Supplementary Material for Bending Graphs: Hierarchical Shape Matching using Gated Optimal Transport

This supplementary material explains further details of the implementation, adds some ablation studies, and provides some complementary experiments with qualitative results to support the paper.

1. Implementation details

Here are further details which were not included in the experiment section of the paper.

- d_{cut}: 7
- Feature size of local graph: 64
- Feature size of shape graph: 64
- Num. of Gated Feature Propagation: 2
- Loss weights:
  - First 30 epochs: $\gamma_D = 1, \gamma_M = 0, \gamma_R = 1$
  - After 30 epochs: $\gamma_D = 0.1, \gamma_M = 1, \gamma_R = 1$

Coarse-to-Fine Dense Matching We implement a simple algorithm based on functional maps to populate our matching to all mesh vertices densely. We provide 200 coarse matches as corresponding landmarks and create Wave Kernel Signature [1] of size 35 to find fine correspondences on all 6890 vertices.

2. Domain Transfer

To demonstrate the generalization ability of our network, we used the same model trained on SURREAL [4] dataset in section 4.3 to test on the TOSCA [3] horse class. The results are shown on Figure 1 and table 1. The results suggest the strong domain transfer capability of our learned model to unseen shapes.

	coarse error (↓)	fine error (↓)
horse	0.1988	0.0146
cat	0.2141	0.0223

Table 1. To evaluate the domain adaptability, we use the model trained with SURREAL dataset to test on TOSCA dataset. By comparing this result to table 1 in the main paper, we find that our method is able to achieve equivalent results compared to the model trained on SMAL dataset with animal shapes.

3. Ablation Studies

3.1. Number of GOT and GFP

The proposed GOT consists of a Sinkhorn optimal transport layer followed by a Gated Feature Propagation (GFP) (See section 3.3). This involves two hyper-parameters, the number of message passing layers and the number of the total GOT operation. We ablate the effect on these two factors following the same setup of the ablation studies we provided in the main paper (Sec. 4.5). The results are shown on table 2 and figure 3

By comparing #2 #4, it can be seen that the number of GFP does not linearly influence the result. Without GFP (#1), the patch features are unaware of the local manifold, thus only focusing on matching similar features, resulting in a reasonable bijection rate and high error. With adequate GFP to enforce the regularity of the adjacent features, the network can achieve the best bijection rate and error. However, when the propagation is performed in several hops, the patch features are bound too widely to the local geometry and cannot provide a correct and distinctive matching. The same phenomenon is observed in the number of GOT operations. By comparing #2, #5, and #6, we can observe that the more GOT operations, the lower the system's performance becomes.

By comparing #3 and #5, it can be seen that using two GOT operations with 2 GFPs is more effective than having a single GOT with 4 GFPs. Note that these two settings are not identical. The former uses different confidence values on each GOT operation, while the latter uses the same confidence value to do 4 GFPs. Re-estimating the confidence values through Sinkhorn algorithm allows more flexible feature propagation, and hence #5 has a better performance than #3. In our network design and the problem setup, the optimized configuration is with one GOT with 2 GFPs (#2).

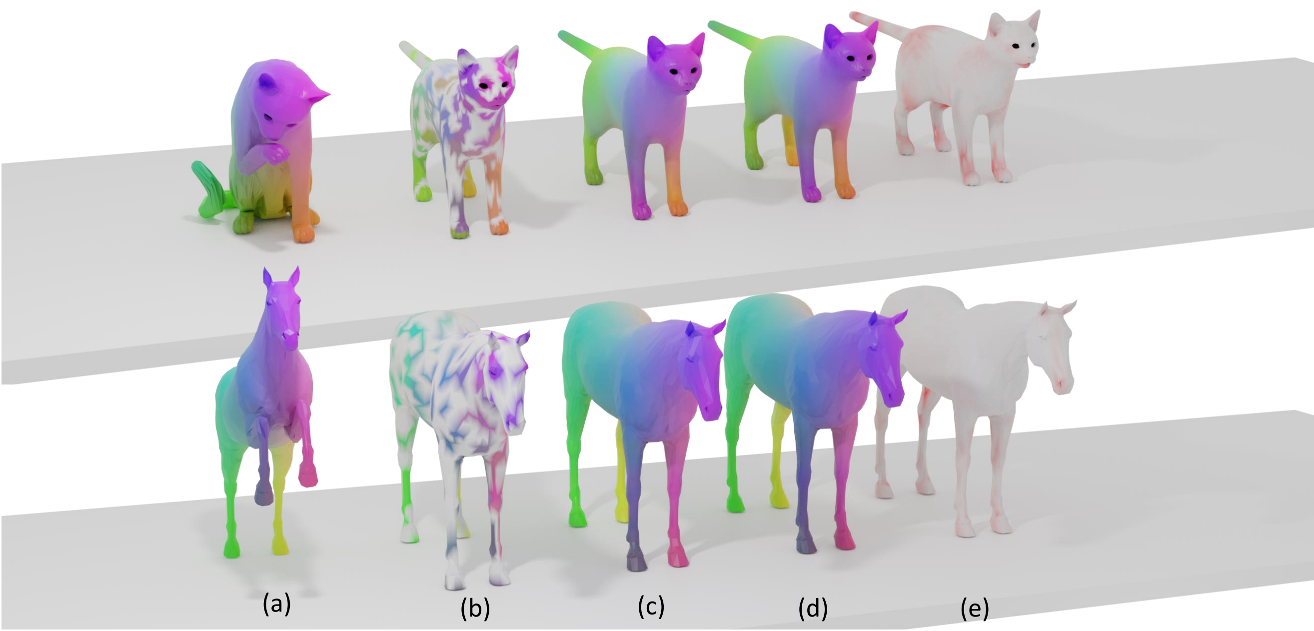

Figure 1. Results trained on SURREAL [4] and tested on Tosca [3]. In this experiment, we show the generalizability of the pipeline when trained on human shapes and tested on animal shapes. a) target shape. b) coarse matches c) source shape with the predicted correspondences d) ground-truth correspondences, e) normalized L2 error map.

	N. GOT	N. GFP	bij. rate (↑)	err. (↓)
#1	0	0	52.75	14.11
#2	1	2	**64.70**	**8.63**
#3	1	4	39.78	10.50
#4	1	8	36.78	10.99
#5	2	2	45.78	9.59
#6	3	2	36.60	11.31

Table 2. The ablation study on the number of gated optimal transport operations and the number of gated feature propagation layers. As proven by bijectivity rate and shortest path error, the study #2 with a single GOT and two layers of GFP performs the best.

3.2. Effect of GFP

To study the effect of the GFP module on matching, we do an ablation on matching results before and after GFP. Here we train our network with one GOT and 2 GFPs and show the matching results of the first Sinkhorn versus the final Sinkhorn. Figure 2 shows the results of coarse matching on the Faust dataset. In Table 3 we also compare coarse and fine geodesic errors with the outputs of each Sinkhorn layer. As visualized, the matching results after the GFP has lower error and less outliars.

	coarse error (↓)	fine error (↓)
Before GFP	0.1315	0.0229
After GFP	0.2896	0.0253

Table 3. We show the effect of GFP during evaluation on FAUST dataset by comparing the matching error before and after the GFP module. Both results are calculated from the output of the Sinkhorn algorithm.

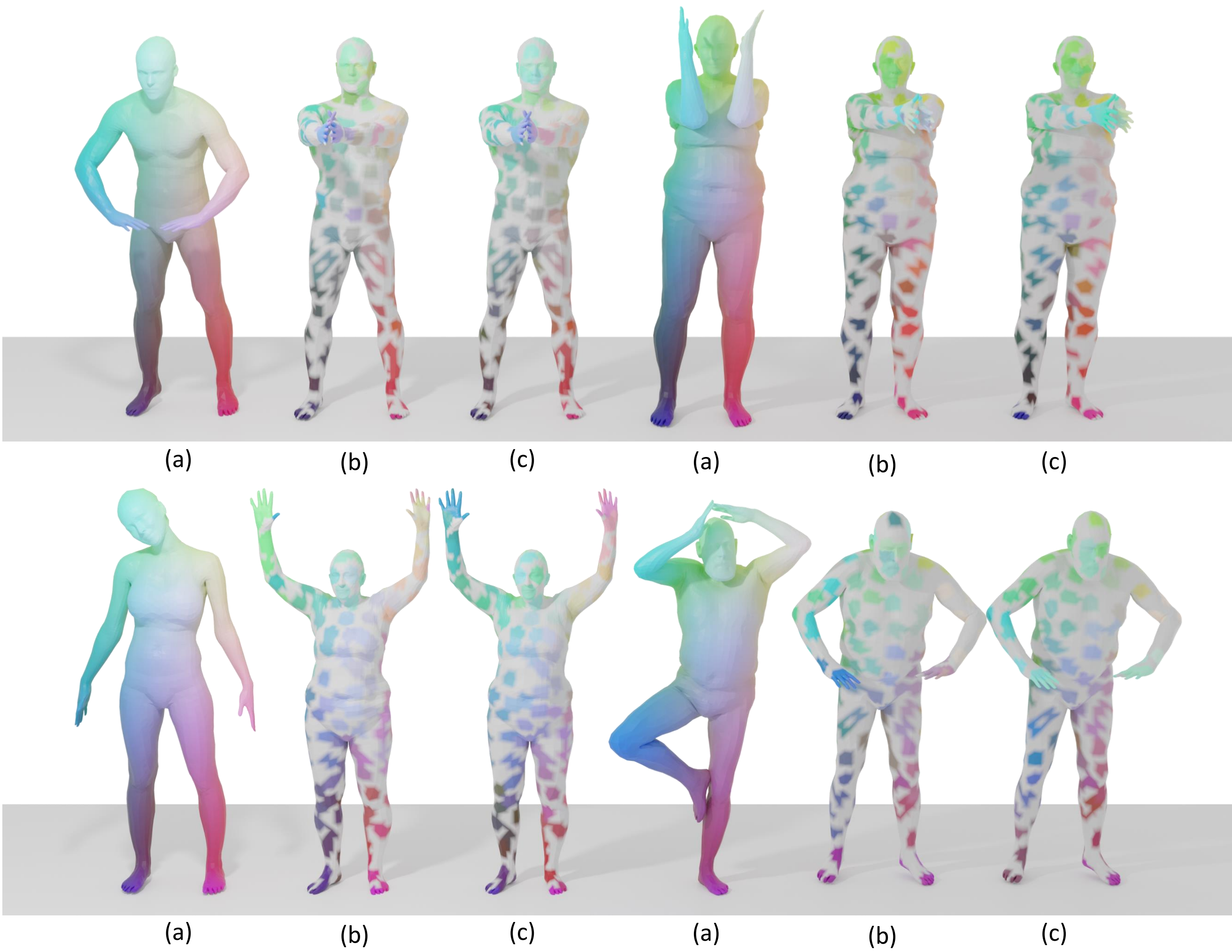

Figure 2. To show the effectiveness of our GFP module, we visualize correspondences before and after the GFP. Here a model is trained on SURREAL [4] and during evaluation on FAUST [2], we visualize the results from the first and final Sinkhorn. a) Correspondence map of target shape. b) Predicted correspondences from the first Sinkhorn (before GFP). c) Predicted correspondences from the final Sinkhorn (after the GFP).

Figure 3. We show the visual difference of the ablation study on the number of GFP layers (corresponding to #1 #2 and #3 on table 2. The human models from left to right are (a) The target shape map (b) The coarse matching result of model #1, (c) the coarse matching result of model #2, (d) the coarse matching result of model #3, (e) the dense matching result of model #2, (f) the ground truth dense matching result and (g) normalized L2 error map of model #2.

References

[1] Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *international conference on computer vision workshops (ICCV workshops)*, pages 1626–1633. IEEE, 2011. 1

[2] Federica Bogo, Javier Romero, Matthew Loper, and Michael J Black. Faust: Dataset and evaluation for 3d mesh registration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3794–3801, 2014. 3

[3] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer Science & Business Media, 2008. 1, 2

[4] Gul Varol, Javier Romero, Xavier Martin, Naureen Mahmood, Michael J Black, Ivan Laptev, and Cordelia Schmid. Learning from synthetic humans. In *CVPR*, pages 109–117, 2017. 1, 2, 3