Shape matching by random sampling*

Helmut Alt, Ludmila Scharf

Abstract

We present a generic probabilistic algorithm based on random sampling for matching shapes which are modelled by sets of curves. The algorithm can be applied to translations, rigid motions, and similarity maps as possible sets of transformations. We analyze which similarity measure is optimized by the algorithm and the number of samples necessary to get a prespecified approximation to the optimal match within a prespecified probability.

1 Introduction

Our research is motivated by the task of automated retrieval of figurative images in large databases, where the evaluation of the similarity of two images is based on their geometric shape and not color or texture. Matching two geometric shapes under transformations and evaluating their similarity is one of the central problems in such retrieval systems, but it is also a problem of independent interest which is widely covered in literature, see [4, 16, 23, 14] for surveys.

We assume that the shapes are modeled by sets of plane curves. As possible classes of transformations we will consider *translations*, *rigid motions* (rotation and translation) and *similarities* (rotation, scaling and translation). Our objective is to develop an algorithm which comes close to human similarity perception and allows an efficient implementation for the retrieval system.

Several similarity measures and algorithms are known to match two curves, especially polygonal curves. One of the "universal" similarity measures is the Hausdorff distance which is defined for any two compact sets A and B. Alt et al. describe in [2, 4] efficient algorithms for computing the Hausdorff distance and minimizing it under translations and rigid motions for arbitrary sets of line segments. One of the drawbacks of the Hausdorff distance is that it is very sensitive to noise. A few similarity measures are defined for pairs of curves, which capture the relative course of two curves: Fréchet distance [4], turning function distance [6], and dynamic time warping distance [11]. There are no generalizations of those distances to sets of curves, although in [5] a generalization of the Fréchet distance to geometric graphs is given, and in [21] Tanase et al. describe an algorithm for matching

*This research was supported by the European Union under contract No. IST-511572-2, Project PROFI.

a set of polygonal curves to a single polygon. A similarity measure which is designed for sets of curves is the reflection visibility distance [13]. The reflection visibility distance is robust against different kinds of disturbances but is expensive to compute. Some of these similarity measures can be modified to valuate partial match, e.g., percentile-based Hausdorff distance [14].

The method we introduce is close to an intuitive notion of "matching", i.e., find one or more candidates for the best transformations, that when applied to the shape B map the most similar parts of the two shapes to each other. The major idea is to take random samples of points from both shapes and give a "vote" for that transformation (translation, rigid motion, or similarity) matching one sample with the other. If that experiment is repeated frequently, we obtain by the votes a certain probability distribution in the space of transformations. Maxima of this distribution indicate which transformations give the best match between the two figures.

The matching step of our algorithm is, therefore, a voting scheme. Voting based methods are widespread in model based pattern recognition. Related approaches include geometric hashing [24], alignment methods [15], and the generalized Hough transform, also called pose clustering [1, 18, 19, 20] and generalized Radon transform [12, 22].

Generalized Hough transform (and Radon transform) is an approach closely related to our algorithm. The general setting is: given a model and an image, the goal is to determine whether the image contains a possibly transformed occurrence of the model. The set of possible transformations is considered as a cluster space according to a suitable parametrization. For each minimum size set of (feature-)points in the image and in the model, that is, sets of points of such size that a transformation mapping one set to the other is uniquely determined, an "evidence" for the transformations mapping model features to the image features is gathered in the cluster space. The size of the feature set depends on the allowed transformation class. When all image features have been considered, each element of the usually discretized cluster space contains a measure of consistency of the transformation.

Most of the work done in this area, though, considers image feature space and sometimes also the model feature space to be discrete, which is natural in the field of computer vision. In our case, the two shapes we want to match are continuous sets of points corresponding to the planar curves. Therefore, the pure enumeration of image features and gathering of evidences for every feature is not applicable. Instead, we perform a randomized approximation of the evidence distribution in the transformation space. Another feature common to most pose clustering methods is the discretization of the transformation space. The usual clustering techniques used together with pose clustering are binning and multidimensional histogramming. In contrast to those methods we do not consider a discrete set of features that describe shapes, but work with continuous curves. Our method is independent of the choice of parameterization and discretization grid in transformation space.

The main contribution of this work is the theoretical analysis of the probability distribution induced in transformation space. We give rigorous bounds on the runtime (number of experiments) necessary to obtain the optimal match within a certain approximation factor with a prespecified probability.

2 Probabilistic matching

We assume that shapes are modelled by sets of curves, and that for each curve a random point under uniform distribution can be generated in constant time. This is the case for line segments, which would be the most common representation in practice, but also for curves for which a natural parameterization is explicitly given. Given two shapes $A, B \subset \mathbb{R}^2$, a class of allowed transformations $\mathcal{T}$ and a certain parameter δ, we want to find a transformation $t \in \mathcal{T}$ which lets the transformed image of B, $t(B)$, *match best* A within a tolerance of δ. Distances in image space are measured with respect to some standard metric L_p, e.g., Euclidean metric (L_2) or maximum distance (L_∞).

The definition of what exactly a good match means is given in section 3 for the general case and in more details for different transformation classes in section 4. Here we follow an intuitive notion of a "good match": two shapes are similar if they can be mapped to each other in such a way that large parts of them are close, this position is then a good match. So we are searching for a transformation that maps the most similar parts of the shapes A and B to each other.

Selecting a random point under uniform distribution from a set of curves can be reduced to selecting a random number under uniform distribution on an interval of real numbers: Select a random curve from the set, where the probability for a curve is weighted with the relative curve length. Then, consider the parameterization by length of the curve and select uniformly a random parameter value from the parameter interval, take the point corresponding to the selected parameter value.

The idea of the *probabilistic approach* is quite simple. We first describe an algorithm for matching under translations:

1. Take a random point a from the shape A and a random point b from B and give one "vote" to the translation t which maps b to a, that is $t = a - b$.

2. Repeat this experiment many times. Then the distribution of votes in the two-dimensional translation space $\mathcal{T}$ approximates a certain probability distribution.

3. For a prespecified neighborhood size δ return the points of $\mathcal{T}$ with the highest number of votes in their δ-neighborhood as candidates for good transformations.

The idea behind this algorithm, is that the transformations, that map large parts of shapes to each other should get significantly more votes than others. The size of the δ-neighborhood influences the quality of the match.

For more complex classes of transformations two points are not sufficient to determine a unique transformation, therefore, several points or a point and a direction vector might be necessary to form a *random sample* in the first step of the algorithm. The size of a random sample within one experiment and the shape and the size of the δ-neighborhoods depend on the class of transformations allowed. Further, a "vote" for a transformation is a *δ-region* in the transformation space, which is defined as a set of transformations that map each element of random sample S_B into a δ-neighborhood of the corresponding element of

S_A. For transformation classes other than translations the shape of a δ-region depends on the sample pair generating it.

Before giving a generic variant of the algorithm we briefly describe the random samples and δ-regions for the different transformation classes:
For *translations*, as described above, a random sample consists of a single randomly selected point of each shape, $a \in A$ and $b \in B$. Two points determine uniquely a translation mapping one point to the other. The transformation space is two-dimensional and the δ-neighborhood of a translation vector t is defined as a set of vectors that have distance at most δ to vector t, where distance is measured with respect to the metric chosen in image space.

In case of *rigid motions* the transformation space is three-dimensional. A rigid motion $t = (\alpha, v_x, v_y)$ is defined by a rotation angle α and a translation $v = (v_x, v_y)$ and maps a point $b \in \mathbb{R}^2$ to a point $t(b) = Mb + v$, where

$$M = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

is the rotation matrix.

We will consider two approaches for randomized matching under rigid motions.

Approach 1: A random sample of a shape within one experiment, that is, in the first step of the algorithm, contains random points $a \in A$ and $b \in B$ and for each of the points an angle θ_a and θ_b respectively defined by the direction of the tangent line at this point, i.e., $S_A = (a, \theta_a)$, $S_B = (b, \theta_b)$. Two such point-angle pairs define uniquely a rigid motion $t = (\alpha, v_x, v_y)$, such that $\theta_b + \alpha = \theta_a$ and $t(b) = a$. Here we have a special case where a sample consists of different types of data: a point and an angle. So it is reasonable to have different values for the tolerance bounds: $\delta = (\delta_1, \delta_2)$, where δ_1 defines a neighborhood of points, and δ_2 restricts the maximal allowed difference of directions. Thus, a δ-region corresponding to a sample pair is the set of rigid motions t' such that the angle between the rotated tangent at b and the tangent at a differ by at most δ_2 and the distance between $t'(b)$ and a is at most δ_1.

Approach 2: We use a single random point of each shape $a \in A$ and $b \in B$ as a sample in one random experiment and record a δ-region in the space of rigid motions as the set of transformations that map the point b into the δ-neighborhood of point a, while all rotation angles are allowed.

The δ-region corresponding to a sample pair (a, b) in the second approach has the shape of a spiral tube extending from 0 to 2π in the direction of the rotation axis, where for each value α the cross-section parallel to the translation plane has the shape of the δ-neighborhood with respect to the chosen metric in image space, see section 4.2 for more details. In the first approach, a δ-region corresponding to a sample pair $((a, \theta_a), (b, \theta_b))$ is a part of the described spiral tube defined by points a, b and tolerance δ_1 which is bounded by the planes $\alpha = \theta_a - \theta_b + \delta_2$ and $\alpha = \theta_a - \theta_b - \delta_2$.

We also experimented with yet other approaches for rigid motions but the two we present here give good results and fit into the general framework of our analysis.

For *similarity maps* the transformation space is four-dimensional. A similarity map $t = (\alpha, k, v_x, v_y)$ is defined by a rotation angle α, scaling factor k and a translation vector $v = (v_x, v_y)$. t maps a point $b \in \mathbb{R}^2$ to a point $t(b) = Mb + v$, where

$$M = \begin{pmatrix} k\cos\alpha & -k\sin\alpha \\ k\sin\alpha & k\cos\alpha \end{pmatrix} = \begin{pmatrix} m_1 & -m_2 \\ m_2 & m_1 \end{pmatrix}.$$

A random sample from the shapes contains two points $S_A = (a_1, a_2)$ of A, and two points $S_B = (b_1, b_2)$ of B, which determine a unique similarity transformation t mapping b_1 to a_1 and b_2 to a_2. Although a standard way to parameterize the space of similarity transformations is by (α, k, v_x, v_y), for computational reasons it is more convenient to use the parameterization (m_1, m_2, v_x, v_y) where $m_1 = k\cos\alpha$ and $m_2 = k\sin\alpha$. For a general L_p metric a δ-region is then bounded by algebraic surfaces, and for the L_1 and L_∞ metrics it is a convex polytope bounded by four pairs of parallel hyperplanes.

A more general class of transformations are the *affine maps*. An affine transformation $t = (M, v)$, where

$$M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}$$

is a matrix and $v = (v_x, v_y)$ is a vector, is defined by six parameters. The transformation space is six-dimensional in this case. An affine transformation t maps a point $b \in \mathbb{R}^2$ to a point $Mb + v$. Three non-collinear points in each shape $a_1, a_2, a_3 \in A$ and $b_1, b_2, b_3 \in B$ determine a unique affine transformation that maps b_i to a_i, $i \in \{1, 2, 3\}$. Therefore, a random sample taken in step one of the algorithm consists of three points of each shape.

Generic probabilistic algorithm: Now we can describe the probabilistic algorithm in a generic way: Given two shapes A and B, a class of allowed transformations $\mathcal{T}$ and a certain parameter δ, we want to find a transformation $t \in \mathcal{T}$ which lets $t(B)$ in some sense match best A within a range of δ:

1. Take random samples S_A from A and S_B from B of an appropriate size s so that there is a unique transformation mapping S_B to S_A (e.g., $s = 1$ for translations and $s = 2$ for similarity maps). Record the δ-region corresponding to this sample pair, that is, the set of transformations that map the elements of S_B into the δ-neighborhoods of their corresponding elements in S_A.

2. Repeat this experiment many times, say N. Then the number of regions covering a point $t \in T$ approximates a certain probability distribution in the transformation space.

3. Take the points of $\mathcal{T}$ covered by the highest number of δ-regions as candidates for good transformations.

In the next section we analyze the probability distribution in transformation space induced by the algorithm and provide the bounds on the number of experiments needed to approximate the maximum of this distribution within a certain factor with a prespecified probability.

3 Generic analysis

3.1 Probability distribution in transformation space

First we introduce some formal notation and definitions. Let Ω denote the sample space, i.e., the set of all sample pairs (S_A, S_B). By the definition of our random experiment, the samples of two figures are drawn independently and uniformly, therefore, we have a uniform distribution on Ω, i.e., for any subset $R \subset \Omega$ the probability of R is $P(R) = \frac{|R|}{|\Omega|}$, where $|\cdot|$ denotes the Lebesgue measure. Further, we say that a transformed sample $t(S_B)$ is in the δ-neighborhood of a sample S_A and write $t(S_B) \in U_\delta(S_A)$ if each element of the sample S_B is mapped into the δ-neighborhood of the corresponding element of S_A by the transformation t. The δ-neighborhoods of the elements can be defined with respect to any metric L_p, in particular with respect to the Euclidean or to the maximum distance.

Let $\mathcal{T} \subset \mathbb{R}^d$ denote the d-dimensional transformation space. Consider the random variable $Y : \mathcal{T} \times \Omega \to \{0, 1\}$ defined as

$$Y(t, (S_A, S_B)) = \begin{cases} 1 & \text{if } t(S_B) \in U_\delta(S_A), \\ 0 & \text{otherwise.} \end{cases}$$

For a transformation t and a sample pair (S_A, S_B) Y indicates whether S_A and S_B match under t, that is, whether t maps S_B into a δ-neighborhood of S_A. By definition, for a fixed transformation t the value of Y is 1 exactly for those sample pairs that are in the δ-neighborhood of each other when the transformation t is applied to B. That means, the pre-image of 1 of the random variable Y for fixed t is the set

$$M_\delta(t) = \{(S_A, S_B) \in \Omega | t(S_B) \in U_\delta(S_A)\} \quad .$$

Therefore, the probability that within one random experiment the value of $Y(t, \cdot)$ is 1 is $P(Y(t, \cdot) = 1) = \frac{|M_\delta(t)|}{|\Omega|}$ which we denote by $p_\delta(t)$.

Thus, we define a similarity measure associated with a transformation t as the Lebesgue measure of the set $M_\delta(t)$ of the possible sample pairs of two shapes that are brought in the δ-neighborhood of each other by t. Intuitively, this should reflect the perceived notion of "closeness" of two shapes, which we could confirm by experiments. We formalize the above observation in the following theorem:

Theorem 3.1. *The probability distribution in the transformation space induced by the generic algorithm described in section 2 has its maximum at the transformation maximizing the Lebesgue measure of the set $M_\delta(t)$ defined as*

$$M_\delta(t) = \{(S_A, S_B) \in \Omega | t(S_B) \in U_\delta(S_A)\} .$$

The role of the parameter δ. In the description of the algorithm we introduced a parameter δ, which defines how far apart two samples are allowed to be that are still considered as being close. The choice of δ, therefore, controls the trade-off between the

quality of match and the size of the parts matched. With a small value of δ our algorithm would find a transformation which maps nearly congruent parts of two shapes to each other. A large value of δ leads to a transformation which gives a rough match but for larger parts of the shapes, see Figure 1.

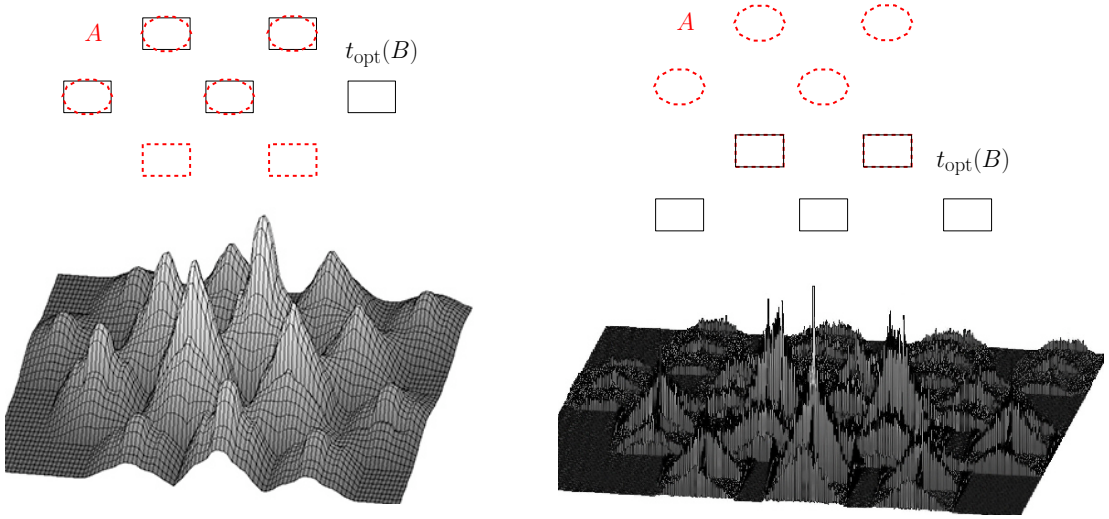


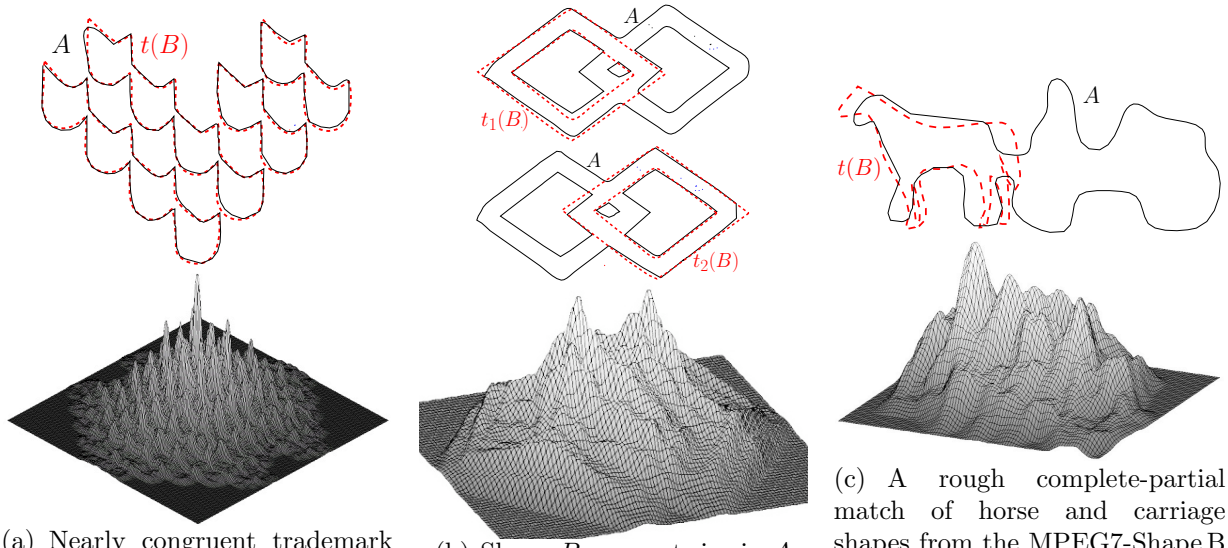
Figure 1: Matching under translations with large (left) and small (right) values of δ and the corresponding probability distributions in translation space.

The value of δ does not alone determine whether the matching is partial or complete or how large the matched parts are, it only specifies how exact the match should be. For nearly congruent figures a small neighborhood size already leads to a complete-complete matching, see Figure 2(a). If figure B is nearly congruent to some parts of A, then still with a small value of δ we detect the occurrences of B in A, that is, find a complete-partial matching as shown in Figures 2(b) and 2(c). The problem of partial-partial matching is not uniquely defined since there is a certain correlation between the quality of match and the size of the matched parts. We address this problem by letting the user specify the quality of match through the choice of δ, for which we then find the matching parts.

For some applications it might be worth to consider several local maxima of the distribution, since they can give us additional information about the shapes. For example, multiple local maxima of the distribution, that are almost equally good, indicate multiple occurrences of one shape, or its parts, depending on the value of the similarity measure, within the other, see Figure 2(b).

3.2 Approximation of the probability distribution

In this section we determine how many samples are needed in order to approximate the probability distribution $p_\delta(t)$ in the transformation space within a certain accuracy ε with high probability and analyze the total running time of the algorithm.



(a) Nearly congruent trademark images

(b) Shape B occurs twice in A

(c) A rough complete-partial match of horse and carriage shapes from the MPEG7-Shape B data set

Figure 2: Matched shapes and the corresponding probability distributions in translation space

In order to find a transformation covered by the highest number of δ-regions corresponding to the samples, we consider the arrangement of these δ-regions. All transformations in the same cell of the arrangement have the same region coverage. Therefore, it is sufficient to traverse the arrangement and take the nodes with the highest number of δ-regions that contain this node.

We first show that the number of δ-regions covering the deepest cell of the arrangement gives a good approximation to the maximum value of the probability distribution. The number of necessary samples is expressed in terms of allowed approximation error ε and maximal probability of failure η. The running time of the algorithm depends on the time needed to determine the deepest cell in the arrangement, which grows exponentially with the dimension of the space, the detailed analysis of running time is given below. We also show that an approximation of the depth of the arrangement as described by Aronov and Har-Peled in [7] results in an approximation of the maximum value of the probability distribution with an error of at most 2ε. The speed-up in determining the deepest cell of the arrangement results directly in a speed-up of the probabilistic matching algorithm.

Let random variable $Z(t)$ denote the sum of the N independent random variables $Y(t,(S_{A1},S_{B1})),\ldots,Y(t,(S_{A\,N},S_{B\,N}))$, as defined in section 3.1, corresponding to the set of N samples in our algorithm. $Z(t)$ counts the number of δ-regions produced by N random experiments that cover t. In the algorithm we find a transformation which is contained in the largest number δ-regions of the sample set, that is, a transformation $t \in \mathcal{T}$ that maximizes the value of $Z(t)$. Let $\tilde{p}_\delta(t)$ denote the ratio of the number of the observed δ-regions that cover t to the total number of samples, that is $\tilde{p}_\delta(t) = \frac{Z(t)}{N}$. $\tilde{p}_\delta(t)$ is an estimate of $p_\delta(t)$.

Using Chernoff bounds, see [17], and the technique described in [9] we will bound the relative error for the estimate of the probability distribution in the transformation space, the proofs are given in section 3.2.1.

The following theorem bounds the number of samples needed for an approximation with a relative error at most ε:

Theorem 3.2. *Given two shapes A and B modeled by finite sets of n curves with total lengths L_A, L_B respectively and a tolerance value $\delta > 0$, for any parameter values ε, η, $0 < \varepsilon, \eta < 1$, the following holds: Let t_{app} be a transformation maximizing $\tilde{p}_\delta(t)$ and t_{opt} a transformation maximizing $p_\delta(t)$. After $N = O\left(\frac{m^2}{\varepsilon^2\delta^2} \ln\left(\max(\frac{1}{\eta}, \frac{m^2}{\varepsilon^2\delta^2})\right)\right)$ random experiments, where $m = \max(L_A, L_B, n\delta)$, the probability that $|\tilde{p}_\delta(t_{\text{app}}) - p_\delta(t_{\text{opt}})| \geq \varepsilon p_\delta(t_{\text{opt}})$ is at most η. (For $\delta = (\delta_1, \delta_2)$ in Approach 1 for rigid motions δ^2 has to be replaced by $\delta_2\delta_1^2$ and in the definition of m, δ by δ_1.)*

Observe that the relative error with respect to $p_\delta(t)$ is also the relative error with respect to $|M_\delta(t)|$, which is the similarity measure underlying our algorithm.

3.2.1 Proof of the relative error bound

We first show that for any transformation vector t and a given threshold ν with high probability we either get a good approximation of $p_\delta(t)$, if this value is at least ν, or otherwise make sure that we do not overestimate it.

Lemma 3.3. *For all $0 < \varepsilon, \nu < 1$ for a sample S of size N and any transformation vector $t \in \mathbb{R}^d$ the following holds:*

- $p_\delta(t) \leq \nu \Rightarrow P(\tilde{p}_\delta(t) > (1+\varepsilon)\nu) \leq e^{-\frac{\varepsilon^2\nu N}{3}}$
- $p_\delta(t) \geq \nu \Rightarrow P(|\tilde{p}_\delta(t) - p_\delta(t)| > \varepsilon p_\delta(t)) \leq 2e^{-\frac{\varepsilon^2\nu N}{4}}$.

Proof. If $p_\delta(t) \leq \nu$:

$$\begin{aligned}
P(\tilde{p}_\delta(t) > (1+\varepsilon)\nu) &= P(Z(t) > (1+\varepsilon)\nu N) = P(e^{rZ(t)} \geq e^{r(1+\varepsilon)\nu N}) \qquad \text{for all } r > 0 \\
&\leq \frac{E(e^{rZ(t)})}{e^{r(1+\varepsilon)\nu N}} \qquad \text{by the Markov inequality [17]} \\
&\leq \frac{e^{(e^r-1)p_\delta(t)N}}{e^{r(1+\varepsilon)\nu N}} \qquad \text{since } r \leq e^r - 1 \\
&\leq \left(\frac{e^{(e^r-1)}}{e^{r(1+\varepsilon)}}\right)^{\nu N} = \left(e^{\varepsilon-(1+\varepsilon)\ln(1+\varepsilon)}\right)^{\nu N} \qquad \text{for } r = \ln(1+\varepsilon) \\
&\leq e^{-\frac{\varepsilon^2\nu N}{3}} \qquad \text{for } 0 < \varepsilon < 1 \ .
\end{aligned}$$

In case $p_\delta(t) \geq \nu$:

$$\begin{aligned} P(|\tilde{p}_\delta(t) - p_\delta(t)| > \varepsilon p_\delta(t)) &= P(|Z(t) - p_\delta(t)N| > \varepsilon p_\delta(t)N) \\ &= P(|Z(t) - E(Z(t))| > \varepsilon E(Z(t))) \\ &\leq e^{-\frac{\varepsilon^2 E(Z(t))}{2}} + e^{-\frac{\varepsilon^2 E(Z(t))}{4}} \end{aligned}$$

by the simplified Chernoff bound [17, Thm. 4.4,4.5]. Since $p_\delta(t) \geq \nu$, we get

$$P(|\tilde{p}_\delta(t) - p_\delta(t)| > \varepsilon p_\delta(t)) \leq 2e^{-\frac{\varepsilon^2 p_\delta(t) N}{4}} \leq 2e^{-\frac{\varepsilon^2 \nu N}{4}},$$

which concludes the proof. □

We associate with each cell C of the arrangement $\mathcal{A}$ of δ-regions a so-called *witness point* p, i.e., a point that lies on a lowest-dimensional face F of $\mathcal{A}$ that contributes to the boundary of C. Observe, that F must be completely contained in the boundary of C and is in general a connected component of the intersection of i boundaries of δ-regions with $1 \leq i \leq d$. Thus, by considering all i-tuples of δ-regions and taking a point in each connected component of their intersection we can be sure to have at least one witness point for each cell of the arrangement.

Now consider a witness point t of the δ-region arrangement corresponding to S. The following lemma states bounds for the witness points of the arrangement.

Lemma 3.4. *For all ε, ν, $0 < \varepsilon, \nu < 1$, and a sample set S of size $N \geq \frac{2d}{\varepsilon\nu} + d$, for a witness point $t \in \mathbb{R}^d$ of the arrangement of the δ-regions corresponding to the samples in S, the following holds:*

- $p_\delta(t) \leq \nu \Rightarrow P(\tilde{p}_\delta(t) > (1+\varepsilon)\nu) \leq e^{-\frac{\varepsilon^2 (N-d)\nu}{12}}$
- $p_\delta(t) \geq \nu \Rightarrow P(|\tilde{p}_\delta(t) - p_\delta(t)| > \varepsilon p_\delta(t)) \leq 2e^{-\frac{\varepsilon^2 \nu (N-d)}{16}}$.

Proof. Observe that Lemma 3.3 cannot be applied to the witness points directly since they depend on the experiment, i.e., the chosen samples. However, since a witness point depends only on $i \leq d$ samples, the remaining $\geq N - d$ samples are "random" for that point and we can apply Lemma 3.3 replacing N by $N - d$. More specifically:

Let $S_1, \ldots, S_i \in S$, $1 \leq i \leq d$, be the sample pairs whose δ-regions induce the witness point t. Consider the sample set $Q = S \setminus \{S_1, \ldots, S_i\}$, $|Q| = N - i$. Let $Z_Q(t)$ and $Z_S(t)$ denote the number of the δ-regions that cover t in sample sets Q and S respectively, and $\tilde{p}_{\delta Q}(t) = Z_Q(t)/(N-i)$, $\tilde{p}_{\delta S}(t) = Z_S(t)/N$. Since we consider closed neighborhoods, $Z_Q(t) = Z_S(t) - i$, $\tilde{p}_{\delta Q}(t) \leq \tilde{p}_{\delta S}(t)$ and

$$\tilde{p}_{\delta Q}(t) = \frac{Z_S(t) - i}{N - i} = \frac{Z_S(t)}{N}\frac{N}{N-i} - \frac{i}{N-i} \geq \tilde{p}_{\delta S}(t) - \frac{i}{N-i} \geq \tilde{p}_{\delta S}(t) - \frac{d}{N-d} \ .$$

Therefore,

$$|\tilde{p}_{\delta S}(t) - p_\delta(t)| \leq |\tilde{p}_{\delta Q}(t) - p_\delta(t)| + |\tilde{p}_{\delta S}(t) - \tilde{p}_{\delta Q}(t)| \leq |\tilde{p}_{\delta Q}(t) - p_\delta(t)| + \frac{d}{N-d} \ .$$

In case $p_\delta(t) \leq \nu$

$$\begin{aligned}
P(\tilde{p}_{\delta S}(t) > (1+\varepsilon)\nu) &\leq P\left(\tilde{p}_{\delta Q}(t) + \frac{d}{N-d} > (1+\varepsilon)\nu\right) = P\left(\tilde{p}_{\delta Q}(t) > (1+\varepsilon)\nu - \frac{d}{N-d}\right) \\
&\leq P\left(\tilde{p}_{\delta Q}(t) > \left(1+\frac{\varepsilon}{2}\right)\nu\right) \qquad \text{for } N \geq \tfrac{2d}{\varepsilon\nu} + d \\
&\leq e^{-\frac{(\varepsilon/2)^2(N-d)\nu}{3}} \qquad \text{by Lemma 3.3} \\
&= e^{-\frac{\varepsilon^2(N-d)\nu}{12}}
\end{aligned}$$

If $p_\delta(t) \geq \nu$:

$$\begin{aligned}
P(|\tilde{p}_{\delta S}(t) - p_\delta(t)| > \varepsilon p_\delta(t)) &\leq P\left(|\tilde{p}_{\delta Q}(t) - p_\delta(t)| + \frac{d}{N-d} > \varepsilon p_\delta(t)\right) \\
&\leq P\left(|\tilde{p}_{\delta Q}(t) - p_\delta(t)| > \frac{\varepsilon}{2} p_\delta(t)\right) \qquad \text{for } N \geq \tfrac{2d}{\varepsilon\nu} + d \\
&\leq 2e^{-\frac{(\varepsilon/2)^2\nu(N-d)}{4}} \qquad \text{by Lemma 3.3} \\
&= 2e^{-\frac{\varepsilon^2\nu(N-d)}{16}}
\end{aligned}$$

□

In the above lemmata we used an additional parameter ν for the smallest value of $p_\delta(t)$ which we want to approximate well enough. Next, we eliminate this parameter and prove Theorem 3.2:

Proof. (Of Theorem 3.2)

First, we show that for an arbitrary sequence of N random experiments performed by the algorithm, the probability that there exists a witness point t' corresponding to this sequence, for which the estimate $\tilde{p}_\delta(t')$ of $p_\delta(t')$ is bad, can be bounded by a parameter η.

As we have seen above it is sufficient to consider k-tuples of the δ-regions with $1 \leq k \leq d$ in order to have at least one witness point in each cell of the arrangement. Any such tuple produces at most a constant number of witness points. We can enumerate all k-tuples of the δ-regions and, thus, all possible witness points corresponding to an arbitrary sequence of N experiments. For an arbitrary witness point t_{ij} (a witness point number j of the i-th δ-region tuple) we can apply Lemma 3.4 and, hence bound the probability that a sequence of N random experiments results in a bad approximation for t_{ij}. Note, that if for some set of random experiments the i-th tuple does not have a witness point number j or does not have any witness points, the statement of the lemma trivially holds. In other words, according to Lemma 3.4 the probability that an arbitrary N-sequence gives a bad approximation for a witness point j produces by i-th tuple is at most $2e^{-\frac{\varepsilon^2\nu(N-d)}{16}}$.

Since there are at most N^d k-tuples of the regions and, therefore, at most $c_1 N^d$ witness points, where c_1 is a constant, we have to apply the lemma at most $c_1 N^d$ times. Then the probability that there exists a witness point t with $p_\delta(t) \geq \nu$ and $|\tilde{p}_\delta(t) - p_\delta(t)| > \varepsilon p_\delta(t)$ or

with $p_\delta(t) < \nu$ and $\tilde{p}_\delta(t') > (1+\varepsilon)\nu$ is, according to Lemma 3.4, at most $N^d 2e^{-\frac{\varepsilon^2\nu(N-d)}{16}}$. A straightforward calculation shows that for $N \geq \frac{c'}{\varepsilon^2\nu}\ln\left(\frac{1}{\varepsilon^2\nu}\right)$ with some suitable constant c' this value is at most $e^{-\frac{\varepsilon^2\nu(N-d)}{32}}$, which is less than η for $N \geq \frac{32}{\varepsilon^2\nu}\ln\frac{1}{\eta} + d$. So the probability that there exists a witness point, for which the estimate of $p_\delta(t)$ is bad in the sense described above, is at most $\eta/2$ for

$$N \geq \frac{c}{\varepsilon^2\nu}\ln\left(\max(\frac{1}{\eta}, \frac{1}{\varepsilon^2\nu})\right) \tag{1}$$

for some constant c. Observe, that this is a combinatorial result which does not depend on the spatial position of witness points corresponding to a certain sequence of N experiments.

We now show that in case of translations and rigid motions for any two shapes and a parameter value δ we can find a value ν^* such that the maximum of the probability function p_δ is greater or equal ν^*. The statement holds also for similarity transformations but we will prove it using a different technique.

Clearly, for translations and rigid motions the support $\mathcal{T}'$ of the function $p_\delta(t)$ is bounded. Let $V_{\mathcal{T}'}$ denote its volume. Also, the δ-regions in the transformation space have the same measure V_δ for any sample. If every transformation would have equal probability to be covered by a δ-region, i.e., there would be a uniform distribution of the neighborhoods in the transformation space, then this probability would be $\nu^* = \frac{V_\delta}{V_{\mathcal{T}'}}$.

Consequently, for any arbitrary distribution there exists at least one transformation t such that $p_\delta(t) \geq \nu^*$. Since we are interested in transformations maximizing the probability it is sufficient to have a good approximation for transformation with coverage probability $\geq \nu^*$, and not to overestimate the probability for the other transformations.

In the case of translations, the support of $p_\delta(t)$ is $\mathcal{T}' = (A \oplus -B) \oplus C_\delta$, where C_δ denotes a disk of radius δ. Consider the connected components $A_1, \ldots, A_k$ of A and $B_1, \ldots, B_l$ of B, $k, l \leq n$, let $s_1, \ldots, s_k$ and $r_1, \ldots, r_l$ denote their lengths. Since $\mathcal{T}' = \bigcup_{i,j}(A_i \oplus (-B_j) \oplus C_\delta)$, its area is at most $\sum_{i,j}|A_i \oplus (-B_j) \oplus C_\delta|$. For a pair of components A_i, B_j the corresponding area in translation space is maximized if A_i and B_j both are a straight line segment and these segments are orthogonal. Then the area of $\mathcal{T}'$ is bounded by

$$\begin{aligned} V_{\mathcal{T}'} &\leq \sum_{i,j}(s_i + 2\delta)(r_j + 2\delta) = \sum_i (s_i + 2\delta)\sum_j (r_j + 2\delta) = (L_A + 2k\delta)(L_B + 2l\delta) \\ &\leq (L_A + 2n\delta)(L_B + 2n\delta) \leq 16m^2, \text{ where } m = \max(L_A, L_B, n\delta). \end{aligned}$$

The δ-regions in translation space have equal shape and size, which correspond to the δ-neighborhoods in the image space. The area of a δ-neighborhood with respect to the L_p-metric is $V_\delta = c_p\delta^2$, where c_p is a constant dependent on the choice of the metric, $2 \leq c_p \leq 4$ for $1 \leq p \leq \infty$. Then $\nu^* \geq \frac{\delta^2}{8m^2}$.

In case of rigid motions, for any fixed rotation angle α the cross-section of $\mathcal{T}'$ corresponds to the support for translations and, therefore, its area is bounded by $16m^2$. By Cavalieri's principle, the volume $V_{\mathcal{T}'}$ is then bounded for both approaches by $32\pi m^2$. For both approaches the cross-section of a δ-region corresponds to a δ-region for translations,

i.e., a δ-neighborhood. Therefore, its area is at least $2\delta_1^2$ in the first approach and $2\delta^2$ in the second approach for any underlying L_p-metric, $1 \leq p \leq \infty$. Consequently, $V_\delta \geq 4\delta_1^2\delta_2$ for the first approach and $V_\delta \geq 4\pi\delta^2$ for the second approach. Hence, we have $\nu^* \geq \frac{\delta_1^2\delta_2}{8\pi m^2}$ and $\nu^* \geq \frac{\delta^2}{8m^2}$, respectively.

For the similarities we use a different approach since the support of the probability distribution in transformation space cannot be bounded as above. Instead, we argue as follows: Suppose that the shape B is scaled by factor $\frac{\delta}{D_B}$, where D_B is the diameter of B, so that the diameter of the scaled shape B is δ. If A contains a connected component of length at least δ, then we can place the scaled shape B in such a way that for any point of a part of A of length δ the distance to any point of the scaled B is at most δ. Therefore, the measure of the set $M_\delta(t)$ for that t is at least $L_B^2 \cdot \delta^2$, details about the set $M_\delta(t)$ for similarity maps are given in section 4.3. The corresponding value of p_δ is $p_\delta(t) = \frac{|M_\delta(t)|}{|\Omega|} \geq \frac{L_B^2 \cdot \delta^2}{L_A^2 L_B^2} = \frac{\delta^2}{L_A^2}$. Otherwise, the largest connected component of A must have length at least $\frac{L_A}{n}$. For the transformation t that maps the scaled B to the largest component of A the measure of $M_\delta(t)$ is then at least $\frac{L_A^2}{n^2} L_B^2$ and $p_\delta(t) \geq \frac{1}{n^2}$. Therefore, $\nu^* \geq \frac{\delta^2}{m^2}$.

Let t^* be a witness point of the cell of the arrangement containing t_{opt}. Plugging the corresponding value of ν^* in the formula (1) for N we obtain that after

$$N = O\left(\frac{m^2}{\varepsilon^2\delta^2} \ln\left(\max(\frac{1}{\eta}, \frac{m^2}{\varepsilon^2\delta^2})\right)\right)$$

experiments for all witness points, in particular for t^* and t_{app}, and for t_{opt} it holds with probability at least $1-\eta/2$ that $|\tilde{p}_\delta(t) - p_\delta(t)| \leq \varepsilon p_\delta(t)$. Combining these error bounds we get

$$\begin{aligned}
\tilde{p}_\delta(t_{\text{app}}) &\geq \tilde{p}_\delta(t^*) && \text{since } t_{\text{app}} \text{ maximizes } \tilde{p}_\delta(t) \\
&= \tilde{p}_\delta(t_{\text{opt}}) && \text{for } t_{\text{opt}} \text{ is in the cell witnessed by } t^* \\
&\geq (1-\varepsilon) p_\delta(t_{\text{opt}}) && \text{with probability } 1-\eta/2 \text{ by Lemma 3.3}
\end{aligned}$$

and

$$\begin{aligned}
\tilde{p}_\delta(t_{\text{app}}) &\leq (1+\varepsilon) p_\delta(t_{\text{app}}) \\
&\leq (1+\varepsilon) p_\delta(t_{\text{opt}}) && \text{since } (t_{\text{opt}}) \text{ maximizes } p_\delta(t)
\end{aligned}$$

Therefore, $|\tilde{p}_\delta(t_{\text{app}}) - p_\delta(t_{\text{opt}})| \leq \varepsilon p_\delta(t_{\text{opt}})$ with probability at least $1-\eta$. □

3.2.2 Running time

The running time of the algorithm consists of the time needed to generate N random samples denoted by $T_{\text{gen}}(n, N)$, where n is the number of curves in the shape, and the time needed to determine the depth of the arrangement of N δ-regions denoted by $T_{\text{depth}}(N)$.

Given a natural parameterization of a curve, a random point on that curve can be generated in constant time. For generating a random point from a set of n curves we first select a curve randomly with probability proportional to the relative length of the curve and then take a random point from the selected curve. The selection of a random segment can be trivially done in time linear in n. If we first compute the relative curve lengths and record the corresponding probabilities to allow for binary search during the generation process we get preprocessing time linear in n and $O(\log n)$ generation time for a single point. Therefore, $T_{\text{gen}}(n, N) = O(n + N \log n)$.

In order to determine the depth of the arrangement of N δ-regions we can construct this arrangement and during the construction keep record of the depth of the cells. Then at the end of the construction algorithm we know the depth of the deepest cell. For general metrics L_p and the considered classes of transformations the boundaries of δ-regions are algebraic hypersurfaces. By Basu et al. [8], the arrangement of such surfaces can be constructed and traversed in $T_{\text{depth}}(N) = O(N^{d+1})$ time.

Summarizing these results and using Theorem 3.2 we obtain the following theorem:

Theorem 3.5. *Given two shapes A and B modeled by sets of n curves in the plane of total lengths L_A and L_B respectively, and parameters ε, η, $0 < \varepsilon, \eta < 1$. Let t_{opt} denote the transformation maximizing $p_\delta(t)$. In time $O(n + N \log n + N^{d+1})$, where d is the dimension of the transformation space, the generic probabilistic algorithm computes a transformation t_{app} such that $|\tilde{p}_\delta(t_{\text{app}}) - p_\delta(t_{\text{opt}})| \leq \varepsilon p_\delta(t_{\text{opt}})$ for $N = O\left(\frac{m^2}{\varepsilon^2\delta^2} \ln\left(\max(\frac{1}{\eta}, \frac{m^2}{\varepsilon^2\delta^2})\right)\right)$, where $m = \max(L_A, L_B, n\delta)$. (For $\delta = (\delta_1, \delta_2)$ in Approach 1 for rigid motions δ^2 has to be replaced by $\delta_2\delta_1^2$ and in the definition of m, δ by δ_1.)*

Observe that, at least for sufficiently small values of δ, the runtime of the algorithm depends much more on the parameters ε and η than on the combinatorial input size n, which is only needed in the preprocessing and the drawing of random samples.

The running time of the algorithm is actually better than that stated in Theorem 3.5 for translations and for similarities in combination with the L_1 and L_∞ metrics. In case of translations, the δ-regions are pseudo-disks and their arrangement can be constructed straightforwardly in time $O(N^2)$. For similarities in combination with the L_1 or L_∞ metric the δ-regions in transformation space are bounded by a constant number of 3-dimensional hyperplanes. Using the algorithm of Edelsbrunner et al. [10] the arrangement of N such δ-regions can be constructed in $O(N^4)$ time.

In order to achieve a speed up of the algorithm a combination with the depth approximation algorithm by Aronov and Har-Peled [7] is possible. Given a set of N object in $\mathbb{R}^d$ whose arrangement has depth $\mathcal{D}$ and a prespecified parameter $\varepsilon > 0$, their algorithm finds a point of depth at least $(1 - \varepsilon)\mathcal{D}$ in $O\left(N + T_{\text{DT}}(N, \varepsilon^{-2} \log n) \log n\right)$ expected time, where $T_{\text{DT}}(N, k)$ is the running time of a depth thresholding algorithm. A depth thresholding algorithm takes a set $\mathcal{S}$ of N objects and an integer $k > 0$ and returns the depth of the arrangement of $\mathcal{S}$ together with a witness point if this depth is at most k, or tells that the depth is greater than k.

It is easy to verify that in combination with depth approximation the probabilistic algorithm gives an approximation of $p_\delta(t_{\text{opt}})$ with an error at most 2ε. Thus, provided a fast

depth thresholding algorithm exists, this combination would result in a significant speed up of the probabilistic algorithm. For translations, that is for arrangement of pseudo-disks in the plane, a fast thresholding algorithm is known to exist [7], which results in running time $T_{\text{depth}}(N) = O(N\varepsilon^{-2}\log N)$.

4 Detailed analysis for different transformation classes

Common to all transformations is the definition of the δ-neighborhood of a point in image space. The δ-neighborhood of a point $a \in \mathbb{R}^2$ is defined as the set of points in $\mathbb{R}^2$ that have distance at most δ to a. Formally: $U_\delta(a) = \{p \in \mathbb{R}^2 | \text{dist}(p, a) \leq \delta\}$, where $\text{dist}(p, a)$ denotes the distance between two points according to a chosen metric. Typically we would use Euclidean distance (L_2 metric) or maximum norm distance (L_∞ metric) for easier computations.

4.1 Translations

The simplest class of transformations are translations. The translation space is two-dimensional. Two points $a, b \in \mathbb{R}^2$ define uniquely a translation t that maps b to a, $t = a - b$. Therefore, a random sample of each shape in step one of our algorithm consist of a single point and the sample space is $\Omega = A \times B$. The set $M_\delta(t)$ is then the set of the point pairs of the two shapes that are mapped to the δ-neighborhood of each other by the transformation t, formally, $M_\delta(t) = \{(a, b) \in A \times B | t(b) \in U_\delta(a)\}$. To maximize the measure of this set means to find a transformation that maps largest possible parts of the shapes into proximity of each other, which comes close to the intuitive notion of matching of two shapes.

Note, that there is a direct connection between the set $M_\delta(t)$ and a free space diagram, which was first defined for polygonal curves by Alt and Godau in [3]. Let $f : I \to \mathbb{R}^2, g : J \to \mathbb{R}^2$ be two curves, with parameter intervals $I, J \subset \mathbb{R}$. The set $F_\delta(f, g) := \{(s, r) \in I \times J \mid \text{dist}(f(s) - g(r)) \leq \delta\}$ denotes the *free space* of f and g, where $\text{dist}(\cdot, \cdot)$ denotes the distance measure in image space. The partition of $I \times J$ into regions belonging or not belonging to $F_\delta(f, g)$ is called the *free space diagram*. Obviously, we can parameterize a set of curves over the interval $[0, 1] \subset \mathbb{R}$ and, thus, view the shapes A and B as functions $A : [0, 1] \to \mathbb{R}^2$, $B : [0, 1] \to \mathbb{R}^2$. Then, the free space of the set A and the set B transformed by t is defined as $F_\delta(A, t(B)) := \{(s, r) \in [0, 1]^2 \mid \text{dist}(A(s) - t(B(r))) \leq \delta\}$, see Figure 3 for an example.

It is easy to see, that there is a one-to-one correspondence between the set $M_\delta(t)$ defined above and the set $F_\delta(A, t(B))$. Furthermore, the measure of the set $F_\delta(A, t(B))$ is exactly the measure of $M_\delta(t)$ normalized by the total measure of sampling space, that is, $p_\delta(t) = \frac{|M_\delta(t)|}{|A \times B|} = |F_\delta(A, t(B))|$.

Another observation that we make is, if a translation t corresponds to the sample pair $s = (a, b)$ and a transformation t' is covered by the δ-region corresponding to s, then the distance between vectors t and t' is at most δ under the same distance measure as for

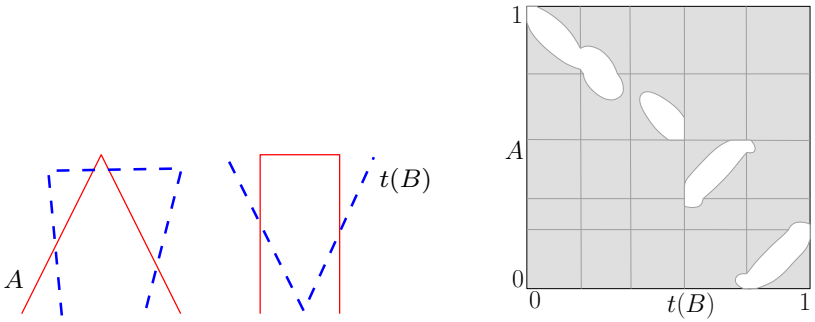


Figure 3: Free space diagram of two sets of curves. White regions denote the free space $F_\delta(A, t(B))$.

points. That means every sample pair produces a δ-region of the same size and shape in the transformation space. Thus, in case of translations we do not need to record a sample pair (a, b), but only the translation $t = a - b$ and then consider the δ-neighborhoods of the translation vectors, which are defined in the same way as the δ-neighborhoods of points.

For a general L_p-metric, $1 \leq p \leq \infty$, the arrangement of the δ-regions is an arrangement of pseudo-disks in the plane, which can be computed in $O(N^2)$ time, where N is the number of objects.

If for each sample pair (a, b) we record a translation that maps b to a instead of a δ-region, we get a certain distribution of "votes" in translation space. It turns out, that the density function of this probability distribution is exactly the weighted generalized Radon transform of the shape A with respect to shape B, [12, 22]. The distribution induced by δ-regions corresponds to a "smoothed" version of the generalized Radon transform. In fact, this distribution results from a convolution with a function that is constant over a δ-neighborhood of the origin and zero elsewhere and integrates to one.

4.2 Rigid motions

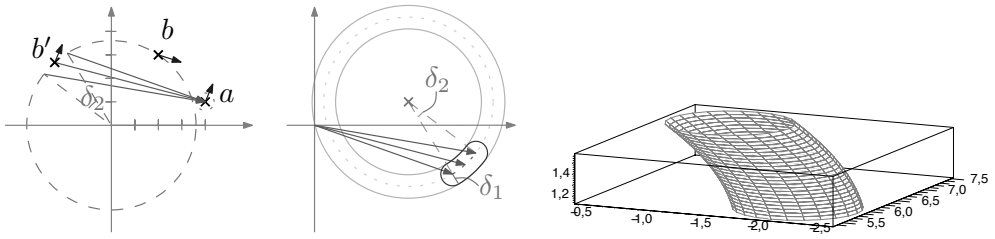
The space of rigid motions $\mathcal{T}$ is three dimensional, $\mathcal{T} = [0, 2\pi) \times \mathbb{R}^2$. For two points in the plane there is no unique rigid motion that maps one point to the other, rather for every rotation angle we can find a unique translation vector such that the resulting rigid motion performs the desired mapping. For two points the set of rigid motions that map one of the points to the other is, therefore, a one dimensional curve in three dimensional transformation space.

In Approach 1 for the rigid motions as described in section 2 a sample of a shape taken within one random experiment in the first step of the generic algorithm consists of a random point of the shape and the angle defined by the (interpolated) direction of the tangent line at that point. The sample space Ω is a subset of $A \times B \times [0, 2\pi]^2$.

We defined a δ-region in the transformation space corresponding to a sample pair as the set of transformations that map each component of one sample into a δ-neighborhood

of the corresponding component of the other sample. So far we interpreted δ as a distance parameter for points but it does not make much sense for the directions. Therefore, in this case we need two parameters $\delta = (\delta_1, \delta_2)$, where δ_1 controls the distance that points are allowed to have to be considered close, and δ_2 represents the maximum difference in the directions that are still considered to be similar.

A pair of samples $S_A = (a, \theta_a), S_B = (b, \theta_b)$ gives us a following neighborhood in the space of rigid motions: the rotation angles are restricted to the interval $I = [\alpha - \delta_2/2, \alpha + \delta_2/2]$, where $\alpha = \theta_a - \theta_b$, and for each $\alpha' \in I$ the allowed translations are v' such that $\text{dist}(\alpha'(b) + v', a) \leq \delta_1$, see Figure 4 for an illustration.



(a) (b) (c)

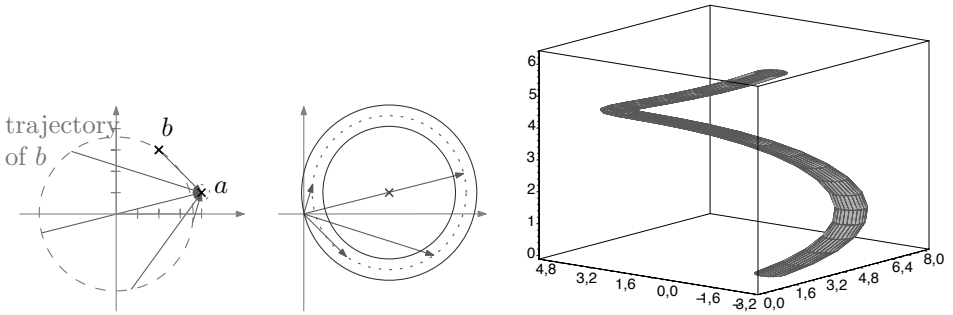
Figure 4: δ-region in the space of rigid motions corresponding to a pair of points $a \in A$ and $b \in B$. (a) Points a and b with translation vectors corresponding to some rotated positions of b. (b) Projection of the δ-region to the translation plane for $\delta_1 = 1$. (c) δ-region in the 3-dimensional space of rigid motions.

Then by taking the transformation that is covered by the most neighborhoods we maximize the measure of the set of point pairs, that are close to each other (have distance at most δ_1) and have similar tangent directions (the directions differ by at most δ_2). Note, that this definition of a good match is different from that we used for the translations, since for translations we did not take the direction of curves into account.

Although, there is still a connection to the free space diagram of the shapes. Consider again shapes to be parameterized over the interval $[0, 1]$, $A : [0, 1] \to \mathbb{R}^2, B : [0, 1] \to \mathbb{R}^2$. Let function $g : [0, 1]^2 \to \mathbb{R}$ denote the distance between the corresponding points on the shapes with respect to chosen metric, $g(s, r) = \text{dist}(A(s), B(r))$. The free space $F_{\delta_1}(A, B)$ in the free space diagram is defined as the set of pairs (s, r) for which $g(s, r) \leq \delta_1$. The additional condition about the closeness of tangent directions can be expressed by a function $h : [0, 1]^2 \to \mathbb{R}$ which assigns to two parameter values the absolute difference of the tangent directions of the corresponding points. Let $\theta_A(s), \theta_B(r)$ denote the angles of slope of the tangent lines to the points $A(s)$ and $B(r)$ respectively. Then the function h can be written as $h(s, r) = |\theta_A(s) - \theta_B(r)|$. Similar to the free space F_{δ_1} with respect to distance function g we can define free space H_{δ_2} with respect to direction distance function h, $H_{\delta_2}(A, B) = \{(s, r) \in [0, 1]^2 \mid h(s, r) \leq \delta_2\}$. The set of the "good" samples for a fixed rigid motion t then corresponds to the intersection of the two free spaces: $M_{\delta_1,\delta_2}(t) = \{(A(s), \theta_A(s), B(r), \theta_B(r)) \mid (s, r) \in F_{\delta_1}(A, t(B)) \cap H_{\delta_2}(A, t(B))\}$. The proba-

bility for a rigid motion t to be covered by a δ-region corresponding to a random sample pair is then $p_\delta(t) = \frac{|M_\delta(t)|}{|\Omega|} = |F_{\delta_1}(A, t(B)) \cap H_{\delta_2}(A, t(B))|$.

In Approach 2 we use a single random point of each shape $a \in A$ and $b \in B$ as a sample in one random experiment and record a δ-region in the space of rigid motions as a set of transformations that map the point b into the δ-neighborhood of point a. The sample space is in this case $\Omega = A \times B$. A δ-region corresponding to a sample $(a, b) \in \Omega$ has a shape of a spiral tube, which projection to the translation vector plane has a shape of a circular ring with center a and radius $\left\|\bar{b}\right\|$ of width δ as shown in Figure 5.



(a) (b) (c)

Figure 5: δ-region in the space of rigid motions corresponding to a pair of points $a \in A$ and $b \in B$. (a) Points a and b with translation vectors corresponding to some rotated positions of b. (b) Projection of the δ-region to the translation plane for $\delta = 1$. (c) δ-region in the 3-dimensional space of rigid motions.

Now by taking a rigid motion covered by the most neighborhoods we find a transformation that maximizes the measure of the set of point pairs that are mapped into δ-neighborhood of each other by this transformation, just as in case of translations. That is, the sample space Ω is $A \times B$, the set of sample pairs that "vote" for a rigid motion t is $M_\delta(t) = \{(a, b) \in A \times B \mid t(b) \in U_\delta(a)\}$ and the probability of t to be covered by a region corresponding to a randomly selected sample pair is $p_\delta(t) = \frac{|M_\delta(t)|}{|A \times B|} = |F_\delta(A, t(B))|$.

4.3 Similarity maps

In case of similarity maps a sample taken from one shape consists of two random points, since two pairs of points in the plane determine uniquely a similarity transformation that maps one pair to the other. The sample space is then $\Omega = A^2 \times B^2$. According to our analysis the similarity map with maximum coverage of the δ-regions is the one maximizing the measure of the set

$$M_\delta(t) = \left\{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 | \text{dist}(t(b_1), a_1) \leq \delta \text{ and } \text{dist}(t(b_2), a_2) \leq \delta\right\} \ ,$$

which is the set of pairs of point pairs that are at most δ apart. This property is no longer intuitive with respect to matching shapes. The following simple consideration shows, however, that maximizing the Lebesgue measure of $M_\delta(t)$ means also to

maximize the Lebesgue measure of the set of point pairs that have distance at most δ, $M'_\delta(t) = \{(a,b) \in A \times B | \mathrm{dist}(t(b), a) \leq \delta\}$. Up to the order of the elements $M_\delta(t)$ is the same as $M'^2_\delta(t)$. Additionally, in our random experiment we exclude samples where $a_1 = a_2$ or $b_1 = b_2$ because in these cases there is either no similarity transformation that maps one sample to the other or it is not uniquely defined. Those pairs are, of course, not excluded in the set $M'^2_\delta(t)$, but they make up a subset of dimension six in an eight dimensional space (recall, we have four points each of dimension two) and have, therefore, Lebesgue measure zero. So, the measure of the set $M_\delta(t)$ is exactly $|M'^2_\delta(t)| = |M'_\delta(t)|^2$. Since measure of a set is always non-negative, both functions have maxima at the same values of t.

Note, that if $\tilde{p}_\delta(t)$ is a good approximation of $p_\delta(t)$, more precisely, an approximation with a relative error at most ε, then $\tilde{m}_\delta(t) = \sqrt{\tilde{p}_\delta(t)\,|\Omega|}$ is an approximation of $m_\delta = |M'_\delta(t)|$ with a relative error at most ε: We have that $\tilde{p}_\delta(t) \geq (1-\varepsilon)p_\delta(t)$ and $\tilde{p}_\delta(t) \leq (1+\varepsilon)p_\delta(t)$, then

$$\begin{aligned}
\tilde{m}_\delta(t) &= \sqrt{\tilde{p}_\delta(t)\,|\Omega|} \\
&\geq \sqrt{(1-\varepsilon)p_\delta(t)\,|\Omega|} \\
&= \sqrt{(1-\varepsilon)\,|M_\delta(t)|} \\
&= \sqrt{(1-\varepsilon)}m_\delta(t) \\
&\geq (1-\varepsilon)m_\delta(t) && \text{since } 0 \leq 1-\varepsilon \leq 1
\end{aligned}$$

and

$$\begin{aligned}
\tilde{m}_\delta(t) &= \sqrt{\tilde{p}_\delta(t)\,|\Omega|} \\
&\leq \sqrt{(1+\varepsilon)p_\delta(t)\,|\Omega|} \\
&= \sqrt{(1+\varepsilon)}m_\delta(t) \\
&\leq (1+\varepsilon)m_\delta(t) && \text{since } 1+\varepsilon \geq 1
\end{aligned}$$

Therefore, for similarities as well as for translations and rigid motions the probabilistic algorithm finds a transformation that approximately maximizes the measure of the set of point pairs that are in a δ-neighborhood of each other.

Although a standard way to parameterize the space of similarity transformations is by rotation angle α, scaling factor k and a translation vector $v = (v_x, v_y)$ in order to avoid trigonometric functions in the definition of δ-regions it is more convenient to use the parameterization (m_1, m_2, v_x, v_y) where $m_1 = k\cos\alpha$ and $m_2 = k\sin\alpha$. For general L_p metric a δ-region is then bounded by algebraic surfaces, and for the L_1 and L_∞ metrics it is a convex polytope bounded by four pairs of parallel hyperplanes. An arrangement of N convex polytopes in d dimensions can be computed and traversed in time $O(N^d)$, see [10], which is $O(N^4)$ for similarity maps.

4.4 Affine maps

According to our analysis, an affine transformation which is covered by the most δ-regions maximizes the measure of the set

$$M_\delta(t) = \left\{(a_1, a_2, a_3, b_1, b_2, b_3) \in A^3 \times B^3 | \text{dist}(t(b_i), a_i) \leq \delta, i \in \{1, 2, 3\}\right\}.$$

With a similar observation as for similarity maps, we find that the measure of this set is maximized exactly if the measure of the set of point pairs which are in the δ-neighborhood of each other is maximized.

5 Conclusions

In this paper we presented a probabilistic approach for matching two shapes which attempts to come close to the human notion of match and is easy to implement. The algorithm is robust to noise, deformations and cracks in the representation of shapes and does not require shapes to be modelled by a single contour line. It is applicable to the problem of complete and partial matching. We achieved encouraging matching results in experiments with the MPEG7-Shape B data set and a selection of trademark images.

References

[1] A. S. Aguado, E. Montiel, and M. S. Nixon. Invariant characterisation of the Hough transform for pose estimation of arbitrary shapes. *Pattern Recognition*, 35:1083–1097, 2002.

[2] H. Alt, B. Behrends, and J. Blömer. Approximate matching of polygonal shapes. *Annals of Mathematics and Artificial Intelligence*, 13:251–265, 1995.

[3] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. *Internat. J. Comput. Geom. Appl.*, 5:75–91, 1995.

[4] H. Alt and L. J. Guibas. Discrete geometric shapes: Matching, interpolation, and approximation. In J.-R. Sack and J. Urrutia, editors, *Handbook of computational geometry*. Elsevier Science Publishers B.V. North-Holland, Amsterdam, 1999.

[5] H. Alt, G. Rote, C. Wenk, and A. Efrat. Matching planar maps. *J. of Algorithms*, pages 262–283, 2003.

[6] E. M. Arkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, and J. S. B. Mitchell. An efficiently computable metric for comparing polygonal shapes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(3):209–216, March 1991.

[7] B. Aronov and S. Har-Peled. On approximating the depth and related problems. In *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 886–894, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[8] S. Basu, R. Pollack, and M.-F. Roy. Computing roadmaps of semi-algebraic sets. In *STOC '96: Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 168–173, New York, NY, USA, 1996. ACM Press.

[9] O. Cheong, A. Efrat, and S. Har-Peled. On finding a guard that sees most and a shop that sells most. In *Proc. 15th ACM-SIAM Sympos. Discrete Algorithms (SODA)*, pages 1091–1100, 2004.

[10] H. Edelsbrunner, J. O'Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM J. Comput.*, 15(2):341–363, 1986.

[11] A. Efrat, Q. Fan, and S. Venkatasubramanian. Curve matching, time warping, and light fields. *Journal of Mathematical Imaging and Vision*, to appear.

[12] M. v. Ginkel, C. Luengo Hendriks, P. Verbeek, and L. van Vliet. The generalized radon transform: Sampling, accuracy and memory considerations. *Pattern Recognition*, 38(12):2494–2505, December 2005.

[13] M. Hagedoorn, M. Overmars, and R. Veltkamp. A new visibility partition for affine pattern matching. In *Proc. Discrete Geometry for Computer Imagery conference, DGCI 2000*, pages 358–370, Berlin, 2000. Springer-Verlag.

[14] M. Hagedoorn and R. Veltkamp. State-of-the-art in shape matching. Technical Report UU-CS-1999-27, Utrecht University, the Netherlands, 1999.

[15] D. Huttenlocher and S. Ullman. Recognizing solid objects by alignment with an image. *International Journal of Computer Vision*, 5(2):195–212, 1990.

[16] L. J. Latecki and R. C. Veltkamp. Properties and performances of shape similarity measures. In *Proceedings of the IFCS06 Conference Data Science and Classification*, 2006.

[17] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[18] S. Moss and E. R. Hancock. Pose clustering with density estimation and structural constraints. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 2085–2091, 1999.

[19] C. F. Olson. Efficient pose clustering using a randomized algorithm. *Int. J. Comput. Vision*, 23(2):131–147, 1997.

[20] G. Stockman. Object recognition and localization via pose clustering. *Computer Vision, Graphics, and Image Processing*, 40:361–387, 1987.

[21] M. Tanase, R. C. Veltkamp, and H. Haverkort. Multiple polyline to polygon matching. In *Proceedings of the 16th Annual Symposium on Algorithms and Computation (ISAAC 2005), LNCS 3827*, pages 60–70, 2005.

[22] P. Toft. *The Radon Transform - Theory and Implementation*. PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, 1996.

[23] R. C. Veltkamp. Shape matching: Similarity measure and algorithms. In *Shape Modelling International*, pages 188–197, Genova, Italy, May 2001. IEEE Press.

[24] H. J. Wolfson and I. Rigoutsos. Geometric hashing: An overview. *IEEE Computational Science and Engineering*, 04(4):10–21, 1997.