Probabilistic Matching of polygons*

Helmut Alt† Ludmila Scharf† Daria Schymura†

Abstract

We analyze a probabilistic algorithm for matching plane compact sets with sufficiently nice boundaries under translations and rigid motions (rotation and translation). Given shapes A and B, the algorithm computes a transformation t such that with high probability the area of overlap of $t(A)$ and B is close to maximal. We give a time bound that does not depend on the number of vertices in the case of polygons.

1 Introduction

Shape matching is a problem that occurs in various areas of computer science and in various flavors. The variant considered here comes from computer vision.

We model shapes as bounded polygonal regions in the plane. Given two shapes A and B, as well as a set of transformations F and a distance measure d, we look for the transformation $t \in F$ such that $t(A)$ and B match optimally with respect to d. Two shapes are similar if there is a transformation f for that the distance between $t(A)$ and B is small.

This is a well-studied problem in the case of rigid motions and the Hausdorff distance, if the shapes are modeled as sets of line segments, for example. For a survey on the topic, see [3].

We study the case where F is the set of translations or rigid motions (rotation and translation) in the plane and the distance measure is the area of the symmetric difference, that is the area that belongs to exactly one of the shapes. Minimizing the area of the symmetric difference under translations or rigid motions is the same as maximizing the area of overlap, and that is what we will speak about from now on. The area of overlap is a well-known similarity measure that is insensitive to noise.

There are efficient algorithms that maximize the area of overlap under translations. Mount et al. [9] show that the maximal area of overlap of a simple n-polygon with a translated simple m-polygon can be computed in $O(n^2m^2)$ time. Recently, Cheong et al. [7] gave a general probabilistic framework that computes an approximation with prespecified absolute error ϵ in $O(m + \frac{n^2}{\epsilon^4} log(n)^2)$ time for translations and $O(m + \frac{n^3}{\epsilon^4} log(n)^5)$ time for rigid motions. De Berg et al. [6] consider the case of convex polygons and give a $O((n+m)log(n+m))$ time algorithm that maximizes the area of overlap. Alt et al. [2] give a linear time constant factor approximation algorithm for minimizing the area of the symmetric difference under translations and homotheties (scaling and translation).

Surprisingly little is known about maximizing the area of overlap in case of rigid motions and similarities.

Here, we analyze a probabilistic algorithm that approximates the maximal area of overlap under translations and rigid motions. Given an allowable error ϵ and a desired probability of success p, we show bounds on the required number of random experiments, guaranteeing that the difference between approximation and optimum is at most ϵ with probability at least p.

This algorithm is a special case of a probabilistic algorithmic scheme for approximating an optimal match of compact sets under a subgroup of affine transformations. Alt and Scharf [4] analyzed an instance of this algorithmic scheme that compares polygonal curves under translations, rigid motions, and similarities.

2 The Algorithm

The idea of the algorithm is to draw random points from the shapes, to compute a transformation that maps the points onto each other, and to keep this transformation, called a "vote", in mind. This is repeated very often; in each step, we grow our collection of "votes" by one. Clusters of "votes" indicate transformations that map large parts of the shapes onto each other.

Now we state the algorithm for translations.

Given: shapes A and B, integer n, positive real δ.

1. Do the following experiment n times:
   Draw uniformly distributed random points $a \in A$ and $b \in B$. Give one "vote" to the unique translation that maps a onto b.

2. Determine and return one of the translations whose δ-neighborhood obtained the most "votes".

The term δ-neighborhood refers to the maximum norm; we identify each translation with its translation vector and equip the translation space $\mathbb{R}^2$ with the maximum norm.

*This research was supported by the European Union under contract No. FP6-511572, Project PROFI.

†Institute of Computer Science, Freie Universität Berlin, {alt,scharf,schymura}@inf.fu-berlin.de

The algorithm captures the intuitive notion of matching. Translations that many pairs of points vote for should be "good" translations since many points from A are mapped onto points from B. Figure 1 illustrates this idea.

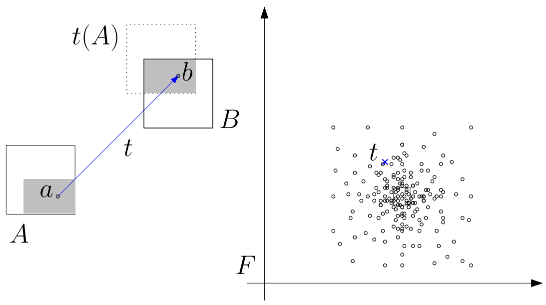

Figure 1: We compare two copies of a square under translations. The area of overlap of $t(A)$ and B corresponds to the chance of choosing a point pair $(x, y) \in A \times B$ such that $y - x = t$.

Now we explain the algorithm for rigid motions. The space of rigid motions R is given as $[-\pi, \pi) \times \mathbb{R}^2$, equipped with the maximum norm. A point $(\alpha, t) \in R$ denotes the rigid motion

$$x \mapsto M_\alpha x + t, \quad M_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

For matching under rigid motions, we draw in each step uniformly distributed an angle α and random points $a \in A$ and $b \in B$. We give one "vote" to the unique rigid motion with counterclockwise rotation angle α that maps a onto b, namely

$$x \mapsto M_\alpha x + (b - M_\alpha a).$$

When speaking of transformations, we refer to translations and rigid motions.

The algorithm does not require the shapes to be polygons; it works for measurable sets in $\mathbb{R}^2$, provided there is a method to draw uniformly distributed random points from them and the density function is Lipschitz continuous (see Section 4.3). The latter is fulfilled, for example, if the shapes' boundaries are unions of piecewise differentiable simple closed curves. The algorithm can be directly applied on bitmap data as well.

3 Main Result

We study the transformation space with the probability distribution that is implicitly given by the random experiment and the fact that we draw the random points from the shapes in a uniformly distributed way.

The distribution of "votes" in the transformation space approximates the probability distribution that results from the experiment; the fraction of "votes" that fall in a set approximates its probability. The output of the algorithm is a transformation t whose δ-neighborhood has approximately highest probability.

The probability Pr of an event E and the density function g are by definition related by

$$Pr(E) = \int_E g(s) ds.$$

If the density function is uniformly continuous and δ is small enough, the transformations whose δ-neighborhood have highest probability and the transformations at which the density function is maximal are the same. Then, the output is a transformation t at which the density function is close to the maximum with high probability.

It turns out, that in our case the density function is the function mapping a transformation vector to the area of overlap of the transformed shape A and B divided by a constant.

Let μ be the Lebesgue measure on $\mathbb{R}^2$, which for polygons equals the area. We always assume the shapes not to be degenerate.

Lemma 1 *The density function of the probability distribution on the translation space that results from the experiment is given by*

$$f(t) = \frac{\mu(t(A) \cap B)}{\mu(A)\ \mu(B)}.$$

In case of rigid motions, the density function is given by

$$g(r) = \frac{\mu(r(A) \cap B)}{2\pi\ \mu(A)\ \mu(B)}.$$

The main result is the following approximation theorem, which states that, for each allowable error ϵ and each desired probability of success p, there is a number of experiments N guaranteeing approximation of the maximal area of overlap with error at most ϵ and with probability at least p.

Theorem 2 *Let t^* be a transformation that is output of the algorithm and t^{opt} a transformation that maximizes the area of overlap of $t^{opt}(A)$ and B. Let $\epsilon > 0$ and $p < 1$. There are a positive real $\delta = O(\epsilon)$ and an integer N such that with probability at least p*

$$|\mu(t^*(A) \cap B) - \mu(t^{opt}(A) \cap B)| < \epsilon.$$

if N is the number of random experiments.

In case of translations,

$$N = O(\max\{-\log(1-p), -\log(\epsilon^6)\}/\epsilon^6).$$

In case of rigid motions,

$$N = O(\max\{-\log(1-p), -\log(\epsilon^8)\}/\epsilon^8).$$

The constants depend on the shapes.

A sketch of the proof is given in Section 4.4.

In the case of polygons, after having triangulated, we can draw random points in constant time. The runtime is bounded trivially by $O(N^2)$ for translations and by $O(N^3)$ for rigid motions, if N is the required number of experiments, since the arrangement of δ-spheres whose center is a "vote" can be built and traversed within this time bound. In contrast to all other known approaches, the runtime does not depend on the number of vertices; it is insensitive to fine sampling.

4 Analysis of the algorithm

For a transformation t, let $B_\delta(t)$ be the δ-neighborhood of t in the maximum norm. Recall that $B_\delta(t)$ is two-dimensional in the case of translations and three-dimensional in the case of rigid motions. Denote by $X_n^\delta(t)$ the fraction of "votes" that lies δ-close to t after n experiments. Intuitively, it is not surprising that the following relations hold

$$X_n^\delta(t) \underset{n\ large}{\approx} Pr(B_\delta(t)) \underset{\delta\ small}{\approx} \mu(B_\delta(t))\ g(t).$$

4.1 Estimate the probability of a fixed δ-ball

The estimate $X_n^\delta(t)$ is called naive estimator in the theory of density estimation [10]. The next lemma states that for each transformation t the estimate $X_n^\delta(t)$ is close to $Pr(B_\delta(t))$ with high probability; it can be proven by using the Chernoff bound, as stated in [7].

Lemma 3 *For each transformation t and for all $0 < \epsilon < 1$ the following holds*

$$Pr(|X_n^\delta(t) - Pr(B_\delta(t))| > \epsilon) < 2e^{-\frac{\epsilon^2 n}{2}}.$$

Later, we will have to show that the output of the algorithm is a transformation that approximately maximizes $Pr(B_\delta(t))$. The latter is no obvious corollary of Lemma 3 since the output transformation t is a random vector depending on the sequence of experiments.

4.2 Density function

In this section we show the proof of Lemma 1 for rigid motions; the proof for translations is a simpler version of it.

We will use the following special case of a transformation formula for density functions of random variables, which can be found in most introductory books about probability theory, for example in [8].

Theorem 4 *Let $X : \mathbb{R}^n \to \mathbb{R}^n$ be a random variable with density function f_X and $h : \mathbb{R}^n \to \mathbb{R}^n, h : x \mapsto Mx$ a linear map with $\det(M) \neq 0$. Then $h \circ X$ has the density function*

$$f_{h\circ X}(y) = f_X(M^{-1}y)\ |\det(M^{-1})|.$$

Proof. (of Lemma 1) Our random experiment consists in drawing uniformly distributed points from $\Omega = I \times A \times B$ where $I = [-\pi, \pi)$. We are interested in the density function f_X of the random variable

$$X : \Omega \to R, \quad X : (\alpha, a, b) \mapsto (\alpha, b - M_\alpha a).$$

Drawing the counterclockwise rotation angle uniformly distributed in I corresponds to the random variable id_I with density function $f_I(\alpha) = \frac{1}{2\pi}$.

Determining the translation vector t depends on the rotation angle α. First, we compute the density function f_α of the random variable X_α that is defined as follows:

$$X_\alpha : A \times B \to \mathbb{R}^2, \quad (a, b) \mapsto b - M_\alpha a.$$

Understanding f_α as conditional density $f_X(\alpha, \cdot)$ on $\mathbb{R}^2$ gives then

$$f_X(\alpha, t) = f_I(\alpha)\ f_\alpha(t).$$

Therefore it suffices to compute f_α.

The function $h : \mathbb{R}^4 \to \mathbb{R}^4, h : (a, b) \mapsto (a, b - M_\alpha a)$ is a linear map with determinant 1. Let $\pi : \mathbb{R}^4 \to \mathbb{R}^2$ be the projection on the third and forth coordinate, then $X_\alpha = \pi \circ h \circ id_{A\times B}$. For a set $E \subset X$, let $\chi_E : X \to \{0, 1\}$ be its characteristic function that is one iff $x \in E$. We know

$$f_{id_{A\times B}}(a, b) = \frac{\chi_A(a)\chi_B(b)}{\mu(A)\ \mu(B)}.$$

Using Theorem 4, we get

$$f_{h\circ id_{A\times B}}(a, b) = \frac{\chi_{A\cap(M_{-\alpha}(B-b))}(a)}{\mu(A)\ \mu(B)}.$$

Now we can compute f_α, which proves the theorem:

$$\begin{aligned} f_\alpha(t) &= \int_A f_{h\circ id_{A\times B}}(a, t) da \\ &= \frac{\mu((M_\alpha A + t) \cap B)}{\mu(A)\ \mu(B)}. \end{aligned}$$

□

4.3 Lipschitz continuity of the density function

A function f from a metric space M to $\mathbb{R}$ is called Lipschitz continuous if there is a constant L such that for all $x, y \in M$ holds

$$\|x - y\| < \delta \Longrightarrow |f(x) - f(y)| < L\delta$$

Denote by μ_δ the Lebesgue measure of the δ-neighborhood of a transformation t in the metric induced by the maximum norm. The number μ_δ does not depend on t since the Lebesgue measure is invariant under translations and rotations. For translations, $\mu_\delta = 4\delta^2$ and for rigid motions, $\mu_\delta = 8\delta^3$. We are interested in the density functions to be Lipschitz continuous because then holds

$$|Pr(B_\delta(t)) - \mu_\delta\ f(t)| \leq L\mu_\delta\delta.$$

Translations. Let $\Delta_{A,B}$ the sum of the lengths of the boundaries of A and B.

It is easy to show, that for the density function f on the translation space, $\sqrt{2}\Delta_{A,B}/(\mu(A)\mu(B))$ is a Lipschitz constant. Observe that the constant depends heavily on the shapes.

Rigid Motions. Let D_B is the minimum radius of a ball around the origin that contains B and D_A the analogue for A. Standard geometric arguments show that for the density function g on the space of rigid motions, $(2D_B + D_A)\Delta_{A,B}/(2\pi\mu(A)\mu(B))$ is a Lipschitz constant.

4.4 Sketch of the main result's proof

We have already seen that for fixed t

$$X_n^\delta(t) \quad \underset{n\ large}{\approx} \quad Pr(B_\delta(t)) \quad \underset{\delta\ small}{\approx} \quad \mu_\delta\ g(t).$$

Obviously, in the approximation process two errors are involved. One we call the sampling error; it becomes smaller if the number of experiments increases and can be bounded by the Chernoff inequality. The other we call the smoothing error; it becomes smaller if δ decreases and can be bounded by the Lipschitz continuity of the density function. Instead of smoothing we could discretize the shapes and would end up with a discretization error added to the sampling error. This would simplify the analysis a little but could not be generalized so nicely to other transformation groups.

Now we need to analyze what happens if the transformation vector is determined by the sequence of random experiments, namely the vector whose δ-neighborhood obtains the most "votes", and thus is a random vector itself.

The output of the algorithm can be modeled as random variable

$$Z_n^\delta = \max_{t\in\mathbb{R}^2} X_n^\delta(t).$$

Let $S = (s_1, \ldots, s_n)$ be a sequence of transformations from the random experiments. Consider the arrangement induced by the boundaries of $B_\delta(s_1), \ldots, B_\delta(s_n)$, which are the δ-spheres of the points in S. The depth of a cell is defined as the number of $B_\delta(s_i)$ it is contained in. The candidates for the output of the algorithm are the transformations corresponding to the deepest cells in this arrangement. A transformation t lies in the intersection of k of the neighborhoods if and only if its neighborhood contains k "votes".

The next lemma can be proven using an idea of [7].

Lemma 5 *Let V be a set of points such that V contains for each cell of the arrangement induced by the δ-spheres with centers in S one point of its lowest-dimensional face. There is such a V that contains at most n^2 points in the case of translations and n^3 in case of rigid motions. For each $\epsilon > 0$ and all $n \geq \frac{6}{\epsilon} + 2$ it holds that*

$$Pr(\exists t \in V : |X_n^\delta(t) - Pr(B_\delta(t))| > \epsilon)$$

is less than $2n^2 e^{-\frac{\epsilon^2(n-2)}{8}}$ in case of translations and less than $2n^3 e^{-\frac{\epsilon^2(n-3)}{16}}$ in case of rigid motions.

Using Lemma 5 and the Lipschitz continuity of the density functions, the main result can be proven.

We note that the proof provides explicit bounds for the required number of experiments to ensure approximation with error at most ϵ with probability at most p.

References

[1] H.-K. Ahn, O. Cheong, C.-D. Park, C.-S. Shin and A. Vigneron. *Maximizing the overlap of two planar convex sets under rigid motions.* Computational Geometry, 37(1):3–15, 2007.

[2] H. Alt, U. Fuchs, G. Rote and G. Weber. *Matching convex shapes with respect to the symmetric difference.* Algorithmica, 21:89–103, 1998.

[3] H. Alt and L. Guibas. *Discrete Geometric Shapes: Matching, Interpolation, and Approximation. A Survey.* Handbook of Computational Geometry, eds. J.-R. Sack and J. Urrutia, pages 121–153. Elsevier Science Publishers B.V. North-Holland, Amsterdam, 1999.

[4] H. Alt and L. Scharf. *Shape matching by random sampling* Technical Report Freie Universität Berlin B 07-xx, 2007.

[5] H. Alt, L. Scharf and S. Scholz *Probabilistic matching of sets of Polygonal curves* Proceedings of the 22nd European Workshop on Computational Geometry (EWCG), 107–110, March 2006, Delphi, Greece.

[6] M. de Berg, O. Devillers, M.J. van Kreveld, O. Schwarzkopf and M. Teillaud. *Computing the maximum overlap of two convex polygons under translations.* Theory of computing systems, 31:613–628,1998.

[7] O. Cheong, A. Efrat and S. Har-Peled. *Finding a guard that sees most and a shop that sells most.* Discrete and Computational Geometry, 37(4):545–563, 2007.

[8] A. Klenke. *Wahrscheinlichkeitstheorie.* Springer, 2006, Berlin.

[9] D.M. Mount, R. Silverman and A.Y. Wu. *On the area of overlap of translated polygons.* Computer Vision and Image Understanding: CVIU, 64(1):53–61, 1996.

[10] B.W. Silverman. *Density Estimation for Statistics and Data Analysis.* Chapman & Hall, 1986, New York.