Test questions Geometric Pattern Recognition 2015/2016

1. Is it true that if the Douglas-Peucker algorithm is run on a polygonal line with two different threshold values ε_1 and ε_2 with $\varepsilon_1 < \varepsilon_2$, then the remaining vertices when using ε_2 is a subset (or the same) of the remaining vertices when using ε_1? Argue why this is true or give a counterexample.

2. Discuss the advantages and disadvantages of the Imai-Iri line simplification algorithm with respect to the Douglas-Peucker algorithm.

3. What is the running time of the Imai-Iri algorithm on a polygonal line with n vertices in the worst case, expressed in $O(..)$ notation?

4. What are the risks associated with analyzing aggregated data?

5. Explain how the area of a simple polygon can be computed.

6. Give two different options for summarizing the location of a simple polygon.

7. One way of summarizing the orientation of a simple polygon is by the orientation of the diameter (the two points furthest apart of that polygon). Another way is to use the orientation of the two parallel lines that enclose the polygon and are closest together. Show that these two possibilities can give a substantially different orientation for some polygons.

8. Which of the three methods listed to capture the orientation of a simple polygon appears to represent orientation best if the polygon can have holes (enclosed regions within the polygon that are not part of it)? Explain your answer.

9. An alternative is to compute the minimum enclosing ellipse and use the direction between their focal points. An advantage of this method is that the distance between the focal points, in relation to the chord length (radius) of the ellipse, gives an indication of how elongated the polygon is.

   Suggest a different way, based on the known ways to measure the orientation, to determine how elongated a polygon is.

10. Describe how the curvature along a polygonal line can be defined.

11. Explain why (point) density is a scale-dependent measure.

12. Suggest a definition of density for points and polygonal lines simultaneously.

13. What is spatial auto-correlation? Why is it important to know the range of spatial auto-correlation?

14. Sketch an efficient algorithm to compute the buffer of a polygonal line (given a buffer distance value).

15. A function that takes two arguments and returns a value may be a metric. Under what conditions is it a metric?

16. Is the Manhattan distance (also called L_1-distance) a metric? Give the arguments why (like a proof sketch), or give an example where one of the conditions for being a metric is violated.

17. Measures can use aggregation of multiple values to determine a final value. Such aggregation may be bottleneck, sum, and sum-of-squares. Explain what it means if the aggregation is done in a bottleneck manner.

18. What measure would you use to determine how much a set of points resembles a simple polygon?

19. Explain how the Frechet distance is defined.

20. What are the main differences between the Frechet distance and Dynamic Time Warping?

21. What is the asymptotic running time of Dynamic Time Warping on a polygonal line with n vertices and a polygonal line with m vertices?

22. Give an example of two polygonal lines where the Hausdorff distance is small but the Frechet distance is large. Can the reverse be the case? Give an example or argue why not.

23. Given the following points with their weights [picture], determine their Earth Mover's Distance.

24. Design a measure that expresses, for a planar subdivision, how similar in size its cells are.

25. Same as the above, but give both a bottleneck measure and a sum measure.

26. Given a set P of points in the plane, design (and define) a score measure for a curve that gives high values if the curve bends much close to points of P but not much otherwise.

    Discuss how robust your measure is (sensitive to outliers in P).

27. Explain how two scores of an object can be converted in a single score that captures both aspects. Give two different possibilities.

28. Explain the empty circle property for Voronoi diagrams.

29. What is the difference between a Gabriel Graph and a Relative Neighborhood Graph?

30. Argue that the edges of a Relative Neighborhood Graph are a subset of the edges of the Delaunay triangulation.

31. What influence does the value of α have on the shape of the α-shape?

32. Given the following disk size and Voronoi diagram inside the disks [picture], draw the alpha-complex.

33. The persistence of a loop in a data set of points, when considering the alpha-complex with growing alpha, is the difference in alpha values when the loop was first created and when the loop was completely filled with triangles.

    Explain that the persistence allows us to distinguish noise from features (like loops) in a data set.

34. In a 3-dimensional alpha-complex, and growing alpha, how can tunnels in the alpha-complex be created and how can they disappear?

35. The rectangle problem concerns detecting whether a sample of points has the shape of a rectangle. There are two conditions to be fulfilled. Suppose there is some occlusion in the data collection, causing that in some parts of the sample, points were not collected. Which criterion must be dropped to still be able to infer a rectangle shape in the points, and why?

    Can you extend the rectangle problem (just the definition of the problem, not the algorithmic solution) to a formulation that can handle a small amount of occlusion? (Changing the value of the parameter r is not a good answer.)

36. Suppose that in a RANSAC method, the model to be retrieved in a set of points in the plane is a square of fixed size but unknown location and orientation. How many points do we sample, and how do we get a model from these points?

37. Suppose that for a LiDAR scanner, the distance of a point (reflection of the laser on a surface) is less precise than the direction in which the point lies. How can this situation be incorporated in a RANSAC method?

38. How does the number of points needed to define a model affect the number of iterations needed to find the model with the highest inlier/point-set ratio?

39. How is the normal of a point used to determine support in the RANSAC method?

40. Suppose that a basic RANSAC algorithm is implemented, which does not use estimated normals and which does not use data structures to compute the supporting points of a model.

    If the RANSAC method works on n points and has i iterations, what is the running time, expressed in $O(..)$-notation?

41. Explain how the Hough transform works.

42. Suppose that the Hough transform to detect lines in a point set in 2D gives rise to the following matrix of counts from the points:

3	7	17	10	6	7	3
4	6	9	9	5	8	2
6	8	3	7	3	23	7
4	4	6	4	3	6	2
5	6	9	9	5	3	2
5	14	15	4	5	3	4
6	8	3	7	3	8	7

    How many lines do you think are in the point set, and why?

43. What is the main disadvantage of the Hough transform with respect to RANSAC?

44. Suppose we want to exactly match two point sets under translation and scaling in 2D. Describe how this can be done efficiently.

45. Same question, but now translation, rotation, and scaling in 3D.

46. Same question, but now in 3D.

47. How can point matching between two sets of points be solved when there is noise in the locations of the points?

48. How can point matching between two sets of points be solved when there are outliers in the locations of the points?

49. Why is it not a good idea in trajectory segmentation to use absolute values across which a segment boundary must be used?

    Explain what alternative segmentation criteria exist for segmentation on speed.

50. Give an example of a criterion that is not monotone, for trajectory segmentation.

51. A distance measure for two trajectories is the time-aware distance. Assuming two trajectories are given in exactly the same time interval, the time-aware distance is the maximum distance that is realized between the two moving entities (maximum distance over time). Is this distance measure a metric? Explain why or why not.

52. Give the definition of a sum-of-squares type of measure based on the time-aware distance between two trajectories.

53. What parameters are needed to define groups in a set of moving entities?

54. How can the grouping definition be adapted to ensure that the group is not a stationary group, but moving?

55. Explain the Reeb graph to represent the grouping structure of a set of moving entities.