A Dual Iterative Refinement Method for Non-rigid Shape Matching

Rui Xiang
Department of Mathematics
UC Irvine
xiangr1@uci.edu

Rongjie Lai
Department of Mathematics
Rensselaer Polytechnic Institute
lair@rpi.edu

Hongkai Zhao
Department of Mathematics
Duke University
zhao@math.duke.edu

Abstract

In this work, a simple and efficient dual iterative refinement (DIR) method is proposed for dense correspondence between two nearly isometric shapes. The key idea is to use dual information, such as spatial and spectral, or local and global features, in a complementary and effective way, and extract more accurate information from current iteration to use for the next iteration. In each DIR iteration, starting from current correspondence, a zoom-in process at each point is used to select well matched anchor pairs by a local mapping distortion criterion. These selected anchor pairs are then used to align spectral features (or other appropriate global features) whose dimension adaptively matches the capacity of the selected anchor pairs. Thanks to the effective combination of complementary information in a data-adaptive way, DIR is not only efficient but also robust to render accurate results within a few iterations. By choosing appropriate dual features, DIR has the flexibility to handle patch and partial matching as well. Extensive experiments on various data sets demonstrate the superiority of DIR over other state-of-the-art methods in terms of both accuracy and efficiency.

1. Introduction

Nonrigid shape matching is one of the most basic and important tasks in computer vision and shape analysis, e.g., shape registration, comparison, recognition, and retrieval. Different from 2D images, shapes are usually represented as 2-dimensional manifolds embedded in $\mathbb{R}^3$. Many challenges for nonrigid shape matching come from embedding ambiguities, i.e. shapes sharing exactly the same metric (isometry) or very similar metrics (nearly isometry) can have drastically different representations in xyz-coordinates in $\mathbb{R}^3$.

To overcome these representation ambiguities, one of the commonly used strategies is to extract features which are isometrically invariant and robust to small perturbations. Along with the idea, many successful and popular approaches are based on spectral geometry [40, 31, 53, 11, 28, 44]. Theoretically, the Laplace-Beltrami (LB) operator is isometrically invariant. As a generalization of Fourier basis functions from Euclidean domains to manifolds, the eigensystem of LB operator provides complete intrinsic spectral information of the underlying manifold. Moreover, from lower eigen-modes to higher eigen-modes, LB eigen-functions also provide a multi-scale characterization of the underlying manifold from coarse to fine resolution. Although using spectral geometry removed possible non-rigid embedding ambiguities, new ambiguities emerge in the spectral domain due to non-uniqueness of the LB eigen-system, e.g., sign ambiguity for eigen-functions, the ambiguity of choosing a basis for the LB eigen-space corresponding to a non-simple LB eigen-value (due to symmetry), the ambiguity of ordering for close eigen-values (due to small perturbations). To handle these ambiguities and use spectral features accurately and robustly, a proper linear transformation (a rigid transformation for exact isometry) needs to be found to align the spectral modes between two shapes first. This linear transformation is typically computed through some matching/correlation based on given (prior) correspondence, e.g., landmarks [36, 2, 29]. Moreover, to resolve fine details and acquire accurate correspondence between two shapes, high eigen-modes need to be used. However, the use of more and higher eign-modes will not only require more computation costs, but, more importantly, can also cause instability with respect to small perturbations.

To tackle the instability issue of using high eigen-modes directly, one natural multi-scale approach is to start from a correspondence at a coarse scale using a few low modes and iteratively refine the correspondence at a finer and finer scale by adding more and more, higher and higher modes gradually. The main motivation is that the linear transformation (a small matrix) on a coarse scale between two truncated spectral spaces spanned by a few low eigen-modes can be determined efficiently and stably from an initial approximate or limited correspondence. Once low eigen-modes are aligned well, an improved correspondence, es-
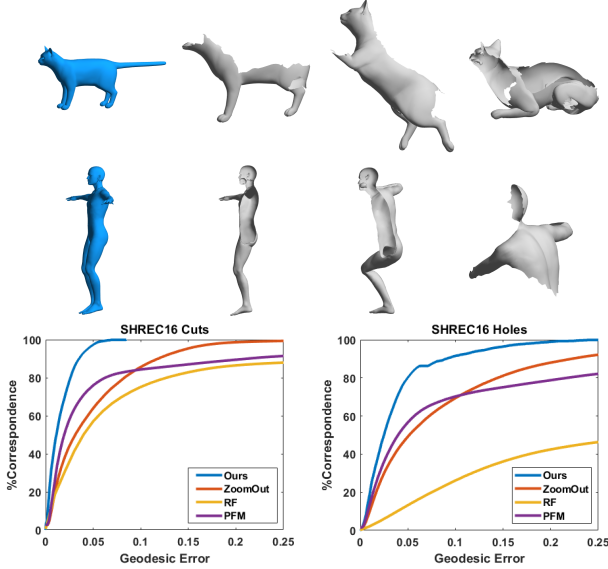
Figure 1: *Top:* First row shows examples from SHREC'16 holes, and second row shows examples from SHREC'16 cuts. We map the partial shapes in gray to full shapes in blue. *Bottom:* Geodesic error on SHREC'16 cuts and holes data set with comparison to other state-of-the-art methods.

pecially between smooth parts of the two shapes, is likely obtained. The improved correspondence is then used to determine the linear transformation for the next iteration which involves more and higher eigen-modes. Such a multi-scale idea for shape correspondence has been proposed in [29] for multi-scale registration using rotation-invariant sliced-Wasserstein distance and in [35] as a Zoom-out process. However, for the above straightforward multi-scale approach in the spectral domain, there are two key issues. First, in each iteration, the determination of the linear transformation between the truncated spectral spaces of two shapes using current correspondence of all points, many of which are incorrect, indiscriminately might be problematic. In Theorem 1, we show that the linear transformation between two spectral spaces determined using all points from an inaccurate correspondence will most likely lead to errors. These errors could be very significant and can cause either a failure for later refinement or slow convergence shown in the appendix. The other issue is the lack of a systematic and data-adaptive way to determine how many eigen-modes can be aligned accurately and stably by the current correspondence. Thus, it is hard to decide what is the appropriate jump in the number of eigen-modes for the next refinement after each iteration to achieve fast convergence. Previously, an increment of one mode was typically used to maintain convergence in Zoom-Out [35], while a prefixed sequence of eigenmodes was proposed in [29] based on the rule of thumb.

In this work, we propose a simple and efficient dual iterative refinement strategy. This method combines complementary information, such as spatial and spectral, or local and global, to simultaneously refine the correspondence between approximately isometric shapes in an effective way. More specifically, spatial matching via local mapping distortion (see Definition 1) is applied point-wisely to choose well-matched correspondence–anchor pairs, at the current stage. Once anchor pairs are selected, they are used to find 1) the maximal dimension of spectral space that can be robustly and accurately determined by these anchor pairs based on their distribution, and 2) the linear transformation that aligns spectral basis at the next scale, in spectral mapping which will lead to a refined correspondence at next stage. This remarkable simple strategy addresses the aforementioned two critical issues in previous multi-scale approaches using only spectral matching and allows one to utilize all well-matched pairs from the current step to jump to the next step in an accurate, efficient, stable, and data-adaptive way. We use extensive numerical experiments to show that our simple strategy, DIR, outperforms start-of-the-art model-based methods in terms of both accuracy and efficiency markedly. By choosing appropriate and application-specific dual features, DIR enjoys the flexibility to handle different scenarios, such as raw point clouds, patch and partial matching (as illustrated in Figure 1). We defer a detailed explanation of this Figure in Section 5.

Contributions. We summarize our main contributions of this paper as follows:

- By combining spatial matching via local mapping distortion and spectral matching via LB eigen-basis, the proposed DIR goes from coarse to fine resolution stably and effectively in a data-adaptive way.

- By generalizing this dual refinement strategy to other local and global features such as local mapping distortion and global geodesic distance, DIR handles patch and partial matching.

The rest of this paper is organized as follows. In Section 2, we give a brief review of some related work. Then we present our proposed method, DIR, in detail followed by a discussion about a few possible extensions of our method in Section 3. After that, we conduct extensive experiments on various benchmark data sets and compare DIR to other state-of-the-art methods in terms of both accuracy and efficiency in Section 5. We conclude the paper in Section 6.

2. Related work

Deigning good shape descriptors is crucial for shape registration or correspondence. In general, descriptors can be

categorized as pointwise or pairwise. Extrinsic pointwise descriptors [50, 21, 15, 41] are easy to compute but usually not very accurate, especially when non-rigid transformation is involved. To deal with non-rigid transformation, different intrinsic descriptors either in spatial domain, such as geodesics distance signatures [54], heat kernel signatures [48] and wave kernel signatures [4], or in spectral domain using eigensystem of LB operator [31, 53, 11, 28, 36, 29], are proposed. Then various nearest neighbor searching or linear assignment methods are used in the descriptor space to find the dense point correspondence. For point-wise descriptors in spectral domain, such as functional map [36], a proper linear transformation is needed to align spectral basis to remove ambiguities in the eigensystem of LB operator. The determination of the linear transformation itself is typically based on some prior knowledge, e.g., given landmarks, region correspondence or orientation preserving properties [39], and can be challenging to achieve good accuracy especially when the dimension of spectral embedding is high, which is needed to resolve fine shape features to achieve accurate correspondence. On the other hand, spatial domain point-wise descriptors are usually defined smoothly on shapes which is difficult to provide accurate correspondence as well. Our method aims at tackling these issues raised in aligning shapes using pointwise descriptors.

Intrinsic pairwise descriptors, such as pairwise geodesic distance [7, 10, 56] and kernel functions [33, 46, 55] have also been proposed for non-rigid shape correspondence problems. Although the matching of pairwise descriptors is stricter and may be more robust and accurate, it requires to solve a quadratic assignment problem (QAP) which is NP-hard. Various methods have been proposed to solve the QAP approximately in a more computational tractable way e.g sub-sampling [49], coarse-to-fine [57], geodesic distance sparsity enforcement methods [19] and various relaxation approaches [1, 8, 25, 30, 13, 17]. One popular approach is to relax the nonconvex permutation matrix (representing pointwise correspondence) constraint in the QAP to a doubly stochastic matrix (convex) constraint [1, 13]. However, both the pairwise descriptors and the doubly stochastic matrix are dense matrices, which make the relaxed QAP still challenging to solve even for a modest size problem. Recently, a novel local pairwise descriptor and a simple, effective iterative method to solve the resulting quadratic assignment through sparsity control was proposed in [58] for shape correspondence between two approximately isometric surfaces. The pairwise descriptor is based on the stiffness and mass matrix of finite element approximation of the LB operator, which is local in space, sparse to represent, and extremely easy to compute while containing global information. First, pairs with good correspondence based on local mapping distortion are selected as anchor pairs. Then a sparsity-enforced quadratic assign-

ment problem only in the neighborhoods of selected anchor pairs is solved. The two strategies lead to fast improvement and increase of anchor pairs while reducing the computation cost in each quadratic assignment iteration significantly.

3. Dual Iterative Refinement Method

3.1. Functional map

Spectral geometry is widely used in shape analysis [40, 31, 53, 11, 28, 36, 29, 45, 22]. As an important member in spectral geometry, functional map is introduced in [36] to solve non-rigid shape correspondence problem. It improves earlier point-based spectral methods [23, 24, 34, 37] which directly matches the spectral embeddings of shapes. We first provide a brief introduction to the concept of functional map which is used in our multi-scale dual iterative refinement process.

Given $(\mathcal{M}_1, g)$ as a closed 2-dimensional Riemannian manifold, the LB operator uniquely determines the underlying manifold up to isometry [6]. Eigen-functions of LB operator form an orthonormal basis on the underlying manifold and can be used as intrinsic and multi-scale descriptors for shapes. In discrete setting, the manifold $\mathcal{M}_1$ is usually represented as a triangulated mesh with vertices $\{\mathbf{x}_i\}_{i=1}^n$. The LB matrix is given by $\mathcal{L}_{\mathcal{M}_1} = \mathcal{A}_1^{-1}\mathcal{W}_1$ [38], where $\mathcal{A}_1$ is the diagonal element area matrix of $\mathcal{M}_1$ and $\mathcal{W}_1$ is the standard cotangent weight matrix. The discrete truncated k-dimensional spectral embedding of $\mathcal{M}_1$ can be expressed as a matrix $\Phi_{\mathcal{M}_1}^k \in \mathbb{R}^{n \times k}$ whose rows are the first k LB eigen-functions evaluated at each point, $\Phi_{\mathcal{M}_1}^k(x) = (\phi_1^{\mathcal{M}_1}(x), ..., \phi_k^{\mathcal{M}_1}(x))$. Similarly, we can define the embedding for $\mathcal{M}_2$ as $\Phi_{\mathcal{M}_2}^k = \{(\phi_1^{\mathcal{M}_2}(x), ..., \phi_k^{\mathcal{M}_2}(x)) \mid x \in \mathcal{M}_2\}$.

The key idea of functional map is to lift a shape correspondence $T : \mathcal{M}_1 \to \mathcal{M}_2$ to a linear map, functional map, between the function spaces $\mathcal{C}(\mathcal{M}_1)$ and $\mathcal{C}(\mathcal{M}_2)$. The functional map for a given correspondence T can be represented as a linear transformation between two given bases of $\mathcal{C}(\mathcal{M}_1)$ and $\mathcal{C}(\mathcal{M}_2)$, which becomes a matrix when the shapes are discretized and bases are truncated. In practice, LB eigen-functions are commonly used as the basis for the function space. For example, given a permutation matrix Π representing a point-to-point map T from $\mathcal{M}_1$ to $\mathcal{M}_2$, the functional map C is defined as:

$$C = \arg\min_C \|\Phi_{\mathcal{M}_1}C - \Pi\Phi_{\mathcal{M}_2}\|_F^2 = \Phi_{\mathcal{M}_1}^\dagger \Pi\Phi_{\mathcal{M}_2} \quad (1)$$

where $\dagger$ denotes the Moore-Penrose pseudo-inverse. Theoretically, LB eigen-functions are isometrically invariant. In other words, if $\mathcal{M}_1$ and $\mathcal{M}_2$ are isometric, then the corresponding LB eigen-functions are the same up to possible ambiguities caused by sign switch and non-simple eigenvalues, which form an orthonormal group. This motivates a

constrained version of the following functional map [29]

$$C = \arg\min_{C \in \mathfrak{R}} \|\Phi_{\mathcal{M}_1} C - \Pi \Phi_{\mathcal{M}_2}\|_F^2 = \text{Proj}_{\mathfrak{R}}(\Phi_{\mathcal{M}_1}^\top \Pi \Phi_{\mathcal{M}_2})$$

(2)

where we write the set of orthonormal matrices as $\mathfrak{R} = \{C \mid C^\top C = Id\}$ and the projection to $\mathfrak{R}$ is provided as $\text{Proj}_{\mathfrak{R}}(A) = UV^\top$ with a singular value decomposition (SVD) of $A = U\Sigma V^\top$. Since this paper aims at tackling nearly isometric shape matching problem, we adopt (2) which is equivalent to (1) for the isometric case.

Typically, the functional map is obtained via a least square optimization with various constraints and regularizations, such as preservation of given landmarks, communitivity with LB operator, and sparsity [46, 39]. Once the optimized functional map C is computed, the correspondence map T can be obtained by solving the following spectral matching problem:

$$T(p) = \arg\min_{q \in \mathcal{M}_2} \|\Phi_M(p) - \Phi_N(q)C^\top\|_F^2, \ \forall p \in \mathcal{M}_1$$

(3)

However, determination of an accurate functional map C can be difficult when limited prior knowledge or a poor correspondence is used. On one hand, it becomes harder when more eigen-modes are involved since the degrees of freedom of the functional map grow quadratically and high eigen-modes are less computationall stable with respect to perturbations or noises. On the other hand, confining to low eigen-modes limits resolution of the spectral representation which leads to inaccurate shape correspondence. A natural multi-scale idea is to start the functional map from a coarse resolution in spectral embedding involving a few low eigen-modes and construct the correspondence T. Then the constructed correspondence at coarse scale is used to help determining the functional map at a finer scale which leads to a more refined correspondence via spectral matching. This process can be iterated until a fine enough resolution is achieved. The main motivation is that the functional map at a coarse scale – a small matrix, can be determined efficiently and stably from an initial approximate or limited correspondence, which is then used to help improving the correspondence at next iteration. This strategy leads to iteratively computing (2) and (3), or (1) and (3), which is the key idea proposed in [29] using rotation-invariant sliced Wasserstein distance and in the Zoom-out process [35]. As a crucial component of these iterative refinement methods, the new functional map C is obtained from the current correspondence of all points. Since intermediate correspondence, especially at the beginning, can be quite inaccurate, this naive strategy may lead to significant errors.

Next, we provide both theoretical and experimental evidences on the effect of correspondence error on spectral alignment, i.e., the functional map C in (2), which can be problematic for any iterative refinement procedure based on spectral geometry.

Assume we have two perfectly isometric manifold $\mathcal{M}_1$ and $\mathcal{M}_2$ and their corresponding discrete truncated spectral embedding $\Phi_{\mathcal{M}_1}$, $\Phi_{\mathcal{M}_2} \in \mathbb{R}^{n \times k}$, where n is the number of points and k is spectral dimension. Without loss of generality, we assume Id is the ground truth correspondence between $\mathcal{M}_1$ and $\mathcal{M}_2$ (otherwise, we can shuffle row vectors of $\Phi_{\mathcal{M}_2}$ according to the ground truth correspondence). The ground truth functional map between $\Phi_{\mathcal{M}_1}$ and $\Phi_{\mathcal{M}_2}$ is an orthonormal matrix $C_T \in \mathfrak{R}$, i.e. $\Phi_{\mathcal{M}_2} = \Phi_{\mathcal{M}_1} C_T$ and $C_T = \text{Proj}_{\mathfrak{R}}(\Phi_{\mathcal{M}_1}^\top \Phi_{\mathcal{M}_2})$.

Theorem 1. *Given $\Phi_{\mathcal{M}_2} = \Phi_{\mathcal{M}_1} C_T$ with $C_T \in \mathfrak{R}$. Let's assume a one-to-one correspondence Π is an inaccurate correspondence which maps a portion of $\Phi_{\mathcal{M}_1}$ accurately to the corresponding part of $\Phi_{\mathcal{M}_2}$, while the rest part of Π is inaccurate. Without loss of generality, we write $\Phi_{\mathcal{M}_1} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ and $\Phi_{\mathcal{M}_2} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ where X_1, $Y_1 \in \mathbb{R}^{n_1 \times k}$, X_2, $Y_2 \in \mathbb{R}^{n_2 \times k}$ and $n_1 + n_2 = n$. We let $\Pi \Phi_{\mathcal{M}_2} = \begin{pmatrix} Y_1 \\ \sigma Y_2 \end{pmatrix}$ for a permutation matrix $\sigma \in \mathbb{R}^{n_2 \times n_2}$. Let*

$$C_a = \arg\min_{C \in \mathfrak{R}} \|\Phi_{\mathcal{M}_1} C - \Pi \Phi_{\mathcal{M}_2}\|_F^2 = \text{Proj}_{\mathfrak{R}}(\Phi_{\mathcal{M}_1}^\top \Pi \Phi_{\mathcal{M}_2})$$

Then it is most likely that the spectral norm $\|C_a - C_T\|_2 > 0$ with probability at least $1 - \eta$ where $\eta = \sum_{j=0}^{\lfloor n_2/2 \rfloor} \frac{1}{2^j j! (n_2 - 2j)!}$

Notice that η decreases rapidly as n_2 grows. For example, when $n_2 = 25$, $\eta \approx 10^{-12}$. We provide a detailed proof of this theorem in appendix A where we also numerically verify that using all points from an inaccurate correspondence may lead to significant errors in the functional map, which can cause failure for iterative refinement.

Another critical issue for multi-scale approach in spectral domain is how to increase the resolution, i.e., the dimension of spectral embedding, during the iteration in an accurate, stable and data adaptive way. Most approaches [55, 29, 35] just adopt an empirical or prefixed increasing sequence.

Motivated from the above difficulties and limitations of iterative refinement approaches in spectral domain, we propose a dual iterative refinement method which iteratively updates on the spatial and spectral domains. The first key idea is to choose well-matched pairs using local spatial matching from current correspondence, called anchor pairs, and only use them to determine the functional map by which will, in turn, help to construct a much more improved correspondence for the next iteration. In order to choose high-quality anchor pairs from a given correspondence, we zoom in at each corresponding pair and measure local mapping distortion. This step integrates local spatial information in

the spectral refinement process. The second key idea is to find the maximal dimension of spectral embedding that can be robustly and accurately determined by these anchor pairs according to their distribution in spectral domain using singular value analysis on their correlation matrix.

3.2. Local mapping distortion

In order to choose candidates for well-matched pairs, or to filter out wrong correspondences, to compute the functional map C more accurately, we use the following *local mapping distortion* introduced in [58] to measure the quality of matching of a given correspondence pair and choose those pairs with small local mapping distortion. With relying on any information of the ground truth correspondence, this distortion provides a quantitative measurement to check the accuracy of the map through its continuity and distance preserving properties. i.e. nearby points on the source surface should be mapped to nearby points on the target surface.

Definition 1 (Local mapping distortion). *Let $T : \mathcal{M}_1 \rightarrow \mathcal{M}_2$ be a map between two manifolds. For any point $\mathbf{x} \in \mathcal{M}_1$, consider its γ-geodesic ball in $\mathcal{M}_1$ as $\mathcal{B}_\gamma(\mathbf{x}) = \{\mathbf{y} \in \mathcal{M}_1 \mid d_{\mathcal{M}_1}(\mathbf{x}, \mathbf{y}) \leq \gamma\}$. Local mapping distortion of T at $\mathbf{x}$ is defined as:*

$$\mathcal{D}_\gamma(T)(\mathbf{x}) = \frac{1}{|\mathcal{B}_\gamma(\mathbf{x})|} \int_{\mathbf{y} \in \mathcal{B}_\gamma(\mathbf{x})} DE_T(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{y} \quad (4)$$

where $DE_T(\mathbf{x}, \mathbf{y}) = \frac{1}{\gamma}|d_{\mathcal{M}_1}(\mathbf{x}, \mathbf{y}) - d_{\mathcal{M}_2}(T(\mathbf{x}), T(\mathbf{y}))|$ and $|\mathcal{B}_\gamma|$ is the volume of $\mathcal{B}_\gamma$.

Based on the above definition of local mapping distortion, it is straightforward to check that the following properties are true:

1. If T is an isometric map, $\mathcal{D}_\gamma(T)(\mathbf{x}) = 0, \forall \mathbf{x} \in \mathcal{M}_1, \gamma > 0$.

2. If $\mathcal{D}_\gamma(T)(\mathbf{x}) = 0, \forall \mathbf{x} \in \mathcal{M}_1$ for some $\gamma > 0$, T is isometric.

In discrete setting, equation (4) can be approximated by:

$$\mathcal{D}_\gamma(T)(\mathbf{x}_i) \approx \frac{\sum_{\mathbf{x}_j \in \mathcal{B}_\gamma(\mathbf{x}_i), \mathbf{x}_j \neq \mathbf{x}_i} \mathcal{A}_1(j) DE_T(\mathbf{x}_i, \mathbf{x}_j)}{\left(\sum_{\mathbf{x}_j \in \mathcal{B}_\gamma(\mathbf{x}_i), \mathbf{x}_j \neq \mathbf{x}_i} \mathcal{A}_1(j)\right)} \quad (5)$$

where $\{\mathbf{x}_i\}_{i=1}^n$ and $\{\mathbf{y}_i\}_{i=1}^n$ represent $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively, and $\mathcal{A}_1$ is the area element of the discrete mesh representing $\mathcal{M}_1$.

Remark: The computation complexity of local mapping distortion is $O(n)$ as we only need to check a fixed radius neighborhood for each point. Moreover, by normalizing the
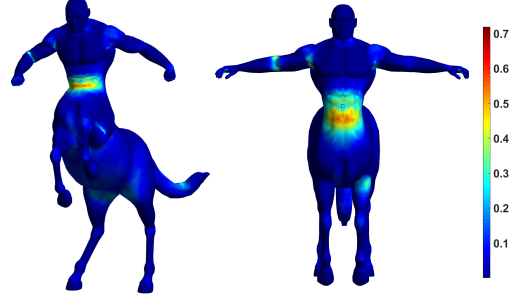

Figure 2: Local mapping distortion on ground truth correspondence: The color map indicates that local mapping distortion is small everywhere except for areas where non-isometric distortion is large.

distance with respect to γ, the local mapping distortion is robust to global scaling as well as local sampling variation. Figure 2 shows that, for the ground truth correspondence, local mapping distortion is very small almost everywhere except in regions where non-isometric distortion is large. Hence the proposed local mapping distortion also serves as a good unsupervised error metric when ground truth is not available.

3.3. Our Method

We first precompute K leading eigen-functions for the discrete LB operator for each shape. K can be determined by computation cost limitation, the noise level in the data, or desired accuracy of the correspondence. In any case, it should not exceed what the mesh resolution can support for each discretized shape, i.e., a few mesh points are needed in each nodal domain.

Our method can start with an initial correspondence provided by SHOT [50] based on extrinsic point-wise features, or a few given landmarks used as initial anchor pairs which can be fixed or updated in later refinement, or by any other (fast but not necessary very accurate) methods. Then we start DIR which includes the following three simple steps, explained in detail later, in each iteration.

- Choose anchor pairs from current correspondence using local mapping distortion criterion;

- Determine the spectral dimension and the corresponding functional map based on the selected anchor pairs;

- Update the correspondence using the updated functional map.

We enforce two stopping criteria, whichever is satisfied first. The first is the total number of iterations, which limit the computation cost. The second is when the spectral dimension supported by anchor pairs reaches K. Here are more detailed descriptions of each step.

5

Step 1 Selecting anchor pairs.

By setting a proper local mapping distortion threshold ϵ, the set of pairs $\{(x_\ell, T(x_\ell)|\mathcal{D}_\gamma(T)(\mathbf{x}_\ell) < \epsilon\}$ are selected as anchor pairs from the current correspondence which will be used to guide next refinement. It is important to note that both the threshold ϵ and anchor pairs are updated in later refinement. By decreasing ϵ with iterations, the quality of anchor pairs is also improved.

Step 2 There are two components described as follows.

(1) **Determining the proper spectral dimension based on anchor pairs.**

Given a set of anchor pairs, an important question is to find a proper spectral dimension that can be determined accurately and stably according to the distribution of the anchor pairs in spectral embedding. Due to possible degeneracy of the distribution, the dimension can be quite smaller than the number of anchor pairs. We use singular value decomposition (SVD)[51], a.k.a principal component analysis, to find the dimension well expanded by a set of anchor pairs in spectral domain. Consider the spectral embedding of a set of anchor pairs $\{x_\ell\}_{\ell=1}^m$ and $\{T(x_\ell)\}_{\ell=1}^m$ as $\Phi_{\mathcal{M}_1}^K(\{x_\ell\}_{\ell=1}^m)$, and $\Phi_{\mathcal{M}_2}^K(\{T(x_\ell)\}_{\ell=1}^m)$, respectively. We compute SVD of the correlation matrix of all anchor pairs

$$U\Sigma V^\top = \Phi_{\mathcal{M}_1}^K(\{x_\ell\}_{\ell=1}^m)^\top \Phi_{\mathcal{M}_2}^K(\{T(x_\ell)\}_{\ell=1}^m). \quad (6)$$

We threshold the singular values to determine the proper dimension. One can use a simple thresholding strategy. In our implementation, we adopt a normalization strategy to make the threshold adaptive to the data and noise level. After normalizing all singular values by the mean of 10 largest singular values, the dimension cut is set at where the sum of ten consecutive normalized singular values is smaller than a threshold (0.1 for all our tests). We point out that since the spectral embedding is defined by each eigen-function of the LB operator, instead of the collectively spanned space, the accuracy and stability of each singular vectors is more relevant in our case. Perturbation analysis for singular values and vectors in [32] can be used to justify our strategy as well.

(2) **Computing functional map based on anchor pairs.**

Once the proper spectral dimension, k, is determined by the selected anchor pairs $\{(x_\ell, T(x_\ell) \mid \ell = 1, \cdots, m\}$, we compute the functional map for k dimensional spectral embedding only based on the anchor pairs. In other words, we consider the following problem

$$C = \arg\min_{C\in\mathfrak{R}} \sum_{\ell=1}^m \| \Phi_{\mathcal{M}_1}^k(x_\ell) - \Phi_{\mathcal{M}_2}^k(T(x_\ell))C\|_F^2$$
$$= \mathrm{Proj}_{\mathfrak{R}} \left(\left(\Phi_{\mathcal{M}_1}^k(\{x_\ell\}_{\ell=1}^m)\right)^\top \Phi_{\mathcal{M}_2}^k(\{T(x_\ell)\}_{\ell=1}^m) \right) \quad (7)$$

This restricted version to compute a functional map has twofold advantages. On one hand, as we discussed in the

Section 3.1 and appendix, it excludes potential corruption from inaccurate correspondence and leads to a better estimation of the functional map. On the other hand, it reduces the computation cost.

Step 3 Construct the new correspondence.

Using the functional map computed from selected anchor pairs in the properly enlarged spectral embedding space, a refined correspondence is constructed by solving the assignment problem (3), where a KNN search method is applied.

In summary, the full procedure is described in Algorithm 1.

ALGORITHM 1: Dual Iterative Refinement (DIR)

input An initial correspondence, truncated discrete spectral embedding $\Phi_{\mathcal{M}_1}^K$ and $\Phi_{\mathcal{M}_2}^K$ through K LB leading eigen-functions for both shapes, local mapping distortion error threshold $\{\epsilon_i\}_1^N$, singular value threshold, maximum iteration number N,

repeat

1. Find anchor pairs $\{(x_\ell, T(x_\ell) \mid \mathcal{D}_\gamma(T)(x_\ell) < \epsilon_i\}, l = 1, \ldots, m_i$ via local mapping distortion criterion from current correspondence T.

2. Find the proper spectral dimension k_i from SVD of correlation matrix $\Phi_{\mathcal{M}_1}^K(\{x_\ell\}_{\ell=1}^{m_i})^\top \Phi_{\mathcal{M}_2}^K(T(\{x_\ell\}_{\ell=1}^{m_i}))$ of anchor pairs.

3. Update functional map
$C = \mathrm{Proj}_{\mathfrak{R}} \left(\left(\Phi_{\mathcal{M}_1}^{k_i}(\{x_\ell\}_{\ell=1}^{m_i})\right)^\top \Phi_{\mathcal{M}_2}^{k_i}(\{T(x_\ell)\}_{\ell=1}^{m_i}) \right)$

4. Update $T(p) = \arg\min_q \|\Phi_{\mathcal{M}_1}^{k_i}(p,:) - \Phi_{\mathcal{M}_2}^{k_i}(q,:)C^\top\|^2$ for all $p \in \mathcal{M}_1$.

until N is reached, or k_i reaches K

Computation Complexity The complexity for computing K leading eigen-vectors of a $n \times n$ sparse matrix corresponding to the discretized LB operator is $O(Kn\log n)$. The complexity for checking local mapping distortion is $O(n)$ since geodesic distance is only computed in a fixed neighborhood, e.g., first ring or second ring, at each point. To compute functional map, the complexity of matrix multiplication is at most $O(nK^2)$ and SVD decomposition is at most $O(K^3)$ since the number of anchor pairs is at most n and the spectral dimension k is at most K which is prefixed and far less than n. KNN search is used to solve Equation (3) which has a complexity of $O(n\log(n))$ [52]. Hence altogether, our method is of complexity $O(n\log(n))$. If we use other global features instead of spectral features, such as geodesic distance to anchor points, it is still $O(n\log(n))$ since we limit the maximal number of anchor points (as the number of LB eigen-functions). The complexity of computing geodesic distance and solving the resulting assignment problem is still $O(n\log(n))$.

4. Discussion

In this section, we discuss other possible extensions of our approach and a few specific applications.

4.1. Combination of local and global features

For shape correspondence problem, an efficient and robust approach should rely on both local and global features. So far we have mainly talked about using both spatial and spectral features, which are perfectly complementary in the sense of local and global information. Anchor pair selection by the local mapping distortion criterion is based on local spatial features while the spectral features are global. We would like to point out that the proposed strategy of dual iterative refinement process can be extended to other combinations of appropriate local and global features in different applications. For examples, one may alternatively use Heat Kernel Signature [12], Wave Kernel Signature [4] or Geodesic Distance Signature [54] as global features. For shapes with holes and boundaries, we choose Geodesic Distance Signature which is less sensitive to local mesh distortions and boundaries for most interior points. However, if accurate spectral information is available, taking advantage of the multi-scale representation in spectral embedding leads to better accuracy and efficiency according to our tests.

4.2. DIR with limited initial landmarks

In our previous discussion, the first collection of anchor pairs is selected from a given initial correspondence such as the one obtained from comparing SHOT features. Our method does enjoy the flexibility to incorporate given landmarks. This is applauded in applications like shape matching in the morphological study in medical imaging where landmarks could be annotated based on specific tasks [20]. Once a few human-annotated and required landmarks are given, which are taken as the initial anchor pairs and fixed in later iterations, DIR will converge to a stable solution. The fewer landmarks are provided, the more iterations are usually needed. The final convergence performance of our numerical tests, as shown in Figure 7, indicates that DIR can provide accurate correspondence based on only four landmarks and it is stable with respect to different initialization.

4.3. Matching of point clouds without mesh

In real applications, point clouds with well-constructed global triangulation is usually hard to obtain. For point clouds without global mesh, we use the local mesh method, first proposed in [27] and improved in [58], to compute the LB eigenvalues and local mapping distortion. DIR works well on matching raw point clouds as shown in our numerical experiments.

4.4. Patch/partial matching

Patch/partial matching often comes with difficulties due to artificial boundaries, different sizes and topological perturbations. Spectral information is either unavailable or unreliable to use. However, our local feature, local mapping distortion, is not affected by the above difficulties at interior points. By replacing spectral features with geodesic distance to anchor points, which is global and less sensitive to the above difficulties for most interior points, DIR can handle patch or partial matching well as shown in our numerical experiments.

5. Experiments

In this section, we conduct comprehensive experiments to evaluate the performance of our method on various data sets including TOSCA [9], SCAPE [3], SHREC'16 [14] and SHREC'20 [16] as well as patches extracted from TOSCA. TOSCA and SCAPE are data sets of approximately isometric shapes; SHREC'16 is also a data set of approximately isometric shapes but includes cuts and holes. SHREC'20 is a data set from real-world 3D scans which includes various deformation and noise.

In all experiments, no pre-processing, such as a low resolution model or pre-computing the geodesic distance matrix, is required in our algorithm. Raw point clouds (with or without meshes) are directly used. In our comparisons, all geodesic errors from existing methods are obtained from the associated error curve data appeared in the papers. Computation using ZoomOut is produced from the code on GitHub shared by the authors [35]. All experiments using our methods are conducted in Matlab with 16GB RAM and Intel i7-6800k CPU.

Error Metric We use the geodesic error as our error metric in most experiments. Suppose the constructed correspondence maps $\mathbf{x} \in \mathcal{M}_1$ to $\mathbf{y} \in \mathcal{M}_2$ while the true correspondence is $\mathbf{x}$ is to $\mathbf{y}^*$, we measure the quality of our result by computing the geodesic error defined as $e(x) = \frac{d_{\mathcal{M}_2}(y, y^*)}{diam(\mathcal{M}_2)}$, where $diam(\mathcal{M}_2)$ is the geodesic diameter of $\mathcal{M}_2$.

Hyperparameters We set neighborhood size for local mapping distortion criterion as the second ring; maximum iteration number is 10; and the local mapping distortion threshold is set to [0.26, 0.22, 0.18, 0.14, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1] at each iteration. The initial guess is given by the KNN search based on SHOT features [50]. These hyperparameters, especially the local mapping distortion, are not sensitive to different data sets. We use the same hyperparameters on TOSCA and SCAPE data sets. For SHREC'16, SHREC'20 and patch matching, we set the maximum number of anchor pairs served as the reference points for geodesic distance feature to 800.

TOSCA and SCAPE TOSCA data set consists of 76 shapes in 8 different categories. The vertex number is from

4k to 50k. It contains both human and animal shapes. We present the results of three experiments on TOSCA data. The first one uses point clouds with its original mesh as input, and computes 1000 spectral basis as the maximal spectral dimension; the second one also uses point clouds with its original mesh as input but computes 500 spectral basis as the maximal spectral dimension; the last one uses only point clouds (no mesh) as input, and computes 1000 spectral basis (using local mesh method) as the maximal spectral dimension. SCAPE data set has 72 shapes of the same person with different poses, and each contains 12,500 vertices. We perform the same experiments as on TOSCA data.

The results on TOSCA and SCAPE data sets are illustrated in Figure 3. We compare our method with the following methods: Blended [26], Best Conformal [26], GMDS [8], Kernel Marching [55], SEQA [58], RSWD [29], ZoomOut [35], and Divergence-Free Shape Interpolation [18]. Our method with 1000 spectral basis outperforms all methods, our method with 500 spectral basis outperforms almost all methods, and our method on purely point cloud inputs still achieves high accuracy. Moreover, the run time of our method DIR is also significantly less than other methods (see Table 1).

SHREC'16   SHREC'16 Partial Correspondence benchmark data set consists of 8 types of isometric human or animal shapes in different poses with regular cuts and irregular holes. We test our method by matching each partial shape to the corresponding full shape. Since spectral basis is sensitive to mesh 'cuts' and 'holes', we use geodesic distance to anchor points as global features. We still use SHOT for the initialization.

The results are illustrated in Figure 1. We compare our method with ZoomOut [35], Partial Functional Maps [42] and Random Forests [43]. Since SHREC'16 only provides ground truth map with barycentric coordinates, we compute the geodesic error at each point as the weighted average of geodesic errors to the three vertices in the target triangle element. The results show that our method is quite flexible and robust to handle shape matching for different scenarios. It again outperforms other state-of-the-art methods markedly and achieves high accuracy on this challenging data set.

SHREC'20   SHREC'20 is a data set from 3D scans of a real-world stuffed toy rabbit that contains deformations caused by stretching, indent, twisting and inflation. To make it even more challenging, the rabbit is inflated by different internal materials and there are partial scans and missing parts caused by occlusion. Since SHREC'20 does not provide ground truth, hence we only show the visual result as illustrated in Figure 4. We also present final local mapping distortion heat maps to indicate local deformation at each point. These heat maps indicate that our method can still produce accurate results although the target shape may have nontrivial metric deformation from the
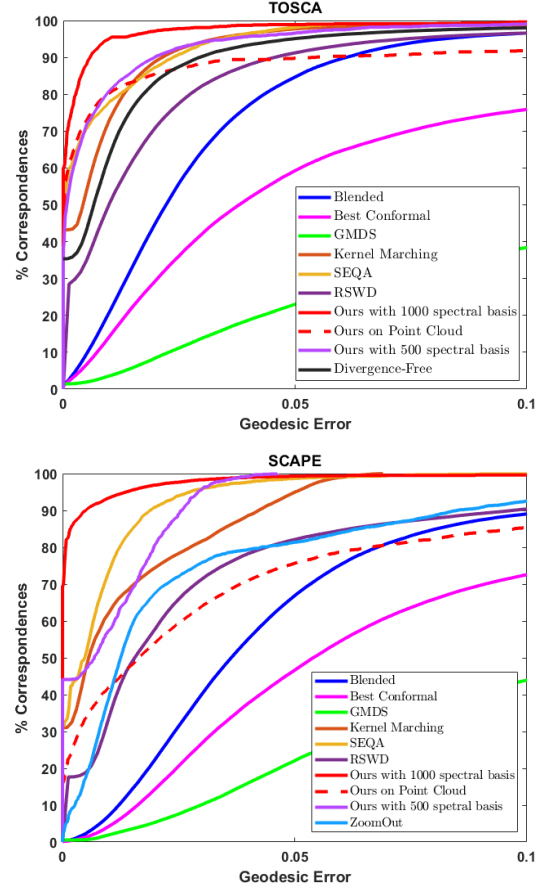

Figure 3: Geodesic error on the TOSCA and SCAPE data sets with comparison to other methods.

source shape. Moreover, the local mapping distortion provides a very good measurement to quantify the mapping quality without knowing the ground truth correspondence. As we can observe from the second row in Figure 4, the mapping quality of shapes in the second and fourth columns is satisfactory while the third one has relatively large distortion in some regions. This is because the target shape is deformed from the source shape by a very distorted metric deformation. In our future work, we would certainly extend our investigation to finding high-quality mappings between shapes with large intrinsic deformations.

Patch Matching   Since there is no standard data set for patch matching, so instead we report several experiments using patches cut from TOSCA data set. The first case contains two examples with artificial boundaries, different sizes (partial patching) and topological changes simultaneously. In the first example, we take a portion of an arm (finger tips also removed) from one of two isometric centaur shapes, and then map the partial arm onto the other entire centaur. In the second example, we map a body patch onto the whole
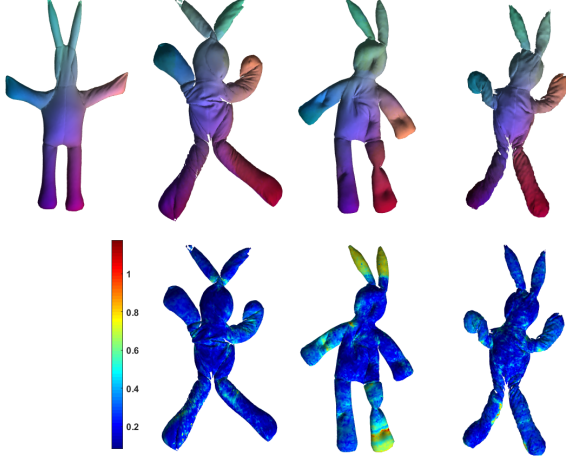
Figure 4: Qualitative result of our method on SHREC'20. In the first row, the rabbit in the first column is mapped to the rest rabbits in the first row. In the second row, final local mapping distortion is computed on the target shapes as a heat map. It is clear that area where local mapping distortion is lower shows a better visual mapping result.

shape. The original centaur is a closed mesh surface with no holes, while the arm and the body patch are not closed and have holes and boundaries. Our simple method again performs well for this challenging example as shown in Figure 5.

The second case is matching two patches which contain both overlap and non-overlap parts. We match an arm without the hand to a portion of an arm with the hand, where forearm is the common part. Since there is no correspondence between the non-overlap parts, a post local mapping distortion test is added to prune out those points. The result is shown in Figure 6. Our method still performs quite well. These experiments on patch matching indicate that our method is robust to size differences, artificial boundaries and topological changes.

Experiments Using Limited Landmarks We select 4 pairs of centaur shapes from TOSCA, initialize our method with different number of randomly chosen landmark pairs (without using SHOT for initialization) and then plot the average geodesic error curve. These landmark pairs are fixed and always belong to our selected anchors pairs in later iterations. After different number of iterations, the final performance is illustrated in Figure 7. Even with four initial random landmarks, the performance is excellent although more iterations are needed. It shows stability of our method with respect to initial correspondence.

Growth of Anchor Pairs and Spectral Dimensions One key feature of our method is the selection of well matched pairs which are used to determine spectral dimension and the corresponding functional map at next iteration in an au-
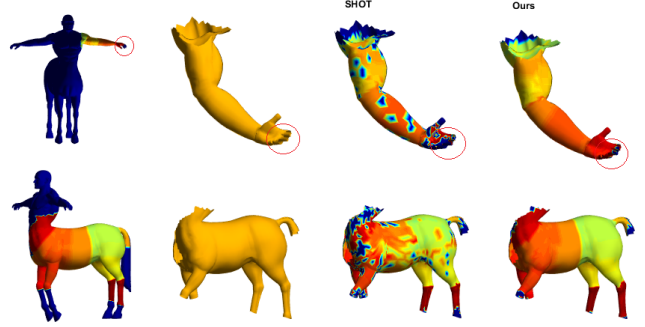

Figure 5: Example of patch matching. Removed finger tips are highlighted by red circles. First column is the entire centaur and its non-blue area is the ground truth color map. Extra points are colored in blue. Arm or body part in the second column is mapped onto the entire shape. The third column is the color mapping result using SHOT, and the last column is the color mapping result using our method.
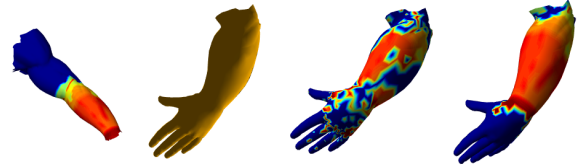

Figure 6: Examples of patch matching. Patch in the second column is mapped to the patch in the first column. Extra points in the first patch is colored in blue and its non-blue area is the ground truth color map. The third column shows mapping result using SHOT, and the last column is the mapping result using our method. Blue color in column three and four indicates points which do not pass the post local mapping distortion test.
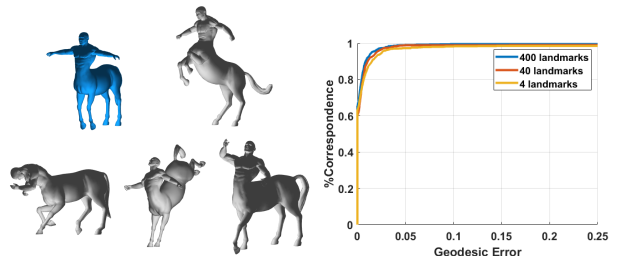

Figure 7: Average convergence performance of our method testing on several pairs of centaur shapes from TOSCA (blue centaur is mapped to gray centaurs), given different number of initial random annotated landmarks.

tomatic and data-adaptive way. A steady and fast growth and improvement (by decreasing the mapping distortion threshold with iterations) of anchors pairs (updated in each iteration) means a steady and fast growth of the spectral di-
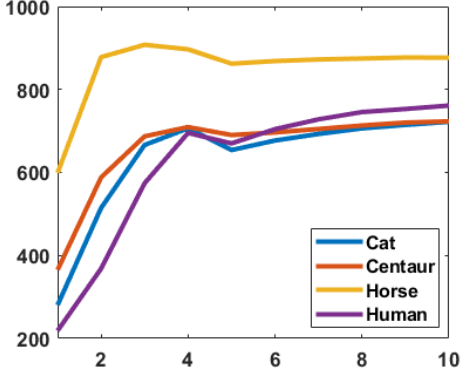
Figure 8: The average number of used eigenfunctions (with a maximal number of eigenfunctions set to 1000) for cat, horse centaur class from TOSCA data set and human from SCAPE data set at different iteration.

Iteration 1

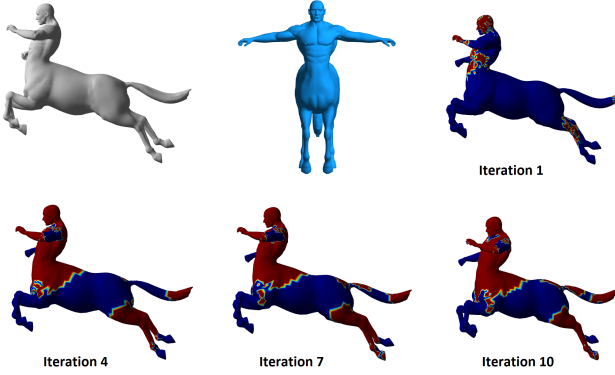Iteration 4   Iteration 7   Iteration 10

Figure 9: We map the gray centaur to blue centaur and record the area of anchor pairs in each iteration where the selected anchor pairs are colored in red. The anchor pairs detected at iteration 1 are directly from matching SHOT descriptor.

mension and increase of resolution which leads to a steady and fast refinement of the correspondence. Figure 8, shows the average number of used LB eigenfunctions (spectral dimensions) for several class of shapes from TOSCA or SPACE data set and Figure 9 shows the growth pattern of anchor pairs. Both results indicate that our method can extract anchor pairs and thus use spectral information in a steady and fast way.

Run Time   In Table 1, we show the average run time of our method on several examples from TOSCA data set which includes shapes with vertices ranging from 4,000 to more than 50,000 and indicates the computation efficiency of our methods. CUP time of our method in the table reports total time of all steps including computation of LB eigenfunctions. As mentioned in Section 3.3, the complexity of

our method is $O(n \log(n))$. But in practice, when computing point clouds with more than 20,000 vertices, our computer suffers a computation speed slow-down from the vast RAM usage because of our limited RAM capacity. Hence, for Cat and David example from TOSCA, the run time is higher than expected. Most state-of-the-art approaches have a computation complexity of $O(n^2)$ and do not report run time for shapes over 10,000 vertices. Our method is very efficient compared with state-of-the-art methods which report the time.

6. Conclusion

We propose a simple and efficient iterative refinement strategy utilizing dual features, spatial and spectral or local and global. By only relying on selected well-matched pairs (selected from local mapping distortion using local features) to guide the next refinement (using spectral or other global features), our proposed method combines complementary information in an optimal and data-adaptive way. Our method shows superior performance on extensive tests compared to other state-of-the-arts methods in terms of accuracy, efficiency and stability.

Acknowledgement

References

[1] Yonathan Aflalo, Alexander Bronstein, and Ron Kimmel. On convex relaxation of graph isomorphism. *Proceedings of the National Academy of Sciences*, 112(10):2942–2947, 2015. 3

[2] Yonathan Aflalo, Anastasia Dubrovina, and Ron Kimmel. Spectral generalized multi-dimensional scaling. *International Journal of Computer Vision*, 118(3):380–392, 2016. 1

[3] Dragomir Anguelov, Praveen Srinivasan, Hoi-Cheung Pang, Daphne Koller, Sebastian Thrun, and James Davis. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *Advances in neural information processing systems*, pages 33–40, 2005. 7

[4] Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 1626–1633. IEEE, 2011. 3, 7

[5] Omri Azencot, Anastasia Dubrovina, and Leonidas Guibas. Consistent shape matching via coupled op-

Model	Wolf	Centaur	Horse	Cat	David
Number of Vertices	4344	15768	19248	27894	52565
SEQA (s)	59	531	801	929	1681
Kernel Matching (s)	60	NA	NA	NA	NA
Consistent Shape Matching (s)	483	NA	2118	3299	NA
Our Method with 1000 spectral basis (s)	36	144	191	474	1774
Our Method with 500 spectral basis (s)	19	106	131	330	1301

Table 1: Average run time for shapes from TOSCA comparing with SEQA [58], Consistent Shape Matching [5] (without including precomputation time for the geodesic distance matrix) and Kernel Matching [55].

timization. In *Computer Graphics Forum*, volume 38, pages 13–25. Wiley Online Library, 2019. 11

[6] Pierre Bérard, Gérard Besson, and Sylvain Gallot. Embedding riemannian manifolds by their heat kernel. *Geometric & Functional Analysis GAFA*, 4(4):373–398, 1994. 3

[7] Alex Bronstein, Michael Bronstein, and Ron Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proceedings of the National Academy of Sciences*, 103(5):1168–1172, 2006. 3

[8] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proceedings of the National Academy of Sciences*, 103(5):1168–1172, 2006. 3, 8

[9] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer Science & Business Media, 2008. 7

[10] Alexander M Bronstein, Michael M Bronstein, Ron Kimmel, Mona Mahmoudi, and Guillermo Sapiro. A gromov-hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching. *International Journal of Computer Vision*, 89(2-3):266–286, 2010. 3

[11] M. M. Bronstein and I. Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1704–1711, 2010. 1, 3

[12] Michael M Bronstein and Iasonas Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1704–1711. IEEE, 2010. 7

[13] Qifeng Chen and Vladlen Koltun. Robust nonrigid registration by convex optimization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2039–2047, 2015. 3

[14] Luca Cosmo, Emanuele Rodolà, Jonathan Masci, Andrea Torsello, and Michael Bronstein. Matching deformable objects in clutter. In *3D Vision (3DV), 2016 Fourth International Conference on*, pages 1–10. IEEE, 2016. 7

[15] Anastasia Dubrovina and Ron Kimmel. Approximately isometric shape correspondence by matching pointwise spectral features and global geodesic structures. *Advances in Adaptive Data Analysis*, 3(01n02):203–228, 2011. 3

[16] Roberto Dyke. *Non-Rigid Shape Correspondence of Physically Based Deformations*, 2020. 7

[17] Nadav Dym, Haggai Maron, and Yaron Lipman. Ds++: A flexible, scalable and provably tight relaxation for matching problems. *arXiv preprint arXiv:1705.06148*, 2017. 3

[18] Marvin Eisenberger, Zorah Lähner, and Daniel Cremers. Divergence-free shape interpolation and correspondence. *arXiv preprint arXiv:1806.10417*, 2018. 8

[19] Andrea Gasparetto, Luca Cosmo, Emanuele Rodola, Michael Bronstein, and Andrea Torsello. Spatial maps: From low rank spectral to sparse spatial functional representations. In *2017 International Conference on 3D Vision (3DV)*, pages 477–485. IEEE, 2017. 3

[20] Xianfeng Gu, Yalin Wang, Tony F Chan, Paul M Thompson, and Shing-Tung Yau. Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE transactions on medical imaging*, 23(8):949–958, 2004. 7

[21] Stefan Gumhold, Xinlong Wang, and Rob S MacLeod. Feature extraction from point clouds. In *IMR*. Citeseer, 2001. 3

[22] Li Han, Shuning Liu, Bing Yu, Shengsi Xu, and Rui Xiang. Orientation-preserving spectral correspon-

dence for 3d shape analysis. *Journal of Imaging Science and Technology*, 2019. 3

[23] Varun Jain and Hao Zhang. Robust 3d shape correspondence in the spectral domain. In *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on*, pages 19–19. IEEE, 2006. 3

[24] Varun Jain, Hao Zhang, and Oliver van Kaick. Nonrigid spectral correspondence of triangle meshes. *International Journal of Shape Modeling*, 13(01):101–124, 2007. 3

[25] Itay Kezurer, Shahar Z Kovalsky, Ronen Basri, and Yaron Lipman. Tight relaxation of quadratic matching. In *Computer Graphics Forum*, volume 34, pages 115–128. Wiley Online Library, 2015. 3

[26] Vladimir G Kim, Yaron Lipman, and Thomas Funkhouser. Blended intrinsic maps. In *ACM Transactions on Graphics (TOG)*, volume 30, page 79. ACM, 2011. 8

[27] Rongjie Lai, Jiang Liang, and Hongkai Zhao. A local mesh method for solving pdes on point clouds. *Inverse Problems & Imaging*, 7(3), 2013. 7

[28] Rongjie Lai, Yonggang Shi, Kevin Scheibel, Scott Fears, Roger Woods, Arthur W Toga, and Tony F Chan. Metric-induced optimal embedding for intrinsic 3d shape analysis. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2871–2878. IEEE, 2010. 1, 3

[29] Rongjie Lai and Hongkai Zhao. Multiscale nonrigid point cloud registration using rotation-invariant sliced-wasserstein distance via laplace–beltrami eigenmap. *SIAM Journal on Imaging Sciences*, 10(2):449–483, 2017. 1, 2, 3, 4, 8

[30] Marius Leordeanu and Martial Hebert. A spectral technique for correspondence problems using pairwise constraints. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 2, pages 1482–1489. IEEE, 2005. 3

[31] B. Levy. Laplace-beltrami eigenfunctions: Towards an algorithm that understands geometry. *IEEE International Conference on Shape Modeling and Applications, invited talk*, 2006. 1, 3

[32] Jun Liu, Xiangqian Liu, and Xiaoli Ma. First-order perturbation analysis of singular vectors in singular value decomposition. *IEEE Transactions on Signal Processing*, 56(7):3044–3049, 2008. 6

[33] Xiuwen Liu, Arturo Donate, Matthew Jemison, and Washington Mio. Kernel functions for robust 3d surface registration with spectral embeddings. In *2008 19th International Conference on Pattern Recognition*, pages 1–4. IEEE, 2008. 3

[34] Diana Mateus, Radu Horaud, David Knossow, Fabio Cuzzolin, and Edmond Boyer. Articulated shape matching using laplacian eigenfunctions and unsupervised point registration. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. 3

[35] Simone Melzi, Jing Ren, Emanuele Rodolà, Abhishek Sharma, Peter Wonka, and Maks Ovsjanikov. Zoomout: Spectral upsampling for efficient shape correspondence. *ACM Transactions on Graphics (TOG)*, 38(6):155, 2019. 2, 4, 7, 8, 14

[36] 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):30:1–30:11, 2012. 1, 3

[37] Maks Ovsjanikov, Quentin Merigot, Facundo Memoli, and Leonidas Guibas. One point isometric matching with the heat kernel. *CGF*, 29(5):1555–1564, 2010. 3

[38] Ulrich Pinkall and Konrad Polthier. Computing Discrete Minimal Surfaces and their Conjugates. *Experimental mathematics*, 2(1):15–36, 1993. 3

[39] Jing Ren, Adrien Poulenard, Peter Wonka, and Maks Ovsjanikov. Continuous and orientation-preserving correspondences via functional maps. *ACM Transactions on Graphics (TOG)*, 37(6), 2018. 3, 4

[40] M. Reuter, F.E. Wolter, and N. Peinecke. Laplace-Beltrami spectra as Shape-DNA of surfaces and solids. *Computer-Aided Design*, 38:342–366, 2006. 1, 3

[41] Emanuele Rodola, Alex M Bronstein, Andrea Albarelli, Filippo Bergamasco, and Andrea Torsello. A game-theoretic approach to deformable shape matching. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 182–189. IEEE, 2012. 3

[42] Emanuele Rodolà, Luca Cosmo, Michael Bronstein, Andrea Torsello, and Daniel Cremers. Partial functional correspondence. *Computer Graphics Forum*, 36(1):222–236, 2017. 8

[43] Emanuele Rodolà, Samuel Rota Bulò, Thomas Windheuser, Matthias Vestner, and Daniel Cremers. Dense non-rigid shape correspondence using random forests. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4177–4184. IEEE, 2014. 8

[44] Raif M Rustamov. Laplace-beltrami eigenfunctions for deformation invariant shape representation. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, pages 225–233. Eurographics Association, 2007. 1

[45] Stefan C Schonsheck, Michael M Bronstein, and Rongjie Lai. Nonisometric surface registration via conformal laplace-beltrami basis pursuit. *arXiv preprint arXiv:1809.07399*, 2018. 3

[46] Alon Shtern and Ron Kimmel. Iterative closest spectral kernel maps. In *2014 2nd International Conference on 3D Vision*, volume 1, pages 499–505. IEEE, 2014. 3, 4

[47] S Skiena. Involutions. *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*, pages 32–33, 1990. 13

[48] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Computer graphics forum*, 28(5):1383–1392, 2009. 3

[49] Art Tevs, Alexander Berner, Michael Wand, Ivo Ihrke, and H-P Seidel. Intrinsic shape matching by planned landmark sampling. In *Computer Graphics Forum*, volume 30, pages 543–552. Wiley Online Library, 2011. 3

[50] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In *European conference on computer vision*, pages 356–369. Springer, 2010. 3, 5, 7

[51] Lloyd N Trefethen and David Bau III. *Numerical linear algebra*, volume 50. Siam, 1997. 6

[52] Pravin M Vaidya. An o (n logn) algorithm for the all-nearest-neighbors problem. *Discrete & Computational Geometry*, 4(2):101–115, 1989. 6

[53] B. Vallet and B. Levy. Spectral geometry processing with manifold harmonics. *Computer Graphics Forum (Proceedings Eurographics)*, 2008. 1, 3

[54] Oliver Van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. A survey on shape correspondence. *Computer Graphics Forum*, 30(6):1681–1707, 2011. 3, 7

[55] Matthias Vestner, Zorah Lähner, Amit Boyarski, Or Litany, Ron Slossberg, Tal Remez, Emanuele Rodolà, Alex Bronstein, Michael Bronstein, and Ron Kimmel. Efficient deformable shape correspondence via kernel matching. In *3D Vision (3DV), 2017 International Conference on*, pages 517–526. IEEE, 2017. 3, 4, 8, 11

[56] Matthias Vestner, Roee Litman, Emanuele Rodolà, Alex Bronstein, and Daniel Cremers. Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space. In *Proc. CVPR*, pages 6681–6690, 2017. 3

[57] Chaohui Wang, Michael M Bronstein, Alexander M Bronstein, and Nikos Paragios. Discrete minimum distortion correspondence problems for non-rigid shape matching. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 580–591. Springer, 2011. 3

[58] Rui Xiang, Rongjie Lai, and Hongkai Zhao. Efficient and robust shape correspondence via sparsity-enforced quadratic assignment. *arXiv preprint arXiv:2003.08680*, 2020. 3, 5, 7, 8, 11

A. The Effect of Correspondence Error on Spectral Alignment

Proof of Theorem 1 We first show that a necessary condition for $\|C_a - C_T\|_2 = 0$ is that permutation σ is an involution, i.e. symmetric matrix.

Since $\Phi_{\mathcal{M}_2} = \Phi_{\mathcal{M}_1} C_T$, we have $Y_i = X_i C, i = 1, 2$. It is clear that $\Phi_{\mathcal{M}_1}^\top \Pi \Phi_{\mathcal{M}_2} = (X_1^\top X_1 + X_2^\top \sigma X_2) C_T$. Assume that there is a SVD decomposition $X_1^\top X_1 + X_2^\top \sigma X_2 = U \Sigma V^\top$. Then we have a SVD decomposition $\Phi_{\mathcal{M}_1}^\top \Pi \Phi_{\mathcal{M}_2} = U \Sigma V^\top C_T$ because C_T is an orthonormal matrix. Therefore, $C_a = \text{Proj}_{\mathfrak{R}}(\Phi_{\mathcal{M}_1}^\top \Pi \Phi_{\mathcal{M}_2}) = U V^\top C_T$. This yields

$$\|C_a - C_T\|_2 = \|U V^\top C_T - C_T\|_2 = \|U V^\top - \text{Id}\|_2$$

Thus, $\|C_a - C_T\|_2 = 0$ implies $U = V$. This leads to σ is symmetric. From the fact that there are $n_2!$ permutation matrix of size $n_2 \times n_2$ and there are $n_2! \sum_{j=0}^{\lfloor n_2/2 \rfloor} \frac{1}{2^j j! (n_2 - 2j)!}$ symmetric permutation matrix of size $n_2 \times n_2$ [47]. This concludes the proof. $\square$

One obvious observation is that $\eta = \sum_{j=0}^{\lfloor n_2/2 \rfloor} \frac{1}{2^j j! (n_2 - 2j)!}$ decreases rapidly as n_2 grows. For example, when $n_2 = 25$, $\eta \approx 10^{-12}$. However, to give a more quantitative characterization of the perturbation for an arbitrary shuffling is difficult since it depends not only on σ but also on X_1 and X_2. Instead, we conduct a few numerical experiments to demonstrate how an inaccurate correspondence Π will affect the correlation matrix $\Phi_{\mathcal{M}_1}^\top \Pi \Phi_{\mathcal{M}_2}$, which is the essential information to align spectral basis, e.g., functional map, between two shapes.

In our experiments, $\Phi_{\mathcal{M}_1}, \Phi_{\mathcal{M}_2} \in \mathbb{R}^{n \times k}$, where n is the total number of points and k is the spectral dimension. We map a human shape with 12,500 vertices to itself, which is a perfect isometric shape correspondence problem with identity map as the ground truth. Theoretically, $0 \le \|C_a - C_T\|_2 \le 2$, since both C_a and C_T are orthonormal. The first experiment tests the behavior of $\|\mathcal{C}_T - \mathcal{C}_a\|_2$ with respect to the ratio of $\frac{k}{n_2}$ for two fixed spectral dimension, $k = 50$ and $k = 100$; second experiments tests the behavior of $\|\mathcal{C}_T - \mathcal{C}_a\|_2$ with respect to n_2 for two fixed $\frac{k}{n_2}$
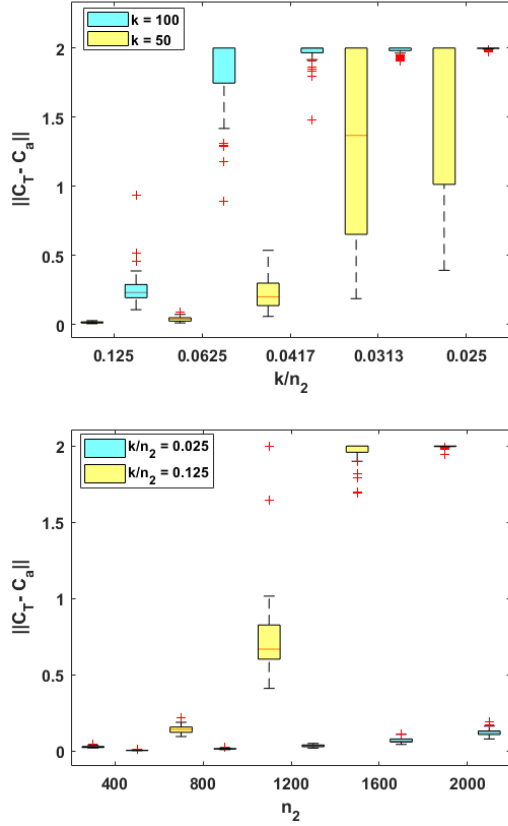
Figure 11: *Left:* The blue centaur is mapped to the gray centaur from TOSCA data set. *Right:* Geodesic error of our method and ZoomOut using different initial guess.

tional map for the first 4 spectral modes from the initial correspondence and then use the code provided by the authors on GitHub to run the experiment. Test results are plotted in Figure 11. These experiment further verify our analysis on possible issues using simple iterative spectral refinement without filtering out wrong correspondences.

Figure 10: *Top:* Error statistics for the relation between functional map error and $\frac{k}{n_2}$. *Bottom:* Error statistics for the relation between functional map error and number of points in X_2.

ratio. A random permutation is imposed on Y_2 to compute C_a, and we independently run 50 trails for each parameter combination. Box-plots are used to illustrate the statistics of our computation in Figure 10. The experiments indicate that using inaccurate correspondences from the current level will most likely introduce error to the functional map for the next level and the error can be quite significant (as big as the worst case, $\|C_a - C_T\|_2 = 2$). This could cause a failure or slow convergence for an iterative refinement strategy.

We further test a real example on one pair of shapes from TOSCA data set using ZoomOut [35], which is a simple iterative refinement method based on functional map in spectral domain. A fixed increment of spectral dimension is pre-specified for each iteration. Due to the use of current correspondence at all points to construct the functional map, the following iterative refinements may not lead to a satisfactory result in the end. In this test, we use 1000 LB eigenfunctions for ZoomOut (and our method) with two different initial setups: (1) correspondence provided by SHOT, or (2) 4 given landmarks. Fo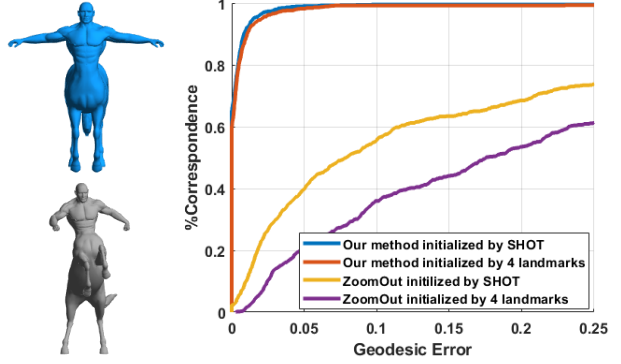r ZoomOut, we first compute a func-