Matching two-dimensional articulated shapes using generalized multidimensional scaling

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

Department of Computer Science,
Technion - Israel Institute of Technology, 32000 Haifa, Israel
{alexbron, bronstein}@ieee.org, {freddy, ron}@cs.technion.ac.il

Abstract. We present a theoretical and computational framework for matching of two-dimensional articulated shapes. Assuming that articulations can be modeled as near-isometries, we show an axiomatic construction of an articulation-invariant distance between shapes, formulated as a generalized multidimensional scaling (GMDS) problem and solved efficiently. Some numerical results demonstrating the accuracy of our method are presented.

1 Introduction

Recognition of two-dimensional shapes (silhouettes) is an important problem with a wide range of applications, extensively addressed in computer vision literature (see e.g. [1–3]). One of the main difficulties in shape recognition arises from the fact that natural objects are non-rigid. A simplified model capturing to some degree this flexibility is the *articulated shape* model, assuming that the object is composed of rigid parts, each of which has a certain freedom to move. Such a model appears to be applicable to many objects in nature, for example, humans, animals, tools, etc [4].

Recently, Ling and Jacobs [5] proposed to use the inner (geodesic) distances for recognition of articulated shapes. The main claim is that the geodesic distances are insensitive to articulations and therefore can be used as robust descriptors of the shape. This approach is related to previous works of Elad and Kimmel on bending-invariant representations of 3D objects [6], in which multidimensional scaling (MDS) was applied to the geodesic distances measured on the shape in order to obtain its intrinsic-geometric representation.

Our current paper is strongly motivated by the study of Ling and Jacobs. Using the model presented in [5], we describe articulations as isometric (distance-preserving) transformations of the shape. The main contribution of this paper is an axiomatic construction of a distance that allows to discern between geometrically different articulated shapes while being articulation-invariant. Our distance is free of error introduced by approaches based on Euclidean MDS [6] and also allows matching of partially occluded shapes. The computation of our distance is formulated as a generalized MDS problem (GMDS) and can be solved efficiently.
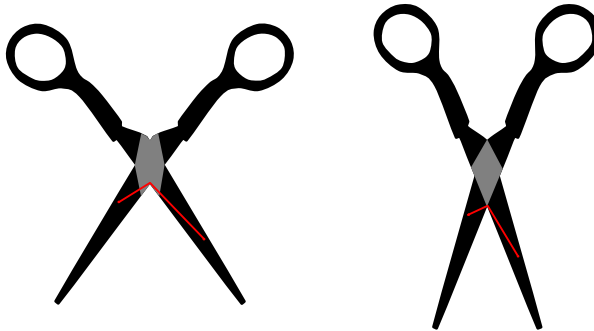
Fig. 1. Example of an articulated shape, consisting of four parts (black) and one joint (gray). The geodesic distance between two points is shown in red. Note that the geodesic distances change is bounded by the diameter of the joint.

This paper consists of five sections. In Section 2, we present the isometric model of articulated shapes and our articulation-invariant distance. Section 3 deals with numerical computation of the distance between articulated shapes using the GMDS. In Section 4, we present an experimental validation of our approach. Section 5 concludes the paper.

2 Isometric model for articulated shapes

Let $\mathcal{S}$ be a shape, represented as a compact, connected, flat two-dimensional manifold with boundary. The metric on $\mathcal{S}$ is assumed to be Euclidean. Following Ling and Jacobs [5], we represent $\mathcal{S}$ as a union of K disjoint *parts* $\mathcal{S}_1, ..., \mathcal{S}_K$ and L *joints* $\mathcal{J}_1, ..., \mathcal{J}_L$ (Figure 1). We call such $\mathcal{S}$ an *articulated shape*. The minimal geodesics (shortest paths) on $\mathcal{S}$ consist of linear segments and portions of the boundary [5]. The geodesic distances between two points $s_1, s_2 \in \mathcal{S}$ are denoted by $d_{\mathcal{S}}(s_1, s_2)$. An articulated shape with $\sum_{i=1}^{L} \mathrm{diam} \mathcal{J}_i \leq \epsilon$ is called an ϵ-*articulated shape*. We denote by $\mathbb{M}_\epsilon$ the space of all ϵ-articulated shapes; $\mathbb{M}$ denotes $\mathbb{M}_\infty$.

An *articulation* is a mapping $f : \mathcal{S} \to \mathcal{S}' \subset \mathbb{R}^2$, which transforms each part $\mathcal{S}_i$ in a rigid manner and preserves the topology of the whole shape, such that different parts remain disjoint. For an ϵ-articulated shape, articulations are ϵ-*isometries*, i.e., have *distortion*

$$\mathrm{dis}\, f \equiv \sup_{s_1, s_2 \in \mathcal{S}} |d_{\mathcal{S}}(s_1, s_2) - d_{\mathcal{Q}}(f(s_1), f(s_2))| \leq \epsilon. \tag{1}$$

An ideal articulated shape has point joints ($\epsilon = 0$) and its articulations are true isometries. In practice, $\epsilon > 0$, yet, the joints can be often assumed significantly smaller compared to the parts [5]. We call this assumption the *isometric model* of articulated shapes.

The shape $\mathcal{S}' = \mathcal{S} \cap \mathcal{Q}$ produced by cutting $\mathcal{S}$ with a planar shape $\mathcal{Q}$, such that $\mathcal{S}'$ has the same topology of $\mathcal{S}$ is said to be a *cut* of $\mathcal{S}$; if $\mathcal{Q}$ is convex, $\mathcal{S}'$ is said to be a *convex cut*. Note that in general, any articulated shape can be cut of the plane, assuming the cutting shape is sufficiently complicated. The intrinsic geometries of $\mathcal{S}$ and $\mathcal{S}'$ may be different in this case. However, a convex cut appears to preserve the intrinsic geometry, in the sense that for every $s_1, s_2 \in \mathcal{S}'$, $d_{\mathcal{S}'}(s_1, s_2) = d_{\mathcal{S}}(s_1, s_2)$.

In practical applications, articulated shapes are usually represented as discrete binary images sampled at a finite number of points (pixels). A finite set $\mathcal{S}_r = \{s_1, ..., s_N\} \subset \mathcal{S}$ is said to be an *r-sampling* of $\mathcal{S}$, if $\cup_{i=1}^{N} B(s_i, r) = \mathcal{S}$, where $B(s_i, r)$ denotes the Euclidean ball of radius r centered at s_i. Since the shapes are assumed to be compact, every ϵ-articulated shape has a finite r-sampling for every $r > 0$.

2.1   Measuring distance between articulated shapes

Comparison of articulated shapes can be performed by defining a *distance* $d_{\mathbb{M}}$: $\mathbb{M} \times \mathbb{M} \mapsto [0, \infty)$. Here, we develop an axiomatic approach, requiring $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q})$ to obey the following set of axioms:

A1. *Articulation invariance*: $d_{\mathbb{M}}(\mathcal{S}, f(\mathcal{S})) \leq \epsilon$ for all $\mathcal{S} \in \mathbb{M}_\epsilon$ and all articulations f of $\mathcal{S}$.
A2. *Dissimilarity*: if $d_{\mathbb{M}}(\mathcal{S}_1, \mathcal{S}_2) > \epsilon$, then there does not exist $\mathcal{S} \in \mathbb{M}_\epsilon$ and two articulations f_1, f_2 of $\mathcal{S}$, such that $\mathcal{S}_1 = f_1(\mathcal{S})$ and $\mathcal{S}_2 = f_2(\mathcal{S})$.
A3. *Partial matching*: for every $\mathcal{S} \in \mathbb{M}_\epsilon$ and its convex cut $\mathcal{S}'$, $d_{\mathbb{M}}(\mathcal{S}, \mathcal{S}') = 0$.
A4. *Triangle inequality*: for every $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3 \in \mathbb{M}_\epsilon$, $d_{\mathbb{M}}(\mathcal{S}_1, \mathcal{S}_2) + d_{\mathbb{M}}(\mathcal{S}_2, \mathcal{S}_3) \geq d_{\mathbb{M}}(\mathcal{S}_1, \mathcal{S}_3)$.
A5. *Sampling consistency*: for every r-samplings $\mathcal{S}_r$ of $\mathcal{S}$ and $\mathcal{Q}_r$ of $\mathcal{Q}$, $|d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) - d_{\mathbb{M}}(\mathcal{S}_r, \mathcal{Q}_r)| \leq 2r$.

In simple words, axioms A1–A2 guarantee that $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q})$ is a good similarity measure, assigning large distances for dissimilar shapes and small distances for similar shapes, while being insensitive to articulations. Note that we do not demand the converse of A1 to hold. In fact, two different ϵ-articulated shapes with intrinsic geometry differing by less than ϵ cannot be discerned in the framework of the isometric model. Axiom A3 allows us to match a portion of a shape to its whole. In order to make the partial matching well-defined, we restrict the cut to be convex. Axiom A4 provides basic metric properties. Note that demanding A3, $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q})$ cannot be made symmetric and thus the triangle inequality holds only in a non-symmetric manner. Finally, Axiom A5 enables a discretization and a numerical computation of $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q})$.

Here, we use the following distance between articulated shapes

$$d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) = \inf_{\varphi : \mathcal{Q} \mapsto \mathcal{S}} \text{dis}\, \varphi, , \tag{2}$$

which essentially measures the least possible distortion of embedding shape $\mathcal{Q}$ into shape $\mathcal{S}$. This distance is intimately related to the Gromov-Hausdorff

distance [7–9]. A very similar distance has been proposed in [10] for bending-invariant matching of three-dimensional objects.

Theorem 1. *$d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q})$ in (2) obeys axioms A1-5.*

Proof. A1: Let $\mathcal{S}$ be a planar shape and $f : \mathcal{S} \to \mathcal{Q}$ a surjective mapping with dis $f \leq \epsilon$. Define $\varphi : f(S) \to S$ by assigning to every $q \in Q$ an arbitrary point $s \in f^{-1}(q)$ in the pre-image of q. Since $f(\varphi(q)) = f(s) = q$, one has $|d_{\mathcal{Q}}(q, q') - d_{\mathcal{S}}(\varphi(q), \varphi(q'))| \leq$ dis $f \leq \epsilon$ for every $q, q' \in \mathcal{Q}$. Consequently, $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) \leq \epsilon$.

A2: Let there be two planar shapes $\mathcal{S}_1$ and $\mathcal{S}_2$ such that $d_{\mathbb{M}}(\mathcal{S}_1, \mathcal{S}_2) > \epsilon$. Assume that there exists a mapping $\varphi : \mathcal{S}_2 \to \mathcal{S}_3$ with dis $\varphi \leq \epsilon$. Then, $|d_{\mathcal{Q}}(q, q') - d_{\mathcal{S}}(\varphi(q), \varphi(q'))| \leq$ dis $f \leq \epsilon$ for every $q, q' \in \mathcal{Q}$ and, clearly, $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) \leq \epsilon$ in contradiction to the assumption. Hence, $\mathcal{S}_1$ and $\mathcal{S}_2$ are not ϵ-isometric.

A3: Let there be a planar shape $\mathcal{S}$ and $\mathcal{S}' \subset \mathcal{S}$ a convex cut of $\mathcal{S}$. Since for every $s, s' \in \mathcal{S}'$, $d_{\mathcal{S}'}(s, s') = d_{\mathcal{S}}(s, s')$, the identity mapping $\varphi : \mathcal{S}' \to \mathcal{S}$ yields $|d_{\mathcal{S}'}(s, s') - d_{\mathcal{S}}(\varphi(s), \varphi(s'))| = 0$. Hence, $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) \leq$ dis $\varphi = 0$.

A4: Let there be three planar shapes $\mathcal{S}_1, \mathcal{S}_2$ and $\mathcal{S}_3$ such that $d_{\mathbb{M}}(\mathcal{S}_1, \mathcal{S}_2) < \epsilon_1$ and $d_{\mathbb{M}}(\mathcal{S}_2, \mathcal{S}_3) < \epsilon_2$. Then, there exist two mappings $\varphi_1 : \mathcal{S}_2 \to \mathcal{S}_1$ and $\varphi_2 : \mathcal{S}_3 \to \mathcal{S}_2$ with dis $\varphi_1 < \epsilon_1$ and dis $\varphi_2 < \epsilon_2$. Denote by $\psi = \varphi_1 \circ \varphi_2 : \mathcal{S}_3 \to \mathcal{S}_1$. Invoking the triangle inequality for real numbers, one has

$$
\begin{aligned}
|d_{\mathcal{S}_3}(s, s') - d_{\mathcal{S}_1}(\psi(s), \psi(s'))| &\leq \\
&\leq |d_{\mathcal{S}_3}(s, s') - d_{\mathcal{S}_2}(\varphi_2(s), \varphi_2(s'))| + |d_{\mathcal{S}_2}(\varphi_2(s), \varphi_2(s')) - d_{\mathcal{S}_1}(\psi(s), \psi(s'))| \\
&\leq \text{dis } \varphi_2 + \text{dis } \varphi_1 < \epsilon_1 + \epsilon_2
\end{aligned}
$$

for every $s, s' \in \mathcal{S}_3$. Hence, dis $\psi < \epsilon_1 + \epsilon_2$, implying $d_{\mathbb{M}}(\mathcal{S}_1, \mathcal{S}_3) \leq d_{\mathbb{M}}(\mathcal{S}_1, \mathcal{S}_2) + d_{\mathbb{M}}(\mathcal{S}_2, \mathcal{S}_3)$.

A5: Using the (non-symmetric) triangle inequality, one has $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) \leq d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}_r) + d_{\mathbb{M}}(\mathcal{Q}_r, \mathcal{Q}) \leq d_{\mathbb{M}}(\mathcal{S}_r, \mathcal{Q}_r) + d_{\mathbb{M}}(\mathcal{S}, \mathcal{S}_r) + d_{\mathbb{M}}(\mathcal{Q}_r, \mathcal{Q})$ and, similarly, $d_{\mathbb{M}}(\mathcal{S}_r, \mathcal{Q}_r) \leq d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) + d_{\mathbb{M}}(\mathcal{S}, \mathcal{S}_r) + d_{\mathbb{M}}(\mathcal{Q}_r, \mathcal{Q})$, yielding $|d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q}) - d_{\mathbb{M}}(\mathcal{S}_r, \mathcal{Q}_r)| \leq d_{\mathbb{M}}(\mathcal{S}, \mathcal{S}_r) + d_{\mathbb{M}}(\mathcal{Q}_r, \mathcal{Q})$. Since $\mathcal{S}_r \subset \mathcal{S}$ and $d_{\mathcal{S}_r} = d_{\mathcal{S}}|_{\mathcal{S}^r}$, according to (A3), $d_{\mathbb{M}}(\mathcal{S}, \mathcal{S}_r) = 0$. It is therefore sufficient to show that $d_{\mathbb{M}}(\mathcal{Q}_r, \mathcal{Q}) \leq 2r$. Let us define a mapping $\varphi : \mathcal{Q} \to \mathcal{Q}_r$ as $\varphi(q) = \arg\min_{q' \in \mathcal{Q}_r} d_{\mathcal{Q}}(q, q')$ (the mininmum exists, since $\mathcal{Q}_r$ can be replaced by a finite sub-covering). Since Q_r is an r-covering, $d_{\mathcal{Q}}(q, \varphi(q)) \leq r$ for every $q \in \mathcal{Q}$. If q, q' are both in $\mathcal{Q}_r$, then $|d_{\mathcal{Q}}(q, q') - d_{\mathcal{Q}}(\varphi(q), \varphi(q'))| = 0$. If $q \in \mathcal{Q}_r$ and $q' \in \mathcal{Q}$, then $|d_{\mathcal{Q}}(q, q') - d_{\mathcal{Q}}(\varphi(q), \varphi(q'))| = |d_{\mathcal{Q}}(q, q') - d_{\mathcal{Q}}(q, \varphi(q'))| \leq d_{\mathcal{Q}}(q', \varphi(q')) \leq r$. If both $q, q' \in \mathcal{Q}$, then $|d_{\mathcal{Q}}(q, q') - d_{\mathcal{Q}}(\varphi(q), \varphi(q'))| \leq d_{\mathcal{Q}}(q, \varphi(q)) + d_{\mathcal{Q}}(q', \varphi(q')) \leq 2r$. $\square$

In practice, it is useful to replace $d_{\mathbb{M}}(\mathcal{S}, \mathcal{Q})$ by an L_p-norm analog,

$$
d_{\mathbb{M}}^p(\mathcal{S}, \mathcal{Q}) = \left(\frac{1}{A_{\mathcal{Q}}^2} \inf_{\varphi: \mathcal{Q} \mapsto \mathcal{S}} \int \int_{\mathcal{Q} \times \mathcal{Q}} (d_{\mathcal{Q}}(q, q') - d_{\mathcal{S}}(\varphi(q), \varphi(q')))^p \, dq dq' \right)^{1/p}, \quad (3)
$$

where dq is the standard area measure in $\mathbb{R}^2$ and $A_{\mathcal{Q}} = \int_Q dq$. In the limit $p \to \infty$, $d_{\mathbb{M}}^p$ is just $d_{\mathbb{M}}$.
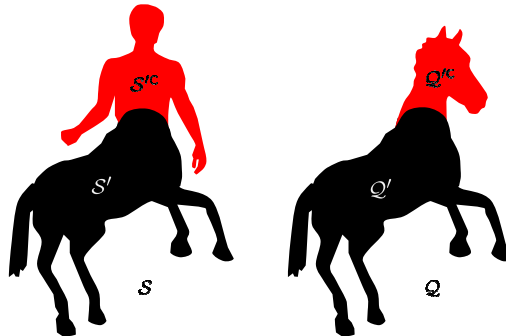
Fig. 2. A centaur (left) and a horse (right) share the bottom part of the body and differ in the upper part. This example is inspired by [11].

Apart from giving a quantitative measure of similarity of two shapes, computation of $d_{\mathbb{M}}^p(\mathcal{S}, \mathcal{Q})$ also yields a *correspondence* φ between $\mathcal{S}$ and $\mathcal{Q}$. Such correspondence may be useful in many applications including tracking of silhouettes in video sequences and alignment of articulated shapes.

2.2   Comparison of partially overlapping shapes

By virtue of axiom A3, $d_{\mathbb{M}}$ allows to compare between a shape and its portion. However, in a more general setting of partial matching, one shape does not necessarily has to be a portion of the other. As a motivating example, consider two planar shapes $\mathcal{S}$ and $\mathcal{Q}$ in Figure 2, which share some large similar portions $\mathcal{S}' \subset \mathcal{S}$ and $\mathcal{Q}' \subset \mathcal{Q}$, yet also have dissimilar portions $\mathcal{S}'^{c} = \mathcal{S} \setminus \mathcal{S}'$ and $\mathcal{Q}'^{c} = \mathcal{Q} \setminus \mathcal{Q}'$. We now outline a method to handle this setting as well.

Let us assume that the computation of $d_{\mathbb{M}}^p(\mathcal{S}, \mathcal{Q})$ gives us a minimum-distortion mapping $\varphi : \mathcal{Q} \mapsto \mathcal{S}$.[1] We define the *local distortion* at a point q as

$$\mathrm{dis}(q; \varphi) = \left(\frac{1}{A_{\mathcal{Q}}} \int_{\mathcal{Q}} \left(d_{\mathcal{Q}}(q, q') - d_{\mathcal{S}}(\varphi(q), \varphi(q')) \right)^p da \right)^{1/p}. \qquad (4)$$

This allows to attribute each point in $q \in \mathcal{Q}$ a quantitative measure of metric distortion introduced by the mapping φ to the distances between all pairs of the form (q, q'), $q' \in \mathcal{Q}$. We can define a portion $\mathcal{Q}'(\rho) = \{q : \mathrm{dis}(q; \varphi) \leq \rho\}$ consisting of all points with local distortion below some threshold ρ. This allows to segment $\mathcal{Q}$ to regions similar to $\mathcal{S}$ and those dissimilar to $\mathcal{S}$.

Omitting technical details, we can measure the relative area of the complement of $\mathcal{Q}'(\rho)$, $\mu(\rho) = 1 - \frac{1}{A_{\mathcal{Q}}} \int_{\mathcal{Q}'(\rho)} da$, and construct a generalized distance function $d'_{\mathbb{M}}(\mathcal{S}, \mathcal{Q})$ assigning to each pair of shapes $(\mathcal{S}, \mathcal{Q})$ a monotonically decreasing function $\mu : [0, \mathrm{diam}\, \mathcal{Q}] \mapsto [0, 1]$. Such a function, essentially similar to a

---

[1] We omit here some technical details: in reality, φ does not necessarily exist, yet $\mathrm{dis}(q; \varphi)$ can still be defined using a sequence of mappings φ_n with convergent distortion.

receiver operator characteristic (ROC) curve, allows many definitions of a partial order relation, which is necessary for measuring the similarity of the shapes. For example, given that the objects subject to comparison are ϵ-articulated shapes, we can set $\rho = \epsilon$ and use the relative area $\mu(\epsilon)$ of the dissimilar portions as the similarity measure. A dual approach is to fix some μ_0 (say, 80% of the shape area) and use ρ for which $\mu(\rho) = \mu_0$ as a measure similarity.

3   Generalized multidimensional scaling

We now address the issue of practical computation of $d_{\mathbb{M}}^p$. Let $\mathcal{S}_r = \{s_1, ..., s_M\}$ and $\mathcal{Q}_r = \{q_1, ..., q_N\}$ be finite r-samplings of articulated shapes $\mathcal{S}$ and $\mathcal{Q}$ (for example, r can be the pixel size when the shapes are represented as binary images) and $\mathbf{\Delta}_\mathcal{S} = (d_\mathcal{S}(s_i, s_j))$ and $\mathbf{\Delta}_\mathcal{Q} = (d_\mathcal{Q}(q_i, q_j))$ be the $M \times M$ and $N \times N$ matrices of geodesic distances between the samples of $\mathcal{S}_r$ and $\mathcal{Q}_r$, respectively. The distances are computed numerically using the *fast marching method* (FMM) [12, 13].

In this discrete setting, $d_{\mathbb{M}}^p$ can be formulated as

$$d_{\mathbb{M}}^p(\mathcal{S}_r, \mathcal{Q}_r) = \left(\min_{s_1', ..., s_N'} \sum_{i,j=1}^N a_i a_j \left| d_\mathcal{Q}(q_i, q_j) - d_\mathcal{S}(s_i', s_j') \right|^p \right)^{1/p}, \qquad (5)$$

for $p < \infty$, and

$$d_{\mathbb{M}}^\infty(\mathcal{S}_r, \mathcal{Q}_r) = \min_{\tau \geq 0, s_1', ..., s_N'} \tau \quad \text{s.t} \quad \left| d_\mathcal{Q}(q_i, q_j) - d_\mathcal{S}(s_i', s_j') \right| \leq \tau \qquad (6)$$

for $p = \infty$, where $s_i' = \varphi(q_i)$ denote the image of q_i under the mapping φ. The weights a_i are selected as the normalized areas of the Voronoi cells of q_i. In practice, when the sampling is sufficiently regular, the simple choice $a_i = 1/N$ appears to be a more convenient alternative.

Problems (5) and (6) can be considered as a generalization of *multidimensional scaling* (MDS) [14] to general metric spaces. We call it the *generalized* MDS or GMDS for short. The optimization is performed directly on the images $s_i' = \varphi(q_i)$, in the spirit of MDS. Since s_i' may fall between the samples of $\mathcal{S}$, one has to compute the geodesic distances $d_\mathcal{S}$ between any two arbitrary points in $\mathcal{S}$. For this purpose, we use the *three-point geodesic distance approximation*, a numerical procedure is to produce a computationally efficient $\mathcal{C}^1$-approximation for $d_\mathcal{S}$ and its derivatives, interpolating their values from the matrix $\mathbf{\Delta}_\mathcal{S}$ of pairwise geodesic distances in $\mathcal{S}$ [9].

The numerical solution of the GMDS problem consists of finding an unconstrained minimum of the following *generalized stress function*

$$\sigma(\mathbf{u}_1, ..., \mathbf{u}_N) = \sum_{i,j=1}^N w_{ij} \left| \delta_{ij} - d_\mathcal{S}(\mathbf{u}_i, \mathbf{u}_j) \right|^p, \qquad (7)$$

where $\delta_{ij} = d_{\mathcal{Q}}(q_i, q_j)$ denote the elements of $\mathbf{\Delta}_{\mathcal{Q}}$, $w_{ij} = a_i a_j$, and $\mathbf{u}_i \in \mathcal{S}$ are vectors of coordinates in $\mathbb{R}^2$ representing s'_i. When $p = \infty$, constrained minimization is used. In our implementation, we used a gradient descent algorithm safeguarded by inexact linesearch (Armijo rule) [15]. The complexity of the stress and its gradient computation is $\mathcal{O}(N^2)$. Typically, N varies between tens to hundreds of points, therefore GMDS is computationally efficient.

Like the traditional MDS, GMDS is a non-convex optimization problem, and therefore convergence to local minima rather than to the global one is possible[14]. Nevertheless, convex optimization is widely used in the MDS community if some precautions are taken in order to prevent convergence to local minima. Here, we use a multiscale optimization scheme that in practical applications shows good global convergence [9, 16].

4 Results

In order to assess the proposed approach, three experiments were performed. In the first experiment, the Tools A dataset[2] consisting of 35 shapes of 7 different tools, was used (see Figure 3). The tools were classified into 4 groups: scissors, pliers, pincers, cutters and knife. All the tools excepting the knife have four parts and one joint. The knife has three parts and two joints. GMDS was used to compute $d_{\mathbb{M}}^p$ with $p = 2$ between the shapes. We used a multiresolution optimization scheme, initialized at 5 points at the coarsest resolution. A total of $N = 25$ points were used. Figure 4 visualizes these distances as Euclidean similarity pattern. One can observe that the shapes are clearly distinguishable and form groups corresponding to their classification (e.g. two different shapes of scissors and pliers are close to each other). Note that different articulations are also distinguishable, such that one (at least theoretically) can infer the articulation constant ϵ of each shape.

In the second experiment, three partial probes for each of the seven tools from the Tools A dataset were used in matching against the set of 35 full shapes. Figure 5 presents the three first closest matches; due space limitations, only representative results are shown. In all cases, the first match was found correctly.

In the third experiment, the Tools B dataset consisting of three instances with minute modifications of small details of four objects from the Tools A set were used. GMDS was used to compute the correspondence φ between the shapes; the embedded shapes were discretized at $N = 50$ points. Figure 6 depicts the local distortion maps, obtained by embedding various shapes to five references models from the Tools B dataset. Note that local distortion maps manifest high distortions in dissimilar regions, which allow to capture most of the local differences between the shapes; we attribute some misses to sampling errors.

---

[2] All the data and codes will be available at `http://tosca.cs.technion.ac.il`
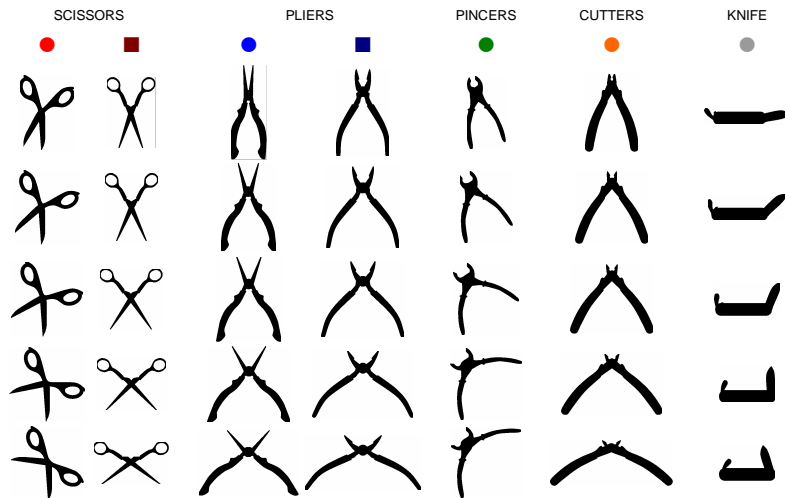
Fig. 3. Articulated shapes from the Tools A data set.

5   Conclusions

We presented a generic framework for the recognition of articulated two-dimensional shapes based on the isometric model. According to this model, articulations arise from near-isometric transformations and therefore inflict small changes to the geodesic distances measured inside the shape. We showed a distance able to distinguish between shapes insensitively to their articulations. This distance is also capable of performing partial matching of shapes and finding local dissimilarities between them. The distance computation is formulated as an MDS-like problem, which is efficiently solved using smooth optimization techniques.

References

1. D. Geiger R. Basri, L. Costa and D. Jacobs. Determining the similarity of deformable shapes. *Vision Research*, 38:2365–2385, 1998.
2. Y. Gdalyahu and D. Weinshall. Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes. *IEEE Trans. PAMI*, 21:1312–1328, 1999.
3. S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape context. *IEEE Trans. PAMI*, 24:509–522, 2002.
4. J. Zhang, R. Collins, and Y. Liu. Representation and matching of articulated shapes. In *Proc. CVPR*, volume 2, pages 342 – 349, June 2004.
5. H. Ling and D. Jacobs. Using the inner-distance for classification of articulated shapes. In *Proc. CVPR*, 2005.
6. A. Elad and R. Kimmel. Bending invariant representations for surfaces. In *Proc. CVPR*, pages 168–174, 2001.
7. M. Gromov. *Structures métriques pour les variétés riemanniennes*. Number 1 in Textes Mathématiques. 1981.
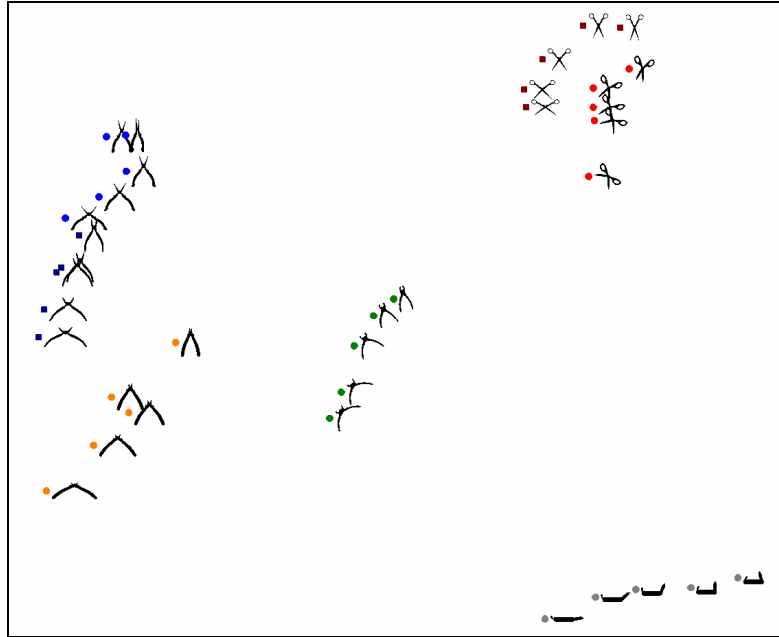
Fig. 4. Visualization of distances between the Tools A shapes.

8. F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 2005. to appear.
9. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Efficient computation of isometry-invariant distances between surfaces. Technical report, Dept. of Computer Science, Technion, Israel, 2005. submitted.
10. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proc. National Academy of Sciences*, 103(5):1168–1172, January 2006.
11. D. Jacobs, D. Weinshall, and Y. Gdalyahu. Class representation and image retrieval with non-metric distances. *IEEE Trans. PAMI*, 22:583–600, 2000.
12. J. A. Sethian. A review of the theory, algorithms, and applications of level set method for propagating surfaces. *Acta numerica*, pages 309–395, 1996.
13. R. Kimmel and J. A. Sethian. Computing geodesic on manifolds. In *Proc. US National Academy of Science*, volume 95, pages 8431–8435, 1998.
14. I. Borg and P. Groenen. *Modern multidimensional scaling - theory and applications*. Springer-Verlag, Berlin Heidelberg New York, 1997.
15. D. Bertsekas. *Nonlinear programming*. Atlanta Scientific, 2 edition, 1999.
16. M. M. Bronstein, A. M. Bronstein, R. Kimmel, and I. Yavneh. Multigrid multidimensional scaling. *Numerical Linear Algebra with Applications (NLAA)*, 13:149–171, March-April 2006.
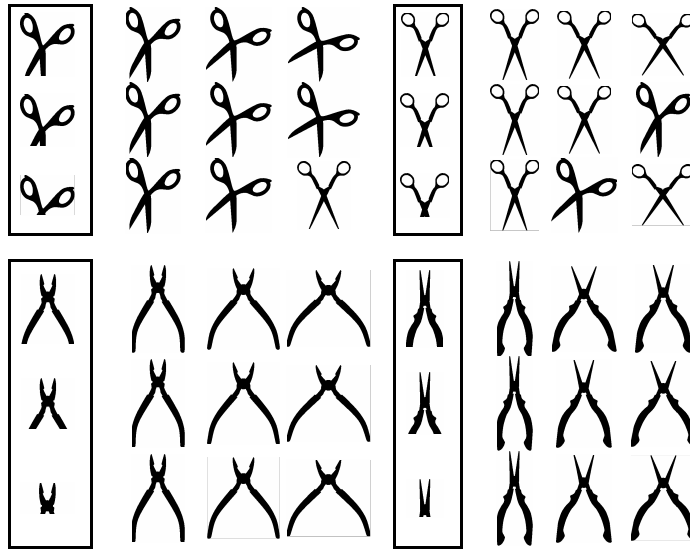
Fig. 5. Retrieval with partial probes: first three closest matches found for different partial probes (outlined).
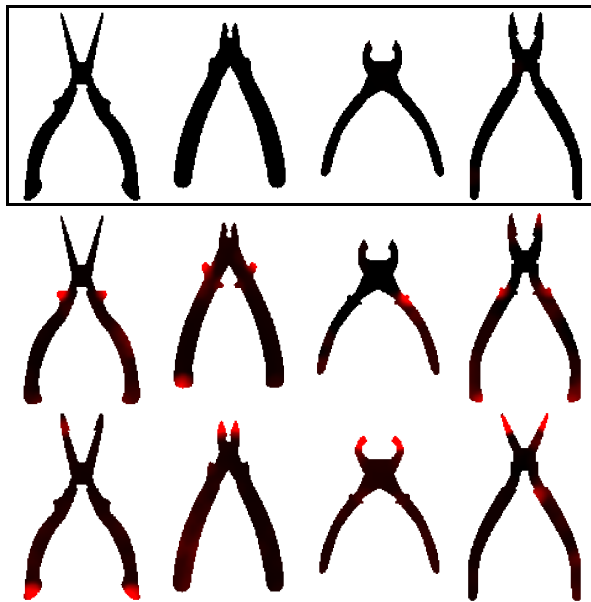

Fig. 6. Local distortion maps obtained by embedding two probes into model shapes from the Tools B dataset (top row). Distortion is represented in shades of red (high distortion) and black (low distortion).