Research Article

Polygonal Approximation Using an Artificial Bee Colony Algorithm

Shu-Chien Huang

Department of Computer Science, National Pingtung University, Pingtung City 90003, Taiwan

Correspondence should be addressed to Shu-Chien Huang; schuang@mail.nptu.edu.tw

Received 8 August 2014; Revised 11 December 2014; Accepted 15 December 2014

Academic Editor: Debasish Roy

Copyright © 2015 Shu-Chien Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A polygonal approximation method based on the new artificial bee colony (NABC) algorithm is proposed in this paper. In the present method, a solution is represented by a vector, and the objective function is defined as the integral square error between the given curve and its corresponding polygon. The search process, including the employed bee stage, the onlooker bee stage, and the scout bee stage, has been constructed for this specific problem. Most experiments show that the present method when compared with the DE-based method can obtain superior approximation results with less error norm with respect to the original curves.

1. Introduction

Polygonal approximation is a common representation of digital curves [1]. A polygonal representation of a digital curve can significantly reduce the amount of data needed to be processed while at the same time preserving important information about the curve. It can be applied in shape analysis, pattern classification, image understanding, 3D reconstruction, and CAD applications. In recent decades, a number of methods have been proposed for approximation. In general, they can be categorized into three approaches: (1) given the number of breakpoints, \( m \), find the approximation with \( m \) breakpoints such that the approximation error is minimized; (2) given an error \( \varepsilon \), find the approximation with the minimum number of breakpoints that is distant from the given curve by no more than \( \varepsilon \); and (3) dominant point-detection approaches. In this paper, the present method is categorized into the first approach for polygonal approximation.


In recent years, researchers have attempted to apply nature-inspired methods to the problem of pattern recognition and image processing. In this paper, the new artificial bee colony (NABC) algorithm [11] is employed for approximation of digital curves with line segments. The present method can output solutions very close to the optimal one in a relatively short time. The remainder of this paper is organized as
follows. In Section 2, the problem is illustrated. In Section 3, the NABC algorithm is described. In Section 4, the NABC algorithm for polygonal approximation is presented. In Section 5, the experimental studies for the comparison of the present method with the DE-based method are shown. Finally, conclusions are given in Section 6.

2. Problem Statement

Given a curve with \( n \) clockwise-ordered points, denoting \( C = \{ p_0, p_1, p_2, \ldots, p_{n-1} \} \), let \( \overline{p_x p_y} \) and \( \overline{p_x p_{y'}} \) represent the arc and the chord starting at point \( p_x \) and continuing to point \( p_y \) in the clockwise direction along the original curve \( C \). The approximation error of line segment \( \overline{p_x p_y} \), denoted as \( e_{x,y} \), is defined as

\[
e_{x,y} = \sum d^2 (p_x, p_x p_y), \quad \forall p_1 \in \overline{p_x p_y},
\]

where \( d(p_x, \overline{p_x p_y}) \) is the perpendicular distance of arc point \( p_x \) to the chord \( \overline{p_x p_y} \).

The objective of the polygonal approximation problem is to find a polygon, assuming that the set of its breakpoints is \( C^* = \{ p_{a(1)}, p_{a(2)}, \ldots, p_{a(m)} \} \), where \( 0 < a(1) < a(2) < \cdots < a(m) \leq n-1 \), such that the integral square error between the polygon and the given curve is minimized. The integral square error (ISE) is defined as follows:

\[
ISE = e_{a(1),a(2)} + e_{a(2),a(3)} + \cdots + e_{a(m-1),a(m)} + e_{a(m),a(1)}. \tag{2}
\]

Therefore, there exists \( C_{n,m} \) \((= n!/m!(n-m)!)]\) combinations to consider when searching for the \( m \) breakpoints of the given curve. Generally, this problem is not trivial and always turns into a time-consuming iterative search.

3. The New Artificial Bee Colony (NABC) Algorithm

The artificial bee colony (ABC) algorithm [12–14] was proposed by Karaboga for optimizing numerical problems. Karaboga and Ozturk [15] applied it to training feed-forward neural networks to classify nine data sets. Xu and Duan [16] proposed an artificial bee colony algorithm with edge potential function to accomplish the target recognition task for aircraft. Zhang et al. [17] used the ABC algorithm for clustering. Zhu and Kwong [18] proposed the Gbest-guided ABC (GABC) algorithm, which takes advantage of the information of the global best solution to guide the search for new candidate solutions in order to improve the exploitation. A set of numerical benchmark functions were tested, and the results show that the GABC algorithm can outperform the ABC algorithm in most of their experiments. Pan et al. [19] proposed a discrete artificial bee colony (DABC) algorithm to solve the lot-streaming flow shop scheduling problem. Horng [20] proposed the multilevel thresholding algorithm for image segmentation based on the technology of the artificial bee colony algorithm, while Gao et al. [21] proposed a modified ABC algorithm, which is based on each bee searching only around the best solution of the previous iteration in order to improve the exploitation. Draa and Bouaziz [22] used an artificial bee colony algorithm for image contrast enhancement. In their method, the grey levels of the input image are replaced by a new set of grey levels. Talatahari et al. [23] presented a new optimization method based on the artificial bee colony algorithm for solving parameter identification problems. Sarkar et al. [24] presented a perturbed martingale approach to global optimization, and the results show that their method appears to substantively improve upon the performance of the particle swarm optimization method.

The ABC algorithm simulates the intelligent foraging behavior of honey bee swarms. The colony of artificial bees contains three groups of bees: employed bees, onlooker bees, and scout bees. In the ABC algorithm, the first half of the colony consists of employed bees while the second half is constituted of the onlooker bees. A bee associated with a food resource is called an employed bee, while a bee waiting in the dance area to get information concerning food sources is called an onlooker bee. An employed bee whose food source is exhausted by the employed and onlooker bees becomes a scout. The scout bee carries out random searches to discover new sources. The number of food sources, denoted by SN, is equal to the number of employed bees. The position of a food source represents a possible solution to the optimization problem, while the nectar amount of a food source corresponds to the quality (fitness) of the associated solution. Each solution \( x_i \) \((i = 1, 2, \ldots, SN)\) is an \( m \)-dimensional vector. Here, \( m \) is the number of optimization parameters and is also the number of breakpoints of a polygon in this paper.

In [11], an improved version of ABC, called NABC, was proposed. A set of twelve benchmark functions were used in their experiments. The simulation results showed that the NABC algorithm is significantly better or at least comparable to the original ABC algorithm. In the initialization stage, NABC generates a randomly distributed initial population of SN solutions (food source positions). The fitness of all solutions is evaluated. Then, a solution pool is constructed by storing the best 100p% solutions in the current swarm with \( p \in (0, 1) \).

In the employed bee stage, the original ABC algorithm generates a candidate solution \( v_i \) in the neighborhood of its present solution \( x_i \) as follows:

\[
 v_{i,j} = x_{i,j} + \phi_j \left( x_{i,j} - x_{k,j} \right), \tag{3}
\]

where \( k \) is a randomly produced index which is different from \( i \) and \( j \) is a random dimension index selected from the set \( \{1, 2, \ldots, m\} \), and \( \phi_j \) is a random number in the range \([-1, 1]\). In [II], the new ABC/best/1 strategy was proposed in the employed bee stage. The NABC algorithm generates a candidate solution \( v_i \) as follows:

\[
 v_{i,j} = x_{i,best,j}^p + \phi_j \left( x_{r1,j} - x_{r2,j} \right), \tag{4}
\]

where \( x_{i,best,j}^p \) is randomly chosen from the solution pool and \( x_{r1,j} \), \( x_{r2,j} \) are two randomly selected solutions from the current swarm with \( i \neq r1 \neq r2 \).

In the onlooker bee stage, an artificial onlooker bee chooses a food source depending on the probability value
associated with that food source, \(\text{prob}_i\), calculated by the following expression:

\[
\text{prob}_i = 0.9 \times \frac{\text{fit}(x_i)}{\max\{\text{fit}(x_1), \text{fit}(x_2), \ldots, \text{fit}(x_{\text{SN}})\}} + 0.1,
\]

where \(\text{fit}(x_i)\) is the fitness value of the solution \(x_i\). In the scout bee stage, if a solution \(x_i\) is abandoned, then the scout bee discovers a new food source to replace \(x_i\).

The difference between the NABC algorithm and the original ABC algorithm is that the NABC algorithm constructs a solution pool and a candidate solution \(v_i\) is generated by using (4) instead of (3). In the NABC approach, bees can search the neighborhood of different best solutions. This can help avoid fast attraction and is helpful for accelerating the convergence speed.

The flowchart of the NABC algorithm is given in Figure 1. The pseudocode of the NABC algorithm is shown in Algorithm 1.

**Algorithm 1**

4. The Proposed NABC Algorithm for Polygonal Approximation

In this section, the representation of solutions and the fitness function are described, and the search process of the NABC algorithm for accomplishing the task of polygonal approximation is also presented.

4.1. Representation of Solutions. In the initialization stage, the SN solution \(x_i (i = 1, 2, \ldots, SN)\) is generated, where \(x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,m})\), \(x_{i,j}\) is the real number, and the
0 ≤ x_{i,j} < x_{i,j+1} < n for all j. Then, the value of cycle is set to 0 and the trial number of each solution x_{i}, trial_{i}, is equal to 0.

4.2. Fitness Function. The solution x_{i} = (x_{i,1}, x_{i,2}, \ldots, x_{i,m}) corresponds to a polygon and the set of its breakpoints is \{p_{a(1)}, p_{a(2)}, \ldots, p_{a(m−1)}, p_{a(m)}\}, where a(j) = floor(x_{i,j}), j ∈ \{1, 2, \ldots, m\}. First, the integral square error (ISE) is computed by using (2). Then, the fitness value can be calculated by the following expression:

\[ \text{fit}(x_{i}) = \frac{1}{(\text{ISE} + 1)}. \] (6)

4.3. The Search Process

4.3.1. The Employed Bee Stage. In this stage, the ith (i = 1, 2, \ldots, SN) employed bee produces a new solution v_{i} by using (4) and computes the fitness value of the new solution. To ensure that v_{i,j} < v_{i,j+1}, the elements of each solution vector are sorted in ascending order of magnitude. If the fitness of the new solution v_{i} is higher than that of the old solution x_{i}, the solution x_{i} is replaced by v_{i} and trial_{i} is set to 0; otherwise, the old solution x_{i} is kept and trial_{i} is increased by 1.

4.3.2. The Onlooker Bee Stage. First, the probability value prob_{i} is calculated by using (5) for the solution x_{i}. Each onlooker bee which selects a solution x_{i} depending on prob_{i} also produces a new solution v_{i} by using (4). Then, it applies the greedy selection process between v_{i} and x_{i}. As described in Section 4.3.1, if the fitness of the new solution v_{i} is higher than that of the old solution x_{i}, the solution x_{i} is replaced by v_{i} and trial_{i} is set to 0; otherwise, the old solution x_{i} is kept and trial_{i} is increased by 1.

4.3.3. The Scout Bee Stage. As described above, if the solution x_{i} is not improved through the employed stage and the onlooker bee stage, the trial_{i} value of solution x_{i} will be increased by 1. If the trial_{i} of solution x_{i} is more than the parameter “limit,” the solution x_{i} is considered to be an abandoned solution; meanwhile, the employed bee will be changed into a scout. The scout randomly produces the new solution to replace x_{i}. Then, the value of trial_{i} is set 0, and this scout is changed into an employed bee.

4.3.4. Check the Termination Criterion. If the cycle is greater than the maximum cycle number (MCN), then the algorithm is finished and the best solution found so far is output.

5. Experimental Results

The present method is experimentally tested and compared with the DE-based method [8]. The program is coded in C language and run on a PC with a Pentium 4 CPU. The control
The DE algorithm is an evolutionary algorithm for solving numerical problems. The NABC algorithm is a novel swarm intelligent algorithm inspired by the foraging behaviors of a honeybee colony. The two methods have attracted the attention of researchers and have been widely used in solving engineering optimization problems. The main difference between the two methods is that they adopt different strategies to produce new solutions. The two methods start with an initial population of SN solutions. In the DE algorithm, new solutions are generated through the operations of mutation, crossover, and selection. In the NABC algorithm, as shown in Figure 1, new solutions are generated after performing a solution pool construction, employed bee stage, onlooker bee stage, and scout bee stage.

In order to verify the performance of the present method, four digital curves, namely, a figure-of-eight curve (see Figure 2(a)), a chromosome-shaped curve (see Figure 3(a)), a leaf-shaped curve (see Figure 4(a)), and a semicircle-shaped curve (see Figure 5(a)), are applied to be tested. The simulation conducted twenty independent runs. The best solutions, average solutions, and the standard deviations (S.D.) of the solutions are reported. The simulation results for the figure-of-eight curve, the chromosome-shaped curve, the leaf-shaped curve, and the semicircle-shaped curve are listed in Tables 1, 2, 3, and 4, respectively. For the average solutions, the present method outperforms the DE-based method. For the best solutions, it is shown that the present method produces better polygonal approximations than the DE-based method in most cases. In addition, the low S.D. reveals that the present method has higher stability than the DE-based method.

In Figures 2–5, Parts (b)–(d) show the best results obtained by the present method with different numbers of breakpoints employed. The computation times for the four digital curves are listed in Table 5. The present method can output solutions immediately because the processing time of all cases is less than 1 second in the experiments.
To demonstrate the applicability of the present method, it is applied to the screwdriver-shaped curve (see Figure 6(a)) and the fish-shaped curve (see Figure 7(a)). The maximum cycle number (MCN) is set to 1,000 for these two relatively larger curves. For the screwdriver-shaped curve, the point number of this curve is 267. Figures 6(b) and 6(c) show the best approximated polygon obtained through the present method with breakpoint numbers of 40 and 45, respectively. The computation times for Figures 6(b) and 6(c) are 2.547 s and 2.656 s, respectively. For the fish-shaped curve, the point number is 537. Figures 7(b) and 7(c) show the best approximated polygon obtained through the present method with breakpoint numbers of 50 and 60, respectively. The computation times for Figures 7(b) and 7(c) are 4.671 s and 4.891 s, respectively. The results of the approximation are tabulated in Table 6. It is seen that the present method can also give good approximation results.

### Table 1: Experimental results for the figure-of-eight curve.

<table>
<thead>
<tr>
<th>Number of breakpoints</th>
<th>DE-based method [8]</th>
<th>Present method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISE (best)</td>
<td>ISE (average)</td>
</tr>
<tr>
<td>6</td>
<td>17.494</td>
<td>17.494</td>
</tr>
<tr>
<td>11</td>
<td>2.869</td>
<td>3.235</td>
</tr>
<tr>
<td>13</td>
<td>2.042</td>
<td>2.509</td>
</tr>
</tbody>
</table>

### 6. Conclusions

A new approach based on the new artificial bee colony (NABC) algorithm has been developed for approximation of digital curves with line segments. Given the number of breakpoints, the algorithm finds an approximation with the
Figure 5: The approximated polygons for the semicircle-shaped curve obtained by the present method. (a) Original curve and (b)–(d) results using the present method.

Table 2: Experimental results for the chromosome-shaped curve.

<table>
<thead>
<tr>
<th>Number of breakpoints</th>
<th>DE-based method [8]</th>
<th>Present method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISE (best)</td>
<td>ISE (average)</td>
</tr>
<tr>
<td>6</td>
<td>23.730</td>
<td>23.730</td>
</tr>
<tr>
<td>12</td>
<td>6.403</td>
<td>6.905</td>
</tr>
<tr>
<td>15</td>
<td>4.232</td>
<td>4.533</td>
</tr>
</tbody>
</table>

Table 3: Experimental results for the leaf-shaped curve.

<table>
<thead>
<tr>
<th>Number of breakpoints</th>
<th>DE-based method [8]</th>
<th>Present method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISE (best)</td>
<td>ISE (average)</td>
</tr>
<tr>
<td>12</td>
<td>80.834</td>
<td>103.944</td>
</tr>
<tr>
<td>17</td>
<td>42.400</td>
<td>47.354</td>
</tr>
<tr>
<td>19</td>
<td>30.862</td>
<td>36.158</td>
</tr>
</tbody>
</table>
Table 4: Experimental results for the semicircle-shaped curve.

<table>
<thead>
<tr>
<th>Number of breakpoints</th>
<th>DE-based method [8]</th>
<th>Present method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISE (best)</td>
<td>ISE (average)</td>
</tr>
<tr>
<td>15</td>
<td>18.531</td>
<td>20.731</td>
</tr>
<tr>
<td>19</td>
<td>12.519</td>
<td>13.377</td>
</tr>
<tr>
<td>22</td>
<td>9.542</td>
<td>10.023</td>
</tr>
</tbody>
</table>

The method has been tested with four different well-known curves. The present method is efficient in dealing with the application because the processing time of all cases is less than 1 second. Compared with the DE-based method, the present method can obtain better approximation results in most cases for both average solutions and best solutions. For the screwdriver-shaped curve and the fish-shaped curve, the present method can also obtain good approximation results. From these experiments, these results have demonstrated the feasibility of the present method for polygonal approximation of digital curves with line segments.
Figure 7: The approximated polygons for the fish-shaped curve obtained by the present method. (a) Original curve and (b)-(c) results using the present method.

Table 5: Processing time for finding various approximation results.

<table>
<thead>
<tr>
<th>Approximation</th>
<th>Number of breakpoints</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2(b)</td>
<td>6</td>
<td>0.216</td>
</tr>
<tr>
<td>Figure 2(c)</td>
<td>11</td>
<td>0.262</td>
</tr>
<tr>
<td>Figure 2(d)</td>
<td>13</td>
<td>0.272</td>
</tr>
<tr>
<td>Figure 3(b)</td>
<td>6</td>
<td>0.273</td>
</tr>
<tr>
<td>Figure 3(c)</td>
<td>12</td>
<td>0.319</td>
</tr>
<tr>
<td>Figure 3(d)</td>
<td>15</td>
<td>0.346</td>
</tr>
<tr>
<td>Figure 4(b)</td>
<td>12</td>
<td>0.544</td>
</tr>
<tr>
<td>Figure 4(c)</td>
<td>17</td>
<td>0.581</td>
</tr>
<tr>
<td>Figure 4(d)</td>
<td>19</td>
<td>0.609</td>
</tr>
<tr>
<td>Figure 5(b)</td>
<td>15</td>
<td>0.488</td>
</tr>
<tr>
<td>Figure 5(c)</td>
<td>19</td>
<td>0.525</td>
</tr>
<tr>
<td>Figure 5(d)</td>
<td>22</td>
<td>0.562</td>
</tr>
</tbody>
</table>

Table 6: Experimental results for the screwdriver-shaped curve and the fish-shaped curve.

<table>
<thead>
<tr>
<th>Curve</th>
<th>Number of breakpoints</th>
<th>ISE (best)</th>
<th>ISE (average)</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screwdriver-shaped curve</td>
<td>40</td>
<td>44.489</td>
<td>47.636</td>
<td>1.562</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>35.475</td>
<td>38.240</td>
<td>1.786</td>
</tr>
<tr>
<td>Fish-shaped curve</td>
<td>50</td>
<td>88.722</td>
<td>93.215</td>
<td>2.657</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>72.864</td>
<td>76.574</td>
<td>2.323</td>
</tr>
</tbody>
</table>

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

 References


Submit your manuscripts at http://www.hindawi.com