

Partial Similarity of Objects, or How to Compare a Centaur to a Horse

Alexander M. Bronstein · Michael M. Bronstein · Alfred M. Bruckstein · Ron Kimmel

Received: 30 September 2007 / Accepted: 3 June 2008 / Published online: 26 July 2008
© Springer Science+Business Media, LLC 2008

Abstract Similarity is one of the most important abstract concepts in human perception of the world. In computer vision, numerous applications deal with comparing objects observed in a scene with some *a priori* known patterns. Often, it happens that while two objects are not similar, they have large similar parts, that is, they are partially similar. Here, we present a novel approach to quantify partial similarity using the notion of Pareto optimality. We exemplify our approach on the problems of recognizing non-rigid geometric objects, images, and analyzing text sequences.

Keywords Shape similarity · Partial similarity · Non-rigid shapes · Gromov-Hausdorff distance · Metric geometry · Deformation-invariant similarity · Correspondence · Levenshtein distance · Edit distance · Pareto optimality · Multicriterion optimization

1 Introduction

Similarity is one of the most important abstract concepts in human perception of the world. For example, we encounter it every day during our interaction with other people whose faces we recognize. Similarity also plays a crucial role in many fields in science. Attempts to understand self-similar or symmetric behavior of Nature led to many fundamental discoveries in physics (Greene 2000). In bioinformatics, a fundamental problem is detecting patterns similar to a sequence of nucleotides in given DNA sequences. In computer

vision, comparing objects observed in a scene with some *a priori* known patterns is a fundamental and largely open problem.

The definition of similarity is, to a large extent, a semantic question. Judging the similarity of faces, one may say that two human faces are similar if they have a common skin tone, while someone else would require the identity of the geometric structure of facial features like the eyes, the nose, and the mouth.

With a slight exaggeration, we can say that all pattern recognition problems boil down to giving a quantitative interpretation of similarity (or equivalently, *dissimilarity*) between objects (Bronstein et al. 2008b). Since there is no unique definition of similarity, every class of objects requires a specific, problem-dependent similarity criterion. Such criteria have been proposed for images (Stricker and Orengo 1995; Gudivada and Raghavan 1995; Chen and Wong 2003; Kim and Chung 2006), two-dimensional shapes (Bruckstein et al. 1992; Geiger et al. 1998, 2003; Latecki and Lakamper 2000; Cheng et al. 2001; Jacobs et al. 2000; Jacobs and Ling 2005; Felzenszwalb 2005; Bronstein et al. 2006e), three-dimensional rigid (Chen and Medioni 1991; Besl and McKay 1992; Zhang 1994; Tal et al. 2001) and non-rigid (Elad and Kimmel 2003; Mémoli and Sapiro 2005; Bronstein et al. 2006a, 2006d) shapes, text (Levenshtein 1965; Hatzivassiloglou et al. 1999), and audio (Foote 1997; Foote et al. 2002). In the face recognition community, extensive research has been done on similarities insensitive to illumination (Hallinan 1994; Bronstein et al. 2004), head pose (Gheorghiadis et al. 2001), and facial expressions (Bronstein et al. 2003, 2005).

In many situations, it happens that, while two objects are not similar, some of their parts are (Berchtold et al. 1997; Veltkamp 2001; Latecki et al. 2005; Boiman and Irani 2006).

A.M. Bronstein (✉) · M.M. Bronstein · A.M. Bruckstein · R. Kimmel
Department of Computer Science, Technion—Israel Institute of Technology, Haifa 32000, Israel
e-mail: bronstein@ieee.org

Such a situation is common, for example, in the face recognition application, where the quality of facial images (or surfaces in the case of 3D face recognition) can be degraded by acquisition imperfections, occlusions, and the presence of

Table 1 Notation and symbols

$\mathbb{R}; \mathbb{R}_+$	Real numbers; non-negative real numbers
\mathbb{R}^m	m -dimensional Euclidean space
\mathbb{R}_+^m	Non-negative m -dimensional orthant
\mathcal{C}	Class of objects
$X; X'; X'^c$	Object; part of object, subset of X ; complement of X
m_X	Fuzzy part of X , membership function
$\delta_{X'}$	Characteristic function of set X' ; Dirac delta function
τ_θ	Threshold function
$T_\theta m_X$	Thresholded fuzzy set
Σ_X	σ -algebra, set of parts of object X
$\Phi; \tilde{\Phi}$	Vector objective function; fuzzy objective function
(X^*, Y^*)	Pareto optimal parts
(m_X^*, m_Y^*)	Pareto optimal fuzzy parts
$\epsilon, \tilde{\epsilon}$	Dissimilarity; fuzzy dissimilarity
$\mu_X, \tilde{\mu}_X$	Measure on X ; fuzzy measure on X
$\lambda, \tilde{\lambda}$	Partiality; fuzzy partiality
$\text{dis}(\varphi, d)$	Distortion of a map φ
$d_X; d_X _{X'}$	Geodesic metric on surface X ; restricted metric
(X, d)	Metric space, non-rigid object
$\text{diam } X$	Diameter of X
B_r	Metric ball of radius r
X_N	Finite sampling of surface X
$T(X_N)$	Triangular mesh built upon vertices X_N
$\mu_X; \mathbf{m}_X$	Discretized measure; discretized membership function
$\text{lcs}(X, Y)$	Longest common subsequence of X and Y
$ X $	Cardinality of set X , length of sequence X
d_{GH}	Gromov-Hausdorff distance
d_{HAM}	Hamming distance
d_{E}	Edit (Levenshtein) distance
d_{MP}	Minimum partiality distance
$d_{\text{P}}; d_{\text{SP}}$	Set-valued Pareto distance; scalar Pareto distance

facial hair (Bronstein et al. 2006c). As an illustration that will help to understand the problem of partial similarity, we give an example from the realm of shape comparison. Figure 1 (inspired by Jacobs et al. 2000; see also Veltkamp 2001; Latecki et al. 2005) shows a centaur—a half-equine half-human mythological creature. From the point of view of traditional shape similarity, a centaur is similar neither to a horse nor to a man. However, large part of these shapes (the upper part of the human body and the bottom part of the horse body) are similar. Semantically, we can say that two object are *partially similar* if they have *large similar parts*. If one is able to detect such parts, the degree of partial similarity can be evaluated (Latecki et al. 2005).

The main purpose of this paper, stated briefly, is to provide a quantitative interpretation to what is meant by “similar” and “large”, and derive a consistent relation between these terms. It allows us to formulate a computationally-tractable problem of finding the largest most similar parts. In our approach, we use the formalism of Pareto optimality and multicriterion optimization. While well-known in fields like information theory and economics, these tools have been explored to a lesser extent in the computer vision and pattern recognition community (for some related concepts, see e.g. Oliveira et al. 2002; Everingham et al. 2002; Dunn et al. 2004).

A narrow setting of the discussed framework was previously presented in Bronstein et al. (2006e, 2008a) in relation to two-dimensional objects, and in Bronstein et al. (2007c) in relation to three-dimensional objects. In this paper, we introduce a more general formalism, which allows us extend the results to generic objects and address problems from other fields as well. We show particular examples of partial similarity of rigid and non-rigid two- and three-dimensional objects and text sequences, and further elaborate the numerical aspects of their computation. In addition, we show the extension of these methods to images and shapes with texture, and discuss an important particular case of partial self-similarity (symmetry) computation.

The paper is structured as follows. In Sect. 2, we give formal definitions of partiality and similarity and the neces-

Fig. 1 Is a centaur similar to a horse? A large part of the centaur is similar to a horse; likely, a large part of the centaur is similar to a human. However, considered as a whole, the centaur is similar neither to a horse, nor to a human

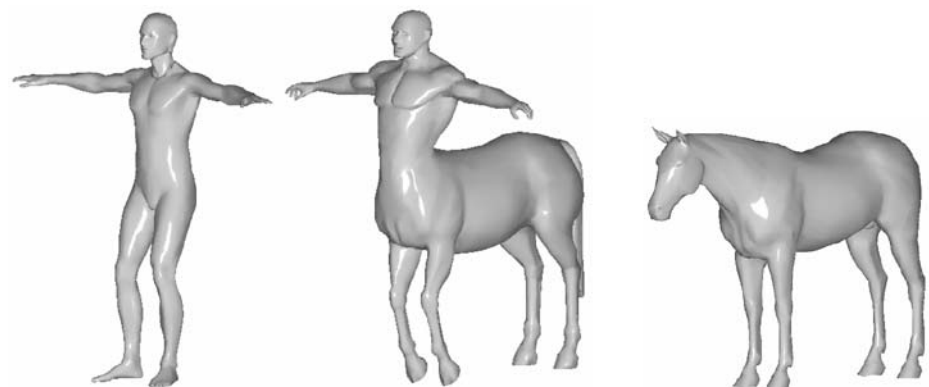


Fig. 2 Illustration of partial similarity problems of different classes of objects (left to right, top to bottom): non-rigid three-dimensional shapes, two-dimensional articulated shapes, images and text sequences



sary generic mathematical background. In Sect. 3, we formulate a multicriterion optimization problem, from which the relation between partiality and similarity is derived using the formalism of Pareto optimality. We represent partial similarity as a set-valued distance and study its properties. Then, we present a few case studies and applications of our partial similarity framework, showing how a specific use of the generic definitions from Sect. 2 can be used in different applications. All the application-specific mathematical background is defined in the beginning of each section.

In Sect. 4, we study the problem of analysis of two- and three-dimensional geometric objects. We show that all these apparently different objects can be modeled as metric spaces, and use the Gromov-Hausdorff distance as the criterion of their similarity. Practical numerical schemes for partial similarity computation are discussed in Sect. 4.6. In Sect. 5, we show how the partial similarity approach generalizes classical results in text sequences analysis. Section 6 is devoted to experimental results. In Sect. 7, we discuss a few extensions of the proposed methods. We address the problem of finding partial symmetries (self similarities) of shapes, which can be considered as a particular case of the partial similarity problem. Another possible generalization of the proposed framework is to textured shapes and two-dimensional images, which is also discussed. In addition, we address the problem of parts regularity and show a way to compute regularized partial similarity. Finally, Sect. 8 concludes the paper. Proofs of some results appear in Appendix.

2 Basic Definitions

In order to give a quantitative interpretation of our definition of partial similarity, we first have to define the terms “part”, “large” and “similar”. We start our construction by defining the class \mathcal{C} of objects we wish to compare: these may be, for instance, shapes, pictures, three-dimensional surfaces, audio and video sequences or words. An object in \mathcal{C} is a set, denoted by X . We assume that X can be decomposed into parts, where a part is modeled as a subset $X' \subseteq X$. The set of all the parts of an object X is described by a σ -algebra Σ_X on X (a subset of the powerset 2^X closed under complement and countable union). We demand that $\Sigma_X \subseteq \mathcal{C}$, or in other word, a part is also an object in \mathcal{C} .

Let us further assume to be given an equivalence relation \sim (a symmetric, reflexive and transitive relation) on the set of all parts of objects from class \mathcal{C} . Given two parts X' and Y' , we will say that they are similar if $X' \sim Y'$ (note that similarity does not necessarily imply that $X = Y$). Many objects have a natural definition of similarity. For example, two rigid shapes are similar if they are congruent, and two non-rigid shapes are similar if they are isometric.

However, two objects may be almost similar, in which case $X' \sim Y'$ does not hold anymore. In order to quantify how similar two objects are, we define a non-negative function $\epsilon : \mathcal{C} \times \mathcal{C} \rightarrow \mathbb{R}_+$, obeying the following properties,

- (D1) *Self-similarity*: $\epsilon(X, Y) = 0$ iff $X \sim Y$;
- (D2) *Symmetry*: $\epsilon(X, Y) = \epsilon(Y, X)$;
- (D3) *Triangle inequality*: $\epsilon(X, Y) + \epsilon(Y, Z) \geq \epsilon(X, Z)$;

for all objects X, Y and Z in \mathcal{C} . We call ϵ a *dissimilarity*, since the greater it is, the less similar are the objects. Property (D1) simply states that $\epsilon(X, Y) = 0$ is equivalent to X and Y being *similar*. Particularly, $\epsilon(X, X) = 0$, which implies that an object is similar to itself. Property (D2) requires similarity to be reflexive; and (D3) expresses the transitivity of similarity: if X is similar to Y , which is in turn similar to Z , then X and Z cannot be dissimilar. Technically speaking, ϵ is a *pseudo-metric* on \mathcal{C} , and a *metric* on the quotient space $\mathcal{C} \setminus \sim$. We will encounter some specific examples of dissimilarities in Sects. 4 and 5.

In order to quantify how *large* a part of an object X is, we define a function $\mu_X : \Sigma_X \rightarrow \mathbb{R}_+$, satisfying the following properties,

- (M1) *Additivity*: $\mu_X(X' \cup X'') = \mu_X(X') + \mu_X(X'')$ for two disjoint parts $X' \cap X'' = \emptyset$.
- (M2) *Monotonicity*: $\mu_X(X'') \leq \mu_X(X')$ for all $X'' \subseteq X' \in \Sigma_X$.

Such a μ_X is called a *measure*. In case of geometric objects, an intuitive example of a measure is the area of the object. Given a real function $u : X \rightarrow \mathbb{R}$, we say that it is Σ_X -*measurable* if $\{x \in X : u(x) \geq \alpha\} \in \Sigma_X$ for all α .

The size of the parts compared to the entire objects is measured using the *partiality* function,

$$\lambda(X', Y') = f(\mu_X(X'^c), \mu_Y(Y'^c)), \tag{1}$$

where $X'^c = X \setminus X'$ and $f : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ is a bivariate non-negative monotonous and symmetric function satisfying $f(0, 0) = 0$. Partiality quantifies how “small” are the parts X' and Y' (the larger is the partiality, the smaller are the parts) and satisfies the following properties,

- (P1) *Partiality of the whole*: $\lambda(X, Y) = 0$.
- (P2) *Symmetry*: $\lambda(X', Y') = \lambda(Y', X')$.
- (P3) *Partial order*: $\lambda(X'', Y'') \geq \lambda(X', Y')$ for every $X'' \subseteq X' \in \Sigma_X$ and $Y'' \subseteq Y' \in \Sigma_Y$.

Hereinafter, we will restrict the discussion to the following partiality function,

$$\lambda(X', Y') = \mu_X(X'^c) + \mu_Y(Y'^c), \tag{2}$$

though other definitions are possible as well.

2.1 Fuzzy Formulation

Anticipating the discussion in Sect. 4.6, we should say that the partial similarity computation problem requires optimization over subsets of X and Y , which in the discrete setting when the objects are represented as finite sets, gives rise to an NP-hard combinatorial problem. In order to cope with this complexity, we extend the above definitions using the fuzzy set theory (Zadeh 1965). This formulation is useful in

numerical computations, and as will be shown in the following sections, allows to relax the combinatorial problem and pose it as a continuous optimization problem.

We define a *fuzzy part* of X as a collection of pairs of the form $\{(x, m_X(x)) : x \in X\}$, where $m_X : X \rightarrow [0, 1]$ is referred to as a *membership function* and measures the degree of inclusion of a point into the subset. A fuzzy part of X is completely described by its membership function m_X ; hereinafter, we use m_X referring to fuzzy parts. A subset $X' \subseteq X$ in the traditional set theoretic sense (called *crisp* in fuzzy set theory) can be described by a membership function $\delta_{X'}(x)$, equal to one if $x \in X'$ and zero otherwise. More generally, a fuzzy part m_X can be converted into a crisp one by *thresholding*, $\tau_\theta \circ m_X$, where

$$\tau_\theta(x) = \begin{cases} 1 & x \geq \theta; \\ 0 & \text{otherwise;} \end{cases} \tag{3}$$

and $0 \leq \theta \leq 1$ is some constant. The corresponding crisp set is denoted by $T_\theta m_X = \{x \in X : \tau_\theta \circ m_X = 1\}$. Given a Σ_X , we define M_X as the set of all the fuzzy parts whose membership functions are Σ_X -measurable. It follows that $T_\theta m_X$ is in Σ_X for all $m_X \in M_X$ and $0 \leq \theta \leq 1$.

Hereinafter, as a notation convention, we will use tilde to denote fuzzy quantities. We define a *fuzzy dissimilarity* as a function of the form $\tilde{\epsilon}(m_X, m_Y)$ satisfying properties (D1)–(D3) with crisp parts replaced by fuzzy ones. We require $\tilde{\epsilon}$ to coincide with ϵ on crisp parts, or in other words, $\epsilon(X', Y') = \tilde{\epsilon}(\delta_{X'}, \delta_{Y'})$. The *fuzzy measure* is defined as

$$\tilde{\mu}_X(m_X) = \int_X m_X(x) d\mu_X, \tag{4}$$

for all $m_X \in M_X$, where μ_X is a (crisp) measure on X . We define the *fuzzy partiality* as

$$\tilde{\lambda}(m_X, m_Y) = \tilde{\mu}_X(m_X^c) + \tilde{\mu}_Y(m_Y^c), \tag{5}$$

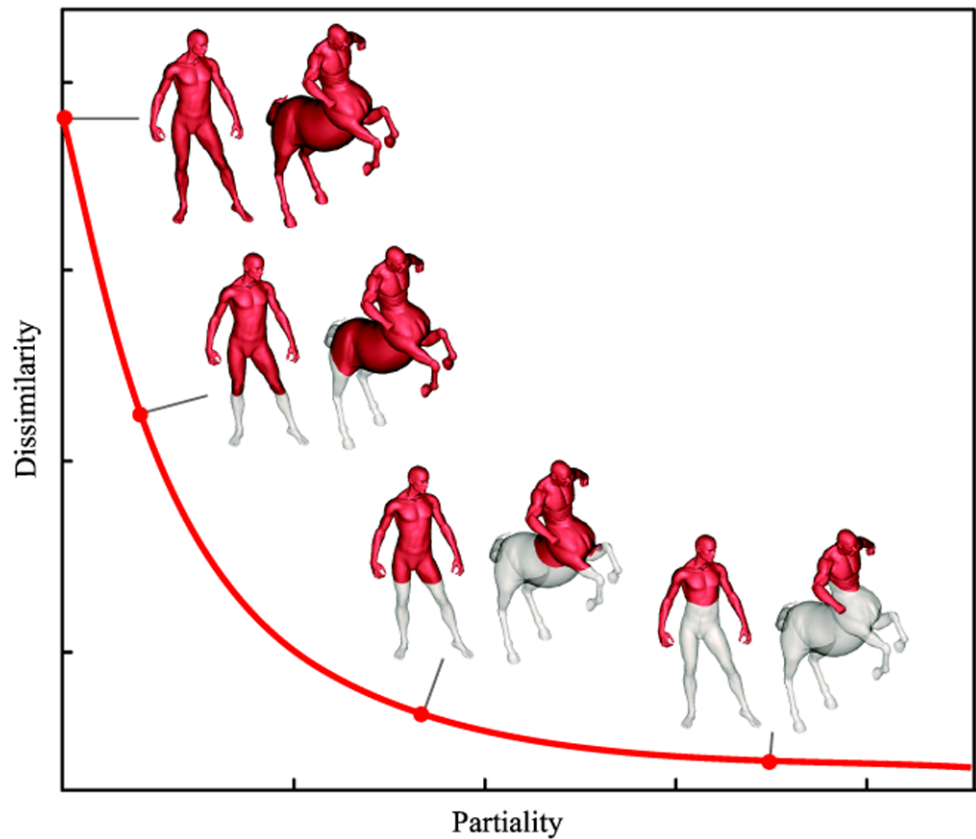
where $m_X^c = 1 - m_X$, similarly to definition (2). The following relation between the fuzzy and the crisp partialities holds,

Proposition 1 (i) $\lambda(X', Y') = \tilde{\lambda}(\delta_{X'}, \delta_{Y'})$; (ii) $\lambda(T_\theta m_X, T_\theta m_Y) \leq \frac{1}{1-\theta} \tilde{\lambda}(m_X, m_Y)$, for all $0 < \theta < 1$.

3 Pareto Framework for Partial Similarity

Using the definitions of Sect. 2, we can now give a quantitative expression to our definition of partial similarity: X and Y are *partially similar* if they have parts X' and Y' with small partiality $\lambda(X', Y')$ (“large”) and small dissimilarity $\epsilon(X', Y')$ (“similar”). We therefore formulate the computation of partial similarity as a *multicriterion optimization* problem: minimization of the *vector objective* function

Fig. 3 Illustration of the notion of Pareto optimality and set-valued distance



$\Phi(X', Y') = (\epsilon(X', Y'), \lambda(X', Y'))$ with respect to the pair (X', Y') over $\Sigma_X \times \Sigma_Y$,

$$\min_{(X', Y') \in \Sigma_X \times \Sigma_Y} \Phi(X', Y'). \tag{6}$$

The values of the criteria $\epsilon(X', Y')$ and $\lambda(X', Y')$ for every (X', Y') can be associated with a point with the coordinates $\Phi(X', Y')$. The set of possible criteria values is described by the region $\Phi(\Sigma_X \times \Sigma_Y)$ in \mathbb{R}^2 , referred to as the *attainable set*. The point $(0, 0)$ is usually not attainable, unless X and Y are fully similar (i.e., $X \sim Y$). For this reason, it is called the *utopia point*.

Since the two criteria are competing, no solution simultaneously optimal for both (i.e., the utopia point) can be found.¹ Thus, the notion of optimality used in traditional scalar optimization must be replaced by a new one, adapted to the multicriterion problem. Since there does not exist a *total order* relation in \mathbb{R}^2 , we generally cannot say which solution is better, for example: is the point $(0.5, 1)$ better than $(1, 0.5)$? Yet, we can introduce *partial order* by coordinate-wise comparison: $\Phi(X', Y')$ is better than $\Phi(X'', Y'')$ if both

$\lambda(X', Y') \leq \lambda(X'', Y'')$ and $\epsilon(X', Y') \leq \epsilon(X'', Y'')$, e.g., the point $(0.5, 0.5)$ is better than $(1, 1)$. By writing $\Phi(X', Y') \leq \Phi(X'', Y'')$, this partial order relation is implied hereinafter.

A solution (X^*, Y^*) is called a *Pareto optimum* (Pareto 1906; de Rooij and Vitanyi 2006; Everson and Fieldsend 2006) of the multicriterion optimization problem, if there exists no other pair of parts $(X', Y') \in \Sigma_X \times \Sigma_Y$ such that both $\epsilon(X', Y') < \epsilon(X^*, Y^*)$ and $\lambda(X', Y') < \lambda(X^*, Y^*)$ hold at the same time. An intuitive explanation of Pareto optimality is that no criterion can be improved without compromising the other. The set of all the Pareto optima, referred to as the *Pareto frontier* and can be visualized as a curve (see Fig. 3). We denote the Pareto frontier by $d_P(X, Y)$ and use it as a set-valued criterion of partial similarity, referred hereinafter as the *Pareto distance*.

When fuzzy quantities are used instead of the crisp ones, the multicriterion optimization problem is defined as the minimization of the vector objective $\tilde{\Phi} = (\tilde{\lambda}, \tilde{\epsilon})$ over the set $M_X \times M_Y$. A Pareto optimum is a point (m_X^*, m_Y^*) , for which there exists no other pair of fuzzy parts $(m_X, m_Y) \in M_X \times M_Y$ such that both $\tilde{\epsilon}(m_X, m_Y) < \tilde{\epsilon}(m_X^*, m_Y^*)$ and $\tilde{\lambda}(m_X, m_Y) < \tilde{\lambda}(m_X^*, m_Y^*)$ hold at the same time. The *fuzzy Pareto distance* $\tilde{d}_P(X, Y)$ is defined as the Pareto frontier, similarly to our previous crisp definition.

¹In information theory, such multicriterion optimization problems are widely known. For example, in statistical estimation, the bias and the variance of an estimator are two competing criteria. In lossy signal compression, distortion and bitrate are competing.

3.1 Scalar-Valued Partial Dissimilarity

Pareto distances are set-valued similarity criteria and should be compared as such. However, since there exists only a partial order relation between our criteria, not all Pareto distances are mutually comparable. In this sense, the notion of partial similarity is considerably different from the standard “full” similarity. We can say that X is more similar to Y than to Z (expressed as $d_P(X, Y) < d_P(X, Z)$, where a coordinate-wise inequality is implied) only if $d_P(X, Y)$ is entirely below $d_P(X, Z)$ (see Fig. 4). Otherwise, only point-wise comparison is possible: we write $(\lambda_0, \epsilon_0) < d_P(X, Y)$, implying that the point (λ_0, ϵ_0) is below $d_P(X, Y)$.

In order to define a total order between partial dissimilarities, we have to “scalarize” the multicriterion optimization problem. We refer to a distance obtained in this way a *scalar Pareto distance* and denote it by d_{SP} . One way to convert the set-valued distance into a scalar-valued one is by selecting a point on the Pareto frontier with a fixed value of partiality,

$$d_{SP}(X, Y) = \min_{(X', Y') \in \Sigma_X \times \Sigma_Y} \epsilon(X', Y')$$

$$\text{s.t. } \lambda(X', Y') \leq \lambda_0,$$

which can be alternatively formulated as an unconstrained problem,

$$d_{SP}(X, Y) = \min_{(X', Y') \in \Sigma_X \times \Sigma_Y} \epsilon(X', Y') + \beta \lambda(X', Y'), \quad (7)$$

where β is the corresponding Lagrange multiplier. Alternatively, we can fix a value of dissimilarity, obtaining the problem

$$d_{SP}(X, Y) = \min_{(X', Y') \in \Sigma_X \times \Sigma_Y} \lambda(X', Y')$$

$$\text{s.t. } \epsilon(X', Y') \leq \epsilon_0,$$

which can be posed equivalently to problem (7). A particular case of $\epsilon_0 = 0$ measures the minimum partiality required to achieve zero dissimilarity. We call such a distance *minimum partiality distance* and denote it by d_{MP} . We will encounter d_{MP} in Sect. 5 when discussing the similarity of text sequences.

The disadvantage of the described scalarization is that it is usually impossible to fix a single threshold value suitable for all the objects. A better and more generic way in which there is no need to fix arbitrary values of λ or ϵ is selecting a point on $d_P(X, Y)$ which is the closest, in the sense of some distance, to the utopia point $(0, 0)$. A Pareto optimum corresponding to such a point is called *Salukwadze optimal* (Salukwadze 1979). We define the scalar Pareto distance between X and Y as

$$d_{SP}(X, Y) = \inf_{(X', Y') \in \Sigma_X \times \Sigma_Y} \|\Phi(X', Y')\|_{\mathbb{R}_+^2}, \quad (8)$$

where $\|\cdot\|_{\mathbb{R}_+^2}$ denotes some norm on \mathbb{R}_+^2 . One example is the family of weighted norms $\|\Phi\|_{\mathbf{w}} = \Phi^T \mathbf{w}$ ($\mathbf{w} \in \mathbb{R}_+^2$). The particular case $\|\cdot\|_{(1,1)}$ coincides with the L_1 -norm. We refer to a d_{SP} constructed in this manner as the *Salukwadze distance*. It generalizes the previous ways of creating scalar-valued distances from the Pareto distance.

4 Geometric Shapes

Our first case study deals with geometric shapes. We attribute to this broad category rigid, articulated and non-rigid two- and three-dimensional objects, shapes with texture and, as a particular case, binary, gray and color images. While the analysis of two-dimensional (Platel et al. 2005) and three-dimensional rigid shapes (Chen and Medioni 1991; Besl and McKay 1992; Zhang 1994) is a well-established field, analysis of non-rigid shapes is an important direction emerging in the last decade in the pattern recognition community and arising in applications of face recognition (Bronstein et al. 2003, 2005), shape watermarking (Reuter et al. 2006), texture mapping and morphing (Bronstein et al. 2006b, 2007a), to mention a few. In many practical problems, it was shown that natural deformations of non-rigid shapes can be approximated as isometries, hence, recognition of such objects requires an isometry-invariant criterion of similarity. A particular case complying to this model are *articulated shapes*, consisting of rigid parts connected by non-rigid joints (Geiger et al. 1998; Jacobs et al. 2000; Elad and Kimmel 2003; Zhang et al. 2004; Ling and Jacobs 2005; Mémoli and Sapiro 2005; Bronstein et al. 2006e). Moreover, in many situations (e.g. in face recognition Bronstein et al. 2006c), due to acquisition imperfections, the objects may be given only partially, i.e., have similar overlapping parts. This makes our partial similarity framework especially useful in such applications.

In this section, following the spirit of (Elad and Kimmel 2003, Bronstein et al. 2006a, 2008a), we consider such objects as metric spaces. We show that such a metric approach provides a unifying framework which allows us to analyze the similarity of two- and three-dimensional shapes and images. We start with discussing similarity and self-similarity of objects, and then extend it using our partial similarity approach.

4.1 Intrinsic and Extrinsic Similarity

A geometric shape is modeled as a *metric space* (X, d) , where X is a two-dimensional smooth compact connected surface (possibly with boundary) embedded into \mathbb{R}^m ($m = 3$ in case of three-dimensional objects and $m = 2$ in case of two-dimensional shapes), and $d : X \times X \rightarrow \mathbb{R}_+$ is some *metric* measuring the distances on X . For the brevity of notation,

we will write shortly X instead of (X, d) when the metric d is implied or not important.

There are at least two natural ways to define the metric d on X . One way is to consider X as a subset of its embedding space \mathbb{R}^m and measure the distances between a pair of points x, x' on X using the *restricted Euclidean metric*,

$$d_{\mathbb{R}^m|_X}(x, x') = d_{\mathbb{R}^m}(x, x'). \tag{9}$$

The Euclidean metric regards the “external” properties of the shape, having to do with the way it is laid out in \mathbb{R}^m . We broadly refer to properties described by $d_{\mathbb{R}^m|_X}$ as the *extrinsic geometry* of X .

Another way is to define the distance between x and x' as the length of the shortest path (*geodesic*) on the surface X connecting x and x' . The length of the path can be computed as the sum of infinitesimally short line segments in the Euclidean space. We call the metric defined in this way the *geodesic metric* and denote it by d_X . Properties described by d_X are called the *intrinsic geometry* of X . Roughly speaking, intrinsic geometry describes the properties of the shape which are invariant to inelastic nonrigid deformations (i.e., deformations which do not “stretch” the surface), and extrinsic geometry is associated with a specific nonrigid deformation. The same shape can be regarded both from the intrinsic and extrinsic point of view by selecting d to be either the geodesic or the Euclidean metric, respectively (Bronstein et al. 2007b).

A transformation $\varphi : X \rightarrow \mathbb{R}^m$ preserving the extrinsic geometry of X is called a *congruence* and X and $\varphi(X)$ are said to be *congruent*. In the Euclidean space, congruences are limited to rigid motions (rotation and translation transformations);² we denote the family of such transformations by $\text{Iso}(\mathbb{R}^m)$. Two shapes X and Y are thus congruent if there exists a bijection $\varphi : X \rightarrow Y$ such that

$$d_{\mathbb{R}^m|_Y} = d_{\mathbb{R}^m|_X} \circ (\varphi \times \varphi). \tag{10}$$

In a similar way, a transformation $\varphi : X \rightarrow \mathbb{R}^m$ preserving the intrinsic geometry of X is called an *isometry* and X and $\varphi(X)$ are said to be *isometric*. Two shapes X and Y are isometric if there exists a bijection $\varphi : X \rightarrow Y$ such that

$$d_Y = d_X \circ (\varphi \times \varphi). \tag{11}$$

The class of isometries can be richer than that of congruences, since any congruence is by definition an isometry. However, for some objects these two classes coincide, meaning that they have no incongruent isometries. Such shapes are called *rigid*, and their extrinsic geometry is completely defined by the intrinsic one.

²Usually, *reflection* transformations are excluded since they have no physical realization.

Congruence is a natural similarity relation for rigid shapes. Congruent shapes have identical extrinsic geometry, or in other words, are the same shape up to a rigid motion. For this reason, we call the similarity relation defined by congruence *extrinsic similarity*. For non-rigid shapes, on the other hand, the natural similarity criterion is the equivalence of intrinsic geometry; two shapes are *intrinsically similar* if they are isometric.

It appears, with a few exceptions (Connelly 1978), that polyhedral surfaces, the most widely used representation of physical objects in geometric modeling, are rigid (Gluck 1974). However, even a rigid object can still have approximate isometries which are incongruent. To this end, we have to relax the requirement (11), making it hold only approximately. In order to measure to which extent (11) does not hold, we define the *intrinsic distortion*,

$$\text{dis}(\varphi, d_X) = \sup_{x, x' \in X} |d_X(x, x') - d_Y(\varphi(x), \varphi(x'))|,$$

of the map φ , and say that X and Y are ϵ -isometric if there exists an ϵ -surjective map $\varphi : X \rightarrow Y$ (i.e., $d_Y(y, \varphi(X)) \leq \epsilon$ for all $y \in Y$) with $\text{dis}(\varphi, d_X) \leq \epsilon$. Such a φ is called an ϵ -isometry (Burago et al. 2001).

For rigid shapes, appealing to the analogy between intrinsic and extrinsic similarity, we define the *extrinsic distortion*,

$$\text{dis}(\varphi, d_{\mathbb{R}^m|_X}) = \sup_{x, x' \in X} |d_{\mathbb{R}^m|_X}(x, x') - d_{\mathbb{R}^m|_Y}(\varphi(x), \varphi(x'))|,$$

where $\varphi \in \text{Iso}(\mathbb{R}^m)$ is a rigid motion. $\text{dis}(\varphi, d_{\mathbb{R}^m|_X})$ measures to which extent (10) does not hold, or in other words, the degree of incongruence between X and Y . We say that X and Y are ϵ -congruent if there exists an ϵ -surjective map $\varphi : X \rightarrow Y$ with $\text{dis}(\varphi, d_{\mathbb{R}^m|_X}) \leq \epsilon$.

4.2 Iterative Closest Point Algorithms

In order to measure how extrinsically dissimilar two shapes X and Y are, we need to find an ϵ -congruence between them with the smallest possible ϵ . A class of methods trying to solve this problem is called *iterative closest point* (ICP) algorithms (Chen and Medioni 1991; Besl and McKay 1992; Zhang 1994). Conceptually, these algorithms minimize the set-to-set distance between the shapes over all the rigid transformations,

$$d_{\text{ICP}}(X, Y) = \inf_{\varphi \in \text{Iso}(\mathbb{R}^m)} d_{\text{H}}^{\mathbb{R}^m}(X, \varphi(Y)), \tag{12}$$

where

$$d_{\text{H}}^{\mathbb{R}^m}(X, Y) = \max \left\{ \sup_{x \in X} d_{\mathbb{R}^m}(x, Y), \sup_{y \in Y} d_{\mathbb{R}^m}(y, X) \right\},$$

is the *Hausdorff distance* measuring how “far” the subsets X and Y of \mathbb{R}^m are from each other, and $d_{\mathbb{R}^m}(x, Y) =$

$\inf_{y \in Y} d_{\mathbb{R}^m}(x, y)$ is the point-to-set Euclidean distance. We call d_{ICP} the *ICP distance*. When X and Y are congruent, $d_{\text{ICP}}(X, Y) = 0$. When X and Y are almost congruent, the ICP distance can be considered as a measure of their incongruence.³

Assume that we can find a *correspondence* $\varphi : X \rightarrow Y$ mapping a point in x on X to the closest point,

$$\varphi(x) = \arg \min_{y \in Y} d_{\mathbb{R}^m}(x, y),$$

on Y , and, analogously,

$$\psi = \arg \min_{x \in X} d_{\mathbb{R}^m}(x, y),$$

to be the closest point on X .⁴ The Hausdorff distance (13) can be rewritten as

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} d_{\mathbb{R}^m}(x, \varphi(x)), \sup_{y \in Y} d_{\mathbb{R}^m}(y, \psi(y)) \right\}. \quad (13)$$

In practical applications, the Hausdorff distance is often replaced by an L_2 approximation,

$$d_H(X, Y) \approx \int_X d_{\mathbb{R}^3}^2(x, \varphi(x)) dx + \int_Y d_{\mathbb{R}^3}^2(y, \psi(y)) dy, \quad (14)$$

which is easier to compute.

Problem (12) can be solved using an iterative two-stage process:

Algorithm 1 (ICP algorithm)

- 1 **repeat**
 - 2 Fix the transformation (R, t) and find the closest-point correspondences φ and ψ between the surfaces X and Y .
 - 3 Fix the correspondences φ and ψ and find a rigid transformation minimizing the Hausdorff distance (13) or its approximation (14) between X and Y with the given correspondences.
 - 4 **until** convergence
-

4.3 Gromov-Hausdorff Distance

The extension of problem (12) for non-rigid shapes is not straightforward. For computing the extrinsic similarity of

³We do not explain formally the analogy between the definition of ϵ -congruence and the ICP distance.

⁴Such a correspondence exists for compact objects were are considering here. In the following, we tacitly assume that infima and suprema can be replaced by minima and maxima.

rigid shapes, it was possible to trivially apply the Hausdorff distance, since $(X, d_{\mathbb{R}^m}|_X)$ and $(Y, d_{\mathbb{R}^m}|_Y)$ were subsets of the same Euclidean space. Unlikely, two non-rigid shapes (X, d_X) and (Y, d_Y) are not parts of the same metric space. However, let us assume that there exists⁵ a metric space $(\mathbb{Z}, d_{\mathbb{Z}})$, into which (X, d_X) and (Y, d_Y) are *isometrically embeddable* by means of two mappings, g and h , respectively. We can now measure the Hausdorff distance in \mathbb{Z} between the images $g(X)$ and $h(Y)$. However, since the metric space \mathbb{Z} was chosen arbitrarily, we will try to find the best one which will minimize the Hausdorff distance between $g(X)$ and $h(Y)$ in \mathbb{Z} .

Using a similar motivation, Mikhail Gromov introduced in Gromov (1981) the *Gromov-Hausdorff distance*,

$$d_{\text{GH}}((X, d_X), (Y, d_Y)) = \inf_{\substack{g:(X, d_X) \rightarrow (\mathbb{Z}, d_{\mathbb{Z}}) \\ h:(Y, d_Y) \rightarrow (\mathbb{Z}, d_{\mathbb{Z}})}} d_{\text{H}}^{\mathbb{Z}}(g(X), h(Y)), \quad (15)$$

where g, h are isometric embeddings (i.e., g is an isometry between (X, d_X) and $(g(X), d_{\mathbb{Z}})$, and h is an isometry between (Y, d_Y) and $(h(Y), d_{\mathbb{Z}})$, respectively). d_{GH} can be regarded as a generalization of the Hausdorff distance: if the Hausdorff distance measures how far two subsets of a metric space are, the Gromov-Hausdorff distance measures how far two metric spaces are.

Particularly, if we used the Gromov-Hausdorff distance with the Euclidean metric, $d_{\text{GH}}((X, d_{\mathbb{R}^m}|_X), (Y, d_{\mathbb{R}^m}|_Y))$, we would obtain an analogous formulation of the ICP distance. The advantage of our formulation is the fact that it uses the same theoretical framework as the intrinsic similarity and boils down to computing the Gromov-Hausdorff distance between metric spaces with different metrics. As a result, the same numerical algorithms can be employed for both intrinsic and extrinsic similarity computation, which will be shown in the next sections.

A practical problem with definition (15) is that its computation involved optimization over a metric space \mathbb{Z} and is thus untractable. For compact surfaces (assumed here), the Gromov-Hausdorff distance has an equivalent formulation using distortion terms,

$$d_{\text{GH}}((X, d_X), (Y, d_Y)) = \frac{1}{2} \inf_{\substack{\varphi: X \rightarrow Y \\ \psi: Y \rightarrow X}} \max\{\text{dis}(\varphi, d_X), \text{dis}(\psi, d_Y), \text{dis}(\varphi, \psi, d_X, d_Y)\},$$

⁵Such a space always exists, the most trivial example being the disjoint union of (X, d_X) and (Y, d_Y) .

where,

$$\text{dis}(\varphi, \psi, d_X, d_Y) = \sup_{\substack{x \in X \\ y \in Y}} |d_X(x, \psi(y)) - d_Y(y, \varphi(x))|.$$

If X and Y are isometric, then $\psi = \varphi^{-1}$, and we have $\text{dis}(\varphi, d_X) = \text{dis}(\varphi, d_Y) = \text{dis}(\varphi, \psi, d_X, d_Y) = 0$. The converse is also true: $d_{\text{GH}}(X, Y) = 0$ if X and Y are isometric. In addition, d_{GH} is symmetric and satisfies the triangle inequality, which means that the Gromov-Hausdorff distance is a metric on the quotient space of objects under the isometry relation. More generally, if $d_{\text{GH}}(X, Y) \leq \epsilon$, then X and Y are 2ϵ -isometric and conversely, if X and Y are ϵ -isometric, then $d_{\text{GH}}(X, Y) \leq 2\epsilon$ (Burakov et al. 2001).

Like in ICP problems, for practical purposes, the Gromov-Hausdorff distance can be approximated in the following way,

$$\begin{aligned} d_{\text{GH}}(X, Y) &\approx \int_{X \times X} |d_X(x, x') - d_Y(\varphi(x), \varphi(x'))|^2 dx dx' \\ &+ \int_{Y \times Y} |d_Y(y, y') - d_X(\psi(y), \psi(y'))|^2 dy dy'. \end{aligned} \quad (16)$$

4.4 Partial Similarity of Geometric Shapes

The common denominator of two- and three-dimensional rigid and non-rigid shapes in our discussion is that they are modeled as metric spaces with an appropriately selected metric, and the criteria of similarity we are considering are between these metric spaces. Hence, from this point on we assume to be given a generic metric space, which can model any of the above objects, and will devise a way to compute partial similarity between metric spaces.

Given an object (X, d) , we define its part as $(X', d|_{X'})$. As the measure μ_X , we use the area of the object (derived from the metric structure of the surface in case of three-dimensional object, or the standard measure on \mathbb{R}^2 in case of two-dimensional objects and images). The partiality is thus interpreted as the portion of the area of the selected parts. As the dissimilarity in our framework, we use the Gromov-Hausdorff distance, wherein the metric is chosen according to the object in hand and the similarity criterion we are interested in (thus, we use d_X when comparing non-rigid objects, and $d_{\mathbb{R}^m}|_X$ when comparing rigid objects).

The Pareto distance $d_P(X, Y)$ measures the tradeoff between the dissimilarity (Gromov-Hausdorff distance) and the area cropped from the objects. The interpretation of the Pareto distance $d_P(X, Y)$ depends on the class to which the objects X, Y belong. In the case of non-rigid objects, in particular, $d_P(X, Y)$ tells us what is the size of the parts that must be removed in order to make X and Y isometric.

By properties of the Gromov-Hausdorff distance, $(\lambda, \epsilon) \in d_P(X, Y)$ implies that there exist $X' \in \Sigma_X$ and $Y' \in \Sigma_Y$ with partiality $\lambda(X', Y')$, such that $(X', d_X|_{X'})$ and $(Y', d_Y|_{Y'})$ are 2ϵ -isometric; and if $(X', d_X|_{X'})$ and $(Y', d_Y|_{Y'})$ are ϵ -isometric, then $(\lambda(X', Y'), 2\epsilon) \in d_P(X, Y)$. For rigid shapes, partial similarity describes the tradeoff between the congruence of the parts and their area.

The set-valued Pareto distance $d_P(X, Y)$ contains significantly more information about the similarity of non-rigid shapes X and Y than the scalar-valued Gromov-Hausdorff distance $d_{\text{GH}}(X, Y)$. In order to illustrate this difference, consider an example with non-rigid shapes shown in Fig. 4. When we compare the shape of a human to a centaur or another human with a spear using $d_{\text{GH}}((X, d_X), (Y, d_Y))$ (point (a) on the Pareto frontier in Fig. 4), we see that these objects are approximately equally dissimilar: the Gromov-Hausdorff distance operates with metric structure, and is thus sensitive to the length of the dissimilar parts (the spear and the bottom part of the horse body). However, if we start removing parts by increasing the partiality, we will see that the Pareto frontier describing the human–spear-bearer distance decreases fast, since the area of the spear is small, whereas the Pareto frontier describing the human–centaur distance decreases slowly, since the area of the horse body is large (Fig. 4b and c, respectively). This suggests that a human is more similar to a spear-bearer than to a centaur.

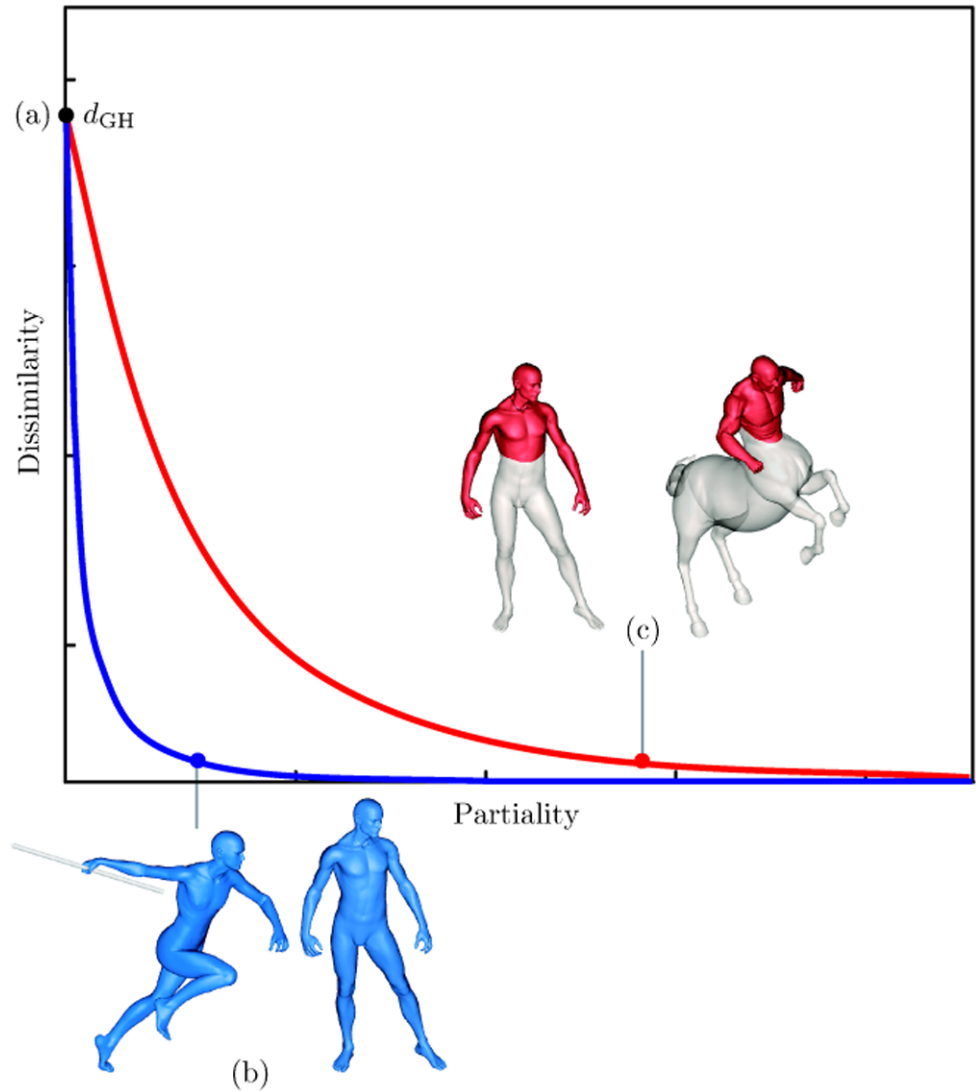
4.5 Fuzzy Approximation

The main problem with the presented approach is that it requires optimization over all the possible parts of the shapes, which, as we anticipatively mentioned in Sect. 2.1, becomes a combinatorial problem in the discrete setting. Therefore, to bring the problem back to the domain of continuous optimization we resort to the fuzzy formulation presented therein. In the fuzzy setting, this problem is avoided by describing the parts X' and Y' by membership functions m_X, m_Y , which obtain a continuous set of values.

The fuzzy partiality is defined according to (5). The fuzzy dissimilarity is a fuzzy version of the Gromov-Hausdorff distance,

$$\begin{aligned} \tilde{d}_{\text{GH}}(m_X, m_Y) &= \frac{1}{2} \inf_{\varphi: X \rightarrow Y} \max_{\psi: Y \rightarrow X} \left\{ \begin{array}{l} \sup_{x, x' \in X} m_X(x)m_X(x')|d_X(x, x') - d_Y(\varphi(x), \varphi(x'))| \\ \sup_{y, y' \in Y} m_Y(y)m_Y(y')|d_Y(y, y') - d_X(\psi(y), \psi(y'))| \\ \sup_{x \in X, y \in Y} m_X(x)m_Y(y)|d_X(x, \psi(y)) - d_Y(\varphi(x), y)| \\ D \sup_{x \in X} (1 - m_Y(\varphi(x)))m_X(x) \\ D \sup_{y \in Y} (1 - m_X(\psi(y)))m_Y(y) \end{array} \right\}, \end{aligned} \quad (17)$$

Fig. 4 Set-valued Pareto distance compared to traditional full similarity. A man and a spear-bearer are as dissimilar as a man and a centaur in the sense of intrinsic full similarity (Gromov-Hausdorff distance, (a)). In order to make a spear-bearer similar to a man, we have to remove a small part (spear, (b)). In order to make a centaur similar to a man, we have to remove the large horse body (c)



where $D \geq \max\{\text{diam}(X), \text{diam}(Y)\}$ and $\text{diam}(X) = \sup_{x, x' \in X} d_X(x, x')$.

Proposition 2 (i) $\tilde{d}_{GH}(\delta'_X, \delta'_Y) = d_{GH}(X', Y')$; (ii) Let $D = \max\{\text{diam } X, \text{diam } Y\} / \theta(1 - \theta)$, where $0 < \theta < 1$ is a parameter. Then, $d_{GH}(T_\theta m_X, T_\theta m_Y) \leq \frac{1}{\theta^2} \tilde{d}_{GH}(m_X, m_Y)$, for all $0 < \theta < 1$.

Combining the results of Propositions 1 and 2, we can connect the crisp and fuzzy partial dissimilarities,

$$\tilde{d}_P(X, Y) \leq (1 - \theta, \theta^{-2}) \cdot d_P(X, Y), \tag{18}$$

where \leq is understood as a partial order relation between vectors in \mathbb{R}^2 , defined in Sect. 3. This result allows us use $\tilde{d}_P(X, Y)$ as an approximation of $d_P(X, Y)$.

As we described in Sect. 3.1, a single point of the set-valued Pareto distance $\tilde{d}_P(X, Y)$ can be computed by fixing a value of partiality $\tilde{\lambda}(m_X, m_Y) \leq \lambda_0$ and minimizing

$\tilde{d}_{GH}(m_X, m_Y)$ with respect to m_X, m_Y subject to this constraint. One can notice that this problem involves two sets of variables: besides the fuzzy parts (membership functions m_X, m_Y), we have the correspondences between the parts (the maps φ and ψ). Optimization over this two sets of variables can be split into the following two-stage alternating scheme:

Algorithm 2 (Fuzzy Pareto distance computation)

- 1 **repeat**
 - 2 Fix the parts m_X and m_Y and find the correspondences φ and ψ .
 - 3 Fix the correspondences φ and ψ and find fuzzy parts m_X, m_Y minimizing the fuzzy Gromov-Hausdorff distance (18) with the given correspondences, subject to constraint $\tilde{\lambda}(m_X, m_Y) \leq \lambda_0$.
 - 4 **until convergence**
-

The analogy with ICP Algorithm 1 is evident. By varying the value of λ_0 , we obtain a set of values of the Pareto distance.

4.6 Numerical Computation

In practice, the computation of the Pareto distance is performed on discretized objects, $X_N = \{x_1, \dots, x_N\}$ and $Y_M = \{y_1, \dots, y_M\}$. The shapes are approximated as triangular meshes $T(X_N)$ and $T(Y_M)$ with vertices X_N and Y_M , respectively. A point on the mesh $T(X_N)$ is represented as a vector of the form $\mathbf{x} = (t, \mathbf{u})$, where t is the index of the triangular face enclosing it, and $\mathbf{u} \in [0, 1]^2$ is the vector of *barycentric coordinates* with respect to the vertices of that triangle.

The geodesic metrics d_X and d_Y are discretized as follows: First, the distances between the vertices $d_X(x_i, x_{i'})$ and $d_Y(y_j, y_{j'})$ (with $i, i' = 1, \dots, N$ and $j, j' = 1, \dots, M$) are numerically approximated using the *fast marching method* (FMM) (Sethian 1996; Kimmel and Sethian 1998). In order to compute the distance $d_X(\mathbf{x}, \mathbf{x}')$ between two arbitrary points \mathbf{x}, \mathbf{x}' on the mesh, interpolation based on the values of $d_X(x_i, x_{i'})$ is used. Here, we employ the *three-point interpolation* scheme (Bronstein et al. 2006d) for this purpose. A more computationally efficient approach can apply FMM “on demand”, using a software cache for geodesic distances.

The measure μ_X can be discretized by assigning to $\mu_X(x_i)$ the area of the *Voronoi cell* around x_i and represented as a vector $\boldsymbol{\mu}_X = (\mu_X(x_1), \dots, \mu_X(x_N))^T$. We use the following approximation,

$$\mu_X(x_i) \approx \frac{1}{3} \sum_{t \in \mathcal{N}(x_i)} a_t,$$

where $\mathcal{N}(x_i)$ denotes the one-ring neighborhood of triangles around the vertex x_i and a_t is the area of triangle t . The discretized membership functions are represented as vectors $\mathbf{m}_X = (m_X(x_1), \dots, m_X(x_N))$ and $\mathbf{m}_Y = (m_Y(y_1), \dots, m_Y(y_M))$.

The main challenge in Algorithm 2 is the computation of the correspondences φ and ψ , which is theoretically an NP-hard problem. Mémoli and Sapiro (2005) proposed a probabilistic approximation scheme for this problem. Bronstein et al. (2006a) introduced a different approach, based on a continuous non-convex optimization problem similar to multidimensional scaling (MDS), dubbed the *generalized MDS* (GMDS). It was shown that the correspondence computation can be formulated as three coupled GMDS problems, which can be solved efficiently using convex optimization (Bronstein et al. 2006d). The result is numerically accurate if global convergence is achieved.

The numerical solution we use here is similar to GMDS (Bronstein et al. 2006a, 2006d) and, in general, to the

spirit of MDS problems (Borg and Groenen 1997; Elad and Kimmel 2003). We express the correspondence as $y'_i = \varphi(x_i)$ and $x'_j = \psi(y_j)$ (note that in our notation, each x'_i is a point anywhere on $T(Y_M)$ and each y'_j is a point on $T(X_N)$). The points $\{y'_1, \dots, y'_N\}$ and $\{x'_1, \dots, x'_M\}$ are represented in barycentric coordinates as matrices \mathbf{Y}' and \mathbf{X}' .

Stage 2 of Algorithm 2 is done with the values of $\mathbf{m}_X, \mathbf{m}_Y$ fixed and minimizing over the correspondences \mathbf{X}' and \mathbf{Y}' ,

$$\begin{aligned} \min_{\substack{\epsilon \geq 0 \\ \mathbf{Y}' \\ \mathbf{X}'}} \epsilon \\ \text{s.t.} \quad & \begin{cases} m_X(x_i)m_X(x_{i'})|d_X(x_i, x_{i'}) - d_Y(\mathbf{y}'_i, \mathbf{y}'_{i'})| \leq \epsilon, \\ m_Y(y_j)m_Y(y_{j'})|d_Y(y_j, y_{j'}) - d_X(\mathbf{x}'_j, \mathbf{x}'_{j'})| \leq \epsilon, \\ m_X(x_i)m_Y(y_j)|d_X(x_i, \mathbf{x}'_j) - d_Y(y_j, \mathbf{y}'_i)| \leq \epsilon, \\ D(1 - m_X(\mathbf{x}'_i))m_X(x_i) \leq \epsilon, \\ D(1 - m_Y(\mathbf{y}'_j))m_Y(y_j) \leq \epsilon. \end{cases} \end{aligned} \tag{19}$$

The values $m_X(\mathbf{x}'_i)$ and $m_Y(\mathbf{y}'_j)$ at arbitrary points of the triangular mesh are computed by interpolation. The distances $d_X(x_i, x_{i'})$ and $d_Y(y_j, y_{j'})$ are pre-computed by FMM; on the other hand, the distances $d_Y(\mathbf{y}'_i, \mathbf{y}'_{i'})$, $d_X(\mathbf{x}'_j, \mathbf{x}'_{j'})$, $d_X(x_i, \mathbf{x}'_j)$ and $d_Y(y_j, \mathbf{y}'_i)$ are interpolated. Numerical solution of problem (19) requires the ability to perform a step in a given direction on a triangulated mesh (such a path is poly-linear if it traverses more than one triangle), computed using an *unfolding* procedure described in Bronstein et al. (2006d). This also ensures that barycentric representation of $\{y'_1, \dots, y'_N\}$ and $\{x'_1, \dots, x'_M\}$ is always correct.

Stage 3 of Algorithm 2 is performed by fixing \mathbf{X}' and \mathbf{Y}' and minimizing with respect to $\mathbf{m}_X, \mathbf{m}_Y$,

$$\begin{aligned} \min_{\substack{\epsilon \geq 0 \\ \mathbf{m}_X \in [0, 1]^N \\ \mathbf{m}_Y \in [0, 1]^M}} \epsilon \\ \text{s.t.} \quad & \begin{cases} m_X(x_i)m_X(x_{i'})|d_X(x_i, x_{i'}) - d_Y(\mathbf{y}'_i, \mathbf{y}'_{i'})| \leq \epsilon, \\ m_Y(y_j)m_Y(y_{j'})|d_Y(y_j, y_{j'}) - d_X(\mathbf{x}'_j, \mathbf{x}'_{j'})| \leq \epsilon, \\ m_X(x_i)m_Y(y_j)|d_X(x_i, \mathbf{x}'_j) - d_Y(y_j, \mathbf{y}'_i)| \leq \epsilon, \\ D(1 - m_X(\mathbf{x}'_i))m_X(x_i) \leq \epsilon, \\ D(1 - m_Y(\mathbf{y}'_j))m_Y(y_j) \leq \epsilon, \\ \mathbf{m}_X^T \boldsymbol{\mu}_X \geq 1 - \lambda, \\ \mathbf{m}_Y^T \boldsymbol{\mu}_Y \geq 1 - \lambda. \end{cases} \end{aligned} \tag{20}$$

If the L_2 approximation of the Gromov-Hausdorff distance is used, the distortion terms can be decoupled and problems (19) and (20) assume the form

$$\min_{\mathbf{X}'} \sum_{i,i'} m_X(x_i)m_X(x_{i'})|d_X(x_i, x_{i'}) - d_Y(\mathbf{y}'_i, \mathbf{y}'_{i'})|^2 + \min_{\mathbf{Y}'} \sum_{j,j'} m_Y(y_j)m_Y(y_{j'})|d_Y(y_j, y_{j'}) - d_X(\mathbf{x}'_j, \mathbf{x}'_{j'})|^2,$$

and

$$\min_{\mathbf{m}_X \in [0,1]^N} \sum_{i,i'} m_X(x_i)m_X(x_{i'})|d_X(x_i, x_{i'}) - d_Y(\mathbf{y}'_i, \mathbf{y}'_{i'})|^2$$

s.t. $\mathbf{m}_X^T \boldsymbol{\mu}_X \geq 1 - \lambda$

$$+ \min_{\mathbf{m}_Y \in [0,1]^M} \sum_{j,j'} m_Y(y_j)m_Y(y_{j'})|d_Y(y_j, y_{j'}) - d_X(\mathbf{x}'_j, \mathbf{x}'_{j'})|^2$$

s.t. $\mathbf{m}_Y^T \boldsymbol{\mu}_Y \geq 1 - \lambda,$

respectively.

Since the above problems are non-convex, optimization algorithms are liable to converge to a local minimum, a caveat widely known in MDS problems (Borg and Groenen 1997). Local convergence can be avoided in practice by using a multiresolution optimization scheme (Bronstein et al. 2006f), in which a hierarchy of problems is constructed, starting from a coarse version of the problem containing a small number of points. The coarse level solution is interpolated to the next resolution level, and is used as an initialization for the optimization at that level. The process is repeated until the finest level solution is obtained. Alternatively, an initialization similar to Gelfand et al. (2005) based on local shape descriptors and a branch-and-bound algorithm can be used (Raviv et al. 2007).

The main computational complexity of the algorithm is finding the correspondence. In our MATLAB implementation, performing GMDS with 50 points takes slightly less than a minute. Since the entire procedure is repeated for a few times in the alternating minimization scheme, computing the partial similarity between two shapes takes a few minutes. These results can be significantly improved by taking advantage of the fact that the correspondences do not change significantly from iteration to iteration, and thus performing full GMDS once followed by an incremental update would result in a much lower complexity.

5 Text Sequences

Another application of our partial similarity framework is the analysis of text sequences. Problems requiring comparison of such sequences arise in linguistics (Gooskens and

Heeringa 2004), web search (Giles et al. 1998; Di Lucca et al. 2002), spell checking (Damerau 1964), plagiarism detection (Wise 1996), speech recognition, and bioinformatics (Kruskal 1999; Bonhoeffer et al. 2004). The basic problem in this field is finding subsets of sequences that are similar to some give pattern—again, a problem fitting nicely into the concept of partial similarity.

The object used in text analysis is a *sequence* $X = (x_n)_{n=1}^N$. Each x_k (called *character*) is an element in some set A , referred to as the *alphabet*. For example, in text analysis A can be the Latin alphabet, and in bioinformatics, A is the set of four nucleotides. A part of a sequence X is a *subsequence* $X' = (x_{n_k})$, where n_k is a strictly increasing subset of the indices $\{1, \dots, N\}$. The σ -algebra Σ_X in this problem is defined as the collection of all the subsequences of X . A natural measure is the subsequence length, $\mu_X(X') = |X'|$.

Given two sequences X and Y , a *longest common subsequence* (LCS) of X and Y is defined as

$$\text{lcs}(X, Y) = \underset{Z \in \Sigma_X \cap \Sigma_Y}{\text{argmax}} |Z|. \tag{21}$$

Note that the LCS may not be unique; for example, the longest common subsequences of AATCC and ACACG are the sequences ACC and AAC.

An *edit* of the sequence X a modification inserting, removing or substituting one of the sequence characters. If X and Y are of equal length, we can define the *Hamming distance* between X and Y as the number of character substitutions required to transform one sequence into another,

$$d_{\text{HAM}}(X, Y) = \sum_{n=1}^{|X|} \delta_{x_n \neq y_n}. \tag{22}$$

For sequences of non-equal length, the Hamming distance can be extended by considering not only the substitution edits, but also character insertions and deletions. A classical tool in text analysis, known as the *edit* (or *Levenshtein distance* and denoted here by $d_E(X, Y)$, is defined as the minimum number of edits required to transform one string to another, where the edits are weighted differently (character deletion or insertion edits add 1 to the distance, and character substitution adds 2)⁶ (Levenshtein 1965; Wagner and Fischer 1974). The edit distance is related to the longest common subsequence by the following formula,

$$d_E(X, Y) = |X| + |Y| - 2|\text{lcs}(X, Y)|. \tag{23}$$

5.1 Partial Similarity of Text Sequences

To define the partial similarity between character sequences, we use d_{HAM} as the dissimilarity. If the subsequences are

⁶In some definitions, character substitution adds 1 to d_E .

not of equal length, ϵ is undefined. The partiality is defined as the total number of characters dropped from the sequences X and Y to obtain the two sub-sequences X' and Y' , $\lambda(X', Y') = |X| + |Y| - (|X'| + |Y'|)$. As the result of the tradeoff between $d_{HAM}(X', Y')$ and $\lambda(X', Y')$, a discrete Pareto frontier $d_P(X, Y)$ is obtained. If the value of $|X| + |Y|$ is even, $d_P(X, Y)$ exists only at even⁷ values of λ ; otherwise, it is defined only at odd values of λ .

We can establish the following relation between the zero-dissimilarity distance and the edit distance:

Proposition 3 (i) $d_{MP}(X, Y) = d_E(X, Y)$; (ii) $d_{MP}(X, Y)$ is realized on subsequences $X' = Y' = lcs(X, Y)$.

In other words, the edit distance is a particular case of our set-valued Pareto distance, obtained by selecting a specific point on the Pareto frontier, corresponding to the minimum partiality obtained requiring that d_{HAM} is zero. However, we may allow for subsequences which are not similar ($d_{HAM} > 0$), yet, have smaller partiality. This brings us to the definition of the Salukwadze distance $d_{SP}(X, Y)$, which may better quantify the partial similarity between two sequences.

⁷Subsequences X' and Y' must be of equal length in order for d_{HAM} to be defined, such that $|X'| + |Y'|$ is always even. If $|X| + |Y|$ is even, an odd value of $\lambda(X', Y')$ implies that X' and Y' are of unequal length and consequently, the Pareto frontier is not defined at this point.

6 Results

In order to exemplify the presented method, three experiments were performed, demonstrating the computation of partial similarity of articulated two-dimensional shapes, non-rigid three-dimensional shapes and text sequences. All the datasets used here are available from <http://tosca.cs.technion.ac.il>. For additional experiments with partial matching of rigid shapes, refer to Bronstein and Bronstein (2008).

6.1 Articulated Two-Dimensional Shapes

The first experiment was performed on the 2D Mythological Creatures database. The database consisted of three articulated shapes: human, horse and centaur, comprising rigid parts and non-rigid joints. Each shape appeared in five different articulations and with additional parts (sword, spear, tail and horns for the human shapes; saddle and wings for the horse shapes; sword, whip and spear for the centaur shapes, see Fig. 5). The shapes were represented as binary images and sub-sampled using the *farthest point* strategy (Hochbaum and Shmoys 1985; Bruckstein et al. 1998; Elad and Kimmel 2003) at approximately 3000 points. The shapes were discretized as described in Sect. 4.6. Thirteen values of λ were used to compute the Pareto distance.

Figure 6 shows the Pareto distances between the shapes. We can say that a human is more similar to a centaur than

Fig. 5 2D mythological creatures database used in the first experiment

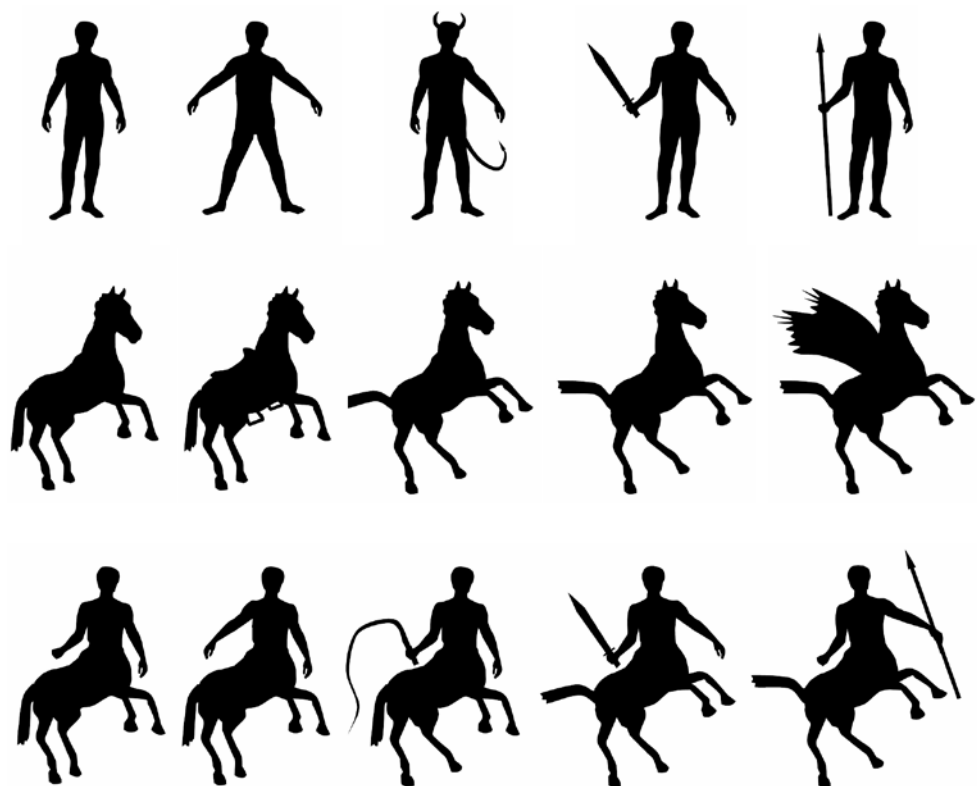


Fig. 6 Pareto distances between two-dimensional mythical creatures

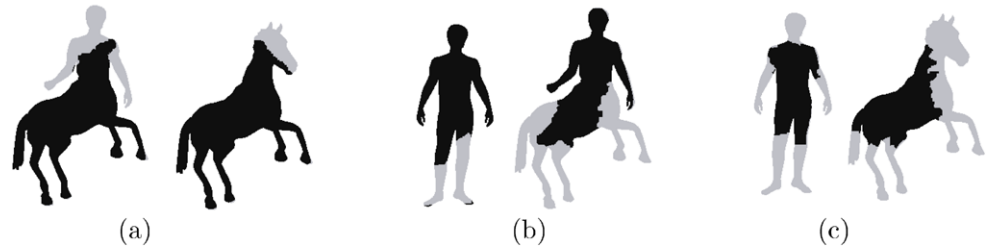
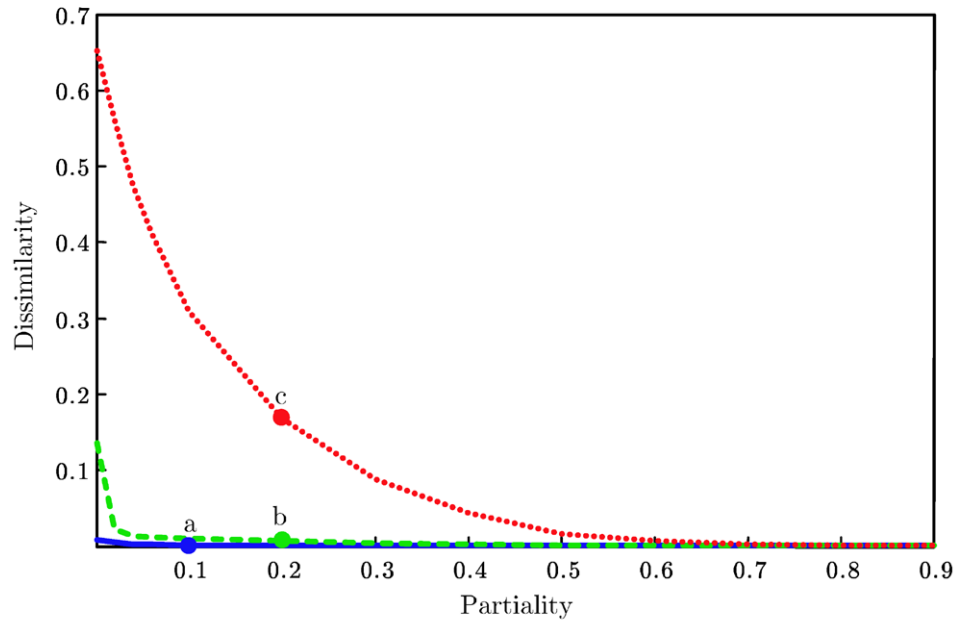
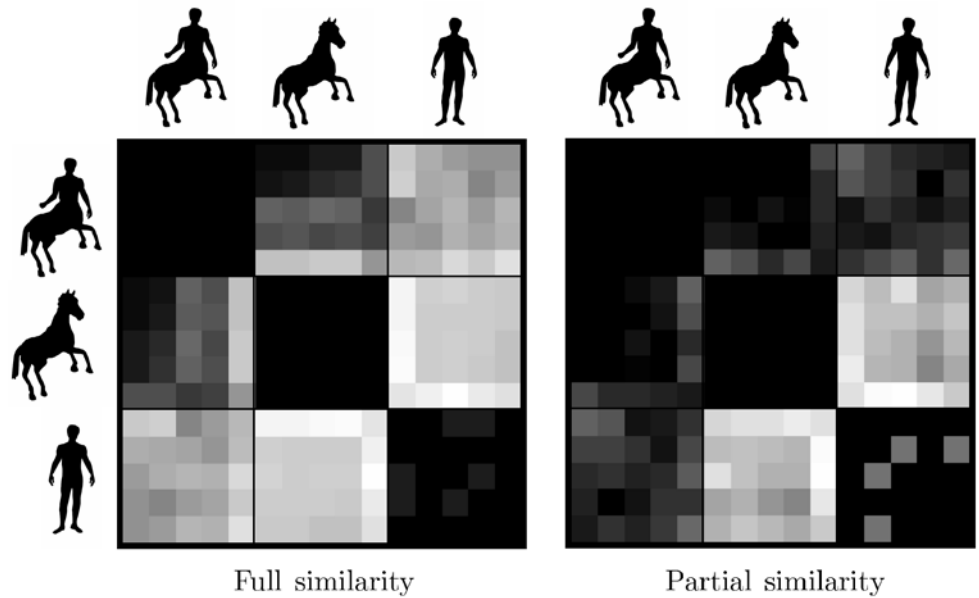


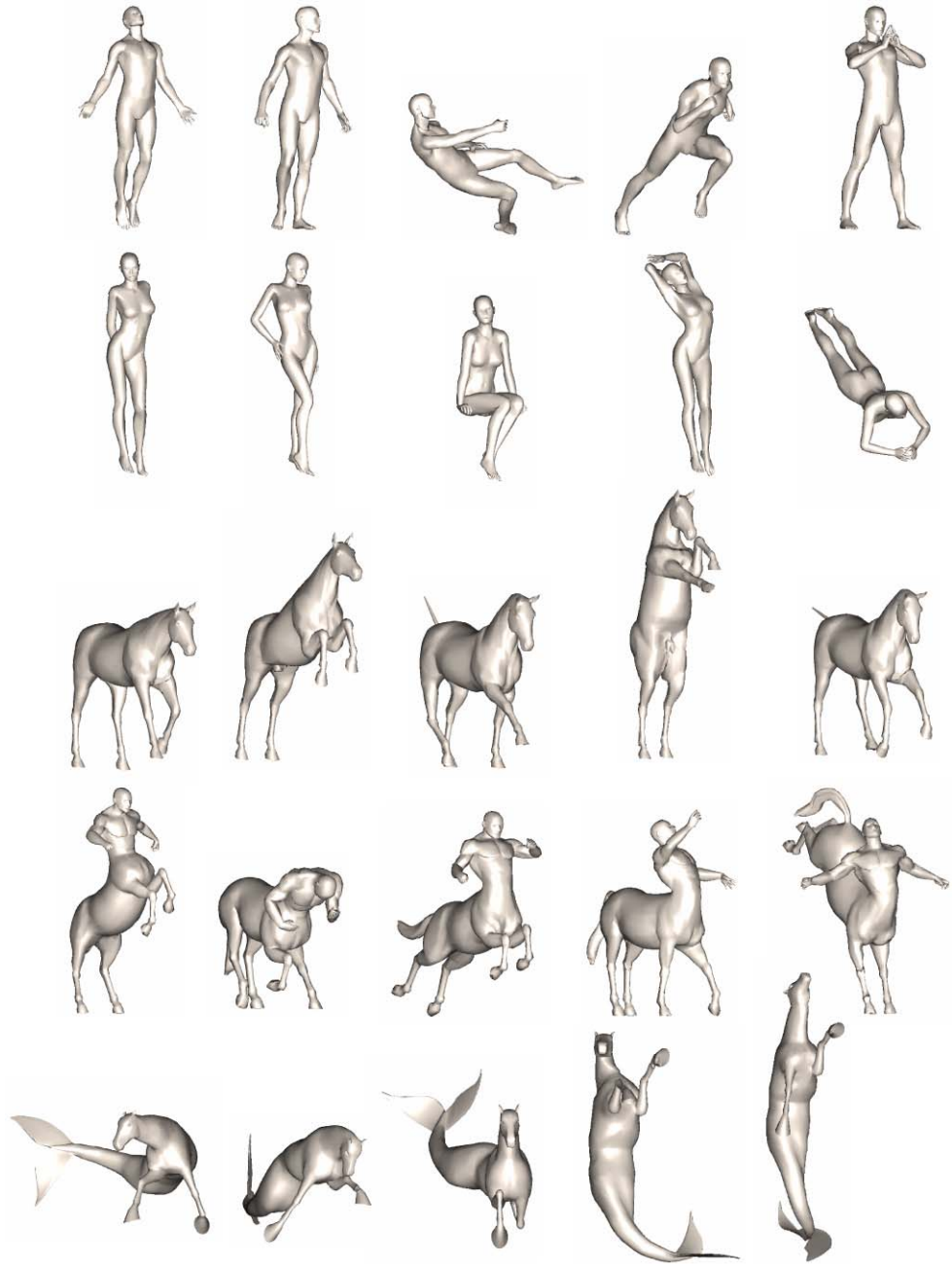
Fig. 7 Dissimilarity matrices representing d_{GH} (left) and \tilde{d}_{SP} (right) between two-dimensional mythical creatures



to a horse, because the Pareto frontier corresponding to the human–centaur comparison (dashed) is below that corresponding to the human–horse comparison (dotted). Figure 7 depicts the full intrinsic similarity (d_{GH}) and the scalar Pareto distance (\tilde{d}_{SP}) between the shapes as dissimilarity

matrices (the color of each element in the matrix represents the dissimilarity; the darker the smaller). In terms of the full intrinsic similarity, different articulations of the same shape are similar, while different shapes are dissimilar. This is observed as a pattern of dark diagonal blocks in Fig. 7 (left).

Fig. 8 3D nonrigid shapes database used in the second experiment



At the same time, the scalar Pareto distance is able to capture the partial similarity of the shapes, i.e., that the centaur is similar to the horse and the human. Additional examples of two-dimensional shape similarity are shown in Bronstein et al. (2008a).

6.2 Non-Rigid Three-Dimensional Shapes

The second experiment is a three-dimensional and a more challenging version of the first one. We used a subset of the nonrigid 3D shapes database, consisting of five objects: male, female, horse, centaur, and seahorse. Each shape ap-

peared in five different instances obtained by non-rigid deformations (Fig. 8). The shapes were represented as triangular meshes, sampled at between 1500 to 3000. We compared d_{GH} and \tilde{d}_{PS} . The distortion terms in the Gromov-Hausdorff distance were computed using 50 samples; the geodesic distances in the embedding spaces were interpolated from all the 1500 to 3000 samples.

The matching results are visualized in Fig. 9 as dissimilarity matrices. Being an intrinsic criterion of similarity, the Gromov-Hausdorff distance captures the intra-class similarity of shapes (i.e. that different instances of the same objects are similar). However, it fails to adequately capture the inter-

Fig. 9 Dissimilarity matrices representing d_{GH} (left) and \tilde{d}_{SP} (right) between three-dimensional mythological creatures

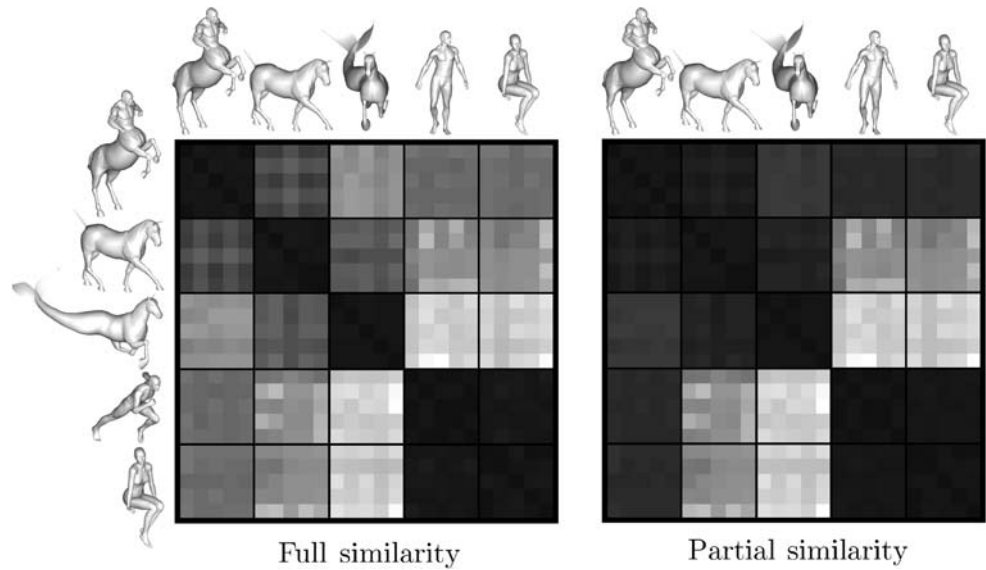
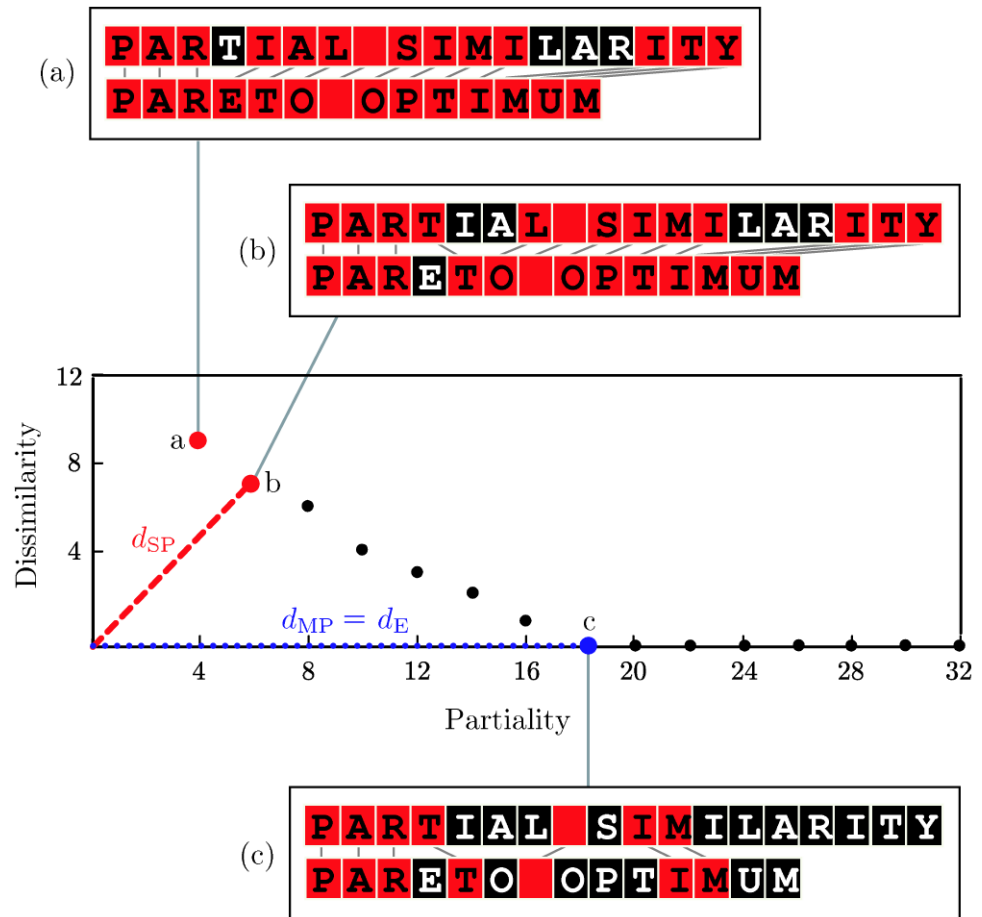


Fig. 10 (Color online) Comparison of sequences of characters. Red denotes the subsequences. Point (b) corresponds to the L_2 -Salukwadze distance (dashed). Point (c) is realized on the LCS. At this point, λ equals the value of d_E and d_{MP} (dotted)



class similarity: the centaur, horse and seahorse appear as dissimilar. On the other hand, the partial similarity approach captures correctly the partial similarity of the centaur, horse and the seahorse. For additional results and examples, see Bronstein et al. (2007c).

6.3 Text Sequences

To demonstrate the partial similarity concept in text analysis, we compare two sequences: $X = \text{PARTIAL SIMILARITY}$ and $Y = \text{PARETO OPTIMUM}$. The obtained discrete Pareto frontier is shown in Fig. 10. Point marked as (a) on the

Pareto frontier in Fig. 10 corresponds to the smallest Hamming distance with the smallest possible partiality ($\lambda = 4$). It is realized on subsequences $X' = \text{PARIAL SIMIITY}$ and $Y' = \text{PARETO OPTIMUM}$, the Hamming distance between which equals 9. Point (b) corresponds to the L_2 -Salukwadze distance ($d_{\text{SP}}(X, Y) = \|(6, 7)\|_2 = \sqrt{85}$). It is realized on subsequences PARTL SIMIITY and PARTO OPTIMUM (highlighted in red in Fig. 10b). Point (c) is the smallest value of partiality ($\lambda = 18$), for which d_{HAM} is zero, i.e., $d_{\text{MP}}(X, Y) = 18$. According to Proposition 3(ii), it is realized on a LCS, which in our example is $\text{lcs}(X, Y) = \text{PART IM}$ (highlighted in Fig. 10c). Using relation (23), it is easy to verify that $d_{\text{E}}(X, Y) = 18$ as well, which is an empirical evidence that Theorem 3(i) holds in this case.

7 Extensions

7.1 Self-Similarity and Symmetry

An important subset of the geometric similarity problem is *self-similarity*, usually referred to as *symmetry*. When we say that an object X is symmetric, we usually imply *extrinsic symmetries*, that is, *self-congruences* of X . The family of all the self-congruences of X , forms a group with the function composition, which can be referred to as the *extrinsic symmetry group*.

In practice, due to acquisition and representation inaccuracies, perfect symmetry rarely exists. Non-symmetric shapes have a trivial extrinsic symmetry group, containing only the identity mapping $\text{id}(x) = x$. However, while not symmetric in the strict sense, a shape can still be *approximately symmetric*. An intuitive way to understand the difference between the two definitions, is by thinking of a non-symmetric shape as obtained by applying a deformation to some other symmetric shape. Such a deformation may break the symmetries of the shape: if previously a symmetry was a self-congruence, we now have mappings which have non-zero extrinsic distortion. This leads to a simple way of quantifying the degree of extrinsic asymmetry of an object as

$$\text{asym}(X, d_{\mathbb{R}^m}|_X) = \inf_{\varphi: X \rightarrow X} \text{dis}(\varphi, d_{\mathbb{R}^m}|_X), \tag{24}$$

which resembles the Gromov-Hausdorff distance. Note, however, that in order to avoid a trivial solution, we require the mapping φ to be a local minimum of the distortion distinct from id .

While being adequate for rigid shapes, the traditional extrinsic notion of symmetry is inappropriate for non-rigid ones. Extrinsic symmetry can be broken as a result of isometric shape deformations, while its *intrinsic symmetry* is preserved. In Raviv et al. (2007) proposed using ϵ -isometries as a generalization of approximate symmetries

for non-rigid shapes. Using our notation, the degree of intrinsic symmetry of an object is quantified as

$$\text{asym}(X, d_X) = \inf_{\varphi: X \rightarrow X} \text{dis}(\varphi, d_X). \tag{25}$$

The problem of partial symmetry is a subset of the partial similarity problem since instead of two objects we have only one, which means that the optimization is performed only on one part, $X' \subseteq X$. This allows for a simpler formulation of the problem as follows. The partiality is simply the measure of the cropped parts, $\lambda(X') = \mu_X(X'^c)$. As the dissimilarity, we use the degree of asymmetry, $\epsilon(X') = \text{asym}(X', d_X|_{X'})$ in case of intrinsic symmetries and $\epsilon(X') = \text{asym}(X', d_{\mathbb{R}^m}|_{X'})$ in the case of the extrinsic ones. The computation of partial symmetry is formulated as the minimization of the vector objective function $\Phi(X) = (\epsilon(X'), \lambda(X'))$ with respect to X' over Σ_X . The part X^* is Pareto optimal if for any $X' \in \Sigma_X$, at least one of the following holds,

$$\begin{aligned} \epsilon(X^*) &\leq \epsilon(X'); \quad \text{or,} \\ \lambda(X^*) &\leq \lambda(X'). \end{aligned} \tag{26}$$

We denote the partial asymmetry by $\text{asym}_p(X)$. By writing $(\lambda, \epsilon) \in \text{asym}_p(X)$, we mean that there exists a part with partiality λ and asymmetry ϵ , such that any other part with smaller partiality have larger asymmetry, or any other part with smaller asymmetry has larger partiality.

7.2 Textured Shapes and Images

Similarity of geometric objects can be extended to textured shapes, in which the objects, in addition to their geometric properties also have some photometric properties. In computer graphics, this is typically modeled by attaching to the object X a texture $I : X \rightarrow C$, where C is some color space (e.g., \mathbb{R} in case of grayscale images and \mathbb{R}^3 in case of RGB images). The metric d_X for such objects can be defined in numerous ways, in general consisting of a *photometric* distance (measuring the dissimilarity between the color of the pixels), or *geometric* distance (i.e., and extrinsic or intrinsic metric discussed before), or a combination thereof. The similarity of textured shapes is posed as a problem of metric spaces comparison in the same way it is done for geometric objects using the Gromov-Hausdorff distance.

A particular case of textured objects are *images*. An image can be considered as a flat rectangular-shaped two-dimensional object with texture. Unlike the common representation of images in image processing community as a graph of function, here an image is modeled as a metric space (X, d) , where $X = [0, M] \times [0, N]$ is a rectangular region and d is a metric representing the similarity between different points in the image. We should note that

images were considered as geometric objects in previous works. For example, Sochen et al. (1998) suggested that images can be considered as Riemannian manifolds embedded into \mathbb{R}^3 (in case of grayscale images) or \mathbb{R}^5 (in case of color images), with the metric structure induced by the embedding. There are several fundamental differences between representing images as Riemannian manifolds and as generic metric spaces. First, being a more crude and less restrictive, a generic metric is not necessarily continuous, unlike the Riemannian one. This is important in images, where discontinuities (edges) play a crucial role. Second, unlike the Riemannian metric which is local, d is global (i.e., it can be used to measure the distance between any two points in the image). Finally, the metric d is arbitrary and not related to a specific embedding.

The choice of d is guided by the requirements posed on the similarity criterion. If one, for example, wishes the similarity between images to be rotation- and translation-invariant, the Euclidean distance $d_{\mathbb{R}^2}(x, x')$ is the easiest choice. If the similarity has to be scale invariant, some kind of photometric similarity is required, e.g., the pixel-wise photometric distance $\|I(x) - I(x')\|_2^2$, or a more general region-wise photometric distance, $(\int_{B_r} \|I(x + \xi) - I(x' + \xi)\|_2^2 d\xi)^{1/2}$, where B_r is a ball of radius r in \mathbb{R}^2 . Recent work in image processing (Tomasi and Manduchi 1998; Buades et al. 2005) suggested that these two distances can be combined, resulting in the following distance,

$$d(x, x') = d_{\mathbb{R}^2}(x, x') + \beta \left(\int_{B_r} \|I(x + \xi) - I(x' + \xi)\|_2^2 d\xi \right)^{1/2},$$

where β is some non-negative constant. Such a distance, referred to as *non-local means*, measures the dissimilarity between two pixels x, x' in the image as the sum of the distance between small patches around x and x' and the Euclidean distance between the locations of x and x' . It appears to be more robust to noise than point-wise comparison of intensities.

The disadvantage of combining the photometric and geometric distances into a single metric is the fact that an isometry in the sense of this metric does not have a clear interpretation. A better alternative is to separately define multiple metrics (for example, a photometric and a geometric one) and minimize a combination of the corresponding distortions.

7.3 Regularization

In our definition of partial similarity, we were interested in finding the largest most similar parts, without saying anything about their shape. This approach is prone to finding

multiple disconnected components, a behavior we sometimes observed is shape comparison. Avoiding this problem is possible by taking into consideration the *regularity* of parts.

There are two ways of doing it. First, some *irregularity* function $r(X')$ can be added as the third criterion into our vector-valued objective function. This new multicriterion optimization problem requires simultaneous minimization of dissimilarity, partiality and irregularity. The Pareto frontier in this case becomes a surface in \mathbb{R}^3 .

Alternatively, instead of partiality we can define *insignificance*, a more complicated criterion of part “quality”, which may include both a measure of the part size and regularity.

The most straightforward definition of irregularity is the length of the boundary $\partial X'$, which allows to define the insignificance of the part as

$$i(X') = \int_{X'} da + \eta \int_{\partial X'} d\ell,$$

where η is some parameter controlling the importance of the regularization term. In the fuzzy setting, the term $\int_{\partial X'} d\ell$ can be approximated using the Mumford-Shah approach (Mumford and Shah 1990). We refer the reader to Bronstein and Bronstein (2008) for additional details.

8 Conclusions

We presented a method for quantifying the partial similarity of objects, based on selecting parts of the objects with the optimal tradeoff between dissimilarity and partiality. We use the formalism of Pareto optimality to provide a definition to such a tradeoff. We demonstrated our approach on problems of analysis of geometric two- and three-dimensional rigid and non-rigid objects and text sequences. In all these problems, our construction has a meaningful interpretation. The set-valued distances resulting from it have appealing theoretical and practical properties. Particularly, in shape matching and text analysis, they can be viewed as a generalization of previous results. Since the presented framework of partial similarity is generic, it can be applied to other pattern recognition problems.

Appendix

Proof of Proposition 1 Part (i) follows trivially from the fact that $\tilde{\mu}_X(\delta_{X'}) = \mu_X(X')$ and $\tilde{\mu}_Y(\delta_{Y'}) = \mu_Y(Y')$. Part (ii): by Chebyshev inequality, we have

$$\begin{aligned} \mu_X((T_\theta m_X)^c) &= \mu_X(\{x \in X : 1 - m_X(x) \geq 1 - \theta\}) \\ &\leq \frac{1}{1 - \theta} \int_X (1 - m_X(x)) d\mu_X \\ &= \frac{1}{1 - \theta} \tilde{\mu}_X(m_X^c). \end{aligned} \quad (27)$$

The same holds for $\mu_Y((T_\theta m_Y)^c)$. Plugging these inequalities into the definition of fuzzy partiality, we have

$$\begin{aligned} \lambda(T_\theta m_X, T_\theta m_Y) &= \mu_X((T_\theta m_X)^c) + \mu_Y((T_\theta m_Y)^c) \\ &\leq \frac{1}{1-\theta} (\tilde{\mu}_X(m_X^c) + \tilde{\mu}_Y(m_Y^c)) \\ &= \frac{1}{1-\theta} \tilde{\lambda}(m_X, m_Y). \quad \square \end{aligned}$$

Proof of Proposition 2 For proof of (i), refer to Bronstein et al. (2008a). Part (ii):

$$\begin{aligned} &\tilde{d}_{GH}(m_X, m_Y) \\ &\geq \frac{1}{2} \inf_{\substack{\varphi: X \rightarrow Y \\ \psi: Y \rightarrow X}} \max \left\{ \begin{array}{l} \sup_{x, x' \in T_\theta m_X} \theta^2 |d_X(x, x') - d_Y(\varphi(x), \varphi(x'))| \\ \sup_{y, y' \in T_\theta m_Y} \theta^2 |d_Y(y, y') - d_X(\psi(y), \psi(y'))| \\ \sup_{x \in T_\theta m_X, y \in T_\theta m_Y} \theta^2 |d_X(x, \psi(y)) - d_Y(\varphi(x), y)| \\ \sup_{x \in T_\theta m_X} \theta D (1 - m_Y(\varphi(x))) \\ \sup_{y \in T_\theta m_Y} \theta D (1 - m_X(\psi(y))) \end{array} \right\}. \end{aligned}$$

From $\tilde{d}_{GH}(m_X, m_Y) \geq \theta D \sup_{x \in T_\theta m_X} (1 - m_X(\varphi(x)))$, and the fact that $\tilde{d}_{GH}(m_X, m_Y) \leq \frac{1}{2} \max\{\text{diam}(X), \text{diam}(Y)\}$, it follows that

$$\begin{aligned} \inf_{x \in T_\theta m_X} m_Y(\varphi(x)) &\geq 1 - \frac{2\tilde{d}_{GH}(m_X, m_Y)}{\theta D} \\ &\geq 1 - \frac{\max\{\text{diam}(X), \text{diam}(Y)\}}{\theta D} \geq \theta. \end{aligned}$$

Consequently, $\varphi(T_\theta m_X) \subseteq T_\theta m_Y$. In the same manner, $\psi(T_\theta m_Y) \subseteq T_\theta m_X$. Therefore,

$$\begin{aligned} &\tilde{d}_{GH}(m_X, m_Y) \\ &\geq \frac{\theta^2}{2} \inf_{\substack{\varphi: T_\theta m_X \rightarrow T_\theta m_Y \\ \psi: T_\theta m_Y \rightarrow T_\theta m_X}} \max \left\{ \begin{array}{l} \sup_{x, x' \in T_\theta m_X} |d_X(x, x') - d_Y(\varphi(x), \varphi(x'))| \\ \sup_{y, y' \in T_\theta m_Y} |d_Y(y, y') - d_X(\psi(y), \psi(y'))| \\ \sup_{x \in T_\theta m_X, y \in T_\theta m_Y} |d_X(x, \psi(y)) - d_Y(\varphi(x), y)| \end{array} \right\} \\ &= \theta^2 d_{GH}(T_\theta m_X, T_\theta m_Y). \quad \square \end{aligned}$$

Proof of Proposition 3 $d_{MP}(X, Y)$ corresponds to the longest subsequences X' and Y' that yield $d_{HAM}(X', Y') = 0$, which is, by definition, $X' = Y' = \text{lcs}(X, Y)$. By definition of partiality, $\lambda(X', Y') = |X'| + |Y'| - 2\text{lcs}(X, Y)$. Using the relation (23), we arrive at $d_{MP}(X, Y) = d_E(X, Y)$. \square

Acknowledgements This research was supported in part by the Center for Security Science and Technology. The authors are grateful to Alexander Brook and Irad Yavneh for valuable comments. This research was partly supported by United States–Israel Binational Science

Foundation grant No. 2004274 and by the Ministry of Science grant No. 3-3414, and in part by Elias Fund for Medical Research.

References

Berchtold, S., Keim, D. A., & Kriegel, H. P. (1997). Using extended feature objects for partial similarity retrieval. *International Journal on Very Large Data Bases*, 6(4), 333–348.

Besl, P. J., & McKay, N. D. (1992). A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239–256.

Boiman, O., & Irani, M. (2006). Similarity by composition. In *Proc. NIPS*.

Bonhoeffer, S., Chappey, C., Parkin, N. T., Whitcomb, J. M., & Petropoulos, C. J. (2004). Evidence for positive epistasis in HIV-1. *Science*, 306, 1547–1550.

Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling—theory and applications*. Berlin: Springer.

Bronstein, A. M., & Bronstein, M. M. (2008). *Partial matching of rigid shapes*. (Technical Report CIS-2008-02). Dept. of Computer Science, Technion, Israel.

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2003). Expression-invariant 3D face recognition. In *Lecture notes on computer science: Vol. 2688. Proc. audio and video-based biometric person authentication* (pp. 62–69). Berlin: Springer.

Bronstein, A. M., Bronstein, M. M., Gordon, E., & Kimmel, R. (2004). Fusion of 3D and 2D information in face recognition. In *Proc. ICIP* (pp. 87–90).

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2005). Three-dimensional face recognition. *IJCV*, 64(1), 5–30.

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2006a). Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *PNAS*, 103(5), 1168–1172.

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2006b). Face2face: an isometric model for facial animation. In *Proc. AMDO* (pp. 38–47).

Bronstein, A. M., Bronstein, A. M., & Kimmel, R. (2006c). Robust expression-invariant face recognition from partially missing data. In *Proc. ECCV* (pp. 396–408).

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2006d). Efficient computation of isometry-invariant distances between surfaces. *SIAM Journal on Scientific Computing*, 28(5), 1812–1836.

Bronstein, A. M., Bronstein, M. M., Bruckstein, A. M., & Kimmel, R. (2006e). Matching two-dimensional articulated shapes using generalized multidimensional scaling. In *Proc. AMDO* (pp. 48–57).

Bronstein, M. M., Bronstein, A. M., Kimmel, R., & Yavneh, I. (2006f). Multigrid multidimensional scaling. *Numerical Linear Algebra with Applications (NLAA)*, 13, 149–171.

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2007a). Calculus of non-rigid surfaces for geometry and texture manipulation. *IEEE Transactions on Visualization and Computer Graphics*, 13(5), 902–913.

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2007b). Rock, paper, and scissors: extrinsic vs. intrinsic similarity of non-rigid shapes. In *Proc. ICCV*.

Bronstein, M. M., Bronstein, A. M., Bruckstein, A. M., & Kimmel, R. (2007c). Paretian similarity for partial comparison of non-rigid objects. In F. Sgallari, A. Murli, & N. Paragios (Eds.), *Proc. scale space and variational methods in computer vision* (pp. 264–275). Berlin: Springer.

Bronstein, A. M., Bronstein, M. M., Bruckstein, A. M., & Kimmel, R. (2008a). Analysis of two-dimensional non-rigid shapes. *IJCV*, 78(1), 67–88.

- Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2008b). *Numerical geometry of nonrigid shapes*. Berlin: Springer.
- Bruckstein, A. M., Katzir, N., Lindenbaum, M., & Porat, M. (1992). Similarity-invariant signatures for partially occluded planar shapes. *IJCV*, 7(3), 271–285.
- Bruckstein, A. M., Holt, R. J., & Netravali, A. N. (1998). Holographic representations of images. *IEEE Transactions on Image Processing*, 7(11), 1583–1597.
- Buades, A., Coll, B., & Morel, J. M. (2005). A non-local algorithm for image denoising. *IEEE Conference on Computer Vision and Pattern Recognition*, 2, 60–65.
- Burago, D., Burago, Y., & Ivanov, S. (2001). *A course in metric geometry. Graduate studies in mathematics* (Vol. 33). Providence: American Mathematical Society.
- Chen, Y., & Medioni, G. (1991). Object modeling by registration of multiple range images. In *Proc. conf. robotics and automation*.
- Chen, Y., & Wong, E. K. (2003). Augmented image histogram for image and video similarity search. *Proceedings of SPIE*, 3656, 523.
- Cheng, S. W., Edelsbrunner, H., Fu, P., & Lam, K. P. (2001). Design and analysis of planar shape deformation. *Computational Geometry: Theory and Applications*, 19(2-3), 205–218.
- Connelly, R. (1978). A flexible sphere. *The Mathematical Intelligencer*, 1(3), 130–131.
- Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors.
- de Rooij, S., & Vitanyi, P. (2006). Approximating rate-distortion graphs of individual data: Experiments in lossy compression and denoising. *IEEE Transactions on Information Theory*, submitted.
- Di Lucca, G. A., Di Penta, M., & Fasolino, A. R. (2002). An approach to identify duplicated web pages. In *Proc. computer software and applications conference* (pp. 481–486).
- Dunn, E., Olague, G., Lutton, E., & Schoenauer, M. (2004). Pareto optimal sensing strategies for an active vision system. In *Proc. congress on evolutionary computation (CEC)*.
- Elad, A., & Kimmel, R. (2003). On bending invariant signatures for surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10), 1285–1295.
- Everingham, M., Muller, H., & Thomas, B. (2002). Evaluating image segmentation algorithms using the Pareto front. In *Proc. ECCV* (pp. 34–48).
- Everson, R. M., & Fieldsend, J. E. (2006). Multi-class ROC analysis from a multi-objective optimization perspective. *Pattern Recognition Letters*, 27(8), 918–927.
- Felzenszwalb, P. F. (2005). Representation and detection of deformable shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2), 208–220.
- Foote, J. (1997). Content-based retrieval of music and audio. *Proceedings of SPIE*, 3229, 138.
- Foote, J., Cooper, M., & Nam, U. (2002). Audio retrieval by rhythmic similarity. *Proc. International Conf. Music Information Retrieval*, 3, 265–266.
- Geiger, D., Basri, R., Costa, L., & Jacobs, D. (1998). Determining the similarity of deformable shapes. *Vision Research*, 38, 2365–2385.
- Geiger, D., Liu, T. L., & Kohn, R. (2003). Representation and self-similarity of shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1), 86–99.
- Gelfand, N., Mitra, N. J., Guibas, L., & Pottmann, H. (2005). Robust global registration. In *Proc. symp. geometry processing (SGP)*.
- Gheorghiadis, A. S., Belhumeur, P. N., & Kriegman, D. J. (2001). From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6), 643–660.
- Giles, C. L., Bollacker, K. D., & Lawrence, S. (1998). CiteSeer: an automatic citation indexing system. In *Proc. 3rd ACM conference on digital libraries* (pp. 89–98).
- Gluck, H. (1974). Almost all simply connected closed surfaces are rigid. In *Proc. conf. geometric topology*.
- Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change*, 16, 189–207.
- Greene, B. (2000). *The elegant universe*. New York: Vintage Books.
- Gromov, M. (1981). *Structures métriques pour les variétés riemanniennes*. Number 1 in Textes Mathématiques.
- Gudivada, V. N., & Raghavan, V. V. (1995). Content based image retrieval systems. *Computer*, 28(9), 18–22.
- Hallinan, P. (1994). A low-dimensional representation of human faces for arbitrary lighting conditions. In *Proc. CVPR* (pp. 995–999).
- Hatzivassiloglou, V., Klavans, J., & Eskin, E. (1999). Detecting text similarity over short passages: exploring linguistic feature combinations via machine learning. In *Proceedings of the joint SIGDAT conference on empirical methods in natural language processing and very large corpora*.
- Hochbaum, D., & Shmoys, D.B. (1985). A best possible heuristic for the k -center problem. *Mathematics of Operations Research*, 10(2), 180–184.
- Jacobs, D., Ling, H. (2005). Deformation invariant image matching. *Proc. ICCV*, 2, 719–726.
- Jacobs, D., Weinshall, D., & Gdalyahu, Y. (2000). Class representation and image retrieval with non-metric distances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(6), 583–600.
- Kim, C. R., & Chung, C. W. (2006). XMage: An image retrieval method based on partial similarity. *Information Processing and Management*, 42, 484–502.
- Kimmel, R., & Sethian, J. A. (1998). Computing geodesic on manifolds. *PNAS*, 95, 8431–8435.
- Kruskal, J. B. (1999). *An overview of sequence comparison*. Chapter: Time warps, string edits, and macromolecules: the theory and practice of sequence comparison. CSLI Publications.
- Latecki, L. J., Lakaemper, R., & Wolter, D. (2005). Optimal partial shape similarity. *Image and Vision Computing*, 23, 227–236.
- Latecki, L. J., & Lakamper, R. (2000). Shape similarity measure based on correspondence of visual parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10), 1185–1190.
- Levenshtein, V. I. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Doklady Akademii Nauk SSSR*, 163(4), 845–848.
- Ling, H., & Jacobs, D. (2005). Using the inner-distance for classification of articulated shapes. In *Proc. CVPR*.
- Mémoli, F., & Sapiro, G. (2005). A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 5(3), 313–347.
- Mumford, D., & Shah, J. (1990). Boundary detection by minimizing functionals. In *Image understanding*.
- Oliveira, L. S., Sabourin, R., Bortolozzi, F., & Suen, C. Y. (2002). Feature selection using multi-objective genetic algorithms for handwritten digit recognition. In *Proc. int'l conf. pattern recognition (ICPR)*.
- Pareto, V. (1906). *Manuale di economia politica*.
- Platel, B., Balmachnova, E., Florack, L. M. J., Kanters, F. M. W., & ter Haar Romeny, B. M. (2005). Using top-points as interest points for image matching. In *Lecture notes in computer science* (Vol. 3753, p. 211). Springer: Berlin.
- Raviv, D., Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2007). Symmetries of non-rigid shapes. In *Proc. workshop on non-rigid registration and tracking through learning (NRTL)*.
- Reuter, M., Wolter, F.-E., & Peinecke, N. (2006). Laplace-Beltrami spectra as shape-DNA of surfaces and solids. *Computer-Aided Design*, 38, 342–366.
- Salukwadze, M. E. (1979). *Vector-valued optimization problems in control theory*. San Diego: Academic Press.
- Sethian, J. A. (1996). A review of the theory, algorithms, and applications of level set method for propagating surfaces. *Acta Numerica*, 309–395.

- Sochen, N., Kimmel, R., & Malladi, R. (1998). A general framework for low level vision. *IEEE Transactions on Image Processing*, 7(3), 310–318.
- Stricker, M., & Orengo, M. (1995). Similarity of color images. *Proceedings of SPIE*, 2420, 381–392.
- Tal, A., Elad, M., & Ar, S. (2001). Content based retrieval of VRML objects—an iterative and interactive approach. In *Proc. eurographics workshop on multimedia*.
- Tomasi, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Proc. ICCV*.
- Veltkamp, R. C. (2001). Shape matching: similarity measures and algorithms. In *International conference on shape modeling and applications* (pp. 188–197).
- Wagner, R. A., & Fischer, M. J. (1974). The string-to-string correction problem. *JACM*, 21(1), 168–173.
- Wise, M. J. (1996). YAP3: improved detection of similarities in computer program and other texts. In *Proc. 27th SIGCSE technical symposium on computer science education* (pp. 130–134).
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.
- Zhang, Z. Y. (1994). Iterative point matching for registration of free-form curves and surfaces. *IJCV*, 13(2), 119–152.
- Zhang, J., Collins, R., & Liu, Y. (2004). Representation and matching of articulated shapes. In *Proc. CVPR* (Vol. 2, pp. 342–349).