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

Lei Li   Souhaib Attaiki   Maks Ovsjanikov

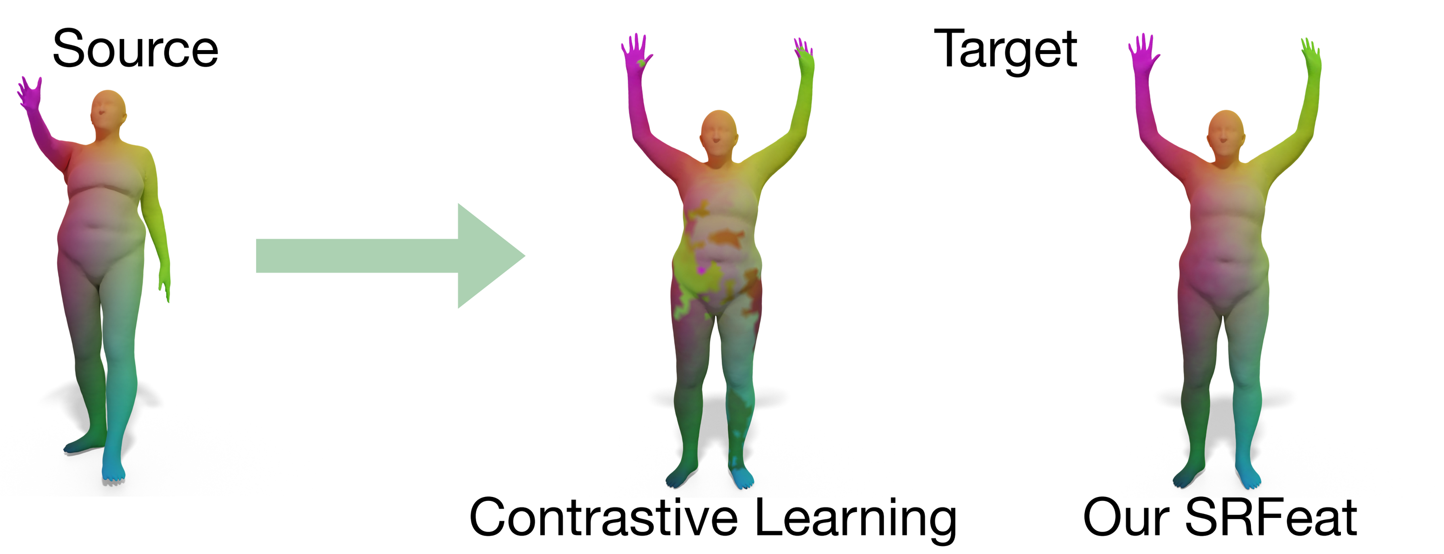
LIX, École Polytechnique, IP Paris

3DV 2022

https://github.com/craigleili/SRFeat

1. Motivation

- **Task:** find corresponding points between deformable shapes via NN search in the local feature space
- **Challenge**: contrastive learning results in correspondences lacking smoothness and consistency



2. Contrastive Learning

- Feature similarity matrix $\mathbf{\Pi} \in \mathbb{R}^{n_1 \times n_2}$:

$$\mathbf{\Pi}^{i,j} = \frac{\exp\left(\mathbf{f}_1^i \cdot \mathbf{f}_2^j / \tau\right)}{\sum_{k=1}^{n_2} \exp\left(\mathbf{f}_1^i \cdot \mathbf{f}_2^k / \tau\right)}$$

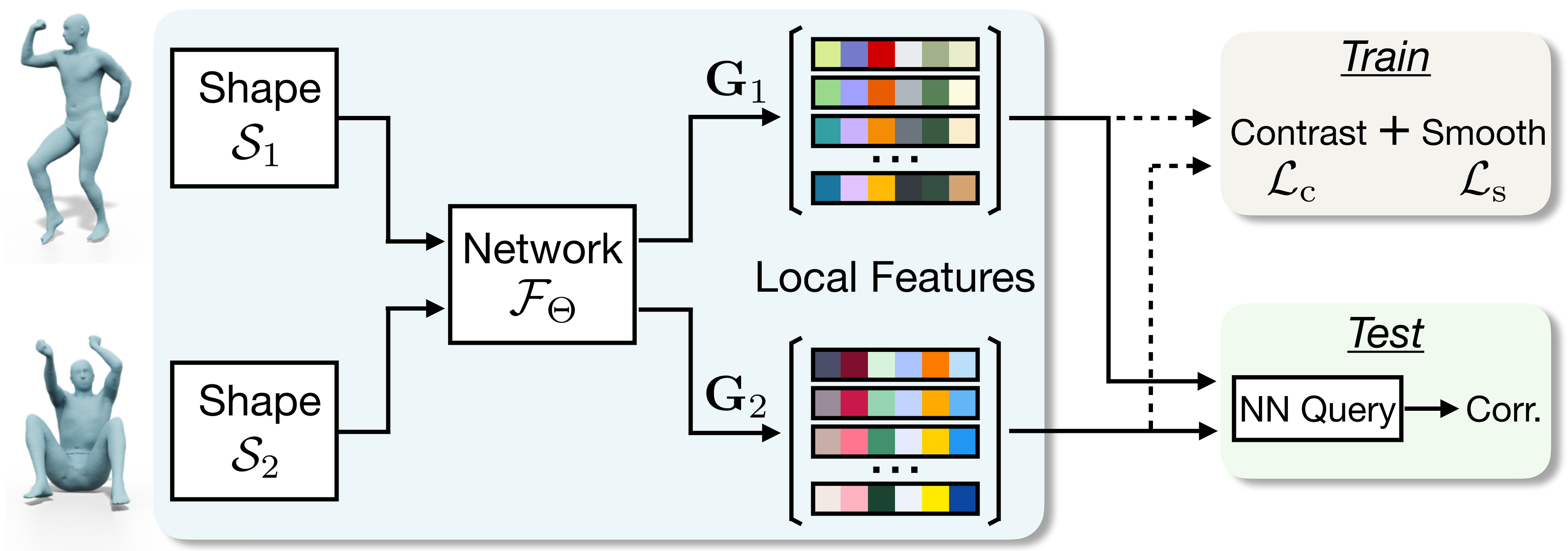
- PointInfoNCE loss: $\mathcal{L}_{\mathrm{c}} = -\sum_i \log\left(\mathbf{\Pi}^{i,\mathrm{GT}(i)}\right)$
- **Issues:** ignore geometric structure of shapes and structural properties of the underlying map

3. Contributions

- Smoothness regularization for contrastive learning leading to more coherent correspondences
- A practical link between contrastive learning and spectral shape matching, through comparison of how they process computed correspondence matrices
- A novel method that extends spectral matching but avoids linear solves inside the network
- SOTA results on both 3D dense correspondence and 2D keypoint matching tasks

4. Our Approach

- Inject geometric information into feature learning
- Maintain the simplicity of contrastive learning



5. Smoothness Regularization

- **Dirichlet Energy Loss:**

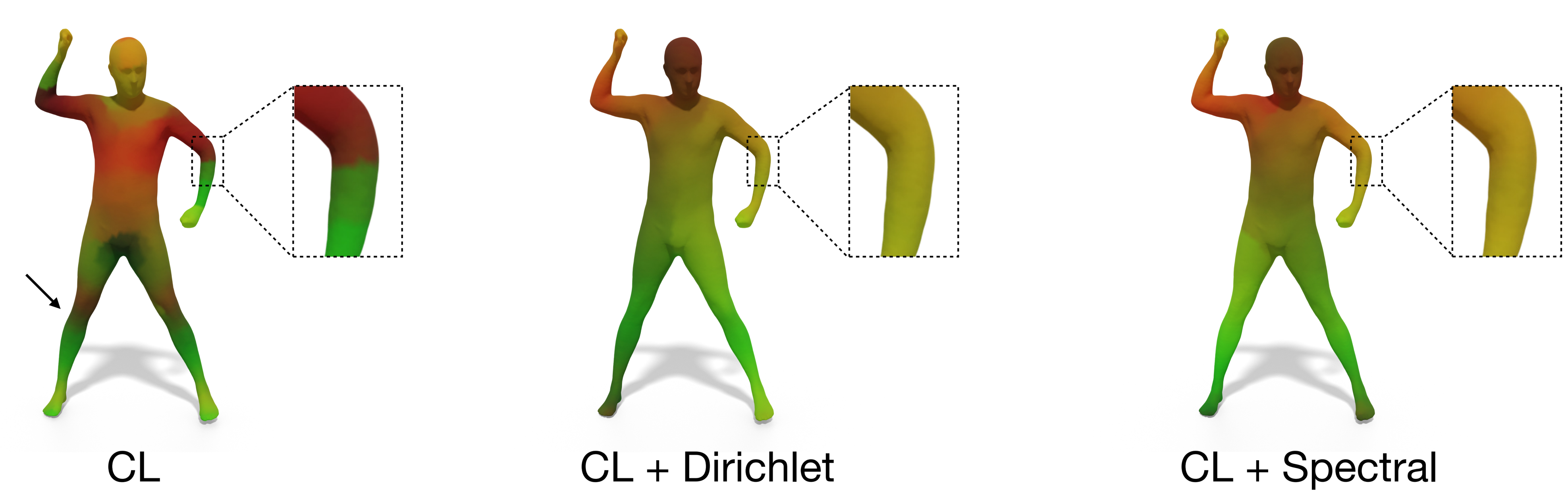
$$\mathcal{L}_{\mathrm{s}} = \frac{1}{2d}\sum_i \mathcal{E}_{\mathrm{d}}(\mathbf{G}_1^i) + \frac{1}{2d}\sum_i \mathcal{E}_{\mathrm{d}}(\mathbf{G}_2^i) \qquad \mathcal{E}_{\mathrm{d}}(g) = \int_{\mathcal{S}} \|\nabla g\|^2 dA$$

Measuring how smooth feature functions are (columns of $\mathbf{G}_1, \mathbf{G}_2$)

- **Spectral Loss:** $\mathcal{L}_{\mathrm{s}} = \|\mathbf{C} - \mathbf{C}_{\mathrm{gt}}\|^2 \qquad \mathbf{C} = \mathbf{\Phi}_1^{\dagger}\mathbf{\Pi}\mathbf{\Phi}_2$

Measuring global consistency in the reduced basis by converting the soft point-wise map $\mathbf{\Pi}$ to a functional map)

- Effect of smoothness regularization on learned local features

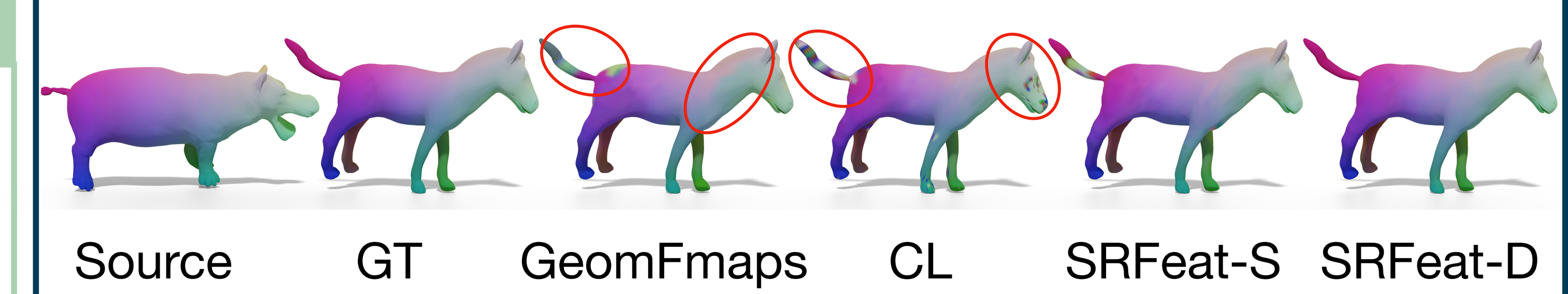


6. Experiments

- 3D Shape Matching:

Human datasets: **F**AUST, **S**CAPE, **S**HREC'**19**
Animal dataset: SMAL

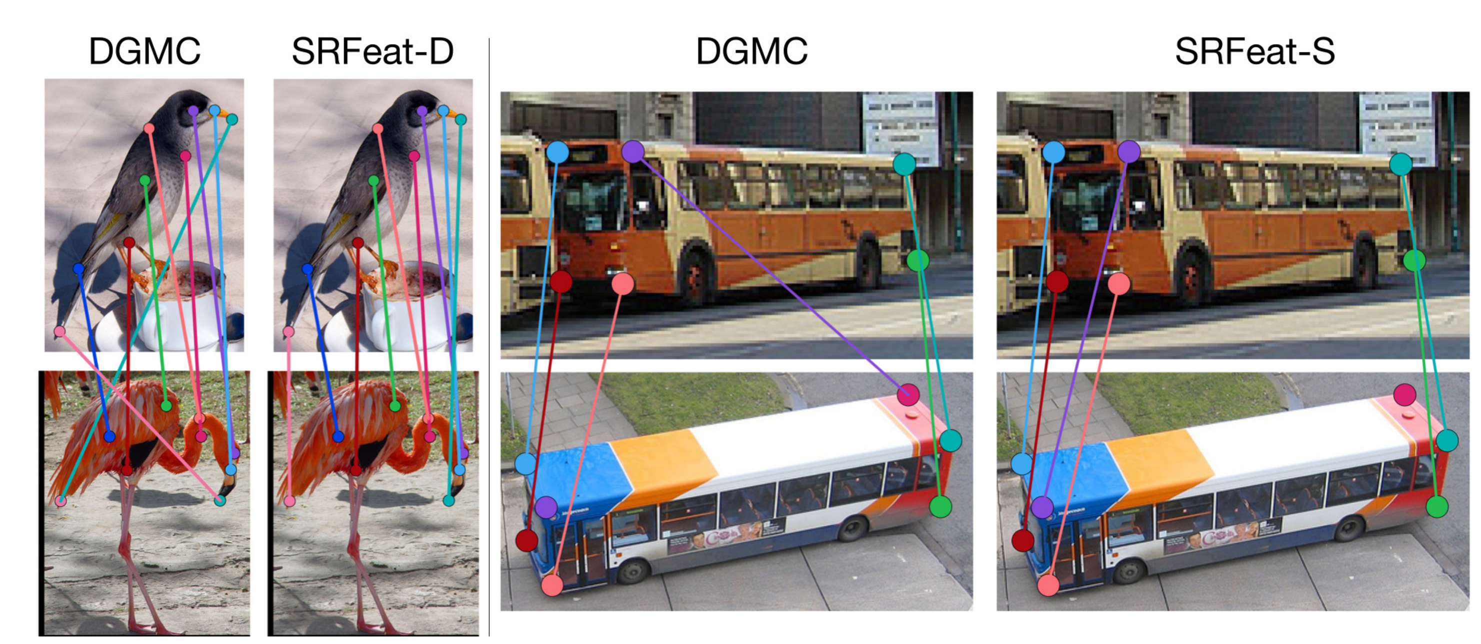
Method	**F**	**S**	Train - Test **F - S**	Train - Test **S - F**	**S19**	SMAL
SURFMNet	15.0	12.0	32.0	32.0	-	-
UnsupFMNet	10.0	16.0	29.0	22.0	-	-
NeuroMorph	8.5	29.9	28.5	18.2	-	-
DeepShells	1.7	2.5	5.4	2.7	21.1	12.6
FMNet	11.0	17.0	30.0	33.0	-	-
3D-CODED	2.5	31.0	31.0	33.0	-	-
HSN	3.3	3.5	25.4	16.7	-	-
ACSCNN	2.7	3.2	8.4	6.0	-	-
DPFM	2.1	2.3	**2.7**	2.5	6.6	6.3
GeomFmaps	2.6	2.9	3.4	3.1	8.5	6.0
CL	**1.1**	**1.9**	6.1	3.7	10.7	13.7
SRFeat-S	**1.1**	2.2	3.9	2.5	6.1	4.5
SRFeat-D	**1.1**	**1.9**	4.3	**2.2**	**5.4**	**3.4**



- Image Keypoint Matching:

PASCAL VOC dataset w/ Berkeley keypoints

Method	Mean Hits@1 (%)
GMN	57.9
PCA-GM	63.8
DGMC	73.0
SRFeat-D	73.9
SRFeat-S	**74.3**



References:
1. S. Xie et al. *PointContrast: Unsupervised pretraining for 3d point cloud understanding*. ECCV 2020.
2. M. Ovsjanikov et al. *Functional maps: a flexible representation of maps between shapes*. TOG 2012.

Acknowledgements:
Parts of this work were supported by the ERC Starting Grant No. 758800 (EXPROTEA) and the ANR AI Chair AIGRETTE