Intrinsic Shape Matching by Planned Landmark Sampling

A. Tevs [1], A. Berner[1], M. Wand[1,2], I. Ihrke[1,2] and H.-P. Seidel[1]

[1]Max-Planck Institut für Informatik, Saarbrücken, Germany
[2]Saarland University, Saarbrücken, Germany

Abstract

Recently, the problem of intrinsic shape matching has received a lot of attention. A number of algorithms have been proposed, among which random-sampling-based techniques have been particularly successful due to their generality and efficiency. We introduce a new sampling-based shape matching algorithm that uses a planning step to find optimized "landmark" points. These points are matched first in order to maximize the information gained and thus minimize the sampling costs. Our approach makes three main contributions: First, the new technique leads to a significant improvement in performance, which we demonstrate on a number of benchmark scenarios. Second, our technique does not require any keypoint detection. This is often a significant limitation for models that do not show sufficient surface features. Third, we examine the actual numerical degrees of freedom of the matching problem for a given piece of geometry. In contrast to previous results, our estimates take into account unprecise geodesics and potentially numerically unfavorable geometry of general topology, giving a more realistic complexity estimate.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Computer Graphics]: Image Processing and Computer Vision—Range data, Registration

1. Introduction

In recent years, the problem of shape matching has gained a lot of interest [ASP*04, MGPG04, GMGP05, HAWG08, RS08]: Given two pieces of geometry, we would like to compute corresponding points on both pieces. A number of variants of this problem exists, such as finding rigid, isometric, conformal, or more general mappings between two surfaces.

In this paper, we consider the problem of approximate intrinsic matching. This means that we are looking for a matching that conserves geodesic distances up to minor errors due to sampling or slight non-isometries in the input. This is a successful and widely adopted model for matching deformable objects such as human or animal body shapes in different poses, where the deformation is naturally very close to isometric. A number of algorithms have been proposed to address this problem [ASP*04, BBK06b, HAWG08, LF09, RS08, SOG09, TBW*09, RPSS10]. The most common approach is to first compute a set of discriminative *feature points* on both shapes along with a local, isometry invariant descriptor and then try to find a matching of these features such that the pairwise geodesic distances between all corresponding pairs of feature points are preserved. Tevs et al. [TBW*09] argue heuristically that a few correct matches are sufficient to fully constrain the problem, and that these can be found using RANSAC. Lipman et al. [LF09] prove that 3 point correspondences are sufficient to fix an isometry on topological discs and spheres, and, very recently, Ovsjanikov et al. [OMMG10] strengthen the result: By analyzing global shape properties encoded in heat-kernel functions, a single

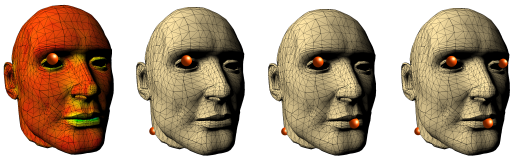


Figure 1: *Landmarks computed for a human face, with one landmark added at each step. The first image shows the distribution of uniqueness (i.e. from blue: most unique, to red: most similar to other) of the first landmark. Further points build up a plan: Each further point is chosen to be of maximal utility for solving the isometry matching problem.*

correspondence is sufficient to fix an isometry if the shape and the point meet certain conditions (implying in particular the absence of continuous sets of self-isometries). These results establish that the intrinsic complexity of finding isometric matches is rather low because the problem does not allow for a large number of degrees of freedom. While these remarkable results solve much of the problem in theory, the practical complexity is still unclear: In real-world scenarios, we have to deal with input data that is undergoing only approximately isometric deformations, there are numerical precision issues, and the intrinsic distances are not reliable due to missing topological information.

For example, if an input shape consists of two pieces connected with a narrow tunnel of very small diameter, fixing intrinsic distances on one piece will in practice not fully constrain the matching on the second piece, even if this was the case if we had a perfectly isometric deformation and infinite numerical resolution. Consequently, the practical implementations of [LF09, OMMG10] employ voting involving multiple correspondence sets in order to obtain stable results in practice. For real-world scenarios, it remains unclear how much correspondence information is necessary to fully constrain the matching problem.

The two questions our paper addresses are the following:

- What is the *practical complexity* of matching two isometrically deformed shapes, assuming a certain amount of uncertainty in intrinsic distance estimates?
- Can we find an algorithm that systematically collects the most relevant correspondence information to find a global shape match as quickly as possible?

Our contributions: To address the problem, we are proposing a new shape matching algorithm based on *planned* random sampling: This means, instead of using a simple random sampling of correspondences as in [TBW*09], we are building a *plan* on how to choose optimized sample points on the source shape that maximize the information gained and therefore minimize the costs for guessing correspondences. Our key idea is to look at the entropy of the posterior distribution of possible matches to assess what are good plan points.

Starting from an empty set of prior knowledge, we incrementally add points that minimize the entropy of the posterior distribution of potential matches and thereby minimize the effort of finding a new piece of information in the next round. A landmark coordinate scheme [HAWG08, TBW*09, OMMG10] is then employed to immediately extend the solution to dense correspondences. The plan can be pre-computed given only the source shape. It can then be reused for matching to a large number of target shapes. This leads to almost real-time matching in situations where a few template shape is matched against a large collection of input data.

Our algorithm leads to significant savings in sampling costs compared to previous random sampling, and gives some insight on the practical complexity of near-isometric shape matching. For many typical example shapes, our plan will actually require a very small set of random samples to obtain a precise, global matching. As a side effect, our new approach also removes the need for any keypoint-detection as a preprocessing step to the matching. In previous work, this has often been a restriction because keypoint detectors are designed for certain data characteristics, such as local bumps in the surface [GMGP05, LG05], that might not be present in sufficient number on some models, thus preventing successful matching. In contrast, our technique fully automatically detects a set of optimized reference points for matching, adapting automatically to the characteristics of the input.

2. Related work

Shape matching: Memoli and Sapiro [MS04] compare two manifolds represented by point clouds using an iterative *Farthest Point Sampling* (FPS) [MD03] algorithm which computes an optimized covering by minimizing an approximate Gromov-Hausdorff-Distance between source and target. In contrast to their work, we first create a plan based on the source surface which is then used to derive an optimized matching strategy. Furthermore, our entropy based criterion directly minimizes the matching uncertainty, taking all available information into account, while FPS sampling approaches this goal only heuristically. Tung and Matsuyama [TM10] as well as Zhang et al. [ZSCO*08] use a set of landmarks which are coarsely sampled to compute a geodesic diffeomorphism. Their method of detecting landmarks is based on extremal points of a geodesic integral function [HSKK01]. Landmarks are placed on such extrema to decrease the error introduced by approximate geodesic distances. In typical situations landmarks are placed on the shape's extremities. In general, however, these points are not sufficient to uniquely determine a diffeomorphism. Hence in case of [TM10] the authors propose to increase the subset of landmarks slightly by randomly sampling additional points on the surface. In general, this can not guarantee a sufficient subset of landmarks. Ruggeri et al. [RPSS10] compute a set of anchor-points of a shape by thresholding critical points of the Laplace-Beltrami [RBG*09] operator. These anchor points are located on geometrically and topologically meaningful regions of the shape and are invariant with respect to isometries. Additionally, they sample reference point by farthest point sampling [ELPZ97]. Although the idea of using the Laplace-Beltrami operator for placing anchor points is different to curvature-based approaches, the authors are still comparing two shapes by the intrinsic distances between these reference points. Reducing the set of points without corrupting the isometric property as proposed in our paper could significantly decrease the computation time of their algorithm.

Feature Saliency: Schmid et al. [SMB98] introduce entropy for feature detection. They compare different feature point detectors by their *distinctiveness*. Entropy has since then been used frequently in follow up work to estimate feature saliency. Our work is different from this line of work in that it not only looks at the entropy of the feature descriptors but on the remaining entropy of the full matching problem. Bronstein et al. [BBCK09] use a statistical measurement of the frequency-inverse document frequency known from the field of document retrieval in the area of shape matching. They measure the statistical significance of a feature point by checking how often it can be found in other shapes. Unique features describe a shape better than those found often in other shapes. Our approach is different in that it examines features within one shape. Again, the most important difference is probably that we capture the interdependence of many interrelated feature matches rather than just the statistics of a single match.

Robustness by randomness: Fischler and Bolles [FB87] describe a *hypothesis-and-check* concept called RANSAC (random sample consensus) which robustly fits a model to observed data. Several correspondence matching algorithms based on RANSAC are published in the literature. Chum and Matas [CM05] show a progressive random consensus, PROSAC, in order to find correspondences between two images (which is a problem very similar to shape matching). The idea is to use linear order on the correspondence set, which is defined by a similarity function of correspondences. The solution samples are drawn from a top-ranked set of correspondences. The convergence time improves on RANSAC, since the probability of choosing bad correspondences is low.

Techniques from the xSAC family have been applied to intrinsic shape matching by Tevs et al. [TBW*09], motivated by the forward search algorithm of Huang et al. for rigid shape matching [HFG*06]. Their technique is also related to the randomized matching algorithm by Memoli et al. [MS04]. Tevs et al. first randomly sample an initial set of correspondences. Starting at random source points, corresponding target points are chosen with probability proportional to the likelihood that the match is correct. Then additional correspondences are added if they do not violate the isometric matching criterion. Our paper extends their idea by not just performing importance sampling on the target point of the match but also carefully planning for which source points matches should be guessed. As we show in our experiments, this usually leads to a reduction in the number of random guesses that are necessary to find the correct solution as well as to an improvement of the quality of a match.

3. Isometry invariant matching

In the following, we assume that we are given two shapes, a source shape $\mathcal{M}_S$ and target shape $\mathcal{M}_T$, both differentiable 2-manifolds embedded in $\mathbb{R}^3$. Both are equipped with the standard intrinsic ("geodesic") metric $d_{\mathcal{M}}(\cdot,\cdot)$ that measures "within-surface" distances. We assume that $\mathcal{M}_S$ and $\mathcal{M}_T$ are isometric to each other, i.e., there exists a bijective function that maps the two shapes onto each other that is an isometry with respect to the intrinsic metric. Our task is to find such a mapping.

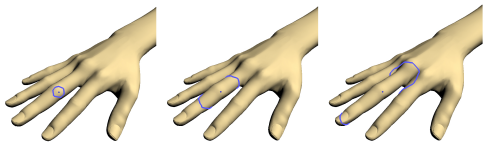


Figure 2: *Our descriptor measures the lengths of the isocurves of increasing distance: the curves spread like a "wave" over the surface.*

In order to solve this problem, we first need to discretize the surfaces. We refer to the discrete version of $\mathcal{M}_S$ and $\mathcal{M}_T$ as M_S and M_T respectively. For resampling, we use a standard uniform Poisson disc sampler. A discretized shape M derived from $\mathcal{M}$ is a set of n_s unordered points and is defined as a maximal set $M = \{\mathbf{x}_i \mid \mathbf{x}_i \in \mathcal{M}\}$ where $\forall i \neq j : |\mathbf{x}_i - \mathbf{x}_j| \geq \varepsilon$, so ε is the minimal distance between two points in the discretization. Furthermore, we assume that topology information is given which is for example the case for triangle meshes. Hence we can compute intrinsic distances as shortest paths along the mesh surface.

We can now pose the approximate intrinsic matching problem as an optimzation problem, finding a function $f : M_S \to M_T$ that minimizes:

$$E(f) := \frac{1}{|M_S|} \sum_{\mathbf{x}_i,\mathbf{x}_j \in M_S} \left| d_{\mathcal{M}_S}(\mathbf{x}_i,\mathbf{x}_j) - d_{\mathcal{M}_T}(f(\mathbf{x}_i),f(\mathbf{x}_j)) \right| \tag{1}$$

Finding the exact minimum of Eq. 1 could lead to a hard combinatorial problem. However, differences in correspondence at the level of the sampling resolution do not matter to us as we use the discretization only as a tool for representing the surface. Therefore, we usually set a fixed error bound of $c \cdot \varepsilon$ with $c \approx 2..3$ and consider a solution correct if it meets this precision bound. The final optimum could then be computed by a locally convergent numerical optimization [HAWG08, TBW*09]. Thus, this precision limit is no restriction in practice. Correspondingly, we focus in the following on finding an approximate solution in this sense as quickly as possible.

3.1. Descriptors

Like most other shape matching techniques, we use local descriptors in order to increase the matching efficiency. While matching local descriptors can rule out most of the infeasible options without costly global optimization.

We use an *intrinsic wave descriptor* in spirit to geodesic fans [GGGZ05]. In contrast to geodesic fans [GGGZ05],

we do not sample extrinsic curvature properties but only use intrinsic lengths of isocurves. Our approach is also independent of the mesh structure. Our approach could be understood as a simplified variant of heat kernel descriptors [SOG09]. The drawback of our simplified approach is that it is not as robust to topological noise and does not come with formal uniqueness guarantees. The advantage is that it is very easy to implement and avoids the substantial computational costs of numerically approximating heat kernels. Needless to say, it is straightforward to plug a different feature description scheme in our technique if necessary.

Let $C_{\mathcal{M}}(\mathbf{x},\rho)$ denote an intrinsic isocurve on manifold $\mathcal{M}$ around point $\mathbf{x} \in \mathcal{M}$: $C_{\mathcal{M}}(\mathbf{x},\rho) = \{\mathbf{y} \in \mathcal{M} \mid d_{\mathcal{M}}(\mathbf{x},\mathbf{y}) = \rho\}$. Starting from a given point $\mathbf{x} \in \mathcal{M}$, we compute all isocurves of distance $\rho \in [0,\rho_{max}]$ and measure their lengths. We normalize each length by dividing by $2\pi\rho$, which is the value expected for a flat (i.e., developable) piece of surface. The descriptor $D_{\mathbf{x}}(\rho)$ for each point $\mathbf{x} \in \mathcal{M}$ is the function that maps ρ to the normalized length of the corresponding isocurve. Fig. 2 illustrates the isocurves for an example model.

In our implementation, we sample 16 values at constant intervals of ρ. To compare two such discrete descriptors $\mathbf{D_x}$ and $\mathbf{D_y}$, we use the sum of squared differences. The maximum distance ρ_{max} is the only parameter of this scheme, and we typically set it to 5% of the longest side of the bounding box of the object. Analogous to heat-kernel signatures [SOG09], a smaller value emphasizes local properties, while a larger one includes more global properties. For example, if we use a small maximal distance on the fingertip of a hand (Fig. 3), we get a similar curve for every fingertip. But with increasing distance, every finger gets a different signature.

3.2. Shape matching by random sampling

We can now formulate the global matching algorithm. Before describing the new planned sampling strategy, we first describe the algorithm for unplanned sampling, mostly following Tevs et al. [TBW*09]: The algorithm starts by picking a random match between points from M_S and M_T. All pairs $(\mathbf{x},\mathbf{y}) \in M_S \times M_T$ are considered as candidate matches with a probability proportional to

$$P_{descr}(\mathbf{x},\mathbf{y}) \sim \exp\left(-0.5\sigma_d^{-2}|\mathbf{D_x} - \mathbf{D_y}|\right). \qquad (2)$$

σ_d is a user parameter that describes the standard deviation of matching descriptors (thus depending on model noise).

Then, subsequent matches are drawn randomly, taking the previous information into account: Points on the source shape M_S are chosen randomly. The likelihood for picking target points from M_T is weighted by the likelihood that the intrinsic distances are preserved: If one correspondence has been fixed, and a second source point has a distance d on the source shape, only points with a distance of roughly d on the target shape are accepted, again following a Gaussian error model. Given we already have k correspondences $\{(\mathbf{x}_1,\mathbf{y}_1),..,(\mathbf{x}_k,\mathbf{y}_k)\}, (\mathbf{x}_i,\mathbf{y}_i) \in M_S \times M_T$, the likelihood for one more correspondence $(\mathbf{x},\mathbf{y}) \in M_S \times M_T$ being correct is given by a Gaussian error model:

$$P_{dist}(\mathbf{x},\mathbf{y}|\mathbf{x}_1,\mathbf{y}_1,..,\mathbf{x}_k,\mathbf{y}_k) \sim \prod_{i=1}^{k} \exp\left(-0.5\sigma_g^{-2}\left(d_{\mathcal{M}_S}(\mathbf{x},\mathbf{x}_i) - d_{\mathcal{M}_T}(\mathbf{y},\mathbf{y}_i)\right)^2\right). \qquad (3)$$

Again, σ_g is a user parameter that describes the standard deviation of matching geodesic distances, thus modeling how "non-isometric" the model can be, either due to numerical imprecision or due to imperfections of the pieces of input geometry themselves. In order to obtain the final sampling probability, this likelihood is multiplied with P_{descr} to account for local descriptor matches, giving the sampling distribution:

$$P_{samp}(\mathbf{x},\mathbf{y}|\mathbf{x}_1,\mathbf{y}_1,..,\mathbf{x}_k,\mathbf{y}_k) = P_{descr}(\mathbf{x},\mathbf{y}) \cdot P_{dist}(\mathbf{x},\mathbf{y}|\mathbf{x}_1,\mathbf{y}_1,..,\mathbf{x}_k,\mathbf{y}_k). \qquad (4)$$

Sampling is continued until no more matching pairs are found that do not exceed the maximum error threshold or until all points are matched. The original algorithm of Tevs et al. uses a small set of fixed feature points which could be exhausted. In our case, we will use plan points, as explained in the next section, to limit the number of samples.

In the following, we refer to a set of matched points $\{(\mathbf{x}_1,\mathbf{y}_1),..,(\mathbf{x}_k,\mathbf{y}_k)\}$ as a solution set $\mathcal{R}$. This set sparsely models an isometric deformation $f_{\mathcal{R}}$ for the complete shape $\mathcal{M}_S$. In order to estimate dense correspondences, Eq. 3 is evaluated with the computed correspondences $\mathcal{R}$ as previous matches. For every point $\mathbf{x} \in M_S$, we compute the probability

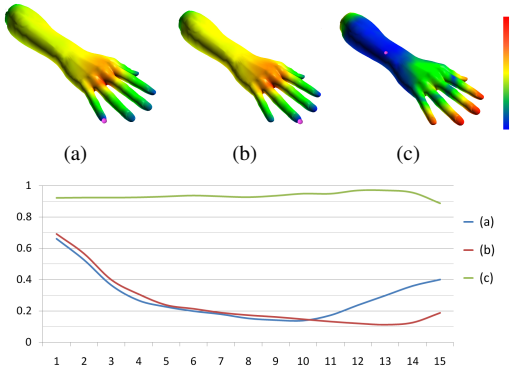
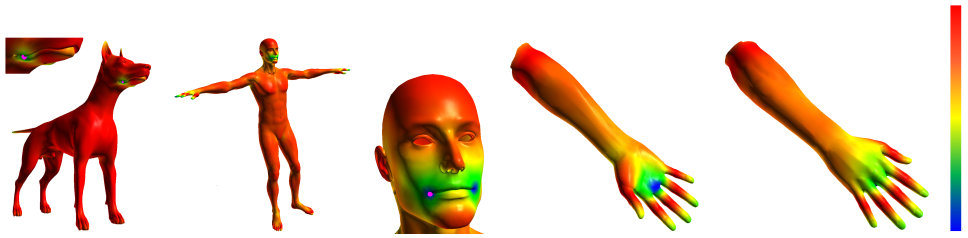


Figure 3: *Top row: descriptor differences - blue is zero and red is the maximum observed difference. Bottom row: descriptor signature: intrinsic distance ρ is mapped to the x-axis and the normed length of the resulting isocurve to the y-axis. The curves corresponds to the points shown in the top row. The points at the two fingertips have similar curves up to the radius of $\rho = 10$. The signature on the arm is, however, entirely different.*



(a) (b) (c) (d) (e)

Figure 4: *Entropy distribution of the most descriptive points, see Sect. 3.3 and Sect. 3.1. Regions with minimal entropy are marked blue, regions with maximal entropy are shown in red. (a), (b) The most descriptive point for the men and dog data sets is in the region of the corner of the mouth. (c) Zoomed view on human's head with most descriptive point shown in purple. (d), (e) Hand data set: the most descriptive point is on the palm in the area of the middle finger.*

distribution $P_{dist}(\mathbf{x},\mathbf{y}|\mathcal{R})$ on all $\mathbf{y}$ of M_T. The point $\mathbf{y}$ yielding the maximum likelihood is chosen as a dense match to $\mathbf{x}$. To compare different solution sets in their ability to characterize the deformation, we evaluate the corresponding deformation function $f_{\mathcal{R}}$ by computing $E(f_{\mathcal{R}})$, Eq. 1.

Intuitively, the score $E(f_{\mathcal{R}})$ measures the average error introduced by the isometry represented by the solution set. The whole sampling scheme is iterated until either a solution set with a score below a given threshold c is found (typical value $1.5\varepsilon \leq c \leq 3.0\varepsilon$) or the maximum number of trials (typically 100) is reached.

3.3. The planned sampling algorithm

Having the basic matching algorithm in place, we can now design a planning scheme that executes the algorithm more rapidly by avoiding uninformative samples. We call our strategy "PLANSAC": PLANned SAmple Consensus, referring to the fact that the algorithm plans ahead in terms of what samples are best to test in order to gain information as quickly as possible. The key idea is to look at P_{samp} (Eq. 4) and try to make the distribution as clearly determined as possible by using a semi-solution set $\{(\mathbf{x}_1^p,\cdot),..,(\mathbf{x}_K^p,\cdot)\}$ that we refer to as a *plan* for the matching step. The plan points $\mathbf{x}_i^p$ are computed such as to maximize the information gained by matching them against a target shape. In order to measure the information content, we look at the *entropy* [Sha51] of the final distribution.

The entropy value is a scalar which represents a statistical measure of the randomness of a discrete random variable Y with possible values $\{y_1,...,y_n\}$ and probability distribution P. It is defined as:

$$H_p(Y) = \sum_{i=1}^{n} p(y_i)\log\frac{1}{p(y_i)}, \tag{5}$$

where $p(y_i)$ is the probability of y_i and we define $p\log p := 0$ for $p = 0$. Intuitively, entropy can be understood as the amount of information contained in a random process described by the probability function. The entropy is maximal for a uniformly distributed probability. In this case we cannot take any advantage of the model described by the random process since every output of the process is equally likely. For a probability function with very few peaks, on the other hand, the entropy takes a low value. In this case the distribution carries a high amount of information. In our case, this means that the correspondences are known more precisely.

Our goal is now to find a set of points $\mathbf{X}^p = \{\mathbf{x}_i^p | \mathbf{x}_i^p \in M_S, i = 1..K\}$ that we use as a plan for a matching step. We begin with an empty set $\mathbf{X}^p$ and first add a most *discriminative* point $\mathbf{x}_k \in M_S$: This a point with a descriptor (Sect. 3.1) that leads to a minimal entropy when compared to all other points. We define a descriptor entropy of a point $\mathbf{x}_k \in M_S$ as $H_{p_k^d}(M_S)$. For the descriptor probability function p_k^d we use:

$$p_k^d(\mathbf{y}) = P_{descr}(\mathbf{x}_k,\mathbf{y}). \tag{6}$$

It computes the probability of a given point $\mathbf{x}_k$ to be equal to any other point $\mathbf{y} \in M_S$. Obviously, a point $\mathbf{x}_k$ that has a very dissimilar descriptor to other points has a low entropy $H_{p_k^d}(M_S)$. In other words, the chosen point is most discriminative among all others, see Fig. 4 for an example.

Having selected the first point, we start to add points to $\mathbf{X}^p$ by searching for points with most information content with respect to both descriptor matching as well as distances to previous points. We model this by checking how well shape $\mathcal{M}_S$ matches onto itself. We do this by assuming a point $\mathbf{x}_k \in M_S$ is the one we look for, i.e. $\mathbf{X}_k^p = \mathbf{X}^p \cup \{\mathbf{x}_k\}$ and computing the matching entropy $H_{p_k^m}(M_S \times M_S)$ of all possible correspondence pairs $(\mathbf{x},\mathbf{y}) \in M_S \times M_S$. The probability function p_k^m, describes how well a particular pair $(\mathbf{x},\mathbf{y})$ is localized by the current candidate plan set $\mathbf{X}_k^p$ and is set to:

$$p_k^m(\mathbf{x},\mathbf{y}) = P_{dist}(\mathbf{x},\mathbf{y}|\mathbf{X}_k^{p\prime}), \tag{7}$$

where $\mathbf{X}_k^{p\prime}$ is a set of trivial matching pairs, i.e. $\mathbf{X}_k^{p\prime} = \{(\mathbf{x}_i^p,\mathbf{x}_i^p)|\mathbf{x}_i^p \in \mathbf{X}_k^p\}$. The trivial matching pairs model the self-matching of shape $\mathcal{M}_S$. Given that we compute:

$$\arg\min_k \left(H_{p_k^m} + \lambda_d H_{p_k^d}\right), \tag{8}$$

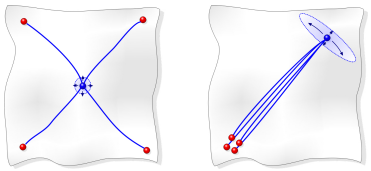


Figure 5: *Blue regions schematically indicate the area of uncertainty given noisy geodesic distances. Left: the subset of red points defines the parametrization of the shape well. Right: A poorly chosen subset increases the uncertainty in the parametrization.*

where λ_d is a factor modeling how much of the descriptor entropy should be considered when searching for optimal points. A typical value is $\lambda_d = 0.1$. The additional descriptor entropy models the choice on the next more discriminative $\mathbf{x}_k$ if the choice is intrinsically ambiguous. The point $\mathbf{x}_k$ which minimizes the sum of entropies is added to our sampling set $\mathbf{X}^p$.

The model of matching $\mathcal{M}_S$ to itself is supported by the fact that we have assumed that the source surface is isometric to any target surface. Therefore, and since we are working with intrinsic surface properties only, the plan developed by our planning algorithm is not affected by isometric deformations.

Finally, we stop building the plan if all remaining points are fixed in their position and can be located uniquely by their intrinsic distance to all points of $\mathbf{X}^p$. Hence, for every remaining point $\mathbf{x}_i \in M_S$ we estimate the variance of $P_{dist}(\mathbf{x}_i, \mathbf{y}|\mathbf{X}^p)$ over all $\mathbf{y} \in M_S$. The variance is computed using weighted principal component analysis (PCA) [Jol00]: We assign a weight $w_i(\mathbf{y}) = P_{dist}(\mathbf{x}_i, \mathbf{y}|\mathbf{X}^p)$ to every point $\mathbf{y}$ and perform a PCA. The largest eigenvalue describes the maximum radius of the uncertainty in the localization, see Fig. 5. If the maximal uncertainty radius for all source points $\mathbf{x}$ is smaller than a specified threshold $c \cdot \varepsilon$, typically $c = 2$, we stop adding plan points. We can now guarantee that our plan can find at least one isometric solution where every point is determined up to a standard deviation of $c \cdot \varepsilon$.

It should also be noted that this incremental algorithm is not optimal; the best choices could be made by looking at all K-tuples simultaneously. However, computing this requires $\mathcal{O}({n_s}^K)$ time, which is very expensive even for moderate numbers of plan points K. Despite the lack of guaranteed optimality, our greedy approximation yields very good results in practice.

4. Results

We evaluate our algorithm by performing shape matches on a number of benchmark models. For the evaluation, we use the standard TOSCA [BBK06a] data set. The data set consist of triangle meshes with approximately 50,000 vertices. In addition, we have constructed a few further synthetic benchmark scenes to examine specific properties of our algorithm.

data set	$\lvert\mathbf{X}^p\rvert$	$\lvert M_S\rvert$	$\lvert M_T\rvert$	$\#it$	$E(f_{\mathcal{R}})$	t_p	t_m
hand	5	164	797	15	1.25ε	175	0.658
dog	9	182	1341	2	1.5ε	432	0.088
			1522	8			0.340
			1394	12	2.0ε		0.572
cat	7	183	1221	8	2.5ε	282	0.298
			1274	10			0.467
			1355	15			0.765
centaur	10	211	1055	12	1.6ε	320	0.731
			909	16			0.853
			998	30			1.830
seahorse	9	126	572	1	1.0ε	19	0.060
			550	1			0.045
			549	2	1.25ε		0.110

Table 1: *Results summarized for different data sets. $\lvert\mathbf{X}^p\rvert$ is the number of points in the sampling plan, $\lvert M_T\rvert$ is the number of points on the target shape for different poses (order as in Fig. 9). #it is the number of iterations required to find a matching solution with an error less than or equal to $E(f_{\mathcal{R}})$ averaged over 50 PLANSAC runs. t_p is the time required to build a plan set $\mathbf{X}^p$, t_m is the average overall matching time. Timings are given in seconds for an Intel Core-2 Xenon 2.8 GHz platform (plan-build step was parallelized over 3 cores).*

For efficiency reasons we down-sample the input meshes. Source shapes have a minimum point spacing of 4..5% of the longest bounding box size. For the target shapes, we use a higher resolution of around 1..2% of the bounding box size to provide sufficient numerical degrees of freedom for the match. The intrinsic wave descriptor is evaluated on the down-sampled source points however propagating waves over the full resolution model. This preserves the *descriptiveness* of the down-sampled points. In the following, we refer to this target mesh sample spacing as ε and specify all accuracy results with respect to this number.

Pairwise matching results: Table 1 summarizes the matching results and timings on the employed data sets, please refer to it for the following discussion. The dog data set, Figs. 9(a)-9(d), represents a case of very descriptive geometry. Note that the average number of iterations required to find a solution is very low, since here we benefit from well-matched descriptors. Fig. 9(d) shows a case where the deformation in the input data does not preserve isometric distances well. Our algorithm, however, is still able to find correct correspondences. Figs. 9(e)-9(h) shows a case where the shape has a low descriptive information. The number of iterations increases as does the localization error of the solution. The centaur data set, Figs. 9(i)-9(l), is the most complex shape that we encountered in our tests. This also shows in the number of iterations that are required to obtain a suitable match. The seahorse data set, Figs. 9(m)-9(p), on the other hand, has very well distributed and descriptive features that explain most of the geometry very well. This, in conjunction with the relatively low resolution, which is due

to the low resolution of the initial model from the data base, enables very fast matching results. Our algorithm is able to find the solution in only a single iteration most of the time. In almost all shapes the most discriminative point was found in the area of the mouth corner. This is due to the fact that these shape models contain throat geometry which is itself very descriptive. Another observation is that landmarks are placed preferably on the shape's extremeties.

Fig. 6 shows our synthetic data-sets. Please note that all used shapes are featureless, however since our technique neither relies on intrinsic nor extrinsic features, we are still able to compute a valid plan set and can guarantee to find a valid isometry.

Examining the Practical Matching Complexity: In the beginning of the paper we posed the question of matching complexity in the case of numerical ill-conditioning, e.q. due to approximate geodesics. We would like to provide a partial answer to this question from a practical point of view. Fig. 6 shows a data set consisting of one, two and three spheres connected through a narrow tube. In theory the number of computed plan points required to isometrically parametrize a sphere is three, which we also found numerically as shown in Fig. 6(a). For two connected spheres, however, not any 3-tuple of points is sufficient to fix an isometry in practice, since fixing intrinsic distances on one piece does not constrain the matching on the second piece. It follows that for a numerically stable match more plan points are needed. As seen in Fig. 6(b) our algorithm automatically identifies this instability and computes two additional points on every piece of geometry. Our experiments show that numerical ill-conditioning can increase the practical matching complexity towards the case of two unconnected objects, where 6 plan points would be required, three for each sphere.

A similar observation can be made for surfaces of higher genus. We tested our algorithm on a torus and double-torus shape as well. The required plan points are 4-5 for the torus, where 5 points lead to more stable matching results. The double-torus requires at least 6 plan points. In practice, we also observe differences between shapes: For example, the centaur requires 10 landmark points to guarantee a matching of reasonable accuracy while the cat requires only 7 to achieve the same accuracy.

Comparison to previous work: We compare our method to two state-of-the-art feature-based techniques, spectral matching [LH05], and the RANSAC method of Tevs et al. [TBW*09]. For fair comparison, the RANSAC method is used with the same maximum number of iterations (100) as our proposed algorithm. We compare the performance on a subset of the models from Table 1 (seahorse and centaur). Fig. 7 shows the correspondences computed on the first pose of the seahorse and centaur data sets. Since both feature-based techniques do not provide dense correspondences, we compute their matching scores $E(f_{\mathcal{R}})$ on the sparse correspondence set only. As can be seen, RANSAC outperforms spectral matching in the number of computed correspondences, while it is unclear which method provides a better matching accuracy. Note however, that our proposed PLANSAC scheme has a significantly higher accuracy of up to 4 times while computing a dense set as compared to only sparse matches in the competing methods.

Additionally, if only the matching phase is considered, i.e. all pre-processing is excluded, our technique is able to outperform both algorithms in computation time as well. The computation time required by spectral matching is 1.2 seconds for the seahorse and around 45 minutes for the centaur data set. For RANSAC, these numbers are 2.5 seconds and around 22 minutes, respectively. In contrast, once the plan is known, our method dramatically reduces computation cost to a fraction of the time required by both techniques, i.e. 70ms for the seahorse data set and 740ms for the centaur data set.

Application Scenarios: As our first scenario we consider a sequence matching application. Fig. 8 shows a running horse sequence [SP04] where the dense correspondences are computed by matching the first frame against all other frames. The matching of one frame against all 50 frames only takes about 52 seconds, i.e. about 1 second per frame. Please note that we do not perform any extrinsic check if a computed match is one of the symmetric solutions, therefore the solution set also contains symmetrically equivalent correspondence matches. The median average number of required PLANSAC iterations is one, their mean number is 4. There are only four frames where more than 10 iterations are required.

As a second scenario we consider a database-like query application. For this scenario we construct a small database consisting of all models in all poses shown in Fig. 9 normalized to exhibit a uniform mean intrinsic distance. Without the latter, the task would be trivial since the models exhibit widely different scales and are thus distinguishable by their mean intrinsic distance alone. Fig. 10 shows the results of our initial experiments. To perform the query, we match the query shape against all shapes in the database. The matches are then sorted by their matching score $E(f_{\mathcal{R}})$. The probability of the best scored shape being the correct one is greater than 95% (here we ignore the match of the given template to itself) over 100 trials. We emphasize that these results are preliminary and only hint at the potential use of our algorithm in database applications.

5. Discussion and Outlook

We have developed a shape matching algorithm based on the novel concept of entropy-based planned random sampling. It enables very fast and reliable matching between similar or different shapes, while simultaneously removing the need for explicit feature detection. Moreover, the algorithm automatically adapts to the input characteristics and chooses an optimized sampling strategy for any given object. Numerical experiments hint at the possibility of characterizing a shape's

intrinsic complexity via the number of plan points necessary to parametrize its surface in a stable way.

The current limitations of our algorithm are that we assume geometry with known, consistent topology. We also assume that there are no measurement holes. Thus to apply our algorithm to real-world scanner data still requires additional work which we plan to perform in the future.

Despite these limitations, we are confident that the ideas put forward in this paper, i.e. the PLANSAC scheme, has a much wider applicability than shape matching. Many related areas in computer graphics and computer vision use random-sampling-based procedures for an abundance of tasks. Introducing a planning step might result in significant performance improvements similar to those demonstrated in our application.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments. This work has been partially supported by the Cluster of Excellence for Multi-Modal Computing and Interaction.

References

[ASP*04] ANGUELOV D., SRINIVASAN P., PANG H.-C., KOLLER D., THRUN S., DAVIS J.: The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *Proc. of NIPS* (2004), pp. 33–40. 1

[BBCK09] BRONSTEIN A. M., BRONSTEIN M. M., CARMON Y., KIMMEL R.: Partial similarity of shapes using a statistical significance measure. *IPSJ Transactions on Computer Vision and Applications 1* (2009), 105–114. 3

[BBK06a] BRONSTEIN A., BRONSTEIN M. M., KIMMEL R.: Tosca dataset. http://tosca.cs.technion.ac.il/data/nonrigid3d.zip, 2006. [Online; accessed 20-June-2010]. 6

[BBK06b] BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *In PNAS 103*, 5 (2006), 1168–1172. 1

[CM05] CHUM O., MATAS J.: Matching with PROSAC - progressive sample consensus. In *Proc. of CVPR* (2005), vol. 1, pp. 220–226. 3

[ELPZ97] ELDAR Y., LINDENBAUM M., PORAT M., ZEEVI Y. Y.: The farthest point strategy for progressive image sampling. *IEEE Trans Image Process 6*, 9 (1997), 1305–1320. 2

[FB87] FISCHLER M. A., BOLLES R. C.: *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography*. Morgan Kaufmann Publishers Inc., 1987, pp. 726–740. 3

[GGGZ05] GATZKE T., GRIMM C., GARLAND M., ZELINKA S.: Curvature maps for local shape comparison. In *Proc. of SMI* (2005), p. 244Ű256. 3

[GMGP05] GELFAND N., MITRA N. J., GUIBAS L. J., POTTMANN H.: Robust global registration. In *Proc. of SGP* (2005), pp. 197–206. 1, 2

[HAWG08] HUANG Q.-X., ADAMS B., WICKE M., GUIBAS L. J.: Non-rigid registration under isometric deformations. In *Proc. of SGP* (2008), vol. 27, pp. 1449–1457. 1, 2, 3

[HFG*06] HUANG Q.-X., FLÖRY S., GELFAND N., HOFER M., POTTMANN H.: Reassembling fractured objects by geometric matching. *ACM Trans. Graphics 25*, 3 (2006), 569–578. 3

[HSKK01] HILAGA M., SHINAGAWA Y., KOHMURA T., KUNII T. L.: Topology matching for fully automatic similarity estimation of 3d shapes. In *Proc. of SIGGRAPH '01* (2001), pp. 203–212. 2

[Jol00] JOLLIFFE I.: *Principal Component Analysis*. Springer, 2000. 6

[LF09] LIPMAN Y., FUNKHOUSER T.: Möbius voting for surface correspondence. In *Proc. of SIGGRAPH '09* (2009), pp. 1–12. 1, 2, 9

[LG05] LI X., GUSKOV I.: Multiscale features for approximate alignment of point-based surfaces. In *Proc. of SGP* (2005), pp. 217–226. 2

[LH05] LEORDEANU M., HEBERT M.: A spectral technique for correspondence problems using pairwise constraints. In *Proc. of ICCV* (2005), vol. 2, pp. 1482–1489. 7, 9

[MD03] MOENNING C., DODGSON N. A.: *Fast Marching farthest point sampling for implicit surfaces and point clouds*. Tech. rep., 2003. 2

[MGPG04] MITRA N. J., GELFAND N., POTTMANN H., GUIBAS L.: Registration of point cloud data from a geometric optimization perspective. In *Proc. of SGP* (2004), pp. 23–31. 1

[MS04] MEMOLI F., SAPIRO G.: Comparing point clouds. In *Proc. of SGP* (2004), pp. 32–40. 2, 3

[OMMG10] OVSJANIKOV M., MÉRIGOT Q., MÉMOLI F., GUIBAS L.: One point isometric matching with the heat kernel. In *Proc. of SGP* (2010), pp. 1555–1564. 1, 2

[RBG*09] REUTER M., BIASOTTI S., GIORGI D., PATANÈ G., SPAGNUOLO M.: Discrete Laplace-Beltrami operators for shape analysis and segmentation. *Computers & Graphics 33* (2009), 381–390. 2

[RPSS10] RUGGERI M. R., PATANÈ G., SPAGNUOLO M., SAUPE D.: Spectral-driven isometry-invariant matching of 3d shapes. *IJCV 89* (2010), 248–265. 1, 2

[RS08] RUGGERI M. R., SAUPE D.: Isometry-invariant matching of point set surfaces. In *Proc. of EG 3DOR* (2008). 1

[Sha51] SHANNON C. E.: Prediction and entropy of printed english. *Bell Systems Technical Journal* (1951), 50–64. 5

[SMB98] SCHMID C., MOHR R., BAUCKHAGE C.: Comparing and evaluating interest points. In *Proc. of ICCV* (Jan. 1998), pp. 230–235. 2

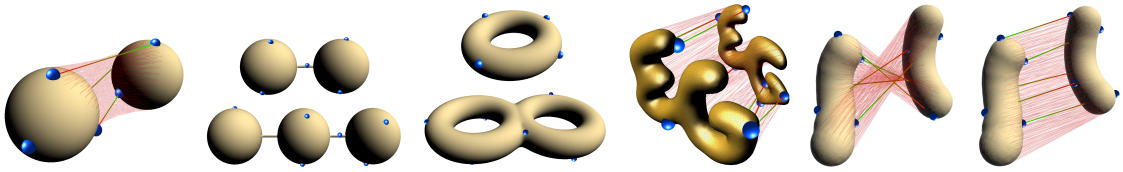
[SOG09] SUN J., OVSJANIKOV M., GUIBAS L.: A concise and provably informative multi-scale signature based on heat diffusion. In *Proc. of SGP* (2009). 1, 4

[SP04] SUMNER R. W., POPOVIĆ J.: Deformation transfer for triangle meshes. In *Proc. of SIGGRAPH '04* (New York, NY, USA, 2004), ACM, pp. 399–405. 7, 9

[TBW*09] TEVS A., BOKELOH M., WAND M., SCHILLING A., SEIDEL H.-P.: Isometric registration of ambiguous and partial data. In *Proc. of CVPR* (2009), pp. 1185–1192. 1, 2, 3, 4, 7, 9

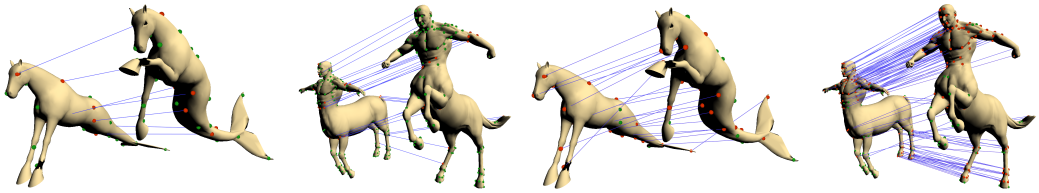
[TM10] TUNG T., MATSUYAMA T.: Dynamic surface matching by geodesic mapping for 3d animation transfer. In *Proc. of CVPR* (2010), pp. 1402–1409. 2

[ZSCO*08] ZHANG H., SHEFFER A., COHEN-OR D., ZHOU Q., VAN KAICK O., TAGLIASACCHI A.: Deformation-drive shape correspondence. *Computer Graphics Forum (Special Issue of SGP) 27*, 5 (2008), 1431–1439. 2



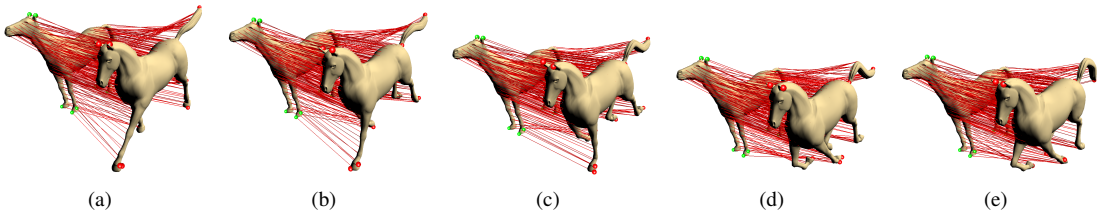
(a) (b) (c) (d) (e) (f)

Figure 6: *A number of featureless objects which can be matched with our method. (a) Match of a simple sphere against itself. Note that the compute plan as numerically chosen by our algorithm consists of 3 points and is thus consistent with theoretical results [LF09]. (b) Set of plan points* $\mathbf{X}^p$ *(marked blue) detected for synthetic data sets consisting of two and three spheres connected by a thin tube. The number of required plan points for stable matching is 5 and 7 respectively. (c) shows the plan points for the torus (5, one invisible) and the double torus (6). (d) EG-Logo as a meta-blob matched against a slightly deformed version of itself. The matching error is* 2ε *or* 8% *of the bounding box size. (e) Synthetic peanut data set consisting of slightly deformed meta-blobs. The solution error is* 1.5ε *or* 3% *of the bounding box size. Although the solution appears incorrect, we have an intrinsically correct match. Extrinsically, the object was flipped along the horizontal axis and mirrored along the vertical axis. (f) Extrinsically plausible match for the peanut data set.*



(a) (b) (c) (d)

Figure 7: *(a) and (b) show the matching for the first pose of the seahorse and centaur data sets computed with spectral matching [LH05]. The localization error* $E(f_{\mathcal{R}})$ *is* 6.5ε *(7 correspondences) and* 4.5ε *(26 correspondences) respectively. (c) and (d) shows the matching computed with RANSAC matching [TBW*09]. The localization errors of the RANSAC results are* 4.3ε *(22 correspondences) and* 5.6ε *(121 correspondences). In contrast, our algorithm achieves a localization error of only* 1.0ε *(seahorse) and* 1.6ε *(centaur) while matching the shapes densely, see Table 1. The corresponding plan matches are shown in Fig. 9 (j) and (n).*



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Figure 8: *(a)-(j) Dense correspondences shown for the galloping horse data set [SP04]. The first 10 out of 50 frames of the data set are shown. The computation time is 52 seconds for 50 frames, the average number of PLANSAC iterations required to match the template against one frame is 4.*

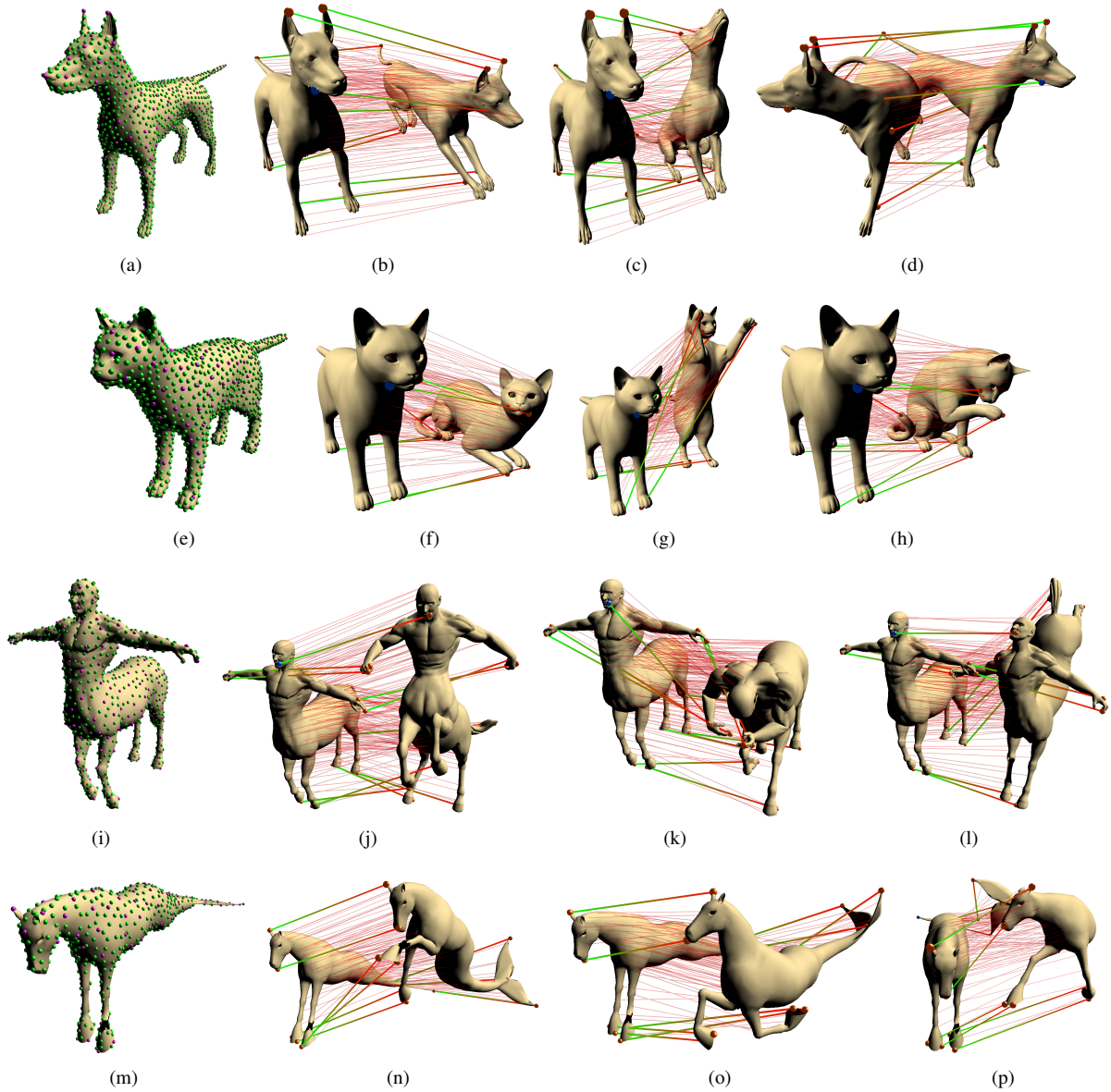


Figure 9: *Matching results for data sets in different poses. (a), (e), (i) and (m) show the discretization of the data. Purple points mark the discretization of the source shape, green points that of the target shape (in the pose of the source shape). The order of the poses is arranged in order of increasing numerical matching complexity, i.e. the average number of iterations required to find a solution.*

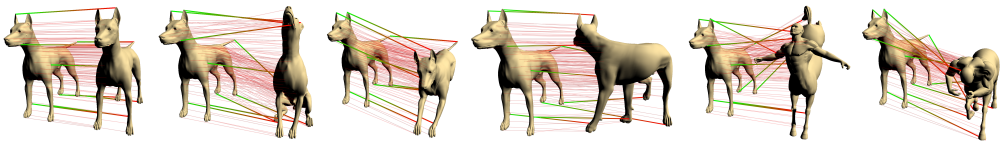


Figure 10: *Results of an exemplary database query. Results are sorted (from left to right) by their matching score* $E(f_{\mathcal{R}})$.