Computers, Environment and Urban Systems 32 (2008) 261–267

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Comparison of region approximation techniques based on Delaunay triangulations and Voronoi diagrams

Michael Barclay *, Antony Galton

School of Engineering, Computer Science and Mathematics, Exeter University, Exeter, UK

ARTICLE INFO

Keywords:
Region
Approximation
Error
Delaunay
Voronoi

ABSTRACT

Region approximation techniques based on constructions from sample data points, i.e. points whose position is known and which are known to be inside or outside the region of interest, can be advantageous in a variety of applications. This paper compares two different constructions and presents results from a Monte Carlo model that shows that the construction based on mid points of edges in a Delaunay triangulation produces the lowest errors. These errors are some 10% less than those produced by the Voronoi diagram construction which appears to be more widely used at present. A consideration of the basic geometries of the different constructions leads to an expression for approximating the expected error in the case of a random point distribution. The expression takes the form

$$E_{\mathrm{RMS}} = k_d\left[\frac{S}{4}+\frac{l_o^2}{32S}\left(\left(\frac{l_c^2}{l_o^2}\right)-1\right)\right],$$

where S is the average point spacing, l_o is the length of region boundary being approximated (the optimum length for the construction), and l_c the length of the constructed line approximating the boundary. The constant k_d accounts for the non-uniform spacing of the points in the distribution and has a value of about 1.1 for a random distribution. Predictions from this expression agree well with the results from the Monte Carlo model.

The case of finite as well as infinite radius of curvature is considered and some possible improvements on the constructions modelled are suggested.

1. Introduction

Region approximation techniques based on constructions from sample data points, i.e., points whose position is known and which are known to be inside or outside the region of interest, can be advantageous in a variety of applications. The technique is introduced in Alani, Tudhope, and Jones (2001) in which approximations to Scottish counties (their boundaries and area) were derived from the locations of towns and villages inside and outside the counties. The technique was extended in Arampatzis et al. (2006) to approximate regions with no defined boundaries such as the English 'Midlands'. Previously Gold (1996) had used Voronoi constructions to approximate region boundaries in an application where the centres of the Voronoi cells could be chosen to be near optimum. Further examples of the use of Voronoi diagrams to approximate regions can be found in geographical data interpretation (e.g., Tatalovich, Wilson, & Cockburn, 2006) and telecommunications network planning (e.g., Navas & Imielinski, 2000).

The techniques are attractive because they will typically require less storage and less computing power to generate representations of regions and answer related queries of the form 'what is the area of X?' and 'what is the length of boundary of X?' than techniques based on arbitrary polygons ('exact' vector representations). They are moreover far more accurate than simplistic region representations such as bounding boxes.

In an application where map data is being transmitted to mobile devices these methods are particularly attractive because the boundary (e.g., of a county) in effect comes free with the transmission of the point locations of (say) towns and villages.

While techniques of this nature are in use as described to approximate boundaries it is not thought that any comparative work on the different possible constructions has been carried out. Also no analysis of the likely errors (as opposed to actual errors for specific cases) exists, so it is not yet possible to say for a real data set (e.g., the case of Alani et al., 2001) that the error is better or worse than expected, or, if the error cannot be calculated (e.g., the case of Arampatzis et al., 2006), what it is likely to be. The

* Corresponding author. Tel.: +44 1392 264058.
E-mail addresses: michael.j.barclay@googlemail.com, mjb231@ex. ac.uk (M. Barclay).

0198-9715/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2008.06.003

results in this paper should enable the likely degree of uncertainty due to the construction of the approximation to be taken into account in future work using these techniques.

In approximating a region the area error is made up of areas modelled as inside the region which are in fact outside (+ve areas) and areas modelled as outside which are in fact inside (−ve areas). The total area error and the total area error as a proportion of the area of the region being approximated can both be misleading if the region is irregularly shaped. A related measure, the RMS (Root Mean Squared) distance between the constructed approximation line and the region boundary (denoted E_{RMS} and with dimension length, l), is considered to be more informative and is used in this paper.

Region approximations based on constructions from data points have an inherent 'scale' that is related to the average point spacing. Region boundary features smaller than the point spacing cannot be reliably modelled. The construction could be considered an approximation to a directly digitised region boundary with the same point spacing or 'scale'. This is in itself an approximation to the 'real' boundary.

The remainder of this paper is organised as follows:

Section 2 contains a brief description of the Voronoi diagram and Delaunay triangulation based constructions under consideration.

Section 3 describes the Monte Carlo model used to estimate the accuracy of the different constructions and the results obtained.

Section 4 looks at the basic geometries of the different constructions and why they should be more or less accurate as approximations to regions. An expression for the expected error in a random point distribution is derived.

Section 5 discusses the implications and notes where further work is required.

2. Constructions considered

This paper compares the following constructions, the methods being illustrated in Fig. 1:

1. Fig. 1a. Delaunay triangulation mid-points method (Arampatzis et al., 2006). The approximator line is formed by joining the mid-points of edges in the triangulation which cross the line to be approximated (i.e., edges that join points inside and outside the region of interest).
2. Fig. 1b. Voronoi diagram method (Alani et al., 2001). The approximator line is formed from the edges in a Voronoi diagram which separate cells around pairs of points, one of which is inside and the other outside the region of interest.

3. Results

3.1. Model

The independent variable, S, used in the investigation is

$$S = \frac{1}{\sqrt{\text{point density}}}, \quad (1)$$

which has dimension l and can be considered as roughly equivalent to the average point spacing. Point densities of 50, 100, 200, 400 and 800 in the unit square have been used.

Results are presented from a Monte Carlo model of the different constructions. 800 random point distributions each with five randomly placed lines were used for each point density. The randomly placed lines equate to portions of a region boundary to be approximated. The model generates a random set of points within the unit square centered on the origin. The Qhull algorithm due to Barber, Dobkin, and Huhdanpaa (1996) is called to generate a Delaunay triangulation and Voronoi diagram for the point set and line. The relevant edges in the Delaunay Triangulation and Voronoi diagram are identified and used to construct the approximation line. The error between this approximation line and the randomly placed 'boundary' line is then calculated.

Edge effects are apparent in both the Delaunay triangulation and Voronoi diagram around a finite point set and are removed by only using inner portions of the diagram.

3.2. Straight line case

The results show that the construction based on mid-points of edges in a Delaunay triangulation produces errors that are some 10% less than those produced by the Voronoi diagram construction which appears to be more widely used at present. The results for the different constructions with varying S (in arbitrary length units) are given in Table 1.

As should be expected the error E_{RMS} can be simply expressed as $E_{RMS} = KS$. The average value of $\frac{E_{RMS}}{S}$ for all the values of S from Table 1 is used, giving

$$K_{\text{Delaunay}} = 0.274,$$
$$K_{\text{Voronoi}} = 0.307.$$

The consistency of the standard deviation would appear to be due to this being a function of the point distribution, not the construction used. The two constructions produce related, though different, area errors, as discussed in Section 4.

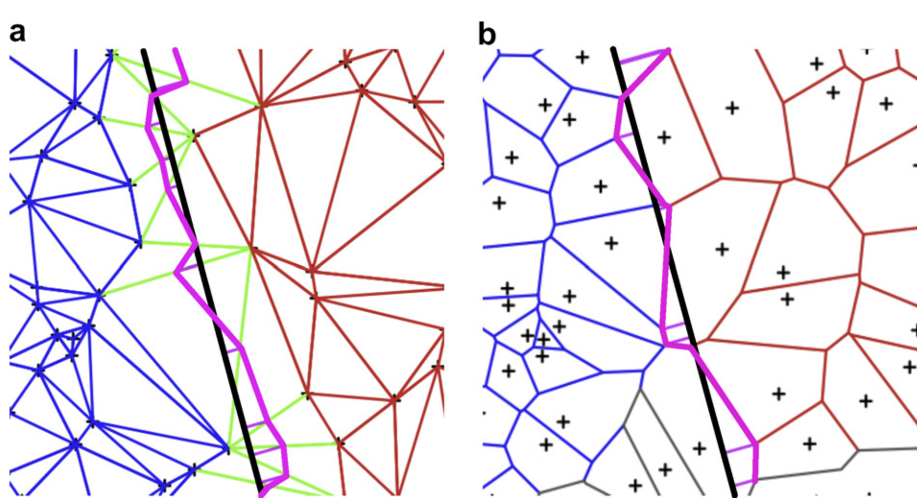

Fig. 1. Approximations to a region boundary line: (a) Delaunay triangulation; (b) Voronoi diagram.