3D Shape Correspondence by Isometry-Driven Greedy Optimization

Yusuf Sahillioğlu and Yücel Yemez
Koç University
Rumelifeneri Yolu, 34450, Sarıyer, Istanbul, Turkey
{ysahillioglu,yyemez}@ku.edu.tr

Abstract

We present an automatic method that establishes 3D correspondence between isometric shapes. Our goal is to find an optimal correspondence between two given (nearly) isometric shapes, that minimizes the amount of deviation from isometry. We cast the problem as a complete surface correspondence problem. Our method first divides the given shapes to be matched into surface patches of equal area and then seeks for a mapping between the patch centers which we refer to as base vertices. Hence the correspondence is established in a fast and robust manner at a relatively coarse level as imposed by the patch radius. We optimize the isometry cost in two steps. In the first step, the base vertices are transformed into spectral domain based on geodesic affinity, where the isometry errors are minimized in polynomial time by complete bipartite graph matching. The resulting correspondence serves as a good initialization for the second step of optimization in which we explicitly minimize the isometry cost via an iterative greedy algorithm in the original 3D Euclidean space. We demonstrate the performance of our method on various isometric (or nearly isometric) pairs of shapes for some of which the ground-truth correspondence is available.

1. Introduction

3D shape correspondence methods aim to find a mapping between the surface points of two given shapes, or more generally, they seek on two given shapes for pairs of surface points that are similar or semantically equivalent. 3D shape correspondence is a fundamental problem in both computer vision and computer graphics with numerous applications such as mesh morphing [1], mesh parameterization [2], shape registration [3], shape matching [4],[5] and analysis of sequential meshes [6]. In this paper we address the problem of establishing correspondence between isometric (or nearly isometric) shapes. Isometric shapes appear in many different contexts such as different poses of an articulated object, models of a mesh sequence representing the motion of a human actor, or two shapes representing different but semantically similar objects (e.g., two different humans or animals).

If two shapes are perfectly isometric, then there exists an isometry, i.e., a distance-preserving mapping, between these shapes such that the geodesic distance between any two points on one shape is exactly the same as the geodesic distance between their correspondences on the other. However, two shapes are hardly ever perfectly isometric, even for different poses of an articulated object due to imperfections of the modeling process and/or geometry discretization errors. Hence the goal of isometric correspondence methods existing in the literature is usually to find an optimal mapping that minimizes the amount of deviation from isometry. Isometry errors are however expensive to compute and optimize in the original 3D Euclidean space [7]. A common strategy to avoid excessive computation is to embed shapes into a different domain where Euclidean distances can approximate geodesic distances so that the isometry errors can efficiently be measured and optimized in the embedding space [5], [8]-[13].

Euclidean embedding can be achieved in various ways. While Jain et al. [9] have used Principle Component Analysis (PCA) on the normalized geodesic affinity matrices to achieve spectral embedding, Elad et al. [5] have employed multidimensional scaling (MDS) to create pose and bending invariant shape signatures. Other related examples are due to Ovsjanikov et al. [8] who define the embedding space by the eigenfunctions of the Laplace-Beltrami operator, and Lipman et al. [12] who apply Möbius transformation in order to transform the given shapes into a canonical coordinate frame on the complex plane where deviations from isometry can be computed based on mutually closest points. Carcassoni et al. [11] have achieved correspondence by using EM algorithm after embedding the problem into spectral domain via alternative proximity weighting matrices, e.g., the sigmoidal proximity matrix, instead of the standard Gaussian proximity matrix. Although the matching process can be enhanced by incorporation of local geometric properties as in [14], a problem common to these embedding-

978-1-4244-6985-7/10/$26.00 ©2010 IEEE

based techniques is that they all produce approximate solutions since they can measure deviations from isometry only approximately in the embedding space. Hence all these techniques have room for improvement. In this paper, we describe a computationally efficient and robust method which finds an optimal correspondence that minimizes deviations from isometry in the original 3D Euclidean space.

2. Problem Description

We cast the 3D shape correspondence problem as a surface correspondence problem. To this effect, we divide the given source and target shapes into surface patches of equal area. We assume that the given shapes are (nearly) isometric and represented as manifold meshes on which geodesic distances can easily be computed. We denote the resulting two sets of surface patches by S and T, where each patch is represented by the point at its center, a base vertex as we call it. The problem is then reduced to searching for an optimal correspondence between the surface patches, or the base vertices of S and T. Note that one can find more than one optimal correspondence for symmetrical objects. The optimal correspondence that we seek for has to meet two requirements: It has to be as complete (not partial, assuming perfect isometry such that $|S| = |T|$) and isometric as possible, i.e., has to minimize deviations from isometry through the following isometry cost:

$$D_{\text{iso}} = \frac{1}{|\S|} \sum_{(s_i,t_k)\in\S} \frac{\sum_{(s_j,t_l)\in\S} |d(s_i,s_j) - d(t_k,t_l)|}{|\S|} \quad (1)$$

where $\S$ denotes the set of correspondences between S and T, and $d(.,.)$ denotes the normalized geodesic distance between two surface patches, or between two base vertices, or more generally between two points on a given surface. We optimize the isometry error in two steps. In the first step, the base vertices are transformed into spectral domain based on geodesic affinity, where the isometry errors are minimized in polynomial time by complete bipartite graph matching. The resulting correspondence serves as a good initialization for the second step of optimization in which we explicitly minimize the isometry cost via an iterative greedy algorithm in the original 3D Euclidean space.

3. Base vertices

The base vertices in S and T are computed by launching the Dijkstra's shortest paths algorithm from an arbitrary source vertex. When a base vertex is selected, all the vertices lying within its patch of radius r are marked. The next base vertex is then selected arbitrarily from the unmarked vertices. When this is repeated until no unmarked vertex is left, we obtain a partitioning of the surface into possibly overlapping patches of equal size, where the patch centers, i.e., the base vertices, are at least at distance r apart from each other [4]. Once the base vertices are paired up accurately between two shapes through the process described in this paper, the surface patches defined by these bases hold valuable information that can be used to expand the obtained correspondence to a denser one if desired. All the base vertices and some of the resulting patches for a given shape are shown in Figure 1 for two different values of r.

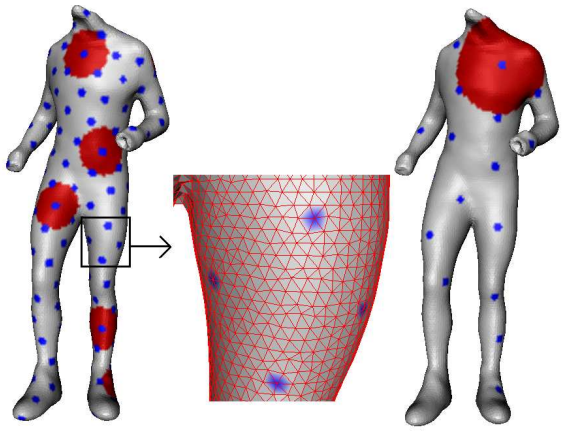

Figure 1. $|S| = 184$ base vertices extracted from a mesh of 16K vertices (left). Zoom on the leg (middle). $|S|$ becomes 48 by simply updating r (right). Some random patches painted as well.

4. Spectral embedding and alignment

Running the Dijkstra shortest paths algorithm from each base vertex yields the geodesic distances between all pairs of bases. These pairwise distances, when exposed to a zero-mean unit-variance Gaussian kernel, form a geodesic affinity matrix, $A_{ij} = \exp(-d^2(i,j)/2)$ for each of the base vertex sets S and T. Each of these base vertex sets is then transformed into the K-dimensional spectral domain using the K leading (scaled) eigenvectors of the associated geodesic affinity matrix [9]. We will denote the base vertices embedded in the spectral domain by $\hat{S}$ and $\hat{T}$, respectively for each shape. The geodesic distances between base vertices in the Euclidean space now approximately correspond to L_2 distances between their K-dimensional embeddings in the spectral domain. Although the same transformation is applied to both shapes, due to arbitrary sign flips of eigenvectors, a disambiguation process is required to test the 2^K different possible embeddings for the best alignment. We measure the alignment of each such embedding $\hat{S}_k$ with the fixed $\hat{T}$ by means of the cost $L_k = \sum_{i=1}^{|\hat{S}_k|}(\|\hat{s}_{i,k} - \hat{t}_i\|)$, i.e., the sum of Euclidean distances between each closest pair $(\hat{s}_{i,k}, \hat{t}_i)$. The embedding $\hat{S}_k$ producing the minimum L_k aligns best with $\hat{T}$. This alignment operation is visualized in Figure 2 for $K = 3$.

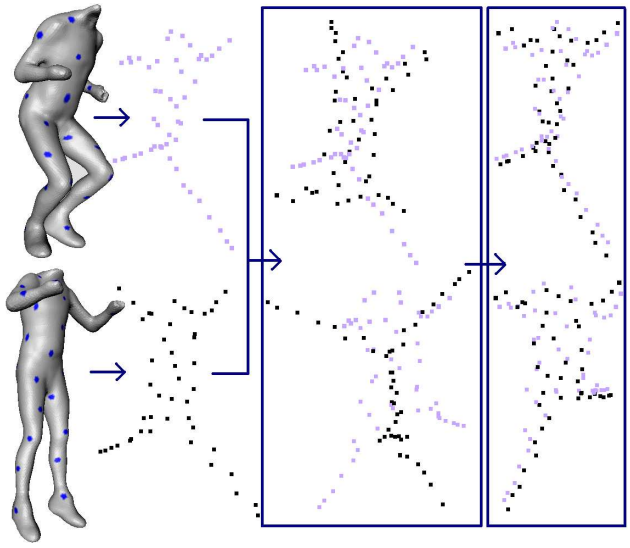

Figure 2. Two shapes along with the spectral embeddings of their base vertices (left), the alignments obtained using an arbitrary permutation of the eigenvectors (left box), and using the best permutation (right box). The boxes display two different views for visual convenience.

5. Optimization

The optimal correspondence is determined in two steps. We first find an initial correspondence in the spectral domain, which then serves as the initialization of the Euclidean greedy optimization performed in the second step, as described in the sequel.

5.1. Initial correspondence

We create a complete bipartite graph G on which the minimum-weight perfect matching is sought. The aligned base vertices $\hat{S}$ and $\hat{T}$ form the disjoint vertex sets of G which is made complete by connecting every vertex of one set to every vertex of the other with edges weighted by $c_{ij} = \|\hat{s}_i - \hat{t}_j\|$, specifying the cost of matching $\hat{s}_i$ and $\hat{t}_j$. Since the cardinalities of disjoint sets must match for a perfect matching, we introduce virtual vertices in the deficient side with connector edges of ∞ weights. Note that the numbers of base vertices are almost equal for a given pair of isometric shapes but need not be exactly the same due to variations from isometry. Hence at the end of the optimization process, some base vertices may be left unassigned. We currently employ the Hungarian algorithm [15] to solve the complete bipartite graph matching problem, which provides us with the optimal initial correspondence when the amount of deviation from isometry is measured in the embedding space.

5.2. Isometry-driven greedy optimization

We have developed an iterative greedy optimization algorithm that minimizes the isometry cost D_{iso} given by Eq. 1 in the original Euclidean space. The greedy optimization starts with the initial correspondence, $\S^0$, found via complete bipartite graph matching in the spectral domain. The algorithm traverses the correspondence list $\S$ and replaces each time the current pair (s_i, t_j) with (s_i, t_c) if this replacement decreases the isometry cost. The accumulation of these greedy decisions, each of which considers a local improvement, eventually leads to an optimal solution on D_{iso} as we re-traverse $\S$ until convergence, i.e., until D_{iso} no longer improves.

The replacement of the correspondence (s_i, t_j) with (s_i, t_c) relies on the votes collected from the correspondences of the base neighbors of s_i. Each of these correspondences votes for its own base neighbors and the most voted base vertex, t_c, replaces t_j if this replacement decreases the following vertex-based isometry cost,

$$v_{\text{iso}}(s_i|t_k) = \frac{1}{|\S|} \sum_{(s_j,t_l)\in\S} |d(s_i, s_j) - d(t_k, t_l)| \qquad (2)$$

with $(s_i, t_k) \in \S$, i.e., if $v_{\text{iso}}(s_i|t_c) < v_{\text{iso}}(s_i|t_j)$, where (rare) ties are resolved by picking the name that yields the minimum D_{iso}. In addition to a potential replacement concerning s_i, we also consider the current match s_i^+ of t_c by replacing (s_i^+, t_c) with (s_i^+, t_j) if $v_{\text{iso}}(s_i^+|t_j) < v_{\text{iso}}(s_i^+|t_c)$. The *Voting* process is demonstrated on an example in Figure 3.

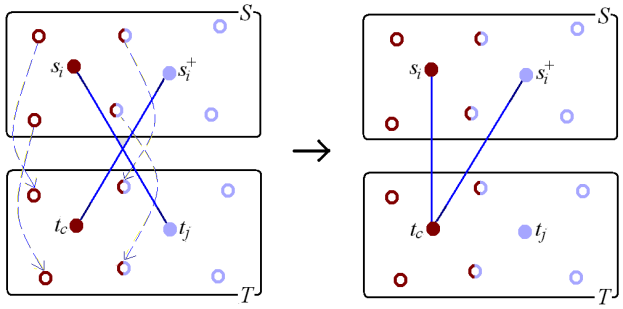

Figure 3. Each base vertex (filled circles) has 4 neighbors (empty circles) and (s_i, t_c) is the ground-truth correspondence. When (s_i, t_j) is in process, *Voting* realizes the correspondences of the base neighbors of s_i (pointed by the dashed arrows) (left-bottom) and takes votes for their neighbors. Since the most voted base t_c is different than t_j in this case, the correspondence (s_i, t_j) is replaced with (s_i, t_c) after checking $v_{\text{iso}}(s_i|t_c)$ vs. $v_{\text{iso}}(s_i|t_j)$. Note that the pair (s_i^+, t_c) is also considered next for a possible replacement with (s_i^+, t_j), though not illustrated in the figure.

The overall greedy optimization algorithm and the *Voting* procedure are given in pseudocode as follows:

Greedy optimization algorithm

Input: Initial Correspondence list $\S^0$
Output: Correspondence list $\S$ that minimizes D_{iso}
$\S = \S^0$
Until convergence (until D_{iso} no longer improves)
  For next pair $(s_i, t_j) \in \S$
    $t_c =$ ***Voting***$((s_i, t_j))$
    If $v_{\text{iso}}(s_i|t_c) < v_{\text{iso}}(s_i|t_j)$
      (s_i, t_c) replaces (s_i, t_j) in $\S$
    Let s_i^+ be the match of t_c, i.e., $(s_i^+, t_c) \in \S$
    If $v_{\text{iso}}(s_i^+|t_j) < v_{\text{iso}}(s_i^+|t_c)$
      (s_i^+, t_j) replaces (s_i^+, t_c) in $\S$

Voting(Pair $(s_i, t_j) \in \S$)
$V = \{\eta(t_r) \mid (s_p, t_r) \in \S$ where $s_p \in \eta(s_i)\}$
  where $\eta(s_i)$ is the set of base neighbors of s_i
$t_c =$ The most voted base listed in V
Return t_c as the candidate replacer

Note that the performance of the voting scheme, hence of the greedy algorithm, highly depends on the assumption that the base neighbors are mostly paired up correctly, which holds thanks to the initial correspondence $\S^0$ that serves as a good initialization.

6. Computational complexity

The base vertices are obtained in $O(V_s \log V_s)$ time because the shortcut Dijkstra shortest paths for each patch adds up to one Dijkstra algorithm spanning V_s, where V_s is the number of vertices in the original mesh. The geodesic affinity matrix is computed in $O(|S|V_s \log V_s)$ time. The eigenanalysis for embedding into the K-dimensional spectral domain takes $O(|S|^3)$ time followed by $O(2^K|S|^2)$ operations for alignment where $K \leq 6$. Repetition of all these operations for T is free of asymptotic cost. Assuming, with no loss of generality, $|S| \geq |T|$, $O(|S|^3)$ time is necessary for the Hungarian assignment which could be replaced with Edmonds' blossom algorithm [16] that requires $O(|S|^2 \log |S|)$ time. The greedy optimization demands $O(|S|^2)$ time since the computation of v_{iso} can be performed in linear time for each of the $|S|$ pairs, and the whole process is repeated until D_{iso} converges where the number of iterations never exceeds 7 in our tests thanks to the good initialization. Under the valid assumption of $|S| \ll V_s$, the overall complexity is then $O(|S|V_s \log V_s)$.

7. Results and Discussion

We have conducted experiments on two mesh sequences, *Jumping Man* [17] and *Dancing Man* [18] as we refer to them, both originally reconstructed from real scenes and each representing the real motion of a human actor. The original meshes of the sequences are all uniform and given at high resolution with fixed connectivity hence we have the ground-truth dense correspondences in both cases. We have also tested our method on a relatively low-resolution Dog-Wolf shape pair from the Nonrigid World 3D database [7].

In addition to visual results which display the computed correspondences in Figures 4-6, we also provide quantitative results in Table 1 which helps assessing the performance in terms of average and vertex-based isometry and ground-truth errors. The average deviation from isometry is measured by the isometry cost D_{iso} (Eq. 1) whereas vertex-based isometry deviation is given by v_{iso} (Eq. 2). On the other hand, the average and vertex-based ground-truth correspondence errors are measured by using the ground-truth correspondence pairs (s_i, t_i) whenever available, respectively as

$$D_{\text{ground}} = \frac{1}{|\S|} \sum_{i,(s_i,t_j)\in|\S|} d(t_i, t_j) \tag{3}$$

$$v_{\text{ground}}(s_i|t_j) = d(t_i, t_j), \text{ where}(s_i, t_j) \in \S \tag{4}$$

In Table 1, we provide, for different cases, the values of the performance measures D_{iso}, v_{iso}, D_{ground} and v_{ground} before and after greedy optimization to emphasize the benefit of the latter. The vertex-based measures are given only for the worst matches, which we denote by v^*_{iso} and v^*_{ground}. The average-based measures are each computed over 10 different runs of the algorithm on 10 different pairs corresponding to different poses of the actor in the corresponding sequence. We have to note that in some cases our algorithm may confuse the symmetrical parts of two given shapes, which we have excluded in our experiments in order not to artificially burst the ground-truth performance errors. All the distance-based measures are given as a factor of the patch radius r for better interpretation of the correspondence errors. Through all experiments, the dimension of the spectral domain was set to be $K = 6$ and the number of base vertices was about 50 as determined in proportion to the surface area. We have also tested the contribution of a local shape descriptor to the initial correspondence, i.e., in addition to L_2-only cost used for complete bipartite graph matching (Section 5.1), we have also tried a weighted cost, $w_1 L_2 + w_2 \kappa$, with $w_1 = w_2 = 0.5$, where κ is the Gaussian curvature [19], computed for each base vertex as an average value over the associated patch. This local feature slightly improves the initial correspondence but eventually has no contribution to the greedy optimization phase as observed in Table 1. We also observe that the correspondence performance is significantly improved after the greedy optimization, for all cases and for all performance measures. It is interesting to see that the improvements in the ground-truth error performances are generally larger as compared to

improvements in isometry, which verifies the intuition that small deviations from isometry can lead to a significant loss in the correspondence performance.

The execution time of our implementation is mainly dominated by the number of the vertices in the original meshes due to the geodesic distance computation. On a 2GB 2.1GHz duo laptop, correspondences for Jumping Man ($V_s = 16K$), Dancing Man ($V_s = 20K$), Jumping-Dancing, and Dog-Wolf ($V_s = 3.4K$) pairs are achieved in 102, 164, 126, and 10 seconds, respectively.

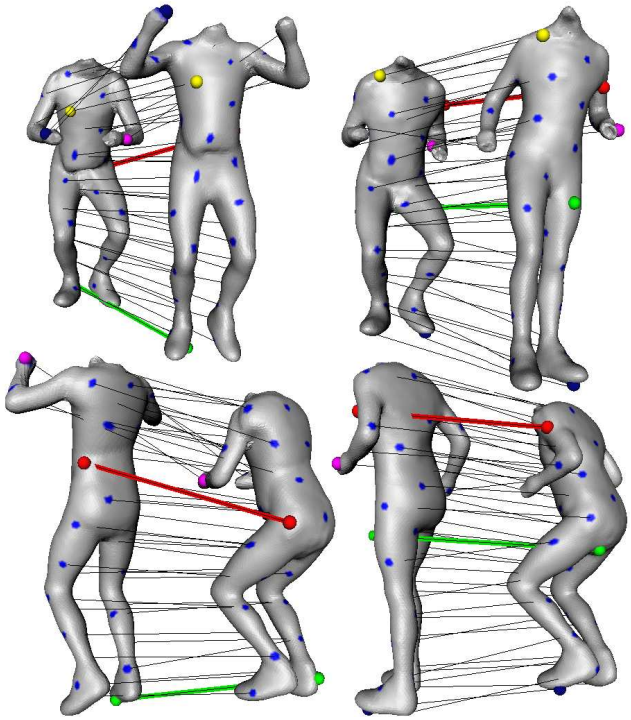

Figure 4. Correspondences displayed from two different views (top-bottom) for two different *Jumping Man* pairs (left-right). Bold green and red lines represent the worst matches w.r.t. ground-truth and isometry costs, respectively. Some other matches are also highlighted with similarly colored spheres.

8. Conclusion

We have presented a novel and computationally efficient algorithm that can achieve robust correspondences between two isometric (or nearly isometric) shapes represented as manifold triangle meshes. We first establish a sufficiently good initial correspondence that minimizes the isometry deviations in the spectral domain by solving a complete bipartite graph matching problem. This initial correspondence is fed to a greedy optimization procedure which further minimizes the isometry cost in the original Euclidean space. In fact, our greedy optimization algorithm can be used to further optimize the output of any isometric surface correspondence method available in the literature.

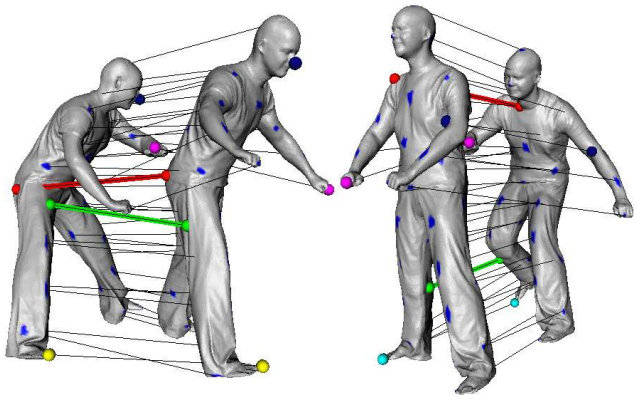

Figure 5. Correspondences displayed for two different *Dancing Man* pairs (left-right). Bold green and red lines represent the worst matches w.r.t. ground-truth and isometry costs, respectively.

As future work, we will expand our relatively coarse correspondence to a denser one, which should be robust and computationally efficient thanks to the information encapsulated within the surface patches that are already matched through base vertices. Another obvious research direction is to extend our method to handle partially isometric shapes.

Acknowledgement

This work has been supported by TUBITAK under the project EEEAG-109E274.

References

[1] M. Alexa. Recent Advances in Mesh Morphing. *Computer Graphics Forum*, vol. 21, pp. 173–196, 2002.

[2] V. Kraevoy and A. Sheffer. Cross-Parameterization and Compatible Remeshing of 3D Models. *SIGGRAPH*, pp. 861-869, 2004.

[3] D. Anguelov, P. Srinivasan, H. Pang, D. Koller, S. Thrun and J. Davis. The Correlated Correspondence Algorithm for Unsupervised Registration of Nonrigid Surfaces. *NIPS*, vol. 37, pp. 33–40, 2004.

[4] M. Hilaga, Y. Shinagawa, T. Kohmura and T. Kunii. Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes. *SIGGRAPH*, pp. 203–212, 2001.

[5] A. Elad and R. Kimmel. On Bending Invariant Signatures for Surfaces. *Pattern Analysis and Machine Intelligence (PAMI)*, vol. 25, pp. 1285–1295, 2003.

[6] A. Zaharescu, E. Boyer, K. Varanasi and R. Horaud. Surface Feature Detection and Description with Applications to Mesh Matching. *Computer Vision and Pattern Recognition (CVPR)*, pp. 373–380, 2009.

[7] A. M. Bronstein, M. M. Bronstein and R. Kimmel. Efficient Computation of Isometry-Invariant Distances Between Sur-

Setup		Initial correspondence		After greedy optimization	
Pair	Cost	$D_{\text{ground}}, v^*_{\text{ground}}$	$D_{\text{iso}}, v^*_{\text{iso}}$	$D_{\text{ground}}, v^*_{\text{ground}}$	$D_{\text{iso}}, v^*_{\text{iso}}$
Jumping Man	L_2	$0.548r, 2.452r$	$0.397r, 0.733r$	$0.376r, 1.312r$	$0.256r, 0.673r$
Jumping Man	$L_2 + \kappa$	$0.582r, 2.229r$	$0.420r, 0.754r$	$0.542r, 2.171r$	$0.401r, 0.657r$
Dancing Man	L_2	$0.544r, 7.485r$	$0.415r, 1.567r$	$0.201r, 1.239r$	$0.244r, 0.776r$
Dancing Man	$L_2 + \kappa$	$0.549r, 7.485r$	$0.470r, 2.702r$	$0.234r, 1.227r$	$0.412r, 1.341r$
Jumping-Dancing	L_2	n/a	$0.571r, 1.548r$	n/a	$0.397r, 0.899r$
Jumping-Dancing	$L_2 + \kappa$	n/a	$0.577r, 1.637r$	n/a	$0.401r, 0.891r$

Table 1. Correspondence performances, where v^* represents the error of the worst match. The patch radius r is approximately 0.1 for all cases, where r is given as normalized with respect to the maximum geodesic distance on the surface.

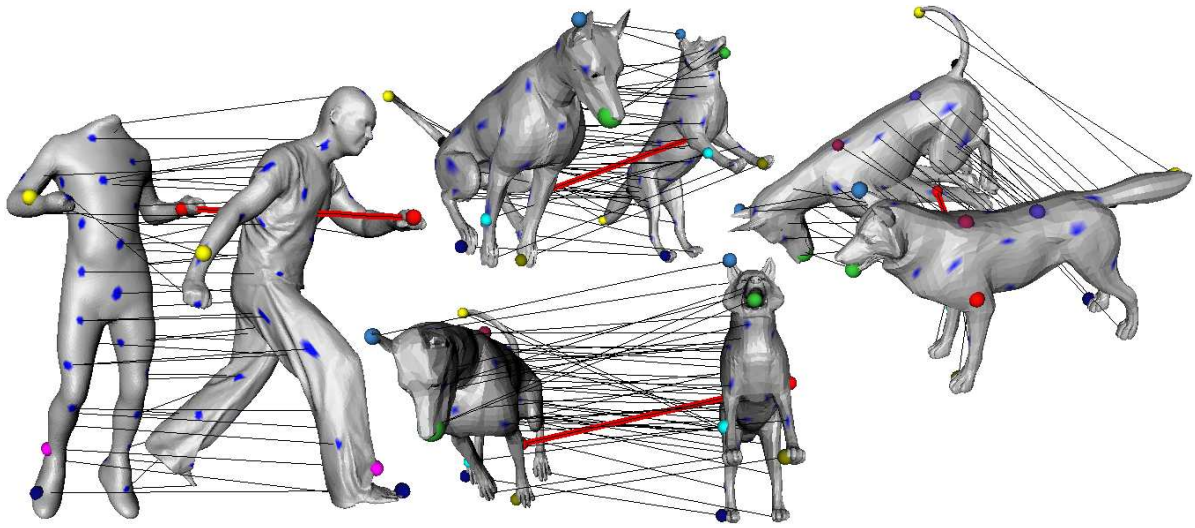

Figure 6. Correspondences for a Jumping-Dancing pair (left) and the Dog-Wolf pair (right). Bold red lines represent the worst matches w.r.t. isometry costs for each case. For the Dog-Wolf pair displayed in three different views, $(D_{\text{iso}}, v^*_{\text{iso}}) = (0.399r, 0.848r)$. The normalized patch radius r is 0.11 in both cases.

faces. *SIAM J. Scientific Computing*, vol. 28(5), pp. 1812–1836, 2006.

[8] M. Ovsjanikov, J. Sun and L. Guibas. Global intrinsic symmetries of shapes. *Computer Graphics Forum*, vol. 27(5), pp. 1341–1348, 2007.

[9] V. Jain and H. Zhang. Robust 3D Shape Correspondence in the Spectral Domain. *Shape Modeling and Applications (SMI)*, pp. 118–129, 2006.

[10] D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin and E. Boyer. Articulated Shape Matching Using Laplacian Eigenfunctions and Unsupervised Point Registration. *Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2008.

[11] M. Carcassoni and E. Hancock. Spectral Correspondence for Point Pattern Matching. *Pattern Recognition*, vol. 36, pp. 193–204, 2003.

[12] Y. Lipman and T. Funkhouser. Mobius Voting for Surface Correspondence. *SIGGRAPH*, vol. 28(3), 2009.

[13] S. Wuhrer, C. Shu and P. Bose. Posture Invariant Correspondence Of Triangular Meshes In Shape Space. *3-D Digital Imaging and Modeling*, 2009.

[14] M. R. Ruggeri, G. Patane, M. Spagnuolo and D. Saupe. Spectral-Driven Isometry-Invariant Matching of 3D Shapes. *Int. J. Computer Vision*, forthcoming.

[15] C. Papadimitriou and K. Steiglitz. Combinatorial Optimization: Algorithms and Complexity. *Prentice-Hall*, 1982.

[16] W. Cook and A. Rohe. Computing Minimum-Weight Perfect Matchings. *Journal on Computing*, vol. 11(2), pp. 138-148, 1999.

[17] P. Sand, L. McMillan and J. Popovic. Continuous Capture of Skin Deformation. *Int. Conf. on Computer Graphics and Interactive Techniques*, pp. 578–586, 2003.

[18] E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H. P. Seidel and S. Thrun. Performance Capture from Sparse Multi-view Video. *SIGGRAPH*, pp. 1–10, 2008.

[19] M. Meyer, M. Desbrun, P. Schröder and A. Barr. Discrete Differential-Geometry Operators for Triangulated 2-Manifolds. *VisMath*, 2002.