Shape Similarity Measures, Properties, and Constructions

Remco C. Veltkamp and Michiel Hagedoorn*
Department of Computer Science, Utrecht University
Padualaan 14, 3584 CH, Utrecht, The Netherlands
{Remco.Veltkamp,mh}@cs.uu.nl

Abstract

In this paper we list a number of similarity measures, some of which are not well known (such as the Monge-Kantorovich metric), or newly introduced (reflection metric). We formulate properties of similarity measures, and introduce new properties. We also give a set of constructions that have been used in the design of some similarity measures, including some new constructions.

1 Introduction

Large image databases are used in many multimedia applications in fields such as entertainment, business, art, engineering, and science. Retrieving images by their content, as opposed to external features, has become an important operation. A fundamental ingredient for content-based image retrieval is the technique used for comparing images. There are two general methods for image comparison: intensity-based (color and texture) and geometry-based (shape). A recent user survey about cognition aspects of object retrieval shows that users are more interested in retrieval by shape than by color and texture [30]. However, retrieval by shape is still considered one of the most difficult aspects of content-based search. Indeed, systems such as IBM's QBIC, Query By Image Content [25], perhaps one of the most advanced image retrieval systems to date, is relatively successful in retrieving by color and texture, but performs poorly when searching on shape. The Alta Vista photo finder [6] shows similar behavior.

There is no universal definition of what shape is. Impressions of shape can be conveyed by color or intensity patterns (texture), from which a geometrical representation can be derived. This is shown already in Plato's work Meno [24]. (This is one of the so-called Socratic dialogues, where two persons discuss aspects of virtue; to memorialize Socrates, one of the figures is called after him.) In this work, the word 'figure' is used for shape. Socrates' description

> "figure is the only existing thing that is found always following color".

does not satisfy Meno, after which Socrates gives a definition in "terms employed in geometrical problems":

> "figure is limit of solid".

*supported by Philips Research Laboratories

In this paper too we consider shape as something geometrical.

Shape similarity measures are an essential ingredient in shape matching. Matching deals with transforming a pattern, and measuring the resemblance with another pattern using some dissimilarity measure. The terms pattern matching and shape matching are commonly used interchangeably. The matching problem is studied in various forms. Given two patterns and a dissimilarity measure:

- (computation problem) compute the dissimilarity between the two patterns,
- (decision problem) for a given threshold, decide whether the dissimilarity between two patterns is smaller than the threshold,
- (decision problem) for a given threshold, decide whether there exists a transformation such that the dissimilarity between the transformed pattern and the other pattern is smaller than the threshold,
- (optimization problem) find the transformation that minimizes the dissimilarity between the transformed pattern and the other pattern.

Sometimes the time complexities to solve these problems are rather high, so that it makes sense to devise approximation algorithms that find an approximation:

- (approximate optimization problem) find a transformation that gives a dissimilarity between the two patterns that is within a specified factor from the minimum dissimilarity.

2 Properties

In this section we list a number of possible properties of similarity measures. Whether or not specific properties are desirable will depend on the particular application, sometimes a property will be useful, sometimes it will be undesirable. Some combinations of properties are contradictory, so that no distance function can be found satisfying them. A shape similarity measure, or distance function, on a collection of shapes S is a function $d : S \times S \to \mathbb{R}$. The following conditions apply to all the shapes A, B, or C in S.

1 (Nonnegativity) $d(A, B) \geq 0$.
2 (Identity) $d(A, A) = 0$ for all shapes A.
3 (Uniqueness) $d(A, B) = 0$ implies $A = B$.
4 (Strong triangle inequality) $d(A, B) + d(A, C) \geq d(B, C)$.

Nonnegativity (1) is implied by (2) and (4). A distance function satisfying (2), (3), and (4) is called a metric. If a function satisfies only (2) and (4), then it is called a semimetric. Symmetry (see below) follows from (4). A more common formulation of the triangle inequality is the following:

5 (Triangle inequality) $d(A, B) + d(B, C) \geq d(A, C)$.

Properties (2) and (5) do not imply symmetry. Similarity measures for partial matching, giving a small distance $d(A, B)$ if a part of A matches a part of B, in general do not obey the triangle inequality. A counterexample is given in figure 1: the distance from the man to the centaur is small, the distance from the centaur to the horse is small, but the distance from the man to the horse is large, so $d(man, centaur) + d(centaur, horse) > d(man, horse)$ does not hold. It therefore makes sense to formulate an even weaker form [14]:

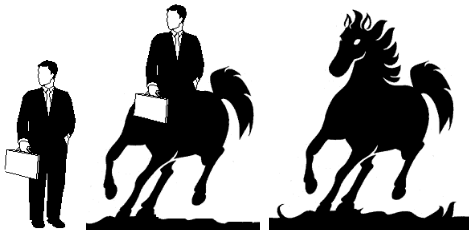

Figure 1: Under partial matching, the triangle inequality does not hold.

6 (Relaxed triangle inequality) $c(d(A,B)+d(B,C)) \geq d(A,C)$, for some constant $c \geq 1$.

7 (Symmetry) $d(A,B) = d(B,A)$.

Symmetry is not always wanted. Indeed, human perception does not always find that shape A is equally similar to B, as B is to A. In particular, a variant A of prototype B is often found more similar to B than vice versa [32].

8 (Invariance) d is invariant under a chosen group of transformations G if for all $g \in G$, $d(g(A), g(B)) = d(A,B)$.

For object recognition, it is often desirable that the similarity measure is invariant under affine transformations, illustrated in figure 2. This however depends on the application, and sometimes a large invariance group is not wanted. For example, Sir d'Arcy Thompson [31] showed that the outlines of two hatchetfishes of different genus, *Argyropelecus olfersi* and *Sternoptyx diaphana*, are shear invariant, so that they cannot be distinguished under shear transformations, see figure 3.

The following four properties are about robustness, a form of continuity. Such properties are useful to be robust against the effects of discretization, see figure 4.

9 (Perturbation robustness) For each $\epsilon > 0$, there is an open set F of deformations sufficiently close to the identity, such that $d(f(A), A) < \epsilon$ for all $f \in F$.

10 (Crack robustness) For each each $\epsilon > 0$, and each "crack" x in the boundary of A, an open neighborhood U of x exists such that for all B, $A - U = B - U$ implies $d(A,B) < \epsilon$.

11 (Blur robustness) For each $\epsilon > 0$, an open neighborhood U of $bd(A)$, the boundary of A exists, such that $d(A,B) < \epsilon$ for all B satisfying $B - U = A - U$ and $bd(A) \subseteq bd(B)$.

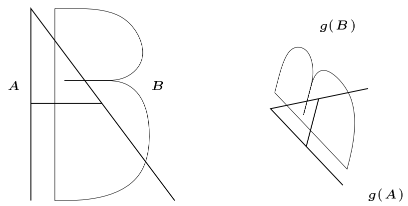

Figure 2: Affine invariance: $d(A,B) = d(g(A), g(B))$.

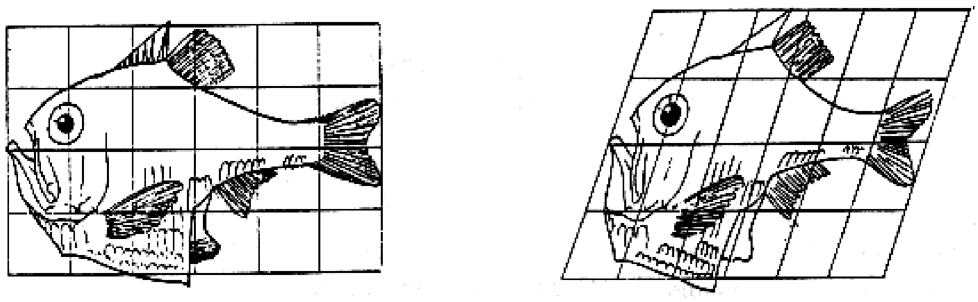

Figure 3: Two shear invariant hatchetfishes of different genus: *Argyropelecus olfersi* and *Sternoptyx diaphana*. After [31].

12 (Noise robustness) For each $x \in \mathbb{R}^2 - A$, and each $\epsilon > 0$, an open neighborhood U of x exists such that for all B, $B - U = A - U$ implies $d(A, B) < \epsilon$.

A distance function is distributive in the shape space if the distance between one pattern and another does not exceed the sum of distances between the one and two parts of the other:

13 (Distributivity) For all A and decomposable $B \cup C$, $d(A, B \cup C) \leq d(A, B) + d(A, C)$.

The following properties all describe variations of discriminative power. The first one says that there is always a shape more dissimilar to A than some shape B. This is not possible if the collection of shapes is finite.

14 (Endlessness) For each A, B there is a C such that $d(A, C) > d(A, B)$.

The next property means that for a chosen transformation set G, the distance d is able to discern A as an exact subset of $A \cup B$. No $g(A)$ is closer to $A \cup B$ than A itself:

15 (Discernment) For a chosen transformation set G, $d(A, A \cup B) \leq d(g(A), A \cup B)$ for all $g \in G$.

The following says that changing patterns, which are already different, in a region where they are still equal, should increase the distance.

16 (Sensitivity) For all A, B with $A \cap U = B \cap U$, $B - U = C - U$, and $B \cap U \neq C \cap U$ for some open $U \subset \mathbb{R}^2$, then $d(A, B) < d(A, C)$.

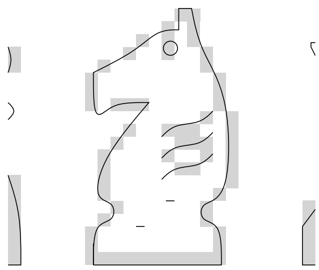

Figure 4: Discretization effects: deformation, blur, cracks, and noise.

The next property says that the change from A to $A \cup B$ is smaller that the change to $A \cup C$ if B is smaller than C:

17 (Proportionality) For all $A \cap B = \emptyset$ and $A \cap C = \emptyset$, if $B \subset C$, then $d(A, A \cup B) < d(A, A \cup C)$.

Finally, the distance function is strictly monotone if at least one of the intermediate steps of adding $B - A$ to A, and $C - B$ to B is smaller than the two steps combined:

18 (Monotonicity) For all $A \subset B \subset C$, $d(A, C) > d(A, B)$, or $d(A, C) > d(B, C)$.

3 Similarity Measures

3.1 Discrete Metric

Finding an affine invariant metric for patterns is not so difficult. Indeed, a metric that is invariant not only for affine transformations, but for general homeomorphisms is the discrete metric:

$$d(A, B) = \begin{cases} 0 & \text{if } A \text{ equals } B \\ 1 & \text{otherwise} \end{cases}$$

However, this metric lacks useful properties. For example, if a pattern A is only slightly distorted to form a pattern A', the discrete distance $d(A, A')$ is already maximal.

3.2 L_p Distances, Minkowski Distance

Many similarity measures on shapes are based on the L_p distance between two points. For two points x, y in $\mathbb{R}^k$, the L_p distance is defined as $L_p(x, y) = (\sum_{i=0}^{k} |x_i - y_i|^p)^{1/p}$. This is also often called the Minkowski distance. For $p = 2$, this yields the Euclidean distance L_2. For $p = 1$, we get the Manhattan, city block, or taxicab distance L_1. For p approaching ∞, we get the max metric: $\max_i(|x_i - y_i|)$.

For all $p \geq 1$, the L_p distances are metrics. For $p < 1$ it is not a metric anymore, since the triangle inequality does not hold.

3.3 Bottleneck Distance

Let A and B be two point sets of size n, and $d(a, b)$ a distance between two points. The bottleneck distance $F(A, B)$ is the minimum over all $1 - 1$ correspondences f between A and B of the maximum distance $d(a, f(a))$. For the distance $d(a, b)$ between two points, an L_p distance could be chosen. An alternative is to compute an approximation $\tilde{F}$ to the real bottleneck distance F. An approximate matching between A and B with $\tilde{F}$ the furthest matched pair, such that $F < \tilde{F} < (1 + \epsilon)F$, can be computed with a less complex algorithm [13].

So far we have considered only the computation problem, computing the distance between two point sets. The decision problem for translations, deciding whether there exists a translation ℓ such that $F(A + \ell, B) < \epsilon$ can also be solved, but takes considerably more time [13]. Because of the high degree in the computational complexity, it is interesting to look at approximations with a factor ϵ: $F(A+\ell,B) < (1+\epsilon)F(A+\ell^*,T)$ [29], where ℓ^* is the optimal translation.

Variations on the bottleneck distance are the minimum weight distance, the most uniform distance, and the minimum deviation distance.

3.4 Hausdorff Distance

In many applications, for example stereo matching, not all points from A need to have a corresponding point in B, due to occlusion and noise. Typically, the two point sets are of different size, so that no one-to-one correspondence exists between all points. In that case, a dissimilarity measure that is often used is the Hausdorff distance. The Hausdorff distance is defined not only for finite point sets, it is defined on nonempty closed bounded subsets of any metric space.

The *directed* Hausdorff distance $\vec{h}(A,B)$ is defined as the lowest upper bound (supremum) over all points in A of the distances to B: $\vec{h}(A,B) = \sup_{a\in A}\inf_{b\in B} d(a,b)$, with $d(a,b)$ the underlying distance, for example the Euclidean distance (L_2). The Hausdorff distance $H(A,B)$ is the maximum of $\vec{h}(A,B)$ and $\vec{h}(B,A)$: $H(A,B) = \max\{\vec{d}(A,B),\vec{d}(B,A)\}$. For finite point sets, it can be computed using Voronoi diagrams [2].

Given two finite point sets A and B, computing the translation ℓ^* that minimizes the Hausdorff distance $H(A+\ell,B)$ is discussed in [9] and [20]. Given a real value ϵ, deciding if there is a rigid motion m (translation plus rotation) such that $H(m(A),B) < \epsilon$ is discussed in [8]. Computing the optimal rigid motion, minimizing $H(m(A),B)$, is treated in [19], using dynamic Voronoi diagrams.

Given the high complexities of these problems, it makes sense to look at approximations. Computing an approximate optimal Hausdorff distance under translation and rigid motion is discussed in [1].

3.5 Partial Hausdorff Distance

The Hausdorff distance is very sensitive to noise: a single outlier can determine the distance value. For finite point sets, a similar measure that is not as sensitive is the partial Hausdorff distance. It is the maximum of the two directed partial Hausdorff distances: $H_k(A,B) = \max\{\vec{h}_k(A,B),\vec{h}_k(B,A)\}$, where the directed distances are defined as the k-th value in increasing order of the distance from a point in A to B: $\vec{h}_k(A,B) = k^{th}_{a\in A}\min_{b\in B} d(a,b)$. The partial Hausdorff distance is not a metric since it fails the triangle inequality. Deciding whether there is a translation plus scaling that brings the partial Hausdorff distance under a given threshold is done in [21] by means of a transformation space subdivision scheme. The running time depends on the depth of subdivision of transformation space.

The subdivision of transformation space is generalized to a general framework by [18]. Here the optimal transformation is approximated. The matching can be done with respect to other transformations as well, for example, similarity (translation, rotation, and scaling), or affine transformations (translation, rotation, scaling, and shear).

3.6 p-th Order Mean Hausdorff Distance

For pattern matching, the Hausdorff metric is often too sensitive to noise. For finite point sets, the partial Hausdorff distance is not that sensitive, but it is no metric. Alternatively, [7] observes that the Hausdorff distance of $A,B \subseteq X$, X having a finite number of elements,

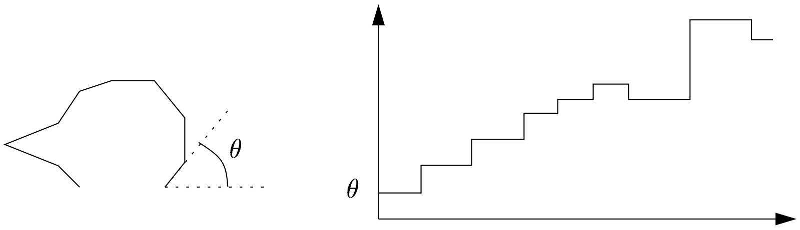

Figure 5: Polygonal curve and turning function.

can be written as $H(A,B) = \sup_{x \in X} |d(x,A) - d(x,B)|$. The supremum is then replaced by an average: $\Delta^p(A,B) = (\frac{1}{|X|}\sum_{x\in X} |d(x,A) - d(x,B)|^p)^{1/p}$, where $d(x,A) = \inf_{a \in A} d(x,a)$. This is a metric less sensitive to noise. This measure can for example be used for comparing binary images, where X is the set of all raster points.

3.7 Turning Function Distance

The cumulative angle function, or turning function, $\Theta_A(s)$ of a polygon A gives the angle between the counterclockwise tangent and the x-axis as a function of the arc length s. $\Theta_A(s)$ keeps track of the turning that takes place, increasing with left hand turns, and decreasing with right hand turns. Clearly, this function is invariant under translation of the polyline. Rotating a polyline over an angle θ results in a vertical shift of the function with an amount θ. For polygons and polylines, the turning function is a piecewise constant function, increasing or decreasing at the vertices, and constant between two consecutive vertices.

In [5] the turning angle function is used to match polygons. First the size of the polygons are scaled so that they have equal perimeter. The L_p metric on function spaces, applied to Θ_A and Θ_B, gives a dissimilarity measure on A and B: $d(A,B) = \left(\int |\Theta_A(s) - \Theta_B(s)|^p \, ds\right)^{1/p}$, see figure 6.

In [33], for the purpose of retrieving hieroglyphic shapes, polyline curves do not have the same length, so that partial matching can be performed. In that case the starting point of the shorter one is moved along the longer one, considering only the turning function where the arc lengths overlap. Partial matching under scaling, in addition to translation and rotation, is more involved [10].

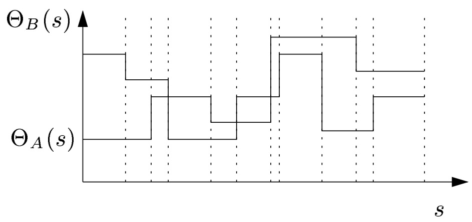

Figure 6: Rectangles enclosed by $\Theta_A(s)$, $\Theta_B(s)$, and dotted lines are used for evaluation of dissimilarity.

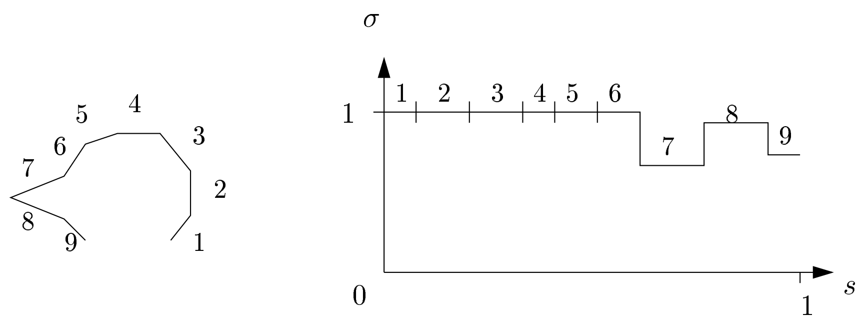

Figure 7: A curve and its signature function.

3.8 Signature Function Distance

A less discriminative function is the so-called signature function. At every point along the curve, the signature function σ value is the arc length of the curve to the left or on the tangent line at that point [23], see Figure 7. It is invariant under similarity: combinations of translation, rotation, and scaling. For convex curves, the signature function has value one everywhere, because at every point, the whole curve lies to the left of the tangent.

For polylines, dissimilarity measures can be used that are based on 'time warps' of sequences of elements (vertices or segments), pairing elements of A to elements of B, which can be computed by dynamic programming [22]. The pairing need not be one-to-one: the pairing of element i of A to element j of B, may be followed by a pairing of i to $j+1$, $i+1$ to j, or $i+1$ to $j+1$.

3.9 Fréchet Distance

The Hausdorff distance is often not appropriate to measure the dissimilarity between curves. For all points on A, the distance to the closest point on B may be small, but if we walk forward along curves A and B simultaneously, and measure the distance between corresponding points, the maximum of these distances may be larger, see Figure 8. This is what is called the Fréchet distance. More formally, let A and B be two parameterized curves $A(\alpha(t))$ and $B(\beta(t))$, and let their parameterizations α and β be continuous functions of the same parameter $t \in [0,1]$, such that $\alpha(0) = \beta(0) = 0$, and $\alpha(1) = \beta(1) = 1$. The Fréchet distance is the minimum over all monotone increasing parameterizations $\alpha(t)$ and $\beta(t)$ of the maximal distance $d(A(\alpha(t)), B(\beta(t)))$, $t \in [0,1]$, see figure 8.

The special case of computing the Fréchet distance for polylines is considered in [4]. A variation of the Fréchet distance is obtained by dropping the monotonicity condition of the

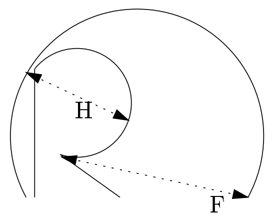

Figure 8: Hausdorff (H) and Fréchet (F) distance between two curves.

parameterization. The resulting Fréchet distance $d(A, B)$ is a semimetric: zero distance need not mean that the objects are the same. Another variation is to consider partial matching: finding the part of one curve to which the other has the smallest Fréchet distance.

Parameterized contours are curves where the starting point and ending point are the same. However, the starting and ending point could as well lie somewhere else on the contour, without changing the shape of the contour curve. For convex contours, the Fréchet distance is equal to the Hausdorff distance [4].

3.10 Nonlinear elastic matching distance

Let $A = \{a_1, \dots, a_m\}$ and $B = \{b_1, \dots, b_n\}$ be two finite sets of ordered contour points, and let f be a correspondence between all points in A and all points in B such that there are no $a_1 < a_2$, with $f(a_1) > f(a_2)$. The stretch $s(a_i, b_j)$ of $(a_i, f(a_i) = b_j)$ is 1 if either $f(a_{i-1}) = b_j$ or $f(a_i) = b_{j-1}$, or 0 otherwise. The nonlinear elastic matching distance $NEM(A, B)$ is the minimum over all correspondences f of $\sum s(a_i, b_j) + d(a_i, b_j)$, with $d(a_i, b_j)$ the difference between the tangent angles at a_i and b_j. It can be computed using dynamic programming [11]. This measure is not a metric, since it does not obey the triangle inequality.

3.11 Relaxed Nonlinear elastic matching distance

The relaxed nonlinear elastic matching distance NEM_r is a variation of NEM, where the stretch $s(a_i, b_j)$ of $(a_i, f(a_i) = b_j)$ is r (rather than 1) if either $f(a_{i-1}) = b_j$ or $f(a_i) = b_{j-1}$, or 0 otherwise, where $r \geq 1$ is a chosen constant. The resulting distance is not a metric, but it does obey the relaxed triangle inequality, property (6) above [14].

3.12 Reflection Distance

The *reflection metric* [17] is an affine-invariant metric that is defined on finite unions of curves in the plane. They are converted into real-valued functions on the plane. Then, these functions are compared using integration, resulting in a similarity measure for the corresponding patterns.

The functions are formed as follows, for each finite union of curves A. For each $x \in \mathbb{R}^2$, the *visibility star* V_A^x is defined as the union of open line segments connecting points of A that

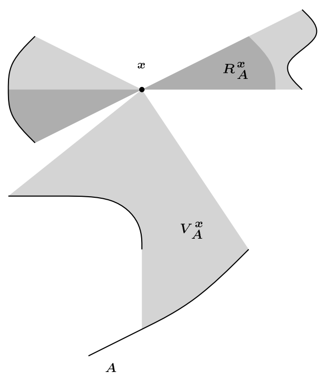

Figure 9: Reflection star at x.

are visible from x: $V_A^x = \bigcup\{\overline{xa} \mid \ a \in A$ and $A \cap \overline{xa} = \varnothing\}$. The *reflection star* R_A^x is defined by intersecting V_A^x with its reflection in x: $R_A^x = \{x+v \in \mathbb{R}^2 \mid x-v \in V_A^x$ and $x+v \in V_A^x\}$, see figure9. The function $\rho_A : \mathbb{R}^2 \to \mathbb{R}$ is the area of the reflection star in each point: $\rho_A(x) = \text{area}(R_A^x)$. Observe that for points x outside the convex hull of A, this area is always zero. The reflection metric between patterns A and B defines a normalized difference of the corresponding functions ρ_A and ρ_B:

$$d(A,B) = \frac{\int_{\mathbb{R}^2} |\rho_A(x) - \rho_B(x)| \, dx}{\int_{\mathbb{R}^2} \max(\rho_A(x), \rho_B(x)) \, dx}.$$

From the definition follows that the reflection metric is invariant under all affine transformations. In contrast with single-curve patterns, this metric is defined also for patterns consisting of multiple curves. In addition, the reflection metric is deformation, blur, crack, and noise robust.

3.13 Area of Overlap

Two dissimilarity measures that are based on the area of the polygons rather than their boundaries, are the area of overlap and the area of symmetric difference (see next subsection). For two compact sets A and B, the area of overlap is defined as $area(A \cap B)$. This dissimilarity measure is a not a metric, since the triangle inequality does not hold. The invariance group is the class of diffeomorphisms with unit Jacobi-determinant.

It turns out that translating the polygons so that their centroids coincide gives an overlap of at least 9/25 of the optimal solution [12]. For translations, the transformation that maximizes the area overlap also minimizes the area of symmetric difference.

3.14 Area of Symmetric Difference, Template Metric

For two compact sets A and B, the area of symmetric difference, also called template metric, is defined as $area((A-B) \cup (B-A))$. Unlike the area of overlap, this measure is a metric.

Translating convex polygons so that their centroids coincide also gives an approximate solution for the symmetric difference, which is at most 11/3 of the optimal solution under translations [3]. This also holds for a set of transformations F other than translations, if the following holds: the centroid of A, $c(A)$, is equivariant under the transformations, i.e. $c(f(A)) = f(c(A))$ for all f in F, and F is closed under composition with translation.

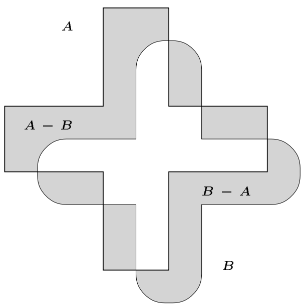

Figure 10: Area of overlap and symmetric difference.

3.15 Banach-Mazur Distance

For any two convex bodies A and B of the Euclidean plane, let $\lambda(A, B)$ be the smallest ratio s/r where $r, s > 0$ satisfy $rB' \subseteq A \subseteq sB''$, and B', B'' are some translates of B. Let $\tilde{B}$ denote class of bodies equivalent to B under translation and positive scaling (the homothets of B). The function $\tilde{\lambda}(\tilde{A}, \tilde{B}) = \log \lambda(\tilde{A}, \tilde{B})$ is a metric on shapes and is called the Banach-Mazur metric. It is invariant under affine transformations [15].

3.16 Monge-Kantorovich Metric, Transport Metric, Earth Mover's Distance

Given two patterns $A = \{(A_1, w(A_1)), \ldots, (A_m, w(A_m))\}$ and $B = \{(B_1, w(B_1)), \ldots, (B_n, w(B_n))\}$, where A_i and B_i are subsets of $\mathbb{R}^2$, with associates weights $w(A_i), w(B_i)$. The distance between A and B is the minimum amount of work needed to transform A into B. (The notion of work as in physics: the amount of energy needed to displace a mass.) This is a form of the Monge-Kantorovich metric used in heat transform problems [26], which is also used in shape matching [16] and color-based image retrieval [27]. The discrete version can be computed by linear programming.

4 Constructions

In this section we discuss a number of constructions that can be used to manipulate similarity measures, in order to arrive at certain properties.

4.1 Remapping

Let $w : [0, \infty] \to [0, \infty]$ be a monotone continuous function with $w(x) = 0$ iff $x = 0$, and which is subadditive: $w(x + y) \leq w(x) + w(y)$. Examples include mappings to $x/(1 + x)$, $\tan^{-1}(x)$, $\log(x)$, $x^{1/p}$, for some $p \geq 1$, and $\min(x, c)$, for some positive constant c. If $d(A, B)$ is a metric, then so is $\tilde{d}(A, B) = w(d(A, B))$. With some of these functions, an unbounded metric d can be mapped to a bounded metric. For the cut-off function $\min(x, c)$, the maximum distance value becomes c, so that property (14) above does not hold. It is used in [7] for comparing binary images. The $\log(x)$ function is used in the Banach-Mazur distance $\log \lambda(\tilde{A}, \tilde{B})$. Without the logarithm it would not satisfy the triangle inequality, and therefore not be a metric.

4.2 Normalization

Normalization is often used to scale the range of values of a similarity measure to $[0, 1]$, but it can also change other properties. For example, normalizing the area of overlap and symmetric difference by the area of the union of the two polygons makes it invariant under a larger transformation group, namely the group of all diffeomorphisms with a Jacobi determinant that is constant over all points [17].

4.3 From Semi-metric to Metric

Let S be a space of objects, and d a semimetric. Identifying all elements A, B of S for which $d(A, B) = 0$, and considering these as a single object, yields another space S'. The semimetric on S then gives a metric $d'(C, D) = d(C, D)$ on S'.

4.4 Semi-metric on Orbits

A collection of patterns S and a transformation group G determine a family of equivalence classes S/G. For a pattern $A \in S$, the orbit is $G(A) = \{g(A) \mid g \in G\}$. The collection of all these orbits forms a space of equivalences classes. A semimetric d invariant under a transformation group G results in a natural semimetric on the orbit set: $\tilde{d} : S/G \times S/G \to \mathbb{R}$ defined by $\tilde{d}(G(A), G(B)) = \inf\{d(g(A), B) \mid g \in G\}$ is a semimetric on the space S/G. Rucklidge [28] used this principle to define a shape distance based on the Hausdorff distance.

4.5 Extension with empty set

A pattern space S not containing the empty set $\varnothing$, with metric d, can be extended with $\varnothing$, by defining $d'(A, B) = d(A, B)/(1 + d(A, B))$, $d'(\varnothing, \varnothing) = 0$, and $d'(A, \varnothing) = 1$ for $A, B \in S$. This gives a bounded metric pattern space such that the restriction of d' to S is topologically equivalent to d [17]. In addition, the invariance group remains the same.

4.6 Vantageing

Let d be some distance function on a space S of patterns, $d : S \times S \to \mathbb{R}$. For some fixed $C \in S$ (vantage object), the function $\tilde{d}_C(A, B) = |d(A, C) - d(B, C)|$ is a semimetric, even if d does not obey nonnegativity, identity, weak triangle inequality, and symmetry.

4.7 Imbedding patterns

Affine invariant pattern metrics can be formed by mapping patterns to real-valued functions and computing a normalized difference between these functions. Affine invariance is desired in many pattern matching and shape recognition tasks.

Let $\mathbf{I}(\mathbb{R}^2)$ be the space of real-valued integrable functions on $\mathbb{R}^2$. Define the L^1 seminorm on $\mathbf{I}(\mathbb{R}^2)$ by $|\mathbf{a}| = \int_{\mathbb{R}^2} |\mathbf{a}(x)| \, dx$.

For a diffeomorphism g the Jacobi-determinant is the determinant of the derivative of g at a given point. We use $j_g(x)$ to denote the absolute value of the Jacobi-determinant of g in x. For real-valued functions $\mathbf{a}, \mathbf{b} : \mathbb{R}^2 \to \mathbb{R}$, $\mathbf{a} \vee \mathbf{b}$ denotes the pointwise maximum. Define the *normalized difference* of two functions with non-zero integrals by $\sigma_n(\mathbf{a}, \mathbf{b}) = |\mathbf{a} - \mathbf{b}|/|\mathbf{a} \vee \mathbf{b}|$. This is a semimetric on the set of non-negative valued functions with non-zero integrals.

A large class of mappings from patterns in $\mathbb{R}^2$ to integrable functions result in invariant semimetrics based on the normalized difference σ_n. Namely, let S be a collection of subsets of $\mathbb{R}^2$. Let each $A \in S$ define a unique function $\mathbf{n}_A : \mathbb{R}^2 \to \mathbb{R}$ in $\mathbf{I}(\mathbb{R}^2)$, and let g be a diffeomorphism with constant Jacobi-determinant. If g determines a number $\delta > 0$ such that $\mathbf{n}_{g(A)}(g(x)) = \delta \mathbf{n}_A(x)$ for all $A \in S$ and $x \in \mathbb{R}^2$, then $\sigma_n(\mathbf{n}_{g(A)}, \mathbf{n}_{g(B)}) = \sigma_n(\mathbf{n}_A, \mathbf{n}_B)$ for all $A, B \in S$ [17]. This was used in the construction of the reflection metric.

5 Conclusions

In this paper, we have identified a number of useful properties of similarity measures. Some properties are conflicting, and the desirability depends on the application at hand. We have listed a number of similarity measures, some of which are not well known (such as the Monge-Kantorovich metric), or new (reflection metric). Finally, we gave a set of constructions that have been used in the design of some similarity measures. It is an interesting problem to design new similarity measures that satisfy new combinations of properties.

References

[1] O. Aichholzer, H. Alt, and G. Rote. Matching shapes with a reference point. In *International Journal of Computational Geometry and Applications*, volume 7, pages 349–363, August 1997.

[2] Helmut Alt, Bernd Behrends, and Johannes Blömer. Approximate matching of polygonal shapes. *Annals of Mathematics and Artificial Intelligence*, pages 251–265, 1995.

[3] Helmut Alt, Ulrich Fuchs, Günter Rote, and Gerald Weber. Matching convex shapes with respect to the symmetric difference. In *Algorithms ESA '96, Proceedings of the 4th Annual European Symposium on Algorithms, Barcelona, Spain, September '96*, pages 320–333. LNCS 1136, Springer, 1996.

[4] Helmut Alt and Michael Godeau. Computing the Fréchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, pages 75–91, 1995.

[5] Esther Arkin, Paul Chew, Daniel Huttenlocher, Klara Kedem, and Joseph Mitchel. An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3):209–215, 1991.

[6] Alta Vista Photo Finder, `http://image.altavista.com/cgi-bin/avncgi`.

[7] A. J. Baddeley. An error metric for binary images. In W. Förstner and S. Ruwiedel, editors, *Robust Computer Vision: Quality of Vision Algorithms, Proceedings of the International Workshop on Robust Computer Vision, Bonn, 1992*, pages 59–78. Wichmann, 1992.

[8] L. P. Chew, M. T. Goodrich, D. P. Huttenlocher, K. Kedem, J. M. Kleinberg, and D. Kravets. Geometric pattern matching under Euclidean motion. *Computational Geometry, Theory and Applications*, 7:113–124, 1997.

[9] Paul Chew and Klara Kedem. Improvements on approximate pattern matching. In *3rd Scandinavian Workshop on Algorithm Theory*, Lecture Notes in Computer Science 621, pages 318–325. Springer, 1992.

[10] Scott D. Cohen and Leonidas J. Guibas. Partial matching of planar polylines under similarity transformations. In *Proceedings of the 8th Annual Symposium on Discrete Algorithms*, pages 777–786, 1997.

[11] G. Cortelazzo, G. A. Mian, G. Vezzi, and P. Zamperoni. Trademark shapes description by string-matching techniques. *Pattern Recognition*, 27:1005–1018, 1994.

[12] Mark de Berg, Olivier Devillers, Marc van Kreveld, Otfried Schwarzkopf, and Monique Teillaud. Computing the maximum overlap of two convex polygons under translation. In *Proc. 7th Annu. Internat. Sympos. Algorithms Comput.*, 1996.

[13] Alon Efrat and Alon Itai. Improvements on bottleneck matching and related problems using geometry. *Proceedings of the 12th Symposium on Computational Geometry*, pages 301–310, 1996.

[14] Ronald Fagin and Larry Stockmeyer. Relaxing the triangle inequality in pattern matching. *International Journal of Computer Vision*, 28(3):219–231, 1998.

[15] Rudolf Fleischer, Kurt Mehlhorn, Günther Rote, Emo Welzl, and Chee Yap. Simultaneous inner and outer approximation of shapes. *Algorithmica*, 8:365–389, 1992.

[16] David Fry. *Shape Recognition using Metrics on the Space of Shapes*. PhD thesis, Harvard University, Department of Mathematics, 1993.

[17] M. Hagedoorn and R. C. Veltkamp. Metric pattern spaces. Technical Report UU-CS-1999-03, Utrecht University, 1999.

[18] Michiel Hagedoorn and Remco C. Veltkamp. Reliable and efficient pattern matching using an affine invariant metric. *International Journal of Computer Vision*, 31(2/3):203–225, 1999.

[19] D. P. Huttenlocher, K. Kedem, and J. M. Kleinberg. On dynamic Voronoi diagrams and the minimum Hausdorff distance for point sets under Euclidean motion in the plane. In *Proceedings of the 8th Annuual ACM Symposium on Computational Geometry*, pages 110–120, 1992.

[20] Daniel P. Huttenlocher, Klara Kedem, and Micha Sharir. The upper envelope of Voronoi surfaces and its applications. *Discrete and Computational Geometry*, 9:267–291, 1993.

[21] Daniel P. Huttenlocher, Gregory A. Klanderman, and William J. Rucklidge. Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machinen Intelligence*, 15:850–863, 1993.

[22] J. B. Kruskal. An overview of sequence comparison: Time warps, string edits, and macromolecules. *SIAM Review*, 25:201–237, 1983.

[23] Joseph O'Rourke. Curve similarity via signatures. In G. T. Toussaint, editor, *Computational Geometry*, pages 295–317. North-Holland, 1985.

[24] Plato. *Meno*. Perseus Encyclopedia, Tuft University, `http://www.perseus.tufts.edu/Texts/chunk_TOC.grk.html#Plato`, 380 B.C.

[25] QBIC project, `http://wwwqbic.almaden.ibm.com/`.

[26] S. Rachev. The Monge-Kantorovich mass transference problem and its stochastical applications. *Theory of Probability and Applications*, 29:647–676, 1985.

[27] Y. Rubner, C. Tomassi, and L Guibas. A metric for distributions with applications to image databases. In *Proceedings of the IEEE International Conference on Computer Vision, Bombay, India*, pages 59–66, 1998.

[28] W. Rucklidge. *Efficient Visual Recognition Using the Hausdorff Distance*. Lecture Notes in Computer Science. Springer, 1996.

[29] Stefan Schirra. Approximate decision algorithms for approximate congruence. *Information Processing Letters*, 43:29–34, 1992.

[30] Lambert Schomaker, Edward de Leau, and Louis Vuurpijl. Using pen-based outlines for object-based annotation and image-based queries. In D. P. Huijsmans and A. W. M. Smeulders, editors, *Visual Information and Information Systems – Proceedings of the Third International Conference VISUAL'99, Amsterdam, The Netherlands, June 1999*, LNCS 1614, pages 585–592. Springer, 1999.

[31] D'Arcy W. Thomsom. Morphology and mathematics. *Transactions of the Royal Society of Edinburgh*, Volume L, Part IV, No. 27:857–895, 1915.

[32] Amos Tversky. Features of similarity. *Psychological Review*, 84(4):327–352, 1977.

[33] Jules Vleugels and Remco C. Veltkamp. Efficient image retrieval through vantage objects. In D. P. Huijsmans and A. W. M. Smeulders, editors, *Visual Information and Information Systems – Proceedings of the Third International Conference VISUAL'99, Amsterdam, The Netherlands, June 1999*, LNCS 1614, pages 575–584. Springer, 1999.