Supplementary Material: Synchronous Diffusion for Unsupervised Smooth Non-Rigid 3D Shape Matching

Dongliang Cao[1], Zorah Lähner[1], and Florian Bernard[1]

University of Bonn

In this supplementary document, we first provide a background on functional maps [66,76] (Section 1) and the deep shape matching block [16] (Section 2) used in our framework. Afterwards, we provide the implementation details of our method in Section 3. Next, we conduct an ablative experiment to investigate the influence of the choice of diffusion time (Section 4). Eventually, we show more qualitative results of our method in Section 5.

1 Deep Functional Maps in a Nutshell

In this section, we provide a brief introduction to the standard pipeline of the deep functional maps method [76]. We consider a pair of 3D shapes $\mathcal{M}$ and $\mathcal{N}$ represented as triangle meshes with $n_{\mathcal{M}}$ and $n_{\mathcal{N}}$ (w.l.o.g. $n_{\mathcal{M}} \leq n_{\mathcal{N}}$) vertices, respectively. Here we summarise the main steps of its pipeline:

1. Compute the Laplacian matrices $\mathbf{L}_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times n_{\mathcal{M}}}, \mathbf{L}_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}}$ [68] and the corresponding first k eigenfunctions $\mathbf{\Phi}_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times k}, \mathbf{\Phi}_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times k}$ and the diagonal eigenvalue matrix $\mathbf{\Lambda}_{\mathcal{M}} \in \mathbb{R}^{k \times k}, \mathbf{\Lambda}_{\mathcal{N}} \in \mathbb{R}^{k \times k}$, respectively.
2. Compute pointwise features $\mathbf{E}_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times d}, \mathbf{E}_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times d}$ defined on each shape via a learnable feature extractor $\mathcal{F}_{\theta}$ with weights θ.
3. Compute the functional map $\mathbf{C}_{\mathcal{M}\mathcal{N}} \in \mathbb{R}^{k \times k}$ associated with the Laplacian eigenfunctions by solving the optimisation problem
$$\mathbf{C}_{\mathcal{M}\mathcal{N}} = \arg\min_{\mathbf{C}} \; E_{\mathrm{data}}\left(\mathbf{C}\right) + \lambda E_{\mathrm{reg}}\left(\mathbf{C}\right). \tag{1}$$
Here, $E_{\mathrm{data}}\left(\mathbf{C}\right) = \left\| \mathbf{C}\mathbf{\Phi}_{\mathcal{M}}^{\dagger}\mathbf{E}_{\mathcal{M}} - \mathbf{\Phi}_{\mathcal{N}}^{\dagger}\mathbf{E}_{\mathcal{N}} \right\|_F^2$ enforces descriptor preservation, while the regularisation term E_{reg} imposes certain structural properties (e.g. Laplacian commutativity $E_{\mathrm{reg}}\left(\mathbf{C}\right) = \left\| \mathbf{C}\mathbf{\Lambda}_{\mathcal{M}} - \mathbf{\Lambda}_{\mathcal{N}}\mathbf{C} \right\|_F^2$ [66]).
4. During training, structural regularisation (e.g. orthogonality, bijectivity [76]) is imposed on the functional maps, i.e.
$$L_{\mathrm{struct}} = \lambda_{\mathrm{bij}} L_{\mathrm{bij}} + \lambda_{\mathrm{orth}} L_{\mathrm{orth}}. \tag{2}$$
5. During inference, the point-wise map $\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}$ is obtained based on the map relationship $\mathbf{C}_{\mathcal{N}\mathcal{M}} = \mathbf{\Phi}_{\mathcal{M}}^{\dagger}\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}\mathbf{\Phi}_{\mathcal{N}}$, e.g. either by nearest neighbour search in the spectral domain or by other post-processing techniques [28,33,61,93].

2 Deep Shape Matching Block

In this section, we provide a brief introduction about our deep shape matching block in Fig. 2 based on the shape matching framework proposed in [16].

2.1 Pointwise Correspondences Based on Feature Matching

In theory, the pointwise map $\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}$ should be a permutation matrix, i.e.

$$\left\{\mathbf{\Pi} \in \{0,1\}^{n_{\mathcal{M}} \times n_{\mathcal{N}}} : \mathbf{\Pi}\mathbf{1}_{n_{\mathcal{N}}} = \mathbf{1}_{n_{\mathcal{M}}}, \mathbf{1}_{n_{\mathcal{M}}}^{\top}\mathbf{\Pi} \leq \mathbf{1}_{n_{\mathcal{N}}}^{\top}\right\}, \tag{3}$$

where $\mathbf{\Pi}(i,j)$ indicates whether the i-th vertex in shape $\mathcal{M}$ corresponds to the j-th vertex in shape $\mathcal{N}$. In our work the pointwise correspondences $\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}$ between shapes $\mathcal{M}$ and $\mathcal{N}$ are obtained based on pairwise similarity of the learned features $\mathbf{E}_{\mathcal{M}}, \mathbf{E}_{\mathcal{N}}$. Following prior works [16,29], we use the softmax operator to produce a soft correspondence matrix, i.e. $\mathbf{\Pi}_{\mathcal{M}\mathcal{N}} = \text{Softmax}\left(\mathbf{E}_{\mathcal{M}}\mathbf{E}_{\mathcal{N}}^{T}/\tau\right)$, where τ is the temperature factor to determine the softness of the pointwise map. The softmax operator is applied in each row to ensure that correspondences are non-negative and $\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}\mathbf{1}_{n_{\mathcal{N}}} = \mathbf{1}_{n_{\mathcal{M}}}$. In this way, $\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}$ can be interpreted as a soft assignment of vertices $\mathbf{V}_{\mathcal{N}}$ in shape $\mathcal{N}$ to vertices $\mathbf{V}_{\mathcal{M}}$ in shape $\mathcal{M}$. Similarly, we can obtain $\mathbf{\Pi}_{\mathcal{N}\mathcal{M}}$ with the roles of $\mathcal{M}$ and $\mathcal{N}$ swapped.

2.2 Spectral Regularisation on Pointwise Correspondences

The predicted pointwise soft correspondences are regularised in the spectral domain based on the relationship to the associated functional map. Specifically, the predicted pointwise maps $\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}, \mathbf{\Pi}_{\mathcal{N}\mathcal{M}}$ are regularised by the coupling relationship between pointwise maps and functional maps [70], i.e.

$$L_{\text{couple}} = \left\|\mathbf{C}_{\mathcal{M}\mathcal{N}} - \mathbf{\Phi}_{\mathcal{N}}^{\dagger}\mathbf{\Pi}_{\mathcal{N}\mathcal{M}}\mathbf{\Phi}_{\mathcal{M}}\right\|_F^2 + \left\|\mathbf{C}_{\mathcal{N}\mathcal{M}} - \mathbf{\Phi}_{\mathcal{M}}^{\dagger}\mathbf{\Pi}_{\mathcal{M}\mathcal{N}}\mathbf{\Phi}_{\mathcal{N}}\right\|_F^2, \tag{4}$$

where $\mathbf{C}_{\mathcal{M}\mathcal{N}}, \mathbf{C}_{\mathcal{N}\mathcal{M}}$ are the functional maps computed by solving the optimisation problem in Eq. (1). To this end, our total loss is a linear combination of our synchronous diffusion regularisation and spectral regularisation related to functional maps, i.e.

$$L_{\text{total}} = L_{\text{diff}} + \lambda_{\text{couple}}L_{\text{couple}} + \lambda_{\text{struct}}L_{\text{struct}}. \tag{5}$$

3 Implementation Details

As we described in Sec. 2, our deep shape matching block is based on the framework proposed in [16]. To this end, we use the official implementation of [16] to build our deep shape matching block. In the context of our regularisation based on synchronous diffusion, the dimension of the random distribution $\mathbf{F}_{\bullet}$ is 128 (i.e. $h = 128$). Regarding to the maximum diffusion time in Eq. (9), we heuristically choice $T = 1.0^{-2}$ for near-isometric, topological noise and partial shape datasets and $T = 1.0^{-4}$ for non-isometric shape datasets, respectively. Since for non-isometric shape pairs, the longer diffusion time causes more inconsistent diffusion processes on both shapes. For the total loss we set $\lambda_{\text{couple}} = 1, \lambda_{\text{struct}} = 1$ in Eq. (5) based on empirical experiments. In the context of input features, we follow previous works [5, 13, 16] by using WKS [6] as input features except for partial shape matching on SHREC'16 dataset, where vertex positions are used. When we evaluate our method on SHREC'16 dataset, we pre-train the feature extractor on complete shapes by using a combination of complete shape datasets (DT4D-H [59], SMAL [102], FAUST [11] and SCAPE [1]).

4 Influence of Diffusion Time

Here, we conduct an ablative experiment to investigate the influence of the choice of the maximum diffusion time on the challenging TOPKIDS dataset and choose the optimal value for our final results. See Table 1. While large values for T are likely to be affected by topological noise, a too small value is not able to provide any benefit due to only small neighbourhood being covered.

Table 1: Qualitative results on TOPKIDS dataset with different maximum diffusion time. A careful choice of the maximum diffusion time is important for our regularisation.

Diff. Time T	1.0	1.0^{-1}	1.0^{-2} (Ours)	5.0^{-3}	1.0^{-3}	1.0^{-4}
Geo. error (×100)	23.7	10.5	**5.4**	6.9	8.5	10.3

5 More Qualitative Results

In the next figures, we provide additional qualitative results of our method on TOPKIDS, SHREC'19, SMAL and DTH4D-H corresponding to the quantitative results reported in the main text.

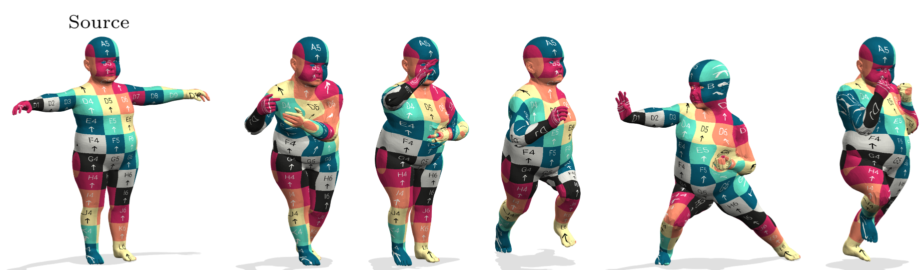

Fig. 1: Qualitative results of our method on TOPKIDS. Our method obtains accurate correspondences for shapes with topological noise.

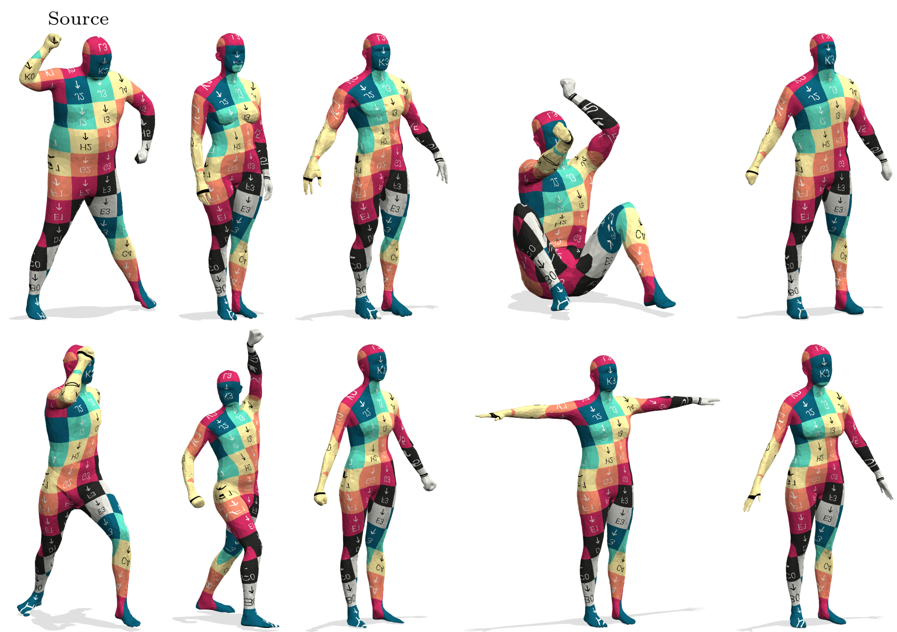

Fig. 2: Qualitative results of our method on SHREC'19. Our method obtains accurate correspondences for human shapes with diverse poses and shapes.

Fig. 3: Qualitative results of our method on SMAL. Our method obtains accurate correspondences for shapes in different classes.

Fig. 4: Qualitative results of our method on DT4D-H inter-class. Our method obtains accurate correspondences for non-isometrically deformed shapes.