

Smooth Shells: Multi-Scale Shape Registration with Functional Maps

Marvin Eisenberger

Zorah Löhner
Technical University of Munich

Daniel Cremers

Abstract

We propose a novel 3D shape correspondence method based on the iterative alignment of so-called smooth shells. Smooth shells define a series of coarse-to-fine shape approximations designed to work well with multiscale algorithms. The main idea is to first align rough approximations of the geometry and then add more and more details to refine the correspondence. We fuse classical shape registration with Functional Maps by embedding the input shapes into an intrinsic-extrinsic product space. Moreover, we disambiguate intrinsic symmetries by applying a surrogate based Markov chain Monte Carlo initialization. Our method naturally handles various types of noise that commonly occur in real scans, like non-isometry or incompatible meshing. Finally, we demonstrate state-of-the-art quantitative results on several datasets and show that our pipeline produces smoother, more realistic results than other automatic matching methods in real world applications.

1. Introduction

The wide selection of affordable 3D scanning devices in recent years has led to an enormous growth in the amount of 3D shapes and scans available. In contrast to synthetic shapes, real-world scans are often noisy and many properties cannot be guaranteed. For example, topological noise might appear in self-touching areas or the meshing density varies depending on scanning conditions. These distortions were proven to be difficult for state-of-the-art shape correspondence methods [24, 34]. Many traditional methods focus only on the (nearly) isometric case, clearly defined extensions of this like partiality [28], or learn matching for different classes of shapes under certain perturbations [18]. Unfortunately, this requires training data and knowledge about what deformations and noise are to be expected.

In general one can distinguish between intrinsic and extrinsic correspondence methods. Intrinsic methods only use surface properties that are independent of the embedding, for example the Laplace-Beltrami operator. On the other hand, extrinsic methods directly use the 3D embedding of

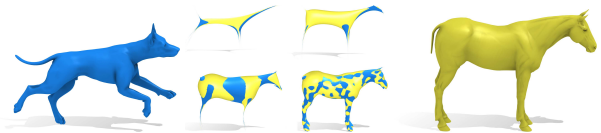


Figure 1: Given a source (left) and target (right) shape we propose a hierarchical smoothing procedure to iteratively align the inputs. First, we align very coarse approximations and then refine until we get correspondences for the original inputs. Among other things, we can handle challenging interclass pairs like matching a dog to a horse and our method is fully automatic, i.e. we do not use any additional information like hand-selected landmarks.

the shapes. While intrinsic methods are invariant to large scale, near-isometric deformations, extrinsic alignment is often more suitable for pairs with topological changes or other non-isometric deformations. A natural step would be to combine both to get the best of both worlds but only few previous approaches venture in this direction [11, 13].

Contribution In this paper we combine intrinsic and extrinsic information by embedding the input shapes into the product space of intrinsic (spectral) and extrinsic (xyz) coordinates. Then, we iteratively align smooth approximations of the two input shapes in this product space which we call smooth shells. Moreover, we propose a Markov chain Monte Carlo initialization strategy to find a meaningful local minimum and disambiguate self-similarities. Overall, we obtain a robust matching pipeline that works out of the box for a broad range of inputs beyond the isometry assumption and in the presence of various types of noise.

2. Related Work

2.1. Shape Correspondence and Registration

Shape correspondence is an extensively studied topic with various applications in Computer Vision and Graphics. Surveys of state-of-the-art methods [55, 47, 52, 46] give a broader overview of existing approaches but here we focus on work that is immediately related to ours.

The Functional Maps [38] paper proposes an elegant

formalism to model shape correspondences. The main idea is to model mappings of functions on the input shape to functions on the output shape instead of point-to-point maps. This allows for a compact matrix representation in a low rank basis. Over the last years, the original framework has been extended and applied to various applications [45, 29, 42, 1]. A major challenge in this context is extracting a point-wise correspondence from a Functional Map [44]. Several methods extend the original formalism but most of them are computationally heavy or make restrictive assumptions about the inputs [44, 42, 37]. Another common approach is to take noisy correspondences obtained from a Functional Map and denoise them [16, 57]. However, this only works if the input map is sufficiently accurate. Finally, all the methods listed above are by design prone to produce faulty matches in the presence of intrinsic self-similarities.

Extrinsic methods explicitly deform and align the input shapes in the 3D embedding space. [36, 31] model the deformation with a linear mapping in a low rank basis on the surface of one shape. Like our approach, [15] alternates between calculating a deformation field and correspondences but the volume-preservation constraint restricts the applicability. Many deformation-based methods require expensive preprocessing to apply the deformation model, for example with a deformation graph [49], structural rods [3] or deep learning [18]. Non-rigid ICP methods iteratively align shapes but they rely on a good initialization [27, 4]. However, for many applications we do not have such a previous alignment and in general there is no trivial way to obtain it.

There exist accurate methods to register certain classes of shapes, e.g. humans [7, 2]. Unfortunately, these are highly specialized and depend on class specific features [33], or learn statistical models from data [41]. While these methods perform extremely well for shapes within their classes, they usually do not generalize to arbitrary examples.

2.2. Shape Approximation and Simplification

The idea of mesh simplification by smoothing is investigated thoroughly in previous work. [54] use manifold harmonics for the smoothing. In surface deformation modeling this is usually a two stage algorithm. First, a smoothed version of a shape is deformed and then the details are added back to the surface, see [8] for an overview. Some classical works on shape modeling with smoothing are [19] and [23]. [9] combines this approach with differential coordinates. Although our smooth shells are related to smoothing technique like [54], none of the mentioned approaches use a series of approximations. We propose a novel hierarchical shape smoothing method that is particularly suitable for coarse-to-fine matching.

Shape skeletons offer a lower-dimensional description of the rough geometry of a shape. A recent survey of 3D skele-

ton methods can be found in [51]. Although the skeletons are usually designed to be easily aligned between different shapes from similar classes, most methods typically only define a single, unique skeleton for each shape. This is useful for a rough matching but does not allow for an iterative refinement of the surface alignment. Similar to our method, [12] extracts a skeleton based on Laplacian-based contraction but aims at getting a unique curve skeleton. Some methods create an entire class of skeletons for each shape [43]. Our method differs from the previously mentioned in that we do not introduce a fixed skeleton for each shape. Instead we construct a whole class of approximations with an increasing level of detail.

3. Background

A correspondence between two input shapes \mathcal{X} and \mathcal{Y} is defined as a point-to-point mapping $\mathbf{P} : \mathcal{X} \rightarrow \mathcal{Y}$. Here, a shape is a 2D Riemannian manifold with an embedding in \mathbb{R}^3 . We use triangular meshes to discretize the surfaces \mathcal{X} and \mathcal{Y} and denote the coordinate matrices as $X \in \mathbb{R}^{N \times 3}$ and $Y \in \mathbb{R}^{M \times 3}$ with N and M vertices respectively.

3.1. Laplace-Beltrami Operator

The Laplace-Beltrami operator $\Delta = \text{div}(\nabla \cdot)$ is an extension of the standard Euclidean Laplacian to manifold domains \mathcal{X} . Computing solutions of $\Delta \phi_k = \lambda_k \phi_k$ yields the Laplace-Beltrami eigenfunctions $\{\phi_k\}_{k \in \mathbb{N}}$ which form an orthonormal basis of $L^2(\mathcal{X})$. This allows for a spectral representation of functions $f \in L^2(\mathcal{X})$:

$$f \approx \tilde{f} = \sum_{k=1}^K \langle f, \phi_k \rangle \phi_k. \quad (1)$$

According to the min-max principle, \tilde{f} is an optimal compact approximation of smooth functions $f \in L^2(\mathcal{X})$ [39] with a fixed basis size K . To compute the Laplace-Beltrami operator on triangular meshes, we use a cotangent discretization $\Delta \in \mathbb{R}^{N \times N}$ with lumped mass matrix [40] and we denote its first K eigenvectors as $\Phi_K = (\phi_1, \dots, \phi_N) \in \mathbb{R}^{N \times K}$ (analogously $\Psi_K \in \mathbb{R}^{M \times K}$ for \mathcal{Y}).

3.2. Functional Maps

The Functional Map framework [38] is a popular approach to solve for correspondences $\mathbf{P} : \mathcal{X} \rightarrow \mathcal{Y}$. In Functional Maps, \mathbf{P} is replaced with a mapping of functions to functions $\mathcal{C} : L^2(\mathcal{X}) \rightarrow L^2(\mathcal{Y})$. \mathcal{C} is linear and can therefore be compactly written as a matrix $\mathbf{C} \in \mathbb{R}^{K \times K}$:

$$\mathcal{C} = \Psi_K \mathbf{C} \Phi_K^\dagger. \quad (2)$$

To compute \mathbf{C} for a pair of input shapes we need additional information to constrain the solution. Given pairs of

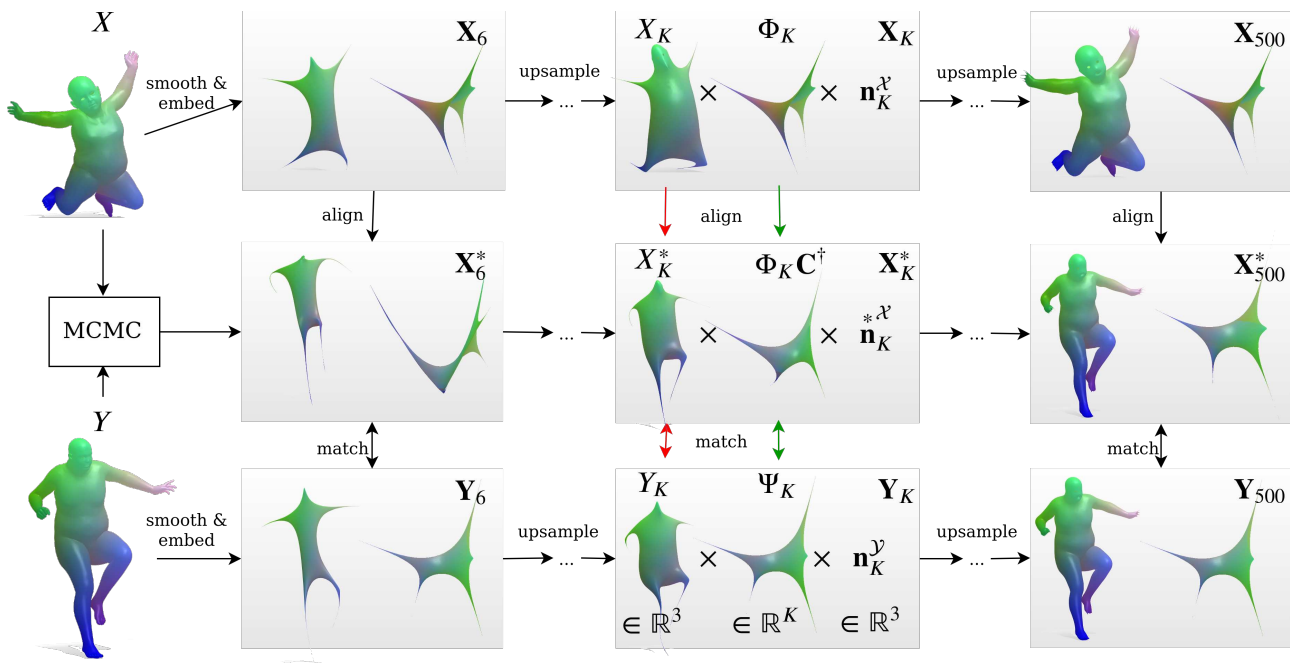


Figure 2: Overview of our pipeline. (Column 1) We initialize our method with the alignment X_6^* from our Markov chain Monte Carlo initialization algorithm, see Section 5 for details. (Columns 2-4) On each level K we embed both shapes in the $K + 6$ dimensional product space of the smoothed extrinsic coordinates $X_K \in \mathbb{R}^3$, the intrinsic spectral coordinates $\Phi_K \in \mathbb{R}^K$ and the outer normals $\mathbf{n}_K^X \in \mathbb{R}^3$. In this space we can align X_K and Y_K by computing an extrinsic morphing $\tau \in \mathbb{R}^{K \times 3}$ for X_K and a Functional Map $\mathbf{C} \in \mathbb{R}^{K \times K}$ for Φ_K . To visualize the spectral embedding Φ_K only the first three dimensions are shown. Finally, using the aligned X_K^* and Y_K we obtain a point-to-point matching $\mathbf{P} : \mathcal{X} \rightarrow \mathcal{Y}$ with a nearest neighbor search in \mathbb{R}^{K+6} . We repeat this process for 50 iterations with smoothing levels on a logarithmic scale between $K = 6 : 500$. Each iteration is initialized with the previous alignment.

corresponding functions $f_i \in \mathbb{R}^N$ and $g_i \in \mathbb{R}^M$ on the two surfaces, an energy to optimize for \mathbf{C} is:

$$E_{\text{feat}}(\mathbf{C}) := \|\mathbf{C}\Phi_K^\dagger F - \Psi_K^\dagger G\|_F^2. \quad (3)$$

Here, F, G are matrices whose columns are the feature functions f_i, g_i . Possible choices for those features range from pointwise descriptors or surface texture to input landmarks. Another common assumption is that the mapping \mathbf{P} is area preserving which leads to orthogonal Functional Maps $\mathbf{C}^\top \mathbf{C} = \mathbf{I}$, see [38, Theorem 5.1].

3.3. Shape deformation

A different approach is to align the surfaces in the embedding space instead of calculating the correspondence directly. We denote the deformed version of \mathcal{X} with \mathcal{X}^* and impose that \mathcal{X}^* should align with \mathcal{Y} . A common choice model is a linear displacement in a low rank basis [36, 31], e.g. the Laplace-Beltrami eigenbasis:

$$X^* = X + \Phi_K \tau. \quad (4)$$

$\tau \in \mathbb{R}^{K \times 3}$ are some displacement coefficients that parameterize the deformation. In the discrete case, the point-

wise correspondence is represented by the matrix $\mathbf{P} \in \{0, 1\}^{M \times N}$ with $\mathbf{P}^\top \mathbf{1} = \mathbf{1}$. Using the aligned shape X^* , we can recover \mathbf{P} by minimizing the following energy:

$$E_{\text{align,3D}}(\mathbf{P}) := \|\mathbf{P}X^* - Y\|_F^2. \quad (5)$$

This is equivalent to a nearest neighbor search in \mathbb{R}^3 . In order to get a meaningful correspondence with this approach we need to additionally regularize the deformations \mathcal{X}^* . One possibility is to assume that the deformations are as-rigid-as-possible on a local scale:

$$E_{\text{arap}}(\tau) := \int_{\mathcal{X}} \int_{\mathcal{N}(x)} \|R(x)(X(x) - X(y)) - (X_\tau^*(x) - X_\tau^*(y))\|_2^2 dy dx. \quad (6)$$

$\mathcal{N}(x)$ denotes the neighborhood of $x \in \mathcal{X}$ and $R : \mathcal{X} \rightarrow \text{SO}(3)$ describes the local rotation, for details see [48].

4. Method

We propose to compute shape correspondences by iteratively aligning a series of coarse-to-fine approximations of

the input surfaces \mathcal{X} and \mathcal{Y} . This is based on the idea that the alignment of two shapes can end up in unwanted local optima in the presence of non-consistent small scale features. In many cases the rough structure of \mathcal{X} and \mathcal{Y} , like the number of extremities, is similar while the fine scale details can differ, see Figure 1. After matching the global features, the local features can be used to refine the alignment. The smooth shells we use as coarse shape approximations are defined in Section 4.1. Section 4.2 explains how we combine extrinsic and intrinsic shape embeddings and Section 4.3 defines our complete matching algorithm.

4.1. Smooth Shells

In this section, we propose a novel shape smoothing operator \mathcal{S}_K that yields smoothed shapes similar to those from spectral surface reconstruction [25]. In comparison, our operator leads to smoother transitions between \mathcal{S}_K and \mathcal{S}_{K+1} which makes it more suitable for a hierarchical alignment.

Spectral Reconstruction Spectral reconstruction [25] smoothes \mathcal{X} by projecting its coordinate function X onto the first K Laplace-Beltrami eigenfunctions:

$$T_K := \mathcal{T}_K(X) = \sum_{k=1}^K (\phi_k \otimes \phi_k) X. \quad (7)$$

Here, $\phi_k \otimes \phi_k$ denotes the outer product of ϕ_k with itself which results in the projection of X onto ϕ_k . Since the eigenfunctions are ordered by frequency, this creates a coarse-to-fine approximation of the original shape. The level of detail is controlled by the number of eigenfunctions K . For small K only the rough geometry is reconstructed, whereas for $K \rightarrow \infty$, T_K converges to the original X .

Shell Operator The gradual smoothing from Eq. (7) is useful for hierarchical shape matching. In each iteration we increase K and use the alignment from the previous iteration as an initialization. However, in many cases the refinement with spectral reconstruction leads to undesirable artifacts. Especially the first few K projections from Eq. (7) cause large disparities between reconstructions. This makes the alignment from the previous iteration less useful for the next step. We introduce the *shell operator* \mathcal{S}_K to circumvent this issue:

$$X_K := \mathcal{S}_K(X) := \sum_{k=1}^{\infty} \frac{1}{1 + \exp(\sigma(k - K))} (\phi_k \otimes \phi_k) X. \quad (8)$$

Just like spectral reconstruction, \mathcal{S}_K smooths X using a projection on ϕ_k . However, instead of truncating the spectral coordinates at a certain K , we introduce a gradual truncation with sigmoid weights. Those are close to 1 if $k \ll K$ and decay to 0 when $k \gg K$. This guarantees that

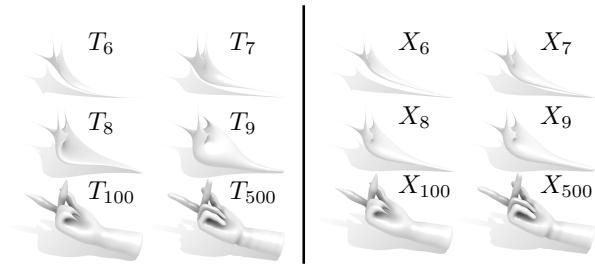


Figure 3: At first glance there is no significant difference between spectral reconstruction T_K and smooth shells $X_K := \mathcal{S}_K(X)$. They both converge to X for $K \rightarrow \infty$, and for high indices $K \gg 50$ they are indistinguishable. The crucial difference lies in their upsampling behaviour. While smooth shells transition smoothly from X_K to X_{K+1} , consecutive shapes T_K tend to have large displacements and are therefore less suitable for iterative alignment methods.

the displacement from \mathcal{S}_K to \mathcal{S}_{K+1} is reasonably bounded. For small σ the transition becomes smoother, whereas for $\sigma \rightarrow \infty$ the sigmoid function converges to the indicator function $\mathbf{1}_{\{k \leq K\}}$ which corresponds to spectral reconstruction. In particular, we can show the following smoothness result for \mathcal{S} :

Theorem 1. (Transition smoothness of \mathcal{S})

Let \mathcal{X} be a shape with coordinate function $X \in L^2$, then the geometric difference of state X_K and X_{K+1} is bounded by the upsampling variance σ in following way:

$$\frac{\|\mathcal{S}_{K+1}(X) - \mathcal{S}_K(X)\|_{L^2}}{\|\mathcal{S}_{K+1}(X)\|_{L^2}} \leq |1 - e^{-\sigma}| = \mathcal{O}(\sigma), \sigma \rightarrow 0. \quad (9)$$

We provide a proof in Appendix B. See Figure 3 for an illustration of the practical implications and Table 1 for a quantitative comparison to spectral reconstruction.

4.2. Intrinsic-extrinsic Embedding

Intrinsic and extrinsic methods are often depicted as opposing viewpoints and, although there are some notable exceptions [11, 13], only few methods try to combine them. Our deformation model combines shape alignment in both intrinsic and extrinsic space. Functional Maps is based on rigid ICP alignment of the spectral coordinates $\Phi_K \in \mathbb{R}^{N \times K}$ of \mathcal{X} and $\Psi_K \in \mathbb{R}^{M \times K}$ of \mathcal{Y} in the K -dimensional spectral domain [38].

$$\Phi_K \mathbf{C}^\dagger \approx \Psi_K. \quad (8)$$

On the other hand, extrinsic methods typically align the 3-dimensional geometry as described in Eq. (4):

$$X_K^* = X_K + \Phi_K \tau \approx Y_K.$$

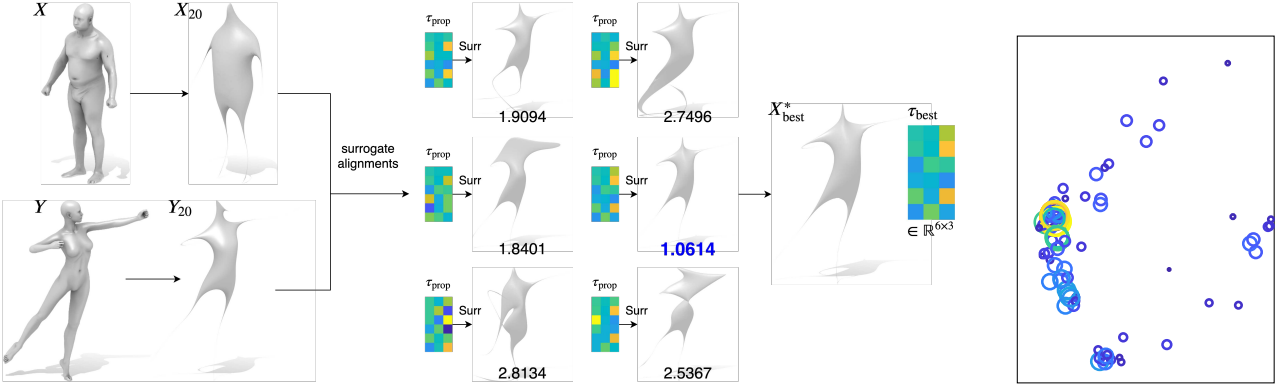


Figure 4: (left) Overview of our MCMC initialization method. We sample potential initial alignments $\tau_{\text{prop}} \in \mathbb{R}^{6 \times 3}$ and rate them using surrogate runs (see Section 5). Each proposal τ_{prop} is assigned a mark $E(\text{Surr}(\tau_{\text{prop}})) > 0$ based on the alignment quality of the current surrogate. In the shown example, the best objective is $E = 1.0614$ and indeed this sample visually shows the tightest alignment of X_{20}^* and Y_{20} . (right) 2D embedding with multi-dimensional scaling of all τ_{prop} used for one initialization. Samples with small objective values have big, yellow circles and big objectives correspond to small, blue circles. Evidently there is a big cluster around the optimal (yellow) circle which shows that our algorithm is able to determine the optimal initialization with high confidence.

We combine intrinsic and extrinsic alignment in order to gain both their advantages. To this end, we embed the inputs \mathcal{X} and \mathcal{Y} in the product space of the *intrinsic* (spectral coordinates Φ_K) and the *extrinsic* (smoothed Cartesian coordinates X_K and outer normals $\mathbf{n}_K^{\mathcal{X}}$ of \mathcal{X}) coordinates:

$$\mathbf{X}_K := (\Phi_K, X_K, \mathbf{n}_K^{\mathcal{X}}) \in \mathbb{R}^{N \times (K+6)}. \quad (10a)$$

$$\mathbf{Y}_K := (\Psi_K, Y_K, \mathbf{n}_K^{\mathcal{Y}}) \in \mathbb{R}^{M \times (K+6)}. \quad (10b)$$

Using the normals makes the embedding more descriptive because they convey information about the inside-outside orientation of each point. Using both the Functional Map \mathbf{C} and the extrinsic deformation τ (see Eq. (4)) now yields the morphed embedding \mathbf{X}_K^* :

$$\mathbf{X}_K^* := (\Phi_K \mathbf{C}^\dagger, X_K + \Phi_K \tau, \mathbf{n}_K^{\mathcal{X}}) \in \mathbb{R}^{N \times (K+6)}. \quad (11)$$

$\mathbf{n}_K^{\mathcal{X}}$ are the normals of $X_K + \Phi_K \tau$. The next section will go into detail on how to compute \mathbf{C} and τ .

4.3. Hierarchical Matching

Putting everything together, we can define a hierarchical correspondence algorithm with the following energy:

$$E(\mathbf{P}, \mathbf{C}, \tau) := \|\mathbf{P} \mathbf{X}_K^* - \mathbf{Y}_K\|_F^2 + \lambda_{\text{feat}} E_{\text{feat}}(\mathbf{C}) + \lambda_{\text{arap}} E_{\text{arap}}(\tau). \quad (12)$$

The regularization terms E_{feat} and E_{arap} are defined in Eq. (3) and Eq. (6) respectively. For the former we use the SHOT [53] and HKS [50] descriptors. To minimize the energy E we choose an alternating optimization strategy. In

particular, we first fix the correspondences \mathbf{P} and optimize for the alignment (\mathbf{C}, τ) and then do the same vice versa. This is a common approach for both intrinsic [38] and extrinsic [31, 32, 36] matching methods. Our overall matching algorithm is the following:

Algorithm 2. (Hierarchical Matching)

1. *Input:* \mathcal{X}, \mathcal{Y}

2. *For* $k \in \{0, \dots, |\mathcal{K}| - 1\}$:

$$2.1 \mathbf{P}_{k+1} := \arg \min_{\mathbf{P}_{mn} \in \{0,1\}, \mathbf{P}^\top \mathbf{1} = \mathbf{1}} E(\mathbf{P}, \mathbf{C}_k, \tau_k).$$

$$2.2 (\tau_{k+1}, \mathbf{C}_{k+1}) := \arg \min_{\tau, \mathbf{C}^\top \mathbf{C} = \mathbf{I}} E(\mathbf{P}_{k+1}, \mathbf{C}, \tau).$$

3. *Output:* $\mathbf{P}_{|\mathcal{K}|}, \mathbf{X}_{|\mathcal{K}|}^*$.

The decomposition of the optimization problem E results in more tractable subproblems. For \mathbf{P} this is a nearest neighbor search, for \mathbf{C} a Procrustes problem and for τ a nonlinear least squares problem. The first two can be solved in closed form, for the last one we use Gauss-Newton optimization. Our method now repeatedly solves those optimization problems with shells of an increasing level of detail $K \in \mathcal{K}$ on a logarithmic scale between $K_{\text{init}} = 6$ and $K_{\text{max}} = 500$. See Figure 2 for a visualization of Algorithm 2.

5. Initialization: Surrogate based Markov chain Monte Carlo Sampling

Self-similarities are still a challenging problem for state-of-the-art shape correspondence methods and many struggle to distinguish them without proper initialization [56, 17,

35]. Even for humans it is difficult to distinguish between the legs/arms of an animal without any context. In other words, our energy from Eq. (12) is highly non-convex with a multitude of local minima. Unfortunately, there is no obvious way to compute a meaningful initial alignment for all classes of shapes. We propose an indirect approach to this using Markov chain Monte Carlo (MCMC) sampling.

Surrogate runs This approach is based on efficiently exploring the space of initial poses instead of heuristically picking one. We assign a probability distribution to the displacement parameter $\tau \in \mathbb{R}^{K_{\text{init}} \times 3}$ and sample from this distribution. In particular, we set the prior for τ to the standard normal distribution $\mathcal{N}(0, \mathbf{I})$ and the negative log likelihood proportional to the objective value E . By design, this yields samples τ that have a high objective value E . Each τ is ranked according to the objective function E from Eq. (12) and the lowest energy result is used to initialize the full pipeline.

To evaluate E , we run a low cost version of the full pipeline, a *surrogate run*, with $K_{\text{max}} = 20$, no regularizers $\lambda_{\text{feat}}, \lambda_{\text{rap}} := 0$ and downsampled versions of the input shapes to 1000 vertices. We evaluate $N_{\text{prop}} = 100$ different proposals τ_{prop} . Those can be run in parallel with an average runtime of 0.46 seconds per surrogate. See Figure 4 for a visualization of this strategy and see Appendix A for pseudo code of our MCMC algorithm as well as the implementation in the supplementary material.

6. Experiments

We apply our pipeline to various, challenging matching tasks using two metrics to measure the quality of a matching. The first one is the *accuracy*, defined as the geodesic distance to ground truth matches, see Section 6.1. The second is the *smoothness* of the correspondence \mathbf{P} which we quantify using the conformal distortion of triangles, see

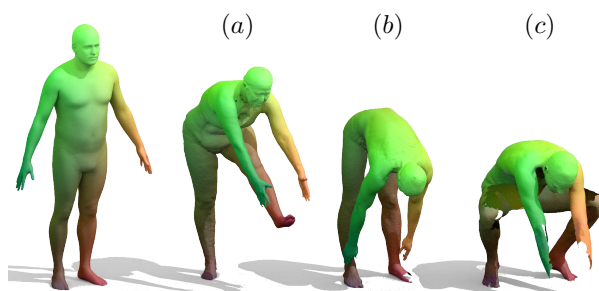


Figure 6: Example matchings for real scans from the FAUST [6] dataset. The shapes have very high resolution (200k vertices) and contain scanning noise. The FAUST interclass challenge consists of (a) different humans that are (b) subject to topological changes and (c) extreme degrees of noise and partiality. Like [18] we match a template (left) to each target. Here, correspondences are color coded such that matching points have the same color.

Section 6.2. To show that our method can be used out of the box, we use the same set of parameters for all experiments and do not require additional information except for the inputs \mathcal{X} and \mathcal{Y} . See our implementation in the supplementary material for more details. Additionally, we perform an ablation study in Section 6.3 and a runtime analysis in the Appendix C to further investigate our method. Finally, there are more qualitative examples of matchings and style transfer in the Appendix.

6.1. Shape correspondence

We evaluate the matching accuracy of our method according to the Princeton benchmark protocol [22] on multiple datasets. Given the ground-truth match (x, y^*) , the error of the calculated match (x, y) is given by the geodesic dis-

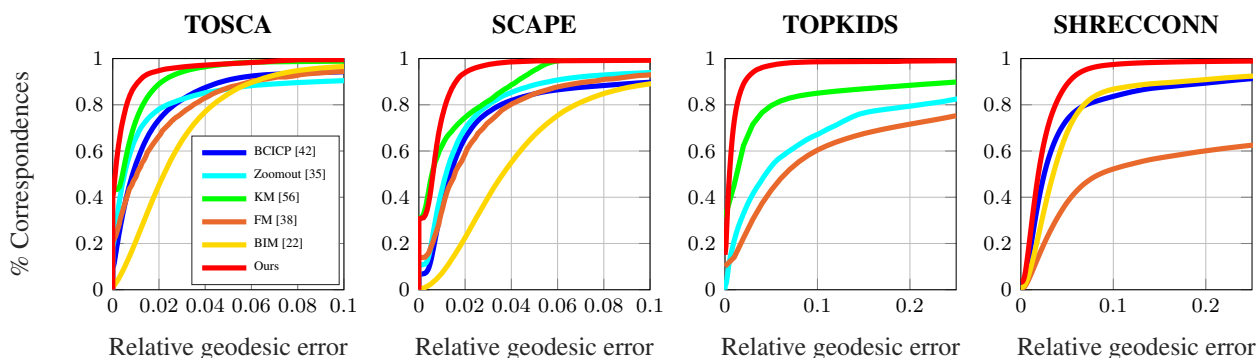


Figure 5: Our matching accuracies for four datasets in comparison to other popular fully automatic shape correspondence methods. For more details on the datasets, see Section 6.1.



Figure 7: Example texture transfers with our method for challenging interclass examples. The texture defined on the source shape (horse) is transferred to two individual target shapes (dog and human).

tance between y and y^* normalized by the diameter of \mathcal{Y} :

$$\epsilon(x) = \frac{d_{geo}(y, y^*)}{\sqrt{\text{area}(\mathcal{Y})}} \quad (13)$$

TOSCA, SCAPE, TOPKIDS, SHRECCONN Two datasets contain synthetic shapes with isometric pairs, TOSCA [10] contains 76 shapes of humans and animals, SCAPE [5] contains 72 poses of the same person. TOPKIDS [24] contains 25 poses of a human child with self intersections. These shapes are also synthetic but topological changes from real scans are simulated with merged meshing at self-touching areas. The SHREC’19 Connectivity [34] dataset contains 430 pairs of human shapes from different classes with severe differences in the meshing ranging from template sized shapes ($N \approx 5000$ vertices) to real scans ($N > 200K$), and varying vertex densities in different areas. We compare our matching accuracy on these datasets to other fully-automatic matching methods, see Figure 5.

FAUST The FAUST [6] dataset contains 300 real scans of different humans in various poses, see Figure 6. Besides being high resolution $N \approx 200K$ with non-compatible meshing, the shapes are noisy and highly non-isometric. Additionally, there are various poses with topological changes due to self touching parts. To address this issue, we do not compute the correspondence directly for a given pair of shapes but, like 3D-CODED [18], FARM [33] and LBS-AE [26], use an intermediate template from [30] to compute correspondences for two scans. This allows our method to separate topological changes and deal with noisy geometry, otherwise the as-rigid-as-possible assumption leads to faulty deformations. The accuracy (in cm) of the best methods for the FAUST [6] interclass challenge reported on the website faust.is.tue.mpg.de are:

Method	3D-CODED	SP	Ours	LBS	FARM	FMNet
Error	2.878 (4.883*)	3.126	3.929	4.079	4.123	4.826

For 3D-CODED (*) refers to the unsupervised version. The striking observation is that our method is on par with the state-of-the-art without specializing on the class of human shapes. Ours is the only method listed here that does not train on human shapes (3D-CODED, FMNet) or makes strong modelling assumptions holding only for humans (Stitched Puppet (SP), LBS-AE, FARM). We did not specifically tune parameters for this challenge.

6.2. Map Smoothness

A map with good accuracy can still produce artifacts when transferring information from surface \mathcal{X} to \mathcal{Y} because small scale noise typically does not have a severe effect on the geodesic matching error (13). This behavior is, however, prohibitive for applications like meshing, texture or normal map transfer, see Figure 7. The conformal distortion of each triangle after deformation measures the local consistency of a matching [21, Eq. (3)]. This allows for a quantification of the smoothness of the map \mathbf{P} .

Figure 8 shows the conformal triangle distortion of our method on the SCAPE dataset. Remarkably, the deformations obtained with our method are even smoother than the ground-truth provided in [5]. The reason for that lies in the way the authors construct this ground-truth. In order to transfer the meshes they use a classical nonrigid registration algorithm [20] to register a template in a canonical pose to 71 noisy scans of a person. This method requires ~ 150 markers to get a faithful alignment, some of which are hand-selected. The main concern was to obtain a possibly tight alignment of the markers. However, in practice the markers are not perfectly placed and these small deviations lead to distorted triangles. In comparison to that, we align the templates without any markers while explicitly using an ARAP penalization term. This evidently leads to smoother deformations and the few triangles that get distorted are typically not artifacts of random noise but rather in meaningful places like the armpits or the abdomen of the person in Figure 8.

6.3. Ablation study

We assess the effect of the different components of our method in the ablation study in Table 1. The main insight is that there is an intricate interplay of the different subparts of our method and the accuracy drops significantly if any part is removed. In particular, the MCMC initialization strategy is vital. Without it our deformation based approach is extremely prone to run into suboptimal local minima which leads to a failure rate of over 83%. Remarkably, even when our rigid initialization strategy (see end of Appendix A) is replaced with random rigid poses the failure rate is only around 38 – 50%. In many cases, our MCMC algorithm is able to find the correct pose, even in the presence of large scale rotational displacements of the inputs.

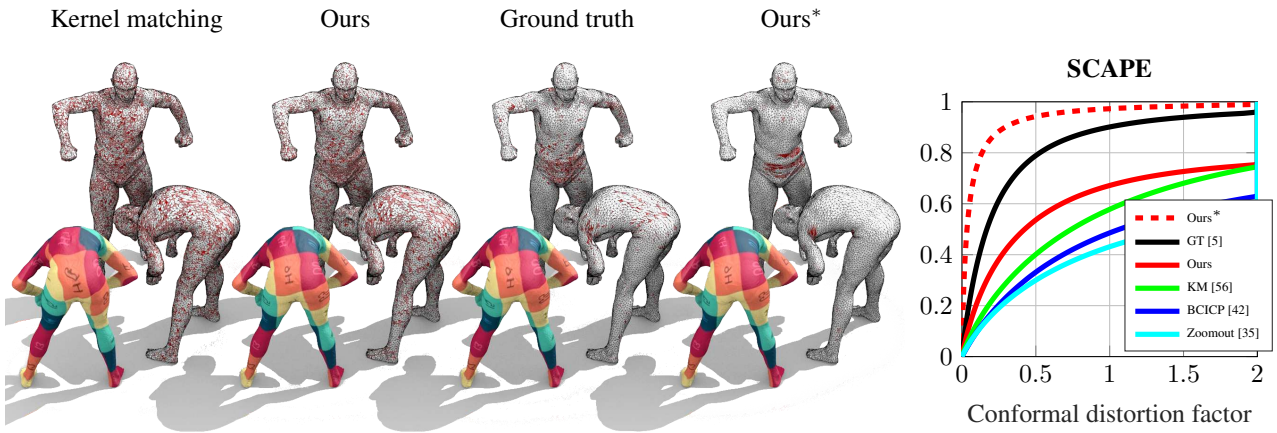


Figure 8: Correspondence smoothness measured on the SCAPE dataset using conformal distortion of triangles. (Left) Qualitative comparison – red triangles are distorted whereas white triangles preserve the angles. Our deformed mesh shows the most meaningful, artifact-free result. Additionally, we provide an example texture transfer to prove that our deformed mesh (Ours*) is the most useful one. (Right) Accumulated distortions for all 71 pairs in the dataset. There are two possibilities to transfer the mesh of \mathcal{X} to the reference pose \mathcal{Y} . Either we use the deformed geometry \mathcal{X}^* directly (Ours*) or we snap it to the surface \mathcal{Y} post alignment using the map $\mathbf{P} : \mathcal{X} \rightarrow \mathcal{Y}$ (Ours). The former is only possible for our method because it is the only one here to calculate a deformation instead of only a correspondence. See Section 6.2 for more details.

SCAPE	Ours	$\lambda_{\text{feat}} = 0$	$\lambda_{\text{arap}} = 0$	Extr. only	Intr. only	w/o normals	w/o MCMC	random rigid	spectral rec.
Avg. error	0.0088	0.0211	0.0147	0.0344	0.0121	0.0115	0.0568	0.1163	0.0139
Failure rate	0	0.2676	0.0282	0.7606	0.2254	0.0141	0.8310	0.4930	0.0282
Avg. Distortion	0.1287	0.1171	0.1604	0.1322	0.1539	0.1310	0.2594	0.2055	0.1305
TOSCA	Ours	$\lambda_{\text{feat}} = 0$	$\lambda_{\text{arap}} = 0$	Extr. only	Intr. only	w/o normals	w/o MCMC	random rigid	spectral rec.
Avg. error	0.0056	0.0075	0.0066	0.0441	0.0205	0.0076	0.1039	0.0694	0.0098
Failure rate	0	0.2083	0.0833	0.7500	0.4028	0.0694	0.8611	0.3889	0.1111
Avg. Distortion	0.1654	0.1641	0.1926	0.1710	0.2239	0.1716	0.3485	0.2829	0.1666

Table 1: Ablation study on TOSCA and SCAPE. We turn off certain parts of the method or replace it with an alternative to assess its necessity and compare the average geodesic error in % of the diameter, the failure rate and the average conformal distortion for each setting. The failure rate is the number of pairs in % where the geodesic error is twice as high than (Ours). $\lambda_{\text{feat}}, \lambda_{\text{arap}} = 0$ turns off the regularizers, *Extr./Intr. only*, *w/o normals* removes one part of the embedding (see Eq. (10)), *w/o MCMC* removes the initialization, *random rigid* replaces our rigid alignment strategy with random rigid poses, and *spectral rec.* replaces smooth shells with spectral reconstruction (see Eq. (7)). The main insight is that the accuracy decreases whenever one of the components is removed. The conformal distortion is rather stable, except when a big percentage of the results are totally broken (e.g. w/o MCMC) or the as-rigid-as-possible regularizer is removed.

7. Conclusion

We have presented a novel approach to shape correspondence that combines geometric and spectral alignment by embedding the input shapes into an extrinsic-intrinsic product space. Our method introduces smooth shells as a coarse-to-fine shape approximation with minimal geometry changes between iterations. This is valuable for hierarchical approaches. Furthermore, we solve the problem of self-similarities by starting with an efficient surrogate based Markov chain Monte Carlo approach in which the deformation energy is used to find the optimal initialization. Finally, our method produces state-of-the-art results on established isometry datasets as well as two datasets which focus on specific noise, namely different meshing and topology

changes. In the FAUST real-scan interclass challenge we are on par with the state-of-the-art although we do not train specifically for this set-up. All results were achieved with the same set of parameters which shows great generality.

Acknowledgements

We would like to thank Emanuele Rodolà for useful discussions and Jing Ren for help with running the ZoomOut experiments. We gratefully acknowledge the support of the ERC Consolidator Grant "3D Reloaded".

References

- [1] Hierarchical functional maps between subdivision surfaces. *Computer Graphics Forum*, 35(5):55 – 73, 2019.

- [2] Non-rigid registration under anisotropic deformations. *Computer Aided Geometric Design*, 71:142 – 156, 2019.
- [3] Ibraheem Alhashim, Kai Xu, Yixin Zhuang, Junjie Cao, Patricio Simari, and Hao Zhang. Deformation-driven topology-varying 3d shape correspondence. *ACM Trans. Graph.*, 34(6):236:1–236:13, Oct. 2015.
- [4] Brian Amberg, Sami Romdhani, and Thomas Vetter. Optimal step nonrigid icp algorithms for surface registration. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.
- [5] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. Scape: shape completion and animation of people. In *ACM transactions on graphics (TOG)*, volume 24, pages 408–416. ACM, 2005.
- [6] Federica Bogo, Javier Romero, Matthew Loper, and Michael J. Black. FAUST: Dataset and evaluation for 3D mesh registration. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Piscataway, NJ, USA, June 2014. IEEE.
- [7] Federica Bogo, Javier Romero, Gerard Pons-Moll, and Michael J. Black. Dynamic FAUST: Registering human bodies in motion. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [8] Mario Botsch and Olga Sorkine. On linear variational surface deformation methods. *IEEE transactions on visualization and computer graphics*, 14(1):213–230, 2007.
- [9] Mario Botsch, Robert Sumner, Mark Pauly, and Markus Gross. Deformation transfer for detail-preserving surface editing. In *Vision, Modeling & Visualization*, pages 357–364. Citeseer, 2006.
- [10] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer, 2008. http://tosca.cs.technion.ac.il/book/resources_data.html.
- [11] Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. Topology-invariant similarity of nonrigid shapes. *International Journal of Computer Vision*, 81:281–301, 2009.
- [12] Junjie Cao, Andrea Tagliasacchi, Matt Olson, Hao Zhang, and Zhinxun Su. Point cloud skeletons via laplacian based contraction. In *Proceedings of the 2010 Shape Modeling International Conference*, Shape Modeling International Conference (SMI), pages 187–197, Washington, DC, USA, 2010. IEEE Computer Society.
- [13] Etienne Corman, Justin Solomon, Mirela Ben-Chen, Leonidas Guibas, and Maks Ovsjanikov. Functional characterization of intrinsic and extrinsic geometry. *ACM Trans. Graph.*, 36(2), Mar. 2017.
- [14] RM Dyke, C Stride, Y-K Lai, Paul L Rosin, Mathieu Aubry, Amit Boyarski, Alexander M Bronstein, Michael M Bronstein, Daniel Cremers, Matthew Fisher, et al. Shape correspondence with isometric and non-isometric deformations. 2019.
- [15] Marvin Eisenberger, Zorah Löhner, and Daniel Cremers. Divergence-free shape correspondence by deformation. In *Computer Graphics Forum*, volume 38, pages 1–12. Wiley Online Library, 2019.
- [16] Danielle Ezuz and Mirela Ben-Chen. Deblurring and denoising of maps between shapes. *Computer Graphics Forum (CGF)*, 36(5), August 2017.
- [17] Danielle Ezuz, Behrend Heeren, Omri Azencot, Martin Rumpf, and Mirela Ben-Chen. Elastic correspondence between triangle meshes. In *Computer Graphics Forum (CGF)*, volume 38, 2019.
- [18] Thibault Groueix, Matthew Fisher, Vladimir G. Kim, Bryan C. Russell, and Mathieu Aubry. 3d-coded: 3d correspondences by deep deformation. In *The European Conference on Computer Vision (ECCV)*, September 2018.
- [19] Igor Guskov, Wim Sweldens, and Peter Schröder. Multiresolution signal processing for meshes. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 325–334. Citeseer, 1999.
- [20] Dirk Haehnel, Sebastian Thrun, and Wolfram Burgard. An extension of the icp algorithm for modeling nonrigid objects with mobile robots. In *IJCAI*, volume 3, pages 915–920, 2003.
- [21] Kai Hormann and Günther Greiner. Mips: An efficient global parametrization method. Technical report, Erlangen-Nuernberg University (Germany) Computer Graphics Group, 2000.
- [22] Vladimir G Kim, Yaron Lipman, and Thomas A Funkhouser. Blended intrinsic maps. *Transactions on Graphics (TOG)*, 30(4), 2011.
- [23] Leif Kobbelt, Swen Campagna, Jens Vorsatz, and Hans-Peter Seidel. Interactive multi-resolution modeling on arbitrary meshes. In *ACM SIGGRAPH*, volume 98, pages 105–114, 1998.
- [24] Zorah Löhner, Emanuele Rodolà, Michael M Bronstein, Daniel Cremers, Oliver Burghard, Luca Cosmo, Andreas Dieckmann, Reinhard Klein, and Yusuf Sahillioglu. Shrec’16: Matching of deformable shapes with topological noise. *Proceedings of Eurographics Workshop on 3D Object Retrieval (3DOR)*, 2:11, 2016.
- [25] Bruno Lévy and Hao Richard Zhang. Spectral mesh processing. In *ACM SIGGRAPH 2010 Courses*, page 8. ACM, 2010.
- [26] Chun-Liang Li, Tomas Simon, Jason Saragih, Barnabás Póczos, and Yaser Sheikh. Lbs autoencoder: Self-supervised fitting of articulated meshes to point clouds. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11967–11976, 2019.
- [27] Hao Li, Robert W. Sumner, and Mark Pauly. Global correspondence optimization for non-rigid registration of depth scans. *Computer Graphics Forum (CGF)*, 27(5), 2008.
- [28] Or Litany, Emanuele Rodolà, Alex Bronstein, and Michael Bronstein. Fully spectral partial shape matching. *Computer Graphics Forum*, 36(2):1681–1707, 2017.
- [29] Or Litany, Emanuele Rodolà, Alex M Bronstein, Michael M Bronstein, and Daniel Cremers. Non-rigid puzzles. *Computer Graphics Forum (CGF), Proceedings of Symposium on Geometry Processing (SGP)*, 35(5), 2016.
- [30] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. SMPL: A skinned multi-person linear model. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, 34(6):248:1–248:16, Oct. 2015.

- [31] Jiayi Ma, Ji Zhao, Jinwen Tian, Alan L Yuille, and Zhuowen Tu. Robust point matching via vector field consensus. *IEEE Transactions on Image Processing*, 23(4):1706–1721, 2014.
- [32] Jiayi Ma, Ji Zhao, and Alan L Yuille. Non-rigid point set registration by preserving global and local structures. *IEEE Transactions on Image Processing*, 25(1):53–64, 2016.
- [33] Riccardo Marin, Simone Melzi, Emanuele Rodolà, and Umberto Castellani. FARM: functional automatic registration method for 3d human bodies. *CoRR*, abs/1807.10517, 2018.
- [34] Simone Melzi, Riccardo Marin, Emanuele Rodolà, and Umberto Castellani. Matching humans with different connectivity. *Proceedings of Eurographics Workshop on 3D Object Retrieval (3DOR)*, 2019.
- [35] Simone Melzi, Jing Ren, Emanuele Rodola, Maks Ovsjanikov, and Peter Wonka. Zoomout: Spectral upsampling for efficient shape correspondence. *arXiv preprint arXiv:1904.07865*, 2019.
- [36] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, 2010.
- [37] Dorian Nogneng, Simone Melzi, Emanuele Rodolà, Umberto Castellani, Michael M Bronstein, and Maks Ovsjanikov. Improved functional mappings via product preservation. *Computer Graphics Forum (CGF)*, 37(2), May 2018.
- [38] Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG)*, 31(4):30, 2012.
- [39] Beresford N Parlett. *The symmetric eigenvalue problem*, volume 20. siam, 1998.
- [40] Ulrich Pinkall and Konrad Polthier. Computing discrete minimal surfaces and their conjugates. *EXPERIMENTAL MATHEMATICS*, 2:15–36, 1993.
- [41] Gerard Pons-Moll, Javier Romero, Naureen Mahmood, and Michael J. Black. Dyna: A model of dynamic human shape in motion. *ACM Transactions on Graphics, (Proc. SIGGRAPH)*, 34(4):120:1–120:14, aug 2015.
- [42] Jing Ren, Adrien Poulenard, Peter Wonka, and Maks Ovsjanikov. Continuous and orientation-preserving correspondences via functional maps. *ACM Trans. Graph.*, 37(6):248:1–248:16, Dec. 2018.
- [43] Dennie Reniers, Jarke van Wijk, and Alexandru Telea. Computing multiscale curve and surface skeletons of genus 0 shapes using a global importance measure. *IEEE Transactions on Visualization and Computer Graphics*, 14(2):355–368, Mar. 2008.
- [44] Emanuele Rodolà, Michael Moeller, and Daniel Cremers. Point-wise Map Recovery and Refinement from Functional Correspondence. In *Vision, Modeling and Visualization*. The Eurographics Association, 2015.
- [45] Emanuele Rodolà, Luca Cosmo, Michael Bronstein, Andrea Torsello, and Daniel Cremers. Partial functional correspondence. *Computer Graphics Forum (CGF)*, 2016.
- [46] Yusuf Sahillioğlu. Recent advances in shape correspondence. *The Visual Computer*, pages 1–17, 2019.
- [47] Joaquim Salvi, Carles Matabosch, David Fofi, and Josep Forst. A review of recent range image registration methods with accuracy evaluation. *Image Vision Comput.*, 25(5):578–596, 2007.
- [48] Olga Sorkine and Marc Alexa. As-rigid-as-possible surface modeling. In *Symposium on Geometry processing*, volume 4, pages 109–116, 2007.
- [49] Robert W. Sumner, Johannes Schmid, and Mark Pauly. Embedded deformation for shape manipulation. 2007.
- [50] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer graphics forum*, volume 28, pages 1383–1392. Wiley Online Library, 2009.
- [51] Andrea Tagliasacchi, Thomas Delame, Michaela Spagnuolo, Nina Amenta, and Alexandru Telea. 3d skeletons: A state-of-the-art report. In *Computer Graphics Forum (CGF) (Proceedings of Eurographics)*, volume 35, 2016.
- [52] Gary K. L. Tam, Zhi-Quan Cheng, Yu-Kun Lai, Frank C. Langbein, Yonghuai Liu, David Marshall, Ralph R. Martin, Xian-Fang Sun, and Paul L. Rosin. Registration of 3d point clouds and meshes: A survey from rigid to nonrigid. *IEEE Transactions on Visualization and Computer Graphics*, 19(7):1199–1217, July 2013.
- [53] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In *Proceedings of European Conference on Computer Vision (ECCV)*, 16(9):356–369, 2010.
- [54] Bruno Vallet and Bruno Lévy. Spectral geometry processing with manifold harmonics. In *Computer Graphics Forum*, volume 27, pages 251–260. Wiley Online Library, 2008.
- [55] Oliver van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. A survey on shape correspondence. *Computer Graphics Forum*, 30(6):1681–1707, 2011.
- [56] Matthias Vestner, Zorah Löhner, Amit Boyarski, Or Litany, Ron Slossberg, Tal Remez, Emanuele Rodolà, Alex M. Bronstein, Michael M. Bronstein, Ron Kimmel, and Daniel Cremers. Efficient deformable shape correspondence via kernel matching. In *International Conference on 3D Vision (3DV)*, October 2017.
- [57] Matthias Vestner, Roei Litman, Emanuele Rodolà, Alex M Bronstein, and Daniel Cremers. Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.