



Coarse-to-Fine Isometric Shape Correspondence by Tracking Symmetric Flips

Y. Sahillioglu and Y. Yemez

Computer Engineering Department, Koç University, Istanbul, Turkey
{ysahillioglu, yyemez}@ku.edu.tr

Abstract

We address the symmetric flip problem that is inherent to multi-resolution isometric shape matching algorithms. To this effect, we extend our previous work which handles the dense isometric correspondence problem in the original 3D Euclidean space via coarse-to-fine combinatorial matching. The key idea is based on keeping track of all optimal solutions, which may be more than one due to symmetry especially at coarse levels, throughout denser levels of the shape matching process. We compare the resulting dense correspondence algorithm with state-of-the-art techniques over several 3D shape benchmark datasets. The experiments show that our method, which is fast and scalable, is performance-wise better than or on a par with the best performant algorithms existing in the literature for isometric (or nearly isometric) shape correspondence. Our key idea of tracking symmetric flips can be considered as a meta-approach that can be applied to other multi-resolution shape matching algorithms, as we also demonstrate by experiments.

Keywords: combinatorial matching, isometric dense shape correspondence, symmetric flips

ACM CCS: I.3.5 [Computer Graphics]: 3D Shape Correspondence—Coarse-to-fine sampling, combinatorial matching, isometric dense shape correspondence, symmetric flips

1. Introduction

Finding correspondences between shapes is a fundamental research problem in computer graphics and vision with various applications [vKZHC01]. Isometry is an important clue for shape correspondence; not only since most real world shape deformations are isometric (translation, rotation and bending), but also because semantically similar shapes have similar metric structures. Isometric shapes appear in many 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) [BBK08].

The shape correspondence problem can be thought of as the problem of finding maps between shapes, which are as isometric as possible, hence an optimization problem where a minimum distortion mapping is sought. The simplest way that guarantees to find the optimal mapping is to explicitly search the space of all maps, which is however computationally intractable due to factorial complexity. Hence, most of the isometric correspondence methods existing in the literature adhere to embedding in some way or other to achieve invariance against isometric deformations and

thereby to reduce the problem into a closest point matching problem. As a result, they all remain as approximate methods due to ambiguities and errors introduced by the embedding process itself [BBK08], where the embedding may be Euclidean [JZ06, CH03, EK03, WSB09, MHK*08, OSG07] or non-Euclidean [OMMG10, SH10, LF09, ZWW*10, DK11, KLF11, OBCS*12]. One exception is the GMDS (generalized multi-dimensional scaling) method proposed in [BBK06b, BBK06a]. The GMDS method eliminates the approximation error by embedding one shape onto the surface of the other, that however requires minimization of a non-convex function which is difficult and expensive to optimize. Another exception is the method presented in [SY12], which minimizes the isometric distortion in the original 3D Euclidean space by using an EM algorithm. Both methods, however, do not scale well to large meshes.

An important problem in shape correspondence is how to achieve dense correspondence, i.e. a matching between all vertices of two given high-resolution meshes. The main bottleneck in achieving dense correspondence is the computational complexity of the existing algorithms. Most embedding-based methods naturally support dense correspondence but their computational load is usually a limiting factor. To the best of our knowledge there exist only two



Figure 1: Our method vs. BIM. The yellow vertex on the face of the source human model is mapped to the hand of the target by BIM (left) and to the face by our method (right). Note also the unmatched target vertices (in grey colour) on the BIM result, which do not exist in our case.

methods in the literature [KLF11, SY11], that can achieve this in $O(V \log V)$ time, where V is the number of vertices in the given meshes to be matched.

One of the best performing dense matching algorithms in the literature appears to be the BIM (blended intrinsic maps) method proposed in [KLF11]. The approach is to search for a continuous blend of multiple low-dimensional maps which can be explored via Möbius transformations. By combining these conformal maps with weights varying smoothly over the surface, a space of maps is obtained, which can be searched in polynomial-time. Thanks to blending, non-isometric deformations are handled better with this method as compared to other Möbius transformation based methods such as in [LF09, ZWW*10]. However, the BIM method introduces various approximations to achieve dense correspondence in $O(V \log V)$ time. First, the isometric distortion of each blending map is estimated on the extended complex plane to which the surfaces to be matched are transformed (flattened) via mid-edge uniformization, hence with some embedding error. Second, geodesic centroids of the blending maps are approximated with Euclidean distances. Third, confidence weights of the blending process are computed over all vertices of a dense surface mesh via interpolation. These approximations may lead to inaccurate matches (see Figure 1 for an example). Moreover the method has restrictions on topology, suffering from non-delaunay triangles during the process of mid-edge uniformization and also requiring zero-genus meshes as input.

In [SY11], the isometric dense correspondence problem is solved by using a multi-resolution strategy. A minimum distortion mapping is searched in the original 3D Euclidean space via coarse-to-fine combinatorial matching, hence as free of embedding errors. A coarse-to-fine approach is employed also in [BBK06a, WBBP11, RDK11] to speed up the optimization process and to improve accuracy. A major drawback of coarse-to-fine shape correspondence algorithms in general is the symmetric flip problem due to few number of samples to be matched at coarse levels (e.g. left arm/leg is

matched to right arm/leg between two human shapes). This paper essentially addresses this initial symmetric flip problem by extending the previous work presented in [SY11]. Our key idea is to keep track of all the optimal solutions, which may be more than one due to symmetry especially at coarse levels, throughout denser levels of the shape matching process. This becomes possible thanks to the coarse-to-fine combinatorial matching process involved in the method described in [SY11], which can sort out all possible mappings with respect to their isometric distortions. We test the performance of our method with extensive experiments on several shape benchmarks in comparison to two state of the art methods mentioned above, namely the GMDS method [BBK06a] and the BIM (blended intrinsic maps) method [KLF11].

This paper has two main contributions. First, it addresses the symmetric flip problem which is actually inherent to all coarse-to-fine correspondence algorithms, and extends the dense isometric shape correspondence method proposed in [SY11] in that respect. Our key idea of tracking symmetric flips can also be considered as a meta-approach which can be applied to other multi-resolution shape matching algorithms such as [BBK06a], as we demonstrate by experiments. Second, the extended method is tested on several shape benchmarks and compared with two state of the art techniques.

We note that the source code and the executables for the method that we present in this paper are publicly available in <http://home.ku.edu.tr/~yyemez/symflip>.

2. Coarse-to-Fine Approach for Shape Correspondence

The basic idea in coarse-to-fine shape correspondence methods is to compute the correspondence at resolution level k based on the correspondence at level $k - 1$ [BBK06a, SY11, WBBP11, RDK11]. Hence, starting from a coarse initial correspondence, a dense matching can efficiently be obtained by iterating this recursion. In this work, we focus on the method proposed in [SY11], which provides an isometric dense shape correspondence algorithm based on combinatorial search. We will refer to this method as coarse-to-fine combinatorial matching (C2FCM). We will also demonstrate the use of our idea of tracking symmetric flips on another well-known technique, namely the multi-resolution GMDS (generalized multi-dimensional scaling) [BBK06a]. In the sequel, we provide a brief summary of the C2FCM technique by highlighting its differences from GMDS.

The idea in C2FCM is to reduce the search space by exploiting the fact that the optimal correspondence maps nearby vertices on the source shape to nearby vertices on the target. Hence, shape matching is performed efficiently on a patch-by-patch basis in a coarse-to-fine manner. The C2FCM algorithm is composed of three basic tasks at each level of detail: sampling, combinatorial matching and merging (see Figure 2). Each of the shape surfaces to be matched is initially regarded as a root patch on which evenly-spaced high-curvature M base vertices are sampled. Each of these base vertices defines a surface patch around itself with a sampling radius. Once the samples, hence the patches, are matched via combinatorial matching by evaluating the isometric distortion of all $M!$ possible mappings, the matched parent patches are recursively subdivided into smaller sub-patches as levels of detail increase, by further sampling M base vertices on each parent. At each level, the samples on each pair

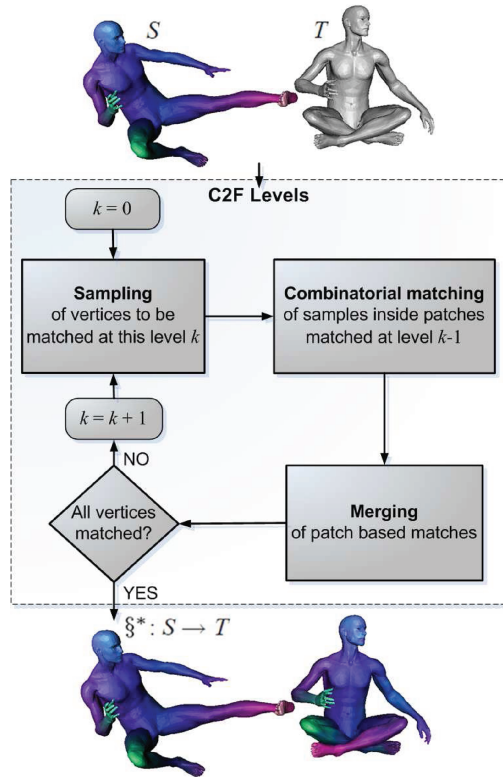


Figure 2: Block diagram of the C2FCM algorithm [SY11].

of corresponding patches are separately matched via combinatorial matching and then merged into one single correspondence. The number M is chosen to be small enough ($M \approx 5$) so that combinatorial matching becomes possible. We note that the algorithm does not require the initial set of samples to include all the shape extremities or to be exactly aligned on the two surfaces since the samples are evenly distributed over each surface and the combinatorial search algorithm guarantees to find the correspondence which is optimal in the minimum-distortion sense at the initial level. As the levels of detail proceed, the samples corresponding to high-curvature vertices are gradually populated and the correspondence accuracy is improved to the extent that the precision of the current level allows, as explained in detail in [SY11]. At the finest level, a complete dense correspondence is obtained, that matches every vertex in either of the shapes with at least one vertex on the other. Hence, the final correspondence is usually a many-to-many mapping. Although the algorithm is built upon the basic assumption that the shapes to be matched are perfectly isometric, the experiments conducted show that it performs well also on nearly isometric shapes. In the case of severe deviations from isometry however, the initially selected M samples can be in very different configurations on the two surfaces so that unintuitive matchings can be generated as the output of the algorithm. We finally note that the method can handle input meshes with arbitrary genus.

While C2FCM uses the coarse-to-fine scheme to gradually populate and fine-tune the correspondence pairs based on proximity clues, the GMDS method uses it to address a difficult non-linear optimization problem, where a non-convex stress function is minimized

by an iterative gradient-type algorithm which requires good initialization. A patch-based vertex hierarchy, which is similar to that of C2FCM, is first constructed via farthest point sampling [ELPZ97]. Then, starting from an initial coarse correspondence, the optimal solution computed at each resolution level is interpolated to the next level and used as an initialization for the non-convex optimization at that level. This process is repeated until the finest level solution is obtained. Hence, the coarse-to-fine scheme utilized in GMDS can be thought of as a smart way of initializing the non-convex optimization problem. In addition to the symmetric flip problem due to coarse sampling at the initial level, one disadvantage of GMDS is that it produces matchings which do not necessarily coincide with the initial sampling and often yields clustered correspondences due to the gradient-type algorithm employed in the optimization process.

3. Symmetric Flip Problem

Given a mapping $\xi: S \rightarrow T$ between two vertex sets S and T sampled from source and target shapes, respectively, we measure its isometric distortion D_{iso} as follows [SY11]:

$$D_{\text{iso}}(\xi) = \frac{1}{|\xi|} \sum_{(s_i, t_j) \in \xi} d_{\text{iso}}(s_i, t_j), \quad (1)$$

where $d_{\text{iso}}(s_i, t_j)$ is the contribution of the individual correspondence (s_i, t_j) to the overall isometric distortion:

$$d_{\text{iso}}(s_i, t_j) = \frac{1}{|\xi'|} \sum_{(s_l, t_m) \in \xi'} |g(s_i, s_l) - g(t_j, t_m)|, \quad (2)$$

and $g(\cdot, \cdot)$ is the geodesic distance between two vertices on a given surface. 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. In its simplest form, $\xi' = \xi - \{(s_i, t_j)\}$, but can also include some known correspondence pairs on the surfaces to be matched, if there are any, which are different than the ones included in ξ . Hence, more generally, $\xi - \{(s_i, t_j)\} \subseteq \xi'$. The correspondence problem can then be formulated as a combinatorial search for an optimal ξ^* over all possible mappings, that minimizes this isometric distortion function. One key observation here is that one can find more than one optimal mapping for symmetrical objects: the true correspondence plus the flipped versions. In the case of coarse-to-fine approach, the symmetric flip problem becomes even more severe due to few number of samples to be matched at coarse levels; a coarsely sampled version of a shape may sometimes appear to be symmetric although the shape is not symmetric when considered at a finer resolution as illustrated in Figure 3. Especially the initial level of the matching process in the C2FCM algorithm (or similarly in multi-resolution GMDS) may exhibit this configuration, and therefore is liable to symmetric flips which cannot be recovered as levels proceed. We note that all purely-isometric shape correspondence methods existing in the literature actually suffer from this symmetric flip problem regardless of the correspondence density since shapes may naturally exhibit intrinsic symmetries. In the coarse correspondence method of [DK11] for instance, the symmetric flip problem is cast to a sign ambiguity problem associated with the eigenfunctions of the Laplace–Beltrami operator to explicitly seek for all possible mappings which are equally or approximately optimal in terms of isometric distortion.

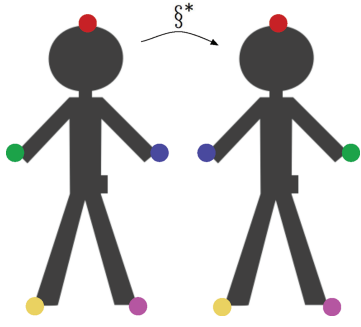


Figure 3: Symmetric flip problem due to coarse sampling: Although the shapes to be matched are not perfectly symmetric, the sampled vertices at the first level appear to be symmetric in terms of geodesic distances. This may yield a flipped coarse correspondence result as shown, which is actually one of the four optimal solutions.

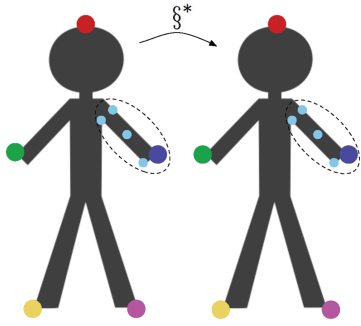


Figure 4: The C2FCM method handling symmetric flip problem at intermediate levels. Suppose that the mapping obtained at the first level is free of symmetric flips. Then at the next level, while matching the base vertices inside the highlighted patches (cyans and blue), the isometric distortions are computed by taking into account also the correspondences obtained at the previous level (other matching colours), which usually resolves the matching ambiguities due to symmetry.

In the C2FCM algorithm, symmetric flips may theoretically appear at any level since base vertices at each level are matched patch by patch, where each patch consists of only about $M = 5$ samples. However, because the C2FCM method expands the correspondence list $\mathcal{S}'$ used in computation of the distortion in Equation 2 at each intermediate level with the known correspondence pairs computed at the previous level, the chance of having flipped correspondences is greatly reduced at intermediate levels, as illustrated in Figure 4. This expansion is not possible at the first level, and the correspondence list $\mathcal{S}'$ in Equation 2 can thus simply be set to $\mathcal{S} - \{(s_i, t_j)\}$.

4. Tracking Symmetric Flips

The discussion in Section 3. basically suggests that by avoiding the symmetric flip problem only at the initial level, we can greatly reduce the chance of having a flipped dense correspondence. We propose to resolve any such ambiguities due to coarse sampling by tracking all the optimal (or near-optimal) solutions of the first level (optimal in

the minimum distortion sense) throughout intermediate levels of the shape matching process. This is possible thanks to the combinatorial matching algorithm that can sort all possible mappings with respect to their distortion values. The algorithm is described in the sequel.

We start by launching the C2FCM method at the first level (level 0), i.e. via combinatorial matching of the initial M samples. Although the original algorithm would pick and pursue only the minimum-distortion mapping of the first level, we keep track of top- N minimum-distortion mappings until some level K_1 , where the number of matched vertices is sufficiently large, which occurs typically at level 4 with about 250 vertices. Hence, for each of these N initial mappings, we separately pursue the C2FCM algorithm and compute the corresponding denser mapping at level K_1 . The choice of N is automatic since we choose it as the first significant jump in the ascendingly sorted sequence of distortion values $D_1, D_2, \dots, D_L$ of all possible mappings $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_L$, respectively, where distortions are computed via Equation (1) with $L = M!$. The first instance on this sorted distortion sequence, where the condition $D_{N+1} - D_N > \delta$ holds, reveals the desired jump in index N which is typically 4 for human shapes. We set the threshold δ as twice the average of the first 10 distortion difference values: $\delta = 2(\sum_{l=2}^{11} (D_l - D_{l-1}))/10$.

This tracking process along with the distortion curve is illustrated in Figure 5 for a pair of human shapes. Among the four mappings which are tracked until level $K_1 = 4$, the blue one ($\mathcal{S}_2$) wins the race, having the least isometric distortion at this intermediate level, and is to be maintained on its own until the final level that gives the dense map. Note that, with the original C2FCM method, the mapping $\mathcal{S}_1$ (in green) would continue alone and eventually converge to a flipped result in the absence of any other flipped competitors. We also observe that $\mathcal{S}_1$ yields flipped hands and legs, whereas $\mathcal{S}_4$ (cyan) has correct hands but flipped legs. The mapping $\mathcal{S}_6$ (orange), that is not actually tracked due to jump condition, would yield hand-to-head matching and flipped legs. Another demonstration of the tracking process is given in Figure 6 for a pair of glass shapes, where N comes out to be 2, a common value for the glasses class.

We note that, in all our experiments, we have used a fixed value, $K_1 = 4$, for the level of decision. One possible way of automating the choice of K_1 could be to keep tracking the N best initial maps up to the level where a clear winner among these maps can be differentiated based on the individual isometric distortions. For this purpose, a jump detection algorithm, that is similar to the one used for choosing N , can be employed.

5. Computational Complexity

The complexity of the original C2FCM method is $O(V \log V)$, where V is the number of vertices in the original meshes to be matched [SY11]. In our case, running the C2FCM method N times for tracking purposes up to some level K_1 , which is typically $K_1 = 4$, i.e. not up to the finest level where the dense map is obtained, incurs no additional asymptotic cost since N is usually a small number varying based on the intrinsic symmetries of the shapes to be matched, which is for example typically 4 for human shapes. Initial sorting of $M!$ mappings for jump detection also comes free of asymptotic cost for $M \leq 8$ but nevertheless, for practical usage,

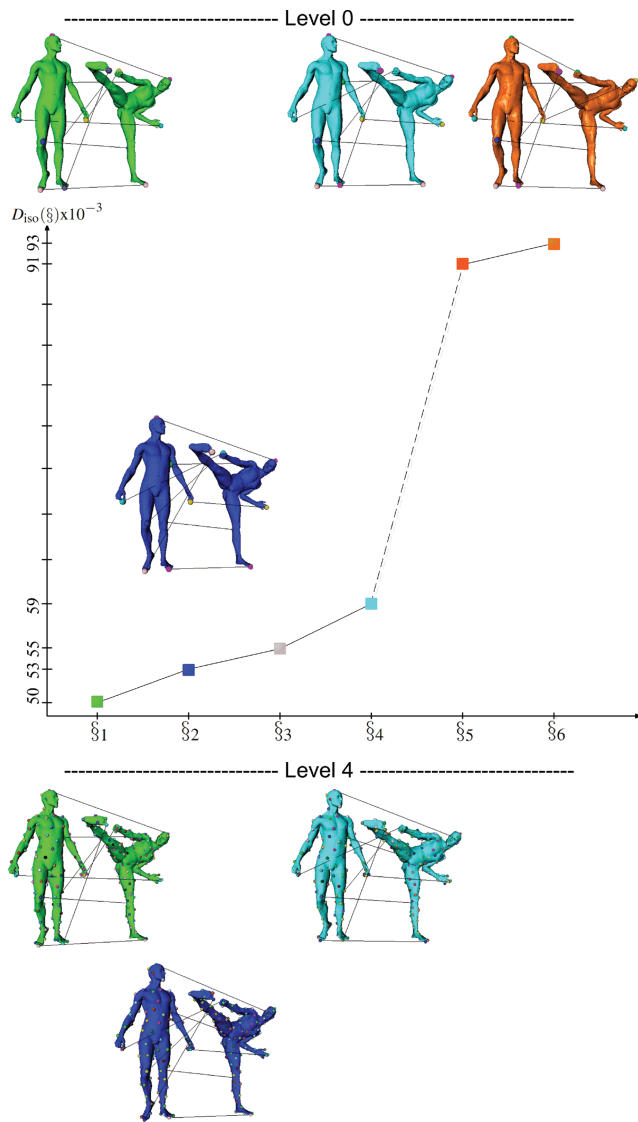


Figure 5: (Top) The distortion plot for the top 6 mappings at the first level along with the visualizations of ξ_1 , ξ_2 , ξ_4 and ξ_6 . (Bottom) The mappings that ξ_1 , ξ_2 and ξ_4 lead to at level $K_1 = 4$. Spheres of matching colours and lines indicate the correspondence pairs (all lines are not drawn at level 4). The mapping ξ_6 (orange) is not tracked since its distortion value appears after the first significant jump. The mapping ξ_3 is actually tracked but not shown in the figure for visual convenience.

we perform the sorting only on mappings with isometric distortion below a threshold value (0.15 in our experiments). This significantly reduces the number of mappings (typically to $M' \approx 300$) to sort with no accuracy loss as the true mapping to be tracked will almost certainly be in the top- M' least distorted mappings. In our experiments, it has always remained within the top 6.

Our overall $O(V \log V)$ complexity outperforms that of other isometric dense shape correspondence algorithms in the literature,

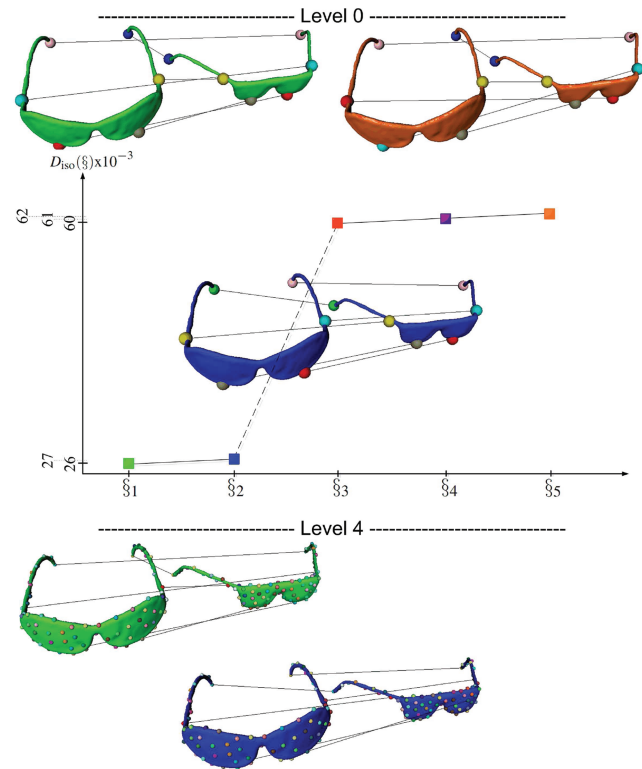


Figure 6: Symmetric flip tracking for a pair of glasses shapes. Same layout with Figure 5 except three out of top 5 mappings at level 0 shown for visual convenience.

Table 1: Percentage values for dense correspondence results without symmetric flips, obtained with different methods. A: Within animals, B: Within humans and C: Across humans. Last three rows show the results on down-sampled versions of classes A–C.

Dataset: TOSCA	% of all results		
	A	B	C
C2FCM	59	61	38
Our method	69	68	78
BIM	100	100	95
C2FCM	37	60	67
Our method	58	75	83
GMDS	37	54	39

such as $O(V^2 \log V)$ complexity of [JZ06] and GMDS [BBK06a], $O(V^2)$ complexity of HKM [OMMG10], $O(V^2 \log V + Y^4 \log Y)$ complexity of Möbius Voting [LF09]. Note that, GMDS and Möbius Voting can achieve a coarse correspondence between Y feature points which can later be interpolated into a dense one with no additional asymptotic cost. Blended intrinsic maps of [KLF11] is faster than these methods with $\max(O(V \log V), O(W^2))$ complexity with W being an upper bound on the number of maps considered.

Table 2: Isometric and ground-truth distortions obtained on TOSCA with our method as compared to BIM. A: Within animals, B: Within humans and C: Across humans.

Class	Our method		BIM	
	$D_{\text{iso}}, d_{\text{iso}}^{\dagger}$	$D_{\text{grd}}, d_{\text{grd}}^{\dagger}$	$D_{\text{iso}}, d_{\text{iso}}^{\dagger}$	$D_{\text{grd}}, d_{\text{grd}}^{\dagger}$
A	0.018, 0.081	0.030, 0.226	0.018, 0.071	0.018, 0.155
B	0.015, 0.052	0.033, 0.202	0.013, 0.266	0.021, 0.417
C	0.019, 0.065	n/a	0.017, 0.068	n/a

Table 3: Isometric distortions obtained on TOSCA downsampled with our method as compared to GMDS. A: Within animals, B: Within humans and C: Across humans.

Class	Our method	GMDS
	$D_{\text{iso}}, d_{\text{iso}}^{\dagger}$	$D_{\text{iso}}, d_{\text{iso}}^{\dagger}$
A	0.024, 0.045	0.017, 0.051
B	0.018, 0.049	0.022, 0.043
C	0.020, 0.081	0.026, 0.057

6. Experimental Results

We have employed a comprehensive test suite consisting of five well-known 3D shape benchmark datasets: TOSCA [BBK08], Watertight [GBP07], SHREC'11 [BBB*11], SCAPE [ASK*05] and Nonrigid World [BBK06a]. Our results are also compared with two state-of-

the-art techniques, namely GMDS [BBK06a] and Blended Intrinsic Maps (BIM) [KLF11] by running their publicly available codes.

We provide symmetric flip ratios resulting from each dataset in Tables 1, 4, 6 and 8. We also demonstrate numerical (Tables 2, 3, 5, 7 and 9) and visual (Figures 7–11) quality of our final dense maps in comparison with the two methods. We measure isometric quality of the final maps by D_{iso} using Equation (1), whereas ground-truth distortion is measured by D_{grd} (whenever ground-truth correspondence pairs are available as $(s_i, t_i) \forall i$):

$$D_{\text{grd}}(\S) = \frac{1}{|\S|} \sum_{(s_i, t_j) \in \S} g(t_i, t_j). \quad (3)$$

We also employ two additional worst-case distortion measures, the maximum ground-truth distortion:

$$d_{\text{grd}}^{\dagger} = \max_{(s_i, t_j) \in \S} g(t_i, t_j), \quad (4)$$

and the maximum isometric distortion:

$$d_{\text{iso}}^{\dagger} = \max_{(s_i, t_j) \in \S} d_{\text{iso}}(s_i, t_j), \quad (5)$$

where d_{iso} is the isometric distortion function given in Equation (2). In computation of the performance measures, we use normalized geodesic distances so that the maximum geodesic distance on a shape surface is taken as 1.0.

As a convention that applies to all figures, for each dataset and a particular shape class, we display the mapping on the pair for which the method whose $d_{\text{iso}}^{\dagger}$ (or $d_{\text{grd}}^{\dagger}$) is worse than the other method for that particular class; a $\dagger$ sign emphasizes this interesting mapping.

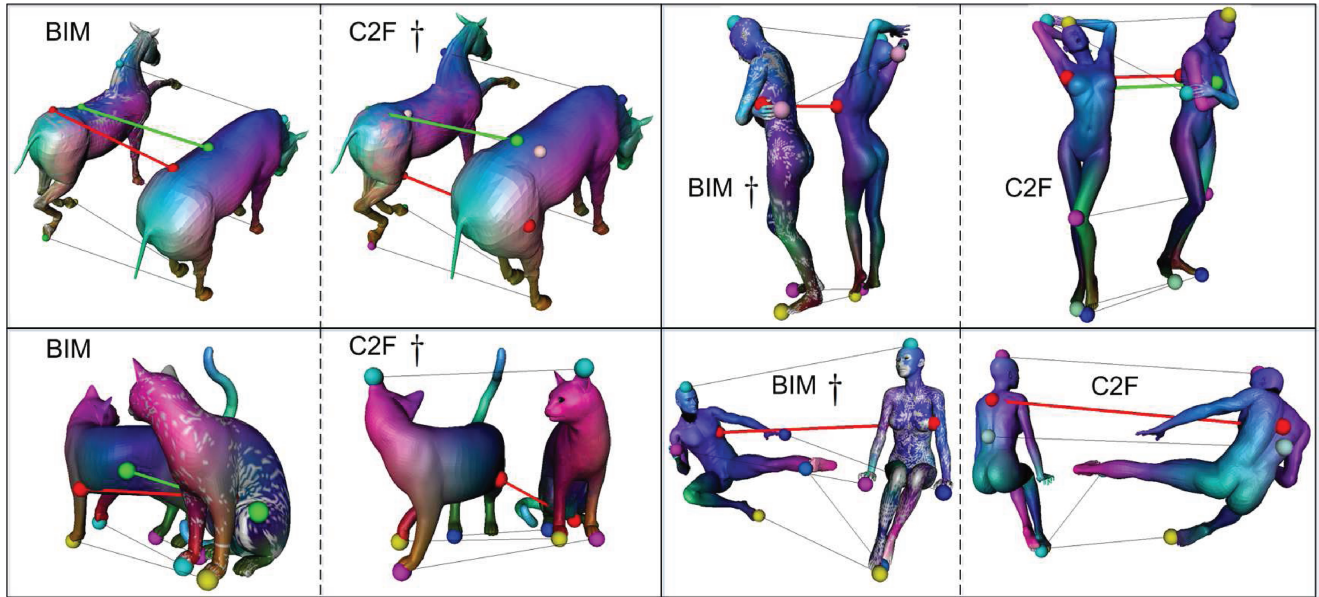


Figure 7: Dense correspondence results on TOSCA, obtained with our method (C2F) vs. BIM. The red and green lines indicate the worst matches w.r.t. isometric and ground-truth distortions, respectively. The worst matchings are in general better in our case when compared to BIM. Notice for example hand-to-breast and breast-to-arm matches (third column) as well as the unmatched regions shown in grey for BIM results. In addition to transferred colours, some lines ending with spheres of matching colours are used to enhance visuals. When red and green lines overlap, only red is shown. This whole representation scheme applies to the subsequent figures as well.

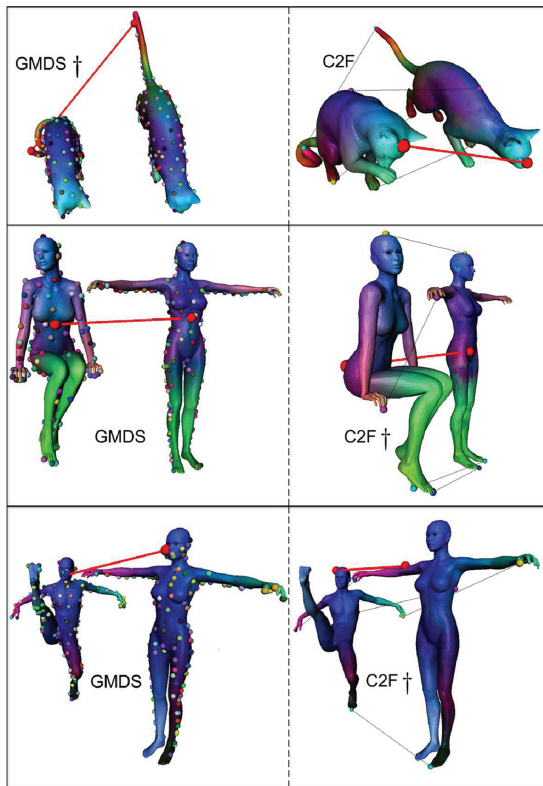


Figure 8: Dense correspondence results on TOSCA downsampled, obtained with GMDS (left) and with our method (right). Generating 150 correspondence pairs for GMDS are shown as spheres of matching colours.

The same pair is also shown with the mapping computed by the other method—expected to be better. In all figures, the dense mapping is visualized by transferring source vertex colours to the target through computed correspondences.

Finally, there is also a downsampled version of each dataset that we use, with meshes of size at most 5K to be able to run the GMDS method which does not scale well to large meshes.

6.1. Dataset: TOSCA

The original TOSCA dataset is explored under three classes: within animals (11 cats and 8 horses), within humans (two different males with 12 and 20 poses each and 12 females), and across humans (using the three human classes above). The within-animal results are obtained by matching every cat (or horse) to a different randomly selected cat (or horse). The same applies to within-human results. Likewise, in the across-human experiments, each human model in the dataset is matched to another model randomly selected from a different human class.

In Table 1—rows 2–4, we see for instance that $100\% - 38\% = 62\%$ of across-human tests start with an initial symmetric flip at level 0, which means that if we had employed the original C2FCM algorithm, then 62% of the resulting dense maps would suffer from

being symmetrically flipped. With our method, on the other hand, only $100\% - 78\% = 22\%$ of all correspondence results is symmetrically flipped at the end. Another interpretation is that 62% initial symmetric flips reduce to just 22% in the final result, an improvement of 40%. The BIM method [KLF11] is 95% free of symmetric flip errors for this class, and outperforms our method for all other classes as well in this set as far as the symmetric flip problem is concerned.

For isometric and ground-truth distortions in this dataset, we observe in Table 2 that $d_{iso}^\dagger$ and $d_{grd}^\dagger$ values favour our method for within and across humans as those classes may include pairs with touching surface parts, e.g. hand-to-breast, which are confused by BIM that approximates geodesic centroid as a weighted Euclidean centroid (see Figure 7 and also Figure 1). As for the average values D_{iso} and D_{grd} , our method is on a par with the results of BIM.

We run the GMDS method of [BBK06a] on a downsampled version of TOSCA database, using random test pairs chosen in the same way as described previously. The number of vertex samples to be matched on the decimated models is set to be 150, a parameter of the publicly available code. The 150 correspondence pairs, which are obtained with GMDS for each shape pair, are then used to interpolate a dense map as described in [KLF11].

The resulting symmetric flip ratios (Table 1—rows 5–7) as well as average isometric distortion values (Table 3) all favour our method in this dataset when compared to GMDS. Our maximum distortion values, however, are slightly worse on human test pairs. Nevertheless, we see that 150 vertices matched with GMDS are not evenly distributed on the shape surfaces, which results in clustered correspondences (Figure 8), hence one of the major shortcomings of the GMDS method. This is mainly due to the fact that the GMDS algorithm employs a gradient-based iterative optimization process that produces sub-vertex 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 sub-vertex coordinates to their nearest vertices on the surface mesh. The effect of this rounding process to the computed distortion values is negligible since the mesh models in this, and all other datasets are almost uniform and at relatively high resolution. We also observe from the visual comparison that the GMDS matching often misses the salient points of a shape (e.g. ear tips of the cats).

6.2. Dataset: Watertight

The part that we have used from the original dataset consists of human, glasses, chair, teddy bear, hand, fish and armadillo classes of cardinality 20 each. We have omitted ant and octopus classes which are too symmetric to handle in the sense that the shapes in these classes do not contain sufficient clues to resolve ambiguities due to symmetry even at denser resolution levels. We also exclude the remaining 10 classes since they exhibit severe non-isometries, such as airplanes and four-leggeds, while our method, as well as GMDS, has been designed to work on isometric shapes. In these experiments, an object fetched from a class is matched to a random object from the same class.

Concerning symmetric flips (Table 4—rows 2–4), our method is as good as BIM which cannot however handle the chair class with

Table 4: Percentage values for dense correspondence results without symmetric flips, obtained with different methods. A: Humans, B: Glasses, C: Chairs, D: Teddy bears, E: Hands, F: Fishes and G: Armadillos. Last three rows show the results on downsampled versions of classes A to G.

Dataset:	% of all results						
	A	B	C	D	E	F	G
Watertight							
C2FCM	28	63	30	50	50	69	60
Our method	89	88	70	67	88	88	90
BIM	89	57	n/a	83	100	85	89
C2FCM	61	29	69	42	50	43	37
Our method	83	71	66	50	83	86	75
GMDS	28	40	n/a	33	50	71	25

nonzero genus objects and is slightly worse than our method on isometric distortion measurements (Table 5—columns 2–3).

As the flow of the transferred colours shows (Figure 9—top), our dense map is not always as smooth as the one generated by BIM which interpolates the feature correspondences in the extended complex plane by nearest neighbour search to generate the conformal maps to be blended. However, these smooth outputs are obtained at the expense of leaving a fair amount of target vertices unmatched (grey), mainly due to the blending step as well as discretization involved while generating discrete conformal maps. Our final map, on the other hand, is always onto, i.e. no vertices are left unmatched.

Table 5: Isometric distortions on Watertight dataset, obtained with different methods. A: Humans, B: Glasses, C: Chairs, D: Teddy bears, E: Hands, F: Fishes and G: Armadillos. Fourth and fifth columns give the results on downsampled versions of classes A–G.

Class	Our method	BIM	Our method	GMDS
	$D_{iso}, d_{iso}^\dagger$	$D_{iso}, d_{iso}^\dagger$	$D_{iso}, d_{iso}^\dagger$	$D_{iso}, d_{iso}^\dagger$
A	0.022, 0.109	0.035, 0.121	0.021, 0.053	0.031, 0.079
B	0.008, 0.020	0.063, 0.248	0.007, 0.015	0.034, 0.147
C	0.024, 0.109	n/a	0.022, 0.022	n/a
D	0.032, 0.108	0.048, 0.225	0.029, 0.081	0.032, 0.119
E	0.034, 0.138	0.046, 0.167	0.027, 0.049	0.029, 0.134
F	0.033, 0.140	0.061, 0.192	0.024, 0.079	0.024, 0.069
G	0.030, 0.185	0.049, 0.193	0.031, 0.098	0.039, 0.135

Downsampled Watertight models, when fed into GMDS in the way described in Section 6.1., are significantly more prone to symmetric flips than our method (Table 4—rows 5–7). Initial GMDS matchings suffer from the clustering issue (Figure 9—bottom), and when interpolated into a dense map, produce worse results as compared to ours in terms isometric distortion (Table 5—columns 4 and 5). Note also that, although a small topological noise connecting index and middle fingers does not fail our method (see Figure 9), noisy connections that alter geodesic distances more severely may cause instabilities that are not expected in diffusion-based methods such as [OMMG10].

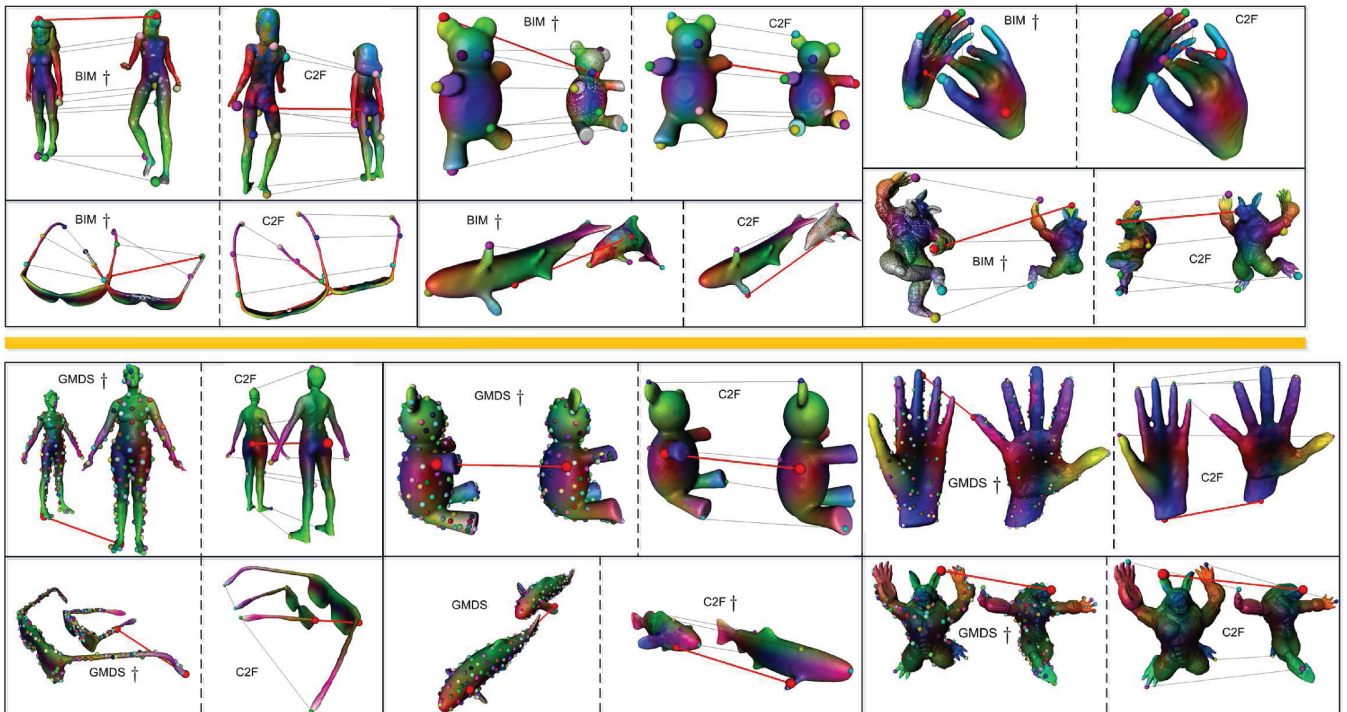


Figure 9: Dense correspondence results on Watertight dataset, obtained with our method vs. BIM (top) and vs. GMDS (bottom, separated by yellow line).

Table 6: Percentage values for dense correspondence results without symmetric flips, obtained with different methods. A: Noise, B: Shotnoise and C: Shotnoise on the downsampled set.

Dataset: SHREC'11	% of all results		
	A	B	C
C2FCM	20	87	54
Our method	80	94	62
BIM	50	38	n/a
GMDS	n/a	n/a	23

Table 7: Isometric distortions on SHREC'11, obtained with different methods. A: Noise, B: Shotnoise and C: Shotnoise on the downsampled set.

Class	Our method $D_{iso}, d_{iso}^{\dagger}$	BIM $D_{iso}, d_{iso}^{\dagger}$	GMDS $D_{iso}, d_{iso}^{\dagger}$
A	0.019, 0.066	0.013, 0.079	n/a
B	0.015, 0.042	0.032, 0.108	n/a
C	0.010, 0.066	n/a	0.026, 0.074

6.3. Dataset: SHREC'11

We employ SHREC'11 dataset to see to what extent our algorithm tolerates noise and holes of various sizes. Due to zero genus constraint of the BIM method, only noise and shotnoise classes are used for the corresponding comparison suite. The GMDS public code, on the other hand, crashes on all test cases but the shotnoise. The results summarized in Tables 6 and 7 generally show that both of our competitors are more sensitive to noise as compared to our method (see Figure 10 as well). Among five different shotnoise levels tested, we fail to achieve stable results for the forth and fifth levels (in increasing order) which possess severe shotnoises.

We give our results on microholes, holes and viewing classes (which cannot be obtained with BIM and GMDS) in the format of $(D_{iso}, d_{iso}^{\dagger})$, as (0.013, 0.059), (0.022, 0.058) and (0.016, 0.068), respectively (Figure 10—bottom).

6.4. Dataset: SCAPE

Each SCAPE model represents the articulated motion of a human actor reconstructed from range data. We match each SCAPE model to a random model from the remaining 71 shapes. Although the BIM method handles symmetric flips better than our method in this dataset (Table 8), the mid-edge uniformization phase of their conformal maps to be blended suffers from non-delaunay triangles on this raw data, hence leaving them slightly behind us in distortion performances (Table 9) when coupled with weighted Euclidean centroids approximation of geodesic centroids for efficiency. The GMDS method is again slightly worse than our method in terms of symmetric flips (Table 8) and distortion measurements $((D_{iso}, d_{iso}^{\dagger}) = (0.025, 0.086)$ of our method versus (0.028, 0.103) of GMDS) mainly due to their clustered generating matches.

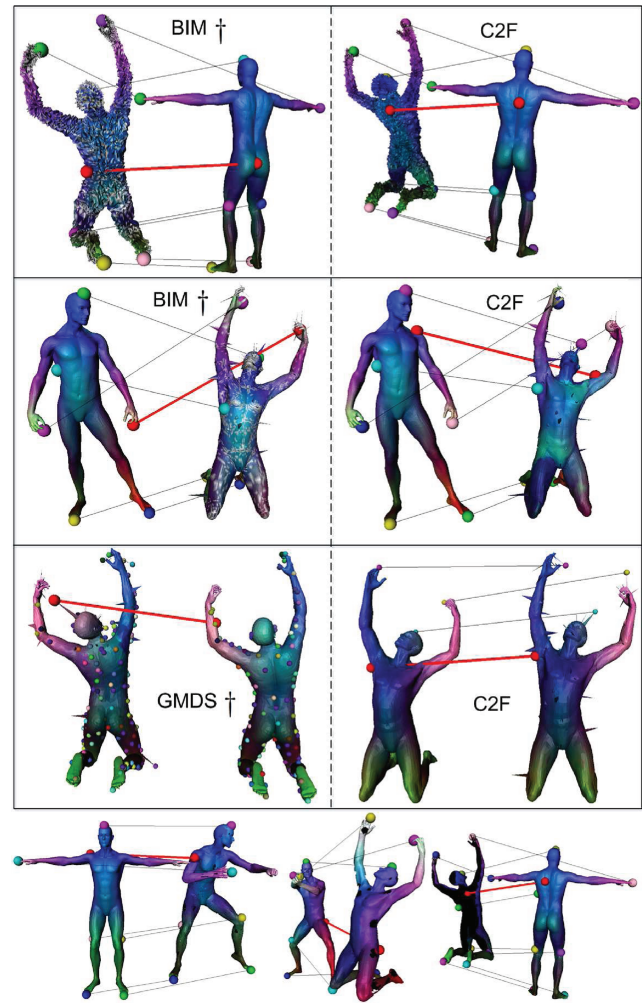


Figure 10: Dense correspondence results on SHREC'11. Sample pairs for noise, shotnoise and downsampled shotnoise classes (top to bottom), obtained with our method (right) and with BIM and GMDS (left). Our result samples for microhole, hole and viewing classes are displayed separately at the bottom (left-right).

6.5. Application to GMDS

In addition to C2FCM [SY11] on which our symmetric flip tracking approach has been verified thus far, we now show how well this method, as a meta-approach, extends to another multi-resolution isometric shape matching algorithm, namely GMDS [BBK06a], whose optimization method is completely different from C2FCM. To this end, we first run the public GMDS code on a part of the Nonrigid World database that contains 8 cats, 6 centaurs and 17 horses (for within animals class) as well as 10 male and 23 female figures (for within and across humans classes), where the classes are formed in the same way as described in Section 6.1. We then slightly modify the original public code so as to initialize GMDS with the qualified maps generated at the first level of our method where qualification is based on the jump detection scheme described in Section 4. Amongst the qualified maps that are tracked until a

Table 8: Percentage values for dense correspondence results without symmetric flips, obtained with different methods. A: SCAPE and B: SCAPE downsampled.

Dataset: SCAPE	% of all results	
	A	B
C2FCM	56	24
Our method	60	61
BIM	92	n/a
GMDS	n/a	42

Table 9: Isometric and ground-truth distortions on SCAPE dataset, obtained with our method in comparison to BIM.

Our method		BIM	
$D_{iso}, d_{iso}^\dagger$	$D_{grd}, d_{grd}^\dagger$	$D_{iso}, d_{iso}^\dagger$	$D_{grd}, d_{grd}^\dagger$
0.019, 0.097	0.045, 0.308	0.027, 0.254	0.039, 0.306

relatively dense GMDS level that matches 250 vertices, we select the one yielding the minimum GMDS distortion as the final map. In Table 10, we provide the symmetric flip ratios on the final maps, which clearly favour the latter GMDS modified by our tracking method (last row). We depict the tracking process in Figure 12 along with the corresponding distortion curve, where the number of tracked (qualified) maps is 8, a common value for centaurs having three different pairs of limbs.

6.6. Timing

With $O(V \log V)$ time complexity, our algorithm scales well to large meshes such as TOSCA and SHREC'11 humans of 53 K vertices, TOSCA animals and Watertight armadillos of 28 K vertices and SCAPE humans of 13 K vertices. The tracking process for symmetric flips (Section 4.) is quite fast; on the largest pair of 53 K vertices, a 2.53 GHz 64-bit workstation takes 10 s to evaluate $M!$ distortions of maps of sizes M each via D_{iso} (Equation 1) for $M = 8$ and sort about 300 of them. This is followed by about 48 s tracking of top- N (typically ≤ 4) mappings until level K_1 (typically 4). The original C2FCM method then continues to bring the qualified mapping to a dense map in 1578 s. For a SCAPE pair with around 13 K vertices, these evaluation-sorting, tracking and dense mapping times become 2, 13 and 391 s, respectively, whereas for a downsampled shape pair, e.g. from TOSCA, with 5 K vertices, the respective execution times are 1, 4 and 66 s.

6.7. Discussion

We summarize our comparative experimental findings as follows:

- 1 In terms of symmetric flip ratio, our method is always better than GMDS which is, as a coarse-to-fine solution, prone to symmetric flips at the initial levels with yet no care taken.

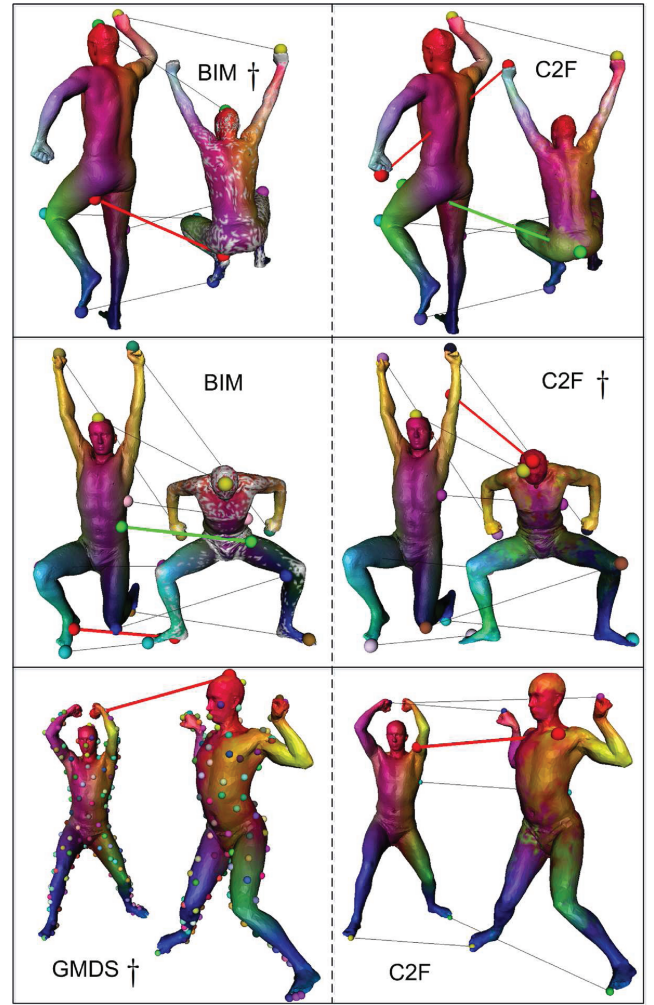


Figure 11: Dense correspondence results on SCAPE. BIM yields the worst $d_{iso}^\dagger$ (top row, red line), our method the worst $d_{grd}^\dagger$ (middle row, green overlapped by red) and GMDS the worst $d_{iso}^\dagger$ (bottom row) on SCAPE downsampled.

While BIM performs as good as our approach on Watertight models, it is worse for SHREC'11 models that are exposed to various types of noises. The results on TOSCA and SCAPE models favour BIM as the conformal maps being blended are area-preserving at every point which proves useful in distinguishing mapping the front of a human to the back.

- 2 In isometric distortion performance, which goes parallel to ground-truth performances (whenever ground-truth correspondences are available), our method outperforms GMDS for all datasets mainly because of the clustered matchings of the latter, that are interpolated into a dense map. The BIM method, on the other hand, is on a par with ours on TOSCA and slightly worse for all other datasets concerning average distortions as the conformal maps to be blended induce embedding errors, and more importantly confidence weights of the blending process are computed over all vertices of a

Table 10: Percentage values for correspondence results without symmetric flips, obtained with the original and modified GMDS methods. A: Within animals, B: Within humans and C: Across humans.

Dataset: Nonrigid world	% of all results		
	A	B	C
GMDS	19	42	47
GMDS modified with tracking	61	70	67

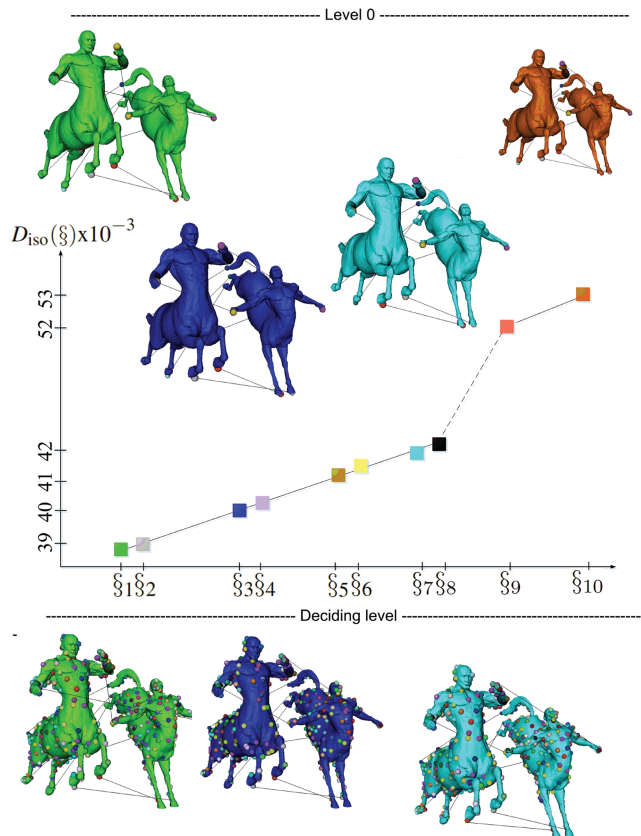


Figure 12: (Top) The distortion plot for the top 10 mappings at the first level along with the visualizations of $\S_1$, $\S_3$, $\S_7$ and $\S_{10}$. (Bottom) The GMDS mappings that $\S_1$, $\S_3$ and $\S_7$ lead to at the dense GMDS level where the map generated by $\S_7$ (cyan) wins the race having the least distortion. The mapping $\S_{10}$ (orange) is not even tracked.

dense surface mesh via interpolation. Recall that our algorithm is embedding and interpolation free. Moreover, the BIM yields unstable results on the datasets that include shapes with touching surface parts since Euclidian centroid is replaced with the geodesic one, with worst distortion behaviour significantly behind ours.

- 3 In the case of GMDS, the uniformity of the evenly-spaced initial samples is not necessarily preserved in the final correspondence output, hence matching salient shape points such as ear tips is not guaranteed. Besides, the algorithm does

not support dense matching without interpolation that incurs more errors. We usually hit salient points at intermediate levels and yet support dense matching without need for any interpolation.

- 4 Sphere topology restriction of the BIM as well as its sensitivity to peculiarities of a given triangulation, e.g. non-delaunay triangles, are avoided in our approach (see especially the experimental results on SHREC'11 and SCAPE datasets).
- 5 The conformal maps in BIM are not usually onto after closest point matching in the embedding domain, which in turn leaves a significant amount of target vertices unmatched after blending. Further interpolation would be required to obtain the full dense mapping that we produce, which is many-to-many. The interpolated conformal maps in BIM, however, generally leads to smoother dense maps than ours.
- 6 Our symmetric flip tracking approach, when applied to the multi-resolution GMDS method, significantly reduces the percentage of the correspondence results with symmetric flips.
- 7 By tracking, we significantly reduce the occurrence rate of symmetric flips as an end result, yet cannot completely solve the problem. In our experiments, in almost all cases where the symmetric flip problem cannot be resolved, the true solution is included in the initial set of qualified mappings, but does not yield the minimum-distortion after tracking (further increasing the resolution of the intermediate dense level does not help either). This is mainly due to two reasons. First, in the case of perfectly isometric and intrinsically symmetric shapes, even at dense resolutions, there are actually more than one correspondence, yielding distortion values which are all close to the global minimum but possibly with some inaccuracy due to discretization and modeling errors. Second, the shapes to be matched are often only nearly isometric while the multi-resolution correspondence algorithms that we have tested are designed to work on perfectly isometric shape pairs. For the rare cases where the true correspondence does not even appear in the initial set of maps to be tracked, one possibility could be to seek for local minima of the distortion function, that would reveal the approximate symmetries more accurately, as suggested in [RBBK09].

7. Conclusion

We have extended our original efficient dense isometric shape correspondence algorithm [SY11] so as to address the symmetric flip problem. To this effect, we couple the coarse-to-fine structure of the original algorithm with a tracking mechanism that brings some candidate initial maps into a finer resolution where their isometric distortions can be more accurately computed. Maintaining only the best one at this stage generally reduces the symmetric flips in the final correspondence, as demonstrated on five well-known benchmark 3D shape datasets in comparison with two state-of-the-art techniques. Our final dense maps are also better than or on a par with our competitors as far as the final isometric and ground-truth distortions are concerned. We perform the isometric matching in the original 3D Euclidean space wherein isometry is defined, and hence free of embedding errors. Our dense map is also interpolation-free as the combinatorial matching continues efficiently until all vertices

are explicitly matched. Note also that, this tracking-based solution can be used until a desired coarseness level of correspondence depending on the application or the method to be compared with. The core idea of tracking symmetric flips (Section 4.) can be adapted to other techniques affected by the same issues (Section 6.5.) as well as to future multi-resolution isometric shape matching algorithms. We finally note that the proposed method can be used for symmetry detection by seeking optimal (or near-optimal) maps from a shape to itself in the same spirit as other symmetry detection methods available in the literature such as [RBBK09, OSG07]. As future work, we plan to make the framework available for partially isometric shapes.

Acknowledgement

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

References

- [ASK*05] ANGUELOV D., SRINIVASAN P., KOLLER D., THRUN S., RODGERS J., DAVIS J.: Scape: shape completion and animation of people. *ACM Transactions on Graphics* 24, 3 (2005), 408–416.
- [BBB*11] BOYER E., BRONSTEIN A. M., BRONSTEIN M. M., BUSTOS B., DAROM T., HORAUD R., HOTZ I., KELLER Y., KEUSTERMANS J., KOVNATSKY A., LITMAN R., REININGHAUS J., SIPIRAN I., SMEETS D., SUETENS P., VANDERMEULEN D., ZAHARESCU A., ZOBEL V.: SHREC 2011: robust feature detection and description benchmark. In *Proceedings of the Eurographics Workshop on 3D Object Retrieval*. Llandudno, UK, (2011).
- [BBK06a] BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: Efficient computation of isometry-invariant distances between surfaces. *SIAM Journal of Scientific Computing* 28, 5 (2006), 1812–1836.
- [BBK06b] BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: Generalized multidimensional scaling: a framework for isometry invariant partial surface matching. In *Proceedings of the National Academy of Science* 103, 5 (2006), 1168–1172.
- [BBK08] BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: *Numerical Geometry of Non-Rigid Shapes*. Springer, New York, 2008.
- [CH03] CARCASSONI M., HANCOCK E. R.: Correspondence matching with modal clusters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 12 (2003), 1609–1615.
- [DK11] DUBROVINA A., KIMMEL R.: Approximately isometric shape correspondence by matching pointwise spectral features and global geodesic structures. *Advances in Adaptive Data Analysis (AADA)* (2011), 203–228.
- [EK03] ELAD A., KIMMEL R.: On bending invariant signatures for surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (2003), 1285–1295.
- [ELPZ97] ELДАР Y., LINDENBAUM M., PORAT M., ZEEVI Y.: The farthest point strategy for progressive image sampling. *IEEE Transactions on Image Processing* 6 (1997), 1305–1315.
- [GBP07] GIORGI D., BIASOTTI S., PARABOSCHI L.: SHREC: Shape retrieval contest: watertight models track (2007), 1–11.
- [JZ06] JAIN V., ZHANG H.: Robust 3d shape correspondence in the spectral domain. *IEEE International Conference on Shape Modeling and Applications (SMI)* (2006), 118–129.
- [KLF11] KIM V., LIPMAN Y., FUNKHOUSER T.: Blended intrinsic maps. *Proceedings of SIGGRAPH* 30, 4 (2011), 1–12.
- [LF09] LIPMAN Y., FUNKHOUSER T.: Möbius voting for surface correspondence. *ACM Transactions on Graphics* 28, 3 (2009), 1–12.
- [MHK*08] MATEUS D., HORAUD R., KNOSSOW D., CUZZOLIN F., BOYER E.: Articulated shape matching using laplacian eigenfunctions and unsupervised point registration. In *Proceedings of the Computer Vision and Pattern Recognition (CVPR)* (2008), 1–8.
- [OBSC*12] OVSJANIKOV M., BEN-CHEN M., SOLOMON J., BUTSCHER A., GUIBAS L.: Functional maps: a flexible representation of maps between shapes. *Proceedings of the SIGGRAPH*. Los Angeles, (2012).
- [OMMG10] OVSJANIKOV M., MÉRIGOT Q., MÉMOLI F., GUIBAS L.: One point isometric matching with the heat kernel. *Computer Graphics Forum* 29, 5 (2010), 1555–1564.
- [OSG07] OVSJANIKOV M., SUN J., GUIBAS L.: Global intrinsic symmetries of shapes. *Computer Graphics Forum* 27, 5 (2007), 1341–1348.
- [RBBK09] RAVIV D., BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: Full and partial symmetries of non-rigid shapes. *International Journal of Computer Vision* 89, 1 (2009), 18–39.
- [RDK11] RAVIV D., DUBROVINA A., KIMMEL R.: Hierarchical matching of non-rigid shapes. In *Proceedings of the Scale Space and Variational Methods* (2011), 604–615.
- [SH10] SHARMA A., HORAUD R. P.: Shape matching based on diffusion embedding and on mutual isometric consistency. In *Proceedings of the Workshop on Nonrigid Shape Analysis and Deformable Image Alignment (NORDIA)*. San Francisco, (2010).
- [SY11] SAHILLIOĞLU Y., YEMEZ Y.: Coarse-to-fine combinatorial matching for dense isometric shape correspondence. *Computer Graphics Forum (Proc. SGP)* 30, 5 (2011), 1461–1470.
- [SY12] SAHILLIOĞLU Y., YEMEZ Y.: Minimum-distortion isometric shape correspondence using EM algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 11 (2012), 2203–2215.
- [vKZHC011] VAN KAICK O., ZHANG H., HAMARNEH G., COHEN-OR D.: A survey on shape correspondence. *Computer Graphics Forum* 30, 6 (2011), 1681–1707.
- [WBBP11] WANG C., BRONSTEIN M. M., BRONSTEIN A. M., PARAGIOS N.: Discrete minimum distortion correspondence problems for non-rigid shape matching. In *Proceedings of Scale Space and Variational Methods*. Ein Gedi, Israel, (2011).

- [WSB09] WUHRER S., SHU C., BOSE P.: Posture invariant correspondence of triangular meshes in shape space. In *Proceedings of the International Conference on 3D Digital Imaging and Modeling (3DIM)*. Kyoto, Japan, (2009).
- [ZWW*10] ZENG Y., WANG C., WANG Y., GU X., SAMARAS D., PARAGIOS N.: Dense non-rigid surface registration using high-order graph matching. *Proc. Computer Vision and Pattern Recognition (CVPR)*. San Francisco, (2010).