

DiscoMatch: Fast Discrete Optimisation for Geometrically Consistent 3D Shape Matching

Paul Roetzer^{*1}, Ahmed Abbas^{*2}, Dongliang Cao¹, Florian Bernard¹, and Paul Swoboda³

¹ University of Bonn

² MPI for Informatics, Saarland Informatics Campus

³ Heinrich-Heine University Düsseldorf

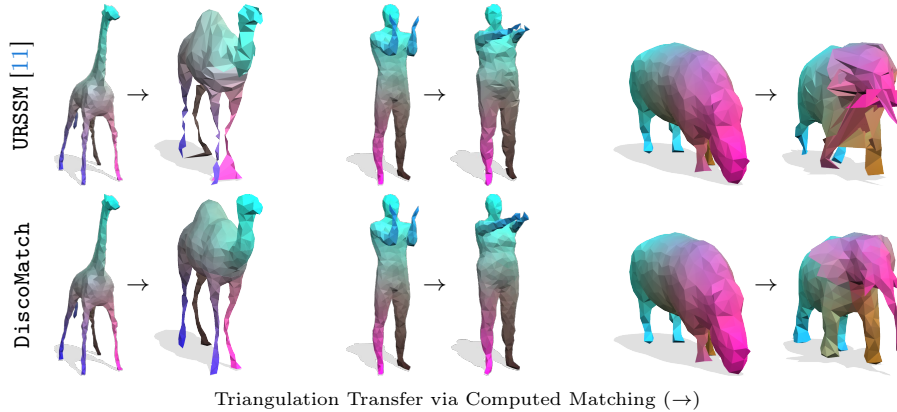


Fig. 1: We efficiently solve a large integer linear program (ILP) to guarantee geometric consistency of matchings between pairs of 3D shapes. Our matchings are smooth and more accurate than the recent state-of-the-art shape matching approach URSSM [11]. Improved solution quality can be seen when transferring triangulation between shapes.

Abstract. In this work we propose to combine the advantages of learning-based and combinatorial formalisms for 3D shape matching. While learning-based methods lead to state-of-the-art matching performance, they do not ensure geometric consistency, so that obtained matchings are locally non-smooth. On the contrary, axiomatic, optimisation-based methods allow to take geometric consistency into account by explicitly constraining the space of valid matchings. However, existing axiomatic formalisms do not scale to practically relevant problem sizes, and require user input for the initialisation of non-convex optimisation problems. We work towards closing this gap by proposing a novel combinatorial solver that combines a unique set of favourable properties: our approach (i) is initialisation free, (ii) is massively parallelisable and powered by a quasi-Newton method, (iii) provides optimality gaps, and (iv) delivers improved matching quality with decreased runtime and globally optimal results for many instances.

Keywords: Shape matching · Discrete optimisation

^{*} These authors contributed equally to this work.

1 Introduction

The shape matching problem, i.e., finding correspondences between two non-rigidly deformed 3D shapes, lies at the heart of many visual computing tasks. Resulting correspondences are relevant for many downstream applications in computer vision (3D reconstruction, shape space modelling), graphics (texture transfer, shape interpolation) or medical image analysis (shape-based segmentation, navigation). Thus, shape matching has received a lot of attention in recent years and has been tackled by numerous variants of both axiomatic and learning-based methods.

Learning-based methods usually provide the best matching performance on the most challenging benchmarks [11]. However, a major downside is that they do not provide any guarantees about crucial properties of the final correspondences, such as *geometric consistency* of matchings: for triangular surface meshes, geometric consistency means that neighbouring triangles of one shape are consistently matched to neighbouring triangles of the other shape (analogous to diffeomorphisms in the continuous domain), cf. Fig. 2 left.

While axiomatic methods can – in principle – take geometric consistency into account, most existing shape matching approaches disregard it [5, 24, 41, 47, 58, 61, 64], mainly due to the resulting non-convexity of the optimisation problem or hard-to-solve constraints. Although there are some exceptions, respective formulations can typically be solved only for small instances, such as [68], or they rely on manual user initialisations to avoid poor local optima and to better constrain the problem [52, 54, 60].

We argue that geometric consistency is of crucial importance and it should be considered as an essential aspect for any shape matching pipeline. Hence, in this work we propose to combine expressive and powerful shape features obtained from a learning-based method with a novel combinatorial solver for the geometrically consistent shape matching formalism [68]. We summarise our main contributions as follows:⁴

- a. We propose a theoretically well-grounded solver for global shape matching that is able to quantify optimality gaps. We are empirically able to certify global optimality in 90% of the considered shape matching instances.
- b. Our solver extends the dual decomposition approach [1] and also addresses specific challenges posed by large shape matching instances. By doing so we demonstrate a 9 to 11 times faster convergence compared to [1] on shape matching problems. Additionally some components of our solver are generally applicable and yield 5 to 10 times runtime improvements over [1].
- c. We outperform the current state-of-the-art learning-based shape matching approach on most considered datasets w.r.t. matching accuracy and smoothness, while at the same time guaranteeing geometric consistency.

⁴ Our code is publicly available at <https://github.com/paul0noah/disco-match>.

2 Related Work

In the following we summarise works that are most relevant to our approach. For an extensive review of shape matching methods we refer to the surveys [48, 64].

Deep Shape Matching. With the rapid development of deep learning, many learning-based approaches are proposed for 3D shape matching, which can be categorised into supervised and unsupervised methods. In the context of supervised learning, some methods [32, 67] formulate shape matching as a classification problem, while others [20, 63] solve the problem based on non-rigid shape alignment. Nevertheless, supervised methods require ground-truth correspondences, which in practice are hard to obtain. Therefore, many unsupervised methods were introduced to get rid of the requirement of ground-truth correspondences. Notably, the functional map framework [41] is one of the most dominant ingredients to enable unsupervised learning.

Earlier works [22, 46] impose isometry regularisation in the spatial and spectral domain, respectively. Follow-up works [14, 55] use point-based networks [42, 62] and lead to better matching performance. Later, learnable implicit diffusion processes pushed the state of the art in shape matching for a broad range settings [57], including near-isometry [2], non-isometry [10, 13, 31], partiality [3], multi-shape matching [8] as well as multi-modal matching [9]. Meanwhile, instead of solely learning features, several works [26, 37] also attempted to learn the basis functions at the same time.

One important ingredient of our work is the utilisation of recent learned feature embeddings. While existing shape matching approaches that are purely learning-based do not provide any guarantees about geometric consistency (i.e., the smoothness of maps), we propose to combine learned embeddings with an axiomatic optimisation approach in order to guarantee geometric consistency via hard constraints.

Geometric Consistency. Several 3D shape matching methods provide geometrically consistent matchings by performing local optimisation on initial maps between two shapes. Continuous maps starting from user-defined correspondences are computed in [54]. Refinement of an initial matching in a geometrically consistent manner using heat diffusion in [56]. 3D shape matching is casted as image matching problem to obtain smooth maps via a smooth surface-to-image parameterisation in [34]. In [66] the so-called product manifold filter is used to find smooth maps via a linearisation of the quadratic assignment problem. Local map refinement of initial maps is done in [16] via denoising, or in [17] via minimising an elastic energy. Common triangulations between two shapes have been utilised in [51] or via intrinsic triangulation in [60]. Mapping of shapes to spheres and then optimising for a homeomorphism is done in [52].

While all of the above methods provide high quality and smooth maps they rely on initial correspondences which are non-trivial to find. Thus these methods can only be considered as *local* shape matching methods. Our approach allows to infer smooth maps (i.e., a discrete analogue to a diffeomorphism) between 3D shapes based on an efficient relaxation solver for integer programs.

Algorithms for Global Shape Matching. Practically relevant and globally optimal 3D shape matching is considered extremely challenging. Nevertheless, in certain cases respective problems are solvable efficiently: for example 2D to 2D matching via dynamic programming [12] analogous to dynamic time warping [49], or via graph cuts [50]; another example is contour-to-image matching via dynamic programming [18] or shortest paths [53]. Matching a 2D contour to a 3D mesh can also be phrased as shortest path problem [29, 44].

For the 3D to 3D matching scenario there exist several sparse matching methods [5, 19, 38], or the popular functional map [41] approach that operates in the spectral domain. Relaxations to bi-stochastic matrices for the QAP are used in [65] to map nearby points on one shape to nearby points on the other shape. Furthermore, an optimisation problem to find a dense discrete diffeomorphism between two 3D shapes is introduced in [68]. While this optimisation problem purely relies on linear constraints to ensure geometric consistency, it is nevertheless hard to solve since it optimises over a large number of binary variables. Thus, various strategies to address this problem have been proposed: a linear programming relaxation [68] combined with a coarse-to-fine scheme [69], as well as an approach using a combinatorial solver with a problem-specific primal heuristic [45]. Still, the mentioned methods do not scale to practically relevant resolutions and oftentimes cannot find good lower-bounds (which as a consequence results in bad primal solutions in many cases). We show that with our combinatorial optimiser we can find solutions much faster and often times find a globally optimal solution due to faster and better lower bound computation.

3 Background

We first introduce the integer linear programming (ILP) formulation [68] for geometrically consistent shape matching by ensuring neighbourhood consistency of matched elements. Further, we touch upon the method [30] (which solves large-scale ILPs based on smaller subproblems), since we build on it. Our notation is summarised in Tab. 1.

Symbol	Description	Symbol	Description
$\mathcal{M} = (V_{\mathcal{M}}, T_{\mathcal{M}})$	Triangle mesh	$x \in \{0, 1\}^{ P }$	Indicator representation of P
$V_{\mathcal{M}} \in \mathbb{R}^{ V_{\mathcal{M}} \times 3}$	Vertices of $\mathcal{M}$	$c \in \mathbb{R}_+^{ P }$	Cost vector
$T_{\mathcal{M}} \in \mathbb{N}^{ T_{\mathcal{M}} \times 3}$	Triangles of $\mathcal{M}$	$A^{\theta}x = \mathbf{0}$	Geo. consistency constraints
$F_{\mathcal{M}} \in \mathbb{R}^{ V_{\mathcal{M}} \times 128}$	Features of $\mathcal{M}$	$A^{\mathcal{M}}x = \mathbf{1}$	Projection on $\mathcal{M}$ constraints
$\mathcal{N} = (V_{\mathcal{N}}, T_{\mathcal{N}})$	Triangle mesh	$A^{\mathcal{N}}x = \mathbf{1}$	Projection on $\mathcal{N}$ constraints
...	(analog. to $\mathcal{M}$)	$\mathcal{I}, \mathcal{J}$	Set of vars., constraints resp.
P	Product space	$\mathcal{I}_j, (\mathcal{J}_i)$	Vars. of j in $\mathcal{J}$ (viceversa)

Table 1: Summary of the notation used in this paper.

3.1 3D Shape Matching as ILP

The matching formalism in [68] aims to find a *discrete diffeomorphism* between two triangle meshes $\mathcal{M}$ and $\mathcal{N}$ by solving a constrained binary integer linear program. The matching is found within the search space P (i.e. the product space) which consists of three types of potential correspondences between $\mathcal{M}$ and $\mathcal{N}$: triangle-triangle, triangle-edge and triangle-vertex, see Fig. 2 right. The latter

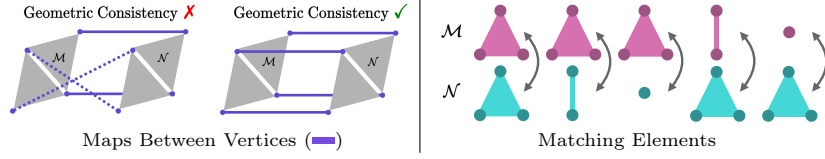


Fig. 2: *Left:* Geometric consistency is an essential property of correspondences between shapes, meaning that neighbouring elements in $\mathcal{M}$ are consistently matched to neighbouring elements in $\mathcal{N}$. *Right:* Matching elements of the approach by [68]: triangle-triangle, triangle-edge and triangle-vertex between shapes $\mathcal{M}$ and $\mathcal{N}$. At least one triangle is involved in every matching element.

two are needed to account for different triangulation as well as stretching and compression of shapes. Employing geometric consistency constraints $A^\partial x = \mathbf{0}$ ensures that neighbouring elements in $\mathcal{M}$ are matched only to neighbouring elements in $\mathcal{N}$, see Fig. 2 left. Additional projection constraints $A^\mathcal{M}x = \mathbf{1}$ and $A^\mathcal{N}x = \mathbf{1}$ ensure that each triangle in $T_\mathcal{M}$ and $T_\mathcal{N}$ respectively are matched exactly once and thus ensure surjectivity. A solution x fulfilling all constraints yields a discrete diffeomorphism between shapes $\mathcal{M}$ and $\mathcal{N}$ [68]. The full optimisation problem reads

$$\min_{x \in \{0,1\}^{|P|}} c^\top x \quad \text{s.t.} \quad \underbrace{\begin{bmatrix} A^\partial \\ A^\mathcal{M} \\ A^\mathcal{N} \end{bmatrix}}_{:=A} x = \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix}}_{:=b}. \quad (\text{P})$$

For a detailed derivation of (P) and more insights we refer readers to [45, 68, 69]. As mentioned, the solution of this optimisation problem yields a discrete diffeomorphism between two triangle meshes. This ensures geometric consistency. Since the problem is a difficult and large combinatorial optimisation problem, only heuristic solvers (cf. [45, 69]) exist to solve the problem for reasonable shape resolutions. We show that by combining deep feature embeddings and adapting a novel combinatorial solver, we can solve the problem significantly faster than before, while at the same time certifying global optimality in most cases.

3.2 Dual Decomposition for ILPs

In order to solve large shape matching integer linear programs (ILPs) (P), we take [1] as a starting point. Its massive parallelism and generic applicability make it a suitable choice. We give an overview of their approach below.

Since ILPs are NP-hard to solve, a relaxation based on Lagrangian dual is proposed in [1]. Formally, denote $\mathcal{I}$ and $\mathcal{J}$ as the set of variables and constraints, respectively, of the ILP (P). Moreover, denote the variables in constraint j as $\mathcal{I}_j$, and rows containing variable i as $\mathcal{J}_i$. Then represent each constraint j as an independent subproblem $\mathcal{S}_j := \{s \in \{0, 1\}^{\mathcal{I}_j} : \sum_{i \in \mathcal{I}_j} a_{ji} x_i = b_j\}$. The relaxed (dual) problem is written as

$$\max_{\lambda} \sum_{j \in \mathcal{J}} \min_{s \in \mathcal{S}_j} s^\top \lambda^j \quad \text{s.t.} \quad \sum_{j \in \mathcal{J}_i} \lambda_i^j = c_i \quad \forall i \in \mathcal{I}, \quad (\text{D})$$

where λ_i^j is the dual variable for primal variable i and row (constraint) j and λ^j is the vector containing all variables of j , i.e., $\lambda^j = (\lambda_i^j)_{i \in \mathcal{I}_j}$. If optima of individual subproblems $\mathcal{S}_j$ agree, then the dual problem solves the original problem (P). In general it provides a lower bound to the original problem. For more details and derivations we refer to [1, 30].

To optimise the dual problem (D) a parallel block coordinate ascent algorithm is proposed in [1]. Their update scheme also guarantees that dual objective improves or stays the same during dual optimisation. To find a feasible solution to (P) (i.e., a matching in our case) it critically relies on optimising the dual problem. Thus, an improved scheme for solving the dual not only helps in finding better lower bounds but also in recovering better solutions of the original problem.

4 Method

We now present our shape matching framework. First, we will discuss improvements to [1] for efficiently solving the ILP (P). Next we propose our improved cost function which helps to obtain more accurate results by injecting more information into the optimisation problem. The dataflow of our proposed matching pipeline is shown in Fig. 3. We note that our approach is initialisation-free in two ways: (i) we maximise a concave optimisation problem (i.e. Problem (D)), so that any local optimum is a global optimum and initialisation of the solver does not play a significant role. (ii) Our approach does not require landmark correspondence [52] or heuristically chosen initial matchings [45].

4.1 Faster Optimisation

Although the method of [1] delivers good performance on large scale ILPs, it exhibits two weaknesses. First, it can be quite slow to converge as it only relies on first-order updates for optimising the dual (D) and indirectly also the associated

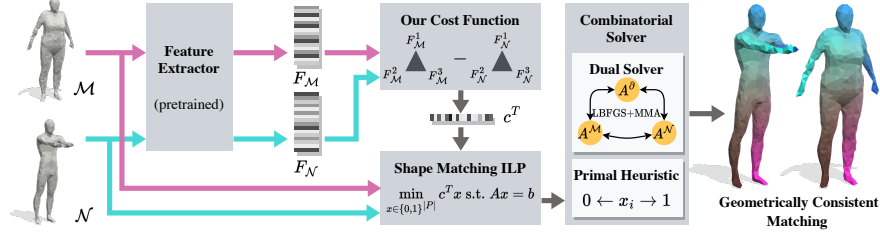


Fig. 3: Illustration of our shape matching pipeline. We use a pretrained feature extractor to define the cost function of ILP (P), which we solve using our combinatorial solver to find a geometrically consistent matching between shapes. The dual solver performs second-order quasi-Newton updates on top of the first-order min-marginal averaging scheme (MMA) of [1]. For primal recovery we adapt the approach of [1] by replacing their dual optimisation procedure by ours leading to faster and better-quality solutions.

primal (P). We aim for better optimisation of the dual problem by leveraging second-order information. Second, the equalities $A^M x = \mathbf{1}$ and $A^N x = \mathbf{1}$ involve a large number of variables (cf. Fig. A.6), thus creating large subproblems, each of which must be processed sequentially, causing bottlenecks when computing on GPUs. To tackle this issue we propose a technique which splits a subproblem into multiple smaller ones that can be solved in parallel.

Quasi-Newton Updates. We incorporate second-order information in optimising the dual problem (D) through quasi-Newton updates. To this end we interleave L-BFGS [35, 40] updates with the first-order min-marginal averaging (MMA) scheme of [1]. An iteration of our scheme is given in Algorithm 1.

Algorithm 1: Quasi-Newton powered MMA (D)

Input: Dual variables λ , Hessian inverse estimate $\hat{H}$, Previous step size γ

- 1 $g^j(\lambda) := \arg \min_{s \in \mathcal{S}_j} s^\top \lambda^j, \forall j \in \mathcal{J}$ // Subgradient of dual obj. (D)
- 2 $\hat{d} = \hat{H} g(\lambda)$ // Compute update direction
- 3 $d_i^j = \hat{d}_i^j - \frac{1}{|\mathcal{J}_i|} \sum_{k \in \mathcal{J}_i} \hat{d}_i^k, \forall i \in \mathcal{I}, j \in \mathcal{J}_i$ // Ensure dual feasibility (D)
- 4 $\gamma \leftarrow \text{FindStepSize}(\lambda, d, \gamma)$
- 5 $\lambda \leftarrow \lambda + \gamma d$ // Update Dual variables
- 6 $\lambda \leftarrow \text{MMA}(\lambda)$ // First-order update of [1]
- 7 $\hat{H} \leftarrow \text{L-BFGS}(\hat{H}, g(\lambda))$ // Update inverse Hessian estimate
- 8 **return** $\lambda, \hat{H}, \gamma$

In detail, we first compute a subgradient of the dual objective by finding a minimising assignment for each subproblem. Next, we calculate the search direction $\hat{d}$ by multiplying the subgradient with an estimate of the inverse Hessian. Naively updating the dual variables in this search direction however, can violate dual feasibility (D). To address this, we project $\hat{d}$ to a feasible search direction d

such that $\sum_j d_i^j$ is zero for all i . Subsequently, a suitable step size is determined that provides sufficient increase in the objective (D). Using this step size, we perform the quasi-Newton update, followed by a min-marginal averaging (MMA) update of [1]. Finally, we update the estimate of the inverse Hessian for use in subsequent iterations. More detailed information about step size selection and inverse Hessian updates can be found in the Appendix. Note that due to its general nature, our second-order update scheme is applicable to all problems which fall under the framework of [1].

In essence, our algorithm combines two distinct update schemes for dual optimisation. It employs a first-order update through min-marginal averaging (MMA), which is fast and guarantees non-decreasing objective in its updates. Our L-BFGS scheme leverages second-order information across MMA iterations and facilitate faster convergence towards the optimum. Empirically, we observe this hybrid scheme to be significantly faster than either L-BFGS or MMA alone.

Constraint Splitting. Similar to [1] we represent each constraint as a binary decision diagram (BDD) [7], a general purpose data structure for encoding functions of binary variables. Essentially, BDDs are directed acyclic graphs and optimal solutions can be found by computing a shortest path. For an example of a BDD see Figure A.5 in the Appendix. Shortest paths and other needed computations require sequentially going through nodes in each of the BDDs. These computations can be parallelized across BDDs on GPUs. However, when a few long BDDs are present, the full parallelism of GPUs can often not be exploited. For this reason we propose a BDD splitting technique that takes a BDD as input, together with the desired smaller chunk size, and outputs a number of sub-BDDs of chunk size length that together are equivalent to the original BDD.

The high level idea is as follows: Let us assume we have a constraint of the form $\sum_{i=1}^k x_i = 2$ and chunk size $k/2$. By introducing auxiliary variables $y_1, y_2 \in \{0, 1\}$ we can split the constraint into two equal-sized sub-constraints $\sum_{i=1}^{\lfloor k/2 \rfloor} x_i - y_1 - y_2 = 0$ and $y_1 + y_2 + \sum_{i=\lfloor k/2 \rfloor+1}^k x_i = 2$. The technical procedure to do this for BDDs representing arbitrary (also non-arithmetic) constraints is detailed in the Appendix, Section A.3.

4.2 Energy Adaption

We use a pretrained feature extractor to predict per-vertex features $F_{\mathcal{M}}$ and $F_{\mathcal{N}}$ for both shapes. These features are used in our adaption of the costs in (P) as

$$c_p = \sum_{v=1}^3 (A_{m_v}^{\mathcal{M}} + A_{n_v}^{\mathcal{N}}) \|F_{m_v}^{\mathcal{X}} - F_{n_v}^{\mathcal{Y}}\|. \quad (1)$$

Here $m_v \in V_{\mathcal{M}}$, $n_v \in V_{\mathcal{N}}$ and $((m_1, n_1), (m_2, n_2), (m_3, n_3))$ forms the p -th product triangle in P . $A_{m_v}^{\mathcal{M}}$ and $A_{n_v}^{\mathcal{N}}$ are the mixed areas according to [39] at vertices m_v and n_v on shapes $\mathcal{M}$ and $\mathcal{N}$, respectively. Using our learned cost function allows to infuse more descriptive information into the problem compared to the previously used hand designed costs based on elastic energy from [68].

Empirically this yields not only in improved matching accuracy but also reduces ILP solver runtimes due to easier optimisation.

5 Experiments

In this section we experimentally analyse on four shape matching datasets the quality of shape matching as well as runtime of our method and finally show how shapes with higher resolutions can be tackled.

Datasets. We consider the challenging remeshed version FAUST_r [14, 43] of FAUST [6], consisting of 100 human shapes (190 pairs in test set). Furthermore, we choose the dataset SMAL [70] with 49 animal shapes of eight species (136 pairs in test set). We also consider the SHREC’20 dataset [15] which consists of 10 highly non-isometric deformed animal shapes. Lastly, we consider DT4D-H [36], which contains nine classes of humanoid shapes sampled from the DeformingThings 4D [33] dataset. We randomly pick 100 intra-class (DT4d-Intra) and 100 inter-class (DT4d-Inter) pairs for evaluation. We split the datasets for training and testing as done in [11] and consequently evaluate only on unseen shapes (except for SHREC’20). We evaluate all shape matching methods by downsampling shapes to 450 faces using [25], unless stated otherwise. We compute features on high-resolution shapes and transfer them to the low-resolution shapes using nearest-neighbour search. Furthermore, we evaluate all methods only on shape pairs which have the same genus, since otherwise there does not exist a solution to (P), i.e., the constraints of the optimisation problem cannot be fulfilled. We repair any defects (non-manifoldness, open boundaries) by [4, 59].

Methods.

LP-Relax [68]: The relaxation based approach for solving the ILP (P). Here, binary constraints are relaxed continuously and the largest element in x is iteratively set to one until the constraints are fulfilled. Since this approach does not scale to larger problems we consider it only in some of the experiments.

Sm-Comb [45]: We compare to the shape matching solver for solving the ILP formulation (P). Here, the dual is optimised to obtain improved costs on the primal which drive a custom primal heuristic.

FastDOG [1]: A general purpose solver for binary optimisation problems including the ILP (P). It shows good scalability on problems from vision and machine learning due to massively parallel dual and primal optimisation routines.

URSSM [11]: The current state-of-the-art for 3D shape matching based on an unsupervised learning-based scheme which, however, does not enforce geometric consistency. We compare against this method to also evaluate if imposing geometric consistency can further improve shape matching performance.

KrnLM [65]: Relaxes a permutation matrix to the set of bi-stochastic matrices while using so-called kernel matrices to efficiently compute solutions of a quadratic assignment problem. For a fair comparison we use the features computed with the pretrained feature extractor of [11].

DiscoMatch: The ILP formulation with our improved cost function (1) based on the pretrained feature extractor of [11]. For solving the dual (D) we use our quasi-Newton scheme along with the BDD splitting approach from Sec. 4.1. For recovering a primal solution we modify the approach of [1] by replacing their dual solver by ours.

We will also utilize our improved cost function (1) in evaluation of all existing ILP based solvers for a fair comparison unless state otherwise. For CPU-based solvers (LP-Relax Sm-Comb) we use an Intel Core i9-12900K CPU and 64GB RAM. For GPU-based solvers we use an NVIDIA A40 GPU with 48GB DRAM.

Metrics. We evaluate shape matching methods w.r.t. the following metrics

Geodesic Distances: We follow the Princeton protocol [27] for evaluation matching quality. To do so, we compute geodesic distances on the meshes and normalise them by the square-root of the area of shape (cf. [27, Sec. 8.2]).

Conformal Distortion: We show conformal distortion errors [23] to evaluate geometric consistency and smoothness of the computed point map. Intuitively, the conformal distortion error evaluates local consistency of triangles after matching (cf. [23, Eq. 3]).

Optimality gaps: To quantify the optimality of a primal solution with objective p , we report the primal dual gap $(p - d_{max})/p$, where d_{max} is the maximum dual objective by the respective solver. Note that the primal dual gap cannot be lower than the integrality gap, since (D) is a relaxation of the original ILP (P). Therefore the primal dual gap provides an upper bound to the optimality gap. We consider a solution to be globally optimal if the primal-dual gap is less than 10^{-2} . Moreover, we evaluate our dual optimisation scheme (Alg. 1) through relative dual gaps $(d^* - d_t)/d^*$, where d_t is the dual objective at time t and d^* is the best known objective for the given instance.

5.1 Results

In Fig. 4, we compare shape matching results w.r.t. mean geodesic errors, conformal distortion errors, and computation times. Results show that the geometric consistency constraints help to decrease mean geodesic errors by sometimes over 50%. At the same time smoothness (by means of conformal distortion errors) of the resulting matching increases consistently across all datasets. Only on SMAL dataset our approach lags behind by a small margin w.r.t. mean geodesic error but still produces a smoother matching. Note that mean geodesic error does not directly quantify geometric consistency and even though results might have small error values, solutions can still be visually implausible (which can also be seen from qualitative results, cf. Fig. 5, Fig. A.3).

From the runtime perspective in Fig. 4 (bottom) our method is by far the fastest method among the ILP-based ones (Sm-Comb, FastDOG, DiscoMatch). Moreover our dual optimisation scheme also helps in solving the primal problem (P). It produces most amount of globally optimal solutions, and fails the least number of times as compared to other ILP solvers, cf. Tab. A.1. Note that

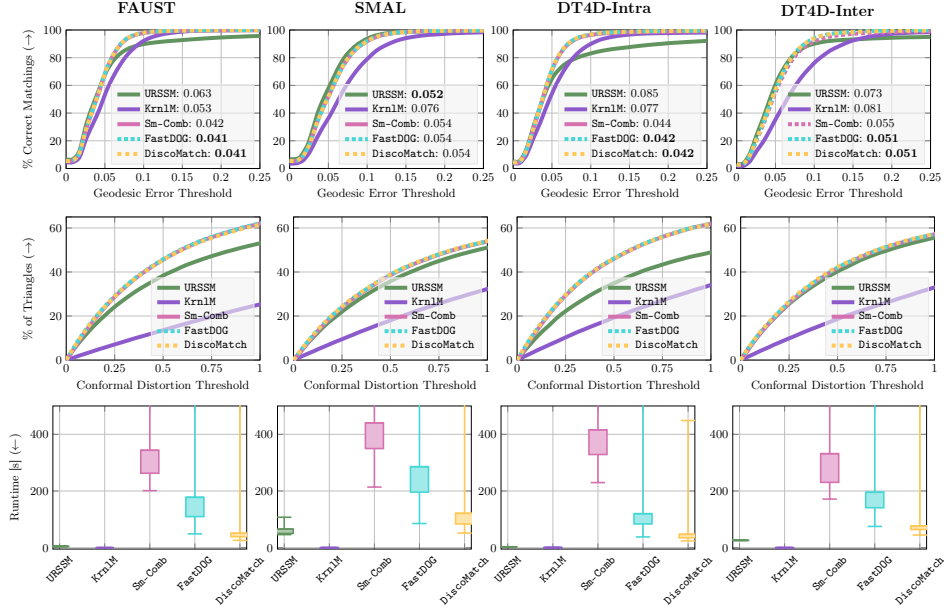


Fig. 4: *Top:* Percentage of correct keypoints w.r.t. geodesic error thresholds on FAUST, SMAL, DT4D-Intra and DT4D-Inter datasets. The numbers in the legends are mean geodesic errors (↓). *Middle:* Conformal distortion errors on FAUST, SMAL, DT4D-Intra and DT4D-Inter datasets. *Bottom:* Statistics of runtime of all methods.

mean geodesic errors of all ILP methods yield very similar matching quality since these methods solve the same optimisation problem.

Furthermore, we observe in Fig. 4 that **KnnLM** struggles on all datasets, likely because it searches over the space of permutation matrices to find shape correspondence which is not a valid assumption for shapes with different discretisation. Thus, we do not include **KnnLM** in further comparisons.

In Fig. 5 and Fig. A.3 we visualize shape matching solutions. For showcasing the effect of geometric consistency we transfer triangulation from one shape to the other via the computed matching. We observe that geometric consistency constraints drastically improve solution quality in our approach as compared to **URSSM** which does not impose geometric consistency.

Ablation: Faster Optimisation. In Fig. 6 we compare the convergence of our solver to the fastest existing approach **FastDOG** for solving the ILP formulation (P). We observe that our algorithm produces much better primal and dual objectives (in significantly less time) compared to **FastDOG**. Moreover, in Tab. A.1 we compare the number of globally optimal and infeasible solutions found by each solver. Our solver finds globally optimal solutions in most cases and the smallest number of infeasible solutions. This shows that our improved dual optimisation also facilitates in primal recovery. In Fig. A.2 we compare runtimes of all ILP-based methods on different problem sizes by varying the res-

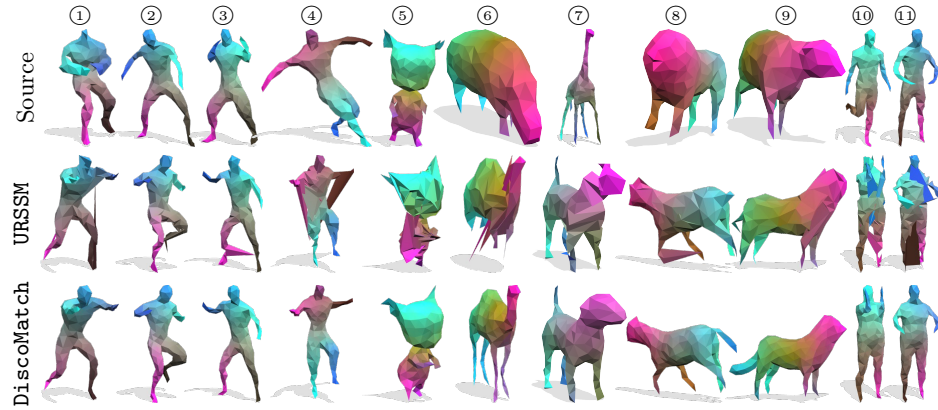


Fig. 5: Qualitative Comparison on DT4D-Inter (①-③), DT4D-Intra (④,⑤), SHREC'20 (⑥,⑦), SMAL (⑧,⑨) and FAUST (⑩,⑪). To better visualise the quality of matchings we transfer triangulation and the color from source to target shapes.

olution (number of triangles) for 5 pairs from the FAUST dataset. We observe that LP-Relax does not scale well with growing problem sizes. While FastDOG is faster than all other existing methods, we outperform it as well.

Note that our solver related improvements from Sec. 4.1 are applicable to ILPs beyond shape matching task as well. To this end we evaluate our solver on some of the difficult datasets used in [1] and obtain 5–10 times runtime improvement. We refer to Sec. A.1 for more details.

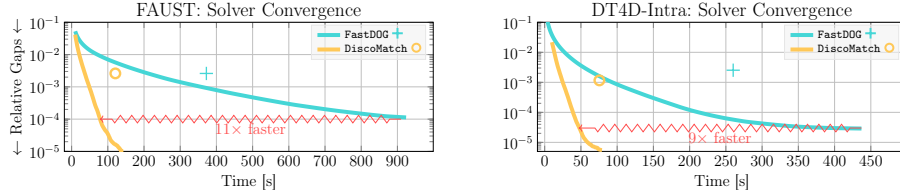


Fig. 6: Convergence plots averaged over a total of 55 instances with varying shape resolutions in $\{500, 550, \dots, 1000\}$. The curves (—) and markers (+, o) represent quality of dual objectives (by relative dual gaps) and primal objectives (by primal-dual gaps), respectively. Our combined solver improvements (Sec 4.1) of quasi-newton updates and constraint splitting give up to $11\times$ improvement in runtime of dual optimisation which also aids in faster primal solutions.

Ablation: Energy Adaptation. In Tab. 2, we compare runtimes and geodesic errors when the ILP (P) is solved with the elastic energy of [69] or with our proposed energy (1). Our solver is the fastest with both energies and furthermore our energy plays a crucial role in decreasing matching errors and runtimes.

	Elastic Energy			Our Energy		
	Time [s.]	Geo. Error	# Inf.	Time [s.]	Geodesic Error	# Inf.
Sm-Comb	648	0.087	12	376	0.042	9
FastDOG	925	0.083	36	327	0.041	4
DiscoMatch	423	0.084	8	134	0.041	2

Table 2: Comparison of solvers for (P) using elastic energy of [69] and our energy (1). We report mean runtime in seconds (Time [s.]), mean geodesic errors (Geo. Error), and number of infeasible solutions (# Inf.) on FAUST dataset (with 450 faces).

5.2 Scaling to Higher Resolutions

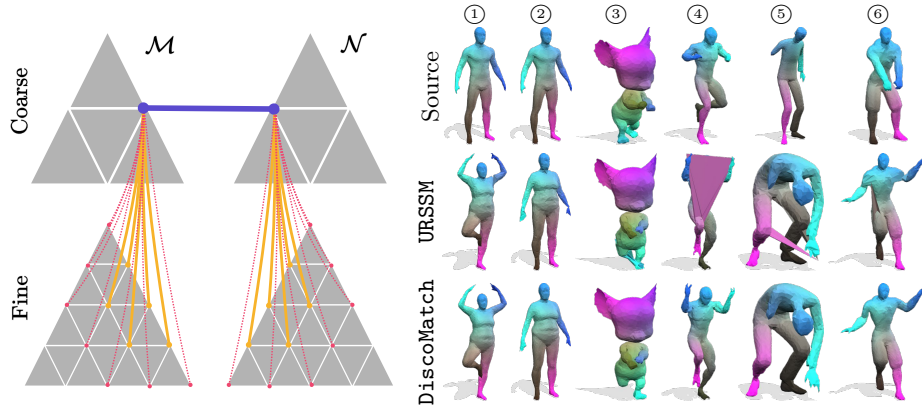


Fig. 7: *Left:* Illustration of our coarse-to-fine pruning strategy. For two matched vertices on the coarse level (—) we only keep potential matchings on the fine level which belong to one-ring (●) and two-ring neighbourhood (●) of respective vertex on coarse level. *Right:* Qualitative matching comparison via triangulation transfer on FAUST (①-②) and DT4D-Intra (③-⑥) datasets for shapes with $4k$ triangles.

As a proof of concept we evaluate our approach on larger shapes resolutions by employing a coarse-to-fine strategy inspired by [69]. We use solutions from coarser shape resolutions to prune (P) on finer ones. We start with a resolution of 500 triangles and prune the subsequent ILPs with underlying shape resolutions of $1k$, $2k$, $3k$ up to $4k$ triangles. For pruning the ILPs we only allow matching between triangles which contain vertices of one-ring and two-ring neighbourhood of a given vertex-vertex matching on the coarser resolution (Fig. 7).

We evaluate on datasets of FAUST, DT4D-Intra (by subsampling shapes to 4000 triangles) with resulting mean geodesic errors of 0.0193/0.0161 for ours and 0.0197/0.0165 for URSSM resp. Note that scale of the geodesic error decreases due to lesser discretisation noise on higher resolutions. The average computation time of the whole coarse-to-fine pipeline is around 20 minutes from which

approximately 30% was spent on dual optimisation. While URSSM requires only few seconds however, it does not provide geometrically consistent matchings. In Fig. 7 (right), we show the corresponding qualitative results. Similar to previous results, our approach leads to better matching quality as compared to URSSM.

6 Discussion & Limitations

Overall, our proposed solver empirically leads to favourable results compared to competing methods: we obtain the largest portion of globally optimal solutions (90%), the smallest amount of infeasible solutions (2%, cf. Tab. A.1), and our solver is up to $11\times$ faster compared to previous approaches (Fig. 6). Despite these positive aspects, there is also a shortcoming: we build upon the product space formalism [68] and also inherit its quadratic space complexity (w.r.t. number of triangles). We found that for our GPU-based solver the memory of modern graphics cards constitutes the main bottleneck when scaling to higher shape resolutions. With a 48GB GPU, we can solve shape matching problems with up to 1000 triangles per shape, which is an order of magnitude smaller compared to purely learning-based shape matching methods (e.g., [11]). As a proof of concept to circumvent the memory bottleneck, we employ a coarse-to-fine scheme allowing us to compute geometrically consistent matchings of shapes with larger resolutions (4000 triangles). Yet, our runtimes for the coarse-to-fine scheme (20 minutes) are still higher than the ones of deep learning methods (few seconds). Nevertheless, we consider our approach to be a large leap forward, especially because our matchings ensure geometric consistency.

7 Conclusion

Our work bridges the gap between learning-based shape matching and purely combinatorial optimisation formulations. While learning-based methods are fast and robust, they do not allow to enforce solutions with certain structural or geometric properties. In contrast, optimisation-based approaches allow for taking such properties into account, but they are generally slower and lead to challenging non-convex problems. To address these limitations, we pursue a two-fold strategy: (i) infuse learning-based feature embeddings into the cost function of a combinatorial problem, and (ii) propose strategies for better optimisation of the Lagrangian dual (D) by quasi-Newton updates and effectively utilising parallelism in modern GPUs. This allows us to solve practically relevant 3D shape matching problems faster as compared to previous approaches, often even to global optimality. We believe that on the one hand our approach is an important contribution for the field of 3D shape analysis per-se. On the other hand, many downstream applications in visual computing that build upon 3D correspondence problems (e.g. autonomous driving and medical image analysis) can benefit from geometrically consistent matchings. Lastly, due to its generic nature our approach can directly benefit from future work in more powerful feature extractors.

References

1. Abbas, A., Swoboda, P.: FastDOG: Fast discrete optimization on GPU. In: CVPR (2022) [2](#), [6](#), [7](#), [8](#), [9](#), [10](#), [12](#), [19](#), [20](#), [23](#), [27](#), [28](#)
2. Attaiki, S., Ovsjanikov, M.: Understanding and improving features learned in deep functional maps. In: CVPR (2023) [3](#)
3. Attaiki, S., Pai, G., Ovsjanikov, M.: DPFM: Deep partial functional maps. In: 2021 International Conference on 3D Vision (3DV). pp. 175–185. IEEE (2021) [3](#)
4. Attene, M.: A lightweight approach to repairing digitized polygon meshes. *The visual computer* **26**, 1393–1406 (2010) [9](#)
5. Bernard, F., Suri, Z.K., Theobalt, C.: MINA: Convex mixed-integer programming for non-rigid shape alignment. In: CVPR (2020) [2](#), [4](#)
6. Bogo, F., Romero, J., Loper, M., Black, M.J.: FAUST: Dataset and evaluation for 3D mesh registration. In: CVPR (2014) [9](#)
7. Bryant, R.E.: Graph-based algorithms for boolean function manipulation. *Computers, IEEE Transactions on* **100**(8), 677–691 (1986) [8](#)
8. Cao, D., Bernard, F.: Unsupervised deep multi-shape matching. In: ECCV (2022) [3](#)
9. Cao, D., Bernard, F.: Self-Supervised Learning for Multimodal Non-Rigid 3D Shape Matching. In: CVPR (2023) [3](#)
10. Cao, D., Eisenberger, M., El Amrani, N., Cremers, D., Bernard, F.: Spectral meets spatial: Harmonising 3d shape matching and interpolation. In: CVPR (2024) [3](#)
11. Cao, D., Roetzer, P., Bernard, F.: Unsupervised learning of robust spectral shape matching. *ACM Transactions on Graphics (TOG)* (2023) [1](#), [2](#), [9](#), [10](#), [14](#), [20](#), [21](#)
12. Coughlan, J., Yuille, A., English, C., Snow, D.: Efficient deformable template detection and localization without user initialization. *Computer Vision and Image Understanding* **78**(3), 303–319 (2000) [4](#)
13. Donati, N., Corman, E., Ovsjanikov, M.: Deep orientation-aware functional maps: Tackling symmetry issues in shape matching. In: CVPR (2022) [3](#)
14. Donati, N., Sharma, A., Ovsjanikov, M.: Deep geometric functional maps: Robust feature learning for shape correspondence. In: CVPR (2020) [3](#), [9](#)
15. Dyke, R.M., Lai, Y.K., Rosin, P.L., Zappalà, S., Dykes, S., Guo, D., Li, K., Marin, R., Melzi, S., Yang, J.: SHREC’20: Shape correspondence with non-isometric deformations. *Computers & Graphics* **92**, 28–43 (2020) [9](#)
16. Ezuz, D., Ben-Chen, M.: Deblurring and denoising of maps between shapes. In: *Computer Graphics Forum*. vol. 36, pp. 165–174. Wiley Online Library (2017) [3](#)
17. Ezuz, D., Heeren, B., Azencot, O., Rumpf, M., Ben-Chen, M.: Elastic correspondence between triangle meshes. In: *Computer Graphics Forum*. vol. 38, pp. 121–134. Wiley Online Library (2019) [3](#)
18. Felzenszwalb, P.F.: Representation and detection of deformable shapes. *TPAMI* **27**(2), 208–220 (2005) [4](#)
19. Gasparetto, A., Cosmo, L., Rodola, E., Bronstein, M., Torsello, A.: Spatial maps: From low rank spectral to sparse spatial functional representations. In: 2017 International Conference on 3D Vision (3DV). pp. 477–485. IEEE (2017) [4](#)
20. Groueix, T., Fisher, M., Kim, V.G., Russell, B.C., Aubry, M.: 3d-coded: 3d correspondences by deep deformation. In: European Conference on Computer Vision (2018) [3](#)
21. Gurobi Optimization, LLC: Gurobi Optimizer Reference Manual (2023), <https://www.gurobi.com> [28](#)

22. Halimi, O., Litany, O., Rodola, E., Bronstein, A.M., Kimmel, R.: Unsupervised learning of dense shape correspondence. In: CVPR (2019) [3](#)
23. Hormann, K., Greiner, G.: MIPS: An efficient global parametrization method. *Curve and Surface Design: Saint-Malo 1999* pp. 153–162 (2000) [10](#)
24. Huang, Q.X., Guibas, L.: Consistent shape maps via semidefinite programming. In: *Computer graphics forum*. vol. 32, pp. 177–186. Wiley Online Library (2013) [2](#)
25. Jacobson, A., Panozzo, D., et al.: libigl: A simple C++ geometry processing library (2018), <https://libigl.github.io/> [9](#)
26. Jiang, P., Sun, M., Huang, R.: Neural Intrinsic Embedding for Non-rigid Point Cloud Matching. In: CVPR (2023) [3](#)
27. Kim, V.G., Lipman, Y., Funkhouser, T.: Blended intrinsic maps. *ACM transactions on graphics (TOG)* **30**(4), 1–12 (2011) [10](#)
28. Knuth, D.E.: *The Art of Computer Programming: Combinatorial Algorithms, Part 1*. Addison-Wesley Professional, 1st edn. (2011) [23](#)
29. Löhner, Z., Rodola, E., Schmidt, F.R., Bronstein, M.M., Cremers, D.: Efficient globally optimal 2d-to-3d deformable shape matching. In: CVPR (2016) [4](#)
30. Lange, J.H., Swoboda, P.: Efficient message passing for 0–1 ilps with binary decision diagrams. In: *International Conference on Machine Learning*. pp. 6000–6010. PMLR (2021) [4](#), [6](#)
31. Li, L., Donati, N., Ovsjanikov, M.: Learning multi-resolution functional maps with spectral attention for robust shape matching. NIPS (2022) [3](#)
32. Li, Q., Liu, S., Hu, L., Liu, X.: Shape correspondence using anisotropic Chebyshev spectral CNNs. In: CVPR (2020) [3](#)
33. Li, Y., Takehara, H., Taketomi, T., Zheng, B., Nießner, M.: 4dcomplete: Non-rigid motion estimation beyond the observable surface. In: ICCV (2021) [9](#)
34. Litke, N., Droske, M., Rumpf, M., Schröder, P.: An image processing approach to surface matching. In: *Symposium on Geometry Processing*. vol. 255, pp. 207–216 (2005) [3](#)
35. Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. *Mathematical programming* **45**(1-3), 503–528 (1989) [7](#), [23](#)
36. Magnet, R., Ren, J., Sorkine-Hornung, O., Ovsjanikov, M.: Smooth non-rigid shape matching via effective dirichlet energy optimization. In: *International Conference on 3D Vision (3DV)* (2022) [9](#)
37. Marin, R., Rakotosaona, M.J., Melzi, S., Ovsjanikov, M.: Correspondence learning via linearly-invariant embedding. In: NeurIPS (2020) [3](#)
38. Maron, H., Dym, N., Kezurer, I., Kovalsky, S., Lipman, Y.: Point registration via efficient convex relaxation. *ACM Transactions on Graphics (TOG)* **35**(4), 73 (2016) [4](#)
39. Meyer, M., Desbrun, M., Schröder, P., Barr, A.H.: Discrete differential-geometry operators for triangulated 2-manifolds. In: *Visualization and mathematics III*. pp. 35–57. Springer (2003) [8](#)
40. Nocedal, J.: Updating quasi-newton matrices with limited storage. *Mathematics of computation* **35**(151), 773–782 (1980) [7](#)
41. Ovsjanikov, M., Ben-Chen, M., Solomon, J., Butscher, A., Guibas, L.: Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG)* **31**(4), 30 (2012) [2](#), [3](#), [4](#)
42. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. NIPS (2017) [3](#)
43. Ren, J., Poulenard, A., Wonka, P., Ovsjanikov, M.: Continuous and orientation-preserving correspondences via functional maps. *ACM Transactions on Graphics (ToG)* **37**, 1–16 (2018) [9](#)

44. Roetzer, P., Löhner, Z., Bernard, F.: Conjugate product graphs for globally optimal 2d-3d shape matching. In: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (2023) [4](#)
45. Roetzer, P., Swoboda, P., Cremers, D., Bernard, F.: A scalable combinatorial solver for elastic geometrically consistent 3d shape matching. In: CVPR (2022) [4](#), [5](#), [6](#), [9](#), [20](#)
46. Roufousse, J.M., Sharma, A., Ovsjanikov, M.: Unsupervised deep learning for structured shape matching. In: ICCV (2019) [3](#)
47. Sahillioglu, Y.: A genetic isometric shape correspondence algorithm with adaptive sampling. ACM Transactions on Graphics (ToG) **37**(5), 1–14 (2018) [2](#)
48. Sahillioglu, Y.: Recent advances in shape correspondence. The Visual Computer **36**(8), 1705–1721 (2020) [3](#)
49. Sankoff, D., Kruskal, J.B.: Time warps, string edits, and macromolecules: the theory and practice of sequence comparison. Addison-Wesley Publishing Company (1983) [4](#)
50. Schmidt, F.R., Töppe, E., Cremers, D.: Efficient planar graph cuts with applications in computer vision. In: CVPR (2009) [4](#)
51. Schmidt, P., Campen, M., Born, J., Kobbelt, L.: Inter-surface maps via constant-curvature metrics. ACM Transactions on Graphics (TOG) **39**(4), 119–1 (2020) [3](#)
52. Schmidt, P., Pieper, D., Kobbelt, L.: Surface maps via adaptive triangulations. Computer Graphics Forum **42**(2) (2023) [2](#), [3](#), [6](#)
53. Schoenemann, T., Cremers, D.: A combinatorial solution for model-based image segmentation and real-time tracking. TPAMI **32**(7), 1153–1164 (2009) [4](#)
54. Schreiner, J., Asirvatham, A., Praun, E., Hoppe, H.: Inter-surface mapping. In: ACM SIGGRAPH 2004 Papers, pp. 870–877 (2004) [2](#), [3](#)
55. Sharma, A., Ovsjanikov, M.: Weakly supervised deep functional maps for shape matching. NIPS (2020) [3](#)
56. Sharma, A., Horaud, R., Cech, J., Boyer, E.: Topologically-robust 3d shape matching based on diffusion geometry and seed growing. In: CVPR 2011. pp. 2481–2488. IEEE (2011) [3](#)
57. Sharp, N., Attai, S., Crane, K., Ovsjanikov, M.: Diffusionnet: Discretization agnostic learning on surfaces. ACM Transactions on Graphics (TOG) (2022) [3](#)
58. Solomon, J., Peyré, G., Kim, V.G., Sra, S.: Entropic metric alignment for correspondence problems. ACM Transactions on Graphics (ToG) **35**(4), 1–13 (2016) [2](#)
59. Sullivan, C.B., Kaszynski, A.: PyVista: 3d plotting and mesh analysis through a streamlined interface for the visualization toolkit (VTK). Journal of Open Source Software **4**(37), 1450 (may 2019). <https://doi.org/10.21105/joss.01450>, <https://doi.org/10.21105/joss.01450> [9](#)
60. Takayama, K.: Compatible intrinsic triangulations. ACM Transactions on Graphics (TOG) **41**(4), 1–12 (2022) [2](#), [3](#)
61. Tam, G.K., Cheng, Z.Q., Lai, Y.K., Langbein, F.C., Liu, Y., Marshall, D., Martin, R.R., Sun, X.F., Rosin, P.L.: Registration of 3d point clouds and meshes: A survey from rigid to nonrigid. IEEE transactions on visualization and computer graphics **19**(7), 1199–1217 (2012) [2](#)
62. Thomas, H., Qi, C.R., Deschaud, J.E., Marcotegui, B., Goulette, F., Guibas, L.J.: Kpconv: Flexible and deformable convolution for point clouds. In: ICCV (2019) [3](#)
63. Trappolini, G., Cosmo, L., Moschella, L., Marin, R., Melzi, S., Rodolà, E.: Shape registration in the time of transformers. Advances in Neural Information Processing Systems (2021) [3](#)

64. Van Kaick, O., Zhang, H., Hamarneh, G., Cohen-Or, D.: A survey on shape correspondence. In: *Computer graphics forum*. vol. 30, pp. 1681–1707. Wiley Online Library (2011) [2](#), [3](#)
65. Vestner, M., Löhner, Z., Boyarski, A., Litany, O., Slossberg, R., Remez, T., Rodola, E., Bronstein, A., Bronstein, M., Kimmel, R., et al.: Efficient deformable shape correspondence via kernel matching. In: *2017 international conference on 3D vision (3DV)*. pp. 517–526. IEEE (2017) [4](#), [9](#)
66. Vestner, M., Litman, R., Rodola, E., Bronstein, A., Cremers, D.: Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3327–3336 (2017) [3](#)
67. Wiersma, R., Eisemann, E., Hildebrandt, K.: Cnns on surfaces using rotation-equivariant features. *ACM Transactions on Graphics (ToG)* **39**(4), 92–1 (2020) [3](#)
68. Windheuser, T., Schlickewei, U., Schmidt, F.R., Cremers, D.: Geometrically consistent elastic matching of 3d shapes: A linear programming solution. In: *ICCV* (2011) [2](#), [4](#), [5](#), [8](#), [9](#), [14](#)
69. Windheuser, T., Schlickewei, U., Schimdt, F.R., Cremers, D.: Large-scale integer linear programming for orientation preserving 3d shape matching. In: *Computer Graphics Forum*. vol. 30, pp. 1471–1480. Wiley Online Library (2011) [4](#), [5](#), [12](#), [13](#)
70. Zuffi, S., Kanazawa, A., Jacobs, D.W., Black, M.J.: 3D menagerie: Modeling the 3D shape and pose of animals. In: *CVPR* (2017) [9](#)