From RGB-D Scans to 3D CAD Models

TUM Visual Computing Group
Prof. Matthias Nießner
Democratizing Content Creation?

3D Reconstructions
Democratizing Content Creation?

3D Reconstructions

Computer Graphics Renderings
3D Reconstructions
Incomplete Scan Geometry

TOG’17 [Dai et al.]: BundleFusion
Completing 3D Shapes

CVPR’17 (spotlight) [Dai et al.]: CNNComplete
Data-driven Shape Completion

CVPR’17 (spotlight) [Dai et al.]: CNNComplete
Shape Completion Results

Results on ShapeNet [Chang et al. 15]

CVPR’17 (spotlight) [Dai et al.]: CNNComplete
What about Entire Scenes?
ScanComplete: Scene Completion

Input Partial Scan

Completed Scan

CVPR’18 [Dai et al.]: ScanComplete
Drawback: SDF + MC -> Oversmoothing

CVPR’18 [Dai et al.]: ScanComplete
CAD-to-Scan Retrieval + Alignment

EG’15 [Li et al.]: DB-assisted Object Retrieval
CAD-to-Scan Retrieval + Alignment

EG’15 [Li et al.]: DB-assisted Object Retrieval
Problem Statement

Semantically same, geometrically different!

Scan (chair)

CAD (chair)
Scan2CAD: Learning CAD Model Alignment in RGB-D Scans
Scan2CAD: Dataset
Scan2CAD Dataset

Annotation Pipeline
Scan2CAD: Dataset

97607 Keypoint Pairs
14225 Aligned CADs
3049 Unique CADs
1506 Scenes

CVPR’19 (Oral) [Avetisyan et al.]: Scan2CAD
Scan2CAD: Alignment Method

Variational 9DoF Optimization

CVPR’19 (Oral) [Avetisyan et al.]: Scan2CAD
Scan2CAD: Alignment Method

For every CAD Model
    For every keypoint in scene
        -> predict heat map

For every CAD Model
    For every set of heat maps
        -> run 9 DoF pose optimization (incl. outlier detect)
## Scan2CAD: Results

<table>
<thead>
<tr>
<th>Method</th>
<th>bath</th>
<th>bookshelf</th>
<th>cabinet</th>
<th>chair</th>
<th>display</th>
<th>sofa</th>
<th>table</th>
<th>trash bin</th>
<th>other</th>
<th>class avg.</th>
<th>avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPFH (Rusu et al.)</td>
<td>0</td>
<td>1.92</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>5.41</td>
<td>2.04</td>
<td>1.75</td>
<td>2</td>
<td>2.57</td>
<td>4.45</td>
</tr>
<tr>
<td>SHOT (Tombari et al.)</td>
<td>0</td>
<td>1.43</td>
<td>1.16</td>
<td>7.08</td>
<td>0.59</td>
<td>3.57</td>
<td>1.47</td>
<td>0.44</td>
<td>0.75</td>
<td>1.83</td>
<td>3.14</td>
</tr>
<tr>
<td>Li et al.</td>
<td>0.85</td>
<td>0.95</td>
<td>1.17</td>
<td>14.08</td>
<td>0.59</td>
<td>6.25</td>
<td>2.95</td>
<td>1.32</td>
<td>1.5</td>
<td>3.3</td>
<td>6.03</td>
</tr>
<tr>
<td>3DMatch (Zeng et al.)</td>
<td>0</td>
<td>5.67</td>
<td>2.86</td>
<td>21.25</td>
<td>2.41</td>
<td>10.91</td>
<td>6.98</td>
<td>3.62</td>
<td>4.65</td>
<td>6.48</td>
<td>10.29</td>
</tr>
<tr>
<td><strong>Ours (best)</strong></td>
<td>36.2</td>
<td>36.4</td>
<td>34</td>
<td>44.26</td>
<td>17.89</td>
<td>70.63</td>
<td>30.66</td>
<td>30.11</td>
<td>20.6</td>
<td>35.64</td>
<td>31.68</td>
</tr>
</tbody>
</table>

CVPR’19 (Oral) [Avetisyan et al.]: Scan2CAD
Scan2CAD: Results

CVPR’19 (Oral) [Avetisyan et al.]: Scan2CAD
Limitations with Scan2CAD

• Run-time: ~10min/scene
• Main reason: Retrieval not efficient
• Try out 400 random CAD models to generate this

3D Scan

Ours
End-to-End Alignment: Method
End-to-End Alignment: Method

Input

CAD Model Pool

3D Scan

Output

End2End 3D CNN

Object Detection

Symmetry-aware Object Correspondences

NN-Lookup CADs

Diff’ Alignment Loss

$\mathcal{L}_{Procrustes}$

ICCV’19 [Avetisyan et al.]: End-to-End Alignment
End-to-End Alignment: Method

ICCV’19 [Avetisyan et al.]: End-to-End Alignment
Symmetry-Aware Object Coordinates (SOCs)

• Dense correspondences
• Map every scan voxel into the unit cube [0,1]^3

ICCV’19 [Avetisyan et al.]: End-to-End Alignment
Network Architecture

ICCV'19 [Avetisyan et al.]: End-to-End Alignment
Symmetries

Prediction degradation through unresolved symmetries

Prediction  GT  Prediction  GT

ICCV'19 [Avetisyan et al.]: End-to-End Alignment
Alignment

9DoF

9DoF
Alignment via Procrustes

$$\min_{R \in SO_3} \| R \cdot x - y \|_2^2$$
End-to-End Alignment

Input Scan

- Anchor Centers + Object Detection + Bbox/Scale regression
- CAD Retrieval for each box
- SOC Prediction for each box
- Differentiable Procrustes
End-to-End Alignment Results

3D Scan

3DMatch

Scan2CAD

Ours

Ground Truth

ICCV’19 [Avetisyan et al.]: End-to-End Alignment
End-to-End Alignment Results

3D Scan | 3DMatch | Scan2CAD | Ours | Ground Truth

ICC’19 [Avetisyan et al.]: End-to-End Alignment
## Ablation Study

<table>
<thead>
<tr>
<th>Variation</th>
<th>Accuracy in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct 9DoF</td>
<td>15.12</td>
</tr>
<tr>
<td>Ours (no SOCs)</td>
<td>29.97</td>
</tr>
<tr>
<td>Ours (no symmetry)</td>
<td>40.51</td>
</tr>
<tr>
<td>Ours (no Procrustes)</td>
<td>35.74</td>
</tr>
<tr>
<td><strong>Ours (final)</strong></td>
<td><strong>50.72</strong></td>
</tr>
</tbody>
</table>

*ICCV’19 [Avetisyan et al.]: End-to-End Alignment*
## CAD Alignment Accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPFH (Rusu et al.)</td>
<td>4.45</td>
</tr>
<tr>
<td>SHOT (Tombari et al.)</td>
<td>3.14</td>
</tr>
<tr>
<td>Li et al.</td>
<td>6.03</td>
</tr>
<tr>
<td>3DMatch (Zeng et al.)</td>
<td>10.29</td>
</tr>
<tr>
<td>Scan2CAD (Avetisyan et al.)</td>
<td>31.68</td>
</tr>
<tr>
<td><strong>Ours</strong></td>
<td><strong>50.72</strong></td>
</tr>
</tbody>
</table>
Unconstrained (In-The-Wild)

3D Scan

In-The-Wild

ICCV’19 [Avetisyan et al.]: End-to-End Alignment
## Timing

<table>
<thead>
<tr>
<th>Scene Size</th>
<th>small</th>
<th>medium</th>
<th>large</th>
</tr>
</thead>
<tbody>
<tr>
<td># Objects</td>
<td>7</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>Scan2CAD</td>
<td>288.60s</td>
<td>565.86s</td>
<td>740.34s</td>
</tr>
<tr>
<td>Ours</td>
<td>0.62s</td>
<td>1.11s</td>
<td>2.60s</td>
</tr>
</tbody>
</table>

ICCV’19 [Avetisyan et al.]: End-to-End Alignment
End-to-End Alignment: Summary

Our contributions:

- Fully-convolutional SOCs prediction pipeline

- Retrieval of CAD models with a scan query

- **Over 250x faster alignment**

- **Over 19% more accurate**
Scan2Mesh: From Unstructured Range Scans to 3D Meshes

CVPR’19 [Dai and Niessner]: Scan2Mesh
Scan2Mesh: From Unstructured Range Scans to 3D Meshes
Scan2Mesh: From Unstructured Range Scans to 3D Meshes
Scan2CAD is a super exciting direction 😊

• Learn better fits of models

• Structural elements in scenes

• Direct prediction of artist modeling steps

• Lighting, material, and textures
Democratizing Content Creation?

3D Reconstructions

Computer Graphics Renderings
Thank You

Armen Avetisyan  Manuel Dahnert  Angel Chang  Manolis Savva  Angela Dai

http://kaldir.vc.in.tum.de/scan2cad_benchmark/

Scan2CAD: Learning CAD Model Alignment in RGB-D Scans

End-to-End CAD Model Retrieval and 9DoF Alignment in 3D Scans
Thank You

Visual Computing Group
Prof. Matthias Nießner

https://niessnerlab.org