Supplementary Document for the Paper Entitled ‘Differentiable Rendering using RGBXY Derivatives and Optimal Transport’

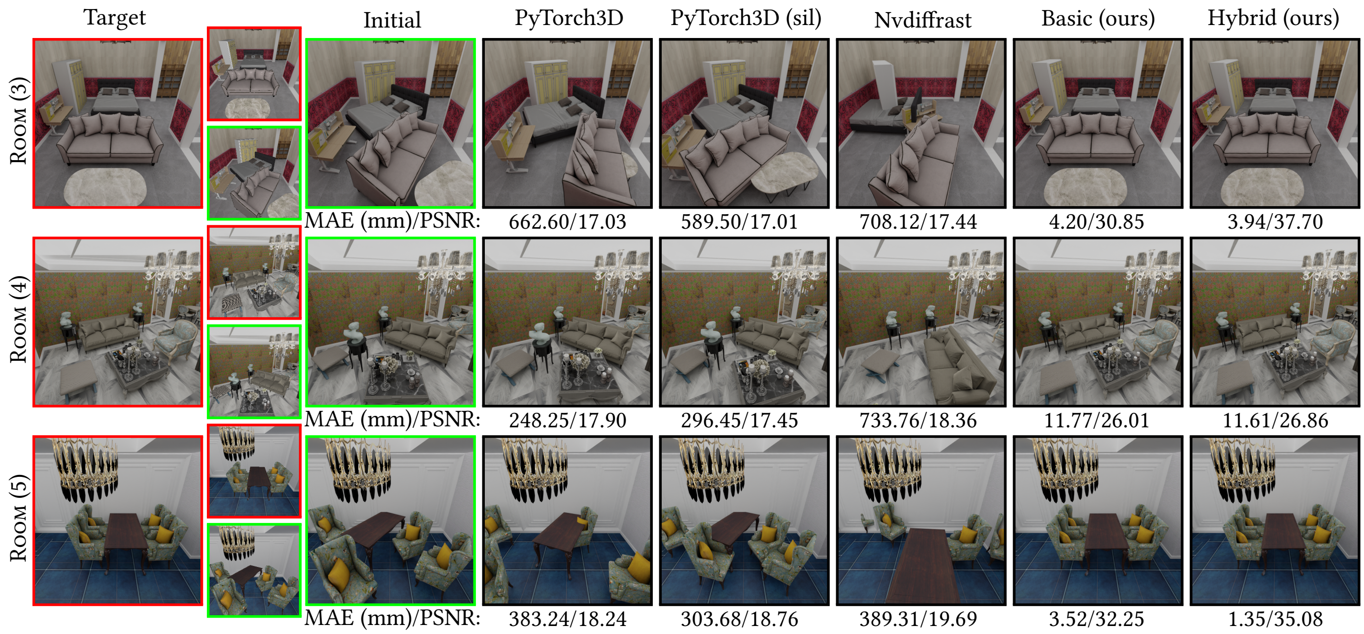

Fig. 1. More examples of the furniture layout adjustment application. Scenes are selected from [Fu et al. 2021]. (Note that optimization is performed using rasterized images under 256 × 256 resolution, as shown in the second column, while other images are re-rendered in a high resolution with ray tracing for higher display quality)

ACM Reference Format:
. 2022. Supplementary Document for the Paper Entitled ‘Differentiable Rendering using RGBXY Derivatives and Optimal Transport’. *ACM Trans. Graph.* 1, 1 (September 2022), 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 MORE EVALUATION DATAS

We provide an extend evaluation table (i.e., extending Table 1 in the paper) in Table 1, which provides all error metrics including MAE, PSNR, SSIM, RMSE, and LPIPS. Furthermore, we provide an equal-time comparison in Table 2. For each scene, the running time limit is set to the time for our basic method (without hybrid strategy) to run for 1000 iterations with weight $\lambda = 0.5$ and match interval $K = 5$. Other settings are the same as those in the paper. Our hybrid method still perform the best under the equal time setting, although Nvdiffrast [Laine et al. 2020] is faster and performs more iterations of optimization.

Author’s address:

0730-0301/2022/9-ART
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 MORE APPLICATION RESULTS

In this section, we show more results of our inverse rendering applications including furniture layout adjustment, human pose fitting and facial expression reconstruction. Please refer to the paper for detailed experiment settings.

Fig. 1 and Fig. 2 show 10 more examples of furniture layout adjustment. In Fig. 3, we show 7 more examples of human pose fitting. In Fig. 4, we show 7 more examples of facial expression reconstruction. All experiments consistently show that our method outperforms baseline methods, including Nvdiffrast [Laine et al. 2020] and PyTorch3D [Ravi et al. 2020]. In the majority of examples, our hybrid strategy further improves reconstruction accuracy.

REFERENCES

Huan Fu, Bowen Cai, Lin Gao, Ling-Xiao Zhang, Jiaming Wang, Cao Li, Qixun Zeng, Chengyue Sun, Rongfei Jia, Binqiang Zhao, et al. 2021. 3D-Front: 3D Furnished Rooms with Layouts and Semantics. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 10933–10942.

Samuli Laine, Janne Hellsten, Tero Karras, Yeongho Seol, Jaakko Lehtinen, and Timo Aila. 2020. Modular primitives for high-performance differentiable rendering. *ACM Transactions on Graphics (TOG)* 39, 6 (2020), 1–14.

Nikhila Ravi, Jeremy Reizenstein, David Novotny, Taylor Gordon, Wan-Yen Lo, Justin Johnson, and Georgia Gkioxari. 2020. Accelerating 3d deep learning with pytorch3d. *arXiv preprint arXiv:2007.08501* (2020).

Table 1. Evaluation results on 8 scenes with 4 or 5 error metrics provided.

Scene		Kitty					Buddha					CornellBox				Teapot			
Metric		MAE	PSNR	SSIM	RMSE	LPIPS	MAE	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS
(a) matching algo.(K = 5,basic)	op.flow	0.010	34.25	0.975	0.011	0.020	**0.342**	22.41	0.771	0.121	0.072	12.77	0.530	0.223	0.211	17.43	0.905	0.187	0.132
	bipart.	0.013	33.68	0.978	0.007	0.021	0.433	23.09	0.772	**0.033**	0.064	14.25	0.620	0.027	0.191	**53.10**	0.997	**0.000**	0.006
(b) param λ (K = 5,basic)	λ = 0.01	0.639	19.16	0.811	0.375	0.129	0.970	18.06	0.571	0.415	0.125	14.71	0.783	0.141	0.113	44.86	0.997	0.007	**0.003**
	λ = 0.1	0.332	20.74	0.867	0.091	0.227	0.775	19.01	0.633	0.107	0.336	16.85	0.840	0.096	0.081	35.95	0.986	0.018	0.017
	λ = 0.9	**0.001**	45.63	0.998	0.006	**0.002**	0.442	22.55	0.789	0.065	0.131	22.89	0.930	0.062	**0.030**	32.39	0.961	0.028	0.032
	λ = 0.99	**0.001**	**47.38**	**0.999**	**0.001**	0.005	0.429	22.73	0.802	0.130	0.063	23.69	0.940	0.025	0.054	33.27	0.969	0.026	0.025
(c) match interval K (λ = 0.5,basic)	K = 0	0.002	42.04	0.995	0.003	0.008	0.457	24.50	0.869	0.088	0.045	22.76	0.932	0.028	0.059	33.88	0.967	0.028	0.024
	K = 1	0.003	41.98	0.995	0.003	0.008	0.460	24.54	0.870	0.088	**0.044**	22.34	0.923	0.030	0.063	33.91	0.967	0.027	0.024
	K = 10	0.002	42.25	0.995	0.003	0.008	0.489	21.66	0.755	0.159	0.075	22.18	0.920	0.034	0.067	33.42	0.967	0.026	0.026
	K = 50	0.011	33.88	0.974	0.012	0.021	0.708	18.81	0.623	0.340	0.115	21.29	0.909	0.033	0.069	31.66	0.962	0.032	0.029
(d) optim. strategy $K = 5,\lambda = 0.5$)	rand	0.419	24.32	0.952	0.155	0.044	0.806	20.49	0.725	0.245	0.084	15.56	0.820	0.152	0.156	49.59	**0.999**	0.001	0.004
	hybrid	0.003	41.53	0.995	0.004	0.008	0.467	24.53	0.867	0.088	0.045	**25.28**	0.952	**0.018**	0.053	49.08	**0.999**	0.001	0.004
	basic	0.002	42.41	0.996	0.003	0.008	0.480	22.88	0.803	0.120	0.063	22.69	0.931	0.030	0.060	33.86	0.968	0.024	0.025
(e) baseline	nvdf.res.128	0.487	24.47	0.957	0.152	0.041	0.857	21.02	0.746	0.215	0.077	14.54	0.796	0.191	0.171	50.16	**0.999**	0.001	0.004
	nvdf.res.1024	0.491	24.56	0.966	0.041	0.028	0.751	20.69	0.787	0.081	0.198	18.02	0.935	0.084	0.077	39.18	0.989	0.011	0.021
	nvdfmul.res	0.404	24.28	0.967	0.039	0.030	0.772	19.47	0.744	0.097	0.242	13.38	0.828	0.197	0.197	41.81	0.987	0.009	0.028
	nvdf.down.res	0.635	24.60	0.968	0.039	0.028	0.449	**25.71**	**0.889**	0.040	0.047	18.91	**0.953**	0.076	0.061	38.05	0.970	0.013	0.045
Scene		Cube					Ginger					Outdoor				Joint			
Metric		MAE	PSNR	SSIM	RMSE	LPIPS	MAE	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS
(a) matching algo.(K = 5,basic)	op.flow	1.020	17.33	0.906	0.180	0.136	0.477	18.93	0.792	0.199	0.113	18.97	0.813	0.194	0.113	18.87	0.840	0.184	0.114
	bipart.	0.171	24.54	0.947	0.020	0.054	0.598	14.33	0.355	0.109	0.192	18.95	0.800	0.054	0.121	16.94	0.540	0.077	0.142
(b) param λ (K = 5,basic)	λ = 0.01	0.920	16.98	0.908	0.214	0.118	0.535	14.80	0.625	0.312	0.182	22.28	0.905	0.045	0.080	16.56	0.731	0.673	0.149
	λ = 0.1	0.007	**38.29**	**0.998**	0.009	**0.002**	0.357	25.18	0.943	0.055	0.045	18.66	0.828	0.121	0.115	16.56	0.731	0.149	0.673
	λ = 0.9	0.063	27.28	0.981	0.040	0.017	0.306	29.68	0.977	0.033	0.033	15.89	0.741	0.162	0.236	25.25	0.886	0.055	0.036
	λ = 0.99	0.066	27.11	0.980	0.018	0.041	0.309	29.14	0.979	0.021	0.035	16.79	0.736	0.274	0.147	25.89	0.907	0.045	0.051
(c) match interval K (λ = 0.5,basic)	K = 0	0.066	27.35	0.982	0.017	0.038	0.302	31.26	0.981	0.013	0.027	17.04	0.779	0.173	0.143	22.84	0.843	0.059	0.072
	K = 1	0.062	27.59	0.983	0.016	0.037	0.308	28.28	0.970	0.033	0.039	17.10	0.781	0.172	0.142	22.05	0.827	0.071	0.079
	K = 10	0.047	29.06	0.987	0.012	0.031	0.300	30.20	0.975	0.012	0.031	17.19	0.785	0.160	0.142	23.96	0.865	0.050	0.063
	K = 50	0.048	28.91	0.986	0.012	0.032	**0.297**	31.34	0.980	0.010	0.027	16.63	0.793	0.169	0.150	20.15	0.779	0.119	0.098
(d) optim. strategy $K = 5,\lambda = 0.5$)	rand	1.666	18.77	0.947	0.254	0.088	0.537	25.85	0.924	0.118	0.051	54.02	**1.000**	**0.001**	0.004	21.69	0.898	0.127	0.082
	hybrid	**0.006**	37.87	**0.998**	**0.002**	0.013	0.305	**35.14**	**0.990**	**0.004**	**0.018**	51.21	**1.000**	**0.001**	0.004	**44.19**	**0.996**	**0.001**	**0.006**
	basic	0.052	28.46	0.986	0.013	0.033	0.304	28.69	0.972	0.037	0.023	16.85	0.779	0.146	0.182	24.00	0.867	0.063	0.054
(e) baseline	nvdf.res.128	1.887	18.98	0.951	0.249	0.086	0.560	24.68	0.913	0.172	0.058	**54.50**	**1.000**	**0.001**	0.004	21.69	0.899	0.126	0.082
	nvdf.res.1024	1.878	19.11	0.967	0.084	0.048	0.551	28.14	0.978	0.039	0.071	53.25	0.999	0.002	**0.001**	23.36	0.950	0.068	0.056
	nvdfmul.res	2.217	17.68	0.951	0.118	0.071	0.653	22.24	0.952	0.077	0.150	44.42	0.986	0.006	0.011	21.38	0.922	0.085	0.082
	nvdf.down.res	1.359	20.00	0.974	0.068	0.037	0.581	29.06	0.973	0.035	0.102	51.49	0.998	0.003	**0.001**	23.36	0.950	0.068	0.055

Table 2. Evaluation results on 8 scenes in an equal-time setting. For each scene, the running time limit is set to the time for our basic method (without hybrid strategy) to run for 1000 iterations with weight $\lambda = 0.5$ and match interval $K = 5$.

Scene		Kitty					Buddha					CornellBox				Teapot			
Metric		MAE	PSNR	SSIM	RMSE	LPIPS	MAE	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS
(a) matching algo.(K = 5,basic)	op.flow	0.015	32.49	0.965	0.024	0.015	**0.373**	21.78	0.745	0.079	0.144	12.53	0.525	0.221	0.243	17.36	0.906	0.133	0.189
	bipart.	0.014	33.34	0.976	0.022	0.008	0.432	23.15	0.776	0.061	**0.032**	13.95	0.607	0.198	0.028	**51.13**	0.997	0.006	**0.000**
(b) param λ (K = 5,basic)	λ = 0.01	0.638	19.07	0.811	0.134	0.375	0.984	18.09	0.572	0.124	0.413	14.71	0.783	0.113	0.140	44.90	0.997	0.007	0.003
	λ = 0.1	0.354	20.75	0.874	0.086	0.231	0.774	19.07	0.638	0.106	0.333	16.91	0.841	0.095	0.082	35.96	0.986	0.018	0.017
	λ = 0.9	0.002	44.50	0.997	0.006	0.002	0.443	22.55	0.791	0.065	0.131	23.16	0.934	0.060	0.028	32.39	0.961	0.028	0.031
	λ = 0.99	**0.001**	**46.95**	**0.998**	0.005	**0.001**	0.434	22.65	0.796	0.064	0.131	23.83	0.942	0.052	0.023	33.12	0.968	0.025	0.026
(c) match interval K (λ = 0.5,basic)	K = 0	0.003	41.68	0.995	0.009	0.003	0.478	24.03	0.847	0.050	0.093	15.63	0.795	0.143	0.115	32.85	0.963	0.026	0.030
	K = 1	0.003	41.85	0.995	0.008	0.004	0.473	24.12	0.850	0.049	0.093	18.96	0.867	0.093	0.054	33.48	0.966	0.025	0.027
	K = 10	0.002	42.37	0.995	0.008	0.003	0.494	22.25	0.778	0.069	0.138	**23.97**	0.942	0.054	0.027	33.72	0.967	0.025	0.025
	K = 50	0.003	41.54	0.995	0.009	0.004	0.590	19.88	0.679	0.099	0.259	23.86	0.940	0.052	0.024	32.84	0.965	0.027	0.028
(d) optim. strategy $K = 5,\lambda = 0.5$)	rand	0.439	24.28	0.949	0.045	0.154	0.788	21.32	0.756	0.076	0.196	15.81	0.828	0.150	0.145	49.86	0.999	**0.004**	0.001
	hybrid	0.002	43.16	0.996	0.007	0.003	0.470	24.66	0.869	0.044	0.088	23.79	0.941	0.062	**0.022**	50.13	0.999	**0.004**	0.001
	basic	0.002	42.41	0.996	**0.003**	0.008	0.480	22.88	0.803	0.120	0.063	22.69	0.931	**0.030**	0.060	33.86	0.968	0.024	0.025
(e) baseline	nvdf.res.128	0.536	24.56	0.953	0.150	0.041	0.789	20.78	0.735	0.080	0.231	13.66	0.760	0.185	0.213	50.22	**1.000**	**0.004**	0.001
	nvdf.res.1024	0.519	24.57	0.968	0.038	0.028	0.803	21.01	0.801	0.076	0.187	19.20	**0.956**	0.068	0.054	40.11	0.987	0.010	0.022
	nvdfmul.res	0.390	24.10	0.963	0.043	0.033	0.830	18.94	0.709	0.109	0.275	12.51	0.761	0.228	0.237	42.00	0.990	0.009	0.023
	nvdf.down.res	0.390	24.14	0.966	0.040	0.030	0.448	**25.72**	**0.889**	**0.040**	0.046	18.94	0.953	0.074	0.060	41.16	0.992	0.011	0.015
Scene		Cube					Ginger					Outdoor				Joint			
Metric		MAE	PSNR	SSIM	RMSE	LPIPS	MAE	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS	PSNR	SSIM	RMSE	LPIPS
(a) matching algo.(K = 5,basic)	op.flow	1.088	15.98	0.885	0.159	0.250	0.397	18.73	0.804	0.116	0.205	18.98	0.814	0.113	0.198	18.95	0.842	0.113	0.186
	bipart.	0.140	25.59	0.958	0.050	0.015	0.600	13.06	0.368	0.222	0.122	18.95	0.800	0.121	0.054	17.67	0.584	0.131	0.069
(b) param λ (K = 5,basic)	λ = 0.01	0.884	17.10	0.909	0.117	0.205	0.539	14.48	0.612	0.189	0.317	22.30	0.906	0.080	0.045	16.56	0.731	0.149	0.673
	λ = 0.1	0.006	**39.52**	**0.999**	**0.007**	**0.001**	0.359	24.59	0.937	0.059	0.061	18.52	0.828	0.122	0.110	16.56	0.731	0.149	0.673
	λ = 0.9	0.072	26.67	0.978	0.043	0.020	0.306	29.55	0.977	0.033	0.034	16.67	0.745	0.150	0.243	25.17	0.885	0.055	0.038
	λ = 0.99	0.079	26.14	0.976	0.045	0.024	0.308	28.89	0.977	0.036	0.027	16.79	0.736	0.147	0.273	25.79	0.905	0.051	0.042
(c) match interval K (λ = 0.5,basic)	K = 0	0.691	16.81	0.900	0.141	0.202	0.342	21.81	0.890	0.081	0.117	16.56	0.779	0.151	0.174	24.41	0.887	0.060	0.044
	K = 1	0.337	19.93	0.937	0.094	0.108	0.312	25.47	0.948	0.053	0.049	16.83	0.781	0.146	0.169	23.15	0.852	0.070	0.056
	K = 10	0.019	33.35	0.995	0.020	0.005	0.302	30.66	0.976	0.029	0.010	17.32	0.785	0.140	0.158	22.48	0.831	0.075	0.065
	K = 50	**0.005**	38.86	0.998	0.012	0.002	**0.301**	30.95	0.978	0.028	0.009	17.14	0.786	0.143	0.163	7.85	0.031	0.405	0.673
(d) optim. strategy $K = 5,\lambda = 0.5$)	rand	1.745	18.84	0.948	0.088	0.253	0.534	26.40	0.934	0.048	0.102	52.35	**1.000**	0.004	**0.001**	21.69	0.896	0.082	0.128
	hybrid	**0.005**	39.23	0.998	0.012	0.002	0.305	**38.48**	**0.993**	**0.012**	**0.003**	**54.20**	**1.000**	0.004	**0.001**	**55.92**	**1.000**	**0.001**	**0.000**
	basic	0.046	29.10	0.987	0.031	0.012	0.304	29.07	0.974	0.035	0.023	17.06	0.782	0.143	0.165	23.34	0.849	0.068	0.060
(e) baseline	nvdf.res.128	1.868	19.09	0.952	0.086	0.247	0.562	27.83	0.950	0.041	0.077	49.91	**1.000**	0.004	**0.001**	21.70	0.898	0.082	0.126
	nvdf.res.1024	0.549	28.77	0.980	0.036	0.067	1.886	19.13	0.968	0.081	0.047	53.25	0.999	**0.002**	**0.001**	23.36	0.950	0.068	0.055
	nvdfmul.res	0.664	21.42	0.927	0.085	0.197	2.159	17.87	0.953	0.112	0.067	31.89	0.940	0.026	0.055	20.26	0.902	0.097	0.114
	nvdf.down.res	2.113	17.03	0.943	0.128	0.084	0.670	20.16	0.918	0.098	0.196	51.36	0.998	0.003	**0.001**	23.34	0.951	0.068	0.054

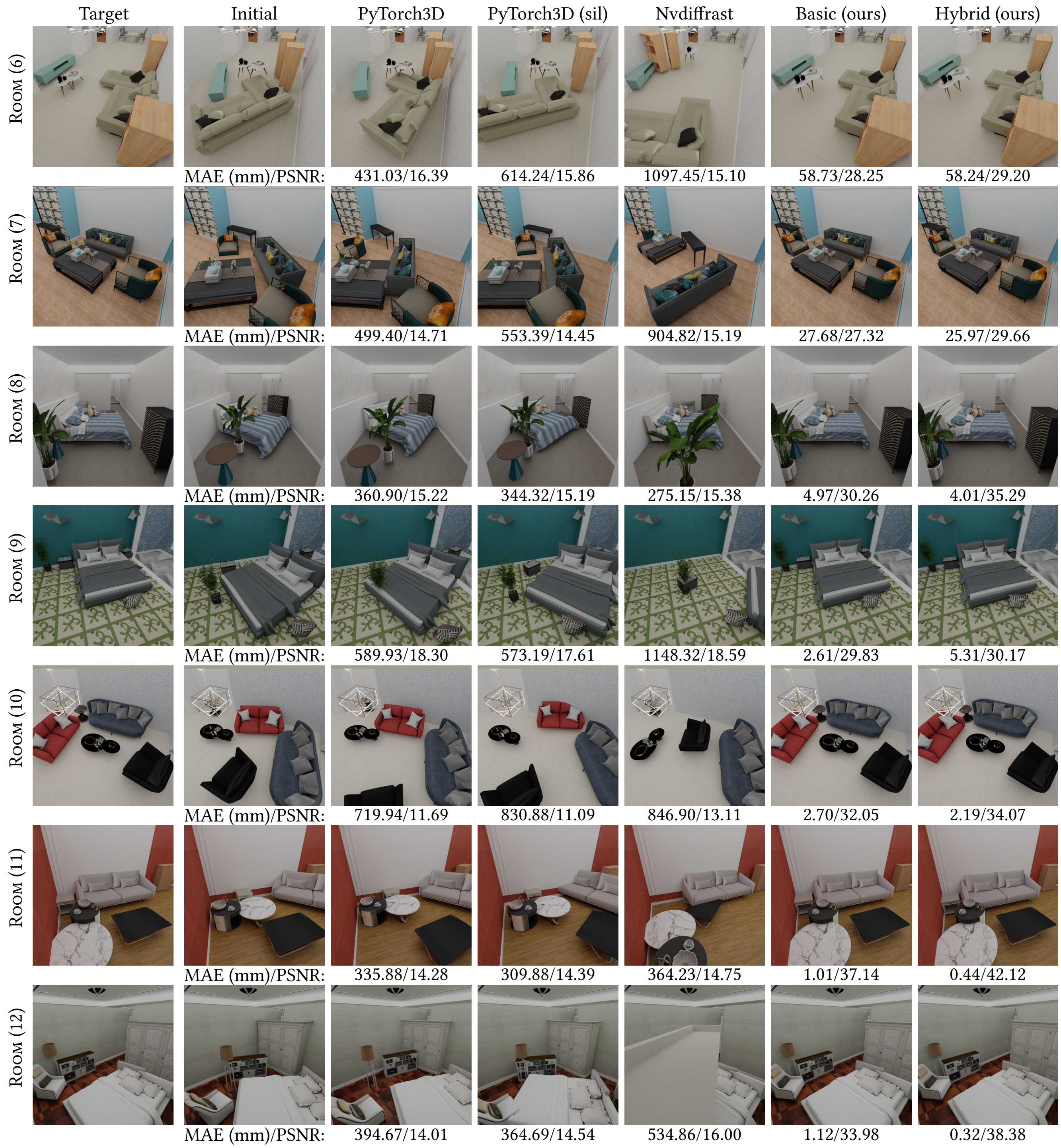

Fig. 2. More examples of the furniture layout adjustment application. Note that optimization is performed using rasterized images under 256 × 256 resolution, while the images shown here are re-rendered in a high resolution with ray tracing for higher display quality.

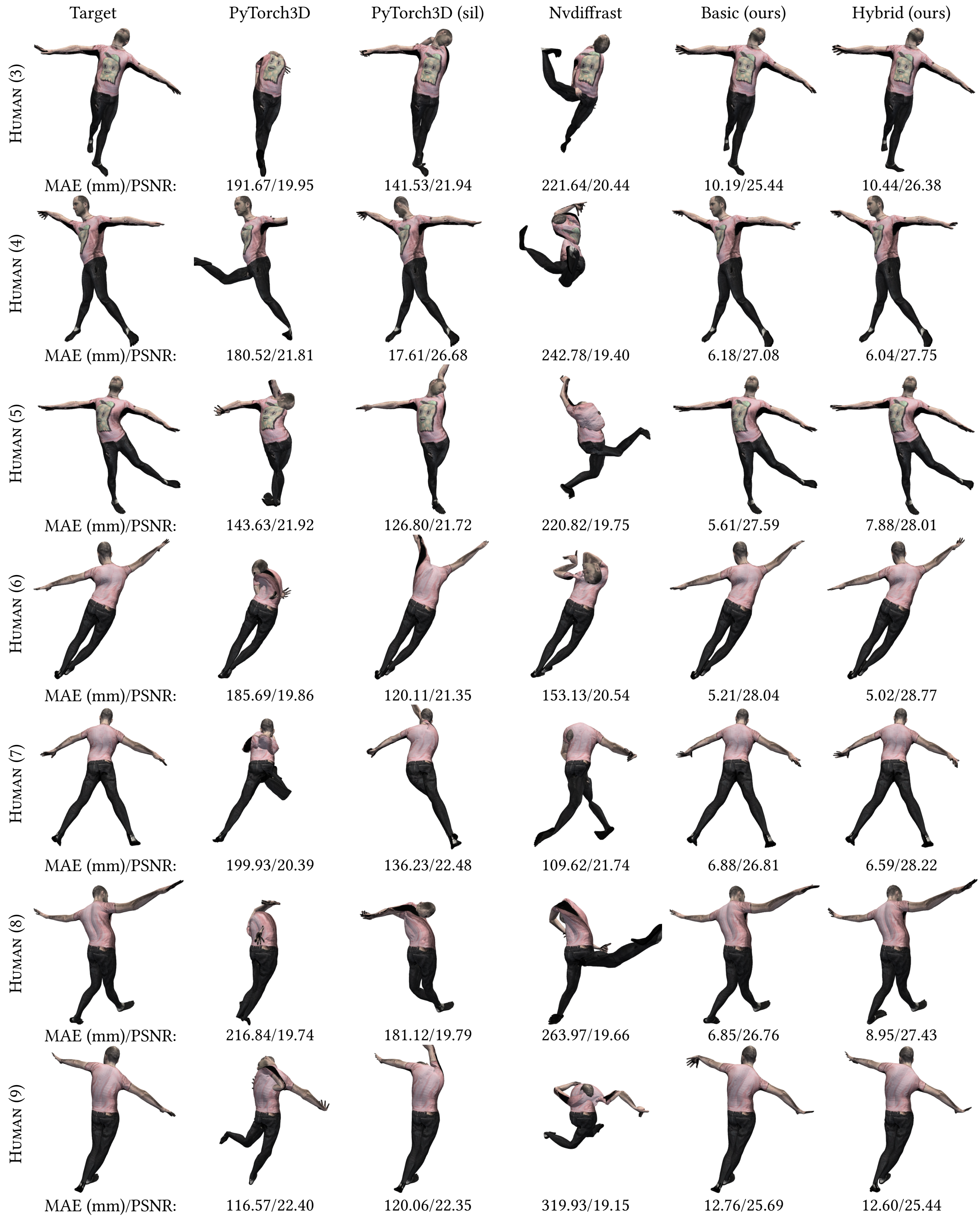

Fig. 3. More examples of the human pose fitting application.

Fig. 4. More examples of the facial expression reconstruction application.