Elastic Net Constraints for Shape Matching

Emanuele Rodolà
The University of Tokyo, TU Munich
rodola@in.tum.de

Andrea Torsello
Università Ca' Foscari Venezia
torsello@dais.unive.it

Tatsuya Harada
The University of Tokyo
harada@mi.t.u-tokyo.ac.jp

Yasuo Kuniyoshi
The University of Tokyo
kuniyosh@isi.imi.i.u-tokyo.ac.jp

Daniel Cremers
TU Munich
cremers@tum.de

Abstract

We consider a parametrized relaxation of the widely adopted quadratic assignment problem (QAP) formulation for minimum distortion correspondence between deformable shapes. In order to control the accuracy/sparsity trade-off we introduce a weighting parameter on the combination of two existing relaxations, namely spectral and game-theoretic. This leads to the introduction of the elastic net penalty function into shape matching problems. In combination with an efficient algorithm to project onto the elastic net ball, we obtain an approach for deformable shape matching with controllable sparsity. Experiments on a standard benchmark confirm the effectiveness of the approach.

1. Introduction

Shape matching is a pervasive problem in computer vision and arises in several different fields ranging from robotics to medical imaging. In its most typical form, it concerns the problem of determining a map $f : X \to Y$ among two given shapes in such a way that their geometrical properties are preserved by the transformation. A particularly challenging instance of this problem occurs when the two shapes undergo general non-rigid deformations. As such, matching of deformable shapes has attracted the interest of researchers during the years and a wide variety of approaches have been proposed (see, *e.g*. [2] and references therein for a recent comparison).

A prominent approach to the matching problem from a metric perspective was introduced in [12], a concept that was explored further in [3] with the introduction of the GMDS framework, where the minimum distortion isometric embedding of one surface onto another is explicitly sought. A different view on the problem stems from the notion of uniformization space [9, 17]. Lipman and Funkhouser [9] proposed to model deviations from isometry by a transportation distance between corresponding points in a canonical domain (the complex plane); the result of this process is a "fuzzy" correspondence matrix, whose values can be given the natural interpretation of confidence levels attributed to each match. This idea of a fuzzy map of assignments is not novel, and can be traced back, for instance, to the softassign method for graph matching [6]. In the specific case of non-rigid shapes, fuzzy schemes are typically adopted to relax the point-to-point mappings [11, 14]. While methods based on uniformization theory are made attractive by the low dimensionality of the embedding domain, they do not behave well with different kinds of deformations (*e.g*., topological changes), and are subject to global inconsistencies in the final mapping. More recently, Windheuser *et al*. [16] gave a linear programming relaxation to the matching problem; the method notably allows to obtain continuous correspondences, but it is sensitive to topological changes and, as noted by the authors, its GPU implementation takes about 2 hours per matching.

In this paper, we consider the widely adopted quadratic assignment problem (QAP) formulation for minimum distortion correspondence between deformable shapes. Notable attempts at relaxing the NP-hard QAP include graduated assignment [6], spectral relaxation [8] and the more recent game-theoretic approach [14]. Motivated by the observation that good accuracy often comes at the price of high sparsity, whereas large cardinality tends to bring distorted matches into the correspondence, we attempt to control the accuracy/sparsity trade-off by introducing a weighting parameter on the combination of two effective relaxations, namely the spectral and game-theoretic techniques, which we relate to their regularizer counterparts from regression analysis. This leads us to the introduction of the *elastic net* penalty function [18] into shape matching problems.

The contributions of this paper are two-fold: First, we provide an interpretation of the correspondence problem from the point of view of regression analysis, yielding a nat-

ural connection between existing approaches and well established regularization techniques. We introduce the family of elastic net constraints into a relaxed QAP formulation, and show how previous relaxation attempts naturally constitute special cases of our formulation. Experiments on a standard benchmark demonstrate densifying behavior while maintaining at the same time high accuracy of the correspondence. Second, we give a solution to the projection problem onto the new set of constraints from the viewpoint of variable selection, giving rise to an especially simple and efficient projection algorithm.

2. Minimum distortion correspondence

We model shapes as compact Riemannian manifolds endowed with an intrinsic metric d. A point-to-point correspondence between two shapes X and Y can be defined as a binary function $c : X \times Y \to \{0, 1\}$ satisfying the mapping constraints

$$\sum_{x \in X} c(x, y) \leq 1 \,, \quad \sum_{y \in Y} c(x, y) \leq 1 \,, \tag{1}$$

for every $x \in X$ and $y \in Y$. Note that these constraints ensure that every point in one shape has *at most* one corresponding point in the other (and vice versa), thus allowing the two shapes to have different size. In the following, we will slightly abuse nomenclature and equivalently refer to the correspondence as the respective collection of matches, that is, the set of pairs $C \subset X \times Y$ for which $c(x, y) \neq 0$.

In order to give a measure of quality to the correspondence, we evaluate the distortion induced by the mapping as measured on the two shapes using the respective metrics d_X and d_Y. In particular, given two matches $(x, y), (x', y') \in C$, the absolute criterion

$$\epsilon(x, y, x', y') = |d_X(x, x') - d_Y(y, y')| \tag{2}$$

directly quantifies to which extent the estimated correspondence deviates from isometry. Following [11, 14], we first relax the correspondence from a discrete to a fuzzy notion by letting $c : X \times Y \to [0, 1]$, effectively setting off the problem from its combinatorial nature and bringing it to a continuous optimization domain. Further, we adopt the so called Gromov-Wasserstein [11] family of metrics, which give rise to a *relaxed* notion of proximity between shapes:

$$D(X, Y) = \frac{1}{2} \min_{C} \sum_{(x,y),(x',y') \in C} \epsilon^p(x, y, x', y') c(x, y) c(x', y') \,. \tag{3}$$

Establishing a minimum distortion correspondence between the two shapes amounts to finding a minimizer of the above distance. To this end, note that the problem can be easily recast as a relaxed QAP,

$$\begin{aligned} \min_{\mathbf{C}} \quad & \mathrm{vec}\{\mathbf{C}\}^T \mathbf{A}\, \mathrm{vec}\{\mathbf{C}\} \\ \text{s.t.} \quad & \mathbf{C}\mathbf{1} \preceq 1,\ \mathbf{C}^T\mathbf{1} \preceq 1,\ \mathbf{C} \succeq 0 \,, \end{aligned} \tag{4}$$

where $\mathrm{vec}\{\mathbf{C}\}$ is the $|C|$-dimensional column-stack vector representation of the correspondence matrix $\mathbf{C}$, $\mathbf{A}$ is a non-negative symmetric cost matrix containing the pairwise distortion terms that appear in (3), $\mathbf{1}$ is a vector of $n = |C|$ ones, and $\succeq$ denotes element-wise inequality. Note that in the standard QAP, function c is taken to be a binary correspondence and the mapping constraints (1) hold with equality (requiring $\mathbf{C}$ to be a permutation matrix).

In the following, we present two existing approaches that relax the mapping constraints in (4) to find a minimum distortion correspondence. Even though originating from distinct motivations, the two methods share a convenient interpretation as partitioning problems in the space of potential assignments. Their introduction here is useful for the construction we will present in Section 3.

2.1. Spectral matching

Taking the point of view of graph clustering, [8] proposed the simplified problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} \\ \text{s.t.} \quad & \|\mathbf{x}\|_2^2 = 1 \,, \end{aligned} \tag{5}$$

where $\mathbf{x} \equiv \mathrm{vec}\{\mathbf{C}\} \in \mathbb{R}^n$ is the vector representation for the correspondence. Following Rayleigh's quotient theorem, this modified QAP is minimized by the eigenvector $\mathbf{x}^\star$ corresponding to the minimum eigenvalue of $\mathbf{A}$. Note that mapping constraints are not imposed in (5). The authors follow a greedy algorithm to impose such constraints only after a solution has been obtained. The method has a tendency to produce matches for each point. This makes it of limited use in real settings, where shapes may undergo partiality deformations. Further, symmetries and structured noise in the data (indeed a characteristic of the non-rigid setting) may lead to unstable eigenvectors [8] and thus unreliable assignments. The spectral method has been applied for non-rigid matching in [13].

The L^2 relaxation to the QAP was introduced mainly because it allows to obtain a global solution to the modified problem in closed form. Here we give another interpretation of this approach as a relaxed two-way partitioning problem [1]. Consider the set of constraints taking the form $\mathbf{x}_i^2 = 1$ for $i = 1 \dots n$; these constraints restrict the values of $\mathbf{x}_i$ to ± 1, so the problem is equivalent to finding the partitioning (as "match" or "non-match") on a set of n elements that minimizes the total cost $\mathbf{x}^T \mathbf{A} \mathbf{x}$. Here, the coefficients $\mathbf{A}_{ij}$ can be interpreted as the cost of having elements i and j in the same partition. Clearly, the new constraints imply $\sum_{i=1}^n \mathbf{x}_i^2 = \|\mathbf{x}\|_2^2 = n$; since this actually allows the $\mathbf{x}_i$ to take on any (small enough) real number, optimizing over this feasible set will yield a *lower bound* on the optimal value of the original partitioning problem.

2.2. Game-theoretic matching

Given the inherent difficulty to solve for a minimum distortion correspondence under general deformations, in a re-

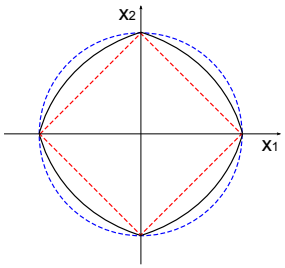

Figure 1. Contour plots of the L^2 (circle), L^1 (diamond), and elastic net (in between) balls in $\mathbb{R}^2$. In this example we set $\alpha = 0.6$. The strength of convexity varies with α.

cent paper [14] we proposed to shift the focus to the search of a maximal group of matches having least distortion, *regardless* of its cardinality. To achieve this, we proposed to optimize over the probability simplex

$$\|\mathbf{x}\|_1 = \mathbf{1}^T\mathbf{C}\mathbf{1} = 1\,,\;\; \mathbf{x} \succeq 0\,. \tag{6}$$

The main benefits of adopting such L^1-type constraint for the matching problem arise from its convenient game-theoretical interpretation, leading to very efficient algorithms for (local) optimization and, most remarkably, in allowing the mapping constraints to be embedded directly into the cost matrix $\mathbf{A}$. Unfortunately, the strong selectivity demonstrated by the game-theoretic approach is hardly desirable for matching problems. While existing algorithms may be applied to densify the few obtained matches, in practice the high spatial locality of the final correspondence does not allow to properly constrain such densification methods [17, 13].

Similarly to the L^2 case, the game-theoretic approach can be regarded as an attempt to solve a partitioning problem where the two partitions are represented by $\mathbf{x}_i = 0$ or 1 for $i = 1 \dots n$. This, in turn, corresponds to imposing a bound on the "counting" norm $\|\mathbf{x}\|_0$, which is relaxed here to the continuous sparsity-inducing counterpart $\sum_{i=1}^{n} |\mathbf{x}_i| = \|\mathbf{x}\|_1 = n$, with $\mathbf{x}_i \geq 0$ for all i.

3. Matching with the elastic net

Both methods presented in the previous section are virtually free from parameters, but their performance directly depends on the specific definition of the distortion function ϵ. It is very difficult, in practice, to give a definition for ϵ that works well for any given pair of shapes. This is, in fact, a difficulty shared by *any* method attempting to minimize (4). Ovsjanikov *et al.* [13] recently introduced the notion of shape condition number. According to this notion, the stability of the matching can be characterized as an intrinsic property of the shape itself, and is related to its intrinsic symmetries as well as the specific choice of a metric.

In order to incorporate a somewhat elusive notion of stability into the matching process, we propose to change the point of view by drawing an analogy between the correspondence problem and model-fitting. Our goal, in this context, is to determine a good approximation of the true relationship between the two shapes: we seek to *fit* or approximate the optimal correspondence $\mathbf{x}^\star$ as closely as possible, with deviation measured in the Gromov-Wasserstein distance, *i.e.*, in the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$. Problems of this kind are often studied with the tools of regression analysis [1]. Here the interest shifts from finding a best fit to analyzing the relationships among the several variables that build up the set of potential assignments $\{\mathbf{x}_i\}_{i=1\dots n}$. These candidate matches act as predictors for the minimum distortion correspondence, and can be given the interpretation of explanatory variables which we observe, while we seek to find the combination that best describes the data in the minimal distortion sense. Since in general these variables hold a certain degree of correlation among them, it is of particular interest to attempt to determine whole groups of highly correlated predictors, as they will likely form consistent groups of matches in terms of the adopted measure of distortion.

In this view, spectral matching can be directly related to ridge regression, whose L^2 penalty is known to generally improve conditioning of the problem, yet always keeping all the predictors in the model. Similarly, the game-theoretic technique finds its equivalent in the lasso, the sparsity-inducing L^1 regularizer performing continuous shrinkage and automatic variable selection simultaneously [1, 18]; one major limitation of the lasso is its tendency to select only one variable from a group of variables among which the pairwise correlations are very high. While none of the two methods dominates the other in all circumstances, both have appealing features. Our aim is to strike a balance between the two. To this end, we adopt a family of constraints known as elastic net [18]. This regularization technique shares with the lasso the ideal property of performing automatic variable selection, and most notably it is able to select entire groups of highly correlated variables. The elastic net criterion is defined as a convex combination of the lasso and ridge penalties:

$$(1-\alpha)\|\mathbf{x}\|_1 + \alpha\|\mathbf{x}\|_2^2\,,\;\; \alpha \in [0,1]\,. \tag{7}$$

It becomes ridge regression for $\alpha = 1$, and the lasso for $\alpha = 0$. This penalty function is singular at 0 and *strictly* convex (differently from the lasso) for $\alpha > 0$, thus possessing the characteristics of both penalties (see Fig. 1).

Strict convexity plays an important role as it guarantees the grouping effect in the extreme situation with identical predictors (that is, whenever the distortion between two matches is exactly 0), and provides a quantitative description of their degree of correlation (proportional in our case to the deviation from isometry) otherwise. Let $\mathbf{x} \in \mathbb{R}^{|C|}$ be the vector representation of some correspondence $C \subset X \times Y$, we expect the elastic net-penalized so-

lution to keep the difference $|\mathbf{x}_i - \mathbf{x}_j|$ small whenever the metric distortion $\epsilon(C_i, C_j)$ between the two matches is small. The trade-off between size of the correspondence and matching error is regulated by the convexity parameter α, which allows to fine tune the model complexity and balance the action of the penalty ranging from the highly selective pure lasso for $\alpha = 0$ to the more tolerant ridge behavior for $\alpha = 1$. This leads to the following family of relaxations for the QAP:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} \\ \text{s.t.} \quad & (1-\alpha)\|\mathbf{x}\|_1 + \alpha \|\mathbf{x}\|_2^2 = 1,\ \mathbf{x} \succeq 0\,, \end{aligned} \tag{8}$$

with $\alpha \in [0, 1]$. The family directly generalizes the spectral and game-theoretic techniques. Similarly to the spectral approach, this formulation does not guarantee the final solution to represent a bijective mapping, which can nevertheless be efficiently obtained *a posteriori* using, for instance, the same greedy technique of [8]. Note that this final step is only performed to emphasize the generalization property of our formulation, as it will transition smoothly from a sparse behavior equivalent to [14] to one similar to [8].

3.1. Optimization

We undertake a projected gradient approach [1] to determine a local optimum for problem (8). The optimization process is governed by the equations

$$\mathbf{x}^{(t+1)} = \Pi\left(\mathbf{x}^{(t)} - \gamma^{(t)} \mathbf{A}\mathbf{x}^{(t)}\right)\,, \tag{9}$$

where $\mathbf{A}\mathbf{x} = \frac{1}{2}\nabla \mathbf{x}^T \mathbf{A} \mathbf{x}$ is a descent direction for the objective, $\gamma > 0$ is the step length taken in that direction, and $\Pi : \mathbb{R}^n \to \mathbb{R}^n$ is a projection operator taking a solution back onto the feasible set.

While efficient methods for projecting onto the L^2 and L^1 balls have been proposed in literature [15], projection onto their convex combination is a more involved task. A detailed explanation of our approach on the computation of Π is deferred to the next Section; nevertheless, we anticipate here that this projection step can be performed in a very efficient manner. This allows us to determine the optimal step size in (9) at each iteration by performing exact line search [1] along the ray $\{\mathbf{x} + \gamma \mathbf{A}\mathbf{x} : \gamma \geq 0\}$ through the application of Newton's method [1], a quadratic fitting algorithm having order two convergence.

Finally, we initialize $\mathbf{x}^{(0)}$ to the barycenter of the elastic net boundary, *i.e.*, for all $i = 1 \ldots n$ we set $\mathbf{x}_i$ to the positive solution of the quadratic equation $\alpha n x^2 + (1-\alpha) n x - 1 = 0$.

3.2. Projection onto the elastic net ball

Computing the Euclidean projection $\Pi(\mathbf{x}_0)$ onto the (positive) elastic net ball boundary amounts to solving the following projection problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x} - \mathbf{x}_0\|_2^2 \\ \text{s.t.} \quad & (1-\alpha)\mathbf{1}^T \mathbf{x} + \alpha \mathbf{x}^T \mathbf{x} = t,\ \mathbf{x} \succeq 0\,, \end{aligned} \tag{10}$$

with $\alpha \in [0, 1]$. Efficient attempts at solving this problem arose only recently [10, 7]. Gong *et al*. [7] formulate the minimization as a root finding problem for a piecewise continuous function. While the method allows to obtain a solution in linear time, the procedure is highly susceptible to numerical errors; these errors are exacerbated when the dimension of the projected vector is high, severely limiting the applicability of the method in several practical settings. Further, the approach draws its major benefits from sparse projected vectors. However, in sparse matching problems vectors tend to be dense, and in this situation the method does not perform as efficiently. Mairal *et al*. [10] propose a linear time projection algorithm based on randomized median search; the algorithm is numerically stable, but it is outperformed by root finding for low dimensions [7].

One major disadvantage of the existing methods lies in their inherently sequential nature. With the advancement of computational technologies and the applications they enable, it has become necessary to provide algorithms that exploit this computational power and scale up with the increased dimensionality of real-world problems. Both methods presented above are unable to address the need. To this end, we take a different point of view and regard the projection problem as one of coordinate selection. Our method is not susceptible to numerical errors, it is at least as efficient as existing methods in its basic form and, most importantly, it easily lends itself to a parallel implementation.

It is immediate to see that the solution to (10) lies in the intersection of two convex sets: a sphere of equation $(1-\alpha)\mathbf{1}^T\mathbf{x} + \alpha \mathbf{x}^T\mathbf{x} = t$ (which we denote by $\mathcal{C}_1$), and the non-negative cone ($\mathcal{C}_2$). Projection onto the intersection of convex sets has been extensively studied in the past; of particular relevance is a result which can be traced back to Dykstra [5], an iterative technique usually referred to as the method of alternating projections.

We determine a closed form projection onto $\mathcal{C}_1$ as follows. Disregarding the non-negativity constraints and introducing Lagrange multiplier $\lambda \in \mathbb{R}$, we obtain the Lagrangian associated to problem (10)

$$L(\mathbf{x}, \lambda) = \|\mathbf{x} - \mathbf{x}_0\|_2^2 + \lambda\left[(1-\alpha)\mathbf{1}^T\mathbf{x} + \alpha \mathbf{x}^T\mathbf{x} - t\right]. \tag{11}$$

The KKT optimality conditions require that the gradient of $L(\mathbf{x}, \lambda)$ with respect to $\mathbf{x}$ vanish at the optimum, *i.e.*,

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) = 2(1 + \lambda\alpha)\mathbf{x} - 2\mathbf{x}_0 + \lambda(1-\alpha)\mathbf{1} = 0, \tag{12}$$

from which we obtain the optimal $\mathbf{x}$ as

$$\mathbf{x} = \frac{\mathbf{x}_0 - \lambda\frac{1-\alpha}{2}\mathbf{1}}{1 + \lambda\alpha}\,. \tag{13}$$

Determining the optimal value for λ is straightforward; imposing the elastic net constraints on $\mathbf{x}$, it must hold

$$\alpha \sum_{i=1}^{n} \left(\frac{\mathbf{x}_{0i} - \lambda\frac{1-\alpha}{2}}{1 + \lambda\alpha}\right)^2 + (1-\alpha)\sum_{i=1}^{n} \frac{\mathbf{x}_{0i} - \lambda\frac{1-\alpha}{2}}{1 + \lambda\alpha} = t. \tag{14}$$

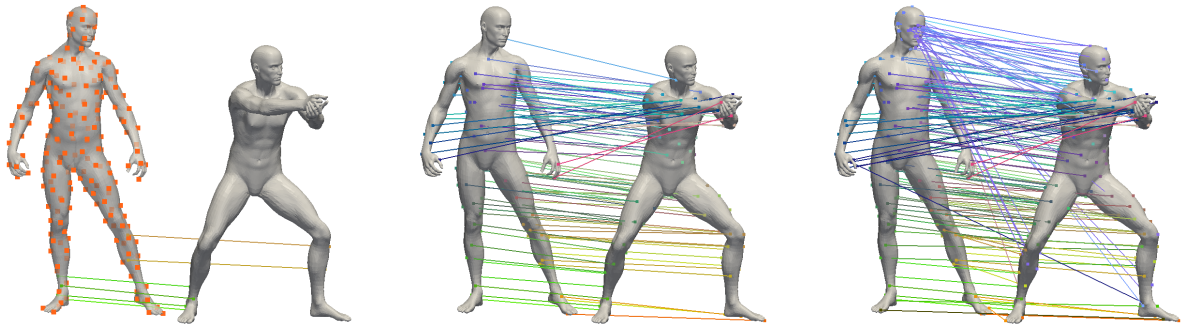

Figure 2. Example of matchings obtained with the game-theoretic, elastic net and spectral techniques. The set of potential assignments is constructed by taking ∼200 farthest points on one shape (shown in the left image for reference), and then building the whole Cartesian product with the *correct* corresponding points from the other shape, after 45% of them have been moved to random positions over the surface. This setup simulates a moderately challenging scenario in which only ∼50% of the shape is matchable with low distortion, and the feasible set comprises all possible assignments between the two shapes. The game-theoretic (L^1) solution is highly selective and only assigns 3% of the shape samples, with geodesic error 3.12 (left image); in contrast, the spectral (L^2) approach favors dense solutions and yields matches for 93% of the points, with total error 62.53 (right image). Elastic net matching (middle) allows to regulate the trade-off between size and distortion: the correspondence is made more dense, and 53% of the points are matched while keeping the error at 15.87 (compare with Table 1). Here we set $\alpha = 0.85$.

From this we obtain a quadratic equation in λ; substituting the λ obtained into Eq. (13) we get the desired projection.

Projection onto the constraint set $\mathcal{C}_1 \cap \mathcal{C}_2$ from problem (10) can be obtained in an iterative manner according to the alternating projections scheme; however, the method of alternated projections is mainly of theoretical importance, while empirical evidence often reveals slow convergence rate and poor scalability. In the following we present a novel iterative approach, in which the original minimum-distance problem is modified with the introduction of a selection term on the optimization variables; the proposed formulation is rather general, and as such it provides a heuristic to a broader class of projection problems.

Let $\mathbf{e} \in \{0,1\}^n$ be an indicator vector used to select inactive coordinates for $\mathcal{C}_2$; consider the minimization problem

$$\begin{aligned} \min_{\mathbf{x},\mathbf{e}} \quad & \|(\mathbf{x}-\mathbf{x}_0)\circ\mathbf{e}\|_2^2 \\ \text{s.t.} \quad & (1-\alpha)\mathbf{x}^T\mathbf{e} + \alpha(\mathbf{x}\circ\mathbf{e})^T(\mathbf{x}\circ\mathbf{e}) = t \\ & \mathbf{e} \in \{0,1\}^n, \end{aligned} \tag{15}$$

where $\circ$ denotes the Hadamard product among two vectors. Note that problems (10) and (15) are equivalent, since they have the same minimizers; however, the addition of the indicator variable allows us to devise an iterative procedure which converges efficiently to the global optimum.

Let us assume we are given an initial guess for vector $\mathbf{e}$, for instance $\mathbf{e}_i^{(0)} = 1$ for all $i = 1 \dots n$. Ruling out $\mathbf{e}$ from the optimization variables, and following similar derivations as in Eq. (11)-(13), we get the solution

$$\mathbf{x} = \frac{\bar{\mathbf{x}}_0 - \lambda(\mathbf{e})\frac{1-\alpha}{2}\mathbf{e}}{1+\lambda(\mathbf{e})\alpha}, \tag{16}$$

where $\bar{\mathbf{x}}_0 = \mathbf{x}_0 \circ \mathbf{e}$, and λ is obtained in a similar manner to Eq. (14) by solving the resulting quadratic equation:

$$\begin{aligned} & c - b\lambda - \frac{\alpha}{2}b\lambda^2 = 0\,, \quad \text{with} \\ & c \equiv (1-\alpha)\mathbf{x}_0^T\mathbf{e} + \alpha(\mathbf{x}_0\circ\mathbf{e})^T(\mathbf{x}_0\circ\mathbf{e}) - t\,, \\ & b \equiv \frac{\mathbf{e}^T\mathbf{e}}{2}(1-\alpha)^2 + 2t\alpha\,. \end{aligned} \tag{17}$$

Eq. (17) always admits two real solutions, only one of which gives the correct projection (we omit the complete derivations for space reasons).

Note that no positivity constraints are imposed on this solution. The key step now consists in deselecting those coordinates getting a non-positive value after an unconstrained projection takes place: this gives us a rule for updating the indicator vector in such a way that projecting again with the new $\mathbf{e}$ will directly put those variables to zero. This leads to an iterative scheme which converges to the unique minimum-distance projection.

Lemma 1. *Let $\mathbf{x}^{(0)} = \mathbf{x}_0$, and $\mathbf{e}^{(0)} = \mathbf{1}$ (a vector of n ones). Then projection problem (10) is minimized by application of the iterative rules*

$$\mathbf{x}^{(t+1)} = \frac{\mathbf{x}^{(0)} - \lambda(\mathbf{e}^{(t)})\frac{1-\alpha}{2}\mathbf{e}^{(t)}}{1+\lambda(\mathbf{e}^{(t)})\alpha} \tag{18}$$

$$\mathbf{e}_i^{(t+1)} = \begin{cases} 0 & \text{if } \mathbf{x}_i^{(t+1)} \le 0 \\ \mathbf{e}_i^{(t)} & \text{otherwise,} \end{cases}, \tag{19}$$

where $\lambda(\mathbf{e}^{(t)})$ is a positive solution to the second order equation (17).

The process terminates when it reaches the fixed point $\mathbf{e}^{(t+1)} = \mathbf{e}^{(t)}$ (equivalently when $\mathbf{x}^{(t+1)} \succeq 0$).

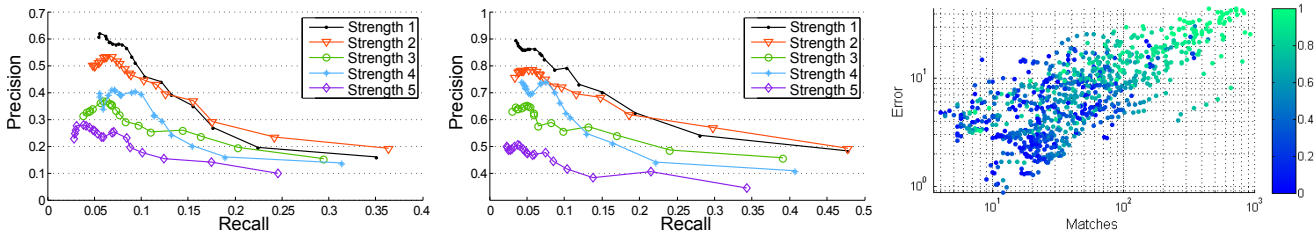

Figure 3. Precision/recall graphs for α ranging over $[0, 1]$, computed with tolerance radii $r = 1$ (shape matching) and $r = 10$ (shape retrieval). On the right, a scatter plot of error vs number of matches over the whole dataset. Color encodes the value of α.

Sketch of proof. In [15] it is shown that the L^1 projection is monotonic, *i.e.*,

$$\mathbf{x}_{0i} > \mathbf{x}_{0j} \Rightarrow \mathbf{x}_i = 0 \rightarrow \mathbf{x}_j = 0\,. \tag{20}$$

Using an analogous derivation we can show that the same holds for the elastic net projection. Further, projection (16) is also monotonic in the active selected coordinates, *i.e.*, those for which $\mathbf{e}_i = 1$. Due to this monotonicity of the projection the entries of $\mathbf{e}$ are eliminated in coordinate rank order, thus the algorithm never eliminates a coordinate that has a non-zero value in the optimal projection and converges in at most k steps, where k is the number of zero entries in the final vector. □

3.3. Performance of the projection

We compared the projection time against other projection methods onto the elastic net, namely piecewise root finding [7], randomized median finding [10], our method (coordinate selection), and its parallel version. The parallel version was obtained compiling the code with OpenMP and using 8 cores on an Intel Core i7. For these comparisons, we generated random vectors of varying size and ran each projection method on the same input. The experiment was repeated 100 times per size, and projection times accumulated and plotted for each method. Fig. 4 show a logarithmic-scale plot of the measured times. While all methods demonstrate linear complexity, we observed that the root finding method produced suboptimal solutions for sizes larger than 10^4. The coordinate selection approach converged up to 4 times faster than previous methods in its single-core implementation, and two orders of magnitude faster in its parallel version, showing the almost linear speedup with the number of cores due to the intrinsically parallel nature of the algorithm. Arguably we can expect even faster speedup with a GPU implementation.

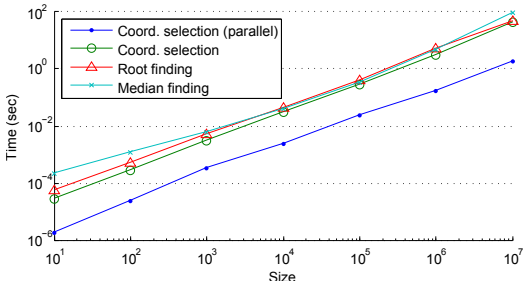

Figure 4. Comparison of projection methods for the elastic net.

4. Experimental results

We performed a wide range of experiments on the SHREC'10 standard dataset [2]. The dataset consists of 3 models of varying resolution (10K-50K points) undergoing 9 different types of deformation (listed in Table 1), each appearing in 5 intensity levels. We measure the matching error of a correspondence $C \subset X \times Y$ as its average geodesic distance from the ground-truth C_g, taking into account possible intrinsic symmetries, as in [2]:

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

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

In order to make the computational task more tractable, only a limited number of samples ($m = 1000$ in our experiments) are considered from one shape, and then potential matches are built with the 5 closest points (in descriptor space) in the other. The descriptor adopted for this step is the scale invariant HKS [4]. Samples are generated via farthest point sampling (FPS) [12, 11] using the extrinsic Euclidean metric, since it gives for large m a good and efficient approximation to the intrinsic measures while being more robust to topological and partiality deformations. Fig. 2 (left) shows an example of FPS with $m = 190$. We emphasize that this step is performed to simplify computations without sacrificing precision, and the goal is not to perform feature detection for finding repeatable interest points. Note also that in the matching process only one of the two shapes is subsampled, while we keep all points in the other.

Finally, for the distortion terms ϵ^p in (3) we set $p = 2$, and adopt for d_X, d_Y the multi-scale diffusion metrics of [14], using time scales $(2^7, \ldots, 2^{16})$.

4.1. Trade-off analysis

Elastic net matching can be seen as an instance of multicriterion optimization, in which we wish to maximize the size of the correspondence while simultaneously bringing

Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	**8.55**	**4.11**	4.90	**6.40**	14.06	$18.25/212$	$22.04/149$	$14.42/212$	$7.25/210$	$18.50/76$
Topology	5.97	9.22	**2.88**	**5.62**	8.58	$19.18/313$	$18.91/321$	$19.44/117$	$18.44/173$	$17.33/99$
Holes	**7.51**	**14.03**	**6.50**	25.85	**8.75**	$19.64/197$	$17.68/278$	$14.99/62$	$17.46/31$	$6.69/10$
Micro holes	**4.06**	8.63	11.36	13.66	**2.98**	$19.35/193$	$16.65/367$	$17.50/170$	$17.48/183$	$9.96/65$
Scale	**4.70**	**2.40**	**1.67**	**4.49**	3.34	$4.23/6$	$17.72/200$	$1.33/10$	$3.38/9$	$2.62/5$
Local scale	**6.54**	**10.89**	**3.47**	59.16	**3.69**	$19.86/178$	$17.42/302$	$19.12/178$	$55.18/19$	$11.19/45$
Sampling	10.31	**9.00**	**6.39**	**9.36**	**8.28**	$16.05/211$	$16.83/374$	$18.63/166$	$14.93/164$	$17.27/38$
Noise	11.20	8.62	53.99	54.96	**5.70**	$15.31/54$	$12.72/218$	$54.03/9$	$69.33/327$	$4.84/4$
Shot noise	7.36	**10.66**	**3.20**	15.95	**3.25**	$16.28/136$	$18.01/380$	$16.05/81$	$15.95/64$	$13.95/54$
Average	**7.36**	**8.62**	**10.49**	21.72	**6.51**	$16.46/167$	$17.55/288$	$19.50/112$	$24.38/131$	$11.37/44$

Table 1. First table: Matching results obtained with $\alpha = 0.65$. Average number of corresponding points is 50. Values in bold indicate better performance than *both* GMDS and the (reiterated) game-theoretic method for the same number of matches. Second table: Matching results (error/matches) obtained by selecting for each case the value of α giving the largest possible error below 20.

the matching distortion to zero. In most cases, it is very difficult to satisfy these competing criteria exactly. Fig. 2 presents an example in which the correct matches have a very small inlier ratio with respect to the set of candidates (a full Cartesian product in this case). In this matching scenario, our method provides a means to select only high-precision correspondences in a situation where there is huge ambiguity in most correspondences. In this view, we might naturally ask how much we must pay in terms of distortion in order to obtain an increase in the number of matches.

Solutions expressing this trade-off are called Pareto optimal [1], and can be visualized as points in a plane whose dimensions correspond to the two scalar objectives. In Fig. 3 (right) we plot each shape-to-shape correspondence over the whole dataset as a point in this plane; for each matching instance, we vary the convexity coefficient α from 0 to 1 (shown by color). The figure suggests that this parameter allows to regulate the compromise between the two criteria, while maintaining a comparable error variance (note that the plot is in log-log scale so we expect an expansion at the lower end of the scale for a fixed variance).

Fig. 5 plots per-class Pareto curves, averaged over all strengths and models. Each curve consists of 21 points, corresponding to as many equally spaced values for α; in particular, the endpoints correspond to the game-theoretic ($\alpha = 0$) and spectral ($\alpha = 1$) solutions. In most cases, there is a point of large curvature where a small increase in the number of matches can only be accomplished by a large increase in matching error. This is the proverbial "knee" of the trade-off curve [1], and can be taken as a good compromise solution; it will correspond, in general, to a different value of α for each shape, which is indeed consistent with the idea of shape condition number (Section 3). The same figure reveals another interesting picture. First, the small loops appearing in the plot indicate that matches could actually be "lost" as α is pushed from 0 to 1; this fact is justified by the presence of rather challenging shapes in the dataset (and the specific choice of the intrinsic metric), which render the Gromov-Wasserstein distance unstable, thus leading to different local optima as the set of constraints is modified slightly. Second, it is evident that smaller values of α do not necessarily lead to better solutions in terms of metric distortion; in particular, the game-theoretic solution will not always be the best choice, even when a sparse solution is being sought. This particular point is also made clearer in Fig. 3 (see the first samples of the first two graphs).

Table 1 (last five columns) reports the results of our method on the whole dataset, averaged over the 3 models. The reported values are geodesic error and number of matches. The aim of this table is not to compare against other methods, but rather to show that elastic net regularization allows to obtain denser correspondences while maintaining a small distortion; thus, in this case, the value for α is selected on a per-case basis as the one keeping the error in line with the state of the art on the same dataset [2]. In particular, this choice gives on average $\sim$4 times more matches than GMDS, and $\sim$20 times more than sparse approaches with the same error.

Finally, Fig. 3 shows precision-recall graphs as α is varied in $[0, 1]$. The five curves correspond to different deformation strengths, averaged over all models and deformations. A match is defined as true positive if it lies within a geodesic radius of $r = 1$ from the ground-truth; likewise, a

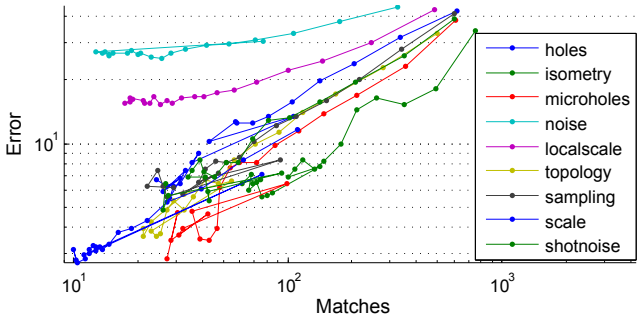

Figure 5. Error versus number of matches, for α varying in $[0, 1]$.

false positive is a match with error larger than r, and false negatives are low-error matches in the candidate set which were not included in the final solution. Note that the small recall we obtain in both graphs does not contradict the fact that for large α we should obtain dense correspondences, as it is due to a filtering step we perform on those matches having a very small value for $\mathbf{x}_i$ (below 10% of the median).

4.2. Comparisons

The adoption of a standard dataset allows us to compare our method directly with other state of the art techniques. Table 1 (first five columns) shows comparisons for the entire dataset averaged over the 3 models, following the same format of previous reports [2, 14]. Since we are able to regulate the size of the final correspondence, we fix the final number of matches to 50 (allowing us to compare directly with methods giving the same number of matches) and determine the (unique) corresponding value of α yielding, on average, the fixed number of matches over the whole dataset. The entries in bold represent values that are better than what is reported in the literature for similar number of correspondences. We can see that the proposed approach is in line and exceeds, on average, the state of the art.

In all our experiments we observed 30-100 gradient descent iterations per matching, choosing as stopping criterion the relative change in the objective value. This amounted to $\sim$30 seconds per matching on average, with most of the variance due to the specific value used for α. In fact, since for small values it yields sparser solutions, the projection step needs more iterations to converge. For a vector of 10,000 elements, we need on average 8 iterations with $\alpha = 0.01$, and 2 iterations with $\alpha = 0.99$. As a reference, the method of alternating projections for the same vector takes up to 1800 iterations to reach a solution. Finally, adopting the parallel version of the projection algorithm into the optimization process lowered the overall convergence time to $\sim$5 seconds per matching.

5. Conclusions

In this paper, we proposed the adoption of the elastic net family of constraints as regularizers for the quadratic assignment problem, which frequently arises in deformable shape matching problems. The approach naturally generalizes existing techniques. It allows to regulate the relative contribution of distortion and size of the correspondence via a single convexity parameter. We provided an efficient and provably optimal solution to the projection problem onto the new set of constraints, and demonstrated on a standard benchmark how the method allows to obtain sparse-to-dense solutions with an accuracy at least as good as the state of the art.

Acknowledgments

The authors wish to thank Asako Kanezaki for helpful discussions. This work was supported by the Japan Society for the Promotion of Science under fellowship program FY2012, and the ERC Starting Grant "ConvexVision".

References

[1] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge Univ. Press, New York, USA, 2004.

[2] A. Bronstein, M. Bronstein, U. Castellani, et al. Shrec 2010: Robust correspondence benchmark. In *Eurographics Workshop on 3D Object Retrieval*, 2010.

[3] A. Bronstein, 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.

[4] M. Bronstein and I. Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *Proc. CVPR*, pages 1704 –1711, 2010.

[5] J. Dattorro. *Convex Optimization & Euclidean Distance Geometry*. Meboo Publ., Palo Alto, CA, USA, 2005.

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

[7] P. Gong, K. Gai, and C. Zhang. Efficient euclidean projections via piecewise root finding and its application in gradient projection. *Neurocomp.*, 74:2754 – 2766, 2011.

[8] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. In *Proc. CVPR*, volume 2, pages 1482–1489, 2005.

[9] Y. Lipman and T. Funkhouser. Mobius voting for surface correspondence. *ACM Trans. on Graphics*, 28(3), 2009.

[10] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online learning for matrix factorization and sparse coding. *J. Mach. Learn. Res.*, 11:19–60, 2010.

[11] F. Mémoli. Gromov-Wasserstein distances and the metric approach to object matching. *Found. Comput. Math.*, 11:417–487, 2011.

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

[13] M. Ovsjanikov, Q.-X. Huang, and L. J. Guibas. A condition number for non-rigid shape matching. *Comput. Graph. Forum*, pages 1503–1512, 2011.

[14] E. Rodolà, A. Bronstein, A. Albarelli, F. Bergamasco, and A. Torsello. A game-theoretic approach to deformable shape matching. In *Proc. CVPR*, 2012.

[15] S. Shalev-Shwartz and Y. Singer. Efficient learning of label ranking by soft projections onto polyhedra. *J. Mach. Learn. Res.*, 7:1567–1599, Dec. 2006.

[16] T. Windheuser, U. Schlickewei, F. Schmidt, and D. Cremers. Geometrically consistent elastic matching of 3d shapes: A linear programming solution. In *Proc. IEEE Intl. Conf. on Comput. Vis.*, pages 2134 –2141, nov. 2011.

[17] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios. Dense non-rigid surface registration using high-order graph matching. In *Proc. CVPR*, pages 382–389, 2010.

[18] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67:301–320, 2005.