Fourier Analysis of the Fitness Landscape for Evolutionary Search Acceleration

Pei, Yan
Graduate School of Design, Kyushu University

Takagi, Hideyuki
Faculty of Design, Kyushu University

http://hdl.handle.net/2324/1808440
Fourier Analysis of the Fitness Landscape for Evolutionary Search Acceleration

Yan PEI
Graduate School of Design
Kyushu University
Fukuoka, 815-8540 Japan
peiyan@kyudai.jp

Hideyuki TAKAGI
Faculty of Design
Kyushu University
Fukuoka, 815-8540 Japan
http://www.design.kyushu-u.ac.jp/~takagi/

Abstract—We propose an approach for approximating a fitness landscape by filtering its frequency components in order to accelerate evolutionary computation (EC) and evaluate the performance of the technique. In addition to the EC individuals, the entire fitness landscape is resampled uniformly. The frequency information for the fitness landscape can then be obtained by applying the discrete Fourier transform (DFT) to the resampled data. Next, we filter to isolate just the major frequency component; thus we obtain a trigonometric function approximating the original fitness landscape after the inverse DFT is applied. The elite is obtained from the approximated surface and used to accelerate the EC search.

Index Terms—evolutionary computation, Fourier transform, fitness landscape, acceleration of convergence, function approximation, filtering

I. INTRODUCTION

Accelerating evolutionary computation (EC) search is an important research topic, and many papers on this topic have been presented [15], [9]. Some of the approaches explored are: altering the EC coding strategy [4], [19], [11], [2], population constitution [21], developing new EC operators [12], [6], and using hybrid and fusion technology [22], [14]. When we develop new methods for accelerating EC, it is important not only to devise new EC operators but also to consider the strategies behind their EC operations. The fitness landscape is one of the important sources of information for supporting these operations and strategies such that the EC search runs correctly.

The Fourier transform is a powerful mathematical tool for analyzing the frequency information of signals using orthogonal trigonometric functions and has been used in signal processing [10], bioinformatics [17], and many other areas. The Fourier transform can also be used in EC as a mathematical tool for studying an EC model, resolving the complexity of a fitness landscape and accelerating EC convergence. It was used to obtain the model parameters of polynomial harmonic models and used for genetic programming in [7]. Taylor series and Fourier series were respectively used for local and global approaches to analyzing the fitness landscape in [18]. EC was also used as an optimization tool to research time-frequency analysis in signal processing [3]. However, little research literature has reported on EC fitness landscape analysis and EC algorithm design where convergence is accelerated using the Fourier transform.

The objective of this paper is to propose a method that analyzes a fitness landscape by considering points on it as signal samples, obtaining the frequency characteristics of same, approximating the fitness landscape by filtering its primary frequency component, and finally accelerating the EC search by obtaining elite from the approximated landscape. Concretely speaking, the method resamples a fitness landscape at uniform intervals, calculates the corresponding fitness values, applies a discrete Fourier transform (DFT) to them, filters only the primary frequency component(s), applies the inverse DFT to the filtered frequency components to obtain an approximate fitness landscape, obtains rough location information for the global optimum from the approximated landscape, and subsequently uses it to accelerate the EC search.

The key points of this proposed method are finding the primary frequency component(s) and filtering it (them) to approximate the original fitness landscape with one or more trigonometric function. We also propose an approach for global approximation, which approximates the whole landscape, and for local approximation, which estimates the landscape around the best individual. Both sampling methods obtain the elite from the approximated surface and use it to accelerate the EC search.

We use a DFT to analyze a fitness landscape on which we can then define primary frequency components. The primary frequency components are used to construct an approximate model for the EC acceleration in section II. We also propose two resampling methods in section II to obtain corresponding frequency components; one approximates the whole fitness landscape and the other a local area in the landscape. We evaluate the proposed methods by applying differential evolution (DE/best/1/bin) to eight benchmark functions in section III and analyze the proposed methods and discuss their future possibilities in section IV. Finally, we conclude this research and describe our future research directions in section V.
II. Fourier Analysis of a Fitness Landscape and Approximation of the Landscape Using Trigonometric Functions

A. Concept of the Proposed Method

EC search is based on the fitness of the individuals. As we do not use information about the EC search surface but rather the fitness of a limited numbers of individuals distributed in the search space, less obtained search information is a restriction on our ability to extending the EC search capability. In a complex search space, the limited number of fitness values cannot adequately express the search space characteristic and direct the EC interaction process.

If we can obtain EC fitness landscape information during EC search, EC search performance and applicability can be improved and extended. Obtaining EC fitness landscape information is a promising area for research [9]. Jin has investigated fitness landscape approximation approaches and basis evolution management strategies [5].

Fig. 1 shows the flow diagram of our proposed method. Frequency characteristics of a fitness landscape are obtained by resampling a search space at regular intervals and applying the DFT to a sequence of fitness values for the resampled points. We can approximate the original fitness landscape with a trigonometric function by filtering a primary frequency component and applying inverse DFT to the filtered frequency components. Our proposed method that aims to accelerate EC search using information of an approximated function obtained by Fourier transform can be considered as a regression model for the trigonometric functions (Eq. 1).

\[ EC(X) = \sum_{i=0}^{N} a_i \sin(2\pi\omega_i X + B_i) \]  

B. Fourier Transform of a Fitness Landscape

Although individuals generated by EC operations are distributed at irregular intervals within the search space, samples used for a DFT must be uniformly sampled. The search space must therefore be resampled, and then their fitness values must be recalculated.

In obtaining the frequency characteristics of a fitness landscape, there are two issues we must first decide: (1) the number of resampling points, (2) the sampling period and DFT base frequency points, i.e. the interval between sample points, which decides frequency resolution.

First, we cannot perform a DFT without deciding on the number of sample points \( M \). The more sample data we use, the higher the frequency resolution we can obtain. However, in deciding upon the number of sample points, we must take into account the balance between computational cost and convergence speed under the conditions of the application task because fitness values must be recalculated for the resampled data and the computational cost \( (M \log 2M) \) of the Fast Fourier Transform (FFT) depends on the number of sample points \( M \). In our experiment in this paper, we set \( M = 16 \) and applied the FFT to a sequence of these 16 fitness values.

Regarding sampling dimension, there are two possible resampling methods we can use. In one method we resample the search space separately for each dimension, and in the other we resample the whole \( n \)-D search space \((n-D sampling)\) at once. Additionally, we have two choices regarding where we choose to resample; in one method we resample the whole search space (global sampling) and in the other we only resample the area around the best individual (local sampling). As a preliminary experiment, we evaluate how these resampling methods influence the approximation of the fitness landscape in the next section.
The second issue, sampling period, depends on the # of sampling points and the sampling area (area to be approximated) which is determined by the the first issue above. Sampling over the entire searching range is one approach, and sampling over a range of distributed individuals and narrowing the range according to the narrowing range of its distribution is another. A third technique would limit the sampling area.

Because both global and local properties are valuable to the EC search, we design two resampling methods. In the first method, the entire search space is uniformly sampled in each dimension to obtain resampled data. We call this GLB. In the second, the resample area is centered on the area around the best individual. We call this LOC. These are respectively our global and local sampling approaches.

The third issue to resolve is in which dimension we should approximate the fitness landscape. If each dimension is being resampled separately, there are an additional two options that must be considered. If it is sufficient to resample in just one dimension, we can resample the 1-D space with a uniform sampling interval and then use parameter values for the best individuals in the other corresponding dimensions. If we are to obtain the full multi-dimensional landscape, we use GLB or LOC. In this case there are a total of 4 ways the sampling methods can be combined (4 = 2 × 2). These combinations are also evaluated below.

C. Filtering in the Frequency Domain

The trigonometric functions that determine the main structure of a fitness landscape are the most important in the regression model of (Eq. 1), and they are determined by the amplitude and phase information at the peaks in the power spectrum. Let us define them as principal frequency components and call them the principal frequency component, the second, and so on according to their maximum power ranking.

Although the original fitness landscape is not always periodical with 2π, we can universally approximate the local landscape with arbitrary accuracy if we use an arbitrary number of trigonometric functions. In practice, however, too many base functions introduce too much computational complexity, so in this paper we will approximate the original fitness landscape with just use one trigonometric function, i.e., we use only the first principal frequency component.

The principal frequency component describes the main structure of a fitness landscape and its shape, and depends in turn on the choice of sampling range and sampling frequency. Table I shows the experimentally obtained principal frequency components for the benchmark functions in Table II. Due to the different sampling methods and sampling spaces, the principal frequency component of each benchmark functions is not the same.

1) Even for the same search space, it cannot be said in general that each dimension has the same landscape; the principal frequency component determined using 1-D sampling (∧ and ♦) may not agree.

<table>
<thead>
<tr>
<th>Func.</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>F2</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>F3</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>F4</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>F5</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>F6</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>F7</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>F8</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
</tbody>
</table>

2) For each benchmark function, most of the principal frequency components determined by 1-D sampling locally and globally (∧ and ♦) are the same, which demonstrates that we can obtain the same original fitness landscape by 1-D sampling locally and globally.

3) For n-D local and global sampling (∧ and ♦), the principal frequency components obtained are different for all of the each eight benchmark functions, and when we use n-D local sampling (∧) to approximate the original fitness landscape, we are at risk of not obtaining correct results.

D. Approximation of a Fitness Landscape by Inverse DFT and the Acceleration of the EC Search

Once a concrete regression model of trigonometric functions is obtained by applying the inverse DFT to the frequency characteristics containing only the filtered principal frequency components, the regression model is used as an approximation of the local or global EC fitness landscape in 1-D dimension or n-D dimensions. With this approximated landscape characteristic, we can obtain more EC search information than that which is directly provided by the individuals who’s population size is limited.

There are several approaches for using the obtained landscape approximation information to accelerate EC search. For example, in one approach we can determine the EC search direction by analyzing the approximated fitness landscape and in another approach can obtain new elites from the approximated landscape and use them for the EC search in the next generation [16]. For optimization problems which involve finding the maximum point, we expect that the global point should be around the points of \( k(1/4ω) + b \) \((x > 0)\) and \( k(3/4ω) + b \) \((x < 0)\); whereas for optimization problems which involve finding the minimum point, we expect that the global point should be around the points of \( k(3/4ω) + b \) \((x > 0)\) and \( k(1/4ω) + b \) \((x < 0)\). We may be able to accelerate the EC search by using the global point as the elite for the next generation. So if we can locate the individuals with this characteristic in the corresponding original search space and put them into the next generation, it should be possible to accelerate the EC. We use this acceleration method using elites obtained from an approximated function in our experimental...
evaluation in the next section.

III. EXPERIMENTAL EVALUATION

A. Experimental Condition

We use the DeJong five standard functions (F1 - F5) [1], Rastrigin function (F6), Schwefel function (F7) and Griewank function (F8) [20] as benchmark functions to evaluate the proposed approaches. The dimensions and search ranges for all parameters are listed in Table II. All these function optimization tasks are posed as minimization problems with the optimal solution being the point with the lowest value. Their fitness landscapes have a variety of characteristics. They include both continuous and discontinuous, convex and non-convex, unimodal and multimodal, and low dimensional, variable separable and non-separable, and high dimensional shapes.

We compare the EC acceleration method proposed in this paper with our previously proposed acceleration method [8] and normal non-accelerated EC as references, using the same benchmark functions.

B. Optimization Methods and Sign Test

We use differential evolution (DE/best/1/bin) as our optimization method to compare the proposed approach with the reference conventional ones. The approaches are applied to 8 benchmark functions for up to 50 generations with 50 trial runs, and a sign test is used to determine if there is a significant difference.

Here, we abbreviate the four variations of our DFT-based methods proposed in this paper as DE-FR-GLB-1D, DE-FR-GLB-nD, DE-FR-LOC-1D, and DE-FR-LOC-nD. These refer, respectively, to DE with elite obtained by global sampling in 1-D, global sampling in n-D, local sampling in 1-D, and local sampling in n-D. We use the same abbreviations in the other EC acceleration approaches we proposed in [8], i.e. DE-LR and DE-LS mean DE with elite obtained by a two-degree Lagrange interpolation and a line power function least squares approximation. Finally, normal DE without any acceleration method is referred to as DE-N. These abbreviations are also used in Figures 2 and 3.

Figure 2 shows the average convergence curves for the best fitness values over 50 trial runs of these methods, and Figure 3 shows the sign test results between DE-N and DE with each acceleration method at each generation.

C. Experimental Results

The following observations can be made from these experimental results.

1) Our proposed methods were able to accelerate the EC well for all benchmark functions except F2.
2) The proposed methods did not accelerate DE convergence well for F2.
3) The performances of the four proposed methods look similar and their relative superiority depends on the task if there even is a difference between their performances.
4) The proposed method, DE-FR-GLB-nD, demonstrated better performance than all the other proposed methods, i.e. global sampling is the best method for obtaining the whole fitness landscape characteristic and efficiently finding the global optimum.
5) Our proposals in this paper have shown better performance for F3, F5, F6, F7, F8 than the acceleration approaches which we previously outlined in [8].

IV. DISCUSSION

We obtained information about a fitness landscape using four proposed methods (DE-FR-GLB-nD, DE-FR-GLB-1D, DE-FR-LOC-nD or DE-FR-LOC-1D) and approximated the landscape by filtering its principal frequency component. From the comparison between global and local sampling methods, we can say that DE-FR-GLB-nD and DE-FR-GLB-1D are suitable for global exploration over a whole search space and DE-FR-LOC-nD and DE-FR-LOC-1D are suitable for local exploration in the local search space near the best individual, because different sampling methods obtain different frequency and phase information from the global/local and 1 dimension/n dimensions.

From the experimental results, the global sampling methods, DE-FR-GLB-nD and DE-FR-GLB-1D, demonstrated better acceleration performance. This is because they can obtain more accurate frequency information about a given fitness
landscapes than the other methods. Even though the current best individuals are in local optimum areas, we can acquire information about the global optimum from DE-FR-GLB-nD and DE-FR-GLB-1D.

From average convergence curves and sign test results of F1, F4, F6, F7, and F8, the local sampling methods, DE-FR-LOC-nD and DE-FR-LOC-1D, are useful for problems in which the fitness landscape has local valleys.

The shape of the global optimum area of F2 is a long, narrow, parabolic, and flat valley. As each of its dimensions has this similar shape, our proposed methods are unavailable for problems with this kind of shape. Although this valley shape looks trivial, convergence to the global optimum is difficult; our proposed methods seem inefficient and leave room to be improved.

F4 is a function with Gaussian noise. As noise made by a Gaussian random number is white Gaussian noise, its spectrum is flat and therefore does not influence the principal frequency components. Consequently, our proposed method can reduce the effect of the noise significantly. Our experimental results indicate that our proposed methods can significantly accelerate F4 in most generations. Local areas of F4 are not approximated well by a cosine curve and significant performance improvement was not obtained for F4 when DE-FR-LOC-nD was used.

For multimodal problems, such as F5, our methods can accelerate the convergence in the initial generations. However, when individuals migrate towards a single local optimum and their diversity is decreased, the performance of the elite acceleration drops. When individuals gather at one point over the generations, the frequency spectrum obtained from the area covered by the individuals no longer has enough information to approximate the fitness landscape.

The objective of this paper is to reduce the complexity of a fitness landscape by obtaining an approximation of it from the frequency characteristics of the fitness landscape. Simplifying
a fitness landscape by reducing the noise that makes a fitness landscape complex realizes this objective, and the subsequent global and local exploration is the biggest contribution of our proposed method for accelerating EC convergence. Experimental results have shown that our proposed methods can accelerate EC.

V. CONCLUSION AND FUTURE WORKS

We proposed a family of landscape approximation methods that analyze the frequency characteristics of a fitness landscape using a DFT, obtain principal frequency components, filter the components, and apply the inverse DFT to the filtered frequency characteristics. Our methods accelerate EC by substituting an elite obtained from the approximated landscape for the worst EC individual. We evaluated the methods with eight benchmark functions and demonstrated that this strategy of approximating fitness landscapes is effective for accelerating EC search. We also analyzed the performance of our proposed methods and discussed them in detail.

The original contribution of this paper is to direct our attention to the frequency characteristics of fitness landscapes, to use this observation to simplify the landscapes, and to apply our simplification so that we can accelerate EC search.

We can increase the precision with which we approximate fitness landscapes and improve acceleration performance by using elites that are located closer to the global optimum by using multiple primary frequency components, although we used only one primary frequency component in our experiments in this paper. Using multiple primary frequency components will be the topic of future research.

It may be possible to extend the proposed approach into evolutionary multi-objective optimization, although this is currently only our expectation. That is, unlike with the conventional Pareto concept, we may be able to handle multiple objectives in one domain, i.e. the frequency domain, by projecting each objective landscape to its corresponding frequency domain. A further investigation of this idea will form the basis of another future study.

TABLE II

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Test function</th>
<th>Range</th>
<th>n</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Sphere</td>
<td>(f(x) = \sum_{i=1}^{n} x_i^2)</td>
<td>[-5.12,5.12]</td>
<td>3</td>
<td>US</td>
</tr>
<tr>
<td>F2</td>
<td>Rosenbrock</td>
<td>(f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2)</td>
<td>[-2.048,2.048]</td>
<td>2</td>
<td>UN</td>
</tr>
<tr>
<td>F3</td>
<td>DeJong-Step</td>
<td>(f(x) = \sum_{i=1}^{n}</td>
<td>x_i</td>
<td>)</td>
<td>[-5.12,5.12]</td>
</tr>
<tr>
<td>F4</td>
<td>Quantic &amp; Noise</td>
<td>(f(x) = \sum_{i=1}^{n} x_i^4 + Gauss(0,1))</td>
<td>[-1.28,1.28]</td>
<td>30</td>
<td>US</td>
</tr>
<tr>
<td>F5</td>
<td>Shekel’s Foxholes</td>
<td>(f(x) = 0.02 + \sum_{j=1}^{25} \sum_{i=1}^{n} \frac{1}{\sqrt{(x_i - a_{ij})^2 + 1}})</td>
<td>[-65.536,65.536]</td>
<td>2</td>
<td>MS</td>
</tr>
<tr>
<td>F6</td>
<td>Rastrigin</td>
<td>(f(x) = (10n) + \sum_{i=1}^{n} x_i^2 - 10\cos(2\pi x_i))</td>
<td>[-5.12,5.12]</td>
<td>5</td>
<td>MS</td>
</tr>
<tr>
<td>F7</td>
<td>Schwefel 2.26</td>
<td>(f(x) = \sum_{i=1}^{n} -x_i \sin(\sqrt{</td>
<td>x_i</td>
<td>}))</td>
<td>[-512,512]</td>
</tr>
<tr>
<td>F8</td>
<td>Griewank</td>
<td>(f(x) = 1 + \sum_{i=1}^{n} \frac{x_i^2}{4000} - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{2}}\right))</td>
<td>[-512,512]</td>
<td>5</td>
<td>MN</td>
</tr>
</tbody>
</table>

ACKNOWLEDGMENT

This work was supported in part by Grant-in-Aid for Scientific Research (23500279). Yan Pei would like to thank the Yoshida Scholarship Foundation for its support of his doctoral research.

REFERENCES


