DISTANCES BETWEEN SETS — A SURVEY

Aura Conci   and   Carlos Kubrusly

ABSTRACT. The purpose of this paper is to give a survey on the notions of distance between subsets either of a metric space or of a measure space, including definitions, a classification, and a discussion of the best-known distance functions, which is followed by a review on applications used in many areas of knowledge, ranging from theoretical to practical applications.

1. INTRODUCTION

Distances between subsets either of a metric space or of a measure space is the reason of this paper, where the main focus is to define, classify, and review the best-known distance functions, and some of their common applications. Most distances between sets, defined either for closed and bounded (in particular, for compact, more particularly, for finite) nonempty sets in a metric space, or for measurable sets with finite measure (in particular, for finite sets under the counting measure) in a measure space, become a metric themselves. Behind the scenes, the main tools are naturally the notions of metric space and measure space. Although the notion of distance between sets has a myriad of conceivable definitions (see, e.g., [16, pp. 46-48, 85-86, 173-184, 298-301, 359-360], there are two main distinct families of them, which we refer to as the *Hausdorff family* and the *Measure Theoretical family*.

Notation, terminology, basic definitions and a few results that will be required in the sequel are posed in Section 2. These are naturally bound to the notions of metric and measure. Section 3 gives a detailed account on the Hausdorff distance (which is a metric for closed and bounded sets) and its many relatives, which are obtained by gradual modifications of the original Hausdorff distance. Four equivalent forms of the original Hausdorff measure are discussed, as well as fifteen variations of it. The measure theoretical approach, dealing with the Fréchet–Nikodým–Aronszajn distance (and also with its normalized version, the Markzewisky–Steinhaus distance) is discussed in Section 4. Section 5 closes the paper with a review of the bibliography (since the 80's) dealing with applications of distance between sets towards innumerable subjects. This is split into three classes, namely, (1) Computational Aspects (with three subclasses:(1.1) distance in graphs, (1.2) distance between polygons, and (1.3) numerical procedures and algorithms), (2) On Distances Between Fuzzy Sets (also with three subclasses: (2.1) Markzewisky–Steinhaus distance, (2.2) Hausdorff distances, (2.3) non-Hausdorff distances), and (3) Distance in Object Analysis (again with three subsections: (3.1) new metrics and comparisons, (3.2) motion — translation and rotation, and (3.3) modified Hausdorff including asymmetries).

2. Notation and Terminology

This section summarizes classical standard topics that will be required in the sequel, which can be found in an infinitude of books dealing with analysis in general. For instance, see [42, Chapter 3] for metric space properties and [41, Chapter 2] for measure space properties, among many others.

Let X be an arbitrary nonempty set, let d be a real-valued function on the Cartesian product $X \times X$ of X with itself,

$$d \colon X \times X \to \mathbb{R},$$

and consider the following properties, holding for arbitrary points x, y, z in X.

(i) $d(x, y) = d(y, x)$ (*symmetry*),
(ii) $d(x, y) \geq 0$ and $d(x, x) = 0$ (*nonnegativeness*),
(iii) $d(x, y) = 0$ implies $x = y$ (*positiveness*),
(iv) $d(x, y) \leq d(x, z) + d(z, y)$ (*triangle inequality*).

A real-valued function d on $X \times X$ that satisfies all properties (i), (ii), (iii) and (iv) is a *metric* in X, and the properties themselves are called the *metric axioms*. A set X equipped with a metric d on $X \times X$ is a *metric space*, also denoted by (X, d). If d satisfies properties (i), (ii) and (iv), but not necessarily property (iii), then it is called a *pseudometric* in X. If d satisfies properties (i), (ii) and (iii), but not necessarily (iv), then it is sometimes called a *semimetric* in X. What is commonly referred to as a *distance function* in X is simply any real-valued function d on $X \times X$ that satisfies properties (i) and (ii) — i.e., any symmetric nonnegative function that vanishes at the identity line.

Remark 2.1. The difference between a metric and a pseudometric is that it is possible for a pseudometric d to vanish at a pair (x, y) even if $x \neq y$. A pseudometric d is trivial if $d(x, y) = 0$ for every $x, y \in X$. However, given a nontrivial pseudometric d in X there is a natural way to obtain a metric space $(\widetilde{X}, \widetilde{d})$, where the set $\widetilde{X}$ is a associated with X and d, and $\widetilde{d}$ (usually denoted again by d) is the natural metric in $\widetilde{X}$ inherited from d. Indeed, every nontrivial pseudometric d induces an equivalence relation $\sim$ on X (given by $x \sim x'$ — read x is equivalent to x' — if and only if $d(x, x') = 0$), and $\widetilde{X}$ is the quotient space $X/\sim$ which is precisely the collection of all equivalence classes $[x] = \{x \in X \colon x' \sim x\}$ with respect to $\sim$ for every element x in X (i.e., $X/\sim$ is a collection of sets $[x]$, called equivalence classes, such that each element in $X/\sim$ is a set consisting of all elements from X that are equivalent to each other), and the metric $\widetilde{d}$ in $\widetilde{X}$ is given by $\widetilde{d}([x], [y]) = d(x, y)$, which does not depend on the representatives $x \in [x]$ and $y \in [y]$ from the equivalence classes.

The *power set* $\wp(X)$ of a given set X is the collection of all subsets of X. Let X be an arbitrary nonempty set. Equip X with a metric $d \colon X \times X \to \mathbb{R}$ and consider the metric space (X, d). A nonempty subset A of X is *bounded* if

$$\sup_{x, y \in A} d(x, y) < \infty,$$

otherwise A is said to be unbounded, which is denoted by $\sup_{x, y \in A} d(x, y) = \infty$. The *diameter* of a nonempty bounded subset A of X is the real number

$$\operatorname{diam}(A) = \sup_{x, y \in A} d(x, y).$$

The *distance from a point $x \in X$ to a nonempty set $A \in \wp(X)$* is the real number
$$d(x, A) = \inf_{a \in A} d(x, a),$$
and the *ordinary distance function between two nonempty sets A and B in $\wp(X)$* is the real number
$$d(A, B) = \inf_{a \in A,\, b \in B} d(a, b).$$
The above expression defines a mere distance function $d \colon \wp(X) \backslash \varnothing \times \wp(X) \backslash \varnothing \to \mathbb{R}$. In fact, such a function d trivially satisfies properties (i) and (ii) for all sets in $\wp(X) \backslash \varnothing$. It is however clear that in general d does not satisfy property (iii), and it does not satisfy property (iv) as well — e.g., take A, B, C in $\wp(X) \backslash \varnothing$ (which may even be pairwise disjoint) such that $d(A, B) \neq 0$ and $d(A, C) = d(C, B) = 0$.

A *σ-algebra* $\mathcal{A}(X)$ of subsets of a given set X is a subcollection of the power set, $\mathcal{A}(X) \subseteq \wp(X)$ (not necessarily a proper subcollection), such that

  the whole set X and the empty set $\varnothing$ belong to $\mathcal{A}(X)$,
  the complement $X \backslash E$ of a set E in $\mathcal{A}(X)$ belongs to $\mathcal{A}(X)$,
  the union of a countable collection of sets in $\mathcal{A}(X)$ belongs to $\mathcal{A}(X)$.

Sets in $\mathcal{A}(X)$ are called *measurable sets*, and the pair $(X, \mathcal{A}(X))$ consisting of a set X and a σ-algebra of subsets of it is referred to as a *measurable space*. A *measure* is an extended real-valued function μ on a σ-algebra $\mathcal{A}(X)$,
$$\mu \colon \mathcal{A}(X) \to \overline{\mathbb{R}},$$
where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ stands for the extended real line, satisfying the following properties (referred to as the *measure axioms*).

  (a) $\mu(\varnothing) = 0$,
  (b) $\mu(E) \geq 0$ for every $E \in \mathcal{A}(X)$,
  (c) $\mu\big(\bigcup_n E_n\big) = \sum_n \mu(E_n)$
      for every countable family $\{E_n\}$ of pairwise disjoint sets in $\mathcal{A}(X)$.

A *measure space* is a triple $(X, \mathcal{A}(X), \mu)$ consisting of an arbitrary set X, a σ-algebra $\mathcal{A}(X)$ of subsets of X, and a measure μ on $\mathcal{A}(X)$. We assume throughout this paper that all measures are nonzero (i.e., $\mu(X) > 0$).

Consider an arbitrary nonempty subcollection $\mathcal{C}(X)$ of $\wp(X)$. If an arbitrary distance function d on $\mathcal{C}(X) \backslash \varnothing \times \mathcal{C}(X) \backslash \varnothing$ is intended for gauging resemblance of sets in $\mathcal{C}(X) \backslash \varnothing$, then it may be useful to control it by preventing too large (and too small) values. In this case d can be normalized in terms of a measure μ on a σ-algebra of subsets of X, bringing forth new distance functions. For instance, suppose μ is any nonzero *finite* measure (i.e., $0 < \mu(X) < \infty$) on any σ-algebra $\mathcal{A}(X)$ of subsets of X, let $d \colon \mathcal{C}(X) \backslash \varnothing \times \mathcal{C}(X) \backslash \varnothing \to \mathbb{R}$ be the ordinary distance function in $\mathcal{C}(X) \backslash \varnothing$, and consider the distance function $d' \colon \mathcal{C}(X) \backslash \varnothing \times \mathcal{C}(X) \backslash \varnothing \to \mathbb{R}$ given by
$$d'(A, B) = \tfrac{1}{\mu(X)} d(A, B) = \tfrac{1}{\mu(X)} \inf_{a \in A,\, b \in B} d(a, b)$$
for every sets A, B in $\mathcal{C}(X) \backslash \varnothing$, or the distance functions in the intersection of $\mathcal{C}(X)$ and $\mathcal{A}(X)$, say $d'', d''' \colon (\mathcal{C}(X) \cap \mathcal{A}(X)) \backslash \varnothing \times (\mathcal{C}(X) \cap \mathcal{A}(X)) \backslash \varnothing \to \mathbb{R}$, given for every sets A, B in $(\mathcal{C}(X) \cap \mathcal{A}(X)) \backslash \varnothing$ by
$$d''(A, B) = \tfrac{1}{\mu(A \cup B)} d(A, B) \quad \text{and} \quad d'''(A, B) = \tfrac{1}{\mu(A) + \mu(B)} d(A, B)$$
if $\mu(A) \neq 0$ or $\mu(B) \neq 0$, otherwise $d''(A, B)$ and $d'''(A, B)$ are defined to be zero.

Let $\#A$ denotes the cardinality of an arbitrary set $A \in \wp(X)$. Let $\mathcal{F}(X) \subseteq \wp(X)$ stand for the collection of all *finite* subsets of the nonempty set X. Consider the collection $\mathcal{F}(X) \backslash \varnothing \subseteq \wp(X) \backslash \varnothing$ of all finite nonempty subsets of X. Take arbitrary finite nonempty sets $A = \{a_i\}_{i=1}^m$ and $B = \{b_i\}_{i=1}^n$ in $\mathcal{F}(X) \backslash \varnothing$, where $m = \#A$ and $n = \#B$ lie in $\mathbb{N}$, the set of all positive integers. Now consider the restriction $d|_{\mathcal{F}(X) \backslash \varnothing \times \mathcal{F}(X) \backslash \varnothing}$ of the ordinary distance function $d \colon \wp(X) \backslash \varnothing \times \wp(X) \backslash \varnothing \to \mathbb{R}$ to pairs of finite nonempty sets, $d \colon \mathcal{F}(X) \backslash \varnothing \times \mathcal{F}(X) \backslash \varnothing \to \mathbb{R}$, denoted again by the same symbol d. Thus in this case, for every $x \in X$ and every $A, B \in \mathcal{F}(X) \backslash \varnothing$,

$$d(x, A) \;=\; \inf_{a \in A} d(x, a) = \min_{a \in A} d(x, a) = \min_{1 \le i \le m} d(x, a_i),$$

$$d(A, B) \;=\; \inf_{a \in A,\, b \in B} d(a, b) \;=\; \min_{a \in A,\, b \in B} d(a, b) \;=\; \min_{1 \le i \le m,\, 1 \le j \le n} d(a_i, b_j);$$

similarly for the normalized versions (to avoid trivialities, suppose $X \in \mathcal{F}(X)$):

$$d'(A, B) = \tfrac{1}{\#X} d(A, B), \quad d''(A, B) = \tfrac{1}{\#(A \cup B)} d(A, B), \quad d'''(A, B) = \tfrac{1}{\#A + \#B} d(A, B).$$

Still in these cases the functions d, d', d'', d''' on $\mathcal{F}(X) \backslash \varnothing \times \mathcal{F}(X) \backslash \varnothing$ are distance functions but it is clear that properties (iii) and (iv) may fail (even if $\#X < \infty$).

3. The Hausdorff Family

Let (X, d) be a metric space. Perhaps the best-known candidate for a metric in a subset of $\wp(X) \backslash \varnothing$ is the Hausdorff function $h \colon \wp(X) \backslash \varnothing \times \wp(X) \backslash \varnothing \to \overline{\mathbb{R}}$, which is defined for every $A, B \in \wp(X) \backslash \varnothing$ by [29, p.293] (see also [30, p.167]).

$$h(A, B) \;=\; \max \Big\{ \sup_{a \in A} d(a, B)\,,\; \sup_{b \in B} d(b, A) \Big\}$$

$$=\; \max \Big\{ \sup_{a \in A} \inf_{b \in B} d(a, b)\,,\; \sup_{b \in B} \inf_{a \in A} d(a, b) \Big\}.$$

Although the Hausdorff function satisfies properties (i) and (ii) and has the property that $h(\{a\}, \{b\}) = d(a, b)$ for every pair of singletons $\{a\}, \{b\} \in \wp(X) \backslash \varnothing$, it is not a distance function because it is not real-valued. For instance, on $\wp(\mathbb{R}) \backslash \varnothing \times \wp(\mathbb{R}) \backslash \varnothing$ we get $h(\{0\}, \mathbb{R}) = +\infty$. Even if it were plausible to admit an extended-real-valued distance function, the function $h \colon \wp(\mathbb{R}) \backslash \varnothing \times \wp(\mathbb{R}) \backslash \varnothing \to \overline{\mathbb{R}}$ would not be a pseudometric since property (iii) fails (e.g., $h(A, B) = 0$ for $A = (0, 1)$, $B = [0, 1]$ in $\wp(\mathbb{R}) \backslash \varnothing$). An equivalent formulation for the Hausdorff function (also called Blaschke function [16, p.48], or Pompeiu–Hausdorff function [50, Example 4.3], [7]) is given below.

Proposition 3.1. *For every* $A, B \in \wp(X) \backslash \varnothing$,

$$h(A, B) = \sup_{x \in X} |d(x, A) - d(x, B)|.$$

Proof. Take $x \in X$, $a \in A$, $b \in B$. Since $d(x, b) \le d(x, a) + d(a, b)$, we get $d(x, B) = \inf_{b \in B} d(x, b) \le d(x, a) + \inf_{b \in B} d(a, b) \le d(x, a) + \sup_{a \in A} \inf_{b \in B} d(a, b)$, and hence $d(x, B) \le \inf_{a \in A} d(x, a) + \sup_{a \in A} \inf_{b \in B} d(a, b) = d(x, A) + \sup_{a \in A} \inf_{b \in B} d(a, b)$. Symmetrically, $d(x, A) \le d(x, B) + \sup_{b \in B} \inf_{a \in A} d(a, b)$. Thus, for every $x \in X$,

$$|d(x, A) - d(x, B)| \le \max \Big\{ \sup_{a \in A} \inf_{b \in B} d(a, b)\,,\; \sup_{b \in B} \inf_{a \in A} d(a, b) \Big\}.$$

In particular, for $x = a \in A$, we get $\sup_{a \in A} \inf_{b \in B} d(a, b) = \sup_{x = a \in A} d(x, B) = \sup_{x = a \in A} |d(x, A) - d(x, B)|$. Symmetrically, by setting $x = b \in B$, we also get $\sup_{b \in B} \inf_{a \in A} d(a, b) = \sup_{x = b \in B} |d(x, A) - d(x, B)|$. Therefore,

$$\max \Big\{ \sup_{a \in A} \inf_{b \in B} d(a, b)\,,\; \sup_{b \in B} \inf_{a \in A} d(a, b) \Big\} \le \sup_{x \in X} |d(x, A) - d(x, B)|. \qquad \square$$

Although the Hausdorff function is not even a distance function when defined on the domain $\wp(X)\backslash\varnothing\times\wp(X)\backslash\varnothing$, if restricted to an appropriate subcollection it becomes a metric. Let $\mathcal{B}(X)\subseteq\wp(X)$ denote the collection of all *closed and bounded* subsets of the metric space (X,d), take the collection $\mathcal{B}(X)\backslash\varnothing\subseteq\wp(X)\backslash\varnothing$ of all nonempty closed and bounded subsets of (X,d), and consider the restriction $h|_{\mathcal{B}(X)\backslash\varnothing\times\mathcal{B}(X)\backslash\varnothing}$ of $h\colon\wp(X)\backslash\varnothing\times\wp(X)\backslash\varnothing\to\overline{\mathbb{R}}$ to $\mathcal{B}(X)\backslash\varnothing\times\mathcal{B}(X)\backslash\varnothing$, denoted by the same symbol,

$$h\colon\mathcal{B}(X)\backslash\varnothing\times\mathcal{B}(X)\backslash\varnothing\to\mathbb{R}.$$

This is not only real-valued but is a metric in $\mathcal{B}(X)\backslash\varnothing$ [18, Problem IX.4.8], [8, Example 1.9.7]. In particular, since $\mathcal{F}(X)\subseteq\mathcal{B}(X)$, when restricted still further, now to $\mathcal{F}(X)\backslash\varnothing\times\mathcal{F}(X)\backslash\varnothing$ (i.e., when restricted to nonempty finite sets), the Hausdorff function (again, denoted by the same symbol h)

$$h\colon\mathcal{F}(X)\backslash\varnothing\times\mathcal{F}(X)\backslash\varnothing\to\mathbb{R},$$

given for every $A=\{a_i\}_{i=1}^m$ and $B=\{b_i\}_{i=1}^n$ in $\mathcal{F}(X)\backslash\varnothing$ by

$$h(A,B)=\max\Big\{\max_{1\leq i\leq m}\ \min_{1\leq j\leq n}d(a_i,b_j)\,,\ \max_{1\leq j\leq n}\ \min_{1\leq i\leq m}d(a_i,b_j)\Big\},$$

is a metric in $\mathcal{F}(X)\backslash\varnothing$. For more on theoretical aspects of Hausdorff distance see, for instance, [57, Chapter 2] (also [56], [58]), and for the topology of Hausdorff distances see [4]. Indeed, the Hausdorff metric has many appealing properties, but it shares some practical drawbacks [5, Section 3.4], [19, Section 2]. For an account on theoretical and practical aspects of some common distances between finite sets, including Hausdorff's see, e.g., [5, 36, 17, 19, 25, 26] and the references therein.

Let $\mathcal{C}(X)$ stand either for $\mathcal{B}(X)$ or $\mathcal{F}(X)$. Associated with each $A\in\wp(X)$ set

$$\text{size}(A)=\begin{cases}\#A, & \text{if }\ \mathcal{C}(X)=\mathcal{F}(X),\\[4pt]\text{diam}(A), & \text{if }\ \mathcal{C}(X)=\mathcal{B}(X)\neq\mathcal{F}(X),\end{cases}$$

(to avoid trivialities suppose $\text{size}(X)>0$) and consider the Hausdorff metric

$$h\colon\mathcal{C}(X)\backslash\varnothing\times\mathcal{C}(X)\backslash\varnothing\to\mathbb{R}.$$

A natural normalization $h'\colon\mathcal{C}(X)\backslash\varnothing\times\mathcal{C}(X)\backslash\varnothing\to\mathbb{R}$ of h is given by

$$h'(A,B)=\tfrac{1}{\text{size}(X)}\,h(A,B)=\tfrac{1}{\text{size}(X)}\max\Big\{\sup_{a\in A}d(a,B)\,,\ \sup_{b\in B}d(b,A)\Big\}$$
$$=\tfrac{1}{\text{size}(X)}\sup_{x\in X}|d(x,A)-d(x,B)|$$

(cf. Proposition 3.1), which is again a metric (since h is a metric) if $\text{size}(X)<\infty$. When dealing with normalized versions of metrics on $\mathcal{C}(X)\backslash\varnothing\times\mathcal{C}(X)\backslash\varnothing$ it is advisable to assume that $\text{size}(X)<\infty$ (i.e., X is bounded if $\mathcal{C}(X)=\mathcal{B}(X)\neq\mathcal{F}(X)$, or X is finite if $\mathcal{C}(X)=\mathcal{F}(X)$) in order to avoid trivial pseudometrics. Further natural normalizations of h, say $h'',h'''\colon\mathcal{C}(X)\backslash\varnothing\times\mathcal{C}(X)\backslash\varnothing\to\mathbb{R}$, lead to the distances

$$h''(A,B)=\tfrac{1}{\text{size}(A\cup B)}\,h(A,B)\quad\text{and}\quad h'''(A,B)=\tfrac{1}{\text{size}(A)+\text{size}(B)}\,h(A,B)$$

if $\text{size}(A)\neq 0$ or $\text{size}(B)\neq 0$; otherwise they are zero. Straightforward modifications of h, namely, $\widetilde{h},\widetilde{h}'\colon\mathcal{C}(X)\backslash\varnothing\times\mathcal{C}(X)\backslash\varnothing\to\mathbb{R}$, where maximum is replaced with sum,

$$\widetilde{h}(A,B)=\sup_{a\in A}d(a,B)+\sup_{b\in B}d(b,A),$$

$$\widetilde{h}'(A,B)=\tfrac{1}{\text{size}(X)}\widetilde{h}(A,B)=\tfrac{1}{\text{size}(X)}\Big(\sup_{a\in A}d(a,B)+\sup_{b\in B}d(b,A)\Big),$$

are still metrics in $\mathcal{C}(X)\backslash\varnothing$ (if $\mathrm{size}(X) < \infty$ — see [59, Section 1]). Again, further natural normalizations $\widetilde{h}'', \widetilde{h}''' \colon \mathcal{C}(X)\backslash\varnothing \times \mathcal{C}(X)\backslash\varnothing \to \mathbb{R}$ of $\widetilde{h}$ are

$$\widetilde{h}''(A,B) = \tfrac{1}{\mathrm{size}(A\cup B)}\,\widetilde{h}(A,B) \quad \text{and} \quad \widetilde{h}'''(A,B) = \tfrac{1}{\mathrm{size}(A)+\mathrm{size}(B)}\,\widetilde{h}(A,B)$$

if $\mathrm{size}(A) \neq 0$ or $\mathrm{size}(B) \neq 0$, otherwise the distances are set down to zero.

Another modified version $h_p \colon \mathcal{F}(X)\backslash\varnothing \times \mathcal{F}(X)\backslash\varnothing \to \mathbb{R}$ of h, where the supremum is replaced with a p-sum, given by

$$h_p(A,B) = \Big(\sum\nolimits_{x \in X} |d(x,A) - d(x,B)|^p\Big)^{\frac{1}{p}}$$

for any real $p \geq 1$ if $X \in \mathcal{F}(X)$, and its normalization $h_p' \colon \mathcal{F}(X)\backslash\varnothing \times \mathcal{F}(X)\backslash\varnothing \to \mathbb{R}$,

$$h_p'(A,B) = \Big(\tfrac{1}{\#X}\sum\nolimits_{x \in X} |d(x,A) - d(x,B)|^p\Big)^{\frac{1}{p}},$$

are referred to as p-Hausdorff distances (see [16, p.48,359]), which in fact are metrics [5, eq.21]). Further variants of h, viz., $h_k, h_k' \colon \mathcal{F}(X)\backslash\varnothing \times \mathcal{F}(X)\backslash\varnothing \to \mathbb{R}$ for $k = 0, 1$ (not all metrics), read as follows. First modify h by replacing supremum with sum,

$$h_0(A,B) = \max\Big\{\sum\nolimits_{a \in A} d(a,B)\,, \ \sum\nolimits_{b \in B} d(b,A)\Big\}.$$

Its averaged version,

$$h_0'(A,B) = \max\Big\{\tfrac{1}{\#A}\sum\nolimits_{a \in A} d(a,B)\,, \ \tfrac{1}{\#B}\sum\nolimits_{b \in B} d(b,A)\Big\},$$

is called modified Hausdorff distance in [16, p.360] (see also [17, eq.6,8]). Next modify h by replacing maximum with sum in h_0,

$$h_1(A,B) = \sum\nolimits_{a \in A} d(a,B) + \sum\nolimits_{b \in B} d(b,A)$$

(considered in [19] with a multiplicative factor of $\tfrac{1}{2}$), and again its averaged version,

$$h_1'(A,B) = \tfrac{1}{\#A}\sum\nolimits_{a \in A} d(a,B) + \tfrac{1}{\#B}\sum\nolimits_{b \in B} d(b,A)$$

(considered in [17, eq.6,9]) with a factor of $\tfrac{1}{2}$). Caution: h_1 and h_1' are different from h_p and h_p' for $p = 1$. A normalized version $h_1'' \colon \mathcal{F}(X)\backslash\varnothing \times \mathcal{F}(X)\backslash\varnothing \to \mathbb{R}$ of h_1,

$$h_1''(A,B) = \tfrac{1}{\#A+\#B}\Big(\sum\nolimits_{a \in A} d(a,B) + \sum\nolimits_{b \in B} d(b,A)\Big),$$

is referred to as geometric mean error between two images in [16, p.360] (see also [17, eq.6,10]). Properties (i) and (ii) hold trivially, and property (iii) is readily verified, and hence all the above functions are semimetrics in $\mathcal{F}(X)\backslash\varnothing$. However, property (vi) may fail. For instance, if $A = \{a\}$, $B = \{b,c\}$, and $C = \{c\}$ with $0 < d(b,c) < d(a,b)$ and $0 < d(a,c) \leq d(a,b)$, then $h_k(A,C) + h_k(C,B) < h_k(A,B)$ for $k = 0,1$. So the semimetrics h_0, h_1 are a not metrics in $\mathcal{F}(X)\backslash\varnothing$. On the other hand, it was show in [25, Theorem 1] that the function $h_1''' \colon \mathcal{F}(X)\backslash\varnothing \times \mathcal{F}(X)\backslash\varnothing \to \mathbb{R}$ given by

$$h_1'''(A,B) = \tfrac{1}{\#(A\cup B)}\Big(\tfrac{1}{\#A}\sum\nolimits_{a \in A}\sum\nolimits_{b \in B\backslash A} d(a,b) + \tfrac{1}{\#B}\sum\nolimits_{b \in B}\sum\nolimits_{a \in A\backslash B} d(a,b)\Big)$$

is a metric in $\mathcal{F}(X)\backslash\varnothing$. For more combinations along these lines, including theoretical aspects or practical applications, see, e.g., [5, 36, 17, 19, 16, 25, 26].

The *envelope of radius ε centered at an arbitrary set* $A \in \wp(X)\backslash\varnothing$, also referred to as the *ε-envelope* (or even the *ε-neighborhood*) of A is the set

$$A_\varepsilon = \big\{x \in X \colon d(x,A) \leq \varepsilon\big\}.$$

If A is closed, so that $d(x, A) = \inf_{a \in A} d(x, a) = \min_{a \in A} d(x, a)$ for every $x \in X$, then $A_\varepsilon = \{x \in X : d(x, a) \leq \varepsilon$ for some $a \in A\}$. Another equivalent formulation for the Hausdorff function reads as follows [40, Problem 4.D] (also [21, Section 9.1]).

Proposition 3.2. *For every* $A, B \in \wp(X) \backslash \varnothing$,

$$h(A, B) = \inf \{\varepsilon \geq 0 : A \subseteq B_\varepsilon \text{ and } B \subseteq A_\varepsilon\}.$$

Proof. Take $A, B \in \wp(X) \backslash \varnothing$ arbitrary, set

$$h_A = \sup_{a \in A} d(a, B) \quad \text{and} \quad h_B = \sup_{b \in B} d(b, A),$$

and observe that

$$A \subseteq B_\varepsilon \iff h_A \leq \varepsilon \quad \text{and} \quad B \subseteq A_\varepsilon \iff h_B \leq \varepsilon.$$

(Indeed, $A \subseteq B_\varepsilon$ if and only if $d(a, B) \leq \varepsilon$ for all $a \in A$). Thus

$$h(A, B) = \max \{h_A, h_B\} \leq \inf \{\varepsilon \geq 0 : A \subseteq B_\varepsilon \text{ and } B \subseteq A_\varepsilon\}.$$

Conversely, note that

$$A \subseteq B_{h_A} \quad \text{and} \quad B \subseteq A_{h_B}$$

(in fact, $A \subseteq \{x \in X : d(x, B) \leq \sup_{a \in A} d(a, B)\} = B_{h_A}$). Therefore,

$$\inf \{\varepsilon \geq 0 : A \subseteq B_\varepsilon \text{ and } B \subseteq A_\varepsilon\} \leq \max \{h_A, h_B\} = h(A, B). \qquad \square$$

This version of the Hausdorff function is particularly useful to verify that the metric space $(\mathcal{B}(X) \backslash \varnothing \times \mathcal{B}(X) \backslash \varnothing, h)$ is complete if $X = \mathbb{R}^n$ equipped with its usual Euclidian metric (see, e.g., [20, p.37]) — this can be extended to any complete metric space (X, d) if $\mathcal{B}(X)$ is swapped with $\mathcal{K}(X)$, the collection of all compact subsets of the metic space (X, d) (where $\mathcal{F}(X) \subseteq \mathcal{K}(X) \subseteq \mathcal{B}(X) \subseteq \wp(X)$). Since $A \subseteq B$ if and only if $A \backslash B = \varnothing$, for any sets A and B, it follows that the equivalent expression for the Hausdorff function in Proposition 3.2 can also be rewritten as

$$h(A, B) = \inf \{\varepsilon \geq 0 : A \backslash B_\varepsilon = \varnothing \text{ and } B \backslash A_\varepsilon = \varnothing\},$$

yielding still another equivalent way to write the same Hausdorff function h which, if acting on $\mathcal{B}(X) \backslash \varnothing \times \mathcal{B}(X) \backslash \varnothing$ (in particular, on $\mathcal{F}(X) \backslash \varnothing \times \mathcal{F}(X) \backslash \varnothing$), is a metric. This last form has been used in some applications (see, e.g., [14]).

4. The Measure Theoretical Family

Let X be any nonempty set and consider the power set $\wp(X)$. The symmetric difference of two sets $A, B \in \wp(X)$ is the set

$$A \triangledown B = (A \backslash B) \cup (B \backslash A) = (A \cup B) \backslash (A \cap B),$$

so that $A = B$ if and only if $A \triangledown B = \varnothing$ (and $A \triangledown \varnothing = \varnothing \triangledown A = A$ for every $A \in \wp(X)$). Let $\mu : \mathcal{A}(X) \to \overline{\mathbb{R}}$ be an arbitrary nonzero measure (i.e., $\mu(X) > 0$) on an arbitrary σ-algebra $\mathcal{A}(X)$ of subsets of X (in particular, $\wp(X)$ may itself be a possible σ-algebra of subsets of X, depending on the measure μ). Consider the measure space $(X, \mathcal{A}(X), \mu)$. Two sets E, F in $\mathcal{A}(X)$ are *equivalent* (or *μ-equivalent*), denoted by $E \sim F$, if $\mu(E \triangledown F) = 0$. The relation $\sim$ is an equivalence relation on $\mathcal{A}(X)$. Define a function $\delta : \mathcal{A}(X) \times \mathcal{A}(X) \to \overline{\mathbb{R}}$ by

$$\delta(E, F) = \mu(E \triangledown F)$$

for every E, F in $\mathcal{A}(X)$. In general, the property

   (iii) $\delta(E, F) = 0$ implies $E = F$

fails; it fails for different sets whose symmetric difference (which is nonempty) is of measure zero (i.e., for distinct equivalent sets). On the other hand, the properties

(i) $\delta(E, F) = \delta(F, E)$,
(ii) $\delta(E, F) \geq 0$ and $\delta(E, E) = 0$,
(iv) $\delta(E, F) \leq \delta(E, G) + \delta(G, F)$,

hold for every $E, F, G \in \mathcal{A}(X)$ [48, p.138], [28, Problems 3-4.9] (also [41, Problem 6.5]). Note that $\delta(\varnothing, E) = \delta(E, \varnothing) = \mu(E)$ for every $E \in \mathcal{A}(X)$. If the measure μ is finite, then the function δ is real-valued. Thus assume from now on that the positive measure μ is finite (i.e., $0 < \mu(X) < \infty$). In this case δ is not only a distance function but is a pseudometric in $\mathcal{A}(X)$.

Remark 4.1. This pseudometric $\delta \colon \mathcal{A}(X) \times \mathcal{A}(X) \to \mathbb{R}$ is referred to as the Fréchet–Nikodým–Aronszajn distance [47, p.319]. In fact, it is sound and clear that the distance is introduced and proved to be a pseudometric in Nikodým's paper [48, Definition 4] as of 1930. As for Aronszajn's credit, Nikodým himself [48, p.134] mentions a work "to appear" which we were not able to trace back. As for Fréchet's own claim of priority, [48, p.134] refers to [23] and [24] as of 1921 and 1934, respectively (see also [49, Section 27.3.4]).

There are essentially two ways of normalizing this pseudometric: either set

$$\delta'(E, F) = \frac{\delta(E,F)}{\mu(X)} = \frac{\mu(E \nabla F)}{\mu(X)},$$

or (since $\{(E \cup F) \backslash (E \cap F), E \cap F\}$ is a partition of $E \cup F$ so that $\mu(E \cup F) = \mu(E \nabla F) + \mu(E \cap F)$), set

$$\delta''(E, F) = \begin{cases} \frac{\delta(E,F)}{\mu(E \cup F)} = \frac{\mu(E \nabla F)}{\mu(E \cup F)} = 1 - \frac{\mu(E \cap F)}{\mu(E \cup F)}, & \text{if } \mu(E \cup F) \neq 0, \\ 0, & \text{if } \mu(E \cup F) = 0, \end{cases}$$

for every sets E, F in $A(X)$. Both functions are bounded by 1 (i.e., $\delta'(E, F) \leq 1$ and $\delta''(E, F) \leq 1$ since $E \nabla F \subseteq E \cup F \subseteq X$). The function $\delta' \colon \mathcal{A}(X) \times \mathcal{A}(X) \to \mathbb{R}$ clearly is a pseudometric in $\mathcal{A}(X)$ (because δ is), and it was proved in [47, Section 1.2] that the function $\delta'' \colon \mathcal{A}(X) \times \mathcal{A}(X) \to \mathbb{R}$ is again a pseudometric in $\mathcal{A}(X)$, which is referred to as the Markzewisky–Steinhaus distance (see also [16, pp.46,175,299]).

It is worth noticing that a normalized real-valued function on $\mathcal{A}(X) \times \mathcal{A}(X)$ given by $\frac{\mu(E \cap F)}{\mu(E \cup F)}$ if $\mu(E \cup F) \neq 0$ is not even a distance function, since $1 = \frac{\mu(E \cap E)}{\mu(E \cup E)} \neq 0$ for every E in $\mathcal{A}(X)$ for which $\mu(E) \neq 0$ (i.e., it is only symmetric and nonnegative but it does not vanish at the identity line, so that property (ii) fails — as well as properties (iii) and (iv)).

The functions $\delta, \delta', \delta'' \colon \mathcal{A}(X) \times \mathcal{A}(X) \to \mathbb{R}$ are pseudometrics in $\mathcal{A}(X) \subseteq \wp(X)$ (nontrivial pseudometrics because μ is a nonzero measure), and so each of them induces a metric on the quotient space $\mathcal{A}(X)/\!\sim$ (of classes of equivalence of sets which differ from each other in the same class only by a set of measure zero: $E \sim F \iff \mu(E \nabla F) = 0$ — cf. Remark 2.1). However, if the finite measure μ is zero only at the empty set, then in this particular case $E \sim F$ if and only if $E = F$, and so $\delta, \delta', \delta''$ are themselves metrics in the σ-algebra $\mathcal{A}(X) \subseteq \wp(X)$.

Set $\mathcal{A}(X) = \wp(X)$ for a given nonempty set X and consider the counting measure $\nu \colon \wp(X) \to \overline{\mathbb{R}}$, which is defined by

$$\nu(E) = \begin{cases} \#E, & \text{if } E \text{ is finite,} \\ +\infty, & \text{if } E \text{ is infinite,} \end{cases}$$

for every set $E \in \wp(X)$. In this case, the only set of measure zero is the empty set. However, the counting measure is not finite if there exist infinite subsets of X. Thus suppose X is finite (i.e., $\#X < \infty$) so that $\nu \colon \wp(X) \to \mathbb{R}$ is given by

$$\nu(E) = \#E$$

for every subset E of the finite set X. Hence $\nu(X) = \#X < \infty$, and ν is a *finite counting measure* on $\wp(X)$, and so the function $\widetilde{\delta} \colon \wp(X) \times \wp(X) \to \mathbb{R}$ defined by $\widetilde{\delta}(E, F) = \nu(E \triangledown F)$, equivalently,

$$\widetilde{\delta}(E, F) = \#(E \triangledown F),$$

for every subsets E and F of the finite set X is indeed a metric in the power set $\wp(X)$. This function $\widetilde{\delta}$ is the finite version of the Fréchet–Nikodým–Aronszajn pseudometric δ (cf. Remark 4.1), now trivially a metric, whose natural normalizations yield the metrics $\widetilde{\delta}', \widetilde{\delta}'' \colon \wp(X) \times \wp(X) \to \mathbb{R}$,

$$\widetilde{\delta}'(E, F) = \frac{\#(E \triangledown F)}{\#X},$$

$$\widetilde{\delta}''(E, F) = \begin{cases} \frac{\#(E \triangledown F)}{\#(E \cup F)} = 1 - \frac{\#(E \cap F)}{\#(E \cup F)}, & \text{if } E \cup F \neq \varnothing, \\ 0, & \text{if } E \cup F = \varnothing, \end{cases}$$

for every $E, F \in \wp(X)$. The function $\widetilde{\delta}''$ is the finite version of the Markzewisky–Steinhaus pseudometric δ'' which, again, is a metric. This particular metric is sometimes called Jaccard distance (see, e.g., [25, eq.4], [26, eq.5], [16, p.299] — see also [43, p.34], [27, p.174]) after the botanist Paul Jaccard, which has been referred to as Tanimoto distance as well [16, pp.46,299], [45, p.263]. For more on metrics based on the cardinality of nonempty *finite* sets see [31].

Observe again that the normalized real-valued function on $\wp(X) \times \wp(X)$ given by $\frac{\#(E \cap F)}{\#(E \cup F)}$ if $E \cup F \neq \varnothing$ and 1 otherwise is not a distance function. In fact, $\frac{\#(E \cap E)}{\#(E \cup E)} \neq 0$ for every nonempty set E in $\wp(X)$, and properties (ii), (iii) and (iv) fail. For instance, consider the sets $E = \{e, f\}$, $F = \{f\}$, and $G = \{g\}$ for pairwise distinct points $e, f, g \in X$, so that $\frac{\#(E \cap F)}{\#(E \cup F)} \neq 0$, $\frac{\#(E \cap G)}{\#(E \cup G)} = \frac{\#(G \cap F)}{\#(G \cup F)} = 0$, and $\frac{\#(H \cap H)}{\#(H \cup H)} = 1$ for every set H in $\wp(X)$. Such a nondistance function has been refereed to as Jaccard similarity or Tanimoto similarity in [16, p.299], and as Jaccard index in [26, eq.4].

5. A Concise Review on Applications

A collection of contributions on applications, classified into three apparently distinct (but certainly not disjoint) classes, is considered in this section. Distance functions are supposed to act on appropriate domains. When we refer to a Hausdorff distance, it is understood that it acts on an admissible domain that makes it well-defined (e.g., on $\mathcal{B}(X) \setminus \varnothing \times \mathcal{B}(X) \setminus \varnothing$ or, in particular, on $\mathcal{F}(X) \setminus \varnothing \times \mathcal{F}(X) \setminus \varnothing$, which make the Hausdorff function into a metric). The distances and metrics of the families $h_{(s)}$ and $\delta_{(s)}$ discussed in Sections 3 and 4 will be freely referred throughout this section, where we will now proceed formally, omitting theoretical details.

The forthcoming reference list bears no claim of completeness. Perhaps a complete list (if this were possible), supporting a brief review, would become unacceptably large, leading to a dull catalog. The objective criterion for selecting the representatives in each class of the present review was mostly based on citations; the subjective one relies on the authors' taste.

Applications of the notion of distances between sets, in a variety of sensible definitions, have been commonly (and naturally) used in many areas of knowledge, ranging from theoretical to practical applications; for instance, from computer science to biological sciences. Roughly speaking, these applications aim at shape analysis in a wide sense (and so they encompass questions involving any sort of procedures towards sets distinction in general). We propose a simple (perhaps too simple) and rough classification of application areas into three classes.

5.1. Computational Aspects. All applications in this subsection deal with the Hausdorff family, most of them exclusively with the plain Hausdorff distance h.

5.1.1. *Distance in graphs*. Computational aspects, with emphasis in graph theory, for the Hausdorff metric h over finite sets and some of its variant distance functions as in Section 3, were considered by Eiter and Mannila [19] in 1997. These were compared both from theoretical and computational points of view, whose comparisons are specially tailored for applications to link distances [19, p.113] in graphs, some of them computed by polynomial time algorithms. In fact the Hausdorff metric is computable in polynomial time and, in spite of its appealing properties, it is shown that it may not be appropriate for some applications, since it does not take into account the entire configuration of some finite sets. A review on some distances in the Hausdorff family is presented together with a review on the literature up to then.

5.1.2. *Distance between polygons*. Polygons (as well as polyhedra) are characterized by the position of a finite number of points, in general in a finite-dimensional Euclidean space. An algorithm for computing the Hausdorff distance h between a pair of convex polygons was proposed by Atallah [2] in 1983. This was extended by Atallah, Ribeiro and Lifischtiz [3] in 1991, where algorithms for computing some Hausdorff-type distances of two possibly overlapping and not necessarily convex polygons was proposed. Algorithms for computing Hausdorff distance h for general polyhedra represented by triangular meshes was considered by Barton, Hanniel, Elber and Ki [6] in 2010, including a literature review regarding applications along this line — the reader is referred to the references therein. For another approach, using the Minimum Norm Duality Theorem (see, e.g. [46, p.136]) regarding the ordinary distance function d between convex sets in a normed space, see [15].

5.1.3. *Numerical procedures and algorithms*. The preceding subsection dealt with numerical aspects and algorithms as well, although this may not have been the main purpose there. Shonkwiler [59] considered in 1989 an algorithm for computing the $\tilde{h}$ distance between two images in linear time. Huttenlocher and Kedem [34] in 1990 and Huttenlocher, Kedem and Kleinberg [35] in 1992 computed translates of the Hausdorff distance h for subsets of the real line and of the Euclidian plane, where the results were also applied for comparing polygons under affine transformations, the main focus being on computational speed. Chew and Kedem [13] in 1998 proceeded along the same line, considering more options for the metric d on an n-dimensional ($n \leq 3$) real space X, such as the sup metric and the d_1 metric in addition to the

usual d_2 Euclidian metric. Numerical comparisons for estimating the Hausdorff distance h between discrete 3-dimensional surfaces represented by triangular meshes, aiming at the reduction of computational effort and memory usage, was considered by Aspert, Santa-Cruz and Ebrahimi [1] in 2002.

5.2. On Distances Between Fuzzy Sets. "There has been a number of papers proposing different extensions of the Hausdorff metric to fuzzy sets. None of these proposals behave as one would intuitively expect" [10].

5.2.1. *Marczewski–Steinhaus distance.* Following the Marczewski–Steinhaus metric $\widetilde{\delta}''$, Gardner, Kanno, Duncan, and Selmic [26] proposed in 2014 an extension of $\widetilde{\delta}''$, bounded by $\widetilde{\delta}''$ itself, which was shown to be a metric and suitable for applications in pattern recognition, image processing, machine learning, and information retrieval. As one would expect, $\widetilde{\delta}''$ would be too much of a metric to be used for fuzzy sets, and so further measure theoretical distances are also considered. Comparisons were implemented involving also the Hausdorff metric.

5.2.2. *Hausdorff distances.* Applications along this lines involving Hausdorff-like distance functions for fuzzy sets had been considered before by Rosenfeld [51] in 1985 and by Chaudhuri and Rosenfeld [11] in 1996, which were followed by Boxer [9] in 1997, by Fan [22] in 1998, and by Chaudhuri and Rosenfeld [12] in 1999. A different approach where the Hausdorff distance is used to generate further similarity measures for fuzzy sets was considered by Hung and Yang [33] in 2004.

5.2.3. *Non-Hausdorff distances.* A critical analysis on applications to fuzzy sets was undertaken by Brass [10] in 2002, where under an intriguing title "on the nonexistence of Hausdorff-like metrics for fuzzy sets" he set about to discuss plausible systems of metric axioms for fuzzy sets. Fujita [25] considered in 2013 some distance functions that can be applied to fuzzy sets, where the Hausdorff metric h and the Marczewski–Steinhaus finite-version metric $\widetilde{\delta}''$ are taken as starting points for yielding the original metric h_1'''.

5.3. Distance in Object Analysis. By "object analysis" we simply mean "image analysis" in a very broad sense, ranging from visible images to binary strings in general. The majority of applications of set distances focuses on problems inside this classification, including all ranges of applications for pattern recognition. There is a very large set of references (most on Hausdorff distance and its relatives) for image processing in such a broad sense. Also nonmetric distances (or nonmetric similarity functions), meaning semimetrics (where the triangle inequality may fail), have been considered for image analysis by Jacobs, Weinshall and Gdalyahu [38] in 2000. Actually, a whole book on visual recognition using Hausdorff distance by Rucklidge [54] has appeared in 1996, emphasizing computational aspects towards applications on imaging processing, including concrete experiments and a large list of references (which goes beyond set distances applied to image processing), to which the reader is referred. We comment on a shorter list (not included and not disjoint with the above-mentioned) with a cutoff roughly after considering some of the most cited articles (but not only) in order keep up with a reasonable-size list.

5.3.1. *New metrics and comparisons.* Baddeley [5] presents in 1992 a rather detailed discussion on the Hausdorff metric h pointing out that, although theoretically attractive, this is too sensitive a metric for image processing purposes, becoming

practically unstable. Thus the metric h'_p is introduced and compared for $p = 2$ with "error measures of current use", which is essentially the metric δ' (and some asymmetric variants of it), and comparisons involving classical synthetic images are also considered. As we have seen in 5.2.1 and 5.2.3, Fujita [25] in 2013, and also Gardner, Kanno, Duncan, and Selmic [26] in 2014, introduced original metrics as well, including discussions on the metrics h and $\widetilde{\delta}''$. Dubuisson and Jain [17] considered in 1994 most of the Hausdorff family of distances (and metrics) of Section 3 towards applications in image processing, comparing 24 combinations of them, including comparisons involving real images. They point out that h'_0 presents the best performance among their experiments.

5.3.2. *Motion – translation and rotation*. Rote [52] proposed in (1991) an algorithm for computing the minimum Hausdorff distance between subsets of the real line under translation. This was also considered by Li, Shen and Li [44] in 2008. Huttenlocher, Klanderman and Rucklidge [36], and Huttenlocher and Rucklidge [37], investigated in 1993 the Hausdorff distance h to evaluate nearness between a model set and an image set, including comparisons under translation and rigid motion (combined translation and rotation), and an algorithm is proposed to compute these distances with examples using real image. Rucklidge [53, 55] in 1995 and 1997 considered a procedure for searching space transformations of a model to match transformations that minimize the Hausdorff distance h between the transformed model and an image, including examples. Hossain, Dewan, Ahn and Chae, [32] also proposed in 2012 an algorithm for computing Hausdorff-like distances with application to moving objects. Also see 5.1.3.

5.3.3. *Modified Hausdorff including asymmetries*. Takács [61] considered in 1998 a procedure for face matching based on the Hausdorff family's distance h'_0. This includes a penalty scheme to ensure that images with large overlap are easily distinguished. Experimental results on a large set of face images are carried out. Sim, Kwon and Park [60] considered in 1999 an asymmetric function associated to the distance h'_1 for object matching, including simulations for comparisons based on synthetic and real images. Jesorsky, Kirchberg and Frischholz [39] also used in 2001 an asymmetric function associated to the distance h'_1, applied for shape comparisons towards face detection, where experiments were carried out with real images. Zhao, Shi and Deng [62] compared in 2005 asymmetric versions of h and h_1 for object matching in two-dimensional images, including experimental results.

References

1. N. Aspert, D. Santa-Cruz and T. Ebrahimi, *Mesh: measuring errors between surfaces using the Hausdorff distance*, Proc. IEEE Int. Conf. Multim. Expo. Lausanne (2002), 705-708.
2. M.J. Atallah, *A linear time algorithm for the Hausdorff distance between convex polygons*, Inform. Process. Lett. **17** (1983), 207–209.
3. M.J. Atallah, C.C. Ribeiro and S. Lifischtiz, *Computing some distances functions between polygons*, Pattern Recognition **24** (1991), 775-781.
4. H. Attouch, R. Lucchetti and R.J.-B. Wets, *The topology of the ρ-Hausdorff distance*, Ann. Mat. Pura Appl. (4) **CLX** (1991), 303–320.
5. A.J. Baddeley, *Errors in binary images and L^p version of the Hausdorff metric*, Nieuw Arch. Wisk. (4) **10** (1992), 157–183.
6. M. Barton, I. Hanniel, G. Elber and M.-S. Kim, *Precise Hausdorff distance computation between polygonal meshes*, Comput. Aided Geom. Design **27** (2010), 580-591.

7. T. Birsan and D. Tiba, *One hundred years since the introduction of the set distance by Dimitrie Pompeiu*, Proc. 22nd IFIP TC7 Conf., Turin (2005), 35–39.

8. L.M. Blumenthal, *Theory and Applications of Distance Geometry*, 2nd ed., Chelsea, New York, 1970.

9. L. Boxer, *On Hausdorff-like metrics for fuzzy sets*, Pattern Recognition Lett. **18** (1997), 115–118; Erratum: Pattern Recognition Lett. **18** (1997), 505–506.

10. P. Brass, *On the nonexistence of Hausdorff-like for fuzzy sets*, Pattern Recognition Lett. **23** (2002), 39–43.

11. B.B. Chaudhuri and A. Rosenfeld, *On a metric distance between fuzzy sets*. Pattern Recognition Lett. **17** (1996), 1157–1160.

12. B.B. Chaudhuri and A. Rosenfeld, *A modified Hausdorff distance between fuzzy sets*, Inform. Sci. **118** (1999), 159–171.

13. L.P. Chew and K. Kedem, *Getting around a lower bound for the minimum Hausdorff distance*, Comput. Geom. **10** (1998), 197–202.

14. A. Conci, S.L. Galvão, G.O. Sequeiros, D.C.M. Saade and T. MacHenry, *A new measure for comparing biomedical regions of interest in segmentation of digital images*, Discrete Appl. Math. **197** (2015), 103–113.

15. A. Dax, *The distance between two convex sets*, Linear Algebra Appl. **416** (2006), 184–213.

16. M.M. Deza and E. Deza, *Encyclopedia of Distances*, Springer, Berlin, 2009.

17. M.-P. Dubuisson and A.K. Jain, *A modified Hausdorff distance for object matching*, Proc. Int. Conf. Pattern Recognition, Jerusalem (1994), 566–568.

18. J. Dugundji, *Topology*, Allyn & Bacon, Boston, 1966.

19. T. Eiter and H. Mannila, *Distance measures for point sets and their computation*, Acta Inform. **34** (1997), 109–133.

20. K.J. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge, 1986.

21. K.J. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 3rd ed., Wiley, Chichester, 2014.

22. J.-L. Fan, *Note on Hausdorff-like metrics for fuzzy sets*, Pattern Recognition Lett. **19**, (1998), 793–796.

23. M. Fréchet, *Sur le diverses modes de convergence d'une suite de fonctions d'une variable*, Bull. Calcutta Math. Soc. **11** (1921), 187–306.

24. M. Fréchet, *Sur la distance de deux ensembles*, Bull. Calcutta Math. Soc. **15** (1924), 1–8.

25. O. Fujita, *Metrics based on average distance between sets*, Japan J. Indust. Appl, Math. **30** (2013), 1–19.

26. A. Gardner, J. Kanno, C.A. Duncan and R. Selmic, *Measuring distance between unordered sets of different sizes*, Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Columbus (2014), 137–143.

27. G. Gilbert, *Distance between sets*, Nature **239** (1972), 174.

28. P.R. Halmos, *Measure Theory*, Van Nostrand, New York, 1950.

29. F. Hausdorff, *Grundzüge der Mengenlehre*, Veit, Leipzig, 1914.

30. F. Hausdorff, *Set Theory*, 2nd ed., Chelsea, New York, 1962.

31. K.J. Horadam and M.A. Nyblom, *Distances between sets based on set commonality*, Discrete Appl. Math. **167** (2014), 310–314.

32. M.J. Hossain, M.A.A Dewan, K. Ahn and O. Chae, *A linear time algorithm of computing Hausdorff distance for content-based image analysis*, Circuits Systems Signal Process. **31** (2012), 389–399.

33. W.-L. Hung and M.-S. Yang, *Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance*, Pattern Recognition Lett. **25** (2004), 1603–1611.

34. D. Huttenlocher and K. Kedem, *Computing the minimum Hausdorff distance for point sets under translation*, Proc. 6th ACM Symp. Comp. Geometry, Berkeley (1990), 340–349.

35. 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*, Proc. 8th ACM Symp. Comp. Geometry, Berlin (1992), 110–119.

36. D.P. Huttenlocher, G.A. Klanderman and W.J. Rucklidge, *Comparing images using the Hausdorff distance*, IEEE Trans. Pattern Anal. Mach. Intell. **15** (1993), 850–863.

37. D.P. Huttenlocher and W.J. Rucklidge, *A multi-resolution technique for comparing images using the Hausdorff distance*, Proc. IEEE Conf. Comput. Vis. Pattern Recognit., New York (1993), 705–706.

38. D.W. Jacobs, D. Weinshall and Y. Gdalyahu, *Classification with nonmetric distances: image retrieval and class representation*, IEEE Trans. Pattern Anal. Mach. Intell. **22** (2000), 583–600.
39. O. Jesorsky, K.J. Kirchberg and R.W. Frischholz, *Robust face detection using the Hausdorff distance*, Proc. 3rd Int. Conf. AVBPA, Halmstad (2001), 90–95.
40. J.L. Kelley, *General Topology*, Van Nostrand, New York, 1955.
41. C.S. Kubrusly, *Measure Theory: A First Course*, Academic Press-Elsevier, San Diego, 2007.
42. C.S. Kubrusly, *The Elements of Operator Theory*, Birkhäuser-Springer, New York, 2011.
43. M. Levandowsky and D. Winter, *Distance between sets*, Nature **234** (1971), 34–35.
44. B. Li, Y. Shen and B. Li, *A new algorithm for computing the minimum Hausdorff distance between two point sets on a line under translation*, Inform. Process. Lett. **106** (2008), 52–58.
45. A.H. Lipkus, *A proof of the triangle inequality for the Tanimoto distance*, J. Math. Chem. **26** (1999), 263–265.
46. D.G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.
47. E. Marczewski and H. Steinhaus, *On a certain distance of sets and the corresponding distance of functions*, Colloq. Math. **6** (1958), 319–327.
48. O. Nikodym, *Sur une généralisation des intégrales de M.J. Radon*, Fund. Math. 15, (1930), 131–179.
49. J.-C. Pinoli, *Mathematical Foundations of Image Processing and Analysis* 2, Wiley, London, 2014.
50. R.T. Rockafeller and R.J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
51. A. Rosenfeld, *Distances between fuzzy sets*, Pattern Recognition Lett. **3** (1985), 229–233.
52. G. Rote, *Computing the minimum Hausdorff-distance between two point sets on a line under translation*, Inform. Process. Lett. **38** (1991), 123–127.
53. W.J. Rucklidge, *Locating objects using the Hausdorff distance*, Proc. 5th Int. Conf. on Comp. Vis., Cambridge-MA (1995), 457–464.
54. W. Rucklidge, *Efficient Visual Recognition Using Hausdorff Distance*, Springer, Berlin, 1996.
55. W.J. Rucklidge, *Efficiently locating objects using the Hausdorff distance*, Int. J. Comput. Vis. **24** (1997), 251–270.
56. B. Sendov, *Some questions of the theory of approximations of functions and sets in the Hausdorff metric*, Russ. Math. Surv. **24** (1969), 143–183.
57. B. Sendov, *Hausdorff Approximations*, Kluwer, Dordretcht, 1990.
58. B. Sendov, *Hausdorff distance and image processing*, Russ. Math. Surv. **59** (2004), 319–328.
59. R. Shonkwiler, *An image algorithm for computing the Hausdorff distance efficiently in linear time*, Inform. Process. Lett. **30** (1989), 87–89.
60. D.-G. Sim, O.-K. Kwon and R.-H. Park, *Object matching algorithms using robust Hausdorff distance measures*, IEEE Trans. Image Process. **8** (1999), 425–429.
61. B. Takács, *Comparing face images using the modified Hausdorff distance*, Pattern Recognition **31** (1998), 1873–1881.
62. C. Zhao, W. Shi and Y. Deng, *A new Hausdorff distance for image matching*, Pattern Recognition Lett. **26** (2005), 581–586.

Fluminense Federal University, Niteroi, Brazil
E-mail address: `aconci@ic.uff.br`

Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil
E-mail address: `carlos@ele.puc-rio.br`