A Game-Theoretic Approach to Deformable Shape Matching

Emanuele Rodolà
Dip. di Scienze Ambientali, Informatica e Statistica
Università Ca' Foscari, Venezia
rodola@dsi.unive.it

Alex M. Bronstein*
School of Electrical Engineering
Tel Aviv University
bron@eng.tau.ac.il

Andrea Albarelli, Filippo Bergamasco, Andrea Torsello
Dip. di Scienze Ambientali, Informatica e Statistica
Università Ca' Foscari, Venezia
albarelli,bergamasco,torsello@dsi.unive.it

Abstract

We consider the problem of minimum distortion intrinsic correspondence between deformable shapes, many useful formulations of which give rise to the NP-hard quadratic assignment problem (QAP). Previous attempts to use the spectral relaxation have had limited success due to the lack of sparsity of the obtained "fuzzy" solution. In this paper, we adopt the recently introduced alternative L^1 relaxation of the QAP based on the principles of game theory. We relate it to the Gromov and Lipschitz metrics between metric spaces and demonstrate on state-of-the-art benchmarks that the proposed approach is capable of finding very accurate sparse correspondences between deformable shapes.

1. Introduction

Finding correspondences between shapes is a fundamental problem in computer vision with a wide variety of applications ranging from robotics to medical imaging. Particularly challenging is the setting of *deformable* shape correspondence, in which shapes may undergo non-rigid deformations under which the correspondence has to be invariant. In the past decade, significant attention has been devoted to problems related to deformable shape correspondence. A large corpus of research makes use of the notion of intrinsic geometry – an umbrella term referring to geometric structures that remain invariant under non-rigid bendings and other types of transformations. In [7, 12, 4, 11, 18] and followup studies, it was proposed to use the distortion of intrinsic metrics as a measure of the correspondence quality. Finding a minimum distortion correspondence can be rigorously formulated in geometric terms and cast as an optimization problem. Several particularly useful instances of minimum distortion correspondence problems can be reduced to quadratic assignment problems (QAP). However, the combinatorial nature of QAPs makes them challenging computationally.

Different relaxations of the NP-hard QAP have been explored in the computer vision literature. In their seminal work, Gold and Rangarajan [8] relax the assignment and solve the optimization problem through a gradient method over the set of bistochastic matrices. In [10], a spectral approach to correspondence finding was presented where mapping constraints are met by iteratively removing inconsistent or weak assignments until an optimum binarized solution is found. The procedure has been successfully applied to 2D matching and recognition, and subsequently extended to other contexts such as isometry-enforcing 3D nonrigid matching [15].

A similar setup has been recently considered in [2], with a cardinal difference of replacing the L^2 constraint $\mathbf{u}^{\mathrm{T}}\mathbf{u} = 1$ by $\mathbf{u}^{\mathrm{T}}\mathbf{1} = 1$, $\mathbf{u} \geq \mathbf{0}$. This modification makes the assignment problem more combinatorial in nature, and, like most types of L^1 constraints, promotes sparsity of the solution. The authors proposed to solve the resulting quadratic program by an iterative procedure based on a game-theoretic evolutionary process, which proved to be very effective and robust. Their method has been extended to other tasks such as 3D rigid surface alignment [1] and facial point localization [14].

In this paper, we adopt the game-theoretic approach as the means to find minimum distortion intrinsic correspondences between non-rigid shapes. The contributions of this paper are three-fold: First, we show an interpretation of the QAPs commonly used in shape matching from the point of view of Gromov and Lipschitz distances between met-

*Research supported by the Israeli Science Foundation and the German-Israeli Foundation.

ric spaces. Second, we adapt the game-theoretic framework to efficiently solve the resulting optimization problems, and again show the relation between different heuristics used by this framework to distances between metric spaces. Third, we propose a method to aggregate multiple sparse solutions obtained using the game-theoretic solver into a denser correspondence. Though the proposed approaches are general and work with any intrinsic distances, in this study we focus on the family of diffusion distances that has a natural scale-space interpretation. We show how to aggregate information from different scales into a single distortion functional.

2. Intrinsic geometries

We model shapes as compact smooth Riemannian manifolds equipped with an *intrinsic metric* d and the standard measure induced by the volume form. By the term intrinsic metric we refer to a distance function on the manifold that depends only on the Riemannian structure and is independent of the way it is embedded in the ambient space. One of the straightforward constructions of an intrinsic metric is the geodesic metric measuring the length of the shortest path (minimal geodesic) connecting a pair of points on the surface. Such a metric is invariant to inelastic bendings, that is, such deformations that do not stretch or tear the shape. A serious disadvantage of the geodesic geometry is its extreme sensitivity to topological noise. In fact, even a point topological change has a great influence on the length of the shortest path. Generally, this influence does not decay as one goes away from the affected point, limiting the practical applicability of geodesic distances.

A partial remedy to this problem has been found in another family of intrinsic geometries introduced by Coifman and Lafon [6] under the name of *diffusion geometry*. Diffusion geometry is an umbrella term referring to intrinsic distances and other geometric quantities based on the properties of diffusion processes on the surface. Diffusion processes are described by the heat equation

$$\Delta f(x,t) + \tfrac{\partial}{\partial t} f(x,t) \quad = \quad 0, \tag{1}$$

with $f(x,t)$ denoting the distribution of heat on the surface at point x at time t, and Δ being the Laplace-Beltrami operator. The equation has the initial condition $f(x, t = 0)$ describing the initial heat distribution; boundary conditions apply in case the manifold has a boundary.

The solution of the heat equation at point x at time t initialized with a point distribution at x' is called the *heat kernel* and is denoted by $h_t(x,x')$. The heat kernel describes the proximity of two points x and x' at different scales t. This notion of proximity can be used to define a family of intrinsic metrics

$$d_t^2(x,x') \quad = \quad \int_X (h_t(x,y) - h_t(x',y))^2 dy \tag{2}$$

called the *diffusion metrics*. The family is parameterized by the scale parameter t and naturally forms a scale space: diffusion metric with small t are sensitive to small features while being rather indiscriminative at larger scale; on the contrary, d_t with large values of t is insensitive to small features, yet captures the global geometry of the shape.

In order to make diffusion distances commensurable and comparable across different scales, they are often normalized by the trace of the heat kernel (the heat trace),

$$H_{X,t} \quad = \quad \frac{1}{\mathrm{Vol}(X)} \int_X h_t(x,x) dx, \tag{3}$$

where $\mathrm{Vol}(X)$ stands for the total area of X. This results in the family of normalized metrics,

$$\hat{d}^2_{X,t}(x,x') \quad = \quad \frac{d^2_{X,t}(x,x')}{H_{X,t}}. \tag{4}$$

The framework of diffusion geometry also allows to define intrinsic point-wise feature descriptors (or signatures) on the surface. In [17], it was shown that under mild technical assumptions, the diagonal $\{h_t(x,x)\}_{t>0}$ of the heat kernel contains full information about the shape's intrinsic geometry (*i.e.*, fully describes the underlying Riemannian structure). The authors proposed to associate each point of the surface with a vector-valued descriptor $\mathbf{h}(x) = (h_{t_1}(x,x), \ldots, h_{t_k}(x,x))$, dubbed as the *heat kernel signature* (HKS). A scale-invariant version of the HKS (SIHKS) was consequently introduced in [5]. Since the Laplace-Beltrami operator is an intrinsic property of the shape, quantities associated with it such as the heat kernel and descriptors based on it are also intrinsic. Being constructed from the same geometric quantities, both diffusion metrics and corresponding signatures are related in that nearly isometric shapes in the sense of the diffusion metrics will also be described by similar HKS and vice versa.

In what follows, we are going to use diffusion geometric quantities to formalize the notion of correspondence between shapes. Most of the presented discussion is however valid for any type of intrinsic metrics.

3. Intrinsic shape correspondence

We define a *correspondence* between two shapes X and Y as the subset $U \subset X \times Y$ satisfying: 1) for every $x \in X$, there exists (at least one) $y \in Y$ such that $(x,y) \in U$; and, vice versa, 2) for every $y \in Y$, there exists $x \in X$ such that $(x,y) \in U$. This relation can be thought of as a generalization of the notion of a function, and can be alternatively formulated as the binary function $u : X \times Y \to \{0,1\}$ satisfying for every $x \in X$ and $y \in Y$,

$$\max_{y \in Y} u(x,y) \; = \; \max_{x \in X} u(x,y) \quad = \quad 1. \tag{5}$$

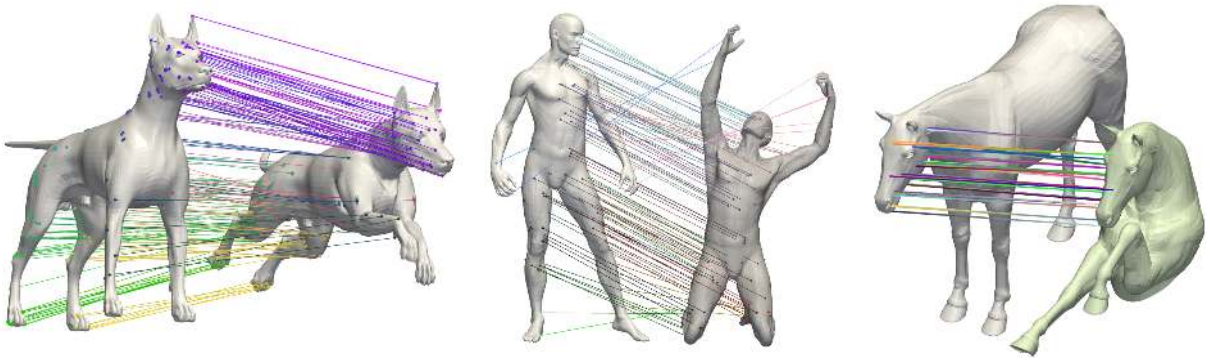

Figure 1. Examples of correspondences obtained with our method. The game-theoretic approach produces a sparse (around 1% of the shape is matched), yet very accurate correspondence which can be used as a robust initialization for subsequent refinement (first two images). The last image presents a case of partial matching, where the second shape additionally underwent a local scale deformation. In this case we applied the merging approach on 5 groups, resulting in 51 matches with an average ground-truth error of 2.57 (see section 4.2).

Suppose two pairs of points (x, y) and (x', y') are in correspondence. Then, we can quantify the quality of the correspondence by measuring to which extent the distance between x and x' measured on X using d_X matches the distance between the corresponding points y and y' measured on Y using d_Y,

$$\epsilon(x, y, x', y') \quad = \quad |d_X(x, x') - d_Y(y, y')|. \tag{6}$$

The worst-case distortion of the metric caused by the correspondence U is given by

$$\|\epsilon\|_{L^\infty(U\times U)} = \sup_{(x,y),(x',y')\in U} \epsilon(x, y, x', y'). \tag{7}$$

Minimizing the distortion over all possible correspondences between X and Y yields a distance

$$D(X, Y) \quad = \quad \frac{1}{2} \inf_U \|\epsilon\|_{L^\infty(U\times U)} \tag{8}$$

between X and Y called the *Gromov-Hausdorff distance*. If the infimum is realized by some U^*, the latter is called a *minimum distortion correspondence* (note that more than one minimum distortion correspondence might exist if the shape possesses intrinsic symmetries). By using intrinsic metrics d_X and d_Y, the obtained correspondence is also intrinsic. In particular, this implies invariance to inelastic bending of the shapes.

It is worthwhile noting that taking the logarithm of the metrics d_X, d_Y, one can replace the absolute distortion (6) with a relative counterpart

$$\begin{aligned} \epsilon(x, y, x', y') \quad &= \quad |\log d_X(x, x') - \log d_Y(y, y')| \quad (9) \\ &= \quad \log \max \left\{ \frac{d_X(x, x')}{d_Y(y, y')}, \frac{d_Y(y, y')}{d_X(x, x')} \right\}. \end{aligned}$$

The resulting distance (8) is called the *Lipschitz distance*. Note that $\epsilon(x, y, x', y') = \infty$ whenever $x = x'$ or $y = y'$, requiring the correspondence u to be bijective. For this reason, the Lipschitz distance is only applicable to topologically equivalent shapes.

Both the Gromov-Hausdorff and the Lipschitz distances constitute a metric on the space of all (homeomorphic in case of the Lipschitz metric) shapes modulo their d-isometries. They naturally express the similarity relation of two shapes being "approximately isometric", and can be consistently discretized [4]. However, the L^∞ formulation makes the Gromov-Hausdorff and the Lipschitz distances of little practical use due to their sensitivity to noise and outliers.

While an L^p relaxation of the distortion (7) would theoretically yield a more robust distance, its direct introduction into (8) results in a distance inconsistent to sampling. A way to overcome this difficulty was proposed by [11]. We first relax the binary notion of correspondence into a *fuzzy* notion allowing the function u to assume a continuum of values between 0 and 1, $u : X \times Y \to [0, 1]$. Condition (5) is relaxed by demanding for every measurable subsets $A \subseteq X$ and $B \subseteq Y$,

$$\int_A \int_Y u dy dx = \int_A dx; \quad \int_B \int_X u dx dy = \int_B dy. \tag{10}$$

In other words, $u(x, y)dxdy$ defines a weighted product measure on $X \times Y$ whose marginals are the measures dx and dy on X and Y, respectively. The quantity $u(x, y)dx$ can be thought of as the infinitesimal amount of mass transported from point x on X to point y on Y, while ϵ^p quantifies the cost of the transport.

Using this relaxed notion of correspondence, a new family of distances can be defined as

$$D(X, Y) \quad = \quad \frac{1}{2} \inf_u \|\epsilon\|_{L^p(u\times u)}, \tag{11}$$

where $1 \leq p \leq \infty$, and

$$\|\epsilon\|^p_{L^p(u\times u)} = \int_{(X\times Y)^2} \epsilon^p(x,y,x',y')u(x,y)u(x',y')dxdydx'dy'. \quad (12)$$

$D(X,Y)$ constitute metrics on the space of equivalence classes of shapes under the isomorphism relation of metric-measure spaces (*i.e.*, measure-preserving isometries). In literature, this class of metrics is usually referred to as Wasserstein or earth mover's distances. Here, following [11] we will refer to them as the Gromov-Wasserstein metrics to emphasize the relation to the Gromov-Hausdorff distances. We note, however, that the two metrics are not equivalent, for the very same reasons the Hausdorff and the earth mover's metrics are not equivalent.

3.1. Multi-scale distortion

In the particular case where diffusion metrics are used to measure distances on X and Y, the selection of the scale parameter is crucial. Small scales alone give excellent feature localization (and hence accurate correspondence), but are not robust globally; on the other hand, large scales alone do not give accurate correspondences, while stabilize the global matching problem. Here, instead of selecting a single scale, we propose to combine several scales into a single distortion criterion,

$$\epsilon^p(x,y,x',y') = \int_{T_1}^{T_2} \left(\hat{d}_{X,t}(x,x') - \hat{d}_{Y,t}(y,y')\right)^p dt, \quad (13)$$

where T_1 and T_2 are parameters determining the range of scales, and $\hat{d}_{X,t}$ are the scale-normalized diffusion distances. Aggregation of multiple scales of spectral distances has been previously successfully used in shape retrieval applications [11].

3.2. Discretization

In the discrete setting, let us assume the shapes X and Y to be represented by m and n points, respectively, with the vectors $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ discretizing the corresponding area elements. The Gromov-Wasserstein metric assumes the form

$$D(X,Y) = \frac{1}{2}\min_{\mathbf{U}} \sum_{i,j,i',j'} \epsilon^p_{iji'j'}\mu_i\nu_j\mu_{i'}\nu_{j'}u_{ij}u_{i'j'}. \quad (14)$$

Absorbing the area elements into the cost term and using matrix notation, we arrive at the following quadratic program

$$\min_{\mathbf{U}\geq\mathbf{0}} \mathrm{vec}\{\mathbf{U}\}^{\mathrm{T}}\mathbf{B}\mathrm{vec}\{\mathbf{U}\} \quad \text{s.t} \quad \begin{cases} \mathbf{U}\mathbf{1} = \mathbf{1} \\ \mathbf{U}^{\mathrm{T}}\mathbf{1} = \mathbf{1} \end{cases} \quad (15)$$

where $\mathrm{vec}\{\mathbf{U}\}$ stands for the mn-dimensional column-stack vector representation of the $m \times n$ correspondence matrix $\mathbf{U}$, $\mathbf{1}$ is a vector of ones of appropriate dimensions, and $\mathbf{B}$ is the $mn \times mn$ cost matrix containing the elements $\epsilon^p_{iji'j'}\mu_i\nu_j\mu_{i'}\nu_{j'}$.

Since our final goal lies in finding the minimum-distortion correspondence rather than computing the Gromov-Wasserstein metric, we are interested in a minimizer rather than a minimum of the above problem. We observe that while the L^1-type constraints are known to favor a sparse solution (*i.e.*, $\mathbf{U}$ will have few strong non-zero elements), it is still a fuzzy correspondence matrix. This may be disadvantageous in matching applications, where usually bijectivity is required. In order to impose bijectivity of the solution, we modify the cost function by setting $\epsilon_{iji'j'} = \infty$ for every $i = i'$ or $j = j'$, exactly as we did in the case of the Lipschitz metric. We denote the modified cost matrix by $\tilde{\mathbf{B}}$.

Finally, observe that the constraints on row and column sums of $\mathbf{U}$ in (15) require it to be a full correspondence (*i.e.*, each point in X corresponds to a point in Y). This is rather a restrictive setting for many applications where a *partial* rather than full correspondence is sought. In order to allow for some points on X to have no corresponding points on Y and vice versa, one has to allow some of the rows or columns of $\mathbf{U}$ to sum to zero. We propose to replace problem (15) by an under-constrained counterpart

$$\min_{\mathbf{U}\geq\mathbf{0}} \mathrm{vec}\{\mathbf{U}\}^{\mathrm{T}}\tilde{\mathbf{B}}\mathrm{vec}\{\mathbf{U}\} \quad \text{s.t} \quad \mathbf{1}^{\mathrm{T}}\mathbf{U}\mathbf{1} = 1. \quad (16)$$

Note that the obtained partial correspondence is bijective by virtue of the modified cost matrix $\tilde{\mathbf{B}}$. In what follows, we show how to efficiently solve the above optimization problem using tools from game theory.

4. Game-theoretic matching

Following [1, 2], we cast the optimization problem in an evolutionary game-theoretic framework. We start by modeling strategies as candidate assignments $(x,y) \in X \times Y$ based on some measure of pointwise similarity among the surface points. Here we use SIHKS [5] descriptors with the standard Euclidean distance since they demonstrate good resilience to a variety of deformations. We emphasize though that the descriptor is only used to constuct the initial set of possible correspondences, while the selection process depends only on the distortion of the intrinsic geometries. For this reason the descriptor need not be extremely robust since this step has the intended effect of reducing the size of the problem and increase the inlier ratio.

We simplify the notation by turning the constraint $\mathbf{1}^{\mathrm{T}}\mathbf{U}\mathbf{1} = 1$ into the equivalent $\mathbf{u} \in \Delta$ where $\mathbf{u} \equiv \mathrm{vec}\{\mathbf{U}\}$ is the correspondence vector, constrained to lie in the standard mn-simplex

$$\Delta = \left\{\mathbf{u} \in \mathbb{R}^{mn} \;:\; \mathbf{u}^{\mathrm{T}}\mathbf{1} = 1 \text{ and } \mathbf{u} \geq 0\right\}.$$

Further, we formulate program (16) as a maximization problem by considering a $mn \times mn$ matrix $\mathbf{A}$ whose elements represent the *similarity* between corresponding pairs of correspondences, a quantity inversely related to the distortion and which can be defined in a variety of ways. In this work we follow [8] and adopt a softmax ansatz to the (relaxed) QAP, which is known to improve the convergence properties of gradient methods. Choosing $p = 2$ in the distortion term, we set $a_{(ij)(i'j')} = \exp(-\alpha\epsilon^2_{iji'j'})$, which incidentally gives $a_{(ij)(i'j')} = 0$ whenever either $i = i'$ or $j = j'$. With these modifications, program (16) is transformed into the maximization problem

$$\max \mathbf{u}^{\mathrm{T}}\mathbf{A}\mathbf{u} \quad \text{s.t} \quad \mathbf{u} \in \Delta\,. \tag{17}$$

While here we are using a Gromov-Wasserstein metric, it is worth noting that in [1] the rigid correspondence problem was solved using the equivalent of a Lipschitz metric.

In our framework, the matching problem is better interpreted as an inlier selection problem in which we operate a search for the most coherent group of matches according to some notion of compatibility between them. In this scenario, pairs of players are repeatedly extracted from a (ideally infinite) population to play a symmetric game. Each player behaves according to a pre-programmed strategy, and receives a payoff from the other players proportional to how compatible his choice is with respect to his opponents. As the game is repeated, players will adapt their behavior pattern to prefer strategies that yield larger payoffs, making for a robust selection process where inconsistent hypotheses are led to extinction.

4.1. Preliminaries on game theory

Let $O = \{1, \cdots, n\}$ be the set of available *pure strategies* and $\mathbf{A} = (a_{ij})$ be a matrix specifying the payoff that an individual playing strategy i receives against an opponent playing strategy j. A *mixed strategy* $\mathbf{u} \in \Delta$ is a probability distribution over O. The expected payoff received by an i-strategist when playing against a player adopting a mixed strategy $\mathbf{u}$ is $(\mathbf{A}\mathbf{u})_i = \sum_j a_{ij}u_j$, hence the expected payoff received by adopting the mixed strategy $\mathbf{v}$ against $\mathbf{u}$ is $\pi(\mathbf{v}|\mathbf{u}) = \mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{u}$. The *best replies* against mixed strategy $\mathbf{u}$ is the set of mixed strategies maximizing the expected payoff against $\mathbf{u}$:

$$\beta(\mathbf{u}) = \{\mathbf{v} \in \Delta \mid \pi(\mathbf{v}|\mathbf{u}) = \max_{\mathbf{z}}(\pi(\mathbf{z}|\mathbf{u}))\}\,.$$

A strategy $\mathbf{u}$ is said to be a *Nash equilibrium* if it is the best reply to itself, *i.e.*, $\forall \mathbf{v} \in \Delta$, $\pi(\mathbf{u}|\mathbf{u}) \geq \pi(\mathbf{v}|\mathbf{u})$. We define the *support* $\sigma(\mathbf{u})$ to be the set surviving strategies, *i.e.*, $\sigma(\mathbf{u}) = \{i \in O \mid u_i > 0\}$. The Nash condition then implies that all strategies $i \in \sigma(\mathbf{u})$ have constant payoff $(\mathbf{A}\mathbf{u})_i = \pi(\mathbf{u}|\mathbf{u})$, while strategies outside the support of $\mathbf{u}$ earn a smaller or equal payoff. A strategy $\mathbf{u}$ is said to be an *evolutionary stable strategy* (ESS) if it is a Nash equilibrium and for all $\mathbf{v} \in \Delta$ for which $\pi(\mathbf{u}|\mathbf{u}) = \pi(\mathbf{v}|\mathbf{u})$ we have $\pi(\mathbf{u}|\mathbf{v}) > \pi(\mathbf{v}|\mathbf{v})$. In other words, deviating from the stable strategies does not pay.

Interestingly, in the special case in which $\mathbf{A}$ is symmetric, there is relationship with optimization theory [20]: Stable states correspond to the strict local maximizers of the average payoff $\pi(\mathbf{u}|\mathbf{u}) = \mathbf{u}^{\mathrm{T}}\mathbf{A}\mathbf{u}$ over Δ, whereas all critical points are related to Nash equilibria. In addition, under mild conditions on the matrix $\mathbf{A}$, the bijectivity constraints on $\mathbf{A}$ are guaranteed since a stable state cannot have in its support pairs of strategies with zero payoff [1, 2]. Assuming a mechanism to reach a stable state is available, this interesting property provides us with a rather flexible and general tool that we can adapt and employ for our purposes. In this paper we adopt a new class of dynamics (called *infection and immunization*) used to evolve the population state to an ESS [16]. This evolution process is characterized by a linear complexity per iteration, providing a very efficient (local) maximization algorithm for problem (17).

The strategy $\mathbf{u}^*$ at equilibrium constitutes a L^1 solution to (17). We note that the final values $u_i = u(x, y)$ can be interpreted as the relative contribution of each strategy to the global coherence of the correspondence, in terms of the distortion measure ϵ^2. The correspondence function u can then be binarized by keeping only the fittest strategies, *e.g.* by setting $u(x, y) = 1$ for the top 80% strategies (with respect to the maximum u_i), and putting the others to zero. In the experimental section, a specific set of experiments analyzing the influence of this parameter on the quality of the final match is presented.

4.2. Merging correspondences

The final correspondence resulting from the local maximization of (17) is characterized by a very strong internal coherence, and typically includes only a small percent (around 5-10%) of matches selected from the initial set of candidates. There exist effective methods to render correspondences denser [19]. Here we repeat application of the game-theoretic scheme in an attempt to "densify" the initial correspondence. This iterative approach is justified by the fact that the extinct strategies of a single game (those not supported by $\mathbf{u}^*$) do not necessarily have a smaller payoff than the extracted (local) maximum (see section 4.1), thus motivating the interest to explore the solution space further.

After an initial solution is obtained, we proceed by invalidating the selected strategies from the set of candidates and play a new (smaller) game with the remaining matches. Once several sets of correspondences are extracted, we need a way to merge correspondences in a manner coherent with the possible intrinsic symmetries. We take the hint from spectral clustering [13] and blended intrinsic maps [9], and formalize this notion of coherence by defining a pair-

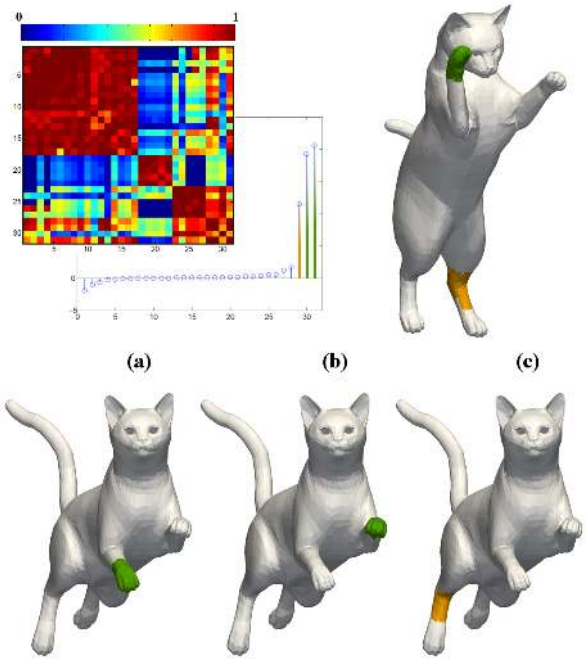

Figure 2. An example of the merging process (with real data) between two isometric shapes. After obtaining 30 correspondences, we compute the spectrum of matrix $\mathbf{S}$ (top left). The dominant eigenvector allows to retrieve the most consistent cluster of correspondences, matching the right paw of the cat (in green) (a); the next eigenvalue is only separated by a very small spectral gap, and the corresponding matches associate the right paw again with a symmetric patch (b); finally, the maximum gap eigenvector represents a reflected correspondence (in orange) with larger error (c). Figure best viewed in color.

wise measure of distortion between groups of matches, and successively operate on the resulting affinity matrix. Let G and H be two correspondence groups $(g_i, g'_i) \in G$ and $(h_j, h'_j) \in H$. We define distortion ζ as:

$$\zeta(G,H) = \frac{1}{mn}\sum_{i,j} w_{g_i g'_i} w_{h_j h'_j} (d_X(g_i,h_j) - d_Y(g'_i,h'_j))^2, \quad (18)$$

where the w_{xy} are proper weights proportional to the point-to-point matching confidence between $x \in X$ and $y \in Y$ (for instance, $w_{xy} = u(x,y)$ before binarization). From this we define the corresponding similarity $\Gamma(G,H) = \exp(-\gamma\zeta(G,H))$, where γ is a scale parameter.

If we play the game k times, we get to the definition of a (non-negative) similarity matrix $\mathbf{S} \in \mathbb{R}^{k\times k}$. The best group separation can then be represented by a selection vector $\mathbf{v}$, which (similarly to [10, 9]) we relax to take continuous values and constrain to have unitary L^2-norm. We get to the quadratic program

$$\max \mathbf{v}^{\mathrm{T}}\mathbf{S}\mathbf{v} \quad \text{s.t} \quad \|\mathbf{v}\|_{L^2}^2 = 1, \quad (19)$$

which is maximized by the leading eigenvector of $\mathbf{S}$. In presence of intrinsic symmetries, program (19) will yield a large energy value for more than one choice of $\mathbf{v}$, corresponding to different groups of coherent matches separated by a small spectral gap $|\mathbf{v}^{\mathrm{T}}\mathbf{S}\mathbf{v} - \tilde{\mathbf{v}}^{\mathrm{T}}\mathbf{S}\tilde{\mathbf{v}}|$ (see Figure 2). This provides us with a robust means to separate symmetric solutions into distinct consistent sets, while at the same time helps to filter out distorted matches that might occur as the game is repeated.

5. Experimental results

We performed a wide range of experiments on the SHREC'10 correspondence dataset [3], for which ground-truth assignments were made available by the authors. The dataset consists of 3 high-resolution (10K-50K vertices) shape classes (human, dog, horse) with simulated transformations, which are split into 9 classes: isometry, topology, small and big holes, global and local scaling, noise, shot noise, sampling. Each transformation class appears in five different strength levels, making for a total of 45 transformations per shape class. When we compute the ground-truth error of correspondence U, we take into consideration reflection intrinsic symmetries by evaluating both the direct and symmetric errors [3]:

$$D(U,U_g) = \frac{1}{|U|}\min\left\{\sum_{k=1}^{|U|} d_X(x_k,x'_k), \sum_{k=1}^{|U|} d_X(x_k,x''_k)\right\},$$

where d_X is a geodesic metric on X and $x'_k, x''_k \in U_g$ are, respectively, the direct and symmetric ground-truth positions of point $x_k \in U$.

5.1. Comparisons

We evaluate the performance of the game-theoretic method in relation to existing techniques. Table 1 reports per-deformation results at all strengths, which can be directly compared with state-of-the-art methods in [3]. Here we used the best parameters determined through the sensitivity analysis that will be presented in the section.

The table shows that the proposed approach provides better accuracy than all of the sparse approaches reported in [3], regardless of transformation class. Further, we achieve near-ideal performance in a number of cases. An interesting instance of surprisingly good behavior is represented by the local scale class, which seems to perform equally well at increasing intensities. This is due to the selective nature of the evolutionary process, which explicitly seeks for the most compatible group of matches in terms of preservation of the metric; in this case, the parts of shape that undergo a local change in scale are filtered out by the selection process, naturally favoring those portions of surface that are mostly left untouched by the transformation.

Transform.	1	≤2	≤3	≤4	≤5
Isometry	**1.47**	**1.73**	**6.83**	**1.77**	**0.68**
Topology	**2.45**	**1.05**	**3.29**	**14.70**	**11.64**
Holes	**3.93**	**3.87**	**3.88**	**7.44**	22.69
Micro holes	**1.09**	**2.59**	**3.70**	**2.34**	**2.87**
Scale	**4.01**	**0.81**	**2.11**	**9.54**	47.99
Local scale	**2.64**	9.12	**8.50**	**8.15**	**8.57**
Sampling	**1.19**	**2.56**	**11.84**	**8.72**	20.25
Noise	**3.74**	**4.34**	8.63	10.72	12.22
Shot noise	**1.46**	**1.06**	**1.09**	**2.06**	14.43
Average	**2.44**	**3.01**	**5.54**	**7.27**	**15.70**

Table 1. Performance of the game-theoretic method using SIHKS and the diffusion metric. Average number of corresponding points is 10. Values in bold indicate better performance than any of the sparse methods reported in [3].

Transform.	1	≤2	≤3	≤4	≤5
Isometry	**9.82**	15.97	**3.28**	**7.52**	**3.26**
Topology	**3.44**	**3.80**	**3.03**	**8.81**	**4.73**
Holes	31.80	**18.13**	**13.49**	**8.07**	49.88
Micro holes	**8.61**	**4.90**	**3.44**	**6.98**	**3.38**
Scale	**11.76**	**6.53**	**8.75**	**8.70**	**3.17**
Local scale	**6.89**	**15.11**	**13.00**	58.76	50.50
Sampling	**6.93**	**26.55**	40.81	**13.20**	**16.06**
Noise	**6.46**	**7.81**	**9.47**	**11.06**	**18.34**
Shot noise	**6.77**	**13.82**	**10.28**	**6.06**	**15.03**
Average	**10.28**	**12.51**	**11.73**	**14.35**	**18.26**

Table 2. Results obtained after merging the correspondences gathered from 25 games. Average number of corresponding points is 50. Values in bold indicate better performance than GMDS method as reported in [3].

By contrast, due to the multi-scale approach followed by our method, global rescaling of the shapes can easily pose problems (compare also with the "scale" curve in Figure 4).

The only approach that provides better accuracy in some instances is the spectral matching algorithm, which also provides a dense correspondence. Note, however, that this approach completely breaks for all topology-modifying transformation classes, *i.e.*, topology, holes, and sampling. On the other hand, our performance is close to that of the spectral matching algorithm for the topology-preserving transformation classes, but is also robust with respect to topology-modifying classes.

We also investigated the effectiveness of the correspondence merging approach presented in section 4.2. For this test, we iteratively generated 25 groups of matches (for each pair of shapes), built the similarity matrix with $\gamma = 10^8$ and kept the principal eigenvector by thresholding it at 60% of its maximum value. Again, the experiments were carried out on the whole dataset and are reported in Table 2. We rule out from direct comparison the spectral graph matching approaches since, as stated before, their idealistic performance is due to the use of identical triangulations in the dataset. The only other algorithm reported in [3] giving the same average number of matches is GMDS [4], while the other approaches produced much sparser matches. For this reason, Table 2 reports a direct comparison with GMDS only.

5.2. Sensitivity analysis

The next set of experiments is aimed at analyzing performance of the game-theoretic method under different parameterizations. In order to limit the size of the problem, we only consider a subset of points from the deformed mesh X. Feature points are detected by computing for all $x \in X$ the HKS function $h_t(x, x)$ for 3 values of t, and keeping points that are 2-ring local maxima across all time scales [17]. The set of strategies is finally built by generating 5 candidate matches per feature point, based on the vicinity of the associated descriptors with points from the model mesh. Finally, diffusion distances in equation (13) were calculated at time scales $(2^7, 2^8, \ldots, 2^{16})$.

Figure 4 shows the results obtained by our method with different choices of payoff coefficient α and of the selection threshold used to determine the final set of matches. We used a threshold of 0.8 for the former experiment, and $\alpha = 10^3$ for the latter. The value of α in these graphs ranges over 50 equally spaced values from 10^3 to 36×10^3. Next, since both the size and quality of the correspondence also depend on the specific set of strategies used, we performed some additional tests with a progressively less aggressive feature detection on the data meshes (Figure 3). The outcome of this experiment suggests that increasing the number of initial samples can be beneficial to the matching pro-

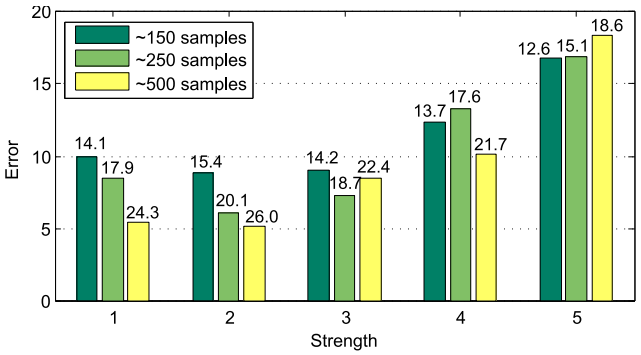

Figure 3. Evaluation of the results obtained under different initial samplings of the transformed mesh, averaged over all deformations of every shape. The initial number of samples has a direct and consistent influence on the final size of the correspondence (noted above each bar), whereas its quality does not appear to be affected at all deformation strengths.

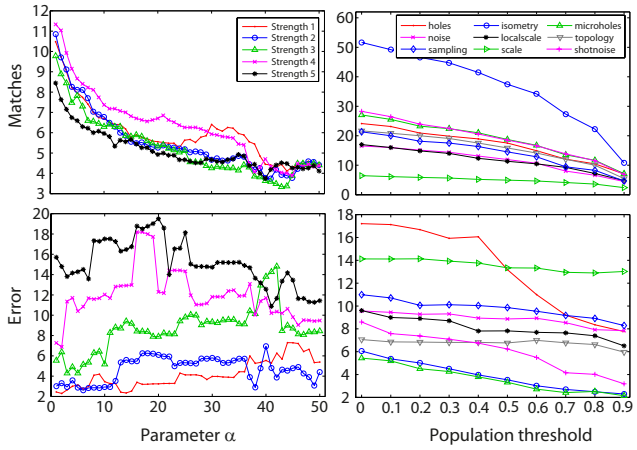

Figure 4. Sensitivity of our method to payoff coefficient α (first column) and the selection threshold used on the final population (second column). Increasing the α parameter reduces the average match distortion at the cost of a smaller correspondence. On the other hand, the population threshold has a more definite effect on size rather than quality of the final correspondence. In particular, while most transformations behave similarly, the "isometry" and "holes" classes appear to be more sensitive to this parameter.

cess; indeed, settling for a selectivity level in the feature detection step is more a matter of memory consumption, while the algorithm is able to extract correspondences in 0.5-4 seconds even with large games with tens of thousands of strategies.

6. Conclusions

We showed a game-theoretic approach to the solution of intrinsic correspondence problems arising in deformable shape analysis. Through the use of multi-scale diffusion metrics, we showed how to fuse information from different scales into a single distortion criterion minimized in search of a minimum distortion correspondence. Evaluation on the SHREC'10 non-rigid shape correspondence benchmark demonstrated that the proposed approach is capable of recovering accurate sparse correspondences between shapes and is robust under a variety of strong deformations.

References

[1] A. Albarelli, E. Rodolà, and A. Torsello. A game-theoretic approach to fine surface registration without initial motion estimation. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 430–437, june 2010. 1, 4, 5

[2] A. Albarelli, E. Rodolà, and A. Torsello. Imposing Semi-Local geometric constraints for accurate correspondences selection in structure from motion: A Game-Theoretic perspective. *International Journal of Computer Vision*, pages 1–18, Mar. 2011. 1, 4, 5

[3] A. M. Bronstein, M. M. Bronstein, U. Castellani, A. Dubrovina, L. J. Guibas, R. P. Horaud, R. Kimmel, D. Knossow, E. von Lavante, D. Mateus, M. Ovsjanikov, and A. Sharma. SHREC 2010: Robust correspondence benchmark. In *Proc. EUROGRAPHICS Workshop on 3D Object Retrieval*, 2010. 6, 7

[4] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proc. National Academy of Science (PNAS)*, 103(5):1168–1172, 2006. 1, 3, 7

[5] M. M. Bronstein and I. Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 1704–1711, 2010. 2, 4

[6] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21:5–30, July 2006. 2

[7] A. Elad and R. Kimmel. On bending invariant signatures for surfaces. *IEEE Trans. Pattern Analysis and Machine Intelligence*, pages 1285–1311, 2003. 1

[8] S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:377–388, 1996. 1, 5

[9] V. G. Kim, Y. Lipman, and T. Funkhouser. Blended intrinsic maps. *Trans. on Graphics (Proc. of SIGGRAPH)*, 2011. 5, 6

[10] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 1482–1489. IEEE, 2005. 1, 6

[11] F. Mémoli. Spectral Gromov-Wasserstein distances for shape matching. In *Proc. NORDIA*, 2009. 1, 3, 4

[12] F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 5:313–346, 2005. 1

[13] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, pages 849–856, 2001. 5

[14] W. Ni, N.-S. Vu, and A. Caplier. An online three-stage method for facial point localization. In *Proc. Computer Analysis of Images and Patterns*, pages 57–64. 2011. 1

[15] M. Ovsjanikov, Q.-X. Huang, and L. J. Guibas. A condition number for non-rigid shape matching. *Comput. Graph. Forum*, 30(5):1503–1512, 2011. 1

[16] S. Rota Bulò and I. M. Bomze. Infection and immunization: A new class of evolutionary game dynamics. *Games and Economic Behavior*, 71(1):193–211, January 2011. 5

[17] J. Sun, M. Ovsjanikov, and L. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Proc. of the Symposium on Geometry Processing*, pages 1383–1392. Eurographics Association, 2009. 2, 7

[18] N. Thorstensen and R. Keriven. Non-rigid shape matching using geometry and photometry. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2009. 1

[19] C. Wang, M. M. Bronstein, A. M. Bronstein, and N. Paragios. Discrete minimum distortion correspondence problems for non-rigid shape matching. In *Proc. Scale Space and Variational Methods*, 2011. 5

[20] J. Weibull. *Evolutionary Game Theory*. MIT Press, 1995. 5