Supplementary material of Stable-SCore: A Stable Registration-based Framework for 3D Shape Correspondence

Haolin Liu[1,2,3*], Xiaohang Zhan[2*], Zizheng Yan[1,4*], Zhongjin Luo[1,4], Yuxin Wen[2], Xiaoguang Han[4,1†]
[*]equal contribution [†]corresponding author
[1]FNii, CUHKSZ [2]Tencent [3]Tencent-Hunyuan3D [4]SSE, CUHKSZ
haolinliu97.github.io/Stable-Score

1. Character in-the-wild Benchmark

To better evaluate non-isometric shape correspondence, we introduce a new benchmark dataset, Character in-the-Wild (CharW). It consists of 100 character meshes collected from various sources, including artists and 3D generative models [11, 19]. Dense correspondences are annotated by manually deforming template meshes from 3DBiCar [13] and SMPL [12] to align with the collected shapes. By using a shared set of templates, ground-truth correspondences can be established between any pair of meshes in the benchmark. We compare CharW with several public character correspondence benchmarks [1, 4, 8, 10, 15] in Table 1. Unlike previous datasets, which are typically re-meshed versions of original mesh, CharW features greater diversity in shape and topology. Visualization examples are shown in Figure 1.

Figure 1. Visualized examples of the diverse shapes in our *Character in-the-wild* benchmark.

Table 1. Quantitative comparison between our *Character in-the-wild* benchmark dataset and other character correspondence benchmark dataset.

Dataset	# of identities	shape variance	Non-isometric	various sources
FAUST [4]	10	small		
SCAPE [1]	1	small		
SHREC'19 [15]	44	small		
TOPKIDS [8]	1	small		
DT4D-H [10]	10	large	✓	
CharW(Ours)	100	large	✓	✓

1.1. CharW Dataset Curation Details

Data Collection We collect meshes from two main sources: generated meshes and artist-crafted meshes. The generated meshes come from Tripo [11] and Rodin [19], while the artist-designed meshes are sourced from Mixamo and Sketchfab. The dataset includes 38 meshes from Mixamo, 23 from Rodin, 22 from Tripo3D, and 40 from Sketchfab. Several criteria were followed during data collection: 1. All meshes represent bipedal characters, including both human and humanoid figures. 2. The meshes must be of high quality, free of significant distortion or artifacts, and avoid low-poly models. 3. We selected characters with diverse body types such as fat, slim, large-headed, and small-headed, to ensure a variety of non-isometric shapes.

Correspondence Annotation CharW benchmark provides correspondence annotations for evaluation. The annotation process begins by selecting the SMPL neutral template mesh and three meshes from 3DBiCar [13] as deformation templates. For each target mesh, a professional artist first selects the template that most closely resembles the target. The artist then annotates key-point correspondences between the source and target meshes, ensuring that the key points are semantically aligned. On average, 60 to 80 key points are annotated per mesh. After annotation, the artist uses ZBrush's warp add-on to align the source mesh with the target. The annotated key points are refined until the results are satisfactory. If needed, the artist manually adjusts the warped mesh using sculpting tools. Finally, the warped mesh is processed with Blender's Shrinkwrap modifier to precisely match the target.

2. Supplementary

2.1. Limitation and Future work

There are several limitations stemming from the 2D correspondence model: (1) significant initial rotation misalignment, which leads to incorrect 2D correspondences and distorts the deformation process; (2) severe occlusion, resulting in missing 2D correspondences; and (3) difficulty handling complex structures due to the low resolution of 2D features, making it challenging to capture details such as fingers, as shown in the first row of Fig 4.

The deformation process also has limitations: (1) topological noise, such as the body and arms merging together, hinders deformation; (2) it inherits limitations from NJF, including the inability to handle partial shapes. An interesting direction for future work is to move beyond the limitations of low-resolution 2D prior models by integrating native 3D prior models, such as large 3D generative models [11, 19, 20], which could potentially address these issues.

2.2. Discussion on Supervision Types

Previous functional map-based methods [2, 3, 5–7, 9, 14, 16, 18] are mostly unsupervised. Supervised methods [5] are rare in shape correspondence, and previous supervised attempts have underperformed compared to unsupervised methods. While supervised methods generally outperform unsupervised ones in most computer vision tasks, this is not the case for shape correspondence. We hypothesize that design flaws in earlier supervised methods, such as strong reliance on LBO's basis or DiffusionNet's prior [17], lead to unsatisfactory results and overfitting problems, suggesting significant room for improvement.

To address this, we propose a supervised, registration-based method that avoids reliance on functional maps or DiffusionNet. Our approach achieves state-of-the-art performance on various non-isometric benchmarks. Although it requires ground-truth (GT) supervision, the annotation cost is relatively low, requiring only sparse 2D/3D key points or dense correspondences via template warping (same as the CharW annotation process). This makes our method highly cost-effective, therefore, the need for GT supervision is no longer a limitation compared to unsupervised approaches.

2.3. Beyond Characters

We also test the applicability of our method to other domains, such as animal shape correspondence, which is commonly evaluated in non-isometric shape matching. Additionally, we compare our method on the SMAL dataset, as shown in Table 2 and Figure 2. The quantitative comparisons are conducted under two setups: training on a character dataset or on the SMAL dataset, with testing performed on the SMAL dataset. "Ours (zero-shot)" refers to a zero-shot version of our method, where feature adapters are removed, and no further fine-tuning is required. Our method outperforms others and demonstrates its ability to generalize to domains where ground truth correspondences are available for training. Furthermore, the zero-shot setup of our method is versatile, achieving strong performance across various tasks.

Table 2. Quantitative comparison between our methods and previous methods on SMAL dataset.

Test on SMAL	Train on character	Train on SMAL
ULRSSM	28.5	3.63
Hybrid ULRSSM	44.0	3.11
Ours (zero-shot)	**8.91**	8.91
Ours (full)	17.01	**2.65**

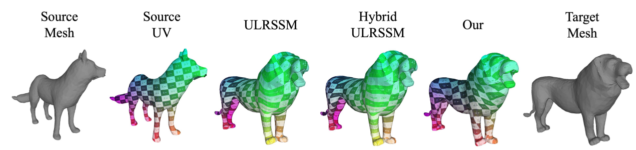

Figure 2. Comparison on SMAL, all methods are trained on SMAL.

3. More results

More results on the DT4D dataset and CharW dataset are shown in Figure 3, 4, 5 and 6. It can be observed that our Stable-Score achieves precise registration while simultaneously preserving the source topology, thereby demonstrating significant potential for re-topology applications.

References

[1] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. Scape: shape completion and animation of people. In *ACM SIGGRAPH 2005 Papers*, pages 408–416. 2005. 1

[2] Souhaib Attaiki and Maks Ovsjanikov. Shape non-rigid kinematics (snk): A zero-shot method for non-rigid shape matching via unsupervised functional map regularized reconstruction. *Advances in Neural Information Processing Systems*, 36:70012–70032, 2023. 2

[3] Souhaib Attaiki and Maks Ovsjanikov. Understanding and improving features learned in deep functional maps. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1316–1326, 2023. 2

[4] 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. 1

[5] Nicolas Donati, Abhishek Sharma, and Maks Ovsjanikov. Deep geometric functional maps: Robust feature learning for shape correspondence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8592–8601, 2020. 2

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

[7] Puhua Jiang, Mingze Sun, and Ruqi Huang. Non-rigid shape registration via deep functional maps prior. In *Advances in Neural Information Processing Systems*, pages 58409–58427. Curran Associates, Inc., 2023. 2

[8] Zorah Lähner, Emanuele Rodolà, Michael M Bronstein, Daniel Cremers, Oliver Burghard, Luca Cosmo, Alexander Dieckmann, Reinhard Klein, Y Sahillioğlu, et al. Shrec'16: Matching of deformable shapes with topological noise. In *Eurographics Workshop on 3D Object Retrieval, EG 3DOR*, pages 55–60. Eurographics Association, 2016. 1

[9] Qinsong Li, Yueyu Guo, Xinru Liu, Ling Hu, Feifan Luo, and Shengjun Liu. Deformable shape matching with multiple complex spectral filter operator preservation. *The Visual Computer*, 40(7):4885–4898, 2024. 2

[10] Yang Li, Hikari Takehara, Takafumi Taketomi, Bo Zheng, and Matthias Nießner. 4dcomplete: Non-rigid motion estimation beyond the observable surface. *IEEE International Conference on Computer Vision (ICCV)*, 2021. 1

[11] Yangguang Li, Zi-Xin Zou, Zexiang Liu, Dehu Wang, Yuan Liang, Zhipeng Yu, Xingchao Liu, Yuan-Chen Guo, Ding Liang, Wanli Ouyang, et al. Triposg: High-fidelity 3d shape synthesis using large-scale rectified flow models. *arXiv preprint arXiv:2502.06608*, 2025. 1, 2

[12] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. SMPL: A skinned multi-person linear model. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, 34(6):248:1–248:16, 2015. 1

[13] Zhongjin Luo, Shengcai Cai, Jinguo Dong, Ruibo Ming, Liangdong Qiu, Xiaohang Zhan, and Xiaoguang Han. Rabit: Parametric modeling of 3d biped cartoon characters with a topological-consistent dataset. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 1

[14] Robin Magnet and Maks Ovsjanikov. Scalable and efficient functional map computations on dense meshes. In *Computer Graphics Forum*, pages 89–101. Wiley Online Library, 2023. 2

[15] Simone Melzi, Riccardo Marin, Emanuele Rodolà, Umberto Castellani, Jing Ren, Adrien Poulenard, Peter Wonka, and Maks Ovsjanikov. Shrec 2019: Matching humans with different connectivity. In *Eurographics Workshop on 3D Object Retrieval*, page 3. The Eurographics Association, 2019. 1

[16] Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (ToG)*, 31(4):1–11, 2012. 2

[17] Nicholas Sharp, Souhaib Attaiki, Keenan Crane, and Maks Ovsjanikov. Diffusionnet: Discretization agnostic learning on surfaces. *ACM Transactions on Graphics (TOG)*, 41(3):1–16, 2022. 2

[18] Mingze Sun, Shiwei Mao, Puhua Jiang, Maks Ovsjanikov, and Ruqi Huang. Spatially and spectrally consistent deep functional maps. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14497–14507, 2023. 2

[19] Longwen Zhang, Ziyu Wang, Qixuan Zhang, Qiwei Qiu, Anqi Pang, Haoran Jiang, Wei Yang, Lan Xu, and Jingyi Yu. Clay: A controllable large-scale generative model for creating high-quality 3d assets. *ACM Transactions on Graphics (TOG)*, 43(4):1–20, 2024. 1, 2

[20] Zibo Zhao, Zeqiang Lai, Qingxiang Lin, Yunfei Zhao, Haolin Liu, Shuhui Yang, Yifei Feng, Mingxin Yang, Sheng Zhang, Xianghui Yang, et al. Hunyuan3d 2.0: Scaling diffusion models for high resolution textured 3d assets generation. *arXiv preprint arXiv:2501.12202*, 2025. 2

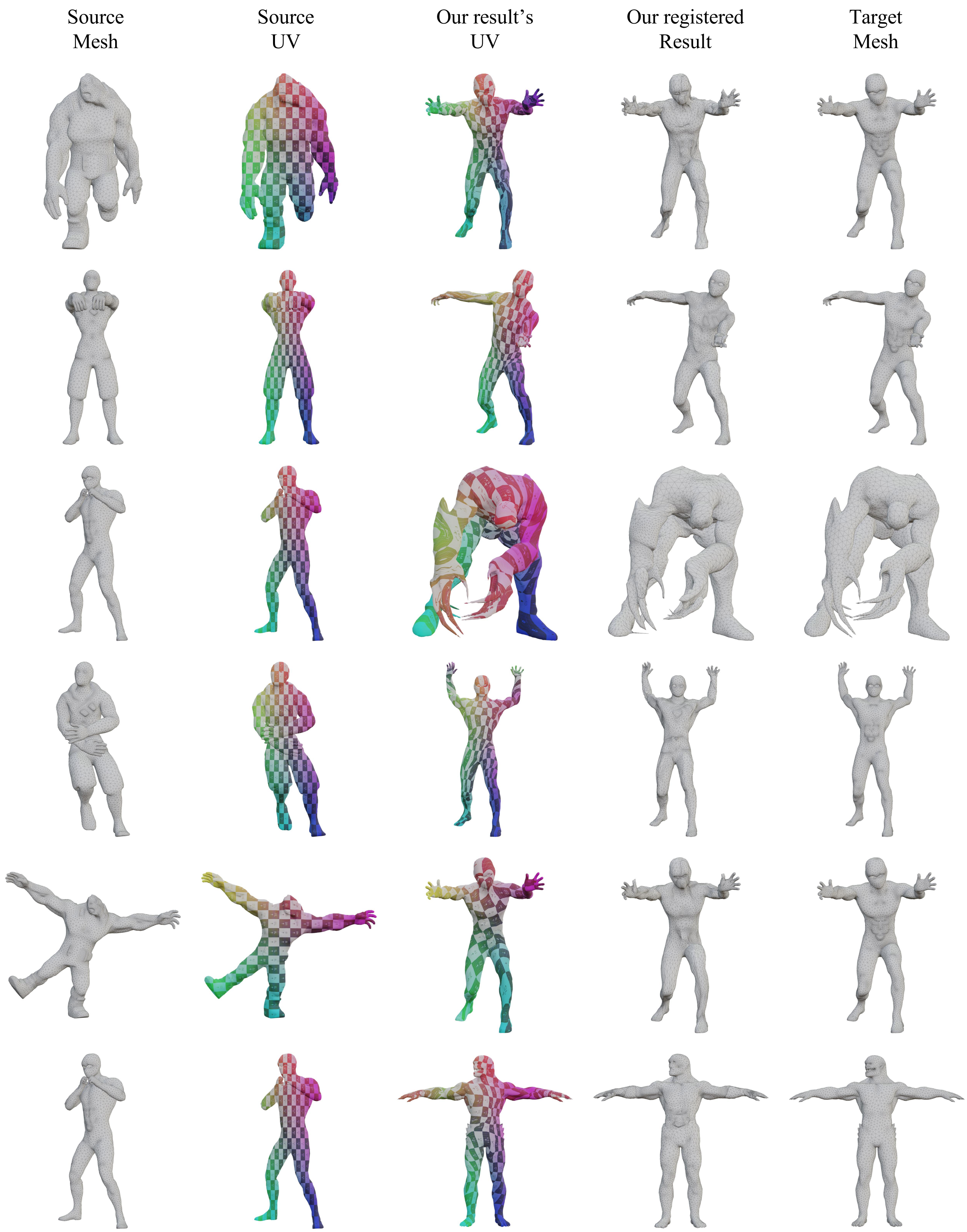

Figure 3. Visualized results produced by Stable-Score on DT4D dataset.

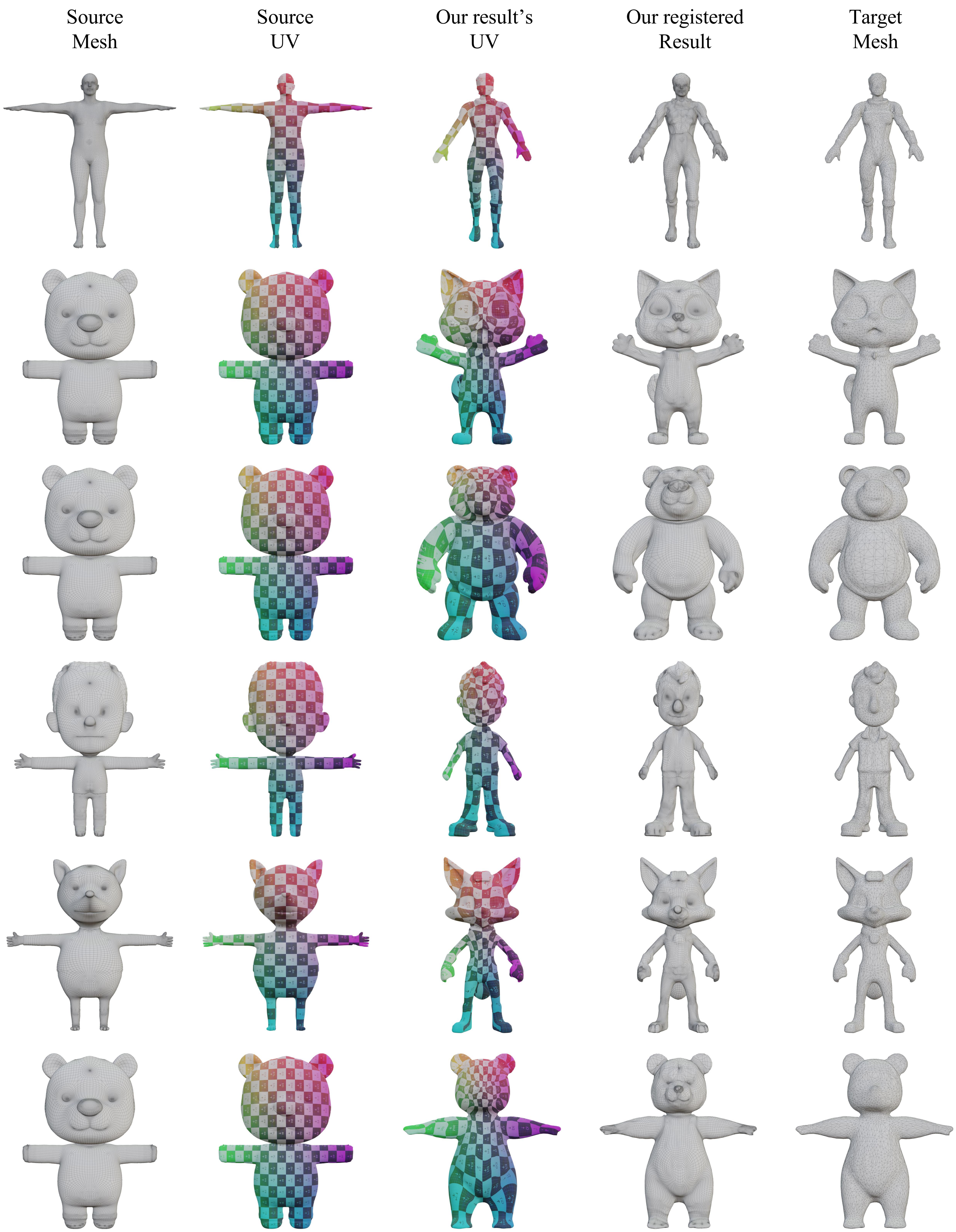

Figure 4. Visualized results produced by Stable-Score on our CharW dataset.

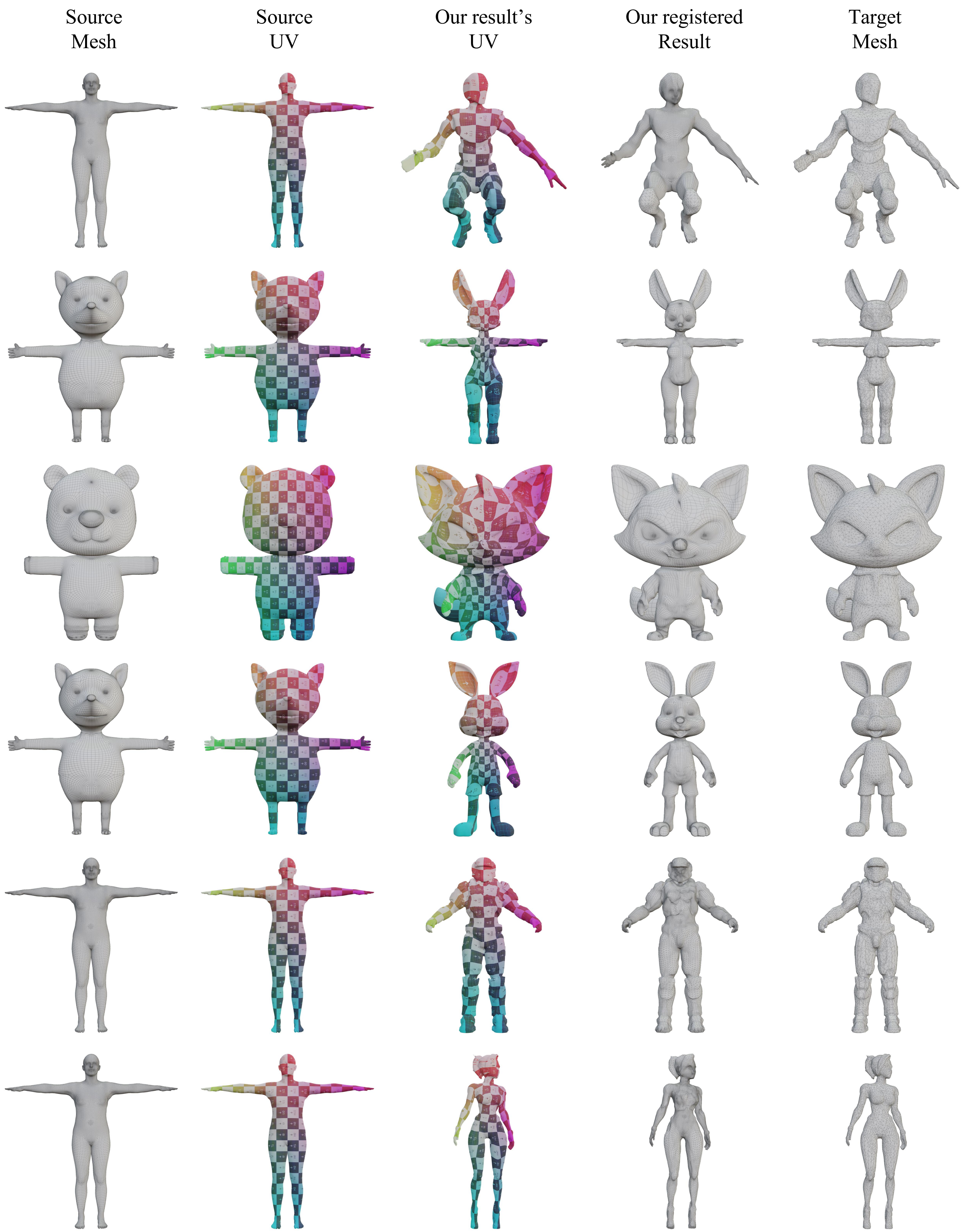

Figure 5. Visualized results produced by Stable-Score on our CharW dataset.

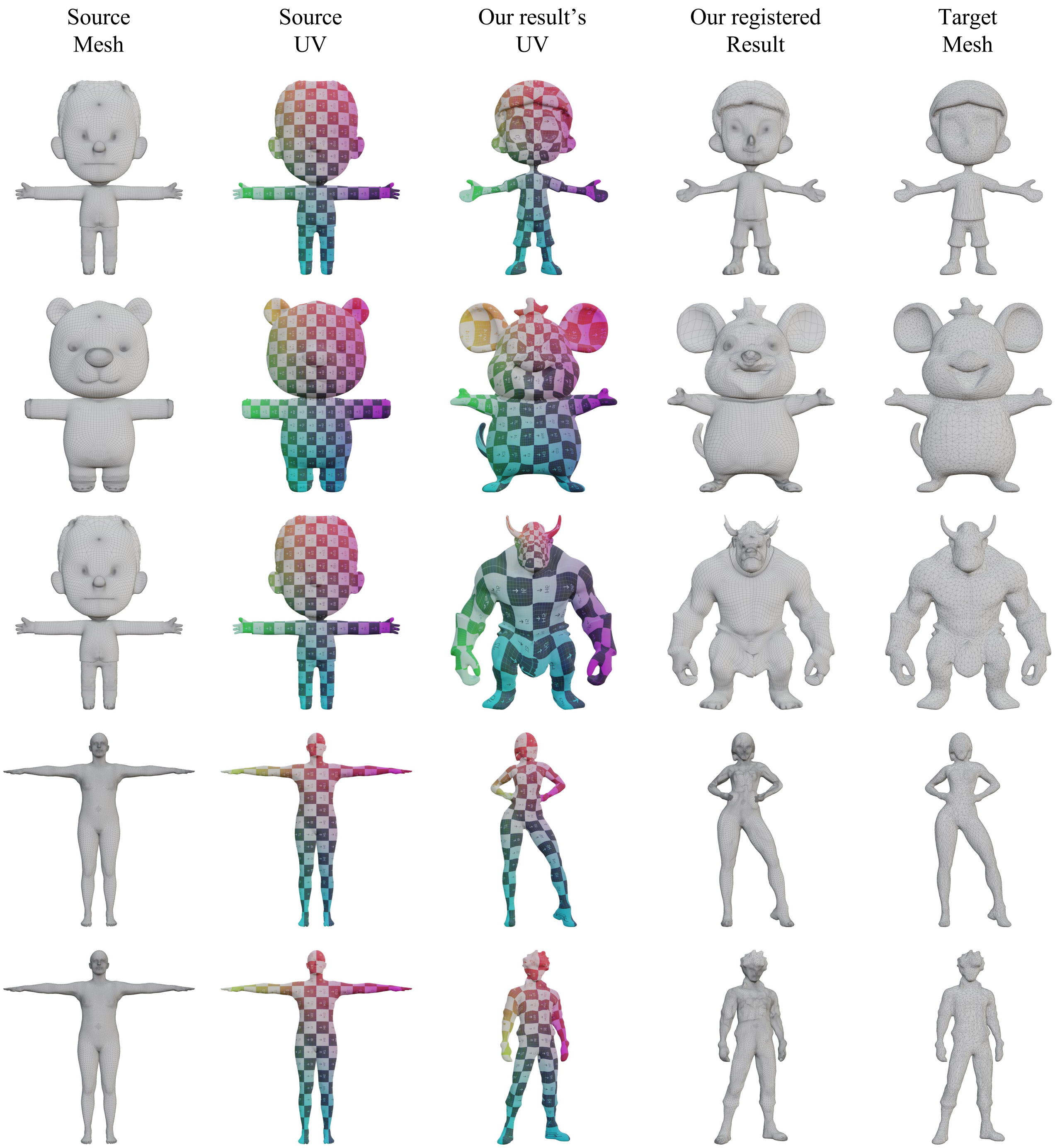

Figure 6. Visualized results produced by Stable-Score on our CharW dataset.