BIOINFORMATICS APPLICATIONS NOTE Vol. 28 no. 13 2012, pages 1802–1804
doi:10.1093/bioinformatics/bts253

Data and text mining

Advance Access publication May 3, 2012

Ka-me: a Voronoi image analyzer

Noppadon Khiripet[1,*], Wongarnet Khantuwan[1] and John R. Jungck[2]

[1,*]Knowledge Elicitation and Archiving Laboratory, National Electronics and Computer Technology Center (NECTEC), 112 Phahon Yothin Rd., Klong 1, Klong Luang, Pathumthani 12120, Thailand and [2]Department of Biology, Beloit College, 700 College Street, Beloit, WI 53511, USA

Associate Editor: John Quackenbush

ABSTRACT

Summary: Ka-me is a Voronoi image analyzer that allows users to analyze any image with a convex polygonal tessellation or any spatial point distribution by fitting Voronoi polygons and their dual, Delaunay triangulations, to the pattern. The analytical tools include a variety of graph theoretic and geometric tools that summarize the distribution of the numbers of edges per face, areas, perimeters, angles of Delaunay triangle edges (anglograms), Gabriel graphs, nearest neighbor graphs, minimal spanning trees, Ulam trees, Pitteway tests, circumcircles and convexhulls, as well as spatial statistics (Clark–Evans Nearest Neighborhood and Variance to Mean Ratio) and export functions for standard relationships (Lewis's Law, Desch's Law and Aboav–Weaire Law).

Availability: Ka-me: a Voronoi image analyzer is available as an executable with documentation and sample applications from the BioQUEST Library (http://bioquest.org/downloads/kame_1.0.rar).

Contact: noppadon.khiripet@nectec.or.th

Received on October 25, 2011; revised on April 5, 2012; accepted on April 24, 2012

Important recent work in computational biology (Bock *et al.*, 2010; Gibson *et al.*, 2006; Mulder, 2008; Patel *et al.*, 2009) on dividing epithelia has focused on the analysis of convex polygonal tessellations, yet no tool exists for doing standard geometric, graph theoretic and spatial statistical analysis of such patterns. A commonly occurring convex polygonal tessellation is a Voronoi tessellation, whose cells are defined as all points in the region around a generator point that are closer to it than any other generator point; thus, cell boundaries perpendicularly bisect the planar space between two generator points. An extensive literature exists on Voronoi tessellations in biology, astronomy, geography, archeology and engineering (Okabe *et al.*, 1992). Allometry and fractals have captured the imagination of biologists as well as amateurs because both apply across biological phenomena from the molecular to the ecological level. Voronoi polygons and polyhedra are less well known, but scale equally well across phylogenetic, spatial and temporal dimensions. Elsewhere, we claimed computational geometric and graph theory based around Voronoi tessellations are particularly applicable in biology because they work well on multiple spatial scales (from proteins to viral capsids to single celled radiolaria to animal coat patterns to a variety of fruits to fish territories to forest canopies), multiple temporal scales (from nanoseconds of protein folding to eons in fossilization of stony coral such as Petoski stones), multiple phylogenetic scales (across viruses, bacteria, protozoans, plants, fungi and animals) and multiple causes (such as diffusion, growth, packing, competition, optimization, sorting and avoidance). Thus, we developed a generally applicable tool that we call 'Ka-me: a Voronoi image analyzer' for biological image analysis. Ka-me is the Japanese word for turtle; turtle shells often have Voronoi tessellations.

We (Khiripet *et al.*, 2011) are generally interested in articulating the close relationship between actual biological specimens and causal theoretical models. Ka-me instantiates this approach by allowing an investigator to choose among a diversity of mathematical and statistical tools for testing causal hypotheses about biological patterns. Ka-me allows the user to import a biological image and then interactively (not automatically) fit individual Voronoi cells to match each convex polygon in the tessellation (Fig. 1). Ka-me uses a quad-edge data structure that allows this to occur instantaneously. Once an image has been completely tessellated, then other analyses are appropriate to examine. We only report statistics for those cells that lie within the convex hull because cells on the periphery often have infinite areas or an indeterminate number of edges.

Our choice of tool sets was informed by such disciplines as developmental biology, ecology, ethology, forestry and clinical pathology. We have included as many of the tools as have seemed profitable in previous studies without overwhelming users or severely cluttering the user interface. The analytical tools include a variety of graph theoretic and geometric tools that summarize the distribution of the numbers of edges per face, areas, perimeters, angles of Delaunay triangle edges (anglograms), Gabriel graphs, nearest neighbor graphs, minimal spanning trees, Ulam trees, Pitteway tests, circumcircles and convex hulls, as well as spatial statistics (Clark–Evans Nearest Neighborhood and Variance to Mean Ratio) and export functions for standard relationships (Lewis's Law, Desch's Law and Aboav–Weaire Law).

Ka-me has been applied to five different kinds of basic biological questions. (1) Spatial point distributions of a single species: John Snow's map of a cholera outbreak in London (Jungck, 2012). Second order statistics like Ripley's *K*-function and a Monte Carlo confidence envelope have been applied to mapped affected individual households over five intervals of a dengue fever outbreak in Argentina. (2) Spatial point distributions of two entities: chi-squared tests (Diggle, 2003) on the occurrence of three types of Delaunay triangulation edges have been used to study successful and unsuccessful red-winged blackbirds nests in

*To whom correspondence should be addressed.

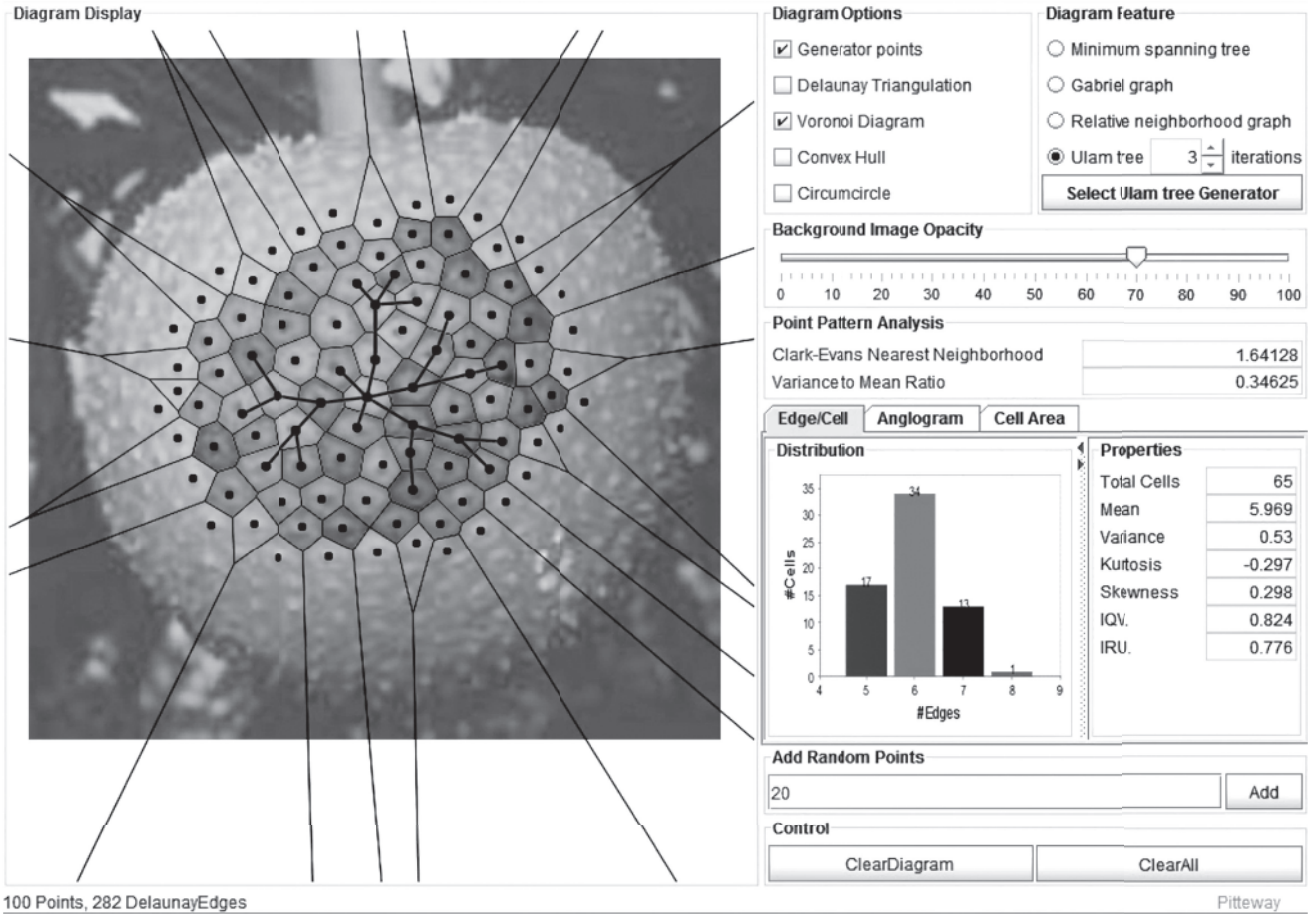

Fig. 1. The Ka-me interface with a Thai breadfruit as the biological image. The distribution of 65 Voronoi cells is illustrated as a histogram of the number of cells with a particular number of sides. From the 'Point Pattern Analysis', we can infer that the generator points used to make Voronoi cells matching polygonal segments on the fruit's surface are distributed slightly over-dispersed toward the limit of being fairly uniformly distributed, but not nearly as uniformly as when all cells are hexagonal. The total number of generator points and Delaunay edges are listed in the lower left hand corner. In the lower right hand corner, Pitteway is not highlighted because there are violations of the Pitteway criterion on the Delaunay triangulation. The Ulam tree graph (degree = 3) is superimposed on the image. More detail is not shown simply to avoid the limits of visual overload in one image

a prairie, on and off Amacrine cells in an eye layer, gall infested versus non-infested *Solidago canadensis* plants, barnacles with and without algae, tests of the self-thinning hypothesis for forests of hickories, oaks and maples as well as old versus new canes in raspberry brambles and Pyknotic versus metaphase nuclei in hamster tumors. (3) Lewis's Law (area versus number of sides), Desch's Law (perimeter versus number of sides) and Aboav–Weaire Law (conservation of topological charge) were used to examine convex polygonal tessellations in a variety of species: radiolaria, plant pollen, *Drosophila melanogaster* eggs, *Arabidopsis thaliana* embryos, apical meristems of *Anagallis arvensis* and petals of *Rosa glauca*. With the development of fluorescent stains for cell membranes, extremely crisp images of polygonal tissue patterns are appearing in contemporary literature. (4) Ten types of Delaunay faces for topological charge distributions were tested on five successive mitotic divisions in *D.melanogaster* embryos. (5) Graph theoretic analyses were performed on tongue cancers [Sudbo *et al*. (2000) and Landini and Othman (2004) described the Ulam tree, a Delaunay triangulation subgraph that was of significant diagnostic value] and interneurons [Minciacchi *et al*. (2009) found that, for some measurements, the number of cells to be analyzed must exceed more than 3000 cells—far more than a typical pathologist can do by inspection]. We believe that Ka-me is the only software package available that implements Ulam trees in an image analysis context as well as robustly and easily perform other graph theoretic, geometric and spatial statistical analyses. Between the built-in functions in Ka-me and the exported data files generated by Ka-me is a broadly applicable tool for analyzing biological images that range from electron micrographs to ecological and epidemiological distribution maps.

ACKNOWLEDGEMENTS

Beloit College students J.P. Rindfleisch, Brian Maunze, Yang Yang, Sijia Liang, Jourdan Posner Erika Moen and Trisha Rettig performed numerous analyses of biological images to test the robustness of Ka-me' in a wide variety of contexts such as faculty development

workshops in mathematical biology offered by the BioQUEST Curriculum Consortium (http://bioquest.org).

Funding: This research is supported by a grant from NECTEC.

Conflict of Interest: none declared.

REFERENCES

Bock,M. *et al*. (2010) Generalized Voronoi tessellation as a model of two-dimensional cell tissue dynamics. *Bull. Math. Biol.,* **72**, 1696–1731.

Diggle,P. (2003) *Statistical Analysis of Spatial Point Patterns*. Oxford University Press, Oxford.

Gibson,M.C. *et al*. (2006) The emergence of geometric order in proliferating metazoan epithelia. *Nature*, **442**, 1038–1041.

Jungck,J.R. (2012) Incorporating quantitative reasoning in common core courses: mathematics for *the ghost map*. *Numeracy*, **5**, Article 7.

Khiripet,N. *et al*. (2011) Morphospace: measurement, modeling, mathematics, and meaning. *Math. Model. Nat. Phenom.*, **6**, 54–81.

Landini,G. and Othman,I.E. (2004) Architectural analysis of oral cancer, dysplastic, and normal epithelia. *Cytometry A*, **61**, 45–55.

Minciacchi,D. *et al*. (2009) Voronoi-based spatial analysis reveals selective interneuron changes in the cortex of FALS mice. *Exp. Neurol.*, **215**, 77–86.

Mulder,B. (2008) On growth and force. *Science* **322**, 1643–1644.

Okabe,A. *et al*. (1992) *Spatial Tessellations: Concepts and Applications of Voronoi diagrams*. Wiley, New York.

Patel,A.B. *et al*. (2009) Modeling and inferring cleavage patterns in proliferating epithelia. *PLoS Comput. Biol.*, **5**, e1000412.

Sudbo,J. *et al*. (2000) Prognostic value of graph theory-based tissue architecture analysis in carcinomas of the tongue. *Lab. Investig.,* **80**, 1881–1889.