PARTICULATE: Feed-Forward 3D Object Articulation

– Supplementary Material –

6. Additional Experimental Results

6.1. Additional Details of Evaluation Metrics

Motivation for imposing penalties on unmatched parts. Recall that we apply Hungarian matching to obtain the correspondence between the articulated parts predicted by the model and the ground-truth parts. Certain parts are unmatched, as the number of predicted parts is not necessarily equal to the number of GT parts. Unlike previous work [21, 32, 33] that simply ignores the unmatched parts, we impose a penalty (corresponding to the worst possible score for each evaluation metric, *e.g.*, -1 for gIoU and 0 for mIoU) on each unmatched part. To motivate this metric design, in Tab. 5 and Tab. 6, we report the evaluation results of the non-penalizing version used in prior work. Note that the naive baseline, which treats the entire object as a single static part, outperforms *all* baseline methods in non-penalizing gIoU, PC and mIoU for both the rest state and the fully articulated state, even though human judges generally find the non-naive baselines' results (Fig. 4) more reasonable. This is because the single part predicted by the naive baseline is matched to the object's base part, which typically occupies most of the volume and is surrounded by smaller adjacent parts, resulting in favorable metric scores. By penalizing missed small parts, our proposed metric is more consistent with human-judged quality.

Motivation for the whole-object Chamfer distance. In Tab. 3, we reported the *whole-object* Chamfer distance (OC), which measures the bi-directional Chamfer distance between the entire predicted and ground-truth point clouds at the fully articulated state. In contrast to the part-averaged metrics (*i.e.*, gIoU and PC), OC does not rely on matching predicted parts to ground-truth parts and instead provides a more *global* evaluation of both part segmentation and motion parameter estimation. In addition, many parts move farther from the base part at the fully articulated state, thus yielding a poorer (and more appropriate) OC value for the naive baseline.

Method	Lightwheel gIoU ↑	Lightwheel PC ↓	Lightwheel mIoU ↑	PartNet-Mobility gIoU ↑	PartNet-Mobility PC ↓	PartNet-Mobility mIoU ↑
Without Penalty						
Naive Baseline	**0.687**	**0.019**	**0.687**	**0.897**	0.003	**0.897**
SINGAPO [33]	0.184	0.044	0.370	–	–	–
SINGAPO [33] (1@10)	0.250	0.033	0.404	–	–	–
PARTICULATE (ours)	0.486 ± 0.007	0.028 ± 0.001	0.541 ± 0.005	0.882 ± 0.006	**0.002** ± 0.000	0.884 ± 0.005
PartField†	0.092	0.099	0.268	0.370	0.040	0.421
P3SAM†	0.421	0.042	0.523	0.390	0.023	0.403
SINGAPO [33]†	0.278	0.051	0.394	0.538	0.008	0.570
SINGAPO(best) [33] (1@10)†	0.407	0.036	0.476	0.563	0.006	0.585
Articulate AnyMesh [43]†	0.567	0.025	0.608	0.581	0.025	0.621
PARTICULATE (ours)†	**0.726** ± 0.052	**0.010** ± 0.003	**0.744** ± 0.047	**0.883** ± 0.003	**0.001** ± 0.000	**0.885** ± 0.003
With Penalty						
Naive Baseline	0.018	0.285	0.413	0.296	0.210	0.612
SINGAPO [33]	-0.116	0.201	0.272	–	–	–
SINGAPO [33] (1@10)	-0.096	0.190	0.277	–	–	–
PARTICULATE (ours)	**0.183** ± 0.005	**0.163** ± 0.001	0.430 ± 0.003	**0.879** ± 0.006	**0.003** ± 0.000	**0.883** ± 0.005
PartField [35]†	0.079	0.106	0.264	0.183	0.123	0.361
P3SAM [39]†	0.122	0.177	0.411	-0.116	0.261	0.267
SINGAPO [33]†	-0.097	0.234	0.273	0.262	0.124	0.468
SINGAPO [33] (1@10)†	-0.050	0.221	0.297	0.271	0.117	0.471
Articulate AnyMesh [43]†	0.172	0.190	**0.452**	0.383	0.104	0.542
PARTICULATE (ours)†	**0.332** ± 0.034	**0.168** ± 0.002	**0.576** ± 0.035	**0.880** ± 0.003	**0.003** ± 0.001	**0.884** ± 0.003

Table 5. **Part segmentation results at the rest state** with and without *penalty* applied to unmatched parts. The metrics without penalty do not sufficiently penalize missing small parts. Under this protocol, the Naive Baseline outperforms *all* baseline methods on *all* metrics. †: leveraging mesh connectivity. 1@10: reporting best out of 10 predictions. Colors: **best** and **second best**. $\pm\sigma$: standard deviation across 10 predictions.

Method	Lightwheel gIoU ↑	Lightwheel PC ↓	Lightwheel OC ↓	PartNet-Mobility gIoU ↑	PartNet-Mobility PC ↓	PartNet-Mobility OC ↓
Without Penalty						
Naive Baseline	**0.680**	**0.056**	0.018	**0.897**	**0.011**	0.027
SINGAPO [33]	0.178	0.173	0.011	–	–	–
SINGAPO [33] (1@10)	0.243	0.096	0.012	–	–	–
PARTICULATE (ours)	0.464 ± 0.007	0.076 ± 0.003	**0.008** ± 0.000	0.845 ± 0.006	0.023 ± 0.003	**0.003** ± 0.000
SINGAPO [33]†	0.271	0.179	0.018	0.523	0.079	0.046
SINGAPO [33] (1@10)†	0.398	0.091	0.019	0.555	0.066	0.043
Articulate AnyMesh [43]†	0.547	0.084	0.010	0.577	0.187	0.022
PARTICULATE (ours)†	**0.692** ± 0.047	**0.062** ± 0.004	**0.009** ± 0.002	**0.846** ± 0.003	**0.021** ± 0.005	**0.003** ± 0.000
With Penalty						
Naive Baseline	0.016	0.293	0.017	0.296	0.216	0.027
SINGAPO [33]	-0.121	0.299	0.011	–	–	–
SINGAPO [33] (1@10)	-0.100	0.238	0.012	–	–	–
PARTICULATE (ours)	**0.165** ± 0.005	**0.200** ± 0.002	**0.008** ± 0.000	**0.842** ± 0.005	**0.024** ± 0.003	**0.003** ± 0.000
SINGAPO [33]†	-0.102	0.329	0.018	0.255	0.184	0.046
SINGAPO [33] (1@10)†	-0.056	0.261	0.019	0.264	0.168	0.041
Articulate AnyMesh [43]†	0.158	0.237	0.010	0.378	0.251	0.022
PARTICULATE (ours)†	**0.305** ± 0.030	**0.208** ± 0.003	**0.009** ± 0.002	**0.843** ± 0.003	**0.022** ± 0.005	**0.003** ± 0.000

Table 6. **Results evaluated using fully articulated geometries** with and without *penalty* applied to unmatched parts. The metrics without penalty do not sufficiently penalize missing small parts. Under this protocol, the Naive Baseline outperforms *all* baseline methods on *all* part-wise metrics (*i.e.*, gIoU and PC). †: leveraging mesh connectivity. 1@10: reporting best out of 10 predictions. Colors: **best** and **second best**. $\pm\sigma$: standard deviation across 10 predictions.

6.2. Per-Category Evaluation Results

Tab. 7 provides the category distribution of our introduced Lightwheel benchmark. Tab. 8 presents the evaluation results on all 14 individual categories. While PARTICULATE outperforms baseline methods on almost all categories, the performance drops significantly on out-of-distribution categories such as stand mixer and stovetop, indicating our approach is still bottlenecked by the limited diversity in the training data.

Articulate AnyMesh [43] performs better specifically on the stand mixer category. This is because stand mixer never appears in its training set and its articulation pattern is distinct from all other training objects. With its strong VLM-

Category	**# Objects**
Blender	13
Coffee machine	25
Dishwasher	13
Electric kettle	13
Microwave	25
Oven	12
Range hood	12
Refrigerator	25
Sink	25
Stand mixer	12
Stove	10
Stovetop	8
Toaster	25
Toaster oven	25
Total	**243**

Table 7. **Category statistics** of our introduced Lightwheel benchmark. Categories highlighted in green indicate out-of-distribution categories absent from PartNet-Mobility and GRScenes.

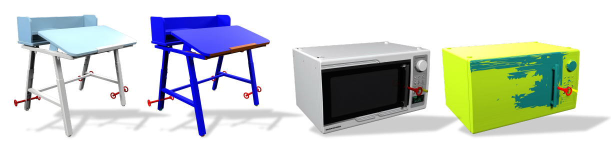

Figure 5. **Failure cases.** Our model can struggle with objects with atypical articulation configurations (**left**) or with AI-generated shapes which lack internal structures (**right**).

based generalization, Articulate AnyMesh handles this unseen category particularly well. For the range hood category, objects often contain a very large base component that occupies the majority of the object's volume. SINGAPO [33] tends to predict oversized parts, especially when the input object falls outside its training distribution. Because the ground-truth base part is already extremely large, SINGAPO's tendency to over-predict part extents accidentally aligns with the dominant fixed region, leading to inflated IoU values for this specific category.

6.3. Failure Cases

We show typical failure cases of PARTICULATE in Fig. 5. Notably, while our model is trained on a large variety of articulated objects, the training dataset remains several orders of magnitude smaller than those used by open-domain 3D generators (*i.e.*, Objaverse [8, 9]). As a result, PARTICULATE can struggle with objects whose articulation configurations deviate significantly from those represented in our training data (**left**). Moreover, many AI-generated 3D assets lack realistic internal structures (**right**), introducing a distribution shift (since all objects in our training set contain well-defined internal parts) that can cause our model to fail. Enhancing the model's robustness to such noisy or unclean inputs is a valuable direction for future work.

Methods	Metrics	Categories													
		Micro-wave	Dish-washer	Sink	Electric Kettle	Refri-gerator	Blender	Oven	Toaster	Stand Mixer	Toaster Oven	Coffee Machine	Stove	Stovetop	Range Hood
P3SAM†	mIoU (rest)	0.492	0.411	0.482	**0.637**	**0.498**	**0.514**	0.364	0.359	**0.569**	0.395	**0.393**	0.253	**0.377**	0.097
	gIoU (rest)	0.252	0.156	0.162	**0.515**	0.229	**0.282**	0.092	0.046	**0.365**	**0.135**	**0.119**	-0.201	**-0.034**	-0.223
	PC (rest)	**0.103**	0.144	0.228	**0.110**	0.162	0.116	**0.170**	**0.187**	**0.146**	**0.152**	**0.156**	0.212	0.183	**0.233**
PartField†	mIoU (rest)	0.408	0.273	0.406	0.320	0.235	0.330	0.202	0.192	**0.445**	0.187	0.205	0.170	0.222	0.056
	gIoU (rest)	0.254	0.087	0.309	0.218	0.068	0.141	-0.029	-0.026	**0.367**	-0.056	**0.048**	-0.166	**-0.023**	-0.204
	PC (rest)	0.103	**0.104**	0.062	**0.090**	**0.070**	**0.072**	**0.129**	**0.146**	**0.063**	**0.125**	**0.083**	0.154	**0.134**	**0.144**
SINGAPO	mIoU (rest)	0.337	0.346	0.423	0.304	0.168	0.050	0.220	0.248	0.071	0.193	0.301	0.225	0.176	**0.443**
	gIoU (rest)	0.003	-0.080	0.135	-0.286	-0.126	-0.328	-0.162	-0.172	-0.335	-0.180	-0.144	-0.190	-0.234	**0.000**
	gIoU (high)	0.001	-0.084	0.135	-0.328	-0.129	-0.331	-0.166	-0.176	-0.337	-0.182	-0.146	-0.196	-0.241	**-0.001**
	PC (rest)	0.138	0.196	0.197	0.313	0.126	0.236	0.187	0.208	0.326	0.192	0.273	0.163	**0.109**	0.259
	PC (high)	0.362	0.401	0.220	0.593	0.176	0.318	0.299	0.281	0.573	0.293	0.342	0.205	**0.134**	0.284
	OC (high)	0.016	**0.013**	0.002	0.059	0.007	0.026	0.008	0.007	0.040	0.009	0.013	0.004	0.001	0.003
SINGAPO†	mIoU (rest)	0.392	0.331	0.056	0.490	0.209	0.102	0.251	0.283	0.067	0.229	0.376	0.279	0.254	**0.515**
	gIoU (rest)	0.080	-0.075	-0.262	0.072	-0.040	-0.271	-0.139	-0.114	-0.349	-0.160	-0.057	-0.132	-0.114	**0.089**
	gIoU (high)	0.079	-0.077	-0.262	0.008	-0.047	-0.271	-0.145	-0.117	-0.349	-0.164	-0.057	-0.136	-0.117	**0.089**
	PC (rest)	0.171	0.225	0.278	0.288	0.124	0.261	0.240	0.241	0.355	0.247	0.289	0.225	0.185	0.265
	PC (high)	0.435	0.425	0.282	0.338	0.182	0.341	0.366	0.326	0.625	0.364	0.329	0.279	0.202	0.283
	OC (high)	0.019	0.013	0.059	0.054	0.007	0.025	0.009	0.007	**0.040**	0.011	0.013	0.005	0.001	0.002
Articulate AnyMesh [43]†	mIoU (rest)	0.671	0.583	**0.493**	0.408	0.362	0.503	0.432	**0.441**	0.425	0.387	0.340	0.369	0.289	0.236
	gIoU (rest)	0.434	0.283	**0.402**	0.092	0.081	0.275	0.122	**0.130**	0.210	0.063	0.000	-0.033	-0.129	-0.104
	gIoU (high)	0.405	0.265	**0.394**	0.053	0.077	**0.247**	0.107	**0.120**	**0.178**	0.057	**-0.005**	-0.040	-0.129	-0.104
	PC (rest)	0.147	0.218	**0.057**	0.218	0.157	0.133	0.235	0.222	0.155	0.236	0.210	0.281	0.273	0.280
	PC (high)	0.216	0.235	**0.068**	0.472	0.215	0.182	0.279	0.258	**0.451**	0.297	**0.227**	0.292	0.289	**0.279**
	OC (high)	0.011	**0.003**	**0.001**	0.064	0.011	**0.021**	0.011	0.009	0.040	0.016	0.014	**0.002**	**0.000**	**0.000**
PARTICULATE (ours)	mIoU (rest)	**0.678**	**0.715**	0.369	0.503	0.488	0.404	**0.491**	0.393	0.267	**0.444**	0.256	**0.392**	0.222	0.208
	gIoU (rest)	**0.472**	**0.593**	0.259	0.218	**0.287**	0.196	**0.199**	0.109	0.210	0.102	-0.144	**0.196**	-0.041	-0.189
	gIoU (high)	**0.439**	**0.554**	0.257	**0.167**	**0.265**	0.173	**0.180**	0.098	**-0.083**	**0.086**	-0.147	**0.186**	**-0.042**	-0.191
	PC (rest)	0.130	**0.088**	0.092	0.208	0.120	**0.111**	0.212	0.188	0.251	0.242	0.224	**0.106**	0.166	0.328
	PC (high)	**0.123**	**0.172**	**0.140**	**0.274**	**0.123**	**0.165**	**0.219**	**0.197**	0.509	**0.244**	**0.267**	**0.110**	**0.179**	0.338
	OC (high)	**0.001**	0.024	**0.002**	**0.018**	**0.001**	**0.019**	**0.005**	**0.001**	0.043	**0.002**	**0.012**	0.002	0.003	**0.002**
PARTICULATE (ours)†	mIoU (rest)	**0.802**	**0.811**	**0.508**	**0.674**	**0.633**	**0.708**	**0.521**	**0.509**	0.284	**0.490**	**0.435**	**0.543**	**0.373**	0.303
	gIoU (rest)	**0.632**	**0.689**	**0.425**	**0.390**	**0.498**	**0.501**	**0.203**	**0.199**	-0.071	**0.134**	0.038	**0.292**	**-0.007**	-0.088
	gIoU (high)	**0.580**	**0.640**	**0.423**	**0.330**	**0.452**	**0.456**	**0.179**	**0.181**	-0.083	**0.110**	**0.033**	**0.274**	**-0.012**	-0.088
	PC (rest)	**0.115**	**0.088**	**0.047**	0.207	**0.072**	0.121	0.234	0.214	0.281	0.252	0.278	**0.141**	0.195	0.273
	PC (high)	**0.116**	**0.177**	0.158	**0.276**	**0.075**	**0.171**	**0.242**	**0.220**	**0.482**	**0.254**	0.321	**0.147**	0.203	**0.276**
	OC (high)	**0.000**	0.027	0.003	**0.018**	**0.001**	**0.019**	**0.005**	**0.001**	**0.038**	**0.002**	**0.012**	**0.001**	**0.001**	0.002

Table 8. **Per-category evaluation results** against baselines on the Lightwheel dataset. We report the results both at the rest state (**rest**) and at the fully articulated state (**high**). The categories are ranked based on gIoU (rest) averaged over all methods (left is higher). †: leveraging mesh connectivity. Categories highlighted in green indicate out-of-distribution categories absent from the training datasets PartNet-Mobility and GRScenes. Colors: **best** and **second best**.