J. Vis. Commun. Image R. 30 (2015) 106–116

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Novel method to obtain the optimal polygonal approximation of digital planar curves based on Mixed Integer Programming ☆

E.J. Aguilera-Aguilera *, A. Carmona-Poyato, F.J. Madrid-Cuevas, R. Muñoz-Salinas

Department of Computing and Numerical Analysis, Córdoba University, 14071-Córdoba, Spain

ARTICLE INFO

Article history:
Received 27 November 2014
Accepted 22 March 2015
Available online 30 March 2015

Keywords:
Polygonal approximation
Digital planar curve
Mixed Integer Programming
Discrete optimization
Dominant points
Breakpoints
Integral square error
Optimization

ABSTRACT

Polygonal approximations of digital planar curves are very useful for a considerable number of applications in computer vision. A great interest in this area has generated a huge number of methods for obtaining polygonal approximations. A good measure to compare these methods is known as Rosin's merit. This measure uses the optimal polygonal approximation, but the state-of-the-art methods require a tremendous computation time for obtaining this optimal solution.

We focus on the problem of obtaining the optimal polygonal approximation of a digital planar curve. Given N ordered points on a Euclidean plane, an efficient method to obtain M points that defines a polygonal approximation with the minimum distortion is proposed.

The new solution relies on Mixed Integer Programming techniques in order to obtain the minimum value of distortion. Results, show that computation time for the new method dramatically decreases in comparison with state-of-the-art methods for obtaining the optimal polygonal approximation.

1. Introduction

Polygonal approximations of digital planar curves are important for a great number of applications, for example image analysis [14], shape analysis [9] and digital cartography [7]. For this reason, a great interest in this area has generated a huge number of methods for obtaining polygonal approximations [22,4,3].

The polygonal approximation problem can be formulated as two different optimization problems [12]:

- min-#: Given a distortion threshold obtain the polygonal approximation with the minimum number of points M.
- min-ε: Minimize the distortion error for a polygonal approximation given the number of points M of the approximation.

Methods in the literature can be classified into suboptimal and optimal, depending on the type of the solution obtained, that is, optimal algorithms guarantee optimal polygonal approximations, whereas, suboptimal methods do not assure the optimality of the solution obtained.

Usually, optimal methods to obtain polygonal approximations are used to compare suboptimal methods. This is done by using the error associated to the optimal polygonal approximation.

The associated error measures the distortion introduced by the polygonal approximation regarding the original digital planar curve. The distortion decreases as the number of points of the polygonal approximation increase, but many of these points may be redundant, that is, they do not provide relevant information. Therefore, a good balance between the distortion and the number of points is a desirable characteristic for polygonal approximations. Due to these two objectives are opposite, achieving a good balance between them is complicated.

A widely used measure of the distortion associated to the polygonal approximation is known as Integral Square Error (ISE). Let's suppose that a curve S has been approximated using a segment composed of points s_i and s_j, then, a distortion measure $\Delta(i,j)$ can be defined as follows:

$$\Delta(i,j) = \sum_{k=i}^{j} d(s_k, \overline{s_i s_j})^2 \quad (1)$$

where the term $d(s_k, \overline{s_i s_j})$ indicates the orthogonal distance from the point s_k to the segment $\overline{s_i s_j}$. Fig. 1 shows a curve that is approximated using a segment $\overline{s_i s_j}$, therefore the distortion $\Delta(i,j)$ is the sum of the squared orthogonal distances from points between point i and j to segment $\overline{s_i s_j}$, that is, $\Delta(i,j) = d_1^2 + d_2^2 + d_3^2$. The distortion

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author. Fax: +34 957218630.

E-mail addresses: i22agage@uco.es (E.J. Aguilera-Aguilera), ma1capoa@uco.es (A. Carmona-Poyato), ma1macuf@uco.es (F.J. Madrid-Cuevas), in1musar@uco.es (R. Muñoz-Salinas).

http://dx.doi.org/10.1016/j.jvcir.2015.03.007
1047-3203/

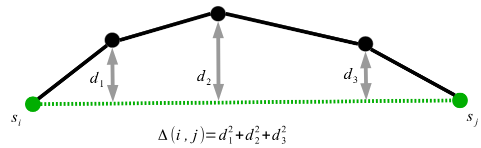

Fig. 1. Distortion sum of an approximated line segment.

measure ISE associated to a polygonal approximation with M points can be defined as follows:

$$\mathrm{ISE} = \sum_{k=0}^{M-1} \Delta(i_k, j_k) \quad (2)$$

where $i_k = j_{k-1}$ (modular arithmetic is assumed). Thus, the distortion associated to each of the M segments of the polygonal approximation are added. The original curve has associated a value of ISE equal to zero. The minimum number of points, that form a polygonal approximation with a value of distortion of zero, are named as breakpoints in the literature.

To compare the different methods to obtain polygonal approximations, Rosin [18] introduced two components named as fidelity and efficiency. Fidelity defines how well the polygonal approximation fits the optimal polygonal approximation in terms of the approximation error. Efficiency measures how compact is the polygonal approximation supplied regarding the optimal polygon with the same error.

Fidelity is defined as:

$$\mathrm{Fidelity} = \frac{\mathrm{E_{opt}}}{\mathrm{E_{approx}}} \times 100 \quad (3)$$

where $\mathrm{E_{approx}}$ is the distortion (ISE) of the polygonal approximation and $\mathrm{E_{opt}}$ is the distortion of the optimal polygonal approximation. Both distortion values, $\mathrm{E_{approx}}$ and $\mathrm{E_{opt}}$, are obtained for the same number of points.

Efficiency is defined as:

$$\mathrm{Efficiency} = \frac{\mathrm{M_{opt}}}{\mathrm{M_{approx}}} \times 100 \quad (4)$$

where $\mathrm{M_{approx}}$ is the number of segments of the polygonal approximation and $\mathrm{M_{opt}}$ is the number of segments that an optimal polygonal approximation would require to obtain the same error.

These two measures may vary depending on the curve. In order to avoid this problem Rosin [18] used a geometric mean of these two measures named as merit. The Rosin's merit is defined as:

$$\mathrm{Merit} = \sqrt{\mathrm{Fidelity} \times \mathrm{Efficiency}} \quad (5)$$

The advantage of using Rosin's merit as a measure for comparing algorithms to obtain polygonal approximations is, that these polygonal approximations may have a different number of points. Therefore, all algorithms can be evaluated in a fair way.

To compute Rosin's merit, the optimal polygonal approximation is required. For this purpose, the method proposed by Perez and Vidal [17] is used. This algorithm relies on the Dynamic Programming (DP) technique, to obtain the optimal polygonal approximation with a fixed number of points M for a digitized planar curve with N points. The main drawback of this method is the complexity, $O(MN^2)$ for closed curves. This complexity also increases because the initial point of the polygonal approximation has to be set as a parameter to the method [17]. Therefore, to obtain the optimal polygonal approximation, the method must try all points as initial point. Taking into account this problem, final complexity for obtaining the optimal polygonal approximation for a closed curve is $O(MN^3)$.

Some authors has proposed great improvements over the Dynamic Programming method for reducing the computation time. Horng and Li [11] proposed a method to determine the initial point of the polygonal approximation. This heuristic method needs two iterations of the Dynamic Programming algorithm to construct a polygonal approximation. The algorithm does not assure that the solution obtained is optimal.

Another attempt to reduce the computation time was introduced by Salotti [20]. This method used the A* algorithm to search in a graph formed by the points of the curve. This graph has a root node, which is the starting point of the curve. Therefore, this solution needs to try all points as initial point (initial node in the graph) to obtain the optimal solution. Nevertheless, this solution has a complexity close to $O(N^2)$[20], where N is the number of points of the curve.

Masood [15] proposed another framework of optimization. This method selects an initial set of points and deletes one point per iteration depending on the error associated to this point. After removing the point a local optimization process search the optimal position of the remaining points that minimizes the distortion. This process does not guarantee that the solution obtained is optimal.

In this paper a new method to optimally solve the min-ε problem is proposed. The new method relies on Mixed Integer Programming (MIP) technique and has some advantages over previous algorithms: (1) no initial point is needed to be set as a parameter, (2) time required to compute optimal solution is significantly lower than the state-of-the-art alternatives and (3) and the solution obtained is optimal.

The rest of this paper is structured as follows: Section 2 describes the proposed method. Section 3 describes the experiments carried out and results obtained by the proposed method. Section 4 discusses some relevant aspects of the proposed method and experimental results. Finally, Section 5 shows the main conclusions.

2. MIP model formulation

The problem of obtaining the optimal polygonal approximation of a planar curve has been solved as an optimization problem, using mainly dynamic programming techniques. We propose to state the polygonal approximation problem as a Mixed Integer Programming problem (MIP).

A MIP problem is a mathematical problem in which an objective function has to be minimized or maximized and is subject to a set of linear constraints. MIP problems may contain a subset of the variables that has also the constraint of being integer.

The problem formulation has an objective function defined as:

$$z = \min c^T x,\ z, x \in \mathbb{R}^n \quad (6)$$

which has to be a linear expression formed by a vector x of decision variables and a cost vector c. This objective function is subject to a set of constraints defined as:

$$Ax \leqslant b \quad (7)$$

where A is called constraint matrix. Decision variables may take values between an upper and a lower bound which is defined as:

$$l \leqslant x \leqslant u \quad (8)$$

Some decision variables are required to take integer values. Integer variables that must take values 0 or 1 are called binary variables and play a special role in MIP modeling and solving.

Solving MIP usually includes two different stages. First, the problem is solved with a relaxation of the integer constraint, that is, the problem is solved by using the Simplex method (introduced by Dantzig [6]) as if there were no integer restrictions. This process