

## A Statistics on the used datasets

In Tables 4 & 5, we summarize statistical information on the number of instances and categories considered in our evaluation. As we require parts annotations as an important ingredient in our deformation, we only select instances in Scan2CAD [5] where the associated parts annotation in PartNet [31] is available, resulting in total in 9 categories (25%), 572 instances (18%), and 1979 annotated correspondences (14%). Note that the vast majority of cases remain within our consideration, keeping our evaluation comprehensive.

Collection	Categories	Instances	Corresp.
Scan2CAD [5]	35	3,049	14,225
w/parts annotations	24	2,477	12,246

**Table 4:** Overall statistics on the numbers of categories, instances, and correspondences present in our datasets.

We further select the most well-presented six shape categories as our core evaluation set, outlined in Table 5. Note that as our method is non-learnable, we can just as easily experiment with the remaining categories, at the cost of somewhat reduced statistical power.

## B Optimization details

Our full experimental pipeline is a sequence of deformation stages with different optimization parameters, and Hessian being recomputed before each stage. Specifically, we perform one *part-to-part* optimization with parameters  $\alpha_{\text{shape}} = 1, \alpha_{\text{smooth}} = 0, \alpha_{\text{sharp}} = 0, \alpha_{\text{data}} = 5 \times 10^4$  for 100 iterations, then we perform 5 runs of *nearest-neighbor* deformation for 50 iterations with parameters  $\alpha_{\text{shape}} = 1, \alpha_{\text{smooth}} = 10, \alpha_{\text{sharp}} = 10, \alpha_{\text{data}} = 10^3$ . Such number of iterations was sufficient to achieve convergence with energy changes less than  $10^{-1}$  in our experiments. Runtime of our method breaks into cost computation ( $\sim 0.3$  s), backward ( $\sim 0.2$  s), and optimization steps containing the main bottleneck (sparse matrix-vector multiplication) ( $\sim 1.2$  s) for a typical  $10^4$  vertices mesh. All operations can be easily further optimized.

## C Qualitative fitting results

In Figure 6, we display a series of qualitative results with a variety of shape deformations with different classes of instances. Comparing to baselines, our framework achieves accurate fit while preserving sufficient perceptual quality.

Name	Scan2CAD		PartNet $\cap$ Scan2CAD	
	corresp.	shapes	corresp.	shapes
<i>Shape categories used in our evaluation:</i>				
chair	4677	652	4351	632
table	2616	830	2594	822
cabinet	1401	310	1258	294
trash bin	1044	89	1042	88
bookshelf	824	150	812	145
display	770	165	762	161
<i>Shape categories NOT used in our evaluation:</i>				
bed	355	50	342	47
file cabinet	294	70	290	68
bag	165	9	165	9
lamp	135	55	135	55
bathtub	474	96	129	25
microwave	99	37	98	36
sofa	577	247	60	20
laptop	51	24	51	24
keyboard	62	11	48	9

**Table 5:** The top 15 most frequent ShapeNet categories in Scan2CAD dataset including a detailed information on those with the availability of the corresponding parts annotations.

Method	bookshelf	cabinet	chair	display	table	trash bin	other	class avg.	avg.
ARAP	52.48	41.77	45.52	51.30	41.77	57.00	39.75	47.08	45.67
Harmonic	64.77	58.74	68.06	64.22	58.26	80.13	61.70	65.12	65.18
Ours: NN only	<b>21.54</b>	<b>23.39</b>	<b>7.31</b>	<b>18.37</b>	<b>18.69</b>	18.13	<b>16.07</b>	<b>17.64</b>	<b>14.14</b>
Ours: p2p only	22.44	24.28	9.51	21.12	18.76	<b>15.30</b>	18.34	18.54	15.57
Ours: w/o smooth	27.15	29.27	14.50	27.05	24.48	24.39	23.26	24.30	20.95
Ours: w/o sharp	26.43	25.98	13.34	24.87	22.47	21.04	21.18	22.19	19.10
CAD-Deform	24.8	24.1	11.4	24.4	21.6	19.4	17.6	20.5	17.2

**Table 6:** Quantitative results of local surface quality evaluation using DAME measure [41] (the smaller, the better, normalized to a maximum score of 100), where our CAD-Deform compares favourably to the baselines across all considered categories. Note, however, how surface quality significantly decreases when smoothness and sharp feature-related terms are dropped.

Table 6 reports the results of surface quality evaluation using deformations obtained with our CAD-Deform vs. the baselines, category-wise. While outperforming the baseline methods across all categories, we discover the smoothness

$\text{EMD} \times 10^{-3}$	bookshelf	cabinet	chair	display	table	trash bin	class avg.	avg.
Ground-truth	77.8	78.9	76.1	77.5	77.3	73.1	76.8	77.0
ARAP [38]	80.3	86.5	88.5	<b>85.4</b>	<b>86.8</b>	98.1	87.6	87.3
Harmonic [7,26]	94.0	110.6	95.8	95.3	103.2	122.6	103.6	101.7
CAD-Deform	<b>79.0</b>	<b>81.1</b>	<b>80.3</b>	91.7	87.0	<b>87.4</b>	<b>84.4</b>	<b>83.8</b>

**Table 7:** Results of LSLP-GAN reconstruction in terms of Earth-Mover’s Distance between reconstructed and original point clouds of mesh vertices.

and sharpness energy terms to be the crucial ingredients in keeping high-quality meshes.

Figure 7 displays visually the deformation results using the three distinct classes, highlighting differences in surfaces obtained using the three methods.

Table 7 reports shape abnormality evaluation results across the six considered categories. Baselines show (Fig. 8) low reconstruction quality as evidenced by a larger number of black points. In other words, comparing to CAD-Deform, the distance from these meshes to undeformed ones is much larger.

In Figure 9, we show a series of examples for CAD-Deform ablation study. Perceptual quality degrades when excluding every term from the energy.

## D Morphing

In this section, we present an additional series of examples of morphing properties (Fig. 10). Every iteration of optimization process gradually increases the quality of fit. With CAD-Deform we can morph each part to imitate the structure of the target shape.

## E PartNet annotation

<b>Accuracy, %</b>	class avg.	avg.
Ground-truth	89.22	90.56
Level 1 (object)	89.25	90.79
Level 2	89.16	91.21
Level 3	89.40	91.05
Level 4 (parts)	<b>91.65</b>	<b>93.12</b>

**Table 8:** Comparative evaluation of our approach in terms of Accuracy on different levels of detail.

This set of experiments shows how quality of fitting depends on mesh vertices labelling. We can provide labels for mesh in different ways depending on the level

in PartNet hierarchy [31]. We observe the increase of fitting quality with greater level of detail (Table 8). Examples presented in Figure 11 are selected as the most distinguishable deformations on different levels. There are minor visual differences in deformation performance of part labeling level.

## F Fitting Accuracy analysis

CAD-Deform deformation framework is sensitive to Accuracy threshold  $\tau$  for the distance between mesh vertices and close scan points. In Figure 12 variation of  $\tau$  threshold is presented and we selected  $\tau = 0.2$  m for fitting Accuracy metric.

## G Perceptual assessment and user study details

Having obtained a collection of deformed meshes, we aim to assess their visual quality in comparison to two baseline deformation methods: as-rigid-as-possible (ARAP) [38] and Harmonic deformation [7,26], using a set of perceptual quality measures. The details of our user study design and visual assessment are provided in the supplementary. To this end, we use original and deformed meshes to compute DAME and reconstruction errors, as outlined in Section 6.1, and complement these with visual quality scores obtained with a user study (see below). These scores, presented in Table 3, demonstrate that shapes obtained using CAD-Deform have  $2\times$  higher surface quality, only slightly deviate from undeformed shapes as viewed by neural autoencoders, and receive  $2\times$  higher ratings in human assessment, while sacrificing only 1.1–4.5 % accuracy compared to other deformation methods.

*Design of our user study.* The users were requested to examine renders of shapes from four different categories: the original undeformed shapes as well as shapes deformed using ARAP, Harmonic, and CAD-Deform methods, and give a score to each shape according to the following perceptual aspects: surface quality and smoothness, mesh symmetry, visual similarity to real-world objects, and overall consistency. Ten random shapes from each of the four categories have been rendered from eight different views and scored by 100 unique users on a scale from 1 (bad) to 10 (good). The resulting visual quality scores are computed by averaging over users and shapes in each category.

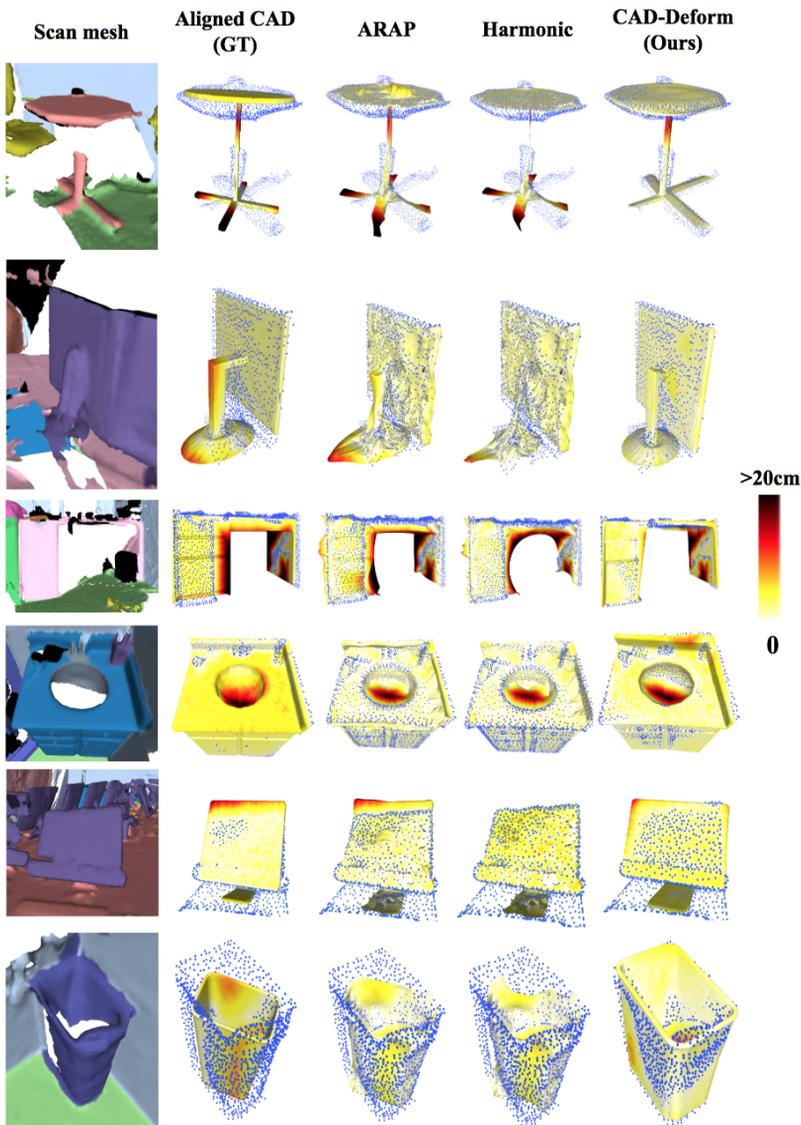
In Figure 13, we present a distribution of user scores over different deformation methods and shapes. It can be clearly seen that users prefer our deformation results to baselines for all of the cases, which is obvious from the gap between histogram of CAD-Deform and ARAP/Harmonic histograms. At the same time, shapes deformed by CAD-Deform are close to undeformed ShapeNet shapes in terms of surface quality and smoothness, mesh symmetry, visual similarity to real-world objects, and overall consistency. Besides, in Tables 9, 10 we provide numbers for evaluation of ARAP/Harmonic deformations w.r.t. the change of Laplacian term weight.

Lap. term weight	Class avg.			Instance avg.		
	GT	S2C [5]	E2E [6]	GT	S2C [5]	E2E [6]
$\alpha_{\text{Lap}} = 10^{-2}$	90.9	81.3	90.0	92.0	80.9	90.8
$\alpha_{\text{Lap}} = 10^{-1}$	91.0	81.3	90.0	92.0	80.9	90.7
$\alpha_{\text{Lap}} = 1$	91.0	81.3	89.9	91.9	80.9	90.7
$\alpha_{\text{Lap}} = 5$	90.9	81.2	89.9	91.9	80.9	90.7
$\alpha_{\text{Lap}} = 20$	90.9	81.2	89.9	91.8	80.8	90.6

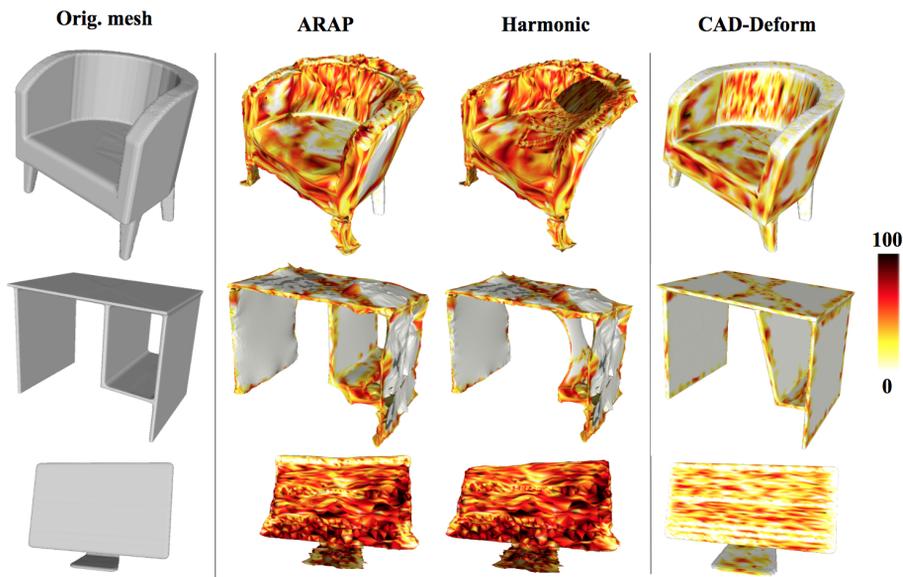
**Table 9:** Comparative evaluation of ARAP deformations w.r.t. the change of Laplacian term weight in terms of Accuracy (%).

Lap. term weight	Class avg.			Instance avg.		
	GT	S2C [5]	E2E [6]	GT	S2C [5]	E2E [6]
$\alpha_{\text{Lap}} = 10^{-2}$	96.3	94.3	96.6	96.7	94.5	96.9
$\alpha_{\text{Lap}} = 10^{-1}$	96.3	94.2	96.6	96.7	94.4	96.9
$\alpha_{\text{Lap}} = 1$	96.3	94.2	96.6	96.7	94.2	96.9
$\alpha_{\text{Lap}} = 5$	96.2	94.0	96.6	96.6	94.0	96.8
$\alpha_{\text{Lap}} = 20$	96.2	93.8	96.5	96.6	93.8	96.7

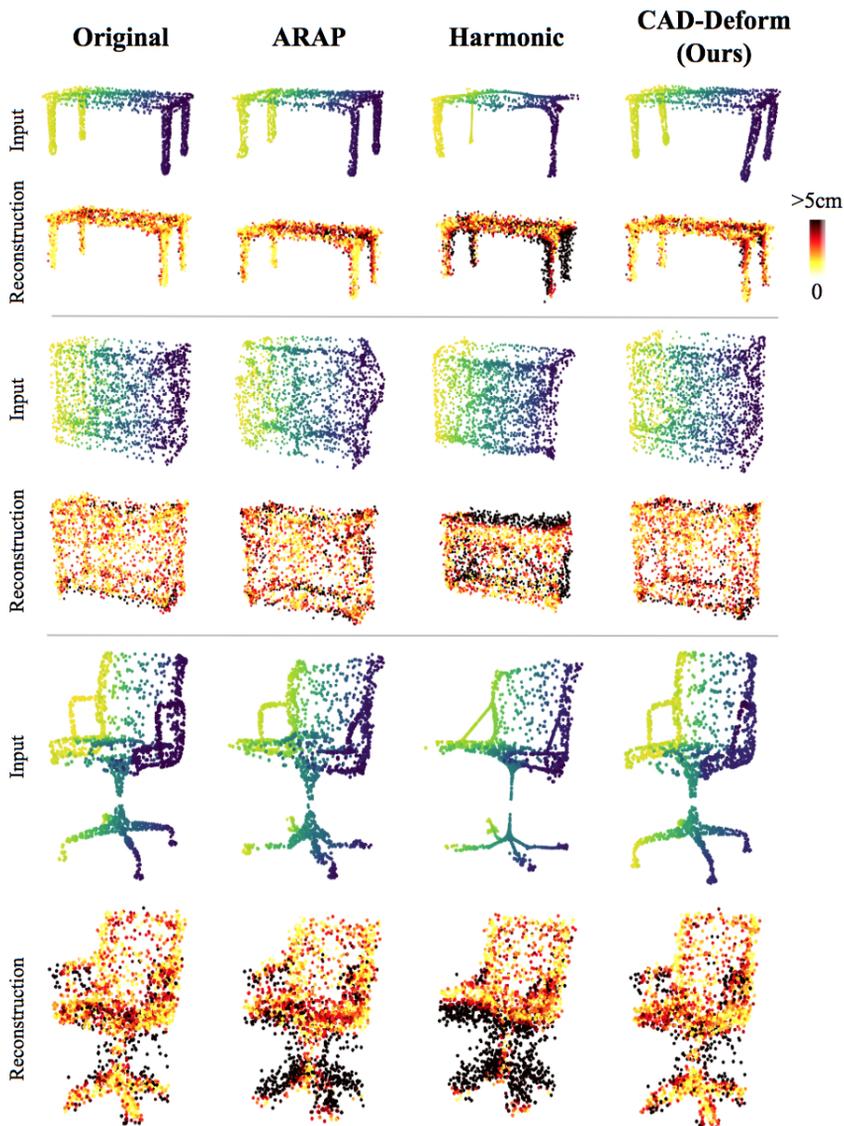
**Table 10:** Comparative evaluation of Harmonic deformations w.r.t. the change of Laplacian term weight in terms of Accuracy (%).



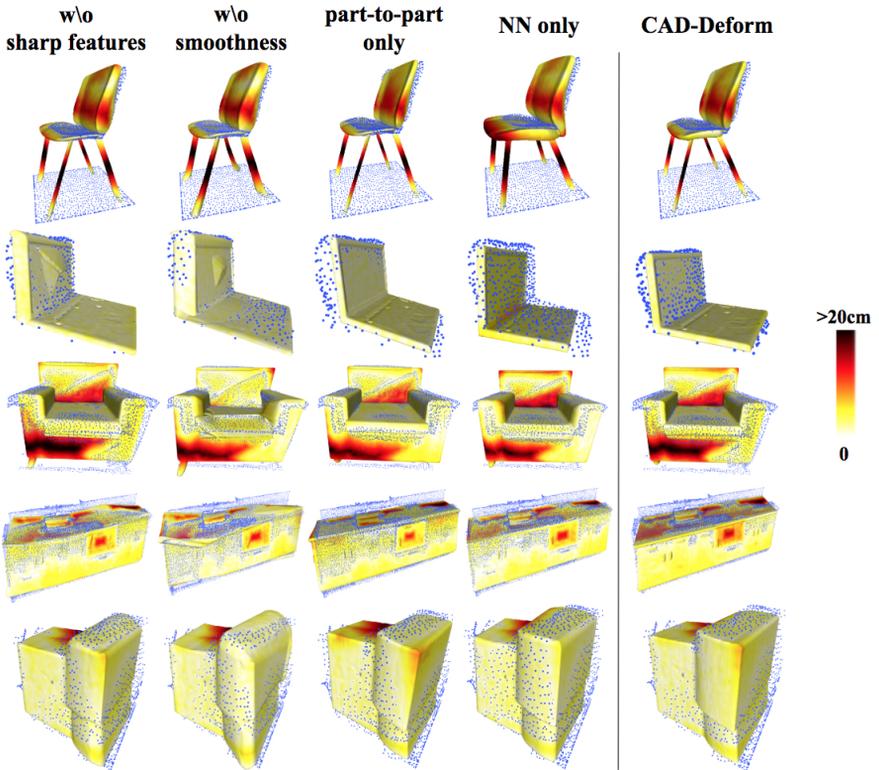
**Fig. 6:** Qualitative shape deformation results using obtained using ARAP [38], Harmonic deformation [7,26], and our CAD-Deform. Mesh surface is colored according to the value of tMMD measure, with darker values corresponding to the larger distance values.



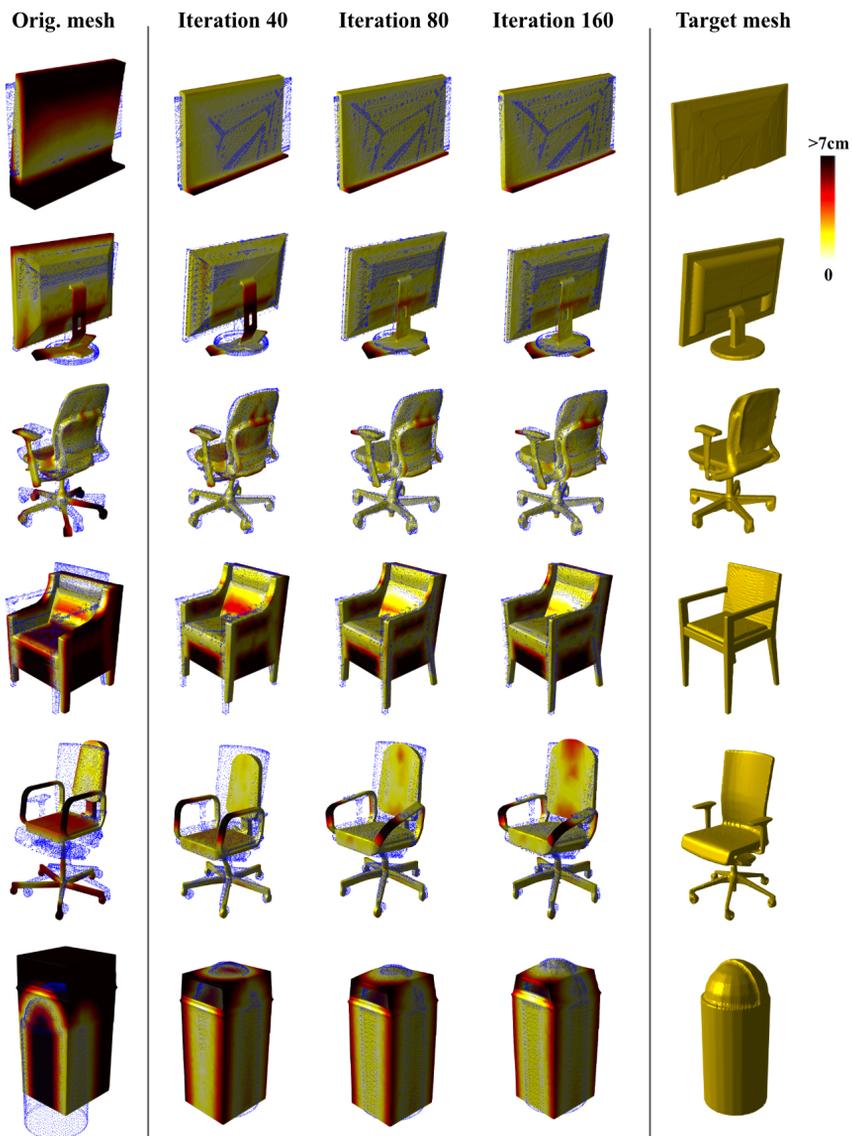
**Fig. 7:** Qualitative comparison of deformations obtained using ARAP [38], Harmonic deformation [7,26], and our CAD-Deform, with shapes coloured according to the value of DAME measure [41]. Our approach results in drastic improvements in local surface quality, producing higher-quality surfaces compared to other deformations.



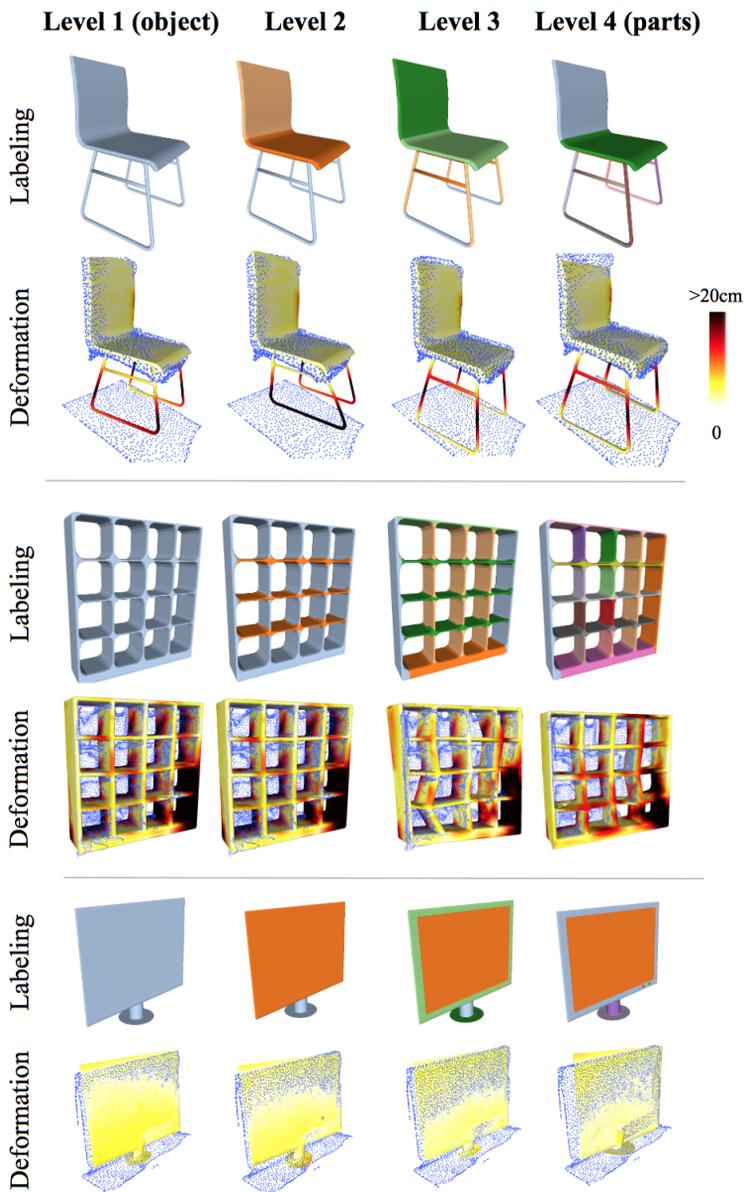
**Fig. 8:** Qualitative comparison of reconstruction of point clouds extracted from mesh vertices. These meshes are obtained using ARAP [38], Harmonic deformation [7,26], and our CAD-Deform, the first column corresponds to original undeformed meshes. The color of reconstructed point clouds is related to Earth-Mover's Distance between reconstructed and original point clouds of mesh vertices.



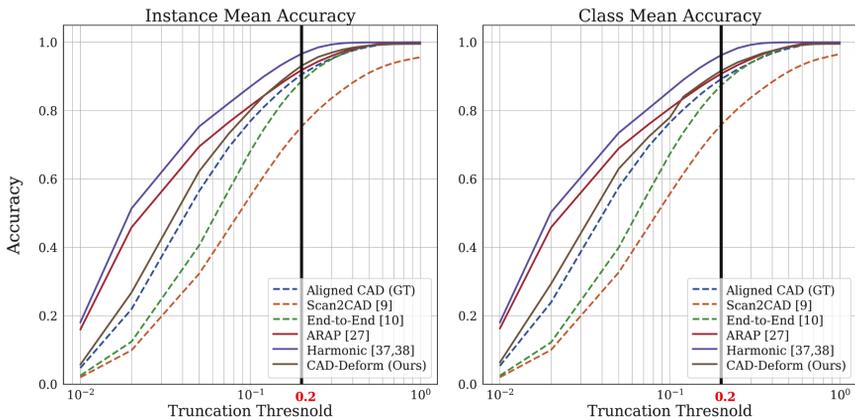
**Fig. 9:** Qualitative results of ablation study using our deformation framework, with mesh coloured according to the value of the tMMD measure.



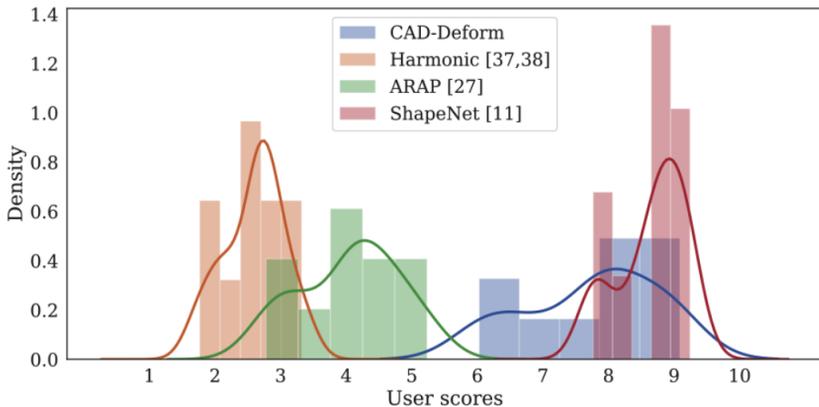
**Fig. 10:** Qualitative shape translation results, interpolating between the original mesh (left) and the target mesh (right).



**Fig. 11:** Deformation performance depending on different level of labelling from the PartNet dataset [31]. Deformed mesh surfaces are colored according to the value of tMMD measure, with darker values corresponding to the larger distance values.



**Fig. 12:** Fitting Accuracy vs. varying  $\tau$  threshold for the distance between mesh vertices and close scan points.



**Fig. 13:** Distribution of user scores averaged by ten shapes from original ShapeNet [9], meshes deformed with ARAP [38], Harmonic [7,26] and CAD-Deform.