

# An Optimization Approach for Extracting and Encoding Consistent Maps in a Shape Collection

Qi-Xing Huang\*

Guo-Xin Zhang<sup>+</sup>

Lin Gao<sup>+</sup>

Shi-Min Hu<sup>+</sup>

Adrian Butscher\*

Leonidas Guibas\*

\*Stanford University

<sup>+</sup>Tsinghua University

## Abstract

We introduce a novel approach for computing high quality point-to-point maps among a collection of related shapes. The proposed approach takes as input a sparse set of imperfect initial maps between pairs of shapes and builds a compact data structure which implicitly encodes an improved set of maps between all pairs of shapes. These maps align well with point correspondences selected from initial maps; they map neighboring points to neighboring points; and they provide cycle-consistency, so that map compositions along cycles approximate the identity map.

The proposed approach is motivated by the fact that a complete set of maps between all pairs of shapes that admits nearly perfect cycle-consistency are highly redundant and can be represented by compositions of maps through a single base shape. In general, multiple base shapes are needed to adequately cover a diverse collection. Our algorithm sequentially extracts such a small collection of base shapes and creates correspondences from each of these base shapes to all other shapes. These correspondences are found by global optimization on candidate correspondences obtained by diffusing initial maps. These are then used to create a compact graphical data structure from which globally optimal cycle-consistent maps can be extracted using simple graph algorithms.

Experimental results on benchmark datasets show that the proposed approach yields significantly better results than state-of-the-art data-driven shape matching methods.

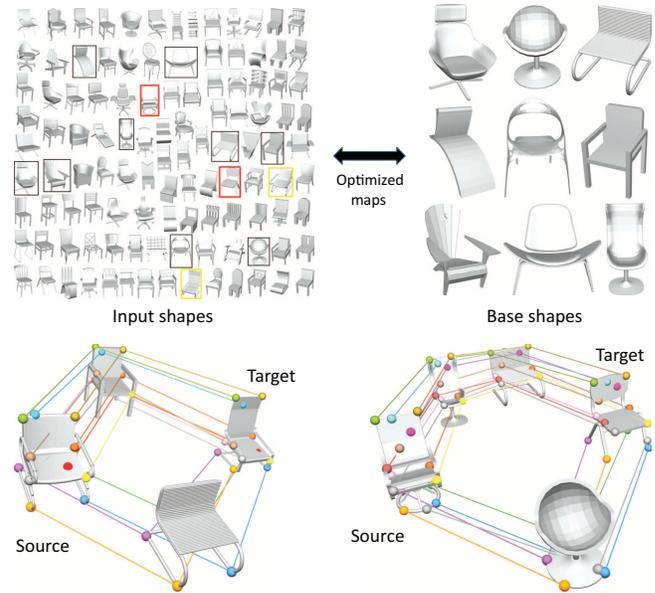
**CR Categories:** I.3.5 [Computing Methodologies]: Computer Graphics—Computational Geometry and Object Modeling;

**Keywords:** Data-driven methods, shape matching, loop closure, geodesic consistency, heat diffusion, quadratic programming

**Links:**  DL  PDF  WEB

## 1 Introduction

With an increasing amount of data describing 3D shapes becoming available, research focus is shifting from processing a single shape to simultaneously processing a collection of shapes, aiming at combining information from multiple sources to improve the processing of each individual shape. In this direction, researchers have studied data-driven shape analysis [Golovinskiy and Funkhouser 2009; Kalogerakis et al. 2010; Huang et al. 2011; Nguyen et al. 2011; Sidi



**Figure 1:** Given a set of input shapes, the proposed approach simultaneously selects a representative subset (the base shapes) and optimizes a set of point-to-point correspondences between each of the base shapes and the entire collection. The final pairwise maps are derived by following shortest paths in the graph specified by these correspondences. (Top) Input shapes and the base shapes selected. (Bottom) Induced maps between example pairs of shapes. For clarity, we only show a subset of correspondences. We see a diversity of paths between the source and target shapes, passing through different intermediate base shapes.

et al. 2011; van Kaick et al. 2011; Kim et al. 2012] and data-driven shape modeling [Chaudhuri et al. 2011; Fisher et al. 2011].

In this paper, we study the shape analysis problem of finding point-to-point maps between shapes belonging to a collection of loosely related shapes. High quality point-to-point maps are crucial to a variety of applications, including information transfer across shapes [Sumner and Popović 2004], shape modeling by assembly [Funkhouser et al. 2004] and detecting the shared structure among a shape collection [James and Twigg 2005]. So far most existing approaches have focused on matching pairs of shapes in isolation. These methods typically find mappings that optimally preserve some invariant property across shapes. One common example is the preservation of geodesic distances amongst near-isometrically deformed shapes, and there are several excellent contributions in this direction [Lipman and Funkhouser 2009; Kim et al. 2011; Ovsjanikov et al. 2012]. These methods work well if the invariant property is at least approximately satisfied but tend to fail when the differences between the input shapes are large. In this case, the search procedure may be trapped in local minima, or the global minimum may be far from the semantic ground truth.

In contrast, considering a collection of shapes together provides additional regularization constraints that help to detect and mitigate these issues. To explain this, we introduce the concept of a

### ACM Reference Format

Huang, Q., Zhang, G., Gao, L., Hu, S., Butscher, A., Guibas, L. 2012. An Optimization Approach for Extracting and Encoding Consistent Maps in a Shape Collection. *ACM Trans. Graph.* 31 6, Article 167 (November 2012), 11 pages. DOI = 10.1145/2366145.2366186 <http://doi.acm.org/10.1145/2366145.2366186>.

### Copyright Notice

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, fax +1 (212) 869-0481, or [permissions@acm.org](mailto:permissions@acm.org).

© 2012 ACM 0730-0301/2012/11-ART167 \$15.00 DOI 10.1145/2366145.2366186 <http://doi.acm.org/10.1145/2366145.2366186>

*model graph* proposed by Huber et al. [Huber 2002], whose vertices represent shapes and whose edges represent maps between pairs of shapes weighted by their “quality.” An important regularizer on such map collections, reflecting semantic consistency in our understanding of the correspondences between the shapes, is *cycle-consistency* — the property that compositions of maps along cycles in the model graph approximate the identity map; or equivalently, that compositions of maps along different paths between two shapes are approximately equal. For instance, it can happen that the direct map between two significantly different shapes, computed via a pair-wise optimization procedure, is further from the ground truth than a composition of such maps through a series of interpolating shapes. Even though the ground truth may be unknown, this failure can still be detected as a break-down in cycle consistency.

In this paper, we develop an optimization framework that takes as input a collection of initial maps between a sparse set of pairs of shapes (so that the model graph is connected), and outputs a new model graph and a new discrete set of correspondences which jointly represent a set of *improved maps* between all pairs of shapes. The improved maps we produce are optimized to be (1) well-aligned with the initial maps, (2) continuous or *neighbor-preserving* (i.e., in the sense that neighboring points are mapped to neighboring points), and (3) cycle-consistent. In other words, our optimization framework exploits global reasoning and regularization constraints enabling us to disentangle the correct correspondence information contained within the initial maps and to produce an improved set.

Directly optimizing the original maps to satisfy the three properties described above is infeasible because the number of constraints involved is very large — the cycle-consistency constraint alone involves an exponential number of loops. Thus the proposed framework favors an indirect solution by combing two key ideas. The first idea is that a complete set of cycle-consistent maps is highly redundant, because each map can be *factorized* as the composition of maps through a single *base shape*. This allows the complete set of maps to be represented as a star tree rooted at the base shape. Therefore we formulate the optimization problem so that the optimization variables are point-to-point maps from the base shape to other shapes and the objective function encourages neighbor-preservation and alignment with the initial maps. We show how to effectively solve this problem in two steps. First we generate multiple candidate correspondences through a diffusion-and-sharpening procedure on the initial maps — diffusion aggregates correspondence information from map compositions through intermediate shapes while sharpening controls the size of the correspondence set. We then extract point-to-point maps by selecting correspondences from these candidates through a joint optimization.

The point-to-point maps from one base shape to all other shapes induce maps between any pair of shapes by composition. However, these maps are not of sufficiently high quality when the input shapes exhibit significant geometric variation. This is because the chosen base shape may not provide sufficient sampling density to “transport” correct correspondences between two non-base shapes without information loss (e.g., in cases where a base shape with a thin part having few samples is used to transport correspondences between shapes having fat corresponding parts with a much larger number of samples; or where the base shape has missing parts).

This brings us to the second key idea: the sequential selection of a modest number of base shapes which jointly cover the given shapes, in the sense that the optimized maps from the base shapes to all other shapes provide a spanning network of maps that contains a sufficiently rich set of correspondences to better approximate all pair-wise maps. In our experiments, we found that a small number of base shapes is sufficient even for shape collections with substan-

tial variation. The selection of the base shapes as “representatives” of the collection is itself an interesting output of our algorithm

Cycle-consistency is an essential ingredient used at this stage to prune spurious correspondences. Only 2-cycles and 3-cycles are considered, but this is sufficient for our purposes. The resulting compact set of weighted correspondences through the base shapes provides a compact *hub-and-spoke correspondence network*<sup>1</sup> from which the final improved maps can be extracted through simple graph algorithms. An important feature of our approach is that the final set of correspondences that represents the improved map between two given shapes may be realized by a collection of correspondence paths that follow different routes in the network in terms of the intermediate shapes visited, depending on the part of a shape being mapped. This path diversity or “mix-and-match” solution naturally exploits partial similarities between the input shapes and is able to use the best parts of each of the initially given maps — without being wholly committed to selecting any one.

We evaluate the performance of the proposed approach on benchmark datasets [Kim et al. 2011; Kim et al. 2012] containing both organic shapes and man-made shapes. Experimental results show that our approach yields significantly better results than state-of-the-art data-driven shape matching approaches [Nguyen et al. 2011; Kim et al. 2012]. We also evaluate different components of our approach and show their advantages over standard alternatives.

In summary, this paper makes the following contributions.

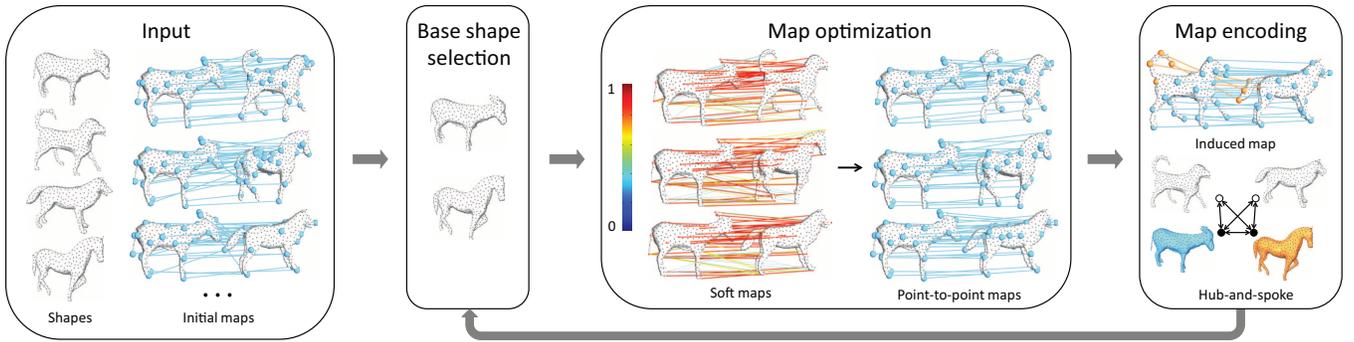
- We present an optimization framework that computes a complete set of new maps which are well-aligned with a sparse-set of initial maps, neighbor-preserving, and cycle-consistent.
- In the process, we introduce a novel correspondence selection algorithm, based on a combined map diffusion-and-sharpening stage, followed by a global combinatorial optimization step.
- We encode our final improved maps through a compact hub-and-spoke network of correspondences, factorizing all pair-wise maps through a small set of sequentially selected base shapes, thereby summarizing the collection.

In the final analysis, we believe that our approach works well because it builds the improved maps indirectly, focusing on first generating and selecting good correspondences. All correspondences originate directly or indirectly from the initially provided maps, but the correspondences that remain in our final hub-and-spoke network have been thoroughly vetted in a wide context of many other correspondences, using both continuity and cycle-consistency criteria involving all the shapes. They are thus well supported by the data (the initial maps) and are known to “play well” with each other. The final maps are computed using simple graph-theoretic criteria from this carefully selected correspondence network, taking advantage of the flexible routing opportunities it provides.

## 1.1 Background

The problem of multiple matching, i.e., finding consistent relations among a collection of objects, is connected with a wide range of scientific problems, including fusing partially overlapped range scans [Huber 2002], re-assembling fractured objects [Huang et al. 2006], solving jigsaw puzzles [Goldberg et al. 2004; Cho et al. 2010], and DNA/RNA sequencing and modeling [Marande and Burger 2007]. It is beyond the scope of this paper to review all the related literature. In the following, we focus on the contributions most related to 3D shapes.

<sup>1</sup> [http://en.wikipedia.org/wiki/Spoke-hub\\_distribution\\_paradigm](http://en.wikipedia.org/wiki/Spoke-hub_distribution_paradigm)



**Figure 2:** The pipeline of our approach is shown. We iteratively select base shapes, from each of which we compute soft maps to other shapes through a sharpening-and-diffusion process. These soft maps are then converted into point-to-point maps through a global optimization step. The optimized point-to-point correspondences from the selected base shapes we have created encode a correspondence graph from which we generate the improved pair-wise point-to-point maps. This procedure is iterated until convergence according to certain criteria.

Most approaches to multiple shape matching take as input an initial model graph derived from matching pairs of shapes, and outputs an optimized model sub-graph whose maps are consistent along loops. Existing approaches fall into two categories. The first category of methods aims at finding a spanning tree in the model graph. In [Goldberg et al. 2004; Huang et al. 2006], the authors propose to use the maximum spanning tree (MST) of the model graph. However, this strategy can easily fail since a single incorrect edge in the MST may break the entire matching result. In the seminal work [Huber 2002], Huber showed that finding the best spanning tree maximizing the number of consistent edges is NP-hard. Although finding the best spanning tree is not tractable, Huber introduced several local operations for improving the score of spanning trees. However, the performance of all these methods relies heavily on the correctness of the edge weights.

The second category of approaches [Zach et al. 2010; Roberts et al. 2011; Nguyen et al. 2011] applies global optimization to select cycle-consistent maps. These approaches are typically formulated as solving constrained optimization problems, where objective functions encode the scores of selected maps, and constraints enforce the consistency of selected maps along cycles. The major advantage of these approaches is that the correct maps are determined globally. However, as the cycle consistency constraint needs to apportion blame along many edges on a cycle, the success of these approaches relies on the assumption that correct maps are dominant in the model graph so that the small number of bad maps can be identified through their participation in many bad cycles.

In contrast, our approach exhibits two advantages. First, we formulate shape matching as selection from a small set of high-quality candidate correspondences. This formulation allows us to combine the strength of both categories of approaches — i.e., the ability to optimize maps globally and the ability to tolerate a large fraction of partially incorrect input maps. Second, existing approaches construct improved maps by composition of initial maps in their entirety. This requires a dominant subset of the initial maps to be fully correct to get good results. In contrast, our method aggregates correspondence information from all initial maps, since the candidate correspondences are created through diffusion-and-sharpening. We thus take advantage of the best parts of each initial map.

The diffusion-and-sharpening operator used in our algorithm follows the line of diffusion-based techniques in geometry processing [Sun et al. 2009; Sidi et al. 2011; Kim et al. 2012], which offer rich tools to relate shapes and to aggregate information from multiple shapes. In particular, Sidi et al [2011] compute diffusion distances between shape parts for clustering. In independent work, Kim et al. [2012] compute fuzzy correspondences based on the

diffusion distance induced from the network of initial maps. In the appendix, we show a mathematical connection between this approach and ours. Both earlier approaches employ spectral analysis to compute diffusion distances. In contrast, our approach is based on exponential maps, involving only sparse matrix multiplication. The new approach is thus more suitable for large-scale data sets.

## 2 Method Overview

### 2.1 Terminology

We assume that shapes are represented as discrete metric spaces [Mémoli and Sapiro 2005], i.e., a shape  $S$  is given by a set of  $M$  samples and a distance matrix  $d_S(\cdot, \cdot)$  that describes all pair-wise distances between samples. For all the examples tested in this paper, we use  $M = 512$  samples and geodesic distances for collections of non-rigid shapes, such as humans or animals; and  $M = 128$  samples and Euclidean distances for more rigid, man-made objects. For shapes given as triangular meshes, we can generate such a representation using the approach proposed in [Lipman and Funkhouser 2009].

Given a source shape  $S$  and a target shape  $S'$ , a correspondence  $c = (s, s') \in S \times S'$  is an oriented link between a pair of points. A point-to-point map  $\phi : S \rightarrow S'$  is equivalent to a set of correspondences of the form  $\{(s, \phi(s)) : s \in S\} \subset S \times S'$ , i.e., each point on the source shape appears in exactly one correspondence. This map is one-to-one if in addition each point from the target shape appears in exactly one correspondence. We encode a set of correspondences from  $S$  to  $S'$  as a sparse matrix  $X \in \mathbb{R}^{M \times M}$  with  $X(s, s') = 1$  if  $(s, s')$  are in correspondence and  $X(s, s') = 0$  otherwise.

We use the notion of *soft map* [Solomon et al. 2012] to describe candidate correspondences. A soft map from  $S$  to  $S'$  associates to each  $s \in S$  a probability distribution over  $S'$  giving the probability that  $s' \in S'$  corresponds to  $s$ . We encode a soft map as a matrix  $C \in \mathbb{R}^{M \times M}$  with  $C(s, s') \in [0, 1]$  and  $\sum_{s' \in S'} C(s, s') = 1$  for each  $s \in S$ . Note that point-to-point maps are a special case of soft maps.

### 2.2 Pipeline

**Input.** Shapes  $S = \{S_i : i = 1, \dots, n\}$  and a (typically sparse) set of initial point-to-point maps along with an associated map weight. We organize this data into a model graph that we denote  $\mathcal{M}_{init}$ . In our evaluations in Section 6, we have used methods adapted from [Kim et al. 2012] to compute the initial model graph.

**Output.** A subset of base shapes  $\mathcal{B} \subset \mathcal{S}$ , a hub-and-spoke correspondence network  $\mathcal{G}$  (in which correspondences between points are routed through base shapes), and a complete set of maps  $\phi_{ij} : S_i \rightarrow S_j$  between all shapes encoded by  $\mathcal{G}$ .

**Components.** We use an iterative strategy involving three separate components to compute this output. These components are:

1. Construction of optimized maps from a single base shape to all other shapes along with their approximate inverses.
2. Construction of a complete set of maps encoded in a hub-and-spoke correspondence network, given a subset of base shapes and maps between each base shape and all other shapes.

In the process, we also derive quality measures that assess the degree to which the resulting complete set of maps is neighbor-preserving and cycle-consistent.

3. Selection of the next base shape.

**Pipeline.** These components combine in the following way. At the beginning of each step of the iteration, we have a current subset of base shapes and maps from these base shapes to all other shapes.

- Add a new base shape and apply the first algorithm component to produce maps from the new base shape to all other shapes.
- Apply the second algorithm component to the augmented subset of base shapes and maps.
- If the quality of these maps improves (as defined in Section 5.3), then proceed to the next iteration.
- Otherwise stop and output the most recent complete set of maps plus the associated correspondence network and base shapes.

### 2.3 Pipeline Components

**Optimized maps from a single base shape.** The input for this component is the model graph  $\mathcal{M}_{init}$  and a choice of a base shape denoted without loss of generality by  $S_1$ . The procedure is divided into two steps: we first compute soft maps  $C_{1i}$  from  $S_1$  to each other shape  $S_i$  along with an associated score  $w_{1i}$  that encodes the global similarity between  $S_1$  and  $S_i$ . We then treat these soft maps as *candidate correspondences* from which we construct point-to-point maps  $X_{1i}$  from  $S_1$  to each other shape  $S_i$  as well as approximate inverse maps  $Y_{i1}$  from  $S_i$  back to  $S_1$  using a global optimization procedure. The output is this set of point-to-point maps.

We use a diffusion-and-sharpening approach to compute the soft maps. The diffusion process aggregates correspondence information contained in the model graph through map compositions. This is motivated by the fact that meaningful correspondences can be realized along many different paths of maps in this graph. The sharpening operator ensures that the resulting soft maps are sparse and neighbor-preserving in the sense that soft maps of neighboring points have similar distributions. The conversion to point-to-point maps is then achieved by optimally selecting the most likely correspondences from the soft maps so that the induced maps between all pairs of shapes are neighbor-preserving and well-aligned with the initial maps. These two steps are presented in Sections 3 and 4.

**Complete set of maps between all shapes.** The input for this component is the current subset of base shapes  $\mathcal{B}$  and the map pairs  $X_{bj}, Y_{jb}$  between each base shape  $S_b \in \mathcal{B}$  and all other  $S_j$ . The output is a correspondence network  $\mathcal{G}$  that is used to encode a complete set of maps between all pairs of shapes, as well as quality measures based on cycle-consistency and neighbor-preservation criteria.

Since there is no guarantee that the full set of maps  $\{X_{bj}, Y_{jb}\}$  is cycle-consistent, we construct the correspondence network  $\mathcal{G}$  by extracting correspondences from a greedily chosen subset of  $\{X_{bj}, Y_{jb}\}$  optimizing the above-mentioned quality measures. The maps encoded by  $\mathcal{G}$  between all pairs of shapes are defined in terms of shortest paths in  $\mathcal{G}$ , which amounts to choosing the appropriate path through base shapes for each correspondence in each map. This step is presented in Section 5.

**Base shape selection.** The first base shape is chosen as the shape having the maximum sum of initial weights to all other shapes:

$$S_{first} = \arg \max_{S_i} \sum_{j \neq i} w_{ji}^{init}.$$

The idea for augmenting the subset of base shapes is to choose the next base shape “farthest” from the existing base shapes; i.e., the new base shape covers shapes that are least similar to existing base shapes. Thus given base shapes  $\mathcal{B} \subset \mathcal{S}$ , the next base shape is

$$S_{next} = \arg \min_{S_i \in \mathcal{S}/\mathcal{B}} \max_{S_b \in \mathcal{B}} w_{bi}$$

where  $w_{ji}$  is the similarity score between shapes  $S_b$  and  $S_i$  computed during the most recent run of the soft map creation step. We stop adding base shapes when the change in the quality measures of the complete set of maps created from  $\mathcal{B}$  is smaller than 0.01 times their absolute value (i.e., the quality measures stabilize).

## 3 Soft Maps by Diffusion-and-Sharpening

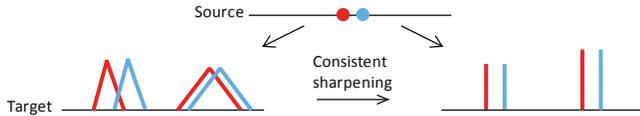
In this section, we introduce an algorithm for computing soft maps from a single base shape to all other shapes. The algorithm takes as input a sparse weighted model graph  $\mathcal{M}_{init}$ , a set of initial maps  $X_{ij}^{init}$  between each pair of shapes  $(S_i, S_j) \in \mathcal{M}_{init}$ , and a choice of base shape (without loss of generality, we let  $S_1$  denote the base shape). The output consists of soft maps  $C_{1i}$  from  $S_1$  to each other shape  $S_i$  and weights  $w_{1i}$  for  $i = 2, \dots, n$ .

Our algorithm is based on the following intuition. Since we can expect the underlying ground truth maps in a shape collection to be cycle-consistent, the point correspondences with high confidence value are those that can be realized by composing maps along many different paths in the model graph — the more such paths, the higher the confidence. Since the analog of composing maps is multiplying soft map matrices (e.g., the composition  $S_i \rightarrow S_j \rightarrow S_k$  corresponds to the matrix product  $\sum_u C_{jk}(t, u)C_{ij}(u, s)$  because this equals the addition of all (independent) probabilities of reaching  $t \in S_k$  from  $s \in S_i$  through  $u \in S_j$ ), a standard diffusion procedure would thus construct a weighted linear combination of products of soft map matrices that represents the contributions of *all* paths in the model graph, with decreasing weights reflecting the fact that long paths carry less reliable information than short paths. However, such a standard formulation is insufficient for our purpose because the diffusion procedure amplifies noise in the soft maps, resulting in non-sparse soft maps that are unsuitable for the point-to-point map optimization procedure that will be carried out in the next section. We therefore introduce a *sharpening operator* that we apply at each diffusion step to counteract this effect.

### 3.1 Algorithm

The diffusion procedure generates a sequence of soft maps  $C_{1i}^{(k)}$  and weights  $w_{1i}^{(k)}$  for  $k = 1, \dots, k_{max}$  representing the soft maps and the global similarity weights derived from paths from  $S_1$  to  $S_i$  of length  $k$ . The choice of  $k_{max}$  will be given below. To begin with, we set

$$w_{1i}^{(1)} = w_{1i}^{init} \quad \text{and} \quad C_{1i}^{(1)} = X_{1i}^{init} \quad \text{for } i = 2, \dots, n.$$



**Figure 3:** The sharpening operator locally perturbs a soft map into a sharpened soft map that favors crisp distributions for each point and similar distributions for neighboring points.

Next, given  $C_{1i}^{(k-1)}$  and  $w_{1i}^{(k-1)}$ , we compute  $C_{1i}^{(k)}$  and  $w_{1i}^{(k)}$  by taking another matrix product with the initial maps:

$$w_{1i}^{(k)} C_{1i}^{(k)} = \sum_{(j,i) \in \mathcal{M}_{init}} (w_{ji}^{init} X_{ji}^{init}) \cdot (w_{1j}^{(k-1)} \mathbb{S}(C_{1j}^{(k-1)})) \quad (1)$$

where  $\mathbb{S} : \mathbb{R}^{M \times M} \rightarrow \mathbb{R}^{M \times M}$  is a column-sum preserving sharpening operator that we will define precisely below. Finally, we set

$$w_{1i}^{output} C_{1i}^{output} = \mathbb{S} \left( \sum_{k=1}^{k_{max}} \alpha_k w_{1i}^{(k)} \mathbb{S}(C_{1i}^{(k)}) \right). \quad (2)$$

Here,  $\alpha_k$  are parameters specifying the importance of aggregating maps along paths of different length. We take  $\alpha_k = t^k/k!$  and have found in our experiments that  $t = 8/G$  where  $G$  is the average valence in  $\mathcal{M}_{init}$  offers a good balance between short and long paths. We also choose  $k_{max}$  as the least  $k$  for which  $\|\alpha_k \mathbb{S}(C_{1i}^{(k)})\|_1 < 0.01$ . Note that equation (1) and (2) define both the soft maps and the weights because matrices with fixed column sums are closed under linear combinations.

### 3.2 Sharpening Operator

We formulate the process of sharpening a soft map  $C$  from shape  $S$  to shape  $T$  as solving a constrained optimization problem. The constraints ensure that correspondences are only aggregated locally during the sharpening process (i.e., we would like to preserve the peaks in the probability distributions represented by the columns of  $C$  while dampening the noise). The objective function favors sparsity (i.e., only a small number of candidate correspondences are retained) and neighbor-preservation (i.e., target points of neighboring source points are also neighbors).

**Constraints.** We parameterize soft map matrices  $\hat{C}$  near  $C$  by adapting the mass-preserving formulation of Earth-Mover's Distance [Rubner et al. 2000]. Let  $\mathcal{A}_T$  be the adjacency graph connecting each point in  $T$  with its  $k = 18$  nearest neighbors. To each edge  $(t, t') \in \mathcal{A}_T$  let  $g_s^{t \rightarrow t'}$  be a positive latent variable giving the contribution of  $t \in T$  to the  $t'$ -entry of the column  $\hat{C}(s, \cdot)$ . That is,

$$\hat{C}(s, t) = \sum_{(t', t) \in \mathcal{A}_T} g_s^{t' \rightarrow t} \quad \forall t \in T. \quad (3)$$

Additionally, we require that the sum of the latent variables of edges emitting from each  $t \in T$  should equal the original entries of the column  $C(s, \cdot)$ . In other words,

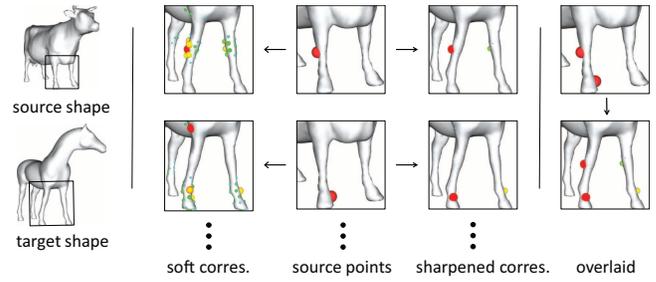
$$C(s, t) = \sum_{(t', t) \in \mathcal{A}_T} g_s^{t' \rightarrow t} \quad \forall t \in T. \quad (4)$$

Note that by construction  $\hat{C}$  has the same column sums as  $C$ . We summarize Equations 3 and 4 in their matrix form as

$$\hat{C}(s, \cdot) = J_{out} \mathbf{g}_s \quad \text{and} \quad J_{in} \mathbf{g}_s = C(s, \cdot) \quad \forall s \in S. \quad (5)$$

where the matrices  $J_{out}$  and  $J_{in}$  encode the connectivity of  $\mathcal{A}_T$  and the vector  $\mathbf{g}_s$  holds the latent variables associated with the point  $s$ .

**Objective function.** Let  $\mathcal{A}_S$  be the adjacency graph connecting each point in  $S$  with its  $k = 6$  nearest neighbors, we define the objective



**Figure 4:** The effect of sharpening on a soft map from a cow model to a horse model. For clarity, we only show the soft maps and sharpened soft maps from two sample points on the source. Note that the sharpened soft maps do not necessarily correspond to local extrema of the original soft maps.

function to be maximized as

$$F(\hat{C}) = \sum_{t, t' \in T} \sum_{(s, s') \in \mathcal{A}_S} \hat{C}(s, t) M(s, s', t, t') \hat{C}(s', t') \quad (6)$$

where  $M(s, s', t, t') = \exp(-(d_S(s, s') - d_T(t, t'))^2 / 2\sigma^2)$  encodes local similarity of geodesic distances and  $\sigma$  estimates the allowed stretching. We set  $\sigma$  equal to 0.05 times the average of the maximum inter-point geodesic distances of the input shapes.

The objective function above achieves both sparsity and neighbor-preservation for the following reasons. First, as the sum of elements of vector  $\hat{C}(s, \cdot)$  is 1 for each point  $s$ , the expected value of the summand in Equation 6 is roughly proportional to  $1/kk'$  where  $k$  and  $k'$  are the number of non-zero elements of  $\hat{C}(s, \cdot)$  and  $\hat{C}(s', \cdot)$ , respectively. Thus, maximizing  $F$  favors a small number of non-zero matrix entries. Second, maximizing  $F$  favors preservation of distances between induced candidate correspondences of neighbor points. We are aware of other possible objective functions — e.g., one could achieve sparsity by optimizing an entropy potential [Chui and Rangarajan 2003].

**Optimization.** Substituting Equation 5 into Equation 6, we have the following linearly constrained quadratic programming problem:

$$(\mathbf{g}_{s_1}^{optimal}, \dots, \mathbf{g}_{s_M}^{optimal}) = \arg \max \sum_{(s, s') \in \mathcal{A}_S} [J_{out} \mathbf{g}_s]^T M_{ss'} J_{out} \mathbf{g}_{s'} \quad (7)$$

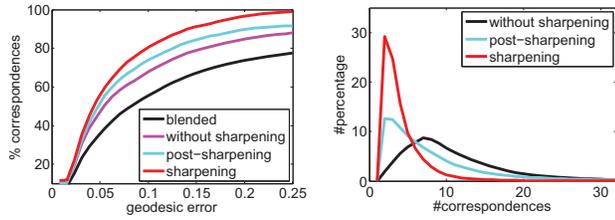
$$\text{s.t.} \quad J_{in} \mathbf{g}_s = C(s, \cdot) \quad \forall s \in S \quad \text{and} \quad \mathbf{g}_s \geq 0 \quad \forall s \in S.$$

The sharpened soft map is then given by

$$\mathbb{S}(C)(s, \cdot) = J_{out} \mathbf{g}_s^{optimal} \quad \forall s \in S.$$

As each variable  $g_s^{t \rightarrow t'}$  appears exactly once in the constraints, this optimization problem can be understood as a quadratic programming relaxation of the standard second-order maximum a posteriori (MAP) problem [Leordeanu and Hebert 2006; Kumar et al. 2009]. In this paper, we use the climbing method introduced in [Leordeanu and Hebert 2006] for optimization due to its memory efficiency (only sparse matrix multiplications are performed).

**Discussion.** We have studied the influence of the sharpening operator on the resulting diffused maps. Figure 4 illustrates the difference between the soft map from a cow model to a horse model and its sharpened soft map. The sharpened soft map contains far fewer correspondences while still preserving the main correspondence clusters. In addition, the sparse correspondences vary in a consistent manner as the source point changes. Figure 5 compares the quality of the best soft correspondences, i.e., the maximum-likelihood correspondence, among (i) diffused maps without sharpening, (ii) sharpened maps after standard diffusion, and (iii) sharpening throughout the diffusion process (i.e., our approach). Exper-



**Figure 5:** Comparison among three different strategies for diffusible maps on the Fourleg data set: sharpening through diffusion; sharpening after diffusion; no sharpening. The former yields the soft maps with sharpest peaks and smallest geodesic error. (Left) Geodesic error of the best soft maps for each point. (Right) The mean number of non-zero entries in the soft map per point. The statistics are averaged over the results obtained by using each input shape as the base shape.

imental results show that our approach yields far better results in quality and sparsity than maps generated through diffusion alone. Moreover, sharpening the soft map throughout the diffusion process prevents large-scale drifting of the correspondences, after which rectification through local modifications alone is impossible.

## 4 Joint Point-to-Point Map Optimization

We now show how to extract maps  $X_{1i} : S_1 \rightarrow S_i$  and their approximate inverses  $Y_{i1} : S_i \rightarrow S_1$  from the soft maps  $C_{1i}$  created above. We perform a joint optimization, formulated as a quadratic integer program, using the entries of  $C_{1i}$  to constrain the number of candidate correspondences from which the correspondences in  $X_{1i}$  and  $Y_{i1}$  are selected. The objective function consists of an alignment term favoring the alignment of the initial maps with the induced maps between all pairs of shapes (i.e., the maps  $X_{1i}$ ,  $Y_{i1}$  and the composite maps  $X_{1j} \cdot Y_{i1} : S_1 \rightarrow S_j$ ), a neighbor-preservation term, and a regularization term favoring  $Y_{i1} \cdot X_{1i}$  approximate the identity map on  $S_1$ .

**Constraints.** To allow only high-probability correspondences, we set a threshold  $\delta \in (0, 1)$  (we take  $\delta = 0.05$ ) and impose

$$X_{1i}(s, s') = 0 \quad \text{whenever } C_{1i}(s, s') \leq \delta. \quad (8)$$

We add additional constraints designed to ensure that the remaining entries of  $X_{1i}$  have a single one in each column, namely

$$\sum_{s'} X_{1i}(s, s') = 1 \quad \forall s \in S_1 \text{ and } \forall i = 2, \dots, n. \quad (9)$$

Similarly, each  $Y_{i1}$  is also constrained by (8) and (9).

**Alignment term.** We score the similarity between the induced maps and the initial maps in  $\mathcal{M}_{init}$  by means of cumulative similarity between their correspondences:

$$\begin{aligned} F^{align}(X_{12}, \dots, X_{1n}, Y_{21}, \dots, Y_{n1}) \\ = \frac{1}{M} \left[ \sum_{(i,1) \in \mathcal{M}_{init}} \overrightarrow{\text{align}}_1(X_{1i}) + \sum_{(i,1) \in \mathcal{M}_{init}} \overleftarrow{\text{align}}_1(Y_{i1}) \right. \\ \left. + \sum_{(i,j) \in \mathcal{M}_{init}} \text{align}_2(X_{1i}, Y_{j1}) \right] \end{aligned} \quad (10)$$

where  $M$  is the number of points for each shape. Here

$$\overrightarrow{\text{align}}_1(X_{1i}) = \sum_{s \in S_1} \sum_{t \in S_i} X_{1i}(s, t) \alpha((s, t), X_{1i}^{init})$$

$$\overleftarrow{\text{align}}_1(Y_{i1}) = \sum_{s \in S_1} \sum_{t \in S_i} Y_{i1}(t, s) \alpha((t, s), X_{1i}^{init})$$

$$\text{align}_2(X_{1i}, Y_{j1}) = \sum_{s \in S_1} \sum_{t \in S_i} \sum_{t' \in S_j} X_{1i}(s, t) Y_{j1}(t', s) \alpha((t, t'), X_{ij}^{init})$$

and  $\alpha((s, t), X)$  scores the alignment between a single correspondence  $(s, t) \in S_i \times S_j$  and an initial map  $X$  from  $S_i$  to  $S_j$  via

$$\alpha((s, t), X) = \max_{t': d_{S_j}(t, t') \leq 2\sigma} w X(s, t').$$

Here  $w$  is the weight associated to  $X$  and  $\sigma$  is 0.05 times the average of the maximum inter-point geodesic distance of each input shape. In other words,  $\alpha((s, t), X) = w$  if the distance between  $t$  and the image point of  $s$  under  $X$  is less than  $2\sigma$ .

**Neighbor-preservation term.** We score neighbor-preservation only for the maps  $X_{1i}$  since optimizing this score implicitly optimizes the neighbor-preservation property of all other induced maps. We count pairs of neighboring points in  $S_1$  mapped to neighboring points in  $S_i$ . Formally, let  $\mathcal{A}_1$  be the adjacency graph connecting each point in  $S_1$  with its  $k = 6$  nearest neighbors and define

$$\begin{aligned} F^{nb}(X_{12}, \dots, X_{1n}) \\ = \frac{1}{E_1} \sum_{i=2}^n \sum_{(s, s') \in \mathcal{A}_1} \sum_{t, t' \in S_i} X_{1i}(s, t) W_i^{nb}(s, s', t, t') X_{1i}(s', t') \end{aligned} \quad (11)$$

where  $E_1$  denotes the number of edges of  $\mathcal{A}_1$  and

$$W_i^{nb}(s, s', t, t') = \begin{cases} 1 & |d_{S_1}(s, s') - d_{S_i}(t, t')| \leq 2\sigma \\ 0 & \text{otherwise.} \end{cases}$$

**Regularization term.** The regularization term promotes the maps  $X_{1i}$  and  $Y_{i1}$  to be inverses of each other. We achieve this with

$$F^{reg}(X_{12}, \dots, X_{1n}, Y_{21}, \dots, Y_{n1}) = \sum_{i=2}^n \text{align}_2(Y_{i1}, X_{1i}). \quad (12)$$

We use the function  $\text{align}_2$  to score the alignment of the composition  $X_{1i} \cdot Y_{i1}$  with the identity mapping of  $S_1$ , i.e., by setting  $X_{ij}^{init} = \text{identity}$  in the definition of  $\text{align}_2$  given above.

**Optimization.** We find the desired point-to-point maps by solving

$$\begin{aligned} (X_{12}, \dots, X_{1n}, Y_{21}, \dots, Y_{n1}) \\ = \arg \max \left( F^{align}(X_{12}, \dots, X_{1n}, Y_{21}, \dots, Y_{n1}) \right. \\ \left. + \lambda \cdot F^{nb}(X_{12}, \dots, X_{1n}) \right. \\ \left. + \mu \cdot F^{reg}(X_{12}, \dots, X_{1n}, Y_{21}, \dots, Y_{n1}) \right) \end{aligned} \quad (13)$$

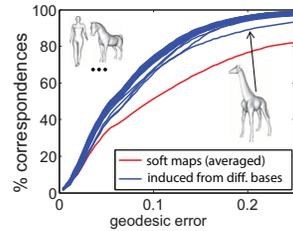
subject to the constraints

$$\begin{aligned} X_{1i}, Y_{i1} \in [0, 1]^{M \times M} \quad \forall i = 2, \dots, n \\ \sum_{s'} X_{1i}(s, s') = 1 \quad \forall s \in S_1, \quad \sum_s Y_{i1}(s, s') = 1 \quad \forall s' \in S_i \\ X_{1i}(s, s') = Y_{i1}(s', s) = 0 \quad \text{whenever } C_{1i}(s, s') \leq \delta. \end{aligned}$$

The parameters  $\lambda$  and  $\mu$  control the importance of the neighbor-preservation and the regularization terms respectively. We have chosen  $\lambda = 10$  and  $\mu = 100$  for all examples tested herein.

The variables in Equation 13 are binary, the objective function is quadratic, and the second constraint can be understood as picking a label between each sample  $s$  and each shape  $S_i$ . Thus the problem (13) is again a second-order MAP problem [Kumar et al. 2009] that we solve using the technique of [Leordeanu and Hebert 2006].

**Discussion.** We have evaluated the performance of the resulting point-to-point maps on the Fourlimb data set consisting of 18 humans and 20 animals. As shown in Figure 6, the resulting maps are much better than those obtained from the most likely correspondences implied by the soft maps. This indicates the need for optimizing maps simultaneously. Moreover, using different base shapes yields roughly the same result, indicating the approximate invariance of our objective function to different base shapes.



**Figure 6:** The quality of all pair-wise induced maps is similar across different shapes.

## 5 Creating a Complete Set of Maps

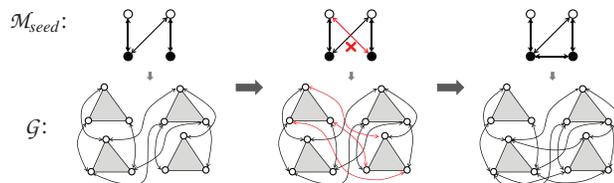
In this section, we show how to create a complete set of maps between all pairs of shapes given a subset of base shapes  $\mathcal{B} \subset \mathcal{S}$  along with the map pairs  $X_{bj}, Y_{jb}$  between  $S_b \in \mathcal{B}$  and all other  $S_j$  for  $j \in \{1, \dots, n\} \setminus \{b\}$ . Let  $\mathcal{M}_B$  be the model graph that represents this data. Our strategy is to construct the complete set of maps in a holistic way in terms of paths of individual points through a hub-and-spoke correspondence network  $\mathcal{G}$  constructed from a greedily chosen subset  $\mathcal{M}_{seed} \subseteq \mathcal{M}_B$  (See Figure 7). We make the choice of  $\mathcal{M}_{seed}$  in such a way to optimize neighbor-preservation and cycle-consistency of the induced complete set of maps, as there is no guarantee that the maps in  $\mathcal{M}_B$  are cycle-consistent. Unlike standard approaches that use composition to define maps between shapes, the idea of using the hub-and-spoke correspondence network allows us to utilize the best parts of the supplied maps.

Below we show how to initialize  $\mathcal{M}_{seed}$  and how to augment it by adding pairs of maps  $X_{bj}, Y_{jb}$ . Then we show how to construct the hub-and-spoke correspondence network  $\mathcal{G}$  given an instance of  $\mathcal{M}_{seed}$  and how to extract maps between all pairs of shapes from  $\mathcal{G}$ . Finally, we introduce the neighbor-preservation and cycle-consistency quality measures used to determine whether a pair of maps should be kept in  $\mathcal{M}_{seed}$  during the augmentation process.

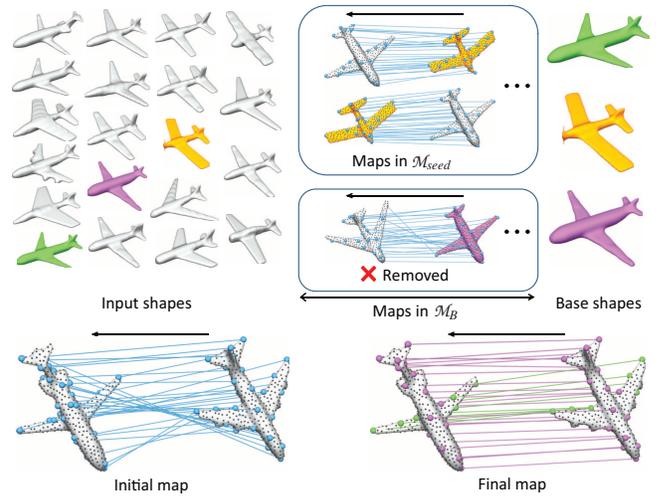
### 5.1 Initialization

We initialize  $\mathcal{M}_{seed}$  by selecting a minimal spanning tree for  $\mathcal{M}_B$  w.r.t. the weights  $1/\text{align}_2(X_{bj}, Y_{jb})$  for each  $X_{bj}, Y_{jb}$  pair. Intuitively, these weights favor map pairs for which  $Y_{jb}$  is close to the inverse of  $X_{bj}$  and imply a high similarity between  $S_b$  and  $S_j$ .

Furthermore, we augment  $\mathcal{M}_{seed}$  during the procedure of the next section by adding map pairs in order of increasing weight.



**Figure 7:** Schematic illustration of the process of creating a complete set of maps.  $\mathcal{M}_{seed}$  is initialized as a minimum spanning tree of  $\mathcal{M}_B$ . We then add remaining maps in  $\mathcal{M}_B$  into  $\mathcal{M}_{seed}$  by monitoring the quality measures of complete set of maps derived from the associated hub-and-spoke correspondence network  $\mathcal{G}$ .



**Figure 8:** Creating a complete set of maps for the Plane dataset. (Top)  $\mathcal{M}_{seed}$  that defines the hub-and-spoke correspondence network. (Bottom) The initial and final point-to-point map between two airplanes. The coloring in the final map represents the intermediate base shape used for defining correspondences.

### 5.2 Constructing the Complete Set of Maps

We first build the hub-and-spoke correspondence network induced by the current  $\mathcal{M}_{seed}$ . This is a graph  $\mathcal{G}$  over all sample points in all shapes  $S_1, \dots, S_n$  using the correspondence information contained in all the  $X, Y$  maps encoded in  $\mathcal{M}_{seed}$ . That is, we let the vertices of  $\mathcal{G}$  be all points  $s \in S_i$  for all  $i \in \{1, \dots, n\}$ . We declare  $(s, t) \in S_i \times S_j$  to be a directed edge in  $\mathcal{G}$  if the correspondence  $(s, t)$  appears in any of these  $X, Y$  maps. In addition, we assign a weight  $\omega(s, t)$  to this edge based on the amount of local geodesic distance distortion caused by these maps. That is,

$$\omega(s, t) = \sum_{s', s'' \in \mathcal{N}(s)} |d_{S_i}(s', s'') - d_{S_j}(t', t'')|, \quad (14)$$

where  $t', t'' \in S_j$  correspond to  $s', s''$  under the appropriate  $X$  or  $Y$  map, and  $\mathcal{N}(s)$  is the set of  $k = 32$  nearest neighbors of  $s$ .

We now use this graph to construct maps between all shapes. Let  $\text{dist}_{\mathcal{G}}$  denote the graph metric of  $\mathcal{G}$  induced by the edge weights defined in (14). For any  $i, j \in \{1, \dots, n\}$  we define  $\phi_{ij} : S_i \rightarrow S_j$  as

$$\phi_{ij}(s) = \arg \min_{t \in S_j} \text{dist}_{\mathcal{G}}(s, t). \quad (15)$$

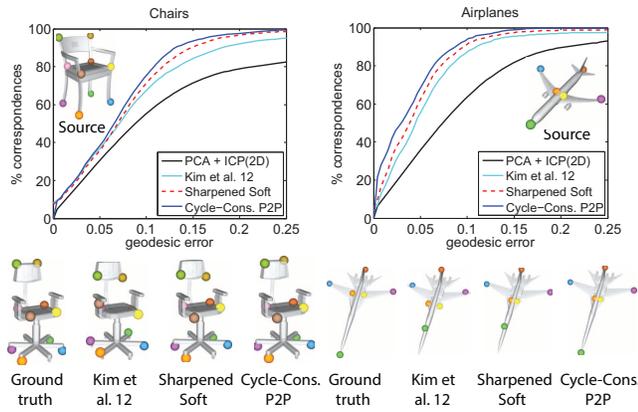
### 5.3 Accepting New Maps Based On Quality Measures

Let  $\Phi = \{\phi_{ij} : i, j = 1, \dots, n\}$  be the complete set of maps between all pairs of shapes created above. It remains to test the quality of  $\Phi$  in order to decide whether or not to accept the additional  $X_{bj}, Y_{jb}$  pair under consideration. We use two quality measures  $\text{qual}_{\text{cycle}}$  and  $\text{qual}_{\text{nb}}$  that evaluate the extent to which these maps fail to be cycle-consistent and neighbor-preserving, respectively. The criterion we use for accepting a new map pair is: if  $\text{qual}_{\text{cycle}}(\Phi) \leq 3$  and  $\text{qual}_{\text{nb}}(\Phi)$  decreases from the previous iteration, then we accept it.

The quality measures we use are defined as follows.

**Cycle-consistency.** Let  $C_i$  denote the union of the 2-cycles  $(\phi_{ij}, \phi_{ji})$  and the 3-cycles  $(\phi_{ij}, \phi_{jk}, \phi_{ki})$  starting at  $S_i$ . Define

$$\text{qual}_{\text{cycle}}(\Phi) = \max_i \max_{c \in C_i} \frac{1}{M} \sum_{s \in S_i} \frac{|d_{S_i}(s, s')|}{\sigma_i}, \quad (16)$$



**Figure 9:** Comparison between various methods on the Chair and Airplane datasets provided in [Kim et al. 2012].

where  $s'$  is the image point of  $s$  under the composition of the maps in the cycle  $c$  (i.e.,  $s' = \phi_{ji} \circ \phi_{ij}(s)$  in the case of a 2-cycle and  $s' = \phi_{ki} \circ \phi_{jk} \circ \phi_{ij}(s)$  in the case of a 3-cycle).

**Neighbor-preservation.** Let  $\mathcal{A}_i$  be the adjacency graph connecting each point in  $S_i$  with its  $k = 6$  nearest neighbors and let  $\sigma_i$  be 0.05 times the average inter-point geodesic distance of  $S_i$ . Define

$$\text{qual}_{nb}(\Phi) = \max_{ij} \sum_{(s,s') \in \mathcal{A}_i} \frac{|d_{S_i}(s, s') - d_{S_j}(\phi_{ij}(s), \phi_{ij}(s'))|}{|\mathcal{A}_i| \sigma_i}. \quad (17)$$

## 5.4 Discussion

The maps  $\phi_{ij}$  that we have created with the procedure explained above are constructed by assembling individual correspondences taken from compositions of  $X$  and  $Y$  maps along different paths. Therefore, a natural question to ask is whether neighbor-preservation and cycle consistency still hold for the output maps.

First and foremost, our experimental results show that these qualities are indeed exhibited by the output maps. Although we do not supply theoretical justification for this behavior, a strong argument can be given as follows. First, if it is the case that the maps in  $\mathcal{M}_{seed}$  are cycle-consistent, then the composite map along any paths between a fixed pair of shapes commute. Thus, these two criteria are satisfied in this case (the composite maps inherit the neighbor-preserving property from the maps in  $\mathcal{M}_{seed}$ ). Since we have greedily chosen  $\mathcal{M}_{seed}$  to satisfy cycle-consistency and neighbor-preservation criteria, we thus expect the output maps to approximately satisfy these criteria in the general case as well.

## 6 Results

### 6.1 Experimental Setup

We have evaluated the presented approach on the benchmark datasets SCAPE [Anguelov et al. 2004], SHREC07 [Giorgi et al. 2007; Kim et al. 2011] and the datasets from [Kim et al. 2012]. The SCAPE dataset consists of 71 models. Each model is meshed with the same topology, providing ground-truth correspondences for evaluation. The SHREC07 benchmark contains 400 models in 20 categories of 20 models each provided with manually aligned features across shapes for evaluation. As in [Kim et al. 2011], we select 11 datasets that are most suitable for computing shape-wise maps. We further divide them into a non-symmetric subset

	$n$	Nb-Pres.		Cycle-Cons.		$ \mathcal{B} $	Tim-ing(s)
		Mean	Max	Mean	Max		
Scape.	71	0.25	2.3	0.45	2.3	3	255.4
Armadil.	20	0.27	2.7	0.65	3.3	2	55.4
Fish	20	0.42	3.5	0.73	3.9	3	287.1
Fourleg	20	0.39	3.5	0.81	3.4	3	71.4
Hand	20	0.35	3.2	0.78	3.3	3	115.1
Human	18	0.28	3.5	0.46	3.4	2	34.2
Fourlimb	38	0.39	3.7	0.87	3.9	4	321.2
Ant	20	0.25	3.2	0.78	3.4	2	282.6
Bird	20	0.34	3.4	0.76	3.5	4	179.1
Glasses	20	0.41	3.5	0.72	3.1	2	172.4
Plane	20	0.37	3.8	0.78	3.3	3	210.6
Plier	20	0.37	2.6	0.62	3.1	2	101.2
Teddy	20	0.34	3.1	0.67	2.8	1	107.6
Airplane	86	0.39	2.2	1.02	2.7	7	981.2
Chair	111	0.36	2.5	1.16	3.2	9	1423

**Table 1:** Statistics on the shape collections provided in SHREC07 [Giorgi et al. 2007] and in [Kim et al. 2012]. From left to right, we show the number of input shapes, the mean and the max of neighbor-preservation distortion, the mean and the max of cycle-consistency distortion, the number of base shapes being used, and the running time.

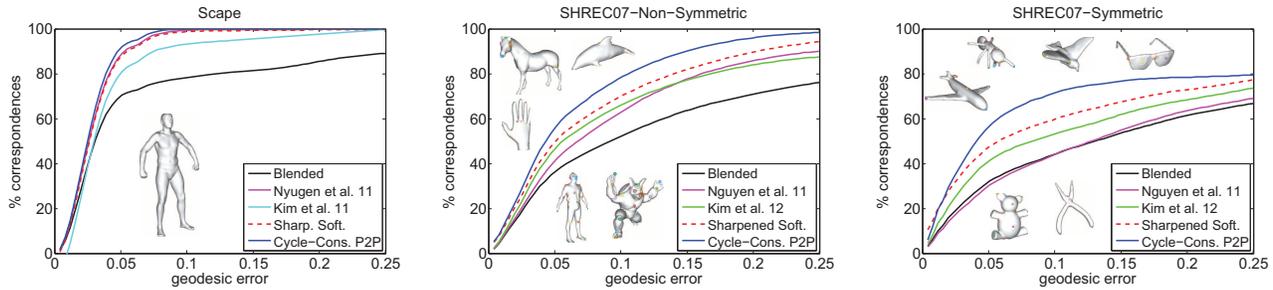
(Armadillo, Fish, Fourleg, Hand, Human), where models do not admit salient rotational intrinsic symmetries, and a symmetric subset (Ant, Bird, Glasses, Plane, Plier, Teddy) containing the remaining models. For the Man-Made datasets from [Kim et al. 2012], we select Chair and Airplane which include manually labeled features.

**Initial maps.** For all the categories provided in the SHREC07 and the SCAPE dataset, we compute blended intrinsic maps [Kim et al. 2011] between all pairs of shapes. For the Chair and the Airplane datasets, we use the same strategy described in [Kim et al. 2012] that performs affine registration between a sparse set of shape pairs that are close to each other in feature space. In order to make a fair comparison with [Kim et al. 2012], we use the same strategy described in [Kim et al. 2012] for weighting each initial map and its correspondences. The weighting scheme essentially considers neighbor-preservation at both the correspondence level and the shape level. Please refer to [Kim et al. 2012] for details.

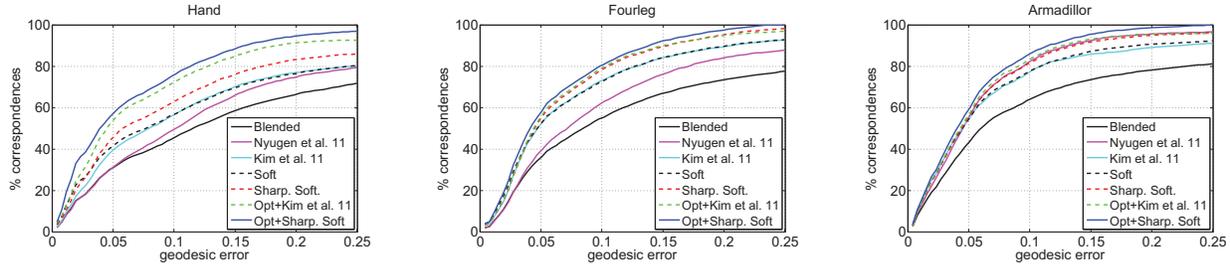
**Evaluation protocol.** We evaluate the quality of the complete set of maps produced by our algorithm in three ways. (1) We evaluate the feature alignment error of the maps when manually labeled features are available. Note that for the SCAPE dataset, we take corresponding vertices on different models. Between each pair of shapes  $S_1, S_2$  and for each pair of corresponding feature points  $f_1, f_2$ , we evaluate the geodesic distance distortion  $d_{S_j}(f_2, X_{ij}(f_1))$ . To visualize the error, we plot the percentage of feature pairs with  $d_{S_j}(f_2, X_{ij}(f_1)) < \epsilon$  against  $\epsilon$  (Figures 10, 11). If  $f_1$  does not lie on a sample point, we snap  $f_1$  into its closest sample. (2) We evaluate cycle-consistency distortion by collecting the mean and maximum of the *per-cycle* quality measure  $\text{qual}_{cycle}(s, c)$  (defined by removing both maxima and the average from (16)). Note that we weigh 2-cycles and 3-cycles properly to equalize their contributions. (3) We evaluate neighbor-preservation distortion by collecting the mean and maximum of the *per-map* quality measure  $\text{qual}_{nb}(X_{ij}, s, s')$  (defined by removing the maximum and the average from (17)) and plot the percentage of maps with  $\text{qual}_{nb}(X_{ij}, s, s') < \epsilon$  against  $\epsilon$  (Figure 12).

### 6.2 Analysis of Matching Results

Table 1 and Figures 9-10 provide the main results of the presented approach. All results were obtained with the same set of parameters described throughout the paper. Overall, both the sharpened soft



**Figure 10:** Performance of various methods on the SCAPE dataset [Anguelov et al. 2004] and on the SHREC07 benchmark [Giorgi et al. 2007]. The sharpened soft maps outperform existing approaches. The cycle-consistent point-to-point maps further improve over the sharpened soft maps. Most of our errors are caused by global intrinsic symmetries existing in the datasets.



**Figure 11:** Comparison between the output of various methods by taking blended intrinsic maps as input. In the plots, (Opt+Fuzzy) stands for using local extrema of fuzzy correspondences [Kim et al. 2011] as input to the global optimization framework, and (Opt+Sharpened Soft) stands for using sharpened soft maps as input to the same framework.

maps and the cycle-consistent point-to-point maps produce results that are much better than the initial maps.

**Feature alignment.** On non-symmetric datasets, the alignment errors between nearly all pairs of features are less than  $0.25d_{max}$  where  $d_{max}$  is the averaged maximum distance between all pairs of points in every shape. In other words, our cycle-consistent point-to-point correspondences can recover the underlying ground truth maps in the global sense, a favorable result. For symmetric shapes, the performance of the presented approach drops. This is expected because both cycle-consistency and neighbor-preservation are preserved if we compose all pairwise maps by the self-symmetries associated with each shape. However, the feature alignment error of our approach is still significantly lower than the initial maps.

**Neighbor-preservation.** Our approach generates maps with low geometric distortion. As shown in Table 1, the averaged neighbor-preservation distortion  $qual_{nb}(X_{ij}, s, s')$  on all datasets tested ranges from 0.25 to 0.41. Since the correspondences between neighboring points may follow different paths, the maximum neighbor-preservation distortion is larger than the sampling density. However, it is still bounded above by 3-4 times the sampling density.

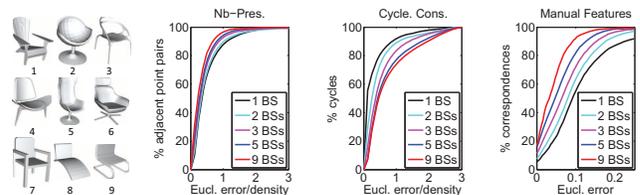
We also studied the behavior of neighbor-preservation when using different numbers of base shapes. There are two potential effects when increasing this number. On one hand, adding more base shapes may provide better base shape selection in building the map between a pair of shapes. On the other hand, a pair of neighboring points might follow different paths in the expanded graph, and thus may be less neighbor-preserving. In practice, we find that the former effect is more dominant than the latter (see Figure 12).

**Cycle-consistency.** The averaged distortion along cycles ranges from 0.45 to 1.16 (the implied geodesic distance distortion is that number times the sample spacing). The cycle distortion increases when variation in the input shapes becomes large, e.g., on Fourlimb, Airplane and Chair datasets. Cycle-consistency distortion increases slightly when using more base shapes (see Figure 12). This is

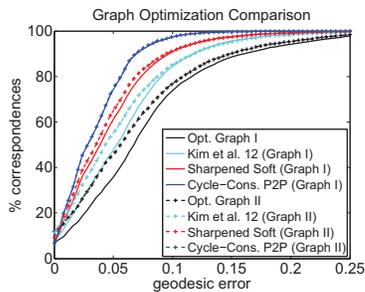
expected since multiple base shapes allow correspondences in the pairwise maps to follow different paths through the correspondence network. However, we find that this effect quickly tapers off.

**Few base shapes are sufficient.** In general, we find that our approach converges in a few iterations. In the ideal case where all input shapes are highly similar, one base shape is enough. For datasets with moderate shape variation, our approach converges in up to four iterations. The number of base shapes increases as the shape variation becomes bigger (7 base shapes for the Airplane dataset and 9 base shapes for the Chair dataset). But still the number of base shapes is much smaller than the number of input shapes.

**Timing.** Our approach runs in a few minutes to dozens of minutes on shape collections with approximately 100 models. Please refer to Table 1 for details. The majority of time is spent on solving the two optimization problems, i.e., sharpening soft maps and converting soft maps into point-to-point maps. The complexity of sharpening one soft map is given by  $O(Mn_c^2 i_{max})$  where  $n_c$  is the averaged number of correspondences with non-zero entries and  $i_{max}$  is the number of climbing methods [Leordeanu and Hebert 2006]. In our experiments, sharpening one soft map takes  $\sim 0.01s$  to  $0.2s$ . The complexity of converting to point-to-point maps is  $O(n \cdot n_c^2 i_{max})$ . This procedure took 100.2 seconds for the Chair dataset.



**Figure 12:** Statistics on the Chair dataset. (a) The base shapes, (b-d) plots respectively of neighbor-preservation, cycle-consistency, feature alignment error when using different numbers of shapes.



**Figure 13:** Comparison between optimized graph I derived from fuzzy correspondences, and optimized graph II derived from cycle-consistent point-to-point maps.

**Limitations.** A limitation of our approach is that the global optimization strategy, which assigns each point on the base shape with exactly one correspondence on each other shape, favors full similarity between shapes. In future, we will consider how to modify the formulation so that it allows partial similarity between shapes.

### 6.3 Comparisons

**Comparison with [Kim et al. 2012].** First, we compare the best fuzzy correspondences [Kim et al. 2012] with best correspondences provide by (sharpened) soft maps computed by our approaches. It is clear that the sharpened soft maps are much better than results from fuzzy correspondences and unsharpened fuzzy correspondences (see Figures 10, 11). This shows that enforcing neighbor-preservation can significantly improve the resulting correspondences and maps. In addition, the un-sharpened soft maps yield similar results as fuzzy correspondences.

Second, we compare the resulting cycle-consistent point-to-point maps with the maps obtained by using local extrema of fuzzy correspondences as input, as well as with the maps obtained by using sharpened soft maps as input. As shown in Figure 11, our global optimization framework improves both methods. However, the one with sharpened soft maps as input yields slightly better results. This is because the fuzzy correspondences are less consistent between neighboring points than the sharpened soft correspondences, decreasing the effectiveness of the neighbor-preservation term.

When evaluating the presented approach on datasets with different quality of initial maps, we can see that the cycle-consistent point-to-point maps are less sensitive to the initial maps than (sharpened) soft maps and best fuzzy correspondences [Kim et al. 2012]. This indicates the stability of our global optimization framework against noise introduced in the initial maps (see Figure 11).

We also compare the effects of feeding the cycle-consistent point-to-point maps into the graph optimization strategy described in [Kim et al. 2012] with that of using the original strategy (see Figure 13). Given the current maps between pairs of shapes, the graph optimization first deletes initial maps that are inconsistent with current maps and then re-computes initial maps between similar shapes suggested by current maps. As the cycle-consistent point-to-point maps provide better alignments of shapes, the soft maps and fuzzy correspondences computed on the optimized graph derived from cycle-consistent maps are better than those computed on the optimized graph generated from fuzzy correspondences. However, the final maps on both optimized graphs are similar, again indicating the stability of our approach.

**Comparison with [Nguyen et al. 2011].** Since the approach presented in [Nguyen et al. 2011] only works for dense graphs, we only present the comparison on SHREC07. As shown in Figures 10, 11,

our approach yields much better results, particularly on datasets where initial maps have moderate or low quality. This is because their approach is based on composing full maps and assumes full maps are correct everywhere when composing maps. In contrast, our approach breaks maps into correspondences which are then re-assembled into new maps through global optimization. We can thus tolerate a significant amount of noise in the initial maps.

## 7 Conclusion

In this paper, we have presented an optimization approach that generates a compact representation of a set of pair-wise maps between all shapes in a collection so that several global consistency criteria are met. The collection of maps is represented in an implicit format, through the extraction of a small set of base shapes that best represents the collection and well-chosen correspondences between each base shape and all the other shapes. This hub-and-spoke network of correspondences reflects our joint understanding of the shape collection and enables us to robustly transport correspondences across the collection. The final map for each pair of shapes  $S$  and  $T$  is obtained by tracing optimal paths through this network, allowing for different intermediate shapes on the paths that transport points of  $S$  to points of  $T$ . This provides the flexibility to benefit from good parts of various maps, without being forced to use a particular map in its entirety or not at all. Experimental results on benchmark datasets show that exploiting the context provided by the shape collection by means of correspondence regularization via consistency constraints enables significant improvements in map quality when compared to extant methods.

The presented work suggests many avenues for future research. For example, it would be desirable to have more formal conclusions about the quality of the final maps (cycle closure, etc), under some assumptions on the quality/number of input maps. It would also be nice to better understand the stability of the computed maps under various perturbations, both to the initial maps and to the set of shapes being given. Ideally, the process of computing the initial maps (and deriving the model graph we start from) should be integrated into our approach to produce a unified pipeline for computing globally consistent maps in a shape collection. Both the selection of the base shapes as the collection representatives, and the combinatorial path classes by which our network routes correspondences among shapes suggest further investigations into ways to summarize shape collections and to recover consistent parts among shapes in the collection. Finally, we also hope that each of our individual technical contributions can be useful in other contexts — for example the diffusion/sharpening approach may also be applicable in processing other kinds of signals, such as images.

**Acknowledgements.** The authors would like to acknowledge the support of NSF grants FODAVA 808515 and CCF 1011228, ONR MURI N0001470710747, the Max Planck Center for Visual Computing and Communications, the KAUST Academic Excellence Alliance, and a Google Research Award. Prof. Shi-Min Hu was supported by the National Basic Research Project 2011CB30220, the National Science Foundation Project 61120106007 and the National High Technology Research and Development Program Project 2012AA011802.

## References

- ANGUELOV, D., SRINIVASAN, P., PANG, H.-C., KOLLER, D., THRUN, S., AND DAVIS, J. 2004. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. *NIPS 17*, 33–40.

- CHAUDHURI, S., KALOGERAKIS, E., GUIBAS, L., AND KOLTUN, V. 2011. Probabilistic reasoning for assembly-based 3d modeling. *ACM Trans. Graph.* 30, 4 (Aug.), 35:1–35:10.
- CHO, T. S., AVIDAN, S., AND FREEMAN, W. T. 2010. A probabilistic image jigsaw puzzle solver. In *CVPR*, 183–190.
- CHUI, H., AND RANGARAJAN, A. 2003. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst.* 89, 2-3 (Feb.), 114–141.
- FISHER, M., SAVVA, M., AND HANRAHAN, P. 2011. Characterizing structural relationships in scenes using graph kernels. *ACM Trans. Graph.* 30, 4 (Aug.), 34:1–34:12.
- FUNKHOUSER, T., KAZHDAN, M., SHILANE, P., MIN, P., KIEFER, W., TAL, A., RUSINKIEWICZ, S., AND DOBKIN, D. 2004. Modeling by example. *ACM Trans. Graph.* 23, 3 (Aug.), 652–663.
- GIORGI, D., BIASOTTI, S., AND PARABOSCHI, L., 2007. Shape retrieval contest 2007: Watertight models track.
- GOLDBERG, D., MALON, C., AND BERN, M. 2004. A global approach to automatic solution of jigsaw puzzles. *Comput. Geom. Theory Appl.* 28 (June), 165–174.
- GOLOVINSKIY, A., AND FUNKHOUSER, T. A. 2009. Consistent segmentation of 3d models. *Computers & Graphics* 33, 3, 262–269.
- HUANG, Q.-X., FLÖRY, S., GELFAND, N., HOFER, M., AND POTTMANN, H. 2006. Reassembling fractured objects by geometric matching. *ACM Trans. Graph.* 25, 3, 569–578.
- HUANG, Q., KOLTUN, V., AND GUIBAS, L. 2011. Joint shape segmentation using linear programming. *ACM Trans. Graph.* 30, 6 (Dec.), 125:1–125:12.
- HUBER, D. 2002. *Automatic Three-dimensional Modeling from Reality*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.
- JAMES, D. L., AND TWIGG, C. D. 2005. Skinning mesh animations. *ACM Trans. Graph.* 24, 3 (July), 399–407.
- KALOGERAKIS, E., HERTZMANN, A., AND SINGH, K. 2010. Learning 3d mesh segmentation and labeling. *ACM Trans. Graph.* 29 (July), 102:1–102:12.
- KIM, V. G., LIPMAN, Y., AND FUNKHOUSER, T. 2011. Blended intrinsic maps. *ACM Trans. Graph.* 30, 4 (Aug.), 79:1–79:12.
- KIM, V. G., LI, W., MITRA, N., DIVERDI, S., AND FUNKHOUSER, T. 2012. Exploring collections of 3d models using fuzzy correspondences. In *ACM SIGGRAPH 2012 papers*, SIGGRAPH '12, to appear.
- KUMAR, M. P., KOLMOGOROV, V., AND TORR, P. H. S. 2009. An analysis of convex relaxations for MAP estimation of discrete MRFs. *Journal of Machine Learning Research* 10, 71–106.
- LEORDEANU, M., AND HEBERT, M. 2006. Efficient map approximation for dense energy functions. *ICML '06*, 545–552.
- LIPMAN, Y., AND FUNKHOUSER, T. 2009. Mobius voting for surface correspondence. *ACM Trans. Graph.* 28, 3 (July), 72:1–72:12.
- MARANDE, W., AND BURGER, G. 2007. Mitochondrial dna as a genomic jigsaw puzzle. *Science* 318 (October), 415.
- MÉMOLI, F., AND SAPIRO, G. 2005. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics* 5, 3, 313–347.
- NGUYEN, A., BEN-CHEN, M., WELNICKA, K., YE, Y., AND GUIBAS, L. 2011. An optimization approach to improving collections of shape maps. *SGP '11*, 1481–1491.
- OVSJANIKOV, M., BEN-CHEN, M., SOLOMON, J., BUTSCHER, A., AND GUIBAS, L. 2012. Functional maps: A flexible representation of maps between shapes. *ACM Transactions on Graphics* 31, 4.
- ROBERTS, R., SINHA, S. N., SZELISKI, R., AND STEEDLY, D. 2011. Structure from motion for scenes with large duplicate structures. In *CVPR*, 3137–3144.
- RUBNER, Y., TOMASI, C., AND GUIBAS, L. J. 2000. The earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vision* 40, 2 (Nov), 99–121.
- SIDI, O., VAN KAICK, O., KLEIMAN, Y., ZHANG, H., AND COHEN-OR, D. 2011. Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering. *ACM Trans. Graph.* 30, 6 (Dec.), 126:1–126:10.
- SOLOMON, J., NGUYEN, A., BUTSCHER, A., BEN-CHEN, M., AND GUIBAS, L. 2012. Soft maps between surfaces. *Computer Graphics Forum* 31, 5, 1617–1626.
- SUMNER, R. W., AND POPOVIĆ, J. 2004. Deformation transfer for triangle meshes. *ACM Trans. Graph.* 23, 3 (Aug.), 399–405.
- SUN, J., OVSJANIKOV, M., AND GUIBAS, L. 2009. A concise and provably informative multi-scale signature based on heat diffusion. *Symposium on Geometry Processing '09*, 1383–1392.
- VAN KAICK, O., TAGLIASACCHI, A., SIDI, O., ZHANG, H., COHEN-OR, D., WOLF, L., AND HAMARNEH, G. 2011. Prior knowledge for shape correspondence. *Computer Graphics Forum* 30, 2, 553–562.
- ZACH, C., KLOPSCHITZ, M., AND POLLEFEYS, M. 2010. Disambiguating visual relations using loop constraints. In *CVPR*, 1426–1433.

## A Connection Between Unsharpened Soft Maps and Spectral Embedding Distance

Let  $A \in \mathbb{R}^{Mn \times Mn}$  be the big initial map matrix for points on all shapes, i.e. each non-zero  $M \times M$  block  $A_{ij} = C_{ij}^{\text{init}}$  is given by the corresponding initial map specified in the model graph. Denote  $C \in \mathbb{R}^{Mn \times Mn}$  as the big soft map matrix that collects all unsharpened soft maps, i.e. the  $M \times M$  block  $C_{ij}$  is given by the unsharpened soft map from shape  $S_i$  to  $S_j$ . Let  $(\lambda_i, \mathbf{u}_i)$ ,  $1 \leq i \leq Mn$  be the eigenvalue-eigenvector pairs of matrix  $A$ , we can write

$$C = \sum_{k=0}^{\infty} \alpha_k A^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} \left( \sum_{j=1}^{Mn} \lambda_j^k \mathbf{u}_j \mathbf{u}_j^T \right) = \sum_{j=1}^{Mn} e^{\lambda_j t} \mathbf{u}_j \mathbf{u}_j^T.$$

Define a distance  $d(s, s')$  between two points  $s$  and  $s'$  as the distance between their corresponding vectors  $C(s, \cdot)$  and  $C(s', \cdot)$ :

$$d^2(s, s') = \|C(s, \cdot) - C(s', \cdot)\|^2 = \sum_{j=1}^{Mn} e^{\lambda_j t} \|\mathbf{u}_j(s) - \mathbf{u}_j(s')\|^2.$$

In other words,  $d(s, s')$  is the distance between  $s$  and  $s'$  in the space spanned by weighted eigenvectors  $\{e^{\lambda_j t} \mathbf{u}_j\}$ . This gives a connection between our approach and [Kim et al. 2012].

