Best Fitting Rectangles

Manuel Abellanas* Ferran Hurtado† Christian Icking‡ Lihong Ma‡
Belén Palop§ Pedro A. Ramos¶

Extended Abstract, January 2003

1 Introduction

We solve an interesting optimization problem motivated by facility location and tolerancing metrology, see [5, 6, 9]: What rectangle fits best a given set of points? This problem also arises in dealing with paper position sensing [3]. Although our problem has some similarities to the problem of the largest empty rectangle for which no $O(n \log n)$ time solution is known, see [4, 8], for our problem there is a simple algorithm which runs within that time if the aspect ratio of the rectangle is given and even in time $O(n)$ if not. Other problems of this kind include the fitting of points by a circle, see [7], and offset polygon problems, see [1, 2].

2 Notations

We are given a set P of n point sites in the plane. Our task is to determine the rectangle which is, in some sense, closest to all of them.

As usual, the distance between a point and an extended object means the distance between the point and the closest point on the object, so the distance between a point p and a rectangle R is

$$d(p, R) = \min_{q \in R} d(p, q) .$$

Here, $d(p, q)$ denotes the distance in the underlying metric. Note that by rectangle we mean the boundary of the rectangle, not the interior. So a point in the interior of a rectangle has a non-zero distance to the rectangle.

For a certain subset, $\mathcal{R}$, of admitted rectangles, we are looking for the best fitting one, i. e., a rectangle such that

$$\max_{p \in P} d(p, R)$$

is minimized over all rectangles R of that kind. In other words, the best fitting rectangle R_{opt} fulfills

$$\max_{p \in P} d(p, R_{opt}) = \min_{R \in \mathcal{R}} \max_{p \in P} d(p, R) .$$

Different kinds of admitted rectangles, $\mathcal{R}$, and different metrics generate different problems to be solved.

An environment of a rectangle is called a frame. More precisely, the set of points whose distance to a rectangle R is less or equal to ε is called the *ε-frame* of R, for short F_R^ε. The two

*Dept. de Matemática Aplicada, Universidad Politécnica de Madrid, Spain
†Dept. de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, Barcelona, Spain
‡Praktische Informatik VI, FernUniversität Hagen, Germany
§Dept. de Informática, Universidad de Valladolid, Spain
¶Dept. de Matemáticas, Universidad de Alcalá, Spain

closed boundaries of F_R^ε are called the *outer and inner ε-offsets* of R. The ε-frame of R is also the Minkowski sum of R and the unit circle of the underlying metric scaled by ε.

In this paper we concentrate on axis-parallel rectangles with or without prescribed aspect ratio, and we will use the L_∞-distance as underlying metric, which has axis-parallel squares as unit circles. Therefore, the outer and inner offsets of a rectangle are also rectangles, see Figure 1.

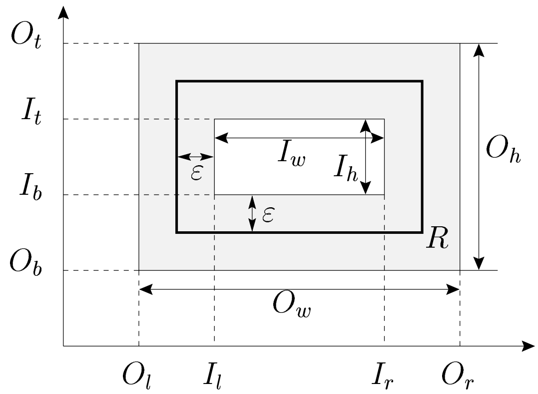

Figure 1: The ε-frame of R and its edges.

The *best fitting rectangle problem* is equivalent to looking for the narrowest covering ε-frame, i. e., the frame with the smallest ε that covers all sites.

3 Arbitrary aspect ratio

If we do not prescribe a certain aspect ratio of the rectangle then to find the best fitting rectangle is an easy problem.

Lemma 1 *Let R be a rectangle whose ε-frame covers P and B the bounding box of P. Then we have for all $p \in P$*

$$d(p, B) \leq 2\varepsilon \ .$$

Proof. Let $p \in P$ and let r be the point of R closest to p (one of them if there are several); we know $d(p, r) \leq \varepsilon$.

The (horizontal or vertical) line through p and r intersects B in two points. It is clear that at least one of the two is not farer than ε away from r because otherwise one of the sites on B would not be covered by the ε-frame of R, thus $d(r, B) \leq \varepsilon$.

We can combine the two inequalities and use the triangle inequality to obtain

$$d(p, B) \leq d(p, r) + d(r, B) \leq 2\varepsilon \ ,$$

which is our claim. □

Now an optimal solution can be obtained as follows.

- Compute B, the bounding box of P, this will turn out to be the outside of an optimal ε-frame.
- Compute $\max_{p \in P} d(p, B)$; call this 2ε.
- Let R be the inner ε-offset of B; this is a best fitting rectangle.

To prove the correctness of the algorithm, which clearly runs in time $O(n)$, it suffices to say that F_R^ε covers all sites and that there is no covering frame of a smaller width, by Lemma 1.

Remark that in most cases the best fitting rectangle with arbitrary aspect ratio is not unique.

4 Given aspect ratio

Now the aspect ratio, $a = \frac{height}{width}$, of the considered rectangles is also given. The problem is more complicated because the bounding box is no longer such a direct key to the solution. Nevertheless, we have the following property.

Lemma 2 *The inner and outer ε-offsets of any best fitting rectangle with given aspect ratio, a, contain a point of P. There is always an optimal solution which contains points of P on at least four of its eight offset edges.*

For the position of the points on the four offset edges, a lot of cases seem to be possible, at first sight. By the next lemma, we reduce the number of cases to three.

As a short and precise notation, we introduce the following abbreviations. For the Y-coordinates of the horizontal offset edges of a certain frame we say O_t, O_b, I_t, and I_b to the outer and inner top and bottom edges, and for the X-coordinates of the vertical offset edges we say O_l, O_r, I_l, and I_r to the outer and inner left and right edges, see Figure 1.

Lemma 3 *There is always a best fitting rectangle with given aspect ratio, a, possibly after a rotation of P by $\pm 90°$ or $180°$, that corresponds to one of three main cases, see Figure 2:*

Case 1 *O_t, O_b, and I_t are determined by points of P.*

Case 2 *O_t, O_b, I_l, and I_r are determined by points of P.*

Case 3 *O_t, I_t, and I_b are determined by points of P.*

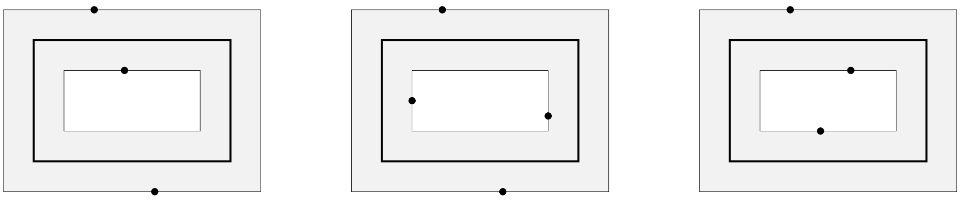

Figure 2: The three main cases for best fitting rectangles.

Our main result says that there is a simple algorithm to solve these cases.

Theorem 4 *A best fitting rectangle with given aspect ratio, a, for n points can be computed in time $O(n \log n)$.*

Proof. Corresponding to the three cases of Lemma 3 and four main orientations (rotations), our algorithm will find a best fitting rectangle in three main steps, each of which must be executed four times. All steps are independent from each other. For the simplicity or the description we assume that no two points of P have identical X- or Y-coordinates. Nevertheless the treatment of the general case is not difficult at all.

Due to the lack of space we sketch only the **algorithm for Case 1**. We perform a sweep from the middle of the interval (O_b, O_t) simultaneously to the top and to the bottom.

$O_t = y_{\max}$, $O_b = y_{\min}$, $O_h = O_t - O_b$, they correspond to the bounding box.
For all points $p = (x, y)$ of P compute $\varepsilon(p) = \frac{1}{2}\min(O_t - y, y - O_b)$.
Sort the points by decreasing ε-value and renumber the points, p_1, p_2, $\ldots$, in that order.

Let $\varepsilon_i = \varepsilon(p_i)$ and $\varepsilon_{start} = \frac{O_t - O_b}{4} \min(1, \frac{2}{a+1})$.
Let T be an empty balanced tree to contain points according to their X-coordinates.
Insert all points p_i with $\varepsilon_i > \varepsilon_{start}$ into T.
For all remaining points $p_i = (x_i, y_i)$ in this order do
  Let $I_w = \frac{O_h}{a} - 2\varepsilon_i(1 + \frac{1}{a})$ be the width of the inner offset.
  Search the two subsequent $x_l, x_r \in T$ with $x_l < x_i < x_r$.
  Let $h_l = \min(x_{\min}, x_i - 2\varepsilon_i, x_r - 2\varepsilon_i - I_w)$
  and $h_r = \max(x_l - 2\varepsilon, x_i - 2\varepsilon_i - I_w, x_{\max} - 4\varepsilon_i - I_w)$.
  If $h_l \leq h_r$ then we have found a narrower covering frame with
    $O_l = h_l$, $I_l = h_l + 2\varepsilon_i$, $I_r = I_l + I_w$, $O_r = I_r + 2\varepsilon_i$,
    $I_t = O_t - 2\varepsilon_i$, and $I_b = O_b + 2\varepsilon_i$.
  Insert x_i into T.

The algorithm chooses the initial value ε_{start} such that the inner offset is just a line segment, and all points lie between $O_b = y_{\min}$ and $O_t = y_{\max}$. Then ε is decreased and more and more points do no longer lie between O_b and $O_b + 2\varepsilon$ or $O_t - 2\varepsilon$ and O_t. These points are stored in T, and we have to find a position of the vertical edges that corresponds to Case 1, i.e., they lie between O_l and $I_l = O_l + 2\varepsilon$ or $I_r = O_r - 2\varepsilon$ and O_r while the current point p_i must lie between I_l and I_r; the test if $h_l \leq h_r$ takes care of exactly this.

The running time of this algorithm (also for the other cases) is in $O(n \log n)$ since for each point we perform just one insert and one query operation. □

References

[1] G. Barequet, A. J. Briggs, M. T. Dickerson, and M. T. Goodrich. Offset-polygon annulus placement problems. *Comput. Geom. Theory Appl.*, 11:125–141, 1998.

[2] G. Barequet, M. T. Dickerson, and M. T. Goodrich. Voronoi diagrams for convex polygon-offset distance functions. *Discrete Comput. Geom.*, 25(2):271–291, 2001.

[3] M. Bern and D. Goldberg. Paper position sensing. In *Proc. 18th Annu. ACM Sympos. Comput. Geom.*, pages 74–81, 2002.

[4] B. Chazelle, R. L. Drysdale, III, and D. T. Lee. Computing the largest empty rectangle. *SIAM J. Comput.*, 15:300–315, 1986.

[5] J. M. Díaz-Bañez, J. A. Mesa, and A. Schöbel. Continuous location of dimensional structures. Report in Wirtschaftsmathematik 79, Department of Mathematics, Universität Kaiserslautern, Germany, 2002.

[6] Z. Drezner and H. W. Hamacher, editors. *Facility Location: Applications and Theory.* Springer-Verlag, Berlin, Germany, 2002.

[7] J. García-López, P. Ramos, and J. Snoeyink. Fitting a set of points by a circle. *Discrete Comput. Geom.*, 20:389–402, 1998.

[8] M. Orlowski. A new algorithm for the largest empty rectangle problem. *Algorithmica*, 5:65–73, 1990.

[9] R. K. Walker and V. Srinivasan. Creation and evolution of the ASME Y14.5.1M standard. *Manufacturing Review*, 7(4):16–23, 1994.