Robust Nonrigid Registration by Convex Optimization

Qifeng Chen
Stanford University

Vladlen Koltun
Intel Labs

Abstract

We present an approach to nonrigid registration of 3D surfaces. We cast isometric embedding as MRF optimization and apply efficient global optimization algorithms based on linear programming relaxations. The Markov random field perspective suggests a natural connection with robust statistics and motivates robust forms of the intrinsic distortion functional. Our approach outperforms a large body of prior work by a significant margin, increasing registration precision on real data by a factor of 3.

1. Introduction

Given two surfaces $\mathcal{S}, \mathcal{T} \subset \mathbb{R}^3$, the nonrigid registration problem calls for computing a physically or perceptually meaningful mapping $f : \mathcal{S} \rightarrow \mathcal{T}$. For example, if $\mathcal{S}$ and $\mathcal{T}$ are two models of people in different poses, a natural mapping associates the corresponding body parts [6]. Such mappings can be used for loop closure detection in dynamic scene reconstruction, regularization in joint shape analysis, propagation of material properties across 3D models, and other applications in computer vision and graphics.

Treatments of the registration problem differ in the mathematical formulation of the registration objective—typically a functional on the space of mappings—and the computational approach to optimizing the objective. A particularly elegant formulation calls for minimizing the intrinsic distortion induced by the mapping [8]. If $\mathcal{S}$ and $\mathcal{T}$ are accurate surface models of near-isometric shapes, a low-distortion mapping can yield plausible correspondences.

Since models produced by 3D scanning systems are imperfect, registration algorithms that operate on such data must be robust to its defects. These defects can be mild, such as surface noise and occlusion gaps, or severe, such as large-scale fusion of adjoining parts. The severe defects are particularly challenging because they violate basic assumptions made by many registration algorithms. These violations commonly lead to registration failure. A recent evaluation of state-of-the-art algorithms on real data with carefully collected ground truth found that none of the evaluated algorithms performed satisfactorily [6].

In this paper, we develop an approach to surface registration that alleviates some of the difficulties encountered by existing techniques. Our guiding observation is that minimization of the intrinsic distortion functional can be cast as MRF optimization [2]. We show that this simple observation has profound implications for surface registration.

The MRF formulation allows us to tap into highly efficient global optimization algorithms based on linear programming relaxations. We show that simply optimizing a variant of the classical intrinsic distortion objective [8] using modern MRF optimization machinery yields state-of-the-art registration accuracy, outperforming far more complex registration pipelines.

Furthermore, the Markov random field perspective suggests a natural connection with robust statistics and motivates robust forms of the embedding objective. We show that optimizing a robust objective stabilizes registration in the presence of major topological violations and substantially increases registration accuracy. Our final model improves registration precision on the challenging FAUST dataset by a factor of 3 relative to the state of the art. Figure 1 illustrates the results. The presented approach is simple and provides a natural baseline for future progress.

2. Background

Nonrigid shape registration was considered by Elad and Kimmel, who identified the utility of isometric embeddings [12]. They proposed matching two shapes by embedding them in an intermediate Euclidean space, such that geodesic distances are approximated by Euclidean ones. This reduces nonrigid registration to rigid registration in an intermediate space, albeit at a loss of accuracy. A number of other methods that use an intermediate space followed, notably the conformal mapping approach [24, 52, 19].

A more direct approach is to embed one surface directly into the other, without intermediate embeddings. This formulation was developed by Bronstein et al. [8], building on the work of Mémoli and Sapiro [28]. They defined the embedding objective in terms of the distance distortion induced by the mapping:

$$E(f) = \sum_{i,j} w_{i,j} \big| d_{\mathcal{S}}(s_i, s_j) - d_{\mathcal{T}}(f(s_i), f(s_j)) \big|^p. \quad (1)$$

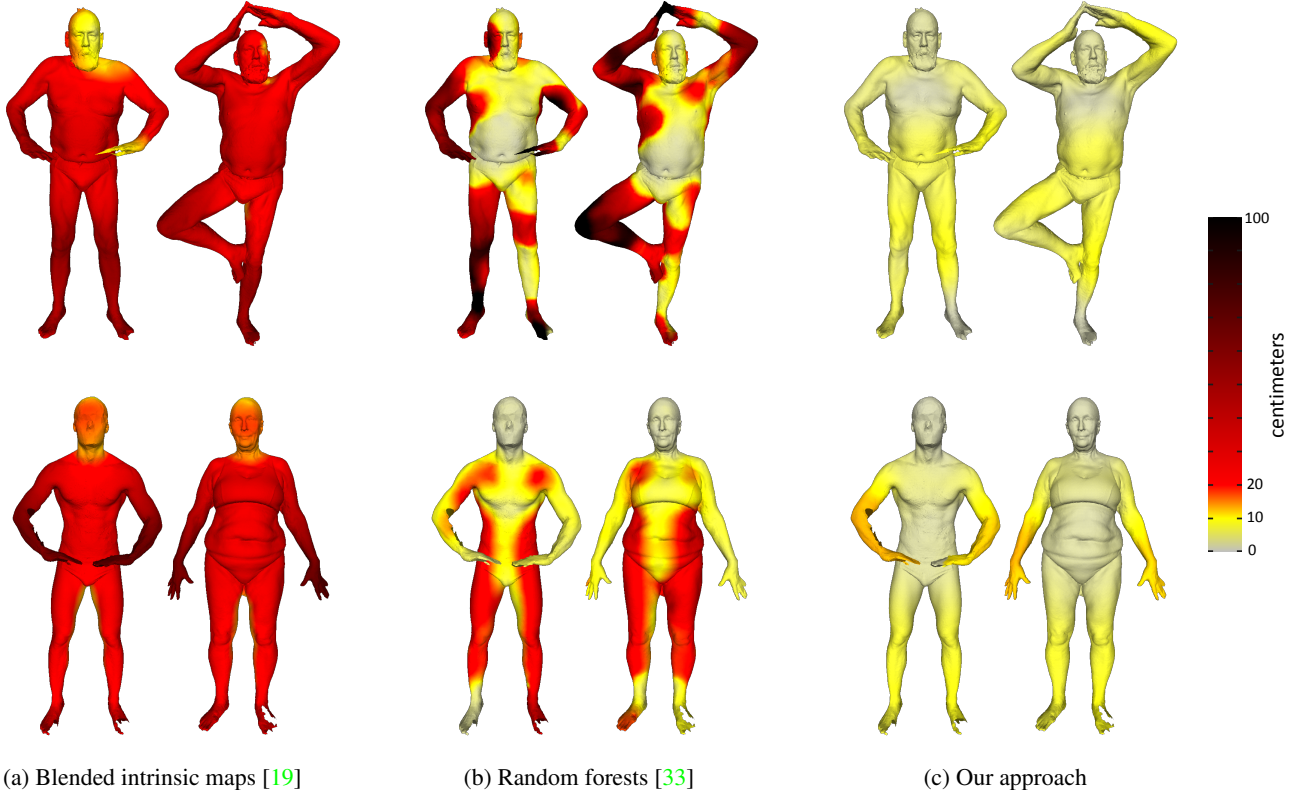| (a) Blended intrinsic maps [19] | (b) Random forests [33] | (c) Our approach |

Figure 1. Nonrigid registration of scans from the FAUST dataset [6]. Two scans of the same subject in the top row, two scans of different subjects below. Error magnitude is coded by color, as shown in the color bar on the right.

Here $d_{\mathcal{S}}(\cdot, \cdot)$ and $d_{\mathcal{T}}(\cdot, \cdot)$ are the geodesic distance functions on $\mathcal{S}$ and $\mathcal{T}$, $S = \{s_i\}_{i=1}^n$ is a fixed discrete sampling of $\mathcal{S}$, and $\{w_{i,j}\}$ are weights that can be used to disable some of the distortion terms. Optimization of (1) is called generalized multidimensional scaling (GMDS).

Objective (1) is in general highly non-convex and non-differentiable. While multiresolution schemes and various search techniques have been proposed, GMDS is generally not considered viable without a good initialization [7, 31, 36, 38]. Our work expresses the GMDS objective as a Markov random field (MRF) and leverages powerful methods developed for MRF optimization over the last decade. The objective admits a linear programming relaxation that can be efficiently solved without an initialization and yields global solutions.

Many approaches guide the registration by computing local shape descriptors [11, 29, 32, 34, 33]. While this strategy has been successful for rigid matching [1], its application to nonrigid registration is complicated by the requirement that the descriptors themselves be invariant to deformations. Many intrinsic descriptors have been proposed [35, 42, 3, 25, 50]. The downside of their invariance to isometric deformations is their sensitivity to gross topological inconsistencies. Such topological violations commonly occur in practice, for example due to self-contact [6]. For this

reason, we do not use keypoints, segments, or shape descriptors. Our work demonstrates that a direct approach that simply optimizes an isometric distortion objective substantially outperforms descriptor-based techniques.

The geodesic distance is itself sensitive to topological inconsistencies. To stabilize the registration, Bronstein et al. [10] propose to replace the geodesic metric by the diffusion distance. Our work is complementary: our approach is agnostic to the distance metric and can be used with the diffusion distance if desired. Yet this is not sufficient in itself. The diffusion distance is robust to some degree of topological noise, but not to large-scale topological violations. We therefore introduce robustness at a higher level, by replacing the L^p norm in (1) with a robust truncated norm. This is motivated by a long-standing connection between MRF modeling and robust statistics [4] and is supported by the optimization strategy. While the plateaus of a truncated norm can stymie gradient-based optimization, they are handled well by our approach.

Windheuser et al. [49] share our goal of a convex optimization approach to nonrigid matching and describe a linear programming relaxation for this problem. However, their formulation is substantially different: it operates on diffeomorphisms and cannot handle topological inconsistencies.

The early work of Anguelov et al. [2], which predates many developments covered earlier in this section, identified the utility of geodesic distance preservation and formulated the nonrigid registration problem in terms of Markov network optimization. However, their formulation is elaborate, incorporating several types of potentials such as local surface signatures in addition to distance preservation, and their distance potentials are binary. In contrast, we show that a simple continuous embedding objective coupled with modern MRF optimization machinery is sufficient.

A number of nonrigid registration pipelines reformulate the objective as an integer quadratic program, which can then be approximately solved using integer programming techniques or by relaxation [11, 31, 32, 34, 47]. These formulations commonly suffer from extremely high computational complexity, which restricts their applicability to a small number of feature points [11, 31, 47]. In particular, Wang et al. [47] use an integer quadratic programming formulation and optimize it using the dual decomposition algorithm of Torresani et al. [44]. This work is related to ours in that the dual decomposition framework also solves a convex relaxation. However, our algorithms are substantially different. Wang et al. rely on an algorithm that computes maximal flows in a graph that has $O(n^2)$ vertices and $O(n^4)$ edges, resulting in computational complexity $O(n^8)$. While the authors discuss a heuristic multiresolution scheme that independently solves small local subproblems, the applicability of the approach is restricted to matching a few points at a time and the authors only explore sparse keypoint matching. In contrast, we do not use a quadratic programming formulation and employ a highly effective optimization strategy with total complexity $O(n^4)$. This allows joint global optimization on dense sample sets, enabling high-accuracy dense registration of challenging real-world data.

3. Optimization

The embedding objective (1) corresponds to a continuous Markov random field with pairwise potentials [5]. To apply optimization techniques developed for discrete Markov random fields, we discretize the label space. Let $T = \{t_j\}_{j=1}^m$ be a set of samples that densely cover $\mathcal{T}$. Let L be a set of m labels, such that each label $l \in L$ corresponds to a distinct sample $t_l \in T$. We seek a labeling $\mathbf{l} : S \to L$. Let $l_i = \mathbf{l}(s_i)$ denote the label assigned to s_i by $\mathbf{l}$. The labeling objective is

$$\underset{\mathbf{l}}{\text{minimize}} \sum_{i,j} w_{i,j} \big| d_{\mathcal{S}}(s_i, s_j) - d_{\mathcal{T}}(t_{l_i}, t_{l_j}) \big|^p. \quad (2)$$

Each labeling $\mathbf{l}$ specifies a correspondence between the sample sets S and T. This correspondence can then be upsampled to a high-resolution mapping from $\mathcal{S}$ to $\mathcal{T}$ as described in Section 5.

Objective (2) corresponds to a discrete MRF with pairwise potentials. Using common MRF notation, we seek a labeling $\mathbf{l}^\star$ that minimizes an energy function:

$$\mathbf{l}^\star = \mathbf{l}^\star(\boldsymbol{\theta}) = \underset{\mathbf{l}}{\arg\min}\, E(\mathbf{l}; \boldsymbol{\theta}),$$

where $\boldsymbol{\theta}$ is the set of potentials:

$$E(\mathbf{l}; \boldsymbol{\theta}) = \sum_{i,j} \theta_{ij}(l_i, l_j), \quad (3)$$

$$\theta_{ij}(l_i, l_j) = w_{i,j} \big| d_{\mathcal{S}}(s_i, s_j) - d_{\mathcal{T}}(t_{l_i}, t_{l_j}) \big|^p.$$

The MRF optimization problem is NP-hard in general. A natural approach to deriving approximate algorithms is to represent the problem as an integer linear program and relax the integer constraints to obtain a linear program [48]. This approach has been a rich source of elegant MRF optimization algorithms. The LP relaxation is

$$\underset{\mathbf{0} \le \mathbf{x} \le \mathbf{1}}{\text{minimize}} \quad \sum_{i,j} \sum_{a,b} \theta_{ij}(a, b) x_{ij}^{ab}$$

$$\text{s.t.} \quad \sum_b x_{ij}^{ab} = x_i^a, \quad \forall i, j \in S,\ \forall a \in L$$

$$\sum_a x_i^a = 1, \qquad \forall i \in S \quad (4)$$

Here $\{x_i^a\}$ and $\{x_{ij}^{ab}\}$ are auxiliary variables that specify a distribution over the space of labelings. The first set of constraints imply that correspondences between pairs of samples must be consistent with correspondences between individual samples.

While standard LP solvers can be used to solve (4), they do not take advantage of the structure of the problem and do not scale well. Over the past decade, a variety of special-purpose MRF optimization algorithms have been developed that can effectively optimize large models with hundreds of thousands of potentials and hundreds of labels. Many of the most successful ones can be interpreted in terms of the dual of (4). This dual corresponds to an optimization problem on the potentials $\boldsymbol{\theta}$. The dual LP is

$$\underset{\boldsymbol{\delta}}{\text{maximize}} \quad \sum_i \min_a \tilde{\theta}_i^a + \sum_{i,j} \min_{a,b} \tilde{\theta}_{ij}^{ab},$$

$$\text{where} \qquad \tilde{\theta}_i^a = \sum_j \delta_{ji}^a \quad (5)$$

$$\tilde{\theta}_{ij}^{ab} = \theta_{ij}(a, b) - \delta_{ji}^a - \delta_{ij}^b$$

The variables $\{\delta_{ij}^a\}$ have a natural interpretation as messages. Optimization of the dual LP yields revised potentials $\tilde{\boldsymbol{\theta}}$, which can be used to obtain an optimized configuration $\tilde{\mathbf{l}}$ [45, 48, 41].

We use tree-reweighted message passing (TRW) to optimize objective (3) and its variants [45, 20]. TRW is one of

the most significant families of algorithms for optimizing Markov random fields. The basic TRW algorithms, their generalizations, and their relationship to other inference algorithms have been extensively analyzed [48, 46, 27, 21]. This interest was motivated in part by the effectiveness of certain forms of TRW in practice. In particular, the convergent sequential algorithm known as TRW-S [20] commonly exhibits both high accuracy and rapid convergence, yielding state-of-the-art results on large energy minimization problems [51, 43, 15]. We have evaluated TRW-S alongside a number of alternative algorithms, including dual decomposition with bundle methods [16], alternating directions dual decomposition [26], adaptive diminishing smoothing [39], and MPLP [40]. All results reported in this paper were obtained using TRW-S.

4. Objective

The optimization approach described in Section 3 does not assume that the distortion terms are convex, differentiable, or have an informative gradient. It is thus broadly applicable and can be used to optimize robust forms of the embedding objective. This significantly increases accuracy on challenging real-world data.

To motivate the revised objective, consider a didactic example of self-contact, shown in Figure 2. In this example, $\mathcal{S}$ and $\mathcal{T}$ are two configurations of a three-bar linkage, such that the bars all have length d. $\mathcal{S}$ is open and $\mathcal{T}$ is closed. Consider a pair of points $s_i, s_j \in \mathcal{S}$. Let $x = d_{\mathcal{S}}(s_i, s_j)$ and $y = d_{\mathcal{T}}(f^*(s_i), f^*(s_j))$, where $f^* : \mathcal{S} \to \mathcal{T}$ is a ground-truth mapping that registers individual bars. It is easy to see that $y = \min(x, 3d - x)$. Thus the distortion $|x - y|^p$ can be as high as $(3d)^p$. This highlights one of the repercussions of topological violations: they can trigger high penalties that distort the cost landscape.
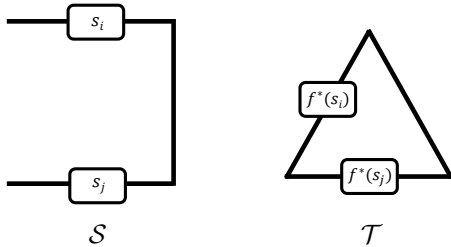

Figure 2. Without robust penalties, self-contact can significantly distort the cost landscape.

To bound penalties imposed on good mappings due to topological violations, we replace the unbounded L^p norm with a robust truncated penalty. (We use truncated L^1.) Furthermore, note that in our didactic example the distortion is non-zero only when $x > 3d/2$; in general, high spurious penalties are only imposed if at least one of the distances, either on $\mathcal{S}$ or on $\mathcal{T}$, is high. We thus additionally apply a Laplace weight to attenuate the contribution of

long-distance pairs to the objective. Our combined penalty function is

$$\rho(x, y) = \exp\left(-\frac{\min(x, y)}{b}\right) \min\left(|x - y|, \tau\right),$$

where b and τ are global parameters, set to 0.1m and 0.2m in our implementation.

We also add a weak extrinsic unary term into the objective to disambiguate intrinsic symmetries [9]. The complete objective is

$$
\begin{aligned}
E(\mathbf{l}) \;=\; & \sum_{i,j=1}^{n} \rho\big(d_{\mathcal{S}}(s_i, s_j), d_{\mathcal{T}}(t_{l_i}, t_{l_j})\big) \\
& + \; \lambda \sum_{i} \|s_i - t_{l_i}\|.
\end{aligned}
\tag{6}
$$

The weight λ balances the strength of the pairwise and unary terms, which have different cardinality.

Objective (6) is a pairwise MRF objective and can be optimized as described in Section 3. For clarity, the LPs in Section 3 were given without unary terms; these can be added without difficulty.

5. Implementation

Preprocessing. Our approach does not assume that the input models are watertight. To simplify the computation of geodesic distances, we begin by applying Poisson reconstruction to the meshes [17]. The two surfaces are covered by sample sets S and T using farthest point sampling [13]. Geodesic distances between all pairs of samples in S and all pairs of samples in T are then precomputed.

Global optimization. The approach described in Section 3 can be used to globally optimize the registration objective for hundreds of samples on each surface on a typical workstation. For $n = 100$ and $m = 400$, the average running time of our implementation on a pair of models from the FAUST dataset is 20 seconds on a workstation with 16GB of RAM and an Intel i7-4960X CPU clocked at 3.60GHz. As shown in Section 6, global optimization of our objective at this sampling resolution already yields state-of-the-art registration accuracy, without any coarse-to-fine processing. Since the primal solution produced by TRW-S is related to the ordering of nodes, we permute the ordering 10 times and choose the best primal solution.

Upsamping. If desired, we can upsample the mapping generated by the global optimization as follows. Consider sample sets $\widehat{S} \supset S$ and $\widehat{T} \supset T$, such that $|\widehat{S}|, |\widehat{T}| \gg n, m$. Since the mapping $\mathbf{l} : S \to T$ is already well-localized, it

can be upsampled to a mapping $\hat{l} : \widehat{S} \to \widehat{T}$ by a lower-complexity algorithm. Specifically, we can produce a high-resolution mapping $\hat{l}$ by optimizing a restricted distance preservation objective:

$$\arg\min_{\hat{l}} \sum_{i=1}^{w} \sum_{j=1}^{n} \rho\big(d_{\mathcal{S}}(s_i, s_j), d_{\mathcal{T}}(t_{\hat{l}_i}, t_{l_j})\big), \quad (7)$$

where $w = |\widehat{S}|$. This decomposes into independent objectives for each point in $\widehat{S} \setminus S$: each can thus be placed independently.

Refinement. We can optionally refine an upsampled mapping $\hat{l}$ by optimizing the complete distance preservation objective that links all pairs of points:

$$\arg\min_{\hat{l}} \sum_{i,j=1}^{w} \rho\big(d_{\mathcal{S}}(s_i, s_j), d_{\mathcal{T}}(t_{\hat{l}_i}, t_{\hat{l}_j})\big). \quad (8)$$

To refine a mapping at high resolution, we use fusion moves [23]. Given a current labeling $\hat{l}$, we generate a proposal l' as follows. For each point $s_i \in \widehat{S}$, we set $t_{l'_i}$ to some point $t \in \widehat{T}$ with probability

$$\exp\left(-\frac{d_{\mathcal{T}}(t_{\hat{l}_i}, t)}{b}\right) / \sum_i \exp\left(-\frac{d_{\mathcal{T}}(t_{\hat{l}_i}, t)}{b}\right).$$

This proposal l' is then fused with the current labeling $\hat{l}$ using QPBO [22, 23]. Our implementation optionally performs fusion moves at resolution $w = 5000$, after which the result is upsampled to the full resolution of the input meshes.

6. Evaluation

6.1. Experimental setup

We conduct extensive experiments on the FAUST dataset. This is a recent dataset of real 3D scans with carefully collected ground-truth correspondences. FAUST has been shown to be much more challenging than earlier datasets such as TOSCA and SCAPE [6].

The FAUST dataset is composed of a training set and a test set. Ground-truth data for the training set is publicly available and allows thorough comparison of different approaches, computation of multiple accuracy measures, and visualization of results. Ground-truth data for the test set is not publicly available and the only way to evaluate an approach is to submit its results to the FAUST server. Since we were not at liberty to submit other groups' results to the server, we focus on the publicly available FAUST training set for detailed comparison to prior work.

The FAUST training set provides 100 scans of 10 human subjects: 10 scans per subject, each in a different pose. We

evaluate on pairs of scans of the same subject in different poses (intra-subject pairs) and on pairs of scans of different subjects (inter-subject pairs). We randomly generate 50 intra-subject pairs such that each scan in the training set is used in exactly one intra-subject pair. We also randomly generate 50 inter-subject pairs such that each scan in the training set is used in exactly one inter-subject pair.

We compare the presented approach to a large number of prior methods for which we could obtain implementations online or by contacting the authors: GMDS [8], Möbius voting [24], blended intrinsic maps (BIM) [19], coarse-to-fine matching (C2F) [36], the EM algorithm [37], coarse-to-fine matching with symmetric flips (C2FSym) [38], sparse modeling (SM) [30], elastic net constraints (ENC) [34], and random forests (RF) [33]. Each method is evaluated on all intra-subject pairs and all inter-subject pairs.

Many of the prior methods require watertight meshes as input. Following Bogo et al. [6], we used Poisson reconstruction for this purpose [17]. The meshes were cleaned up by automatically removing small disconnected components. This enabled us to successfully evaluate GMDS, Möbius, and BIM on all pairs of models, in contrast to the incomplete evaluation reported in [6]. Other methods were also successfully evaluated on all pairs, with the exception of C2F and EM, which crashed on 7 out of 50 intra-subject pairs and 7 out of 50 inter-subject pairs.

A number of prior methods do not scale to full-resolution models. We run these on simplified meshes with approximately 10,000 vertices. Some of the methods, such as GMDS, Möbius, ENC, and EM, produce sparse correspondences for roughly 100 to 200 samples. To maximize the reported accuracy for these prior techniques, we computed error measures only for the sparse correspondences they produced. For methods that do provide complete dense correspondences, including ours, accuracy is reported for all correspondences.

RF is a model-based technique since it involves learning on the dataset, but we report its performance alongside other methods for completeness. To maximize the performance of this approach, we trained a dedicated random forest (15 trees) for each tested pair. For each pair of scans in poses i and j, we trained a forest on all scans in the training set (including scans of the same subjects) that are in different poses ($\neq i, j$): 80 training scans for each forest when $i \neq j$, 90 training scans if $i = j$ in an inter-subject pair. This tests the generalization of the method to new poses, in analogy to the partition of the FAUST dataset into training and test sets, which are composed of scans of the same ten subjects in different sets of poses.

6.2. Results

Results are summarized in Figures 3 and 4. Figure 3 (top) shows the cumulative error distributions for the eval-

(a) Cumulative error distribution, intra-subject

(b) Cumulative error distribution, inter-subject

(c) Average error for each intra-subject pair

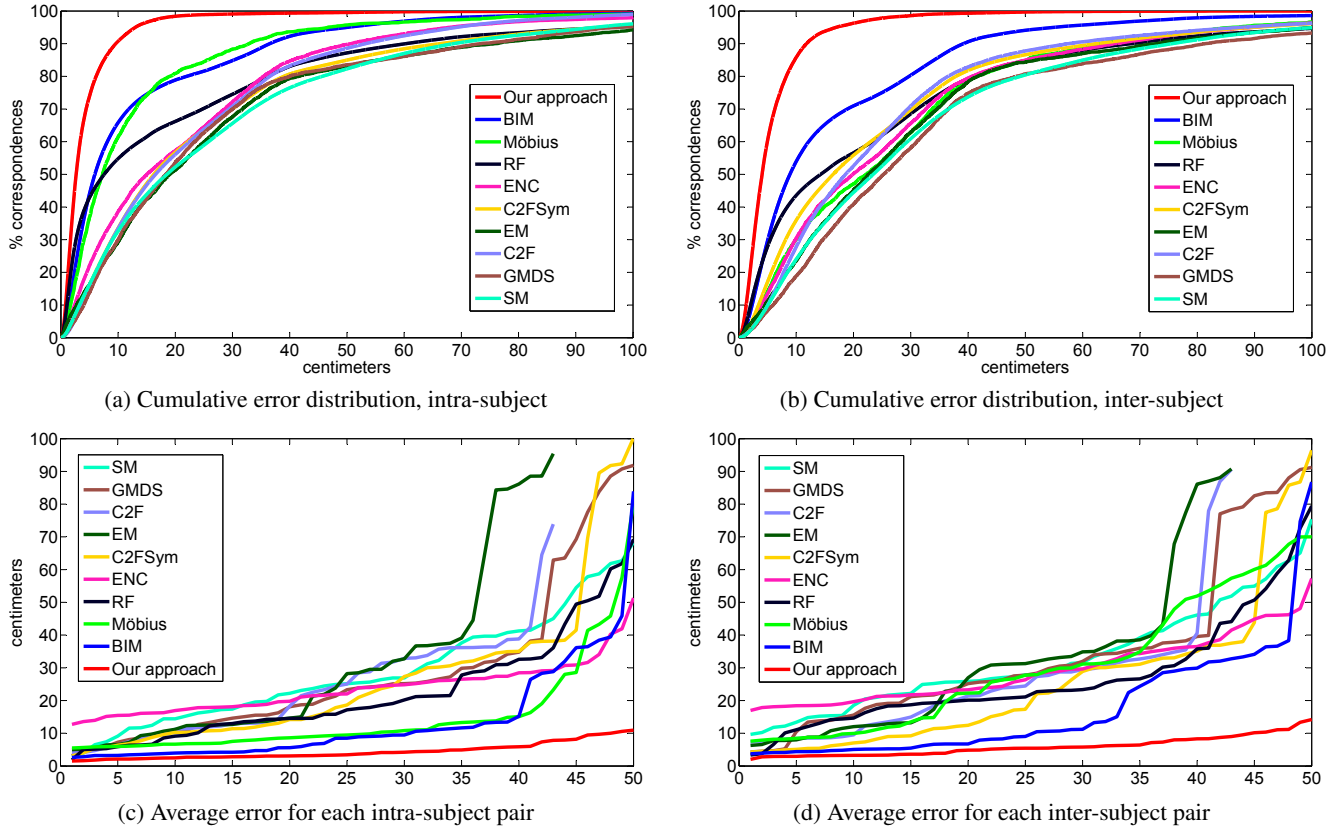(d) Average error for each inter-subject pair

Figure 3. Evaluation on the FAUST dataset. Intra-subject pairs on the left, inter-subject pairs on the right. Top: Cumulative error distributions for the evaluated methods, in centimeters (higher is better). Bottom: Average error for each of the tested pairs, in centimeters, sorted by magnitude (lower is better); the tested pairs are sorted independently for each method, according to the accuracy achieved by the method on each pair: for example, pair #50 is not the same for different methods.

method	AE (cm)	worst AE	10cm-recall
GMDS [8]	28.98	91.84	0.300
Möbius [24]	14.99	80.40	0.614
BIM [19]	13.60	83.90	0.658
C2F [36]	23.63	73.89	0.333
EM [37]	30.11	95.42	0.293
C2FSym [38]	26.87	100.23	0.335
SM [30]	28.81	68.42	0.326
ENC [34]	23.60	51.32	0.385
RF [33]	22.26	69.26	0.548
Our approach	**4.49**	**10.96**	**0.907**

Accuracy on intra-subject pairs

method	AE (cm)	worst AE	10cm-recall
GMDS [8]	35.06	91.21	0.188
Möbius [24]	30.58	70.02	0.300
BIM [19]	17.36	86.76	0.539
C2F [36]	25.51	90.62	0.277
EM [37]	31.25	90.74	0.235
C2FSym [38]	25.89	96.46	0.359
SM [30]	32.66	75.38	0.240
ENC [34]	29.29	57.28	0.303
RF [33]	26.92	79.43	0.435
Our approach	**5.95**	**14.18**	**0.858**

Accuracy on inter-subject pairs

Figure 4. Evaluation on the FAUST dataset. Intra-subject pairs on the left, inter-subject pairs on the right. For each technique, each table reports the average error on all pairs (AE, in centimeters), average error on the worst pair for this technique (worst AE, in centimeters), and the fraction of generated correspondences that are within 10cm of the ground truth (10cm-recall). Our technique reduces the average error by a factor of 3 over the best-performing prior approach on intra-subject pairs, and by a factor of 2.9 on inter-subject pairs. The worst AE is reduced by a factor of 4.7 on intra-subject pairs and by a factor of 4 on inter-subject pairs.

uated methods, in centimeters, aggregated over all tested pairs (intra-subject on the left, inter-subject on the right). For example, the left plot shows that 73.85% of the surface area of the intra-subject pairs was registered to a 5cm accuracy by the presented approach, while the best-performing prior approach (BIM) achieved this accuracy for only 43.99% of the surface area. Figure 3 (bottom) shows the average error for each tested pair, in centimeters, sorted by magnitude. For example, the left plot shows that the highest average error on the worst intra-subject pair for our technique was 10.96cm, whereas all other techniques had average error of at least 51.32cm on some of the tested pairs.

Figure 4 reports the average error of each technique across all tested pairs, the average error on the worst tested pair for each technique, and the 10cm-recall of each technique (i.e., the fraction of correspondences across all tested pairs that were within 10cm of the ground truth). Our technique reduces the average error by a factor of 3 over the best-performing prior approach on intra-subject pairs, and by a factor of 2.9 on inter-subject pairs. The worst average error is reduced by a factor of 4.7 on intra-subject pairs and by a factor of 4 on inter-subject pairs. The tested pairs that suffered the worst average errors produced by our approach are shown in Figure 5.

We have also evaluated the contribution of different elements of the presented approach. If we use the presented optimization approach to optimize the embedding objective with a standard L^1 penalty on distance distortion, the average error is 7.97cm on intra-subject pairs and 8.82cm on inter-subject pairs. This is already considerably more accurate than the results produced by the best-performing prior method. When we replace the L^1 norm with the robust penalty ρ described in Section 4, the average error drops to 5.55cm intra-subject and 6.88cm inter-subject. These results are for single-resolution global optimization. If we add refinement with fusion moves at high resolution as described in Section 5, the average error drops further, to 4.49cm intra-subject and 5.95cm inter-subject. This is illustrated in Figure 6.

Finally, we have also submitted our results to the FAUST server for evaluation on the test set. The average error reported by the server for our model-free approach is 4.86cm on the intra-subject challenge and 8.30cm on the inter-subject challenge.

7. Conclusion

We presented an approach to robust nonrigid registration of 3D surfaces. The approach optimizes a clear objective using well-understood techniques. It yields state-of-the-art results on challenging real-world data, outperforming a large body of prior work by a multiplicative factor. We hope that the simplicity of the presented approach will stimulate further advances in this area. Our results are far from per-
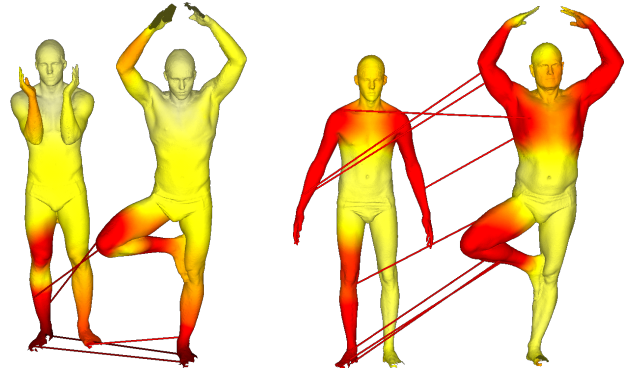

Figure 5. Pairs of scans that set the worst average errors reported for our approach in Figure 4. Worst intra-subject pair on the left (AE 10.96cm), worst inter-subject pair on the right (AE 14.18cm). The color coding is defined in Figure 1.

fect. We hope that future work will build on the presented ideas to advance the theory and practice of surface registration. The ideas developed in this paper may also be useful in the context of partial surface registration and in joint analysis of non-isometric shapes [14, 18].

References

[1] A. Aldoma, Z. Marton, F. Tombari, W. Wohlkinger, C. Potthast, B. Zeisl, R. B. Rusu, S. Gedikli, and M. Vincze. Point cloud library: Three-dimensional object recognition and 6 DoF pose estimation. *IEEE Robotics and Automation Magazine*, 19(3), 2012. 2

[2] D. Anguelov, P. Srinivasan, H. Pang, D. Koller, S. Thrun, and J. Davis. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *NIPS*, 2004. 1, 3

[3] M. Aubry, U. Schlickewei, and D. Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *ICCV Workshops*, 2011. 2

[4] M. J. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *IJCV*, 19(1), 1996. 2

[5] A. Blake, P. Kohli, and C. Rother. *Markov Random Fields for Vision and Image Processing*. MIT Press, 2011. 3

[6] F. Bogo, J. Romero, M. Loper, and M. J. Black. FAUST: dataset and evaluation for 3D mesh registration. In *CVPR*, 2014. 1, 2, 5

[7] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Efficient computation of isometry-invariant distances between surfaces. *SIAM J. Scientific Computing*, 28(5), 2006. 2

[8] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: A framework for isometry-invariant partial surface matching. *Proceedings of the National Academy of Sciences*, 103(5), 2006. 1, 5, 6

[9] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Topology-invariant similarity of nonrigid shapes. *IJCV*, 81(3), 2009. 4

[10] A. M. Bronstein, M. M. Bronstein, R. Kimmel, M. Mahmoudi, and G. Sapiro. A Gromov-Hausdorff framework with

(a) Objective with L^1 penalty   (b) Objective with robust penalty   (c) Refinement with fusion moves
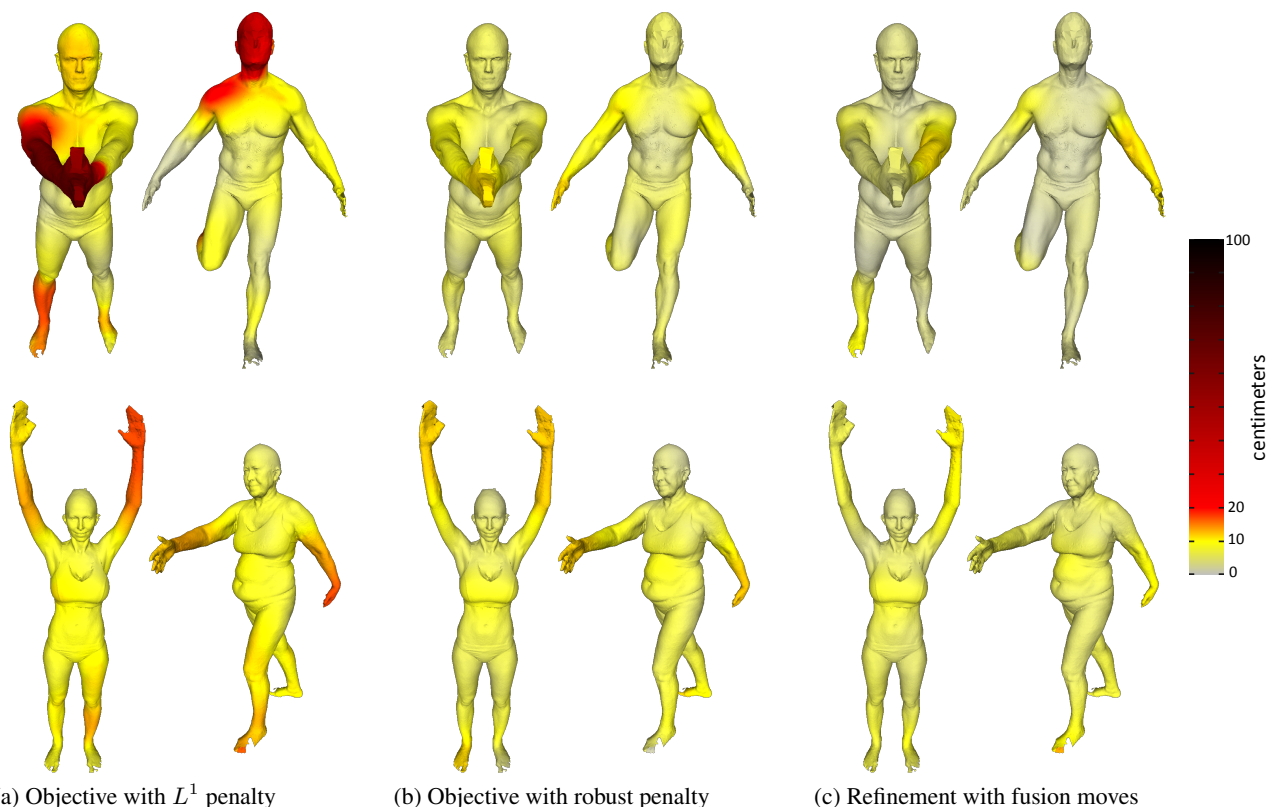
Figure 6. Different elements of the presented approach. An intra-subject pair in the top row, an inter-subject pair below. (a) Optimization of the L^1 embedding objective yields 9.16cm and 9.02cm AE on the presented pairs. (b) Replacing the L^1 norm with the robust penalty ρ reduces AE to 5.73cm and 5.81cm. (c) Refinement with fusion moves at high resolution reduces AE further to 3.92cm and 5.26cm.

diffusion geometry for topologically-robust non-rigid shape matching. *IJCV*, 89(2-3), 2010. 2

[11] A. Dubrovina and R. Kimmel. Approximately isometric shape correspondence by matching pointwise spectral features and global geodesic structures. *Advances in Adaptive Data Analysis*, 3(1-2), 2011. 2, 3

[12] A. Elad and R. Kimmel. On bending invariant signatures for surfaces. *PAMI*, 25(10), 2003. 1

[13] T. F. Gonzalez. Clustering to minimize the maximum inter-cluster distance. *Theoretical Computer Science*, 38, 1985. 4

[14] Q. Huang, V. Koltun, and L. J. Guibas. Joint shape segmentation with linear programming. *ACM Trans. Graph.*, 30(6), 2011. 7

[15] J. H. Kappes, B. Andres, F. A. Hamprecht, C. Schnörr, S. Nowozin, D. Batra, S. Kim, B. X. Kausler, T. Kröger, J. Lellmann, N. Komodakis, B. Savchynskyy, and C. Rother. A comparative study of modern inference techniques for structured discrete energy minimization problems. *IJCV*, 2014. 4

[16] J. H. Kappes, B. Savchynskyy, and C. Schnörr. A bundle approach to efficient MAP-inference by Lagrangian relaxation. In *CVPR*, 2012. 4

[17] M. M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Symposium on Geometry Processing*, 2006. 4, 5

[18] V. G. Kim, W. Li, N. J. Mitra, S. DiVerdi, and T. A. Funkhouser. Exploring collections of 3D models using fuzzy correspondences. *ACM Trans. Graph.*, 31(4), 2012. 7

[19] V. G. Kim, Y. Lipman, and T. A. Funkhouser. Blended intrinsic maps. *ACM Trans. Graph.*, 30(4), 2011. 1, 2, 5, 6

[20] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *PAMI*, 28(10), 2006. 3, 4

[21] V. Kolmogorov. A new look at reweighted message passing. *PAMI*, 37(5), 2015. 4

[22] V. Kolmogorov and C. Rother. Minimizing nonsubmodular functions with graph cuts – a review. *PAMI*, 29(7), 2007. 5

[23] V. S. Lempitsky, C. Rother, S. Roth, and A. Blake. Fusion moves for Markov random field optimization. *PAMI*, 32(8), 2010. 5

[24] Y. Lipman and T. A. Funkhouser. Möbius voting for surface correspondence. *ACM Trans. Graph.*, 28(3), 2009. 1, 5, 6

[25] R. Litman and A. M. Bronstein. Learning spectral descriptors for deformable shape correspondence. *PAMI*, 36(1), 2014. 2

[26] A. F. T. Martins, M. A. T. Figueiredo, P. M. Q. Aguiar, N. A. Smith, and E. P. Xing. AD3: Alternating directions dual decomposition for MAP inference in graphical models. *Journal of Machine Learning Research*, 16, 2015. 4

[27] T. Meltzer, A. Globerson, and Y. Weiss. Convergent message passing algorithms – a unifying view. In *UAI*, 2009. 4

[28] F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 5(3), 2005. 1

[29] M. Ovsjanikov, M. Ben-Chen, J. Solomon, A. Butscher, and L. J. Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Trans. Graph.*, 31(4), 2012. 2

[30] J. Pokrass, A. M. Bronstein, M. M. Bronstein, P. Sprechmann, and G. Sapiro. Sparse modeling of intrinsic correspondences. *Comput. Graph. Forum*, 32(2), 2013. 5, 6

[31] D. Raviv, A. Dubrovina, and R. Kimmel. Hierarchical framework for shape correspondence. *Numerical Mathematics: Theory, Methods and Applications*, 6, 2013. 2, 3

[32] E. Rodolà, A. M. Bronstein, A. Albarelli, F. Bergamasco, and A. Torsello. A game-theoretic approach to deformable shape matching. In *CVPR*, 2012. 2, 3

[33] E. Rodolà, S. R. Bulò, T. Windheuser, M. Vestner, and D. Cremers. Dense non-rigid shape correspondence using random forests. In *CVPR*, 2014. 2, 5, 6

[34] E. Rodolà, A. Torsello, T. Harada, Y. Kuniyoshi, and D. Cremers. Elastic net constraints for shape matching. In *ICCV*, 2013. 2, 3, 5, 6

[35] R. M. Rustamov. Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In *Symposium on Geometry Processing*, 2007. 2

[36] Y. Sahillioglu and Y. Yemez. Coarse-to-fine combinatorial matching for dense isometric shape correspondence. *Comput. Graph. Forum*, 30(5), 2011. 2, 5, 6

[37] Y. Sahillioglu and Y. Yemez. Minimum-distortion isometric shape correspondence using EM algorithm. *PAMI*, 34(11), 2012. 5, 6

[38] Y. Sahillioglu and Y. Yemez. Coarse-to-fine isometric shape correspondence by tracking symmetric flips. *Comput. Graph. Forum*, 32(1), 2013. 2, 5, 6

[39] B. Savchynskyy, S. Schmidt, J. H. Kappes, and C. Schnörr. Efficient MRF energy minimization via adaptive diminishing smoothing. In *UAI*, 2012. 4

[40] D. Sontag, D. K. Choe, and Y. Li. Efficiently searching for frustrated cycles in MAP inference. In *UAI*, 2012. 4

[41] D. Sontag, A. Globerson, and T. Jaakkola. Introduction to dual decomposition for inference. In S. Sra, S. Nowozin, and S. J. Wright, editors, *Optimization for Machine Learning*. MIT Press, 2011. 3

[42] J. Sun, M. Ovsjanikov, and L. J. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Comput. Graph. Forum*, 28(5), 2009. 2

[43] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. F. Tappen, and C. Rother. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. *PAMI*, 30(6), 2008. 4

[44] L. Torresani, V. Kolmogorov, and C. Rother. A dual decomposition approach to feature correspondence. *PAMI*, 35(2), 2013. 3

[45] M. J. Wainwright, T. Jaakkola, and A. S. Willsky. MAP estimation via agreement on trees: message-passing and linear programming. *IEEE Transactions on Information Theory*, 51(11), 2005. 3

[46] M. J. Wainwright and M. I. Jordan. Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning*, 1(1-2), 2008. 4

[47] C. Wang, M. M. Bronstein, A. M. Bronstein, and N. Paragios. Discrete minimum distortion correspondence problems for non-rigid shape matching. In *Scale Space and Variational Methods in Computer Vision*, 2011. 3

[48] T. Werner. A linear programming approach to max-sum problem: A review. *PAMI*, 29(7), 2007. 3, 4

[49] T. Windheuser, U. Schlickewei, F. R. Schmidt, and D. Cremers. Geometrically consistent elastic matching of 3D shapes: A linear programming solution. In *ICCV*, 2011. 2

[50] T. Windheuser, M. Vestner, E. Rodolà, R. Triebel, and D. Cremers. Optimal intrinsic descriptors for non-rigid shape analysis. In *BMVC*, 2014. 2

[51] C. Yanover, T. Meltzer, and Y. Weiss. Linear programming relaxations and belief propagation – an empirical study. *Journal of Machine Learning Research*, 7, 2006. 4

[52] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios. Dense non-rigid surface registration using high-order graph matching. In *CVPR*, 2010. 1