Between Two Shapes, Using the Hausdorff Distance*

Marc van Kreveld[1], Tillman Miltzow[1], Tim Ophelders[2], Willem Sonke[3], and Jordi Vermeulen[1]

1 **Dep. of Information and Computing Sciences, Utrecht University**
`{m.j.vankreveld|t.miltzow|j.l.vermeulen}@uu.nl`

2 **Dep. of Computational Mathematics, Science and Engineering, Michigan State University**
`ophelder@egr.msu.edu`

3 **Dep. of Mathematics and Computer Science, TU Eindhoven**
`w.m.sonke@tue.nl`

Abstract

Given two shapes A and B in the plane with Hausdorff distance 1, is there a shape S with Hausdorff distance 1/2 to and from A and B? The answer is always yes, and depending on convexity of A and/or B, S may be convex, connected, or disconnected. We show a generalization of this result and a few others about Hausdorff distances, middle shapes, and related properties. We also show that the implied morph has a bounded rate of change.

1 Introduction

The Hausdorff distance is a widely used distance metric with many applications. For two sets A and B in $\mathbb{R}^2$, we define the *directed Hausdorff distance* as

$$d_{\vec{H}}(A, B) = \sup_{a \in A} \inf_{b \in B} d(a, b),$$

where d denotes the Euclidean distance. The *undirected Hausdorff distance* is defined as

$$d_H(A, B) = \max(d_{\vec{H}}(A, B), d_{\vec{H}}(B, A)).$$

If A and B are closed sets then $d_H(A, B) = r$ is equivalent to saying that $A \subseteq B \oplus D_r$ and $B \subseteq A \oplus D_r$, where $\oplus$ denotes the Minkowski sum, and D_r is a disk of radius r centered at the origin. Recall that the Minkowski sum of sets A and B is the set $\{a + b \mid a \in A,\ b \in B\}$. As we will use this alternative definition throughout the paper, we will only consider closed sets.

Algorithms to compute the Hausdorff distance are available for many types of input sets, such as points, line segments, polylines, polygons and simplices in k-dimensional Euclidean space [1, 2, 4]. However, the question whether a polynomial time algorithm exists to compute the Hausdorff distance between general semialgebraic sets remains open [5].

In this paper, we consider the problem of finding a third set that lies "between" the two input sets, in a Hausdorff sense. We define a class of sets that smoothly interpolate between the two input sets, giving a morph between them. Unlike many existing morphing algorithms [3, 6], our approach does not require any correspondence between features of the

* Research on the topic of this paper was initiated at the 4th Workshop on Applied Geometric Algorithms (AGA 2018) in Langbroek, The Netherlands, supported by the Netherlands Organisation for Scientific Research (NWO) under project no. 639.023.208. The second author is supported by the NWO Veni grant EAGER. The first and fifth authors are supported by the NWO TOP grant no. 612.001.651

36th European Workshop on Computational Geometry, Würzburg, Germany, March 16–18, 2020.

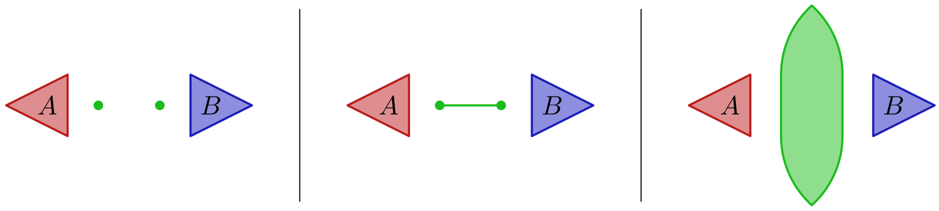

Figure 1 Three possible Hausdorff middles of A and B: two points, a line segment, and $S_{1/2}$.

input to be calculated. However, our approach is unusual in the sense that the intermediate shapes when morphing between e.g. two polygons are not polygons themselves.

Our main contribution is to pioneer the notion of Hausdorff middle and the interpolation between two shapes. We address and solve elementary combinatorial, topological, and algorithmic questions.

2 The Hausdorff middle

Consider two closed compact sets A and B in $\mathbb{R}^2$; we are interested in computing a *Hausdorff middle*: a set C that minimizes the maximum of the undirected Hausdorff distances to A and B. That is,

$$C = \operatorname*{argmin}_{C'} \max(d_H(A, C'), d_H(B, C')).$$

Note that there may be many such sets that minimize the Hausdorff distance; see Figure 1 for a few examples. It might seem intuitive to restrict C to be the minimal set that achieves this distance, but this is ill-defined: the minimal set is not necessarily unique, and the common intersection of all minimal sets is not a solution itself (see Figure 2). However, the maximal set is well-defined and unique. Let $d_H(A, B) = 1$. Then

$$S = (A \oplus D_{1/2}) \cap (B \oplus D_{1/2})$$

is the maximal set satisfying the constraints. We want to show that $d_H(A, S)$ and $d_H(B, S)$ are at most $1/2$. In fact, we can prove something more general where we define maximal sets that may be at other places between A and B, not only halfway.

▶ **Lemma 1.** *Let A and B be two closed sets in the plane with $d_H(A, B) = 1$, and let $S_\alpha = (A \oplus D_\alpha) \cap (B \oplus D_{1-\alpha})$ for $\alpha \in [0, 1]$. Then $d_H(A, S_\alpha) = \alpha$ and $d_H(B, S_\alpha) = 1 - \alpha$.*

Proof. We first show that $d_H(A, S_\alpha) \leq \alpha$ by showing that $d_{\vec{H}}(A, S_\alpha) \leq \alpha$ and $d_{\vec{H}}(S_\alpha, A) \leq \alpha$; the case for $d_H(B, S_\alpha) \leq 1-\alpha$ is analogous and therefore omitted. Then we show equality.

Consider any point $a \in A$; by definition, there is a point $b \in B$ with $d(a, b) \leq 1$. Now consider a point $s \in \operatorname{seg}(a, b)$ with $d(a, s) \leq \alpha$ and $d(b, s) \leq 1 - \alpha$; clearly this point must

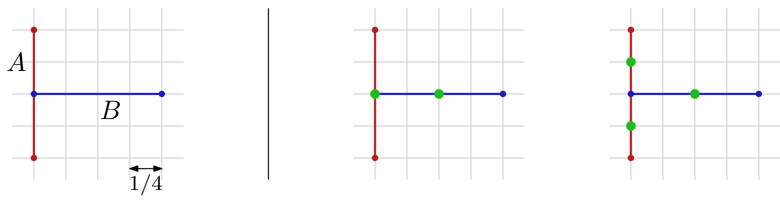

Figure 2 Two different minimal sets achieving minimal Hausdorff distance to A and B.

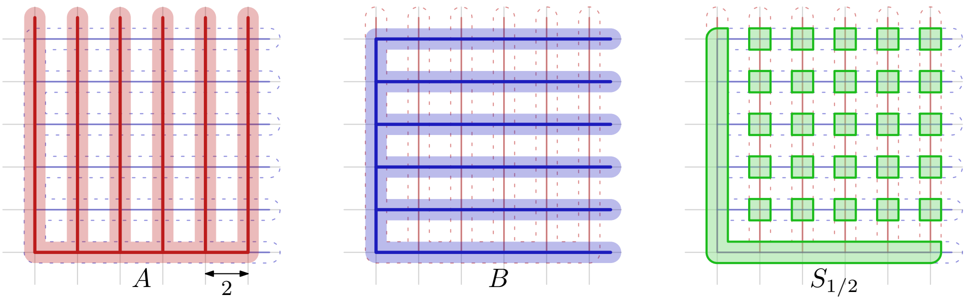

Figure 3 Sets A and B for which $S_{1/2}$ is disconnected. The shaded areas around A and B represent $A \oplus D_{1/2}$ and $B \oplus D_{1/2}$, respectively.

be in S_α, as it is contained in both $A \oplus D_\alpha$ and $B \oplus D_{1-\alpha}$, and it has $d(a, s) \leq \alpha$. As this works for every $a \in A$, it holds that $d_{\vec{H}}(A, S_\alpha) \leq \alpha$. The fact that $d_{\vec{H}}(S_\alpha, A) \leq \alpha$ follows straightforwardly from S_α being a subset of $A \oplus D_\alpha$. Thus, $d_H(A, S_\alpha) \leq \alpha$.

To show equality, assume that the Hausdorff distance is realized by a point $\hat{a} \in A$ with closest point $\hat{b} \in B$, at distance 1. Consider the point $\hat{s} \in \text{seg}(\hat{a}, \hat{b})$ with $d(\hat{a}, \hat{s}) = \alpha$ and $d(\hat{b}, \hat{s}) = 1 - \alpha$. As observed, $\hat{s} \in S_\alpha$. Since $\hat{s}$ is the closest point of S_α to $\hat{a}$, and $\hat{b}$ is the closest point of B to $\hat{s}$, equality follows. ◀

▶ **Lemma 2.** *S_α is the maximal set that satisfies $d_H(A, S_\alpha) = \alpha$ and $d_H(B, S_\alpha) = 1 - \alpha$.*

Proof. Consider any set T for which we have $d_{\vec{H}}(T, A) \leq \alpha$ and $d_{\vec{H}}(T, B) \leq 1-\alpha$. As $A \oplus D_\alpha$ contains all points with distance at most α to A, we have that $T \subseteq A \oplus D_\alpha$; similarly, we have that $T \subseteq B \oplus D_{1-\alpha}$. By the definition of S_α, this implies that $T \subseteq S_\alpha$. As this holds for any T, we conclude that S_α is maximal. ◀

2.1 Properties of S_α

In this section, we study the convexity and connectedness of S_α. Recall that a set $A \subseteq \mathbb{R}^2$ is convex if for any two points $a, b \in A$, the segment $\bar{ab}$ between them is completely contained in A. Also, recall that a set $A \subseteq \mathbb{R}^2$ is connected if for any two points $a, b \in A$, there exists a path from a to b completely contained in A.

1. if A and B are convex, S_α is convex;
2. if A is convex and B is connected, S_α is connected;
3. if A and B are connected, but neither A nor B is convex, S_α may be disconnected.

Property 1 is straightforward: the Minkowski sum of A and B with a disk is convex, and the intersection of convex objects is itself also convex. The example in Figure 3 demonstrates Property 3. We show Property 2 next.

We say a set $A \subset \mathbb{R}^2$ is connected if for any two points $a, b \in A$, there exists a continuous curve $c : [0, 1] \to A$ such that $c(0) = a$ and $c(1) = b$. This type of connectedness is known as path-connectedness. Note that the empty set is trivially connected. The following observation is straightforward:

▶ **Observation 3.** *Let A and B be two connected sets in the plane. If $A \cap B$ is not connected, $A \cup B$ contains a hole.*

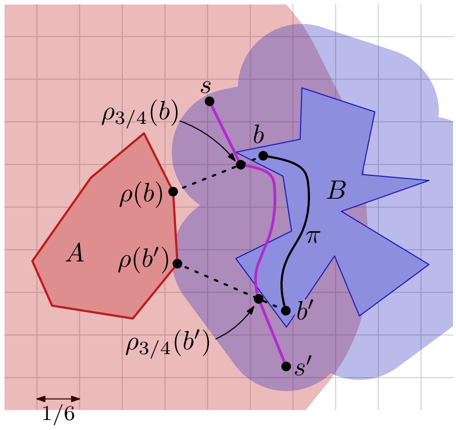

Figure 4 Example of the construction from Lemma 4 showing that S_α is connected if A is convex (sketched here for $\alpha = 3/4$). The shaded areas around A and B represent $A \oplus D_{3/4}$ and $B \oplus D_{1/4}$, respectively, so that the doubly-shaded area is $S_{3/4}$.

The next lemma establishes property 2:

▶ **Lemma 4.** *Let A and B be two connected polygons with Hausdorff distance 1, and A convex. Then $S_\alpha = (A \oplus D_\alpha) \cap (B \oplus D_{1-\alpha})$ is connected.*

Proof. The argument is as follows (see Figure 4 for a sketch):

- Because A is convex, there is a continuous map $\rho\colon B \to A$ that maps each point of B to a closest point (within distance 1) in A.
- For $b \in B$, let $\rho_\alpha(b) = b + \alpha(\rho(b) - b)$. We have $\rho_\alpha\colon B \to S_\alpha$ which is also continuous.
- Now take any two points s and s' in S_α; respectively, they have points b and $b' \in B$ within distance $1 - \alpha$.
- The segments between s and $\rho_\alpha(b)$ and between s' and $\rho_\alpha(b')$ lie completely in S_α.
- Take a path π from b to b' inside B. The image of π under ρ_α connects $\rho_\alpha(b)$ to $\rho_\alpha(b')$ within S_α, so s and s' are connected inside S_α. ◀

We note that S_α may contain holes. Furthermore, S_α is not shape invariant when B is translated with respect to A. For example, let A be the union of the left and bottom sides of a unit square and let B_1 and B_2 be the left and right sides of that same unit square. Then $(A \oplus D_{1/2}) \cap (B_1 \oplus D_{1/2})$ is not a translate of $(A \oplus D_{1/2}) \cap (B_2 \oplus D_{1/2})$. See Figure 5.

2.2 Complexity of S_α

In this section, we describe the complexity of S_α in terms of the number of vertices, line segments, and circular arcs on its boundary, for several types of polygonal input sets. Recall that ∂A denotes the boundary of set A.

▶ **Lemma 5.** *Let A be a convex polygon and B a simple polygon with n, respectively m vertices. Then S_α consists of $\Theta(n+m)$ vertices, line segments and circular arcs in the worst case.*

Proof. There is a trivial worst-case lower bound of $\Omega(n + m)$ by taking $\alpha = 0$ or $\alpha = 1$. Note that if the boundaries of $A^\oplus = A \oplus D_\alpha$ and $B^\oplus = B \oplus D_{1-\alpha}$ would consist of only line segments, the upper bound is easy to show: $A^\oplus$ is convex, and its boundary can therefore intersect each segment of $\partial B^\oplus$ at most twice, making ∂S_α consist of (parts of) segments

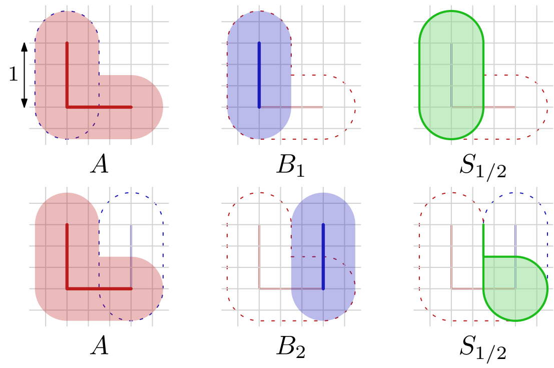

Figure 5 Although B_2 is a translate of B_1, the middle set between A_1 and B_2 is not a translate of the middle set between A_1 and B_1.

from $\partial A^{\oplus}$ and $\partial B^{\oplus}$ and at most $O(m)$ intersection points. The problem is that $\partial A^{\oplus}$ and $\partial B^{\oplus}$ also contain circular arcs, in which case an arc of $\partial B^{\oplus}$ may intersect $\partial A^{\oplus}$ many times.

To show an upper bound of $O(n+m)$, we distinguish two cases. In the first case, we assume $\alpha \geq 1-\alpha$. Note that in this case, the circular arcs that are part of the boundary of $A^{\oplus} = A \oplus D_\alpha$ have a radius larger or equal to those of $B^{\oplus} = B \oplus D_{1-\alpha}$. In this case, we do in fact have that any line segment or circular arc b of $\partial B^{\oplus}$ can intersect $\partial A^{\oplus}$ at most twice: for any circular arc of $\partial A^{\oplus}$, the entire disk that it bounds must be contained in $A^{\oplus}$. This means that b either intersects the same feature of $\partial A^{\oplus}$ twice, or it intersects two different features once.

For the second case, we assume $\alpha < 1-\alpha$. Again, take an arbitrary arc b of $\partial B^{\oplus}$ that intersects some arc a of $\partial A^{\oplus}$. We distinguish two cases: the center point of the disk whose boundary contains a is inside $B^{\oplus}$, or it is outside. If it is outside, b can only intersect $\partial A^{\oplus}$ in two points. If it is inside, $\partial A^{\oplus}$ may intersect b many times. We charge these intersections to the arcs of $\partial A^{\oplus}$. We argue that each arc a of $\partial A^{\oplus}$ is charged at most four times: Consider any α-disk D_α and any $(1-\alpha)$-disk $D_{1-\alpha}$ containing the center of D_α, the latter will cover at least $1/3$ of the perimeter of the former. Hence, the boundary of the union of any number of such $(1-\alpha)$-disks intersects D_α at most four times. The circular arcs of $\partial A^{\oplus}$ cannot be charged more often because they are less than a full circle. ◀

▶ **Lemma 6.** *Let A and B be two simply connected polygons of n and m vertices, respectively. Then S_α consists of $\Theta(nm)$ vertices, line segments and circular arcs in the worst case.*

Proof. We can show a worst-case lower bound of $\Omega(nm)$ by taking A and B to be two "combs" placed at right angles to each other; see Figure 3. In this example, for $\alpha = 1/2$, S_α consists of $\Omega(nm)$ distinct components. The upper bound follows directly from the fact that $A^{\oplus} = A \oplus D_\alpha$ and $B^{\oplus} = B \oplus D_{1-\alpha}$ have complexities $O(n)$ and $O(m)$, respectively. ◀

2.3 S_α as a morph

By increasing α from 0 to 1, S_α morphs from $A = S_0$ into $B = S_1$. The following lemma shows that this morph has a bounded rate of change.

▶ **Lemma 7.** *Let S_α and S_β be two intermediate shapes of A and B with $\alpha \leq \beta$. Then $d_H(S_\alpha, S_\beta) = \beta - \alpha$.*

Proof. We have $d_H(S_\alpha, S_\beta) \geq \beta - \alpha$ because by the triangle inequality, $d_H(A, B) = 1 \leq d_H(A, S_\alpha) + d_H(S_\alpha, S_\beta) + d_H(S_\beta, B) \leq \alpha + d_H(S_\alpha, S_\beta) + 1 - \beta$.

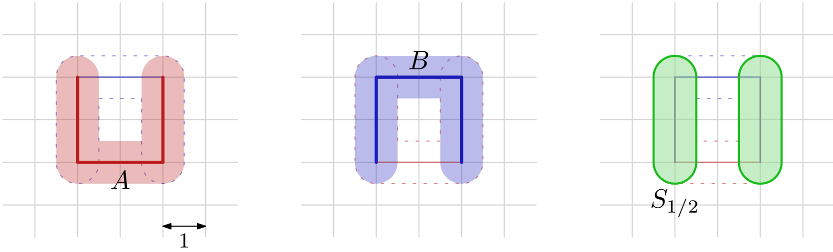

Figure 6 $S_{1/2}$ for the red and blue polygons is shown in green. Any connected Hausdorff middle must cross vertical middle line or stay on one side of it. In both cases, a Hausdorff distance doubles.

It remains to show that $d_H(S_\alpha, S_\beta) \leq \beta - \alpha$. We show that $S_\beta \subseteq S_\alpha \oplus D_{\beta-\alpha}$; the other case is analogous. Let p be some point in S_β. Then, by definition of S_β, there exist some points $a \in A$ and $b \in B$ such that $d(a, p) \leq \beta$ and $d(b, p) \leq 1-\beta$. Let $\bar{p}$ be the point obtained by moving p in the direction of a by $\beta - \alpha$. By the triangle inequality, we then have that $d(a, \bar{p}) \leq \beta - (\beta - \alpha) = \alpha$ and $d(b, \bar{p}) \leq (1 - \beta) + (\beta - \alpha) = 1 - \alpha$. This implies that $\bar{p} \in S_\alpha$. As p was an arbitrary point in S_β, and $d(p, \bar{p}) \leq \beta - \alpha$, we have that $S_\beta \subseteq S_\alpha \oplus D_{\beta-\alpha}$. So $d_H(S_\alpha, S_\beta) \leq \beta - \alpha$. ◀

The lemma implies that, even though the number of connected components of S_α can change when α changes, new components arise by splitting and never 'out of nothing', and the number of components can only decrease through merging and not by disappearance.

2.4 The cost of connectedness

For some applications, it might be necessary to insist that S_α is always connected. However, in the worst case, the cost of connecting all components of S_α can be that its Hausdorff distance to A and B becomes 1. See Figure 6 for an example where this is the case.

3 Conclusion

Besides the maximal middle set, there are other options for a Hausdorff middle. For example, we can choose S_α clipped to the convex hull of $A \cup B$, which is also a valid Hausdorff middle. In Figure 6, the green shape would be reduced to the part inside the square, which may be more natural. This Hausdorff middle can also be used in a morph.

A natural question is whether these results extend to more than two input shapes. The problem would then be to compute some shape S that minimizes the maximum pairwise Hausdorff distance to any of the input shapes, assuming that the pairwaise Hausdorff distance is at most 1. It turns out that in some cases, the minimum Hausdorff distance that

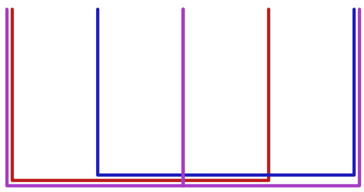

Figure 7 Assuming the pairwise Hausdorff distance is 1, any shape will have Hausdorff distance at least 1 to or from at least one of the three inputs.

can be achieved is 1, even when the input sets are connected; see Figure 7 for an example. When the input shapes are all convex, the worst example we have found so far is three sets of a single point each, forming an equilateral triangle. In this case, the best middle set is the centroid of the triangle, which has Hausdorff distance $1/\sqrt{3}$ to each input set. We conjecture that this is the highest possible optimal value for convex input sets.

References

1 H. Alt, B. Behrends, and J. Blömer. Approximate matching of polygonal shapes. *Annals of Mathematics and Artificial Intelligence*, 13(3):251–265, Sep 1995.

2 H. Alt, P. Braß, M. Godau, C. Knauer, and C. Wenk. Computing the Hausdorff distance of geometric patterns and shapes. In *Discrete and Computational Geometry*, pages 65–76. Springer, 2003.

3 H. Alt and L.J. Guibas. Discrete geometric shapes: Matching, interpolation, and approximation. In *Handbook of Computational Geometry*, pages 121–153. Elsevier, 2000.

4 M. J. Atallah. A linear time algorithm for the hausdorff distance between convex polygons. *Information Processing Letters*, 17(4):207 − 209, 1983.

5 M. G. Dobbins, L. Kleist, T. Miltzow, and P. Rzążewski. $\forall\exists\mathbb{R}$-completeness and area-universality. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pages 164–175. Springer, 2018.

6 C. Gotsman and V. Surazhsky. Guaranteed intersection-free polygon morphing. *Computers & Graphics*, 25(1):67–75, 2001. Shape Blending.