Comparison of region approximation techniques based on Delaunay Triangulations and Voronoi Diagrams

Michael Barclay and Antony Galton
School of Engineering, Computer Science and Mathematics.Exeter University, Exeter, UK
Telephone: +44 1392 264058
michael.j.barclay@googlemail.com

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 et al (2001) and examples of its use are given in Arampatzis et al (2006) and Tatalovich (2005).

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 however far more accurate than simplistic region representations such as bounding boxes.

In the 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 that the error is better or worse than expected, or, if the error cannot be calculated, what it is likely to be. The results in this paper should enable the likely degree of uncertainty 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). A related measure, the RMS (Root Mean Squared) distance between the constructed approximation line and the region boundary, is considered to be more informative.

2 Constructions considered

This paper compares the following constructions:

2. Delaunay triangulation mid-points method. This method is illustrated in figure 1. The approximator line (dark gray) is formed by joining the mid-points of edges (mid gray) in the triangulation which cross the line to be approximated (in black) (i.e., they join points inside and outside the region of interest).

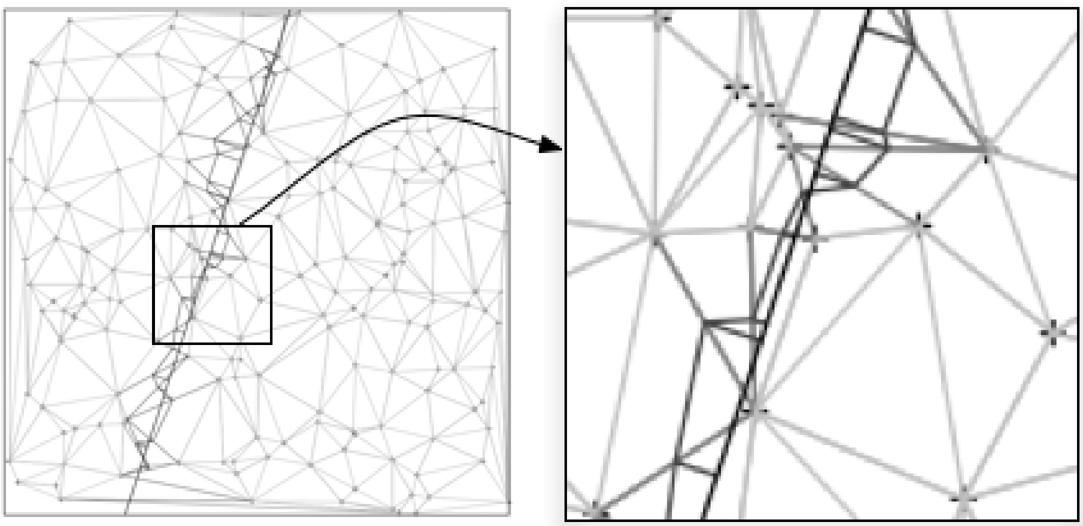

Figure 1: Delaunay mid-point approximation to a region boundary line

3. Further constructions. Various other possibilities for forming the approximator line and limitations on the use of filtering (smoothing) are briefly considered.

3 Results

Results are presented from a Monte Carlo model of the different constructions. 800 random point distributions each with 5 randomly placed lines were used for each point density. A Delaunay triangulation and Voronoi diagram were produced for each random point distribution and line. The results show that the construction based on mid-points of edges in a Delaunay triangulation produces the lowest errors, 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 for differing sample point densities in RMS length units are given in table 1. The sample point density is the stated number of points in a square of unit length.

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 below.

The results from the model in the case of approximating to a circle of fixed radius with varying point density are given in table 2 and table 3. The imbalance between the +ve and $-$ve area errors is discussed below.

4 Discussion

A consideration of the basic geometries of the two constructions suggests an explanation for the difference between them based on the length of the generated approximation line. The geometrically based predictions agree closely with the results from the Monte Carlo model.

Consider the 'canonical' case in figure 2. Here a line segment for each construction is shown along with a segment (L(average)) of the line to be approximated. The average length of line crossing the triangle can be derived for the special case of an equilateral triangle grid by considering the diagram in figure 3. From symmetry it is clear that all

	Delaunay mid-points		**Voronoi diagram**	
No. Points	RMS Error	Std. Dev.	RMS Error	Std. Dev.
50	0.03783	0.01438	0.04046	0.01487
100	0.02763	0.00878	0.03025	0.00897
200	0.01949	0.00483	0.02204	0.00474
400	0.01380	0.00281	0.01584	0.00280
800	0.00976	0.00169	0.01122	0.00169

Table 1: Approximation error with varying sample point density

Delaunay mid-points errors				
No. points	+ve rms error	-ve rms error.	total rms error	std. dev.
50	0.01749	0.02266	0.04020	0.00956
100	0.01273	0.01541	0.02814	0.00576
200	0.00956	0.01018	0.01975	0.00324
400	0.00668	0.00728	0.01395	0.00205
800	0.00475	0.00504	0.00979	0.00118

Table 2: Radius of curvature effect on Delaunay mid-points approximation error

Voronoi diagram errors				
No. points	+ve rms error	-ve rms error.	total rms error	std. dev.
50	0.02232	0.02266	0.04496	0.00952
100	0.01583	0.01615	0.03198	0.00586
200	0.01082	0.01132	0.02214	0.00313
400	0.00797	0.00790	0.01588	0.00199
800	0.00567	0.00561	0.01128	0.00117

Table 3: Radius of curvature effect on Voronoi diagram approximation error

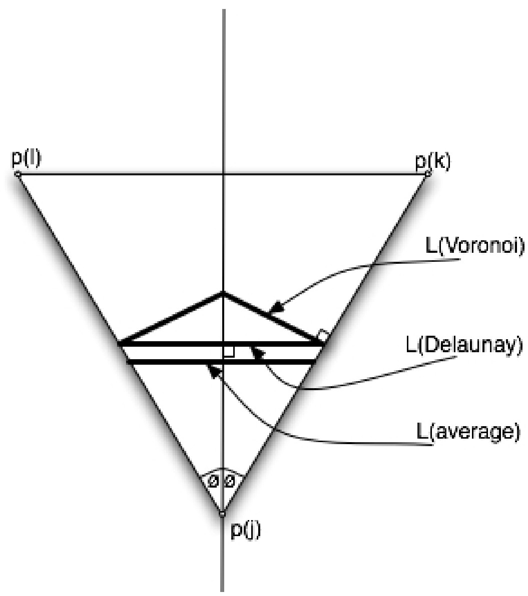

Figure 2: Canonical case of approximation lines constructed

lines P can be considered as being between angle 0 and $\frac{\pi}{6}$ of some arbitrary orientation. For lines in general position (i.e.,not passing through vertices) the number of triangles, N, crossed traversing the hexagon is constant $N = 2n$ (figure 3) and the average length of line, $L(average)$, per triangle of side length d, can be easily calculated as:

$$L(average) = \frac{P(average)}{N} = \frac{3}{\pi}\int_0^{\frac{\pi}{6}}(d\cos(\alpha))d\alpha = 0.477d \tag{1}$$

assuming all angles α to be equally likely. The extra line length generated by the two constructions ($L(Delaunay) = 0.5d$, $L(Voronoi) = 0.577d$) necessarily leads to an extra area error. This is greater for the Voronoi diagram as shown in the Monte Carlo results. A construction which generated the correct line length would still exhibit area errors.

Generalising from this special case requires consideration of the distribution of angles ϕ but it is worth noting that the Delaunay triangulation by its nature tends to minimise variation of ϕ from the equilateral case. Cases where the Voronoi construction does not conform to figure 2 would also need to be considered.

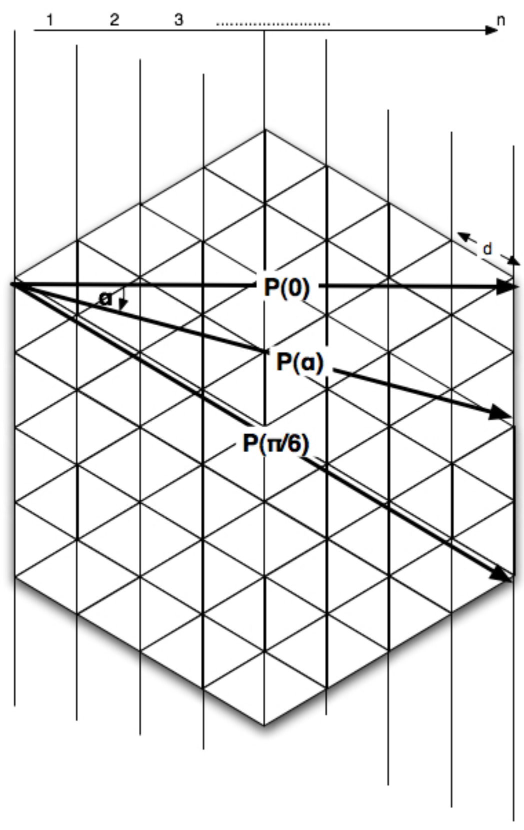

Figure 3: Construction for calculating L(average)

The geometry of the constructions when approximating boundaries with finite radius of curvature is illustrated in figure 4. The effect of the curvature will depend on whether P(j) is inside or outside of the region, however for a closed convex region there will be more cases where it is inside than out (6 more for the equilateral grid special case). Thus we would expect the reduction in positive area error shown in the Monte Carlo results (table 2) for the Delaunay mid-points approximation to a circle. Also as expected the Voronoi diagram approximation to a circle does not exhibit the effect to the same extent (table 3)

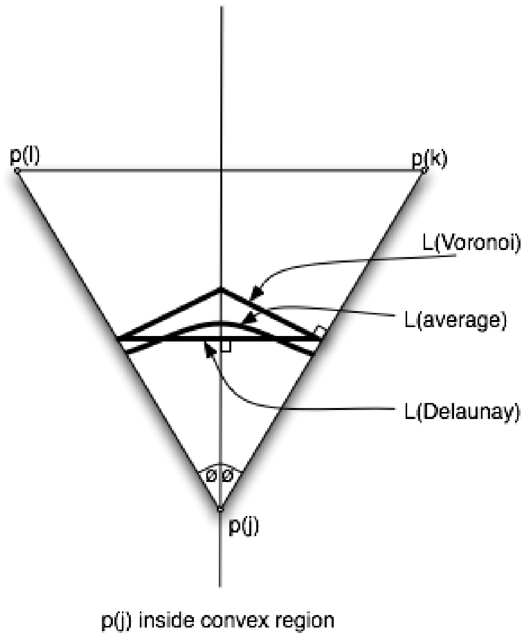

Figure 4: Approximation to curved line segments

Some discrepancies in the size of the errors between the Monte Carlo model and those predicted from consideration of the geometry remain unresolved and are noted as topics for further work.

References

Arampatzis, A., van Kreveld, M., Reinbacher, I., Jones, C. B., Vaid, S., Clough, P., Joho, H., and Sanderson, M. (2006). Web-based delineation of imprecise regions. *Computers, Environment and Urban Systems*, **30**(4), 436–459.

Tatalovich, Z. (2005). A comparison of Thiessen-polygon, Kriging, and spline models of UV exposure. *Proceedings of the University Consortium of Geographical Information Science Summer Assembly*.

Alani, H., Tudhope, D., and Jones, C. (2001). Voronoi-based region approximation for geographical information retrieval with gazetteers. *Int. J. Geographical Information Science*, **15**(4), 287–306.