On Unsupervised Partial Shape Correspondence

Supplementary Material

7 Proof of Theorem 1

Proof. By splitting integrals on $\mathcal{X}$ into a sum of integrals over $\mathcal{Y}$ and $\mathcal{Z}$, we have,

$$\begin{aligned}\hat{\boldsymbol{F}}_x = \langle \psi_i^{\mathcal{X}}, f_j^{\mathcal{X}} \rangle_{\mathcal{X}} &= \int_{\mathcal{X}} \psi_i^{\mathcal{X}*}(x) f_j^{\mathcal{X}}(x) dx \\ &= \int_{\mathcal{Y}} \psi_i^{\mathcal{X}*}(y) f_j^{\mathcal{X}}(y) dy + \int_{\mathcal{Z}} \psi_i^{\mathcal{X}*}(z) f_j^{\mathcal{X}}(z) dz \\ &= \langle \psi_i^{\mathcal{X}}, f_j^{\mathcal{X}} \rangle_{\mathcal{Y}} + \langle \psi_i^{\mathcal{X}}, f_j^{\mathcal{X}} \rangle_{\mathcal{Z}} = \hat{\boldsymbol{F}}^{(\mathcal{Y})} + \hat{\boldsymbol{F}}^{(\mathcal{Z})}. \end{aligned} \tag{20}$$

The FM-layer computes the functional map in Eq. (2). We can plug in $\hat{\boldsymbol{F}}_x = \hat{\boldsymbol{F}}^{(\mathcal{Y})} + \hat{\boldsymbol{F}}^{(\mathcal{Z})}$ into Eq. (2) and obtain $\hat{\boldsymbol{C}}_{yx} = \boldsymbol{C}_{yx} + \boldsymbol{C}_{yx}^{E}$.

8 Proof of Theorem 2

Proof. Let n_x, n_y, and n_z be the number of vertices of $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively. Without loss of generality, we sort the vertices of $\mathcal{X}$ such that those corresponding to the vertices in $\mathcal{Y}$ appear first and in the same order as those of $\mathcal{Y}$. Thereby, we can write,

$$\boldsymbol{\Psi}_x = \begin{bmatrix} \boldsymbol{\Psi}^{(\mathcal{Y})} \\ \boldsymbol{\Psi}^{(\mathcal{Z})} \end{bmatrix}, \tag{21}$$

where $\boldsymbol{\Psi}^{(\mathcal{Y})} \in \mathbb{R}^{n_y \times k}$ and $\boldsymbol{\Psi}^{(\mathcal{Z})} \in \mathbb{R}^{n_z \times k}$.

Denote, by $\boldsymbol{F}_x \in \mathbb{R}^{n_x \times d}$ and $\boldsymbol{F}_y \in \mathbb{R}^{n_y \times d}$ the feature matrices of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Based on our ordering, we have that,

$$\boldsymbol{F}_x = \begin{bmatrix} \boldsymbol{F}^{(\mathcal{Y})} \\ \boldsymbol{F}^{(\mathcal{Z})} \end{bmatrix}. \tag{22}$$

Here, $\boldsymbol{F}^{(\mathcal{Y})}$ are the features extracted from vertices on $\mathcal{X}$ located on the subsurface $\mathcal{Y}$, and similarly for $\boldsymbol{F}^{(\mathcal{Z})}$. We can rewrite $\boldsymbol{\Psi}_x^\top \boldsymbol{A}_x \boldsymbol{F}_x$ using Eqs. (22) and (21) as

$$\boldsymbol{\Psi}_x^\top \boldsymbol{A}_x \boldsymbol{F}_x = (\boldsymbol{\Psi}^{(\mathcal{Y})})^\top \boldsymbol{A}_y \boldsymbol{F}^{(\mathcal{Y})} + (\boldsymbol{\Psi}^{(\mathcal{Z})})^\top \boldsymbol{A}_z \boldsymbol{F}^{(\mathcal{Z})}, \tag{23}$$

since vertices on both $\mathcal{X}$ and $\mathcal{Y}$ are associated to the same area on both shapes, meaning that $\boldsymbol{A}_y = \boldsymbol{A}^{(\mathcal{Y})}$, and likewise $\boldsymbol{A}_z = \boldsymbol{A}^{(\mathcal{Z})}$, where $\boldsymbol{A}^{(\mathcal{Y})} = \boldsymbol{A}_x|_{\mathcal{Y}}$ and $\boldsymbol{A}^{(\mathcal{Z})} = \boldsymbol{A}_x|_{\mathcal{Z}}$ are the restrictions of $\boldsymbol{A}_x$ to the matching part of $\mathcal{Y}$ and $\mathcal{Z}$ on

$\mathcal{X}$, respectively. Plugging Eq. (23) into Eq. (3), $\hat{\boldsymbol{C}}_{xy} = \boldsymbol{C}_{xy} + \boldsymbol{C}^E_{yx}$ is the sum of two terms,

$$\boldsymbol{C}_{yx} = (\boldsymbol{\Psi}^{(\mathcal{Y})})^\top \boldsymbol{A}_y \boldsymbol{F}^{(\mathcal{Y})} \boldsymbol{F}_y^\top \boldsymbol{A}_y \boldsymbol{\Psi}_y \boldsymbol{Q}_y^{-1}, \tag{24}$$

the self-functional map of $\mathcal{Y}$ with its correspondence in $\mathcal{X}$ with respect to the basis $\boldsymbol{\Psi}_x^{(\mathcal{Y})}$ and $\boldsymbol{\Psi}_y$, and,

$$\boldsymbol{C}^E_{yx} = (\boldsymbol{\Psi}^{(\mathcal{Z})})^\top \boldsymbol{A}_z \boldsymbol{F}^{(\mathcal{Z})} \boldsymbol{F}_y^\top \boldsymbol{A}_y \boldsymbol{\Psi}_y \boldsymbol{Q}_y^{-1}, \tag{25}$$

the unavoidable error injected into C_{yx}.

9 Error analysis of the FM-layer

To better understand the two components that form $\hat{\boldsymbol{C}}_{yx}$, let us simplify the problem. Assume that $\mathcal{X}$ is composed of two disconnected sub-surfaces $\mathcal{Y}$ and $\mathcal{Z}$. As the surfaces $\mathcal{Y}$ and $\mathcal{Z}$ are disconnected, the eigenfunctions of the LBO of $\mathcal{X}$ consist of two disjoint sets of functions, assuming different modes. These two disjoint sets are interleaving according to increasing eigenvalues, where one set contains functions that are the eigenfunction of the LBO of $\mathcal{Y}$ extended to $\mathcal{Z}$ by taking zero values for every point on $\mathcal{Z}$, and vice versa for the other set.

Additionally, we assume that we have a feature extractor which is robust to partial shape matching, that is,

$$\boldsymbol{F}_y = \boldsymbol{\Pi}^*_{xy} \boldsymbol{F}_x, \tag{26}$$

where $\boldsymbol{\Pi}^*_{xy}$ is the exact correspondence matrix between $\mathcal{X}$ and $\mathcal{Y}$. Moreover, we assume that the extracted features have rank d. We now look for the functional map $\boldsymbol{C}^*$ that would give us a perfect match between $\mathcal{X}$ and $\mathcal{Y}$. That is,

$$\boldsymbol{C}^* = \boldsymbol{\Psi}_x^\top \boldsymbol{A}_x \boldsymbol{\Pi}^*_{yx} \boldsymbol{\Psi}_y, \tag{27}$$

where, w.l.o.g. we choose $\boldsymbol{\Psi}_x$ and $\boldsymbol{\Psi}_y$ as the eigenfunctions of the LBO of surfaces $\mathcal{X}$ and $\mathcal{Y}$, respectively, and we sort the vertices of $\mathcal{X}$ such that those corresponding to the vertices in $\mathcal{Y}$ appear first and in the same order as those of $\mathcal{Y}$. we have that $\boldsymbol{\Pi}^*_{xy} = \boldsymbol{J}_{n_y}$ where $\boldsymbol{J}_{n_y} = \begin{bmatrix} \boldsymbol{I}_{n_y} \\ \boldsymbol{0} \end{bmatrix}$ is the identity matrix until column n_y and only zeros columns afterwards. Without loss of generality, we sort $\boldsymbol{\Psi}_x$ so that its leading eigenfunctions are those related to $\mathcal{Y}$, and then those related to $\mathcal{Z}$. Thus, we have,

$$\boldsymbol{C}^* = \boldsymbol{\Psi}_x^\top \begin{bmatrix} \boldsymbol{A}_y \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{\Psi}_y = \begin{bmatrix} \boldsymbol{\Psi}_y^\top \boldsymbol{A}_y \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{\Psi}_y = \boldsymbol{J}_{n_y}. \tag{28}$$

Using our previous assumptions in Eq. (6) gives us,

$$\begin{aligned} \boldsymbol{C}_{yx} &= (\boldsymbol{\Psi}^{(\mathcal{Y})})^\top \boldsymbol{A}_y \boldsymbol{F}^{(\mathcal{Y})} \boldsymbol{F}_y^\top \boldsymbol{A}_y \boldsymbol{\Psi}_y \boldsymbol{Q}_y^{-1} \\ &= \begin{bmatrix} \boldsymbol{\Psi}_y^\top \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{A}_y \boldsymbol{F}_y \boldsymbol{F}_y^\top \boldsymbol{A}_y \boldsymbol{\Psi}_y \boldsymbol{Q}_y^{-1} \end{aligned}$$

$$= \begin{bmatrix} \boldsymbol{\Psi}_y^\top \boldsymbol{A}_y \boldsymbol{F}_y \boldsymbol{F}_y^\top \boldsymbol{A}_y \boldsymbol{\Psi}_y \boldsymbol{Q}_y^{-1} \\ \mathbf{0} \end{bmatrix} = \boldsymbol{J}_{n_y}. \tag{29}$$

This is the ideal functional map as it defined by the given mapping between the surfaces $\mathcal{X}$ and $\mathcal{Y}$. Therefore, the second term $\boldsymbol{C}_{yx}^E$ is a an error which is proportional to the area of $\mathcal{Z}$.

10 The softmax operator compared to the FM-layer.

Pipelines using the FM-layer or ours using the Softmax operator all involve the feature product $\boldsymbol{F}^{(\mathcal{Z})}\boldsymbol{F}_y^\top$. Since the feature extractor is unaware of the locations of the missing parts, it cannot attribute trivial features, e.g. $\mathbf{0}$ only values, at the missing parts. Therefore, the term $\boldsymbol{F}^{(\mathcal{Z})}\boldsymbol{F}_y^\top$, present in both pipelines, introduces errors. In the case of the FM-layer the dependence on this product is linear, whereas in our pipeline the error in the correspondence matrix estimated from the softmax operator is proportional to $\boldsymbol{E}^{-1}\exp(\boldsymbol{F}^{(\mathcal{Z})}\boldsymbol{F}_y^\top/\tau)$, where $\boldsymbol{E} = \operatorname{diag}(\sum_i \exp((\boldsymbol{F}_x\boldsymbol{F}_y^\top/\tau)_{ij}))$. For any practical feature extractor, like DiffusionNet [42], there is usually good alignment of features of vertices on the partial shape and their counterpart on the full shape. Comparatively, features on the missing parts may fail to align as well. Thus, the normalisation $\boldsymbol{E}^{-1}$ reduces exponentially the influence of noisy correlations with missing parts. This is particularly the case for $\tau \to 0$ (as in our experiments), as then, the softmax with temperature hyperparameter τ collapses to the max operation. For example, consider three points x, y, and z where $x \in \mathcal{X} \setminus \mathcal{Z}$, $y \in \mathcal{Y}$, and $z \in \mathcal{Z}$, having features $\boldsymbol{f}_x$, $\boldsymbol{f}_y$, and $\boldsymbol{f}_z$ respectively. If $\boldsymbol{f}_z\boldsymbol{f}_y^\top = 0.5$, $\boldsymbol{f}_x\boldsymbol{f}_y^\top = 0.7$, and $\tau = 0.01$, then the error is close to 10^{-9} in the softmax operation, which is negligible to the error of 0.35 induced in the FM-layer.

11 The PFAUST benchmark

We created the PFAUST benchmark from FAUST remeshed [38] similarly to the way SHREC'16 [15] was created from TOSCA [12]. For each shape in FAUST, we randomly chose m vertices, and for each selected vertex we removed all other vertices on the shape within a geodesic radius of r. As these holes may disconnect the shape, we then keep only the connected component with the largest number of vertices. This process constructs the parts of our shapes. For the full shapes, we only chose subjects with ID ending with 0. Shapes created from the train and test splits of FAUST remeshed populate our train and test splits respectively, and we provide for each partial shape its ground-truth correspondence matrix with its full shape. Thus, the training set contains 8 full shapes and 80 partial shapes. The test set is analogous except for its number of shapes; 20 partial and 2 full shapes. We created two datasets, PFAUST-M and PFAUST-H, generated with different choices of m and r to create two levels of missing regions. A high number of small holes significantly changes the topology, which makes the partial shape correspondence task harder than with fewer albeit larger holes. As such,

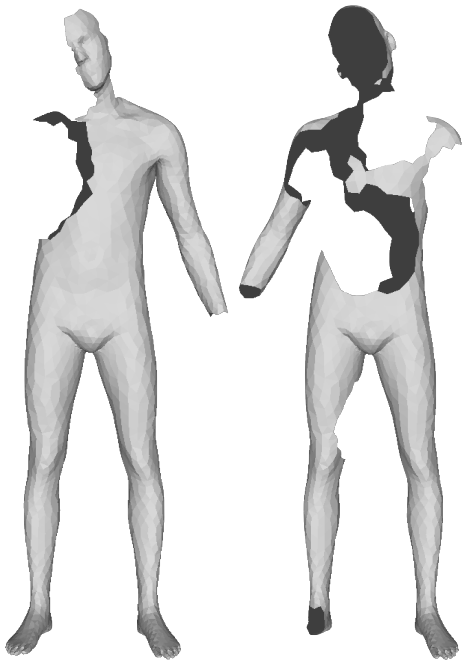

Fig. 4: Example of shapes existing in our new PFAUST benchmark. The left shape is from PFAUST-M, while the right one is from PFAUST-H, which are of medium and hard difficulty, respectively.

PFAUST-M and PFAUST-H are of "medium" and "hard" difficulties, respectively. We created PFAUST-M by taking $r = 0.16$ and $m = 4$, and PFAUST-H by choosing $r = 0.1$ and $m = 13$. See Fig. 4 for plots of example shapes from each dataset.

12 Additional Implementation Considerations

Our method. We trained our method using the Adam optimizer [23], with a learning rate of 10^{-3} and cosine annealing scheduler [28] with minimum learning rate parameter $\eta_{\min} = 10^{-4}$ and maximum temperature of $T_{\max} = 300$ steps. We train for 20000 iterations, and at test time our refinement process has 15 iterations.

On RobustFMnet. The method by Cao et al. [14] was claimed to produce state-of-the-art results when applied to various shape correspondence benchmarks including partial shape matching. And yet, the methodology proposed in [14], in spite of its conceptual beauty, involved difficulties in the evaluation phase. The results reported by Cao et al. were obtained by first pre-training on the full shapes from the TOSCA dataset [12]. The problem is that the shapes and

poses used in the TOSCA full shape dataset also play active part in the test set of SHREC'16. The test set used by Cao et al. was thereby, unfortunately unintentionally, contaminated. Recently, to avoid test set contamination, the authors pretrained the models on four external datasets. However, the standard methodology in the field does not use any external datasets. As such, we did not include results from these updated models for fair comparison with all other methods. It turns out, that without such pre-training, the method struggles to compete with previous approaches like [27, 39] for partial shape matching.

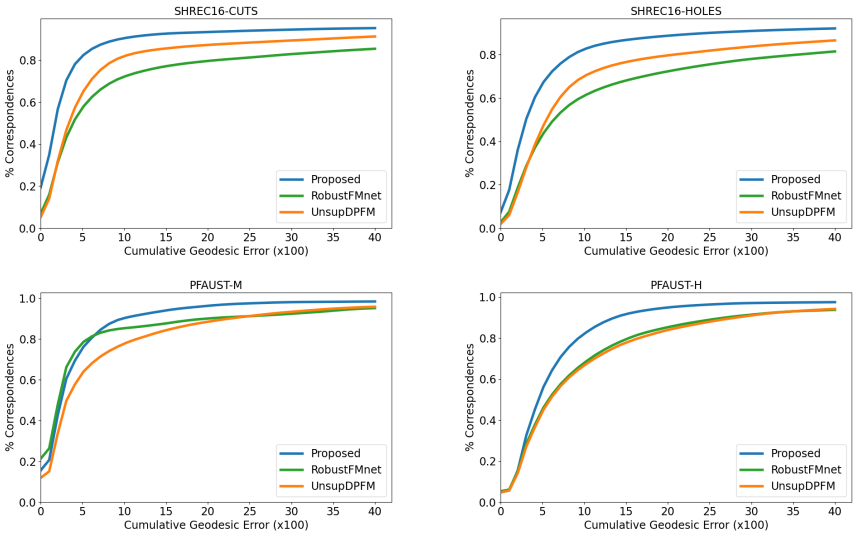

Fig. 5: PCK curves of existing unsupervised methods and ours on the test sets of SHREC'16 CUTS (top left), SHREC'16 HOLES (top right), PFAUST-M (bottom left) and PFAUST-H (bottom right). Our method is systematically superior compared to competing unsupervised approaches.

13 Further results

We provide the Percentage of Correct Keypoints (PCK) curves of ours and existing unsupervised methods, RobustFMnet [14] and UnsupDPFM [4], methods on the SHREC'16 and PFAUST benchmarks in Fig. 5. We also provide additional qualitative results of our method for partial shape matching on the SHREC'16 CUTS (Fig. 6) and HOLES (Fig. 7) datasets. Finally, we provide visual qualitative results on our new PFAUST-M (Fig. 8) and PFAUST-H (Fig. 9) datasets.

Fig. 6: Additional qualitative results on the SHREC'16 CUTS dataset. Zoom in for a better view.

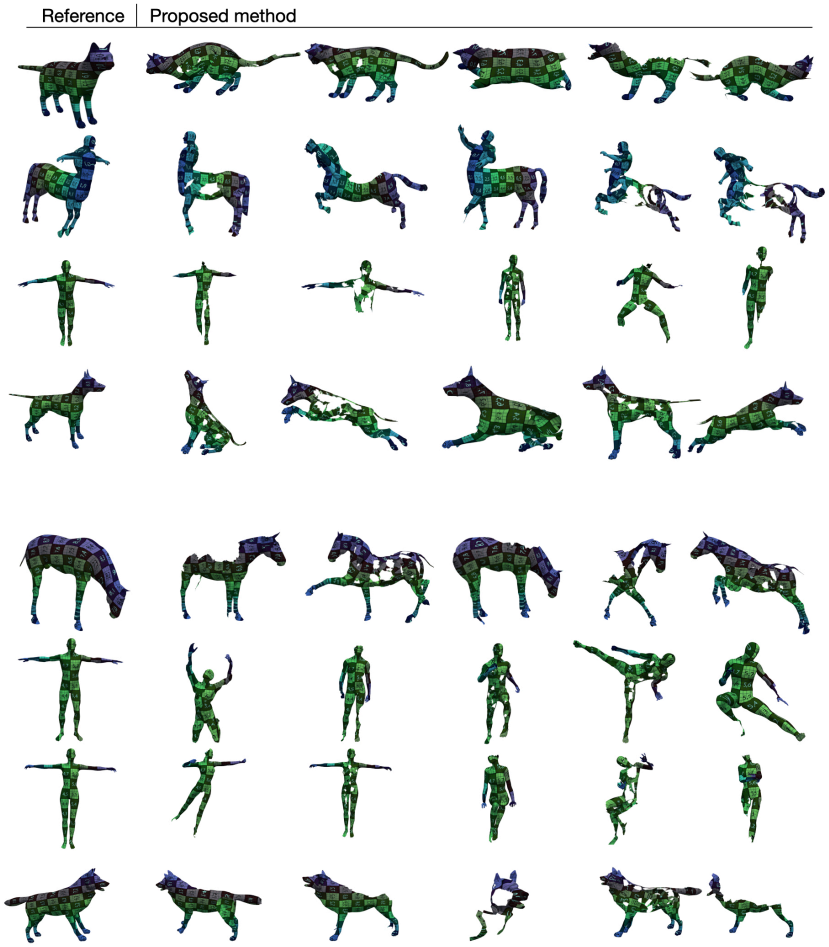

Fig. 7: Additional qualitative results on the SHREC'16 HOLES dataset. Zoom in for a better view.

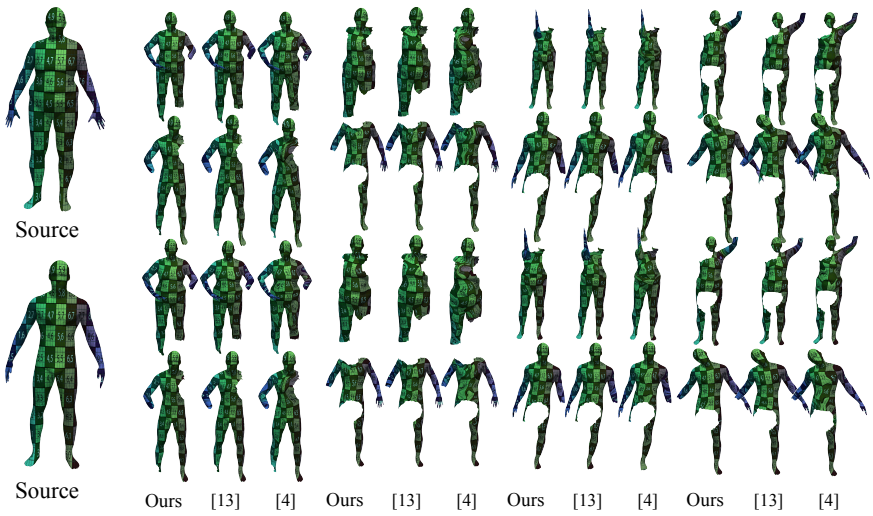

Fig. 8: Qualitative results on PFAUST-M of our method and RobustFMnet [14] and UnsupDPFM [4], zoom in for a better view. We obtain visually appealing results that outperform previous unsupervised methods. This figure also presents the shape partiality present in PFAUST-M, which mostly consists in body parts removal due to the size of the holes created on the original shapes.

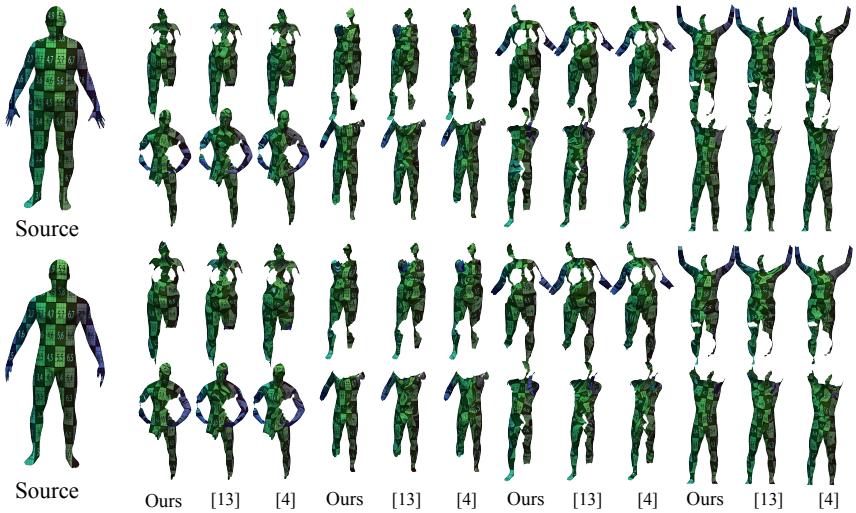

Fig. 9: Qualitative results on PFAUST-H of our method and RobustFMnet [14] and UnsupDPFM [4]. Zoom in for a better view. We obtain visually appealing results that outperform previous unsupervised methods. This figure also presents the shape partiality present in PFAUST-H, which involves extremely challenging topology with 13 holes.