SRFeat: Learning Locally Accurate and Globally Consistent Non-Rigid Shape Correspondence - Supplementary Material

Lei Li Souhaib Attaiki Maks Ovsjanikov
LIX, École Polytechnique, IP Paris

In this supplementary material, we first include more implementation details in Appendix A. Next, we perform further investigation of our spectral loss in Appendix B. Finally, we present more quantitative and qualitative results in Appendix C.

A. Implementation Details

A.1. 3D Shape Matching

We use DiffusionNet [5] as the feature extraction backbone, and its implementation is based on the publicly available codebase[1] released by its authors. The network is composed of four DiffusionNet blocks of width 128. The network takes 3D point positions (*i.e.*, xyz) as input signals and outputs 128-dimensional point-wise features. We set the batch size to 1 and use the ADAM optimizer [3] with an initial learning rate of 0.001. We use servers equipped with NVIDIA A100 and GeForce GTX 1080 GPUs for network training.

In Tab. 4, we show the matching performance w.r.t. λ in Eq. (4) on the SHREC'19 and SMAL datasets, resulting in the choice $\lambda = 10$ for the spectral loss Eq. (9) and $\lambda = 1$ for the Dirichlet energy loss Eq. (7).

In Tab. 5, we show the runtime of shape matching on the SMAL dataset. The statistics were collected on a server with AMD EPYC 7302 CPU, 512GB RAM, and NVIDIA A100 GPU. The columns *Feature*, *FMap*, and *PMap* represent the runtime of feature extraction by DiffusionNet, functional map computation, and point-wise map computation with k-d tree, respectively. We reiterate that the feature matching based methods, *i.e.*, CL, SRFeat-S, and SRFeat-D, do *not* require functional map computation at test time. Note that the distribution of high dimensional features can affect the nearest neighbor search performance of k-d trees used in the point-wise map computation. Nevertheless, SRFeat-D has the best runtime performance.

A.2. Image Matching

We incorporate our proposed smoothness regularization in DGMC [2] for the 2D image keypoint matching task.

[1] https://github.com/nmwsharp/diffusion-net

	SHREC'19		SMAL	
	SRFeat-S	SRFeat-D	SRFeat-S	SRFeat-D
$\lambda = 0.1$	11.2	7.3	14.4	4.8
$\lambda = 1$	10.4	**5.4**	8.6	**3.4**
$\lambda = 10$	**6.1**	7.2	**4.5**	5.0
$\lambda = 100$	11.1	37.2	6.6	10.7

Table 4: Matching performance w.r.t. λ on the SHREC'19 and SMAL datasets (mean geodesic error $\times 100$ on unit-area shapes).

	Feature	FMap	PMap	Total
GeomFmaps	0.0226	0.0437	0.0215	0.0878
CL	0.0227	-	0.0794	0.1021
SRFeat-S	0.0227	-	0.0773	0.1000
SRFeat-D	0.0225	-	0.0419	**0.0644**

Table 5: Runtime (s) per shape pair averaged on SMAL.

Specifically, we build upon the publicly available codebase[2] of DGMC, which is trained with only a contrastive loss, as mentioned in Sec. 3.1 of the main text. In the pre-processing stage, each image is forwarded through a pre-trained VGG16 network, and features of the annotated image keypoints are then extracted on the `relu4_2` and `relu5_1` feature maps through bilinear interpolation and concatenated together. DGMC adopts SplineCNN, a graph neural network, to extract 256-dimensional node-wise features for matching. Delaunay triangulation is used to construct a graph for the keypoints in each image. To incorporate our smoothness regularization in the training loss, we reuse the triangulation result for the Laplacian matrix construction, which is required in Eq. (6) and (8) of the main text. We set the batch size to 512 and use the ADAM optimizer with a learning rate of 0.001. The network is trained for 15 epochs.

[2] https://github.com/rusty1s/deep-graph-matching-consensus

	SHREC'19	SMAL
w/ FMReg [1]	8.5	**6.0**
w/ Eq. (8)	**5.8**	6.1

Table 6: Matching performance of GeomFmaps [1] with different functional map computation schemes (mean geodesic error $\times 100$ on unit-area shapes).

	Feature	FMap	PMap	Total
w/ FMReg [1]	0.0226	0.0437	0.0215	0.0878
w/ Eq. (8)	0.0226	0.0004	0.0193	**0.0423**

Table 7: Runtime (s) of GeomFmaps [1] with different functional map computation schemes on SMAL.

B. Spectral Loss

In Eq. (8) of the main text, we propose to compute a functional map directly from a learned soft point-wise map within the network. In this section, we perform further investigation of Eq. (8) and compare it with the FMReg layer proposed in GeomFmaps [1].

GeomFmaps computes a functional map by treating learned features as probe functions and solving an energy minimization problem in the spectral domain (see Sec. 4.4 of [1]), which is referred as the FMReg layer. This layer, however, needs to solve multiple linear systems within the network, and requires differentiating through the matrix inverse, and thus can be computationally demanding and numerically unstable during training as observed in existing literature [4, 1].

We also compared our proposal based on the definition given in Eq. (8) of the main text, with the FMReg layer introduced in [1]. For this, we directly replace the FMReg layer in GeomFmaps with our Eq. (8) to compute the functional map $\mathbf{C}$, which is compared to the ground-truth $\mathbf{C}_{\text{gt}}$ as the training loss. The rest of the GeomFmaps network is kept the same.

We remark that GeomFmaps w/ Eq. (8) studied in this additional experiment is a variant of the functional map approaches for shape correspondence, which is *different* from the feature matching based methods SRFeat-S and SRFeat-D in our main text. Specifically, GeomFmaps w/ Eq. (8) does not use any contrastive learning losses, and requires the Laplacian basis computation and the functional map estimation at test time.

In Tab. 6, we show the matching performance on the SHREC'19 and SMAL datasets. We observe that Eq. (8) improves GeomFmaps on SHREC'19 and has comparable performance with the FMReg layer on SMAL. Note that the performance of our SRFeat-S and SRFeat-D has been reported in Tab. 1 of the main text. We further show the runtime statistics in Tab. 7 and observe that Eq. (8) significantly speeds up the functional map computation by two orders of magnitude (from 0.0437s to 0.0004s) and reduces the overall runtime by more than a *half*.

C. More Results

In Tab. 8, we show the performance of SRFeat-S-D, which combines CL with the spectral and Dirichlet energy losses. We performed a hyperparameter search to set weights for the spectral and Dirichlet energy losses, resulting in $(0.1, 1)$ for SHREC'19, and $(0.1, 0.1)$ for SMAL. Observe that this variant slightly outperforms SRFeat-S but is comparable to SRFeat-D, indicating that the Dirichlet energy regularization is sufficient for contrastive learning on 3D shapes. SRFeat-S-D requires more hyperparameter tuning, which may be undesirable in practice.

Method	SHREC'19	SMAL
SRFeat-S	6.1	4.5
SRFeat-D	5.4	3.4
SRFeat-S-D	5.3	3.5

Table 8: Matching performance of SRFeat-S-D (mean geodesic error $\times 100$ on unit-area shapes).

In Fig. 8, we present more qualitative results of non-rigid shape matching on the SHREC'19 and SMAL datasets. We note that the matching results are obtained *without* performing any post-refinement, which shows the original matching quality of each method. While SRFeat may not be completely free from correspondence outliers, the results show that our smoothness regularization brings noticeable improvement to the matching quality of CL.

In Fig. 9, we also present more qualitative results of the 2D image keypoint matching task on the PASCAL VOC dataset, demonstrating the improvement of SRFeat over DGMC.

References

[1] Nicolas Donati, Abhishek Sharma, and Maks Ovsjanikov. Deep geometric functional maps: Robust feature learning for shape correspondence. In *CVPR*, 2020. 2

[2] M. Fey, J. E. Lenssen, C. Morris, J. Masci, and N. M. Kriege. Deep graph matching consensus. In *ICLR*, 2020. 1

[3] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015. 1

[4] Jean-Michel Roufosse, Abhishek Sharma, and Maks Ovsjanikov. Unsupervised deep learning for structured shape matching. In *ICCV*, 2019. 2

[5] Nicholas Sharp, Souhaib Attaiki, Keenan Crane, and Maks Ovsjanikov. DiffusionNet: discretization agnostic learning on surfaces. *ACM TOG*, 41(3), 2022. 1

SHREC'19

Source Ground Truth DeepShells DPFM GeomFmaps CL SRFeat-S SRFeat-D

SMAL

Figure 8: More qualitative results from the SHREC'19 and SMAL datasets *without* using any post-refinement. Correspondence is visualized by color transfer.

DGMC SRFeat-D DGMC SRFeat-D DGMC SRFeat-S DGMC SRFeat-S

Figure 9: More qualitative results from the PASCAL VOC dataset. Ground-truth corresponding keypoints have the same color.