Minimum-Distortion Isometric Shape Correspondence Using EM Algorithm

Yusuf Sahillioğlu and Yücel Yemez

Abstract—We present a purely isometric method that establishes 3D correspondence between two (nearly) isometric shapes. Our method evenly samples high-curvature vertices from the given mesh representations, and then seeks an injective mapping from one vertex set to the other that minimizes the isometric distortion. We formulate the problem of shape correspondence as combinatorial optimization over the domain of all possible mappings, which then reduces in a probabilistic setting to a log-likelihood maximization problem that we solve via the Expectation-Maximization (EM) algorithm. The EM algorithm is initialized in the spectral domain by transforming the sampled vertices via classical Multidimensional Scaling (MDS). Minimization of the isometric distortion, and hence maximization of the log-likelihood function, is then achieved in the original 3D euclidean space, for each iteration of the EM algorithm, in two steps: by first using bipartite perfect matching, and then a greedy optimization algorithm. The optimal mapping obtained at convergence can be one-to-one or many-to-one upon choice. 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.

Index Terms—3D isometric shape correspondence, multidimensional scaling, spectral embedding, isometric distortion, greedy optimization, bipartite perfect matching, EM algorithm

✦

1 INTRODUCTION

THREE-DIMENSIONAL shape correspondence methods aim to find a mapping between the surface points of two given shapes or, more generally, they seek two given shapes for pairs of surface points that are similar or semantically equivalent [1]. Shape correspondence is a fundamental problem in both computer vision and computer graphics, with numerous applications such as mesh morphing [2], mesh parameterization [3], deformation transfer [4], shape registration [5], shape matching [6], [7], analysis of sequential meshes [8], and statistical shape modeling [9]. In this paper, we address the problem of establishing correspondence between isometric (or nearly isometric) shapes. Isometric shapes appear in various 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 digital shapes are hardly ever perfectly isometric, even for different poses of a rigid 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 rather to find a mapping that minimizes the amount of deviation from isometry. A common strategy to achieve this is to embed shapes into a different (e.g., spectral) domain where geodesic distances are replaced with euclidean distances so that isometric distortion can be efficiently measured and optimized in the embedding space [7], [10], [11], [12], [13], [14], [15], [16]. However, since the euclidean embedding process itself introduces a distortion [17], deviations from isometry can be measured only approximately in the embedding space. Hence, all these methods mentioned above produce approximate and/or ambiguous solutions, and thus have room for improvement. To remove the approximation error, Bronstein et al. [18] propose embedding one shape into the surface of the other by using generalized multidimensional scaling (GMDS), which, however, requires minimization of a nonconvex function that is difficult and expensive to optimize. The main contribution of our paper is a novel shape correspondence method that minimizes the isometric distortion directly in the 3D euclidean space, i.e., in the domain where isometry is originally defined, with a computationally efficient algorithm.

2 RELATED WORK

Isometry is an important clue for shape correspondence; not only since most real-world deformations are isometric, but also because semantically similar shapes have similar metric structures. There are different ways of exploiting isometry for shape correspondence. One way is to minimize the deviation from isometry, indirectly in some embedding space, where euclidean distances approximate geodesic distances. Euclidean embedding, in the context of shape analysis, can be achieved by using various techniques such as classical MDS (Multidimensional Scaling) [10], [11], [16], least-squares MDS [7], [13], and spectral analysis of the

---

• The authors are with the Department of Computer Engineering, Koç University, Sariyer 34450, Istanbul, Turkey.
E-mail: {ysahillioglu, yyemez}@ku.edu.tr.

graph Laplacian [14] or of the Laplace-Beltrami operator [19]. Some recent methods propose using embeddings such as the heat kernel [20] and the canonical diffusion embedding [21], which are invariant under isometric deformations, but which are not euclidean in the multidimensional scaling sense. Another three recent examples, which can be applied to only shapes with genus zero, are based on the Möbius transformation, which is used for conformal embedding of the given shapes into a canonical coordinate frame on the complex plane where deviations from isometry are approximated based on mutually closest points [15], [22], [23]. A problem common to these embedding-based techniques is that they all produce approximate and/or ambiguous solutions since they can measure deviations from isometry only approximately in the embedding space. In order to eliminate the approximation error, Bronstein et al. [24] propose embedding one of the shapes to be matched into the surface of the other via the generalized MDS, which requires minimization of a nonconvex stress function. To optimize this function, they use an iterative gradient-type algorithm, which, however, results in two major drawbacks. First, in order to avoid convergence to local minima, they employ a coarse-to-fine (multiresolution) optimization strategy [18], which may in turn yield artificial symmetric flips in the computed correspondences due to initial coarse sampling. Second, since the algorithm is based on gradient descent, it produces subvertex matchings which do not necessarily coincide with the initial sampling, yielding clustered correspondence samples.

The embedding process brings in distortion but reduces the correspondence problem to an alignment or point-to-point matching problem which is easier to solve. Jain and Zhang [10], for example, employ the TPS-RPM (Thin-Plate SplineRobust Point Matching) algorithm of [25] for nonrigid alignment of the embedded points. Other examples are the Hungarian algorithm employed in [13] and [16], high-order graph matching in [22], graph-matching based on dual decomposition in [26], Möbius voting scheme in [15], unsupervised point clustering in [14], EM-algorithm in [11] and [14], nearest point matching in [20], and association graphs in [21].

Isometric correspondence methods may incorporate local shape information into their schemes. In [15], for example, Gaussian curvature is used to find an initial set of feature points upon which the rest of the correspondence algorithm relies, whereas in [22], the extrinsic curvature and orientation information (also including texture if available) is used to augment the intrinsic global shape (isometric) information. Some methods rely mainly on local shape similarity, using descriptors such as spherical harmonics as in [27], histogram of oriented gradients as in [8], mean curvature as in [28], and shape contexts as in [29]. Local shape similarity is an important clue for shape correspondence, especially in the case of nonisometric deformations, but otherwise it is considered as less reliable than global shape information such as isometry. The methods which rely only on local geometric information may not perform well when the shapes to be matched exhibit large variations in their local geometry, or may easily confuse surface parts when there are many points that are locally similar. Hence,

some feature-based correspondence search algorithms also include a pruning procedure that takes into account isometric clues by enforcing geodesic consistency [28], [30], [31], [32]. In a recent work, Wang et al. [33] define an objective function in terms of local diffusion information as well as global isometric distortion that also accounts for shape variability in a probabilistic setting. They formulate the correspondence problem as a graph labeling problem and then solve it by graph matching based on the dual decomposition technique proposed in [26]. Although this is a nice mathematical formulation, the optimization process is computationally very demanding.

Another distinction between shape correspondence methods is whether they target sparse or dense correspondence. Most embedding-based methods naturally support dense correspondence but the computational load is usually a limiting factor. There also exist methods which primarily aim to find a small number of feature correspondences [30], [31], [21]. These methods regard the sparse correspondence problem as the main challenge since, based on a sufficient number of reliable landmarks, cross-parameterization [3] or some other form of interpolation technique can always be applied to obtain a dense correspondence [2]. However, in this coarse-to-fine approach, to decide on the degree of sparsity that would lead to a robust dense correspondence is always a problematic issue. A particular example of sparse correspondence methods is the deformation-driven approach of [30]. In this method, shape extremities are first determined based on the average squared geodesic distance field. Then, an optimal correspondence is sought between these extremities via combinatorial tree traversal by pruning the search space according to some criteria based on local shape similarity and geodesic consistency. For each candidate correspondence set, the source shape is deformed to the target based on these small number of landmarks (anchor points), and the correspondence with the smallest distortion gives the best matching. Another similar deformation-driven method is presented also in [31]. Both methods can handle large deformations, but their computational cost is very high due to the repeated deformation process that they involve.

In this paper, we present a purely isometric method that finds an optimal correspondence between two shapes. We first evenly sample high-curvature vertices from the given shapes and then seek a minimum-distortion mapping from one vertex set to the other. Hence, our correspondence method can be regarded as sparse, though the density of the mapping can be increased as desired by adjusting the sampling rate. Since the mapping is sought between two evenly sampled vertex sets, it becomes possible to formulate minimization of the isometric distortion as a combinatorial optimization problem over the domain of all possible mappings. The resulting optimization problem can then be efficiently solved in the original 3D euclidean space as free of embedding errors by using an expectation-maximization (EM) algorithm that we initialize in the spectral domain. In fact, our method can also be used to further improve the performance of any isometric correspondence method existing in the literature. A preliminary version of our method was presented in [16], which we

substantially improve in this paper mainly by solving the combinatorial optimization problem in a probabilistic setting via EM algorithm. We also note that we have very recently proposed a dense shape correspondence method [34] that minimizes the isometric distortion via combinatorial matching in a coarse-to-fine fashion. Since the goal of [34] is to achieve a dense correspondence, i.e., a matching between all vertices of two given high resolution meshes, the focus is rather on computational efficiency, and hence the method is less accurate than the one we present in this paper in achieving sparse correspondence. Moreover, the dense correspondence method in [34] suffers from symmetric flips due to initial coarse sampling like the multi-resolution GMDS method [18].

3 PROBLEM FORMULATION AND OVERVIEW

We address the problem of 3D shape correspondence between two (nearly) isometric shapes. We assume that each shape is represented by a manifold surface mesh of sufficiently high resolution on which geodesic distances can easily be computed. We designate one of the shapes as *source* and the other as *target*. Let S and T denote the two sets of points sampled uniformly on the source and the target, respectively. Since the points are evenly sampled over mesh representations, we refer to them as *base vertices* (Section 4). The problem then reduces to searching for an optimal mapping from the base vertex set S to T. Note that one can find more than one optimal correspondence for symmetrical objects. We require the optimal correspondence (mapping) to have two properties: 1) to be as complete and one-to-one as possible (note that $|S| = |T|$ only in the case of perfect isometry) and 2) to minimize the deviation from isometry, i.e., the isometric distortion function defined below:

$$D_{\mathrm{iso}}(\S) = \frac{1}{|\S|} \sum_{(s_i,t_j)\in\S} d_{\mathrm{iso}}(s_i,t_j), \qquad (1)$$

where $\S$ denotes the set of correspondence pairs between S and T, and

$$d_{\mathrm{iso}}(s_i,t_j) = \frac{1}{|\S|-1} \sum_{\substack{(s_l,t_m)\in\S \\ (\S s_l,t_m)\neq(s_i,t_j)}} |g(s_i,s_l) - g(t_j,t_m)|, \qquad (2)$$

where $g(.,.)$ is the geodesic distance between two base vertices, more generally, between two points on a given surface. Hence, $d_{\mathrm{iso}}(s_i,t_j)$ is the contribution of the individual correspondence (s_i,t_j) to the overall isometric distortion. Both d_{iso} and D_{iso} take values in the interval $[0,1]$ since the function g is normalized with respect to the maximum geodesic distance over the surface. Note that (1) can be seen as a variant of the generalized stress function defined in [24] that we compute between evenly sampled base vertices. Given S and T, finding the optimal correspondence $\S^*$ that minimizes the isometric distortion is actually a combinatorial problem, which we solve in a probabilistic framework using the EM algorithm (Section 5). We assume that the probability of a base vertex s_i being in correspondence with t_j can be defined in terms of isometric
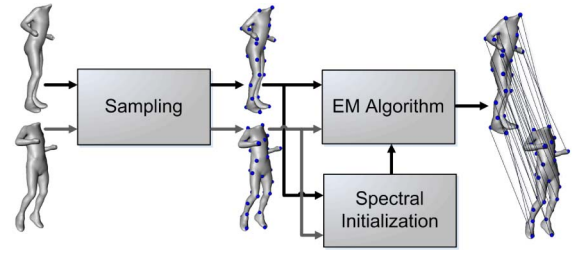

Fig. 1. Overall isometric shape correspondence scheme.

distortion. Given these probabilities, the optimal correspondence can be estimated by maximizing a likelihood function. The probability values, and hence the EM algorithm, are initialized based on the euclidean distances between the vertices embedded into spectral domain through classical MDS (Section 6). Maximization of the likelihood function, and hence minimization of the isometric distortion given in (1), is then achieved in the original 3D euclidean space for each iteration of the EM algorithm in two steps: by first using a bipartite graph matching algorithm (Section 7.1), and then a greedy optimization technique (Section 7.2), both with polynomial time complexity (Section 8). Note that neither the number of vertices of the original source and target meshes nor $|S|$ and $|T|$ need to be exactly equal. The block diagram of the overall shape correspondence scheme is given in Fig. 1. We demonstrate the performance of our method with experiments in Section 9, where we also compare our algorithm with two state-of-the-art techniques: the spectral method of Jain and Zhang [10] and the GMDS of Bronstein et al. [24].

4 SAMPLING

The sampling of base vertices should be as consistent as possible between source and target meshes. To achieve this goal, we sample each mesh separately, but impose a uniformity constraint and favor the vertices which are geometrically salient. Our sampling algorithm is as follows: We initially compute the Gaussian curvature at each vertex of the original mesh [35]. We then sort the vertices into a list in descending order with respect to their curvature values and select the top vertex of the list as the first base vertex. We launch the Dijkstra's shortest paths algorithm from this vertex and mark all the vertices lying within a distance r. The next base vertex is then picked as the first unmarked vertex in the list. When this is repeated until no unmarked vertex is left, we obtain a sampling of the surface where the base vertices are at least at distance r apart from each other.

The sampling algorithm described above is the same as the one proposed by Hilaga et. al in [6] except that the next base vertex is selected arbitrarily from unmarked vertices in [6], whereas we select it as the unmarked vertex with the highest curvature. Hence, our algorithm places the base vertices on local maxima of the Gaussian curvature and thereby generates a more consistent sampling on two shapes, which eventually improves the correspondence performance and yields more intuitive matchings. As verified in our experiments, this vertex sampling heuristic, as we call it the curvature-oriented evenly spaced (COES) sampling, provides a good start for our algorithm. The

resolution of the sampling, and hence the density of the ultimate correspondence, can easily be set thanks to the intuitive parameter r, the *sampling distance*, as we will refer to it. Although COES sampling algorithm cannot in general guarantee the same sampling on two shapes, the parameter r can be set to a sufficiently small value so that all critical vertices of the given meshes (for typical shapes) are included by the sampling process. Moreover, the sampling distance r puts an upper bound to the isometric distortion of the optimal correspondence $\S^*$ that minimizes (1). One such upper bound is given by $D_{\mathrm{iso}}(\S^*) \leq r$ in the case of perfect isometry.

We note that there exist in the literature several other techniques for uniform mesh sampling such as the farthest point sampling technique [36], which is commonly employed by shape correspondence algorithms such as in [15] and [24]. Our COES algorithm is, rather, built upon the sampling method of [6] since it provides us with a convenient framework to develop an efficient saliency-based uniform sampling algorithm.

5 EM FRAMEWORK

Minimization of the isometric distortion in (1) can be formulated in a probabilistic setting as the following log-likelihood maximization problem

$$\S^* = \arg \max_{\S} \log \mathrm{P}(\S|\mathcal{X}, \mathbf{Q}), \qquad (3)$$

where $\mathcal{X} = (S, T)$ is the observed data, i.e., the base vertices, and $\mathbf{Q}$ is the matrix with entries q_{ij}, each representing the probability of source base vertex s_i being in correspondence with target t_j such that $\sum_j q_{ij} = 1$. Note that since we seek an (almost) complete correspondence, we require every $s_i \in S$ to be matched with exactly one $t_j \in T$.

The likelihood maximization problem given in (3) can be solved by using the EM algorithm [37]. Let $\S^{(k)}$ and $\mathbf{Q}^{(k)}$ denote the estimates for $\S$ and $\mathbf{Q}$ at iteration k. Starting from an initial estimate $\mathbf{Q}^{(0)}$, the EM algorithm alternately recomputes the expected value of $\mathbf{Q}$ and the estimate $\S^{(k)}$ through the following E-step (for Expectation) and M-step (for Maximization)"

- E-step: $\mathbf{Q}^{(k)} = E(\mathbf{Q}|\mathcal{X}, \S^{(k-1)})$,
- M-step: $\S^{(k)} = \arg \max_{\S} \log \mathrm{P}(\S|\mathcal{X}, \mathbf{Q}^{(k)})$,

until convergence.

We assume that, for each base vertex s_i, the probability of correspondence with t_j can be modeled as a function of the resulting isometric distortion via the following Gibbs distribution:

$$q_{ij} = \mathrm{P}(t_j|s_i) = \frac{1}{T_i} e^{-\beta d_{\mathrm{iso}}(s_i, t_j)}, \qquad (4)$$

where β is a fixed positive factor that determines the sharpness of the distribution and T_i is a normalizing constant to be chosen such that the constraint $\sum_j q_{ij} = 1$ is satisfied for all i. The term $d_{\mathrm{iso}}(s_i, t_j)$ is the isometric distortion due to the correspondence of s_i with t_j (see (2)), and its expected value at iteration k can be estimated by averaging over all the correspondence pairs in $\S^{(k-1)}$:
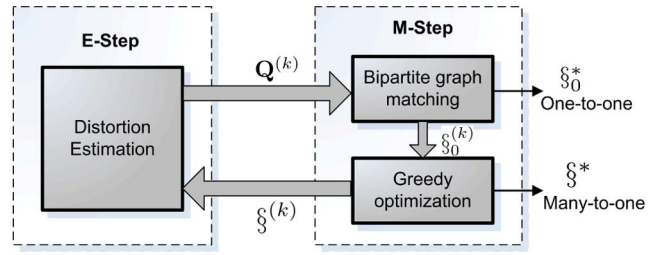

Fig. 2. EM algorithm.

$$
\begin{aligned}
d_{\mathrm{iso}}^{(k)}(s_i, t_j) &= E(d_{\mathrm{iso}}(s_i, t_j)|\mathcal{X}, \S^{(k-1)}) \\
&= \frac{1}{|\S^{(k-1)}| - 1} \sum_{\substack{(s_l, t_m) \in \S^{(k-1)} \\ (s_l, t_m) \neq (s_i, t_j)}} |g(s_i, s_l) - g(t_j, t_m)|, \quad (5)
\end{aligned}
$$

which can then be used to compute $q_{ij}^{(k)}$ in E-step:

$$q_{ij}^{(k)} = E(q_{ij}|\mathcal{X}, \S^{(k-1)}) = \frac{1}{T_i} e^{-\beta d_{\mathrm{iso}}^{(k)}(s_i, t_j)}. \qquad (6)$$

Hence in M-step (assuming statistical independence for assignments of s_i), we have

$$\S^{(k)} = \arg \max_{\S} \log \prod_{(s_i, t_j) \in \S} q_{ij}^{(k)}, \qquad (7)$$

and, substituting (6), we get

$$\S^{(k)} = \arg \max_{\S} \sum_i \log \frac{1}{T_i} - \beta \sum_{(s_i, t_j) \in \S} d_{\mathrm{iso}}^{(k)}(s_i, t_j). \qquad (8)$$

Here the first summation is constant, and hence can be ignored. Since the second term is always negative, using (5), the maximization in (8) becomes equivalent to the following minimization problem:

$$
\begin{aligned}
\S^{(k)} &= \arg \min_{\S} \sum_{(s_i, t_j) \in \S} d_{\mathrm{iso}}^{(k)}(s_i, t_j) \\
&= \arg \min_{\S} \frac{1}{|\S|} \sum_{(s_i, t_j) \in \S} \frac{1}{|\S^{(k-1)}| - 1} \sum_{\substack{(s_l, t_m) \in \S^{(k-1)} \\ (s_l, t_m) \neq (s_i, t_j)}} |g(s_i, s_l) - g(t_j, t_m)|.
\end{aligned}
$$
$$(9)$$

Hence, starting from an initial estimate $\mathbf{Q}^{(0)}$, or equivalently $\{d_{\mathrm{iso}}^{(0)}(s_i, t_j)\}$, the EM algorithm repeatedly computes (2) based on $\S^{(k-1)}$ and minimizes the overall isometric distortion (1) to generate $\S^{(k)}$ (see Fig. 2). The EM algorithm is expected to converge to a local minimum [37], as also verified by our experiments presented in Section 9.

6 INITIALIZATION

In this section, we describe how we initialize the EM algorithm, that is, how we obtain good initial estimates for the entries of the probability matrix $\mathbf{Q}$ or, equivalently, for the isometric distortions $\{d_{\mathrm{iso}}^{(0)}(s_i, t_j)\}$ defined in the previous section.

6.1 Spectral Embedding

We start by computing the geodesic distances between all pairs of bases by running the Dijkstra shortest paths

algorithm from each base vertex. These pairwise distances, when exposed to an exponential kernel, form a geodesic affinity matrix, $A_{ij} = \exp(-g^2(i, j)/2\sigma^2)$, for each of the base vertex sets S and T, with $g(.,.)$ being the geodesic distance between two points on a given surface. We set the kernel width σ to be half of the maximum geodesic distance over the surface. Note that computation of the geodesic affinity matrix does not bring any additional load to our overall correspondence algorithm due to the Dijkstra framework that is already prepared for the COES sampling procedure.

Based on the computed geodesic affinity matrices, we transform each of the base vertex sets into an M-dimensional spectral domain by means of the classical MDS algorithm to gain invariance against rigid transformations as well as shape bending. The classical MDS, introduced by Gower [38], essentially uses the M leading eigenvectors of the affinity matrix to obtain spectral embedding. We scale each of these eigenvectors with the square root of the corresponding eigenvalue (ignoring the first eigenvector since it becomes constant due to normalization) as suggested in [10]. The scaled eigenvectors provide us with an M-dimensional spectral embedding of the base vertices, which we will denote by $\hat{S}$ and $\hat{T}$, respectively, for the source and the target.

6.2 Alignment

The geodesic distances between base vertices in the original 3D space approximately correspond to L_2 distances between their M-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 which tests the 2^M different possible embeddings for the best alignment. We measure the alignment of each such embedding $\hat{S}_m$ with a fixed $\hat{T}$ by means of the cost $C_m = \sum_{i=1}^{|\hat{S}_m|}(\|\hat{s}_{i,m} - \hat{t}_i\|_2)$, i.e., the sum of L_2 distances between mutually closest points in the spectral domain, where each term in the summation is the distance between some $\hat{s}_{i,m}$ and the target $\hat{t}_i$ which is closest to it. The embedding $\hat{S}_m$ producing the minimum cost C_m aligns best with $\hat{T}$. This alignment operation is visualized in Fig. 3 for $M = 3$.

6.3 Isometric Distortion

For each possible (s_i, t_j) pair, the isometric distortion $d_{\text{iso}}(s_i, t_j)$ in (2) can be approximated by the L_2 distance between the embedded coordinates $\hat{s}_i$ and $\hat{t}_j$, which provides us with the initial estimate that we seek:

$$d_{\text{iso}}^{(0)}(s_i, t_j) = \|\hat{s}_i - \hat{t}_j\|_2. \tag{10}$$

7 Optimization

In this section, we describe the M-step of our EM algorithm, which involves minimization of the isometric distortion as given in (9). We solve this optimization problem in two steps. We first find a one-to-one mapping from S to T which is close to the global optimum via bipartite perfect matching, and then in the second step, we refine it locally via a greedy optimization algorithm that relaxes the injectivity constraint and thereby generates a many-to-one
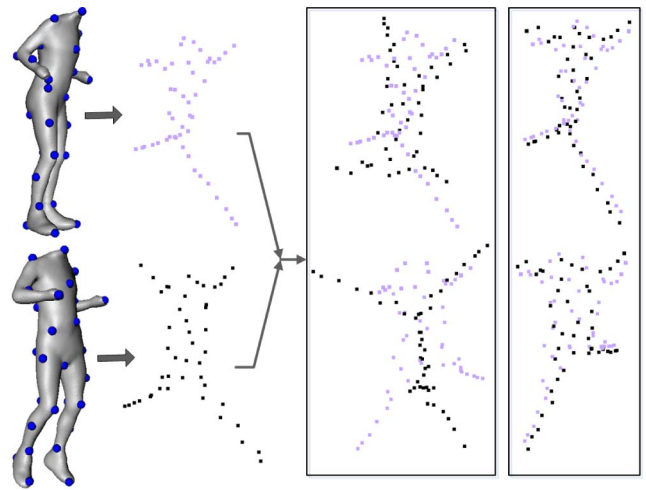

Fig. 3. Two shapes along with their spectral embeddings (left), the alignments obtained using an arbitrary configuration of the eigenvectors (left box), and using the best aligned embedding (right box). The boxes display two different views for visual convenience.

mapping. Both optimization steps are carried out in the original 3D euclidean space.

7.1 Bipartite Graph Matching

At each iteration k of the EM algorithm, we create a complete bipartite graph G on which the minimum-weight perfect matching is sought. The base vertices S and 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 $w_{ij} = d_{\text{iso}}^{(k)}(s_i, t_j)$, specifying the isometric cost of matching s_i with t_j. Since the cardinalities of the disjoint sets must match for a perfect matching, if $|S| \neq |T|$, we introduce virtual vertices 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 possible deviations from isometry. Hence, at the end of the optimization process, some base vertices in T may be left unassigned (we designate source and target so that $|S| \leq |T|$). We employ the efficient Blossom V algorithm [39] to solve the bipartite graph matching problem, hence the minimization in (9). Bipartite perfect matching results in a one-to-one (but not necessarily onto) mapping from S to T, which we denote by $\S_0^{(k)}$.

7.2 Greedy Optimization

The greedy optimization starts with the one-to-one mapping found via bipartite perfect matching of the first step, $\S = \S_0^{(k)}$, and iteratively improves $\S$ by relaxing the injectivity constraint imposed on it. The algorithm traverses the current $\S$ and each time replaces the current pair (s_i, t_j) with some (s_i, t_n) provided that this replacement decreases the isometric distortion. The accumulation of these greedy decisions, each of which considers a local improvement, eventually leads to a local optimum on D_{iso} as we retraverse $\S$ until convergence, i.e., until D_{iso} no longer improves.

For replacement of (s_i, t_j), the algorithm considers a small set of candidates t_m. This candidate set, denoted by T', is formed by the target base vertices (plus their base neighbors) that have been matched with the base vertices in
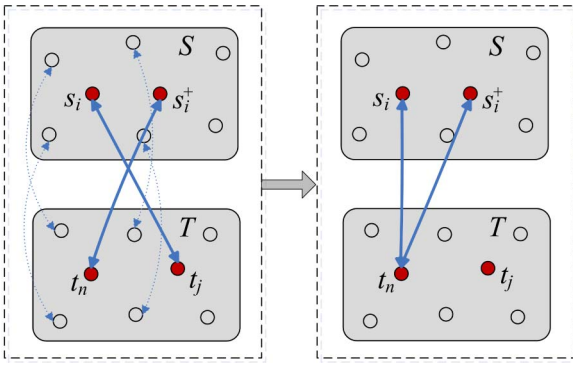
Fig. 4. Illustration of a local improvement by the greedy algorithm. Suppose each base vertex (filled circles) has four neighbors (empty circles) and (s_i, t_n) is the ground-truth correspondence. When $(s_i, t_j) \in \S$ is in process, the target base vertices matched with the base neighbors of s_i (pointed to by the dashed arrows) are tested (the base neighbors of these targets are also tested, though this is not illustrated in the figure). Since the target base t_n with minimum distortion is different than t_j in this case, the correspondence pair (s_i, t_j) is replaced with (s_i, t_n) after checking $d_{\mathrm{iso}}(s_i, t_n)$ versus $d_{\mathrm{iso}}(s_i, t_j)$. Note that the pair (s_i^+, t_n) is also considered next for its possible replacement with (s_i^+, t_j), though this is not illustrated in the figure.

the closed neighborhood of s_i. For each such t_m, $d_{\mathrm{iso}}(s_i, t_m)$ is computed, and (s_i, t_j) is replaced with the pair (s_i, t_n), yielding the minimum distortion, provided also that this replacement brings in some improvement over the current one, i.e., if $d_{\mathrm{iso}}(s_i, t_n) < d_{\mathrm{iso}}(s_i, t_j)$. In addition to this potential replacement concerning s_i, we also consider the current match s_i^+ of the selected t_n by replacing (s_i^+, t_n) with (s_i^+, t_j) if $d_{\mathrm{iso}}(s_i^+, t_j) < d_{\mathrm{iso}}(s_i^+, t_n)$. Note that the function d_{iso} is evaluated each time using the current correspondence list $\S$. The greedy optimization algorithm is illustrated on an example in Fig. 4 and given in pseudocode in Fig. 5.

The greedy algorithm relies on the assumption that the initial correspondence found in the first step is reasonably good, which is so thanks to the spectral matching that serves as a good starting point. This assumes that for most of the samples on the source shape, at least a base neighbor is initially matched close to its optimum on the target. For the remaining samples for which this assumption does not hold, EM iterations are expected to resolve the problem as described next.

7.3 EM Iterations

At each iteration k of the EM algorithm, the M-step first generates a one-to-one mapping from S to T via bipartite perfect matching. This mapping, that we denote by $\S_0^{(k)}$, is then iteratively refined by the greedy algorithm so as to produce a many-to-one mapping, denoted by $\S^{(k)}$. At the E-step of the next iteration, the bipartite matching process uses $\S^{(k)}$ to compute estimates for isometric distortion via (5), and thereby generates a one-to-one mapping $\S^{(k+1)}$ in a globally consistent manner. The final output of the EM algorithm at convergence, $\S^*$, can hence be one-to-one or many-to-one upon choice, depending on at which step the algorithm is stopped. We will provide experimental results for both cases in Section 9.

8   COMPUTATIONAL COMPLEXITY

Let V denote the number of vertices in the original mesh (source or target, whichever has more vertices) and N the

```
Input: §₀⁽ᵏ⁾, one-to-one
Output: §⁽ᵏ⁾, many-to-one
§ = §₀⁽ᵏ⁾
For each pair (sᵢ, tⱼ) ∈ §
    T' = GetCandidateTargets(sᵢ, tⱼ)
    minD = ∞
    For each tₘ ∈ T'
        If d_iso(sᵢ, tₘ) < minD
            tₙ = tₘ
            minD = d_iso(sᵢ, tₘ)
    If d_iso(sᵢ, tₙ) < d_iso(sᵢ, tⱼ)
        (sᵢ, tₙ) replaces (sᵢ, tⱼ) in §
    Let sᵢ⁺ be the match of tₙ, i.e., (sᵢ⁺, tₙ) ∈ §
    If d_iso(sᵢ⁺, tⱼ) < d_iso(sᵢ⁺, tₙ)
        (sᵢ⁺, tⱼ) replaces (sᵢ⁺, tₙ) in §
Until convergence (until D_iso no longer improves)
§⁽ᵏ⁾ = §

GetCandidateTargets(Pair (sᵢ, tⱼ) ∈ §)
T' = ∅
For each sₚ ∈ W(sᵢ) where W(sᵢ) is the closed
                          base neighborhood of sᵢ
    T' = T' ∪ {W(tᵣ) | (sₚ, tᵣ) ∈ §}
Return T'
```

Fig. 5. The greedy optimization algorithm.

number of base vertices sampled on the target ($N = |T|$, recalling that $|S| \leq |T|$). The COES sampling process involves curvature-based sorting with $O(V \log V)$ complexity, followed by the actual base vertex sampling procedure in again $O(V \log V)$ time since the shortcut Dijkstra shortest paths for each base sum up to one single Dijkstra algorithm spanning V vertices. The geodesic affinity matrix is computed in $O(NV \log V)$ time. The eigenanalysis for embedding into M-dimensional spectral domain is $O(N^2)$, followed by $O(2^M N^2)$ operations for alignment, where $M \leq 6$. As for the EM algorithm, the greedy optimization demands $O(N^2)$ time since the computation of d_{iso} can be performed in linear time for each pair in $\S$. The bipartite matching part of the M-step employs the Blossom V algorithm with $O(N^2 \log N)$ complexity for a given $N \times N$ cost matrix which is created in the E-step by computing a d_{iso} value for each entry in linear time. The EM framework hence demands $O(N^3)$ work per iteration until D_{iso} converges, which takes no more than five iterations in our experiments. Under the valid assumption of $N \ll V$, the overall complexity is then given as $O(NV \log V)$, which can be regarded as quite efficient when the other methods in the shape correspondence literature are considered. For example, this is almost equivalent to the computational complexity of the method presented in [10], but better than GMDS [24] with $O(V^2 \log V)$ complexity, than the method in [15] with $O(N^4 \log N + V^2 \log V)$ complexity, and than the method in [30], which involves combinatorial tree traversals and repeated shape deformations.

9   EXPERIMENTS

9.1   Data Sets

We have conducted experiments on four different types of datasets. The first consists of two mesh sequences, *Jumping Man* [40] and *Dancing Man* [41], both originally reconstructed
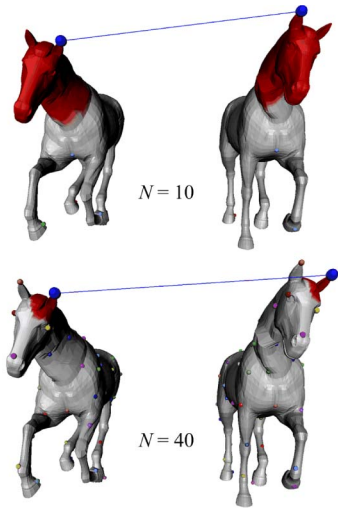
Fig. 6. Two correspondence results on a pair of meshes from the *Horse Gallop* sequence, with $N = 10$ (top) and with $N = 10$ (bottom). The horse has two critical vertices on the tips of the ears, sampled alternately on the source and on the target for $N = 10$. Hence, it is impossible to obtain the ground-truth left-to-left (or right-to-right) ear match at this resolution. The region painted (in red) on each shape represents the surface patch with radius r around the sampled (blue) base vertex. The mismatch problem is resolved when the number of base vertices to be matched is increased to $N = 40$.

from real scenes and each representing the real motion of a human actor. The second data set is the *Horse Gallop*, which is a computer generated synthetical mesh sequence [4]. The original meshes of these sequences are all uniform and given at high resolution with fixed connectivity; hence we have the ground-truth dense correspondences in all three cases. The third data set is the *Ballerina* from the 3D segmentation benchmark [42], which contains five different poses of the same human model, each represented with a mesh model of arbitrary connectivity. The last data set is the Nonrigid World shape database [18], which contains various animal and human mesh models with different poses, where each object has approximately 3K vertices with arbitrary connectivity.

9.2 Evaluation Metrics

We measure the performance of our shape correspondence scheme in terms of deviation from ground truth as well as isometric distortion. To quantify isometric distortion, we use the average distortion measure D_{iso} defined in (1), and also a maximum distortion measure that we denote by $d_{\text{iso}}^{\dagger}$:

$$d_{\text{iso}}^{\dagger} = \max_{(s_i, t_j) \in \S} d_{\text{iso}}(s_i, t_j), \qquad (11)$$

where d_{iso} is the isometric distortion function given in (2). Similarly, we compute average and maximum ground-truth correspondence errors, respectively, by (whenever the ground-truth correspondence is available)

$$D_{\text{grd}} = \frac{1}{|\S|} \sum_{(s_i, t_j) \in \S} g(t_i, t_j), \qquad (12)$$

and

$$d_{\text{grd}}^{\dagger}(s_i, t_j) = \max_{(s_i, t_j) \in \S} g(t_i, t_j), \qquad (13)$$

TABLE 1
One-to-One Mapping Performance for Varying Number N of Base Vertices on *Horse Gallop* Sequence

N	$D_{\text{grd}}, d_{\text{grd}}^{\dagger}$	$D_{\text{iso}}, d_{\text{iso}}^{\dagger}$
40	$1.99r', 4.29r'$	$1.34r', 3.13r'$
80	$0.58r', 2.92r'$	$0.56r', 1.85r'$
160	$0.46r', 2.11r'$	$0.42r', 1.28r'$
320	$0.38r', 1.13r'$	$0.31r', 0.84r'$

r' *denotes the sampling distance for* $N = 80$.

where each (s_i, t_i) stands for a ground-truth correspondence pair.

9.3 Parameter Setting

The correspondence that we find between the sampled base vertices can be regarded as sparse, which can though be expanded to a dense correspondence via cross-parametriza-tion or some other form of interpolation techniques [2]. However, it is always an issue to decide on the optimal sampling that would render the dense matching phase as efficient and robust as possible. The related sampling distance parameter r of our algorithm is the only parameter that we set manually. Our current strategy to set this parameter is to select a sufficiently small value so that all the critical vertices of a given mesh for typical shapes are included by our COES sampling procedure.

In Fig. 6, we demonstrate, on a pair of shapes from the Horse Gallop sequence, a degenerate case when the number of base vertices is significantly low, e.g., $N = 10$. The problems due to such degenerate cases are usually fixed by increasing the base resolution, e.g., to $N = 40$, hence by a decrement on the sampling distance r. The smaller the parameter r, the less the deviation from the ground truth during base vertex sampling, and hence the more accurate the correspondence obtained, as also verified by the quantitative results given in Table 1. In the table, we observe that the isometric distortion and the correspon-dence error both decrease as the sampling distance gets smaller, at the expense of some execution time: 1.5, 2.1, 4.0, and 10.5 seconds for $N \approx 40, 80, 160,$ and 320, respectively, (recall the $O(NV \log V)$ complexity of our algorithm).

9.4 Results

We display some visual examples from the computed correspondences through Figs. 7, 8, 9, 10, 11, and 12, and provide quantitative results in Tables 1 and 2. In all our experiments, the dimension of the spectral domain is set to be $K = 6$, and the number of base vertices is about 80 unless stated otherwise.

In Fig. 7, we display one-to-one matching results on a shape pair from the Horse Gallop sequence for varying resolutions and hence for varying N. We observe that, while the obtained correspondences satisfactorily match the ground truth at all resolution levels, the matching precision increases as the number N of base vertices increases. In Figs. 8 and 9, we provide the visual correspondence results obtained on sample shape pairs, respectively, from the Dancing Man and Ballerina sequences. Fig. 10 displays the correspondence results across two mesh sequences, i.e., on two shapes representing two different humans, which are hence only nearly isometric. We observe that the one-to-one
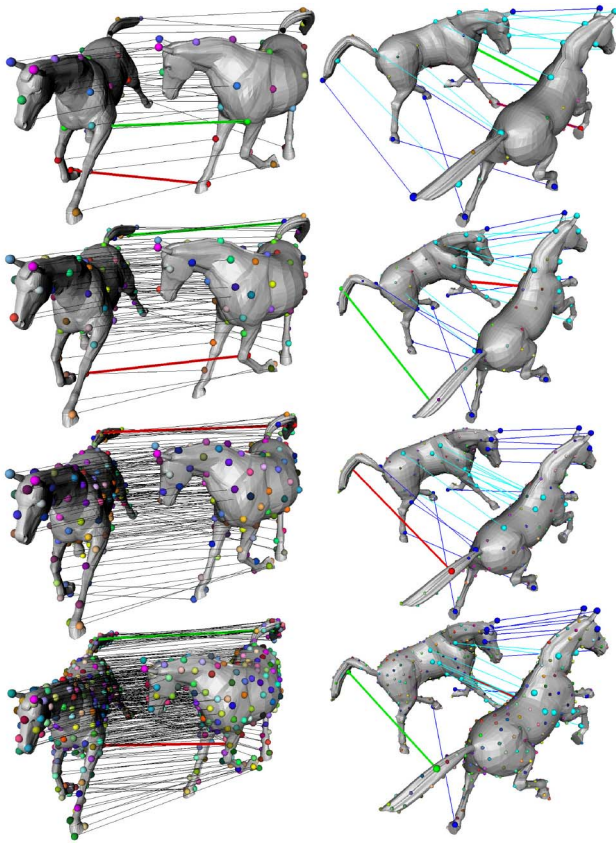
Fig. 7. (From top to bottom) Correspondence results on *Horse Gallop* with $N = 40, 80, 160, 320$, where each row displays the shape pair from two different views. Bold green and red lines indicate the worst matches with respect to ground-truth error and isometric distortion, respectively. Note that the two worst matches coincide for $N = 160$, where red overwrites green. (Left) All available correspondence pairs, each indicated with a line segment drawn between two spheres of the same color at both ends, (right) best-10 matches with respect to isometric distortion and 10 matches for top-10 vertices of the curvature-sorted vertex list, highlighted in cyan and blue, respectively (cyan may overwrite blue). **This whole color representation scheme applies to the subsequent figures as well.**
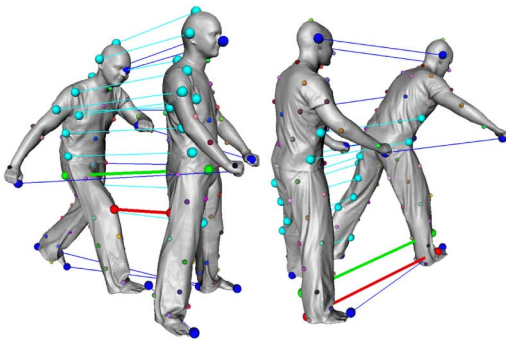

Fig. 8. The final one-to-one mappings obtained for two different shape pairs from *Dancing Man*.

mapping obtained is very accurate despite the missing head of the Jumping Man. In all these figures, two base vertices with the same color indicate a correspondence pair, whereas the worst matchings with respect to isometric distortion and ground-truth correspondence error (whenever available) are highlighted with bold red and green lines, respectively. Note also that some target bases may not be drawn in the figures if not matched with any source base vertex.
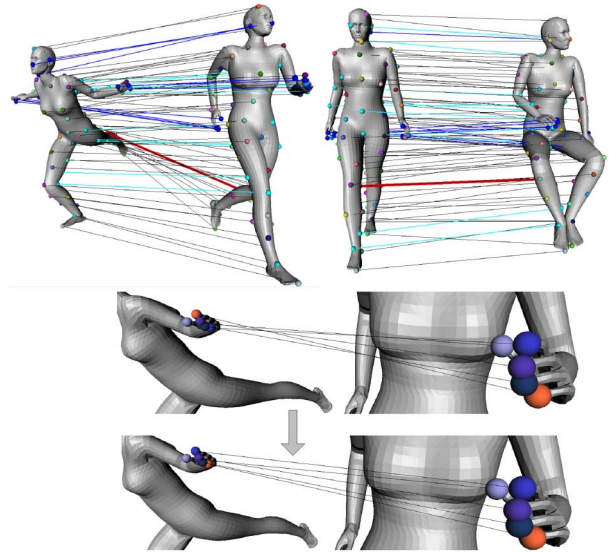

Fig. 9. The final one-to-one mappings obtained for two different *Ballerina* pairs (top row). Note that $N \approx 80$ base vertices include all high-curvature points of a hand, which are then accurately matched (bottom row, left pair shown in zoom). Observe how the initial mismatches on the fingers ($\S_0^{(0)}$, middle row) are healed by our EM-based optimization framework ($\S_0^*$, bottom row).
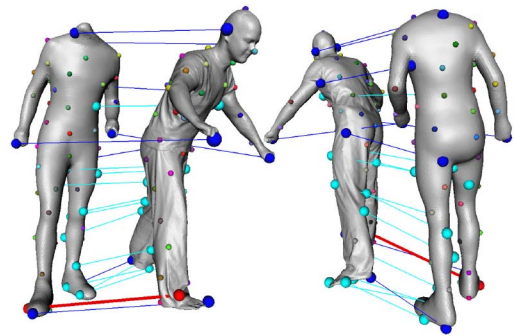

Fig. 10. The final one-to-one mapping obtained for a *Jumping Man-Dancing Man* hybrid pair, displayed from two different views.
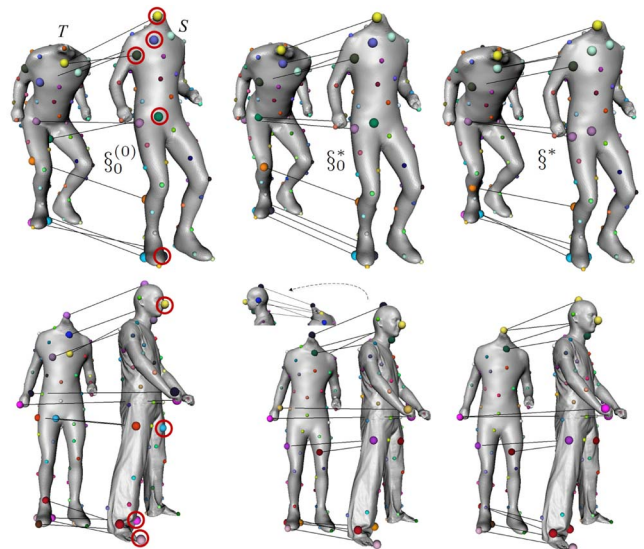

Fig. 11. (Left) The initial one-to-one mapping $\S_0^{(0)}$, (middle) the final one-to-one mapping $\S_0^*$, and (right) the final many-to-one mapping $\S^*$, for an isometric (top) and a nearly isometric pair of shapes (bottom). Some base vertices from the source shapes for which the final correspondences improve significantly as compared to the initial are each marked with a red circle.
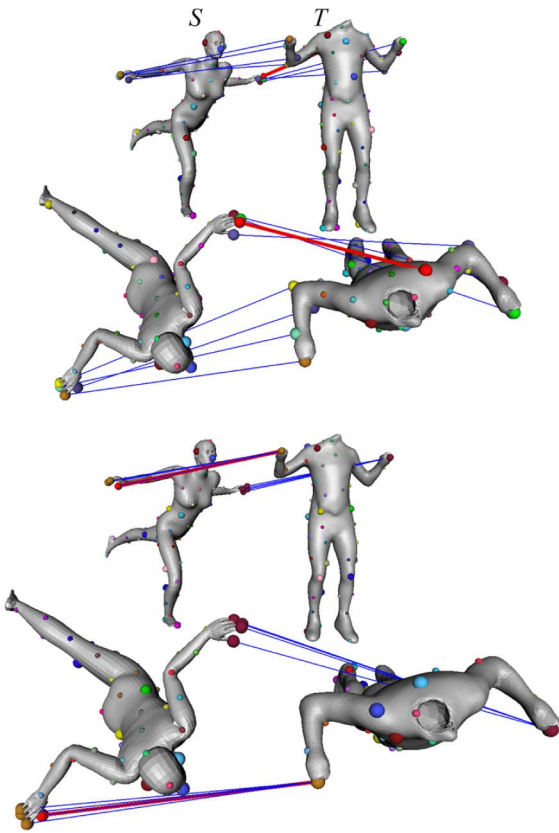
Fig. 12. (Top) The initial one-to-one mapping $\S_0^{(0)}$ and (bottom) the final many-to-one mapping $\S^*$ obtained for a *Jumping Man-Ballerina* hybrid pair, displayed from two different views.

In Fig. 11, we demonstrate the outputs obtained at different stages of our correspondence algorithm on a Jumping Man pair and on a hybrid Jumping Man-Dancing Man pair. For each pair, we display the initial one-to-one mapping $\S_0^{(0)}$ (the output of the first bipartite matching process), the final one-to-one mapping $\S_0^*$ (the output of the last bipartite matching process at convergence), and the final many-to-one mapping $\S^*$ (the output of the last greedy optimization process at convergence). Note the improvements obtained through the different stages of the correspondence algorithm. Fig. 12 demonstrates these improvements on a particular example, i.e., on a Ballerina-Jumping Man hybrid pair which contains local but severe nonisometries due to the fingers of the Ballerina

shape (note that the Jumping Man model does not have any fingers). Hence the Ballerina shape has more base vertices sampled on the hands as compared to the Jumping Man, and since the initial one-to-one mapping forces every source base vertex to match a different target base, the result severely deviates from the ground truth. Our final many-to-one mapping however correctly assigns all base vertices of the fingers to the single base vertex sampled on the corresponding hand of the Jumping Man.

Table 2 provides, on different datasets, the values that the performance measures take at different stages of our algorithm, for $\S_0^{(0)}$, $\S_0^*$, and $\S^*$ (see also Fig. 11). The average-based performance measures, D_{grd} and D_{iso}, and the maximum-based measures, $d_{\mathrm{grd}}^\dagger$ and $d_{\mathrm{iso}}^\dagger$, are each computed over 10 different runs of the algorithm on 10 different pairs, except for one pair for *Dog-Wolf* from the Nonrigid World database and five pairs for *Ballerina*. Each run matches two spatially apart poses of the articulated object in the corresponding sequence. We note that, in some cases, especially at low resolution matchings when N is small, the spectral alignment procedure (Section 6.2) may fail to resolve the sign ambiguities of the eigenvectors due to the symmetries in the object shapes. The results given in Table 2 exclude such cases in order not to artificially burst the ground-truth correspondence errors. All the performance measures are provided as a factor of the sampling distance r for better interpretation of the errors.

The results given in Table 2 can also be assessed by considering the input quality, i.e., the isometric deviations inherent in the original data sets. To this effect, we have measured the isometric distortion within and across our test sequences by computing a D_{iso} value, given the manual match of 10 critical vertices on each shape pair, and obtained the following distortion values (from the most isometric to the least): 0.030 for Jumping Man, 0.034 for Ballerina, 0.036 for Dancing Man, 0.039 for Horse Gallop, 0.046 for Jumping-Dancing pair, 0.047 for Jumping Man-Ballerina pair, and 0.060 for the Dog-Wolf pair. Each of these values can be interpreted as the isometric distortion of the semantic mapping that can be established between the corresponding shapes, i.e., as a measure of how isometric the given shapes to be matched are. In Table 2, we observe that the isometric distortion inherent in a dataset puts roughly an upper bound to the performance; hence a lower bound to the D_{iso} value can be achieved by our method. We

TABLE 2
Quantitative Performance Analysis of Our Shape Correspondence Method in Comparison with the Spectral Method of [10]

Pair	Initial one-to-one, $\S_0^{(0)}$		Final one-to-one, $\S_0^*$		Final many-to-one, $\S^*$		Many-to-one of [10]	
	$D_{\mathrm{grd}}, d_{\mathrm{grd}}^\dagger$	$D_{\mathrm{iso}}, d_{\mathrm{iso}}^\dagger$	$D_{\mathrm{grd}}, d_{\mathrm{grd}}^\dagger$	$D_{\mathrm{iso}}, d_{\mathrm{iso}}^\dagger$	$D_{\mathrm{grd}}, d_{\mathrm{grd}}^\dagger$	$D_{\mathrm{iso}}, d_{\mathrm{iso}}^\dagger$	$D_{\mathrm{grd}}, d_{\mathrm{grd}}^\dagger$	$D_{\mathrm{iso}}, d_{\mathrm{iso}}^\dagger$
Horse Gallop	$0.73r, 7.71r$	$0.64r, 2.75r$	$0.58r, 2.92r$	$0.56r, 1.85r$	$\mathbf{0.35r}, \mathbf{2.17r}$	$\mathbf{0.29r}, \mathbf{0.87r}$	$0.55r, \mathbf{2.17r}$	$0.50r, 1.49r$
Jumping Man	$0.57r, 2.45r$	$0.49r, 0.91r$	$0.52r, 2.45r$	$0.28r, 0.54r$	$0.40r, \mathbf{1.92r}$	$0.24r, 0.47r$	$0.52r, 3.98r$	$0.38r, 0.68r$
Dancing Man	$0.58r, 7.48r$	$0.48r, 2.55r$	$0.44r, 1.88r$	$0.28r, 0.82r$	$\mathbf{0.38r}, \mathbf{1.22r}$	$\mathbf{0.23r}, \mathbf{0.60r}$	$0.44r, 1.41r$	$0.27r, 0.77r$
Ballerina	n/a	$0.50r, 1.46r$	n/a	$0.43r, 1.02r$	n/a	$\mathbf{0.26r}, \mathbf{0.63r}$	n/a	$0.46r, 1.25r$
Jumping-Dancing	n/a	$0.51r, 1.03r$	n/a	$0.46r, 1.12r$	n/a	$\mathbf{0.37r}, \mathbf{0.59r}$	n/a	$0.47r, 1.03r$
Jumping-Ballerina	n/a	$1.05r, 3.32r$	n/a	$0.67r, 2.00r$	n/a	$\mathbf{0.46r}, \mathbf{1.14r}$	n/a	$0.64r, 1.50r$
Dog-Wolf	n/a	$0.73r, 2.38r$	n/a	$0.69r, 1.62r$	n/a	$\mathbf{0.51r}, \mathbf{0.79r}$	n/a	$0.65r, 1.20r$

The best (lowest) performance values on each row are given in bold. In all cases the sampling distance r is approximately 0.07 (the maximum geodesic distance on the mesh is taken as 1.0), which yields about $N = 80$ base vertices.
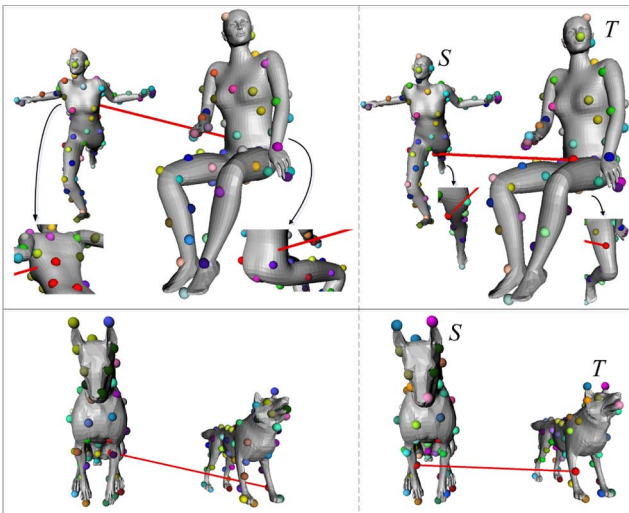
Fig. 13. The many-to-one mappings on the *Dog-Wolf* hybrid pair from the Nonrigid World database (bottom) and on a *Ballerina* pair (top); obtained with our method (right) and with the spectral method of [10] (left).

see that, for each data set, our one-to-one mapping performance achieves a D_{iso} value which is close to this bound, and even exceeds it in the many-to-one setting.

The execution time of our implementation is mainly dominated by the number of vertices in the original meshes due to geodesic distance computation. On a 6 GB 2.53 GHz 64-bit workstation, the overall algorithm takes, for $N \approx 80$ base vertices, 2.1, 5.6, 8.0, 2.0, 6.0, 3.5, and 1.5 seconds, respectively, on Horse Gallop ($V = 9K$), Jumping Man ($V = 16K$), Dancing Man ($V = 20K$), Ballerina ($V = 6K$), Jumping Man-Dancing Man pair ($V = 20K$), Jumping Man-Ballerina pair ($V = 16K$), and Dog-Wolf pair ($V = 3.4K$). We note that the portion of the execution time that belongs to geodesic computation varies between 77 and 90 percent on different data sets. This suggests that our shape correspondence method can be made even more efficient by using faster geodesics computation algorithms such as the one in [43].

9.4.1 Comparison with [10]

The spectral method of Jain and Zhang [10] is one of the state-of-the-art techniques for embedding-based nonrigid shape correspondence. The method generates as output a many-to-one mapping from source to target. In Fig. 13, we compare our many-to-one mapping results with the many-to-one mappings obtained via this spectral method on two distinct cases: on an isometric shape pair from the Ballerina sequence and on the nonisometric Dog-Wolf pair from the Nonrigid World database, which contains two shapes with similar structure but distorted isometry. We observe that our method outperforms the spectral method significantly in both cases, the worst matchings being significantly better as well as the correspondences being generally more accurate. While our method correctly matches all the shape extremities, the tips of the left ears on the Dog-Wolf pair are not, for example, correctly matched with the spectral method. We have obtained the results of the spectral method by using the Matlab code made publicly available by the Jain and Zhang [10]. In the comparison tests, we have

run their code on the same sets of base vertices (about 80) and the same geodesic affinity matrices that we have used for our algorithm.

The last column of Table 2 presents the average and maximum-based performance measures for the spectral method in comparison to ours. These measures are computed over the shape pairs for which none of the two methods results in a symmetric flip. We observe that our many-to-one mapping $\S^*$ outperforms the results of the spectral method significantly in all settings for all data sets. Another interesting observation is that our one-to-one mapping $\S_0^*$ is performance-wise almost on a par (even slightly better in some occasions) with the many-to-one output of the spectral method.

We note that, although the complexity of the spectral method as presented in [10] is $O(V^2 \log V)$ against our $O(NV \log V)$ complexity, it can be reduced to $O(NV \log V)$ by using the Nyström method, which, however, introduces further approximation to the euclidean embedding process [44].

9.4.2 Comparison with GMDS

We have conducted experiments on the Nonrigid World shape database to compare our method with the GMDS method of [24]. For GMDS, we have used the publicly available Matlab code, which uses multiresolution optimization [18]. The part of the Nonrigid World database that we have used contains mesh models of 9 cats, 11 dogs, 3 wolves, 17 horses, 24 female figures, and 2 different male figures containing 15 and 20 poses, respectively. We have excluded some shape classes from the experiments; the gorillas having disconnected shape components, the shark, which has only one mesh model in its class, and the lions, on which the Matlab code for GMDS has crashed.

We have evaluated the performances in three different categories: 1) within human models, 2) within animal models, and 3) across human models. In the experiments belonging to the first and second categories, for each model we pick a random model from the same class and then compute the correspondence in between, e.g., we match each dog to another dog. In the third category, for each human object, we pick a random model from each different human class and then compute the correspondence. We note that both methods can result in symmetric flips, though the symmetry problem with the GMDS method is more severe due to the coarse initialization step, which is based on only eight samples in the available implementation. In our experiments, while our method matched about 40 percent of the shape pairs over the three categories as free of symmetric flip errors, this ratio was only 27 percent in the case of GMDS. We note that almost all the shapes in the Nonrigid World database have intrinsic symmetries (though not perfect), and the number of symmetric flips tend to decrease with our method as we increase the number of samples, which was about 80 in these experiments. In the comparison tests, we included only the pairs that could be matched as free of symmetry problems by both methods.

Like the spectral method, the GMDS method generates as output a many-to-one mapping from the source shape to the target. Hence, also in this case, for comparison we have used our many-to-one mapping results. The GMDS method

TABLE 3
Quantitative Performance of Our Shape Correspondence Method in Comparison with the GMDS Method on the Nonrigid World Data Set

	Many-to-one (GMDS)	Many-to-one (our method)
Category	$D_{\mathrm{iso}}, d_{\mathrm{iso}}^{\dagger}$	$D_{\mathrm{iso}}, d_{\mathrm{iso}}^{\dagger}$
Across humans	$0.625r, 1.701r$	$\mathbf{0.407}r, \mathbf{1.158}r$
Within humans	$\mathbf{0.314}r, 1.139r$	$0.379r, \mathbf{0.885}r$
Within animals	$\mathbf{0.358}r, \mathbf{1.036}r$	$0.405r, 1.231r$

has its own sampling procedure, and the publicly available code can be set so as to generate the same number of vertices that our method produces. Thus, in our experiments for comparison with GMDS, the sampled vertices on a given mesh are different but the same in number.

We present the quantitative results of the comparison tests in Table 3. We observe that our method significantly outperforms GMDS across humans category, and is on a par with it for the isometric categories (recall also that the complexity of GMDS is $O(V^2 \log V)$ against our $O(NV \log V)$ complexity). Note that the ground-truth correspondence information is not available with the Nonrigid World database. We have excluded across animal models as a fourth possible category from our experiments for comparison with GMDS since in that case some shape pairs to be matched would have severe global nonisometries that both methods, which are essentially isometric, would fail to handle. The three categories that we have experimented on correspond to (nearly) isometric instances of the correspondence problem with the following isometric distortion values: 0.035 for across humans, 0.028 for within humans, and 0.030 for within animals (or equivalently, $0.500r$, $0.400r$, and $0.428r$, respectively, where $r = 0.07$). Recall however that the Dog-Wolf pair for which we have presented visual and quantitative results in Section 9.4.1 is from the Nonrigid World database.

We visually demonstrate the comparison tests on three examples in Fig. 14. In the figure, we see that the vertices matched with GMDS are not as evenly distributed as they are in our case, resulting in clustered correspondences. Note, for instance, in the top row, that the samples clustered on the inner part of the left leg of one model have been matched to the samples clustered on the outer part of the left leg of the other model. This is mainly due to the fact that the GMDS algorithm is a gradient-based iterative optimization process that produces subvertex matchings which do not necessarily coincide with the initial sampling. We also note that, while computing the resulting isometric distortion value for a GMDS output, we round subvertex coordinates to their nearest vertices on the surface mesh. The effect of this rounding process to the computed distortion values is, however, negligible since the mesh models in the Nonrigid World database are almost uniform and at relatively high resolution. We also observe from the visual comparison that our method can match the salient points of a shape more successfully than GMDS (e.g., ear tips of the cats), thanks to our saliency-based sampling.
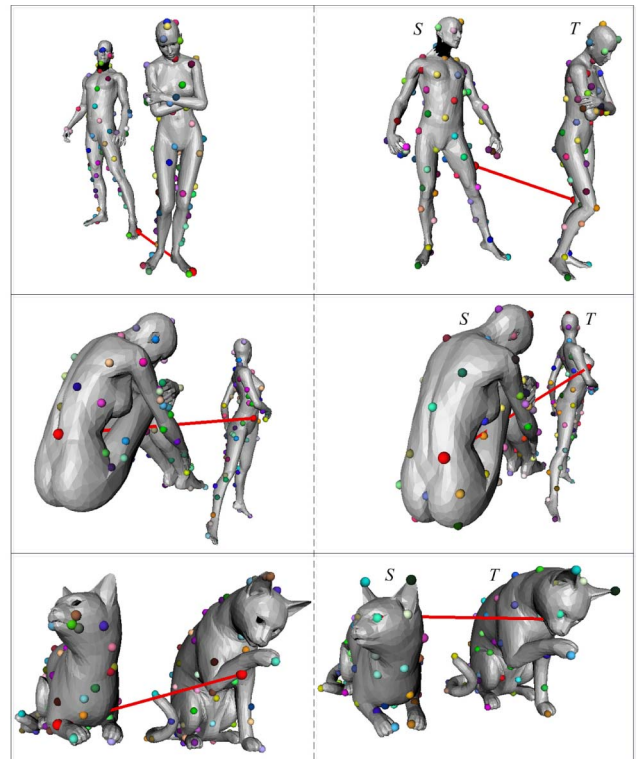

Fig. 14. Many-to-one mappings obtained with GMDS (left) and with our method (right) for sample pairs selected from across humans (top), within humans (middle), and within animals (bottom) categories of the Nonrigid World dataset. Our worst match between the cat models (red line) falls inside the skull and hence is not visible. The vertices matched with our method are always uniformly distributed and at prominent regions such as finger or ear tips, unlike GMDS matches. Note that the correspondences tend to get clustered with GMDS in all three cases.

10 CONCLUSION

We have proposed an isometric shape correspondence method that minimizes the isometric distortion in the original 3D euclidean space by using EM algorithm. We have conducted experiments on various data sets, and our findings can be summarized as follows:

- Our method performs well not only on isometric shapes, but also on pairs of shapes which are nearly isometric, such as mesh representations of two different humans or animals, i.e., for shapes of the objects which are different but semantically and/or structurally similar.
- The COES sampling algorithm that we propose solves the joint sampling problem by evenly sampling high-curvature vertices from both shapes using a sufficiently small sampling distance.
- When the shapes to be matched are given in terms of evenly sampled surface points (base vertices), any isometric correspondence technique in the literature can be used to initialize our EM-based method. In that sense, our method can be used to further improve the results of any embedding-based isometric method, such as [10], that possibly suffers from approximations and/or ambiguities in the embedding domain.
- Our method generates two optimal mappings at a time from one shape to the other, one injective and

the other many-to-one, the former having generally higher isometric distortion than the latter, especially in the presence of severe nonisometries between the shapes.

- Our method is computationally very efficient and can be made even more efficient by using faster algorithms for computation of geodesic distances.
- Our method can be applied to any pair of isometric shapes with arbitrary genus. However, it cannot handle the deformations which include topological changes that would modify geodesic distances drastically.

One limiting factor for our method is that it has been designed to work on isometric shapes, though the experiments show that our algorithm works well also for nearly isometric shapes. Yet, our algorithm may not produce good correspondence results when two shapes contain severe global nonisometries such as some of the hybrid animal shape pairs in the Nonrigid World database. Another limitation of our method is due to the problem of symmetric correspondences, which is inherent in all isometry-based correspondence methods. Hence, if two shapes to be matched include intrinsic symmetries, then our method may fail to find the correct matching and result in symmetric flips. To resolve the symmetry problem, explicit symmetry information is needed and one can resort to methods that can detect global intrinsic symmetries such as [19], [45]. We finally note that, as future work, we plan to extend our method to handle partially isometric shapes.

ACKNOWLEDGMENTS

REFERENCES

[1] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or, "A Survey on Shape Correspondence," *Computer Graphics Forum*, vol. 30, pp. 1681-1707, 2011.

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

[3] V. Kraevoy and A. Sheffer, "Cross-Parameterization and Compatible Remeshing of 3D Models," *ACM Trans. Graphics*, vol. 23, no. 3, pp. 861-869, 2004.

[4] R.W. Sumner and J. Popovic, "Deformation Transfer for Triangle Meshes," *ACM Trans. Graphics*, vol. 23, no. 3, pp. 399-405, 2004.

[5] D. Anguelov, P. Srinivasan, H. Pang, D. Koller, S. Thrun, and J. Davis, "The Correlated Correspondence Algorithm for Unsupervised Registration of Nonrigid Surfaces," *Neural Information Processing Systems*, vol. 37, pp. 33-40, 2004.

[6] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. Kunii, "Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes," *Proc. ACM Siggraph*, pp. 203-212, 2001.

[7] A. Elad and R. Kimmel, "On Bending Invariant Signatures for Surfaces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1285-1295, Oct. 2003.

[8] A. Zaharescu, E. Boyer, K. Varanasi, and R. Horaud, "Surface Feature Detection and Description with Applications to Mesh Matching," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009.

[9] B.C. Munsell, P. Dalal, and S. Wang, "Evaluating Shape Correspondence for Statistical Shape Analysis: A Benchmark Study," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 2023-2039, Nov. 2008.

[10] V. Jain and H. Zhang, "Robust 3D Shape Correspondence in the Spectral Domain," *Proc. IEEE Int'l Conf. Shape Modeling and Applications*, pp. 118-129, 2006.

[11] M. Carcassoni and E.R. Hancock, "Correspondence Matching with Modal Clusters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1609-1615, Dec. 2003.

[12] M. Leordeanu and M. Hebert, "A Spectral Technique for Correspondence Problems Using Pairwise Constraints," *Proc. IEEE Int'l Conf. Computer Vision*, pp. 1482-1489, 2005.

[13] S. Wuhrer, C. Shu, and P. Bose, "Posture Invariant Correspondence of Triangular Meshes in Shape Space," *Proc. Int'l Conf. 3D Digital Imaging and Modeling*, 2009.

[14] D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer, "Articulated Shape Matching Using Laplacian Eigenfunctions and Unsupervised Point Registration," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1-8, 2008.

[15] Y. Lipman and T. Funkhouser, "Mobius Voting for Surface Correspondence," *ACM Trans. Graphics*, vol. 28, no. 3, p. 72, 2009.

[16] Y. Sahillioğlu and Y. Yemez, "3D Shape Correspondence by Isometry-Driven Greedy Optimization," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 453-458, 2010.

[17] A.M. Bronstein, M.M. Bronstein, and R. Kimmel, *Numerical Geometry of Non-Rigid Shapes*. Springer, 2008.

[18] A.M. Bronstein, M.M. Bronstein, and R. Kimmel, "Efficient Computation of Isometry-Invariant Distances between Surfaces," *SIAM J. Scientific Computing*, vol. 28, no. 5, pp. 1812-1836, 2006.

[19] M. Ovsjanikov, J. Sun, and L. Guibas, "Global Intrinsic Symmetries of Shapes," *Computer Graphics Forum*, vol. 27, no. 5, pp. 1341-1348, 2007.

[20] M. Ovsjanikov, Q. Mérigot, F. Mémoli, and L. Guibas, "One Point Isometric Matching with the Heat Kernel," *Computer Graphics Forum*, vol. 29, no. 5, pp. 1555-1564, 2010.

[21] A. Sharma and R.P. Horaud, "Shape Matching Based on Diffusion Embedding and on Mutual Isometric Consistency," *Proc. Workshop Nonrigid Shape Analysis and Deformable Image Alignment*, 2010.

[22] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios, "Dense Non-Rigid Surface Registration Using High-Order Graph Matching," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2010.

[23] V. Kim, Y. Lipman, and T. Funkhouser, "Blended Intrinsic Maps," *ACM Trans. Graphics*, vol. 30, no. 4, Article 79, 2011.

[24] A.M. Bronstein, M.M. Bronstein, and R. Kimmel, "Generalized Multidimensional Scaling: A Framework for Isometry Invariant Partial Surface Matching," *Proc. Nat'l Academy of Sciences USA*, vol. 103, no. 5, pp. 1168-1172, 2006.

[25] H. Chui and A. Rangarajan, "A New Point Matching Algorithm for Non-Rigid Registration," *Computer Vision and Image Understanding*, vol. 89, pp. 114-141, 2003.

[26] L. Torresani, V. Kolmogorov, and C. Rother, "Feature Correspondence via Graph Matching: Models and Global Optimization," *Proc. 10th European Conf. Computer Vision*, pp. 596-609, 2008.

[27] T. Funkhouser and P. Shilane, "Partial Matching of 3D Shapes with Priority-Driven Search," *Proc. Symp. Geometry Processing*, 2006.

[28] A. Tevs, M. Bokeloh, M. Wand, A. Schilling, and H.-P. Seidel, "Isometric Registration of Ambiguous and Partial Data," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009.

[29] S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509-522, Apr. 2002.

[30] H. Zhang, A. Sheffer, D. Cohen-Or, Q. Zhou, O. van Kaick, and A. Tagliasacchil, "Deformation-Driven Shape Correspondence," *Proc. Symp. Geometry Processing*, pp. 1431-1439, 2008.

[31] Q. Huang, B. Adams, M. Wicke, and L. Guibas, "Non-Rigid Registration under Isometric Deformations," *Proc. Symp. Geometry Processing*, pp. 1149-1458, 2008.

[32] O.K.-C. Au, D. Cohen-Or, C.-L. Tai, H. Fu, and Y. Zheng, "Electors Voting for Fast Automatic Shape Correspondence," *Computer Graphics Forum*, vol. 29, no. 2, pp. 645-654, 2010.

[33] C. Wang, M.M. Bronstein, A.M. Bronstein, and N. Paragios, "Discrete Minimum Distortion Correspondence Problems for Non-Rigid Shape Matching," *Proc. Third Int'l Conf. Scale Space and Variational Methods in Computer Vision*, 2011.

[34] Y. Sahillioğlu and Y. Yemez, "Coarse-to-Fine Combinatorial Matching for Dense Isometric Shape Correspondence," *Computer Graphics Forum*, vol. 30, no. 5, pp. 1461-1470, 2011.

[35] M. Meyer, M. Desbrun, P. Schröder, and A. Barr, "Discrete Differential-Geometry Operators for Triangulated 2-Manifolds," *Proc. Int'l Workshop Visualization and Math.*, 2002.

[36] G. Peyré and L.D. Cohen, "Geodesic Remeshing Using Front Propagation," *Proc. IEEE Workshop Variational, Geometric and Level Set Methods,* pp. 33-40, 2003.

[37] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. Royal Statistical Soc.: Series B,* vol. 39, pp. 1-38, 1977.

[38] J.C. Gower, "Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis," *Biometrika,* vol. 53, pp. 325-338, 1966.

[39] V. Kolmogorov, "Blossom V: A New Implementation of a Minimum Cost Perfect Matching Algorithm," *Math. Programming Computation,* vol. 1, pp. 43-67, 2009.

[40] P. Sand, L. McMillan, and J. Popovic, "Continuous Capture of Skin Deformation," *Proc. Int'l Conf. Computer Graphics and Interactive Techniques,* 2003.

[41] E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.P. Seidel, and S. Thrun, "Performance Capture from Sparse Multi-View Video," *ACM Trans. Graphics,* vol. 27, no. 3, 2008.

[42] X. Chen, A. Golovinskiy, and T. Funkhouser, "A Benchmark for 3D Mesh Segmentation," *ACM Trans. Graphics,* vol. 28, 2009.

[43] O. Weber, Y.S. Devir, A.M. Bronstein, M.M. Bronstein, and R. Kimmel, "Parallel Algorithms for Approximation of Distance Maps on Parametric Surfaces," *ACM Trans. Graphics,* vol. 27, no. 4, pp. 1-16, 2008.

[44] R. Liu, V. Jain, and H. Zhang, "Subsampling for Efficient Spectral Mesh Processing," *Proc. 24th Int'l Conf. Advances in Computer Graphics ,* 2006.

[45] V.G. Kim, Y. Lipman, X. Chen, and T. Funkhouser, "Mobius Transformations for Global Intrinsic Symmetry Analysis," *Computer Graphics Forum,* vol. 29, no. 5, pp. 1555-1564, 2010.

Yusuf Sahillioğlu received the BS degree from Bilkent University, Ankara, in 2004, the first MS degree from Koç University, Istanbul, in 2006, and the second from the University of Florida in 2008, all in computer science. He is currently working toward the PhD degree on the topic of 3D shape correspondence at Koç University. His research interests include digital geometry processing, computer graphics, and computervision.
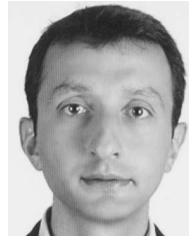
Yücel Yemez received the BS degree from Middle East Technical University, Ankara, in 1989, and the MS and PhD degrees from Boğaziçi University, Istanbul, respectively, in 1992 and 1997, all in electrical engineering. From 1997 to 2000, he was a postdoctoral researcher in the Image and Signal Processing Department of Télécom Paris (ENST). Currently, he is an associate professor in the Computer Engineering Department at Koç University, Istanbul. His research interests include various fields of computer vision and graphics.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.