Computational Geometry 48 (2015) 673–687

Contents lists available at ScienceDirect

Computational Geometry: Theory and Applications

www.elsevier.com/locate/comgeo

Separating bichromatic point sets by L-shapes

Farnaz Sheikhi [a], Ali Mohades [a,*], Mark de Berg [b], Mansoor Davoodi [c]

[a] Laboratory of Algorithms and Computational Geometry, Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran
[b] Department of Computer Science, TU Eindhoven, PO Box 513, 5600 MB, Eindhoven, The Netherlands
[c] Department of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran

ARTICLE INFO

Article history:
Received 21 December 2012
Accepted 3 February 2014
Available online 29 June 2015

Keywords:
Separability
Bichromatic point sets
L-shape

ABSTRACT

Given a set R of red points and a set B of blue points in the plane, we study the problem of determining all angles for which there exists an L-shape containing all points from B and no points from R. We propose a worst-case optimal algorithm to solve this problem in $O(n^2)$ time and $O(n)$ storage, where $n = |R| + |B|$. We also describe an output-sensitive algorithm that reports these angles in $O(n^{8/5+\varepsilon} + k \log k)$ time and $O(n^{8/5+\varepsilon})$ storage, where k is the number of reported angular intervals and $\varepsilon > 0$ is any fixed constant.

1. Introduction

Background. In a separability problem in $\mathbb{R}^2$ one is given two colored sets of objects, a set R of red objects and a set B of blue objects, and a class $\mathcal{S}$ of curves. The curves are usually either infinite curves (such as lines) or closed curves (such as circles), so that they partition the plane into two regions. The goal is now to decide whether there is a curve in $\mathcal{S}$ that is a *separator* for R and B, that is, a curve that partitions the plane such that R and B lie in different regions. If such a placement is possible, one often also wants to compute all separators, or the separator minimizing some cost function. Geometric separability arises in applications where classification is required, such as machine learning and image processing.

There has been a fair amount of work on different kinds of separators, both in the plane and in higher dimensions. For separability in the plane where the objects to be separated are points—this is the topic of our paper—the following results are known. The most basic version of the problem is where the class $\mathcal{S}$ is the class of all lines. The problem of deciding whether the two point sets can be separated by a line can be solved in linear time, as shown by Megiddo [14]. Seara [18] showed how to compute in linear time all orientations for which there exists a line separating the two point sets. O'Rourke et al. [17] studied a different type of separators, namely circles. They presented a linear-time algorithm for deciding whether the two point sets can be separated by a circle, and they also gave algorithms for finding the smallest and the largest separating circle. The problem of finding a convex polygon with minimum number of edges separating the two point sets, if it exists, was solved by Edelsbrunner and Preparata [6] in $O(n \log n)$ time. Fekete [7] showed that the problem of determining a simple polygon with a minimum number of edges separating the two point sets is NP-hard, and a polynomial-time approximation algorithm was provided by Mitchell [15]. Separability problems have been studied for separators in the form of strips and wedges [9] as well. A thorough study is presented by Seara [18].

* Corresponding author.
E-mail addresses: f.sheikhi@aut.ac.ir (F. Sheikhi), mohades@aut.ac.ir (A. Mohades), mdberg@win.tue.nl (M. de Berg), mdmonfared@iasbs.ac.ir (M. Davoodi).

http://dx.doi.org/10.1016/j.comgeo.2015.06.008
0925-7721/

Separability has also been studied for rectangular separators. This has applications in urban scene reconstruction, which seeks to reconstruct buildings from LIDAR data. The idea is, roughly, to first cluster the data points, then project the points from each cluster onto a suitable plane, and finally find the building facets (typically walls and roofs) corresponding to the clusters [11–13]. Since walls and roofs are often rectangles or other rectilinear shapes, the task is to find a suitable rectilinear shape enclosing the points. Sometimes there are also points that are known to not be part of the facet being reconstructed. We then seek a shape that includes the points in the facet (positive samples) while excluding the points known to not be in the facet (negative samples). The distinction between the positive and negative samples is represented by the color assigned to them, thus leading to separability problems for rectilinear shapes. Following this motivation, Van Kreveld et al. [12,13] recently studied the separability problem in the plane for the case where the separator is a (not necessarily axis-aligned) rectangle. They proposed an $O(n \log n)$ time algorithm to compute all orientations for which a rectangular separator exists. They also required that the rectangle tightly fits the blue point set. More precisely, the rectangle should be δ-covered by the blue points, meaning that it is contained in the union of the radius-δ disks centered at the blue points. They mentioned the case of a non-convex separator, namely an L-shape, as an open problem. This is the topic of our paper, except that we do not require the L-shape to be δ-covered. Extending it to take this extra condition into account is an interesting open problem. Next we define the problem more precisely and we state our results.

Exact problem statement and results. We define an *axis-aligned L-shape* to be the set-theoretic difference $M \setminus M'$ of two axis-aligned rectangles M and M' such that the top-right corners of M and M' coincide. More precisely, an L-shape is defined as $\mathrm{Cl}(\mathrm{Int}(M) \setminus \mathrm{Int}(M'))$, where M and M' are closed rectangles with $M' \subsetneq M$ that share their top-right corner. (Here $\mathrm{Cl}(\cdot)$ and $\mathrm{Int}(\cdot)$ denote the closure and interior, respectively, of a planar set.) Note that an L-shape is a closed set, that is, it includes its boundary and that an L-shape can degenerate into a rectangle. An *L-shape with orientation* θ is defined as an axis-aligned L-shape that has been rotated in counterclockwise direction over an angle of θ. Now, given a blue point set B and a red point set R, we wish to find all angles $\theta \in [0, 2\pi)$ for which there exists an L-shape L with orientation θ such that $B \subset L$ and $R \cap L = \emptyset$. From now on, we call such an L-shape a *blue L-shape*. The orientations for which a blue L-shape exists form a collection of subintervals of $[0, 2\pi)$. We present two algorithms for computing this collection of intervals.

The first algorithm is an algorithm running in $O(n^2)$ time and using $O(n)$ storage. We prove that this algorithm is worst-case optimal by showing that there are point sets that admit $\Omega(n^2)$ disjoint intervals for which there exists a blue L-shape. The second algorithm is an output-sensitive algorithm which uses $O(n^{8/5+\varepsilon} + k \log k)$ time and $O(n^{8/5+\varepsilon})$ storage, where k is the number of reported angular intervals and $\varepsilon > 0$ is any fixed constant. (Our results are based on two preliminary papers [20,21]. Compared to these preliminary results, the running time of the first algorithm has been slightly improved so that it is now worst-case optimal. The lower bound is completely new.)

2. Preliminaries

Terminology and notation. Our global strategy will be to do a rotational sweep: we increase θ from 0 to 2π and we report the angular intervals for which there is a blue L-shape while we sweep. It will be convenient to think about the sweep as rotating the coordinate frame. Thus, we define the x_θ-axis and the y_θ-axis as the coordinate axes after the coordinate frame has been rotated over an angle θ in counterclockwise direction. We denote the coordinates of a point p in the rotated coordinate frame by $x_\theta(p)$ and $y_\theta(p)$. Whenever we talk about the top-right corner of a rectangle, we mean the top-right corner with respect to the current coordinate frame.

For an angle θ, we denote the minimum bounding rectangle with orientation θ of the blue point set B by $M_B(\theta)$. In other words, $M_B(\theta)$ is the axis-aligned bounding box of B in the rotated coordinate frame. Let $R_\theta := R \cap M_B(\theta)$ be the set of red points inside $M_B(\theta)$, and define $M_R(\theta)$ to be the smallest rectangle with orientation θ that contains R_θ and shares its top-right corner with $M_B(\theta)$. Note that $L_\theta := M_B(\theta) \setminus M_R(\theta)$ is an L-shape.

To determine whether L_θ contains all points from B, we will define a so-called *step-shape*. We say that a point q *dominates* another point p (at orientation θ) if $x_\theta(q) \geq x_\theta(p)$ and $y_\theta(q) \geq y_\theta(p)$, and we say that a point $b_i \in B$ is *maximal* (at orientation θ) if there is no other point $b_j \in B$ that dominates b_i. The step-shape is now defined as the region in $M_B(\theta)$ consisting of all points that are dominated by a blue point. The step-shape is a closed shape, bounded by the staircase that connects the maximal blue points, and by parts of the boundary of $M_B(\theta)$—see Fig. 1. Using the step-shape, we can characterize when L_θ contains all the blue points. To this end, we define the *red witness* as the lower-left corner of $M_R(\theta)$; note that the red witness is the reflex corner of L_θ. We denote the red witness by w_θ. Then L_θ contains all blue points if and only if the red witness w_θ lies outside or on the step-shape. However, note that despite containing all blue points in this case, L_θ is not blue as it contains red points on its boundary. But, here we can slightly shrink L_θ (by slightly enlarging $M_R(\theta)$) to obtain an L-shape not containing any red points. The shrunk L-shape still contains all blue points if and only if the red witness did not lie on the boundary of the step-shape. So we have the following observation.

Observation 1. *There exists a blue L-shape in the orientation θ if and only if the red witness w_θ lies outside the step-shape.*

The global strategy. As remarked earlier, our global strategy is to perform a rotational sweep. While we sweep, we will maintain the following information: