Jin et al. [7] have indicated that Kriging model construction can be very time-consuming for large sample sizes. In addition, solving an extremely large Kriging model can be numerically unstable as the matrix become nearly singular, in which case, the Kriging predictions are unlikely to be reliable.

* Corresponding author at: Department of Naval Architecture and Marine Engineering, University of Michigan, Room 235, 2600 Draper Drive, Ann Arbor, MI 48109, USA. Tel.: +1 734 764 8422; fax: +1 734 936 8820.
E-mail addresses: mdcoll@umich.edu, lyforgood@gmail.com (M. Collette).

http://dx.doi.org/10.1016/j.asoc.2014.07.022
1568-4946/© 2014 Published by Elsevier B.V.

Researchers have suggested using multiple surrogates in place of the single surrogate model to improve prediction quality. Jin and Sendhoff [8] have proposed using neural network ensembles to improve the performance of surrogate-assisted evolutionary optimization. In the works of Goel et al. [9] and Sanchez et al. [10], the benefits of using multiple surrogates have also been reported empirically. Hamza and Saitou [11] have used polynomial surrogate ensembles in genetic algorithm for vehicle crash-worthiness design. Isaacs et al. [12] used spatially distributed multiple radius basis function surrogates for multi-objective optimization. Within their study, a fraction of the total sample points in each sub-surrogate were used for modeling, while the rest points were used for accuracy validation. However, there is not a solution proposed to deal with the large surrogate modeling problem that variable fidelity optimization scheme has faced.

In this article, multiple surrogate models are used to improve the ability of the proposed VFO method to tackle larger problems. However, some means to determine how to split a single large Kriging model into multiple Kriging models is needed. Clustering is an attractive technique for this purpose and has been widely employed for data mining. In clustering, a large dataset is separated into subsets that are in some sense more closely related to each other than the other members of the overall data set. The advantage of clustering is that the algorithm can perform this separation without external guidance, making it ideal for inclusion in an automated optimization procedure. A *k*-means clustering algorithm is employed to partition the Kriging sample dataset, then multiple Kriging models are built in each of these partitions. This approach helps avoid solving large Kriging models, thus keeps the surrogate-assisted optimization efficient. The proposed clustering implemented method is believed to be helpful in a broader means of large sampling size surrogate model management problems other than Kriging.

The remainder of the paper is organized as follows. Next section outlines the basic of variable fidelity optimization approach with the new proposed clustering algorithm implemented Kriging modeling method. Subsequent section examines the proposed method using a series of benchmarking optimization problems, including comparisons with problems solved by the previous version of the method. *Stiffened panel design* section shows the implementation of the proposed method in a three-objective structure design problem. Conclusions and future work are discussed in last section.

Variable fidelity optimization using multiple Kriging surrogates

Overview

The approach proposed by Zhu et al. [6] is used here as surrogate model management framework to facilitate multi-objective optimization. Usually in the real-world design optimization problems, there are various simulation functions for fitness evaluation with different levels of fidelity. As high-fidelity evaluations are more time-consuming, there is a need to trade fidelity with computational cost in optimization design. In the author's previous work [6], a variant on Haftka's [13] global-local approach was shown to work well in approximating the Pareto front with fewer high-fidelity fitness function calls. In the proposed variable-fidelity approach, the high-fidelity function $f_h(x)$ could be approximated by a global simplified structure method $f_g(x)$ and a Kriging correction model $f_f(x)$, as shown in Eq. (1)

$$f_h(x) = f_g(x) * f_f(x) \tag{1}$$

In this formulation, we proposed that the global approximation mathematical model is a simplified function that runs rapidly with a relatively high coefficient of variation (COV) in the bias of its prediction. The bias of the prediction is defined as:

$$\text{bias} = \frac{\text{predicted}}{\text{actual}} \tag{2}$$

This leads to a bias of 1.0 for perfect approximation, and a bias below or above 1.0 for less accurate estimation methods. The VFO formulation takes advantage of the fact that the genetic algorithm starts off with random population, so lower fidelity can be used to evolve a rough Pareto front. While in this process, interpolation surrogate model can be constructed online. Afterwards, the surrogate model can be used to transition the rough Pareto front to an approximation of the true Pareto front. The detailed implementation of the proposed VFO with clustered multiple Kriging models will be introduced in the following paragraphs.

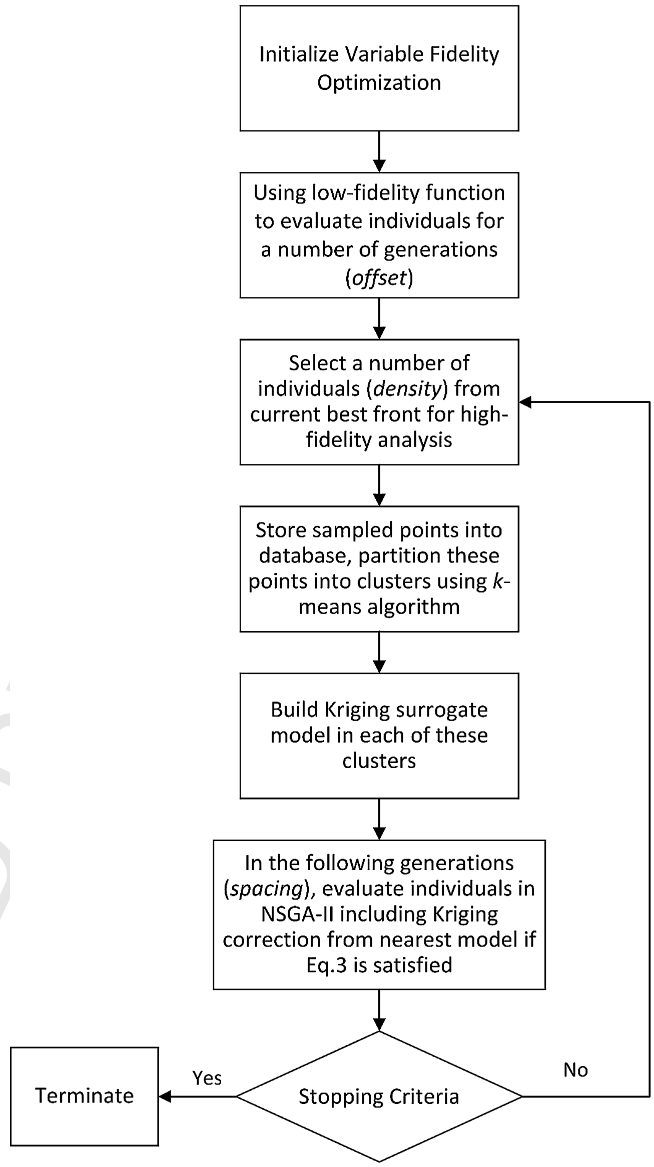

Fig. 1. Flow chart of the method.

Revised variable fidelity scheme

The variable fidelity-updating scheme proposed in the previous section was implemented in the standard NSGA-II multi-objective genetic algorithm [1]. The VFO strategy used here is based on that which was proposed in Zhu et al. [6], parts of the description from that reference are summarize below. The major difference is in the Kriging model construction, where multiple Kriging surrogates are built with the help of *k*-means clustering algorithm. Fig. 1 shows the steps of the method to be followed, each step is explained in detail below. The numerical implementation details of the techniques employed for each of these steps is explained in subsequent subsections.

A detailed process of the revised VFO updating strategy using multiple Kriging models is described as follows:

1. To initialize the optimization, a random population is selected. The first few generations only use the low-fidelity model for fitness evaluation. A rough Pareto front is evolved based on the low fidelity model.