Abstract

This paper deals with questions from convex geometry related to shape matching. In particular, we consider the problem of matching convex figures minimizing the area of the symmetric difference. The main theorem of this paper states, that if we just match the two centers of gravity the resulting symmetric difference is within a factor of $11/3$ from the optimal one. This leads to efficient approximate matching algorithms for convex figures.
1 Introduction

A very common problem arising in application areas like computer vision or pattern recognition is that two "figures" $F_1$ and $F_2$ are given and the question is, how much these figures "look alike". In other words, we want to match $F_1$ and $F_2$ as good as possible and somehow measure the quality of this match. More precisely, assume that $F_1$ and $F_2$ are subsets of the plane, whose boundaries are simple closed curves, that there is a certain set $\mathcal{T}$ of feasible transformations that may be used for matching, and that there is a certain distance function $\delta$ measuring the quality of the match. Reasonable sets of matching transformations could be for example translations, rigid motions i.e. compositions of translations and rotations, similarities, or arbitrary affine mappings. The problem of matching $F_2$ to $F_1$ optimally then means to find a transformation $t_{opt}$ minimizing $\delta(F_1, t(F_2))$ for all $t \in \mathcal{T}$.

Most of the previous work has concentrated on the Hausdorff distance as a distance measure [ABB91, AST94, CGH+93, HKS93, AAR94]. Since solving the optimization problem exactly turned out to be rather difficult, more efficient approximation algorithms have been developed. These algorithms do not necessarily find the optimum but a solution whose quality is within a constant factor of the optimal one. In particular, approximation algorithms have been found using so-called reference points. This means that for each figure a characteristic point can be computed with the property that if two figures are matched optimally, their reference points lie close together. Conversely, if we restrict to matching transformations that map the reference point of $F_2$ onto the one of $F_1$ and optimize with respect to this restricted set, the solution cannot be much worse than the optimal unrestricted one. On the other hand the restricted set has fewer degrees of freedom and the optimization problem is easier to solve. For example, in the case of translations the restricted optimal translation is directly available, namely the vector between the two reference points. In the case of rigid motions, the two reference points are matched and then the optimal position of $F_2$ is sought among rotations around this point.

Formally we shall call a map that assigns to each figure a point in the plane a reference point for $\mathcal{T}$ (with respect to a distance $\delta$) if there is a constant $c \geq 1$ such that for any two figures $F_1$, $F_2$ there exists $r \in \mathcal{T}$ that maps the reference point of $F_2$ onto the reference point of $F_1$ such that for all $t \in \mathcal{T}$

$$\delta(F_1, r(F_2)) \leq c \cdot \delta(F_1, t(F_2)).$$

Two reference points with respect to Hausdorff distance and rigid motions (and also more general classes of transformations) have been found: the centroid of the boundary of the convex hull [ABB91] and the so-called Steiner point [AAR94]. Considering very long rectangles on the one hand and one of the triangles obtained from the rectangle by cutting along the diagonal on the other hand it is easily seen, that the center of gravity or centroid is not a reference point with respect to the Hausdorff distance.

Here we will consider a different distance measure between figures, namely the area of the symmetric difference which in the sequel we will denote by $\delta(F_1, F_2)$. $\delta$ has been investigated only in a few papers so far including [ABGW90], where simplification problems are addressed, and a recent paper by de Berg et al. [dBDvK+96] which is also concerned with matching problems under translations.

We believe that $\delta$ is a quite natural distance measure and in some applications is more appropriate than the Hausdorff distance. On the other hand it seems more difficult to handle. So in this paper we will restrict to convex figures but for them we can show that
the centroid is a reference point for translations, rigid motions, and some other sets of transformations. In particular, we will show for convex figures $F_1, F_2$ that if we translate $F_2$ so that its centroid matches the one of $F_1$ the resulting symmetric difference is at most $11/3$ times as large as the optimal one under translations. On the other hand we will give an example where this constant is assumed, so it is tight. We believe, that this result is already interesting by itself from a convex geometry point of view.

Finally, we will give approximate matching algorithms based on this reference point for various sets of transformations.

2 A reference point for translations

**Lemma 1** Consider a convex body $C$ in the plane and a line $e$ passing the centroid $s$ of $C$. Denote by $\ell_0$ and $\ell_1$ the (lengths of the) two line segments of $e$ inside $C$ separated by $s$. Then $\ell_0 \leq 2\ell_1$, in other words, each of these segments has length at most $2/3$ times the length of the part of $e$ inside $C$.

**Proof:** Denote by $P$ and $Q$ the intersections of $e$ with the boundary of $C$. Let $g$ be the tangent to $C$ through $P$ and $h$ the line through $s$ parallel to $g$. Consider the smallest triangle $T$ containing $Q$ and $C \cap h$, that has one side on $g$. As can be seen in Figure 1 the centroid of $T$ must lie between $h$ and $g$. Denote by $\ell_0'$ and $\ell_1'$ the segments of $e$ given by the projection of $T$’s centroid parallel to $h$ (or $g$). Then $\ell_0 \leq \ell_0' = 2\ell_1' \leq 2\ell_1$. □

**Convention:**
For simplicity, we use the same notation for subsets of $\mathbb{R}^2$ and their area. It should be clear from the context, which one is meant.
Lemma 2 Let $F \subset \mathbb{R}^2$ be convex, $f \subset F$ be measurable and $s_F$, $s_f$ their centroids. Let $w$ be the length of the projection of $F$ onto a line perpendicular to the vector $s_F - s_f$. Then

$$w \cdot |s_F - s_f| \leq \frac{4}{3}(F - f).$$

Proof. We assume w.l.o.g. that $s_F$ and $s_f$ have the same $x$-coordinate so that $w$ is the width of $F$ in $x$-direction. We also assume that the $y$-coordinate of $s_f$ is the larger one. We will transform the sets $f$ and $F$ in four steps into more special sets (see Figure 2). For simplicity the sets will be called $f$ and $F$ throughout the process although their shapes may change. Their area and the width of $F$ in $x$-direction will not change. The centroids will move, but in each step the distance of their $y$-coordinates will not decrease. After the final step their $x$-coordinate will again coincide and we will be able to prove the inequality directly.

Step 1
is a shearing that preserves $x$-coordinates in general and transforms $F$ and $f$ so that the (one) leftmost and (one) rightmost point of $F$, $L$ and $R$, respectively, have the same $y$-coordinate.

Step 2
We apply Steiner Symmetrization (see [BW48], §9) to make $F$ a convex body symmetric to some vertical axis $s$. This operation can be imagined as cutting $F$ into infinitesimally
flat horizontal slices and arranging these slices symmetrically to \( s \). The \( y \)-coordinates of \( s_f \) and \( s_F \) do not change in this step.

**Step 3**

Let \( g \) be a horizontal line and \( H^+ \) and \( H^- \) the upper and lower halfplane, respectively, defined by \( g \). Let \( \overline{f} = F \setminus f \) and choose \( g \) such that the areas \( H^+ \cap \overline{f} = H^- \cap f \). Change \( f \) to \( f \setminus (H^- \cap f) \cup (H^+ \cap \overline{f}) \). The new \( f \) lies completely above \( g \), in fact, \( f = F \cap H^+ \).

Now assume that the line segment \( LR \) does not intersect \( f \), if not, we can just exchange the roles of \( f \) and \( \overline{f} \), since the formula \( s_F = \frac{f}{2} s_f + \frac{\overline{f}}{2} s_{\overline{f}} \) implies that

\[
\frac{s_F - s_f}{F - f} = -\frac{(s_F - s_{\overline{f}})}{F - \overline{f}}.
\]

So we have to show the upper bound of the lemma only either for \( f \) or for \( \overline{f} \).

**Step 4**

The region above \( LR \) is transformed into a triangle. To this end we consider the line segment \( e \) from \( L \) to some point \( A \) on the symmetry axis \( s \) with the property that the portion \( f_1 \) of \( f \) left of \( e \) has the same area as the region \( f_2 \) between \( e \) and \( s \) that lies outside of \( F \) (see Figure 3).

![Figure 3: Step 4](image)

Likewise we have symmetric objects \( e' \), \( f'_1 \), and \( f'_2 \) on the right hand side of the figure. In addition, let \( F_1 \) be the portion of \( F \) left of \( e \) and \( F'_1 \) the set symmetric to \( F_1 \) on the right. Now we create the triangle by “moving” \( f_1 \), \( f'_1 \) into \( f_2 \), \( f'_2 \), respectively. \( F_1 \), \( F'_1 \) are moved into a rectangle \( F_2 \) just below \( LR \) and the part of \( F \) below \( LR \) is attached below the rectangle. Moving \( f_1 \cup f'_1 \) into \( f_2 \cup f'_2 \) increments the \( y \)-coordinate of \( s_f \) by some amount \( \varepsilon \geq 0 \) and the \( y \)-coordinate of \( s_F \) by \( f \cdot \varepsilon \). Moving \( F_1 \cup F'_1 \) into \( F_2 \) (at most) decrements the \( y \)-coordinate of \( s_F \), so the difference \( |s_f - s_F| \) is at most incremented.
For the figure remaining after step 4 we show the claim of the lemma directly. Let \( r, t, d, b, B \) as in Figure 4. Then

![Diagram](image_url)

**Figure 4**: Figures after step 4

\[
F \geq \frac{1}{2} wd
\]

(1)

since the right hand side is the area of the quadrilateral ARBL. By Lemma 1 we have \( r + t \leq \frac{2}{3} d \), so by (1)

\[
r + t \leq \frac{4}{3} \cdot \frac{F}{w}
\]

(2)

since the area of the shaded triangle is \( f = \frac{13}{27} tb \leq \frac{2}{3} tw \) we have

\[
t \geq \frac{4 f}{3 w}
\]

(3)

by (3) and (2) \( wr \leq \frac{4}{3}(F - f). \square \)

Before we get to the main result we note that by straightforward set theoretic considerations for arbitrary sets \( A, B, C \)

\[
A \Delta B \subset (A \Delta C) \cup (B \Delta C)
\]

and hence for measurable sets using the notational convention:

\[
A \Delta B \leq (A \Delta C) + (B \Delta C).
\]

(4)

Consider convex bodies \( F_1 \) and \( F_2 \) in the plane. \( \delta^{opt}(F_1, F_2) \) denotes the minimal area of the symmetric difference between translates of \( F_1 \) and \( F_2 \). \( \delta^H(F_1, F_2) \) denotes the area of the symmetric difference between translates of \( F_1 \) and \( F_2 \) whose centroids coincide.
**Theorem 3** Let $F_1, F_2$ be convex bodies. Then

$$
\delta^H(F_1, F_2) \leq \frac{11}{3} \delta^{opt}(F_1, F_2)
$$

*Proof.* We first assume that $F_2 \subset F_1$, so $\delta^{opt}(F_1, F_2) = F_1 - F_2 = F_1 \Delta F_2$. Let $s_1, s_2$ be the centroid of $F_1, F_2$ respectively. Now suppose that $F_2$ is translated by the vector $s_1 - s_2$ from $F_2$ the optimal position into the heuristic position $F_2'$ where $s_1, s_2$ are matched (see Figure 5). The symmetric difference $F_2' \Delta F_2$ is bounded by the area $c$ that is "swept" by the boundary of $F_2$ during the translation. By Cavalieri's principle $c$ is bounded by twice the length of the translation vector $s_1 - s_2$ times the width of the projection of $F_2$ onto a line normal to this vector. So we have

\[
\delta^H(F_1, F_2) = F_1 \left< F_2' \right> \\
\leq (F_1 \left< F_2 \right>) + (F_2 \left< F_2' \right>) \\
\leq (F_1 \left< F_2 \right>) + 2 \cdot |s_1 - s_2| \cdot w \\
\leq (F_1 \left< F_2 \right>) + 2 \cdot \frac{4}{3} \cdot (F_1 - F_2) \\
= \frac{11}{3} \delta^{opt}(F_1, F_2)
\]

where $w$ denotes the width of the projection of $F_1$ onto a line normal to $s_1 - s_2$.

Next let us consider the general case. Assume that $F_1, F_2$ are in optimal position and let $I = F_1 \cap F_2$. Applying (4) to translates of $F_1, F_2$ and $I$ with coinciding centroids we
get
\[
\delta^H(F_1, F_2) \leq \delta^H(F_1, I) + \delta^H(F_2, I) \\
\leq \frac{11}{3} \delta^{\text{opt}}(F_1, I) + \frac{11}{3} \delta^{\text{opt}}(F_2, I) \quad \text{by the first case} \\
= \frac{11}{3} \delta^{\text{opt}}(F_1, F_2)
\]
which proves the theorem. □

**Theorem 4**

\[
\sup \left\{ \frac{\delta^H(F_1, F_2)}{\delta^{\text{opt}}(F_1, F_2)} : F_1 \text{ and } F_2 \text{ are convex bodies, } \delta^{\text{opt}}(F_1, F_2) > 0 \right\} = \frac{11}{3}
\]

*Proof.* In Theorem 3 we have shown that 11/3 is an upper bound of the set on the left side of the equation. Thus we only have to construct examples which show that 11/3 is approached arbitrarily close by this set. Denote by $F_0$ an isosceles triangle whose base and the corresponding height have unit length. For $\varepsilon > 0$ denote by $F_\varepsilon$ the trapezoid obtained from $F_0$ by cutting off a tip of height $\varepsilon$. (Fig. 6). Clearly,

\[
\delta^{\text{opt}}(F_0, F_\varepsilon) = \frac{\varepsilon^2}{2}, \quad \delta^H(F_0, F_\varepsilon) = \varepsilon^2/2.
\]  \hspace{8cm} (5)

The cut moves the centroid by $\frac{2 \varepsilon^2}{3 + \varepsilon}$ towards the base. The area of the symmetric difference of translates of $F_0$ and $F_\varepsilon$ with the same centroid is shown in Fig. 6. Simple computation yields

\[
\delta^H(F_0, F_\varepsilon) = \varepsilon^2 \frac{33 + 18 \varepsilon - 11 \varepsilon^2}{18(1 + \varepsilon)^2}.
\]  \hspace{8cm} (6)

It follows from (5) and (6):

\[
\lim_{\varepsilon \to 0} \frac{\delta^H(F_0, F_\varepsilon)}{\delta^{\text{opt}}(F_0, F_\varepsilon)} = \frac{11}{3}. \quad \square
\]
3 Transformations other than translations

In many applications more general matching transformations than just translations are considered. These include, for example,

- rigid motions, i.e. combinations of translations and rotations,
- homotheties, i.e. mappings of the form $\mathbf{x} \mapsto \mathbf{a} + \lambda(\mathbf{x} - \mathbf{a})$, for some fixed $\lambda \in \mathbb{R}$ and some fixed center $\mathbf{a} \in \mathbb{R}^2$
- similarities, i.e. combinations of homotheties and rigid motions or
- arbitrary affine mappings.

For an arbitrary set $\mathcal{T}$ of transformations, we define the matching problem as follows:

*Given two figures $F_1$, $F_2$, find a transformation $t^{\text{opt}} \in \mathcal{T}$ minimizing $\delta(F_1, t(F_2))$ for $t \in \mathcal{T}$.***

Our heuristic approach for finding approximate solutions to this matching problem is to consider only those transformations in $\mathcal{T}$ that map the centroid of $F_2$ onto the one of $F_1$. Let $t^H$ be an optimal transformation of this kind. Hence the centroid is a reference point for $\mathcal{T}$ if there is a constant $c \geq 1$ such that for any two figures $F_1$, $F_2$

$$\delta^H(F_1, F_2) \leq c\delta^{\text{opt}}(F_1, F_2),$$

where we define $\delta^H(F_1, F_2) = \delta(F_1, t^H(F_2))$ and $\delta^{\text{opt}}(F_1, F_2) = \delta(F_1, t^{\text{opt}}(F_2))$.

Now the following can be shown:

**Theorem 5** If a set of transformations $\mathcal{T}$ has the following properties:

(i) Equivariance with respect to the centroid, i.e. $t(cg(F)) = cg(t(F))$ for all figures $F$ and all $t \in \mathcal{T}$.

(ii) $\mathcal{T}$ is closed under compositions with translations

Then the centroid is a reference point for $\mathcal{T}$ and the corresponding constant $c$ is $11/3$.

**Proof.** Let $F_1$, $F_2$ be two figures and $F'_2 := t^{\text{opt}}(F_2)$. Translate $F'_2$ so that the resulting figure $F''_2$ has the same centroid as $F'_1$. Then by Theorem 3

$$\delta(F_1, F''_2) \leq c\delta(F_1, F'_2) = c\delta^{\text{opt}}(F_1, F_2)$$

with $c = 11/3$. $F''_2 = t'(F_2)$ for a transformation $t'$ which is a composition of $t^{\text{opt}}$ and a translation. By condition (ii), $t' \in \mathcal{T}$, and by condition (i), $cg(F_1) = cg(F''_2) = t'(cg(F_2))$.

Since $t^H$ is the optimal match of $F_1$ and $F_2$ under transformations satisfying these conditions, we have

$$\delta^H(F_1, F_2) \leq \delta(F_1, t'(F_2)) = \delta(F_1, F''_2) \leq c\delta^{\text{opt}}(F_1, F_2). \quad \square$$

Observe, that many natural sets of transformations satisfy conditions (i) and (ii) of Theorem 5, for example rigid motions, homotheties, similarities, affine transformations, and rigid motions where only a restricted set of rotations is allowed. So for all these sets of transformations we obtain a simplified and (hopefully) easier to solve matching problem, whose optimal solution is an approximate one for the original problem.

We will consider algorithms based on this idea in the next section.
4 Algorithms

The results of the previous sections can be used to design efficient approximately optimal matching algorithms for convex polygons under various sets of transformations. These algorithms will produce a solution which is at most by a factor of $11/3$ worse than the optimal one. Throughout this section we assume that we are given two convex polygons $F_1$ and $F_2$ (by a sorted list of their vertices) which are to be matched. Let $n$ be the total number of vertices of $F_1$ and $F_2$.

4.1 Translations

In the case of translations we just have to compute the centroids $s_1$ and $s_2$ and then to translate $F_2$ by the vector $s_1 - s_2$. The computation of the centroids can be done in linear time by triangulating each figure, determining the centroids and areas of all triangles, and then determining the total centroid as the weighted sum of the triangle centroids.

This gives a matching algorithm of runtime $O(n)$, which is as far as asymptotic runtime is concerned not too big an improvement over the algorithm by de Berg et al. [BDvK+96] which computes the optimal match under translations in $O(n \log n)$ time. But our algorithm may be a viable alternative in practice since it is much less complicated.

Usually, after the two figures have been matched, one also wants to compute the resulting area of the symmetric difference. This can be done by a straightforward sweepline-algorithm. The sweepline is a ray emanating from the centroid which visits the vertices of both polygons in (say) clockwise order. Observe that from the two sorted lists of vertices we can compute the order needed by the sweepline in linear time. The sweepline algorithm also determines all intersection points of both polygons. Between two event points (vertices or intersection points) the symmetric difference is a triangle or trapezoid and, therefore, can be determined in constant time. Consequently, measuring the total area takes time $O(n)$.

4.2 Homotheties

According to Theorem 5 we get an $11/3$-approximate solution by first computing the two centroids $s_1$ and $s_2$, then translating $F_2$ by $s_1 - s_2$ obtaining $F'_2$, and finally stretching $F'_2$ about $s_1$ by a factor minimizing the symmetric difference.

It remains to explain the last step. Suppose w.l.o.g. that $s_1$ is the origin. Consider the ‘wedges’ that are obtained by drawing a ray from the origin through each vertex of $F_1$ and $F'_2$. Within a wedge $W$ the boundary of each of the two figures consists of a line segment. Now, suppose that $F'_2$ is stretched by some factor $\lambda \geq 0$ which is 0 in the beginning and then is continuously increased. Thus we obtain, consecutively, the configurations a), b), and c) of the edges $e_1$ of $F_1$ and $e_2$ of $\lambda F'_2$ shown in Figure 7. In each case the symmetric difference within $W$ is a quadratic function. The symmetric difference within the $i$-th wedge, which we denote by $q_i(\lambda)$, is thus a piecewise quadratic function with three quadratic pieces. The total symmetric difference $q(\lambda)$ is the sum of these $n$ functions $q_i(\lambda)$. It is piecewise quadratic with $2n$ breakpoints.

The minimizer $\lambda^*$ of this function can be determined in $O(n)$ time by the prune-and-search technique: We search for the quadratic piece in which $\lambda^*$ lies by performing a binary search among the $2n$ breakpoints, successively narrowing down the interval $[\lambda_0, \lambda_1]$ in which $\lambda^*$ is known to lie. The decision whether $\lambda^*$ is bigger or smaller than the current
4.3 Rigid Motions

As in the previous case we will first perform a translation such that w.l.o.g. the centroids of $F_1$ and $F_2$ coincide. Now we have to rotate $F_2$ around the common centroid to minimize the symmetric difference.

We consider the same wedges as before. By distinguishing several cases we reduce the problem of computing a formula that computes the symmetric difference as a function of the rotation angle $\varphi$ to the simple case shown in Figure 8. $s$ is the center of the rotation, the distances $a, c$ and the angle $\alpha$ are fixed constants. We get

$$F(\varphi) = \frac{1}{2} \left( c - \frac{a}{\sin(\varphi)} \right)^2 \frac{\sin(\varphi) \sin(\alpha)}{\sin(\alpha + \varphi)}$$

The symmetric difference $\Delta(\varphi)$ of $F_1$ and $F_2$ is the boolean combination of areas of the form $F(\gamma \pm \varphi)$ for some $\gamma$, hence the area can be computed as a linear combination of $O(n)$ terms of the above form. The process of minimization mimics the rotation, while we determine all local minima. During this process there occur $O(n^2)$ event points,
whenever one of the spokes of $F_1$ sweeps over one of the spokes of $F_2$. As above we have to update the formula at each event point. Between any two consecutive event points we have to determine all local minima of $\Delta(\phi)$, i.e. the zeroes of the derivative $\Delta'(\phi)$. This problem can in principle be approached by exact computation, since it can be formulated in such a way that only computations with algebraic numbers are required, see below. However, this approach is very expensive and moreover, it does not make sense in the present context, where the computed value is only a rough approximation. Therefore the minimum must be computed numerically which involves highly nontrivial numerical problems, whose detailed investigation goes beyond the scope of this paper. Just to get an idea of the number of computational steps necessary let us use $t = \tan(\phi/2)$ as a parameter instead of $\phi$. Then $\sin(\phi) = 2t/(1+t^2)$ and $\cos(\phi) = (1-t^2)/(1+t^2)$, so between any two event points $\Delta$ can be written as a rational function in $t$ whose numerator and denominator have degree $O(n)$. Consequently, between two event points $\Delta$ has $O(n)$ local minima. Suppose that we determine these zeroes of $\Delta'$ by Newton iteration. Each step involves evaluating $\Delta$ and $\Delta'$ which costs $O(n)$ arithmetic operations. This gives $O(n^4)$ arithmetic operations but we have to multiply this number with the number of iteration steps performed by one application of the Newton iteration. This number depends on the coefficients of the rational function and the desired precision of the result and is together with the problem of finding suitable starting points for the iteration part of the numerical problems mentioned before. For an intensive treatment of this problem with respect to bit complexity see [Sch82].

Acknowledgements

We would like to thank Stefan Felsner and Emo Welzl for helpful discussions and hints concerning this research.

References


