Learning to Regress Bodies from Images using Differentiable Semantic Rendering

Sai Kumar Dwivedi¹, Nikos Athanasiou¹, Muhammed Kocabas¹,², Michael J. Black¹

¹Max Planck Institute for Intelligent Systems, Tübingen, ²ETH Zürich

Motivation

Hard to estimate 3D human pose and shape for people with clothing as clothes vary a lot in shape and size

Previous Methods

- Key points
  - Do not supervise body shape
  - Do not match projected minimally-clothed SMPL

- Silhouette / Part Segmentation
  - Do not supervise clothing

SOTA Method

Our Method

High-Level Semantics

- Supervise SMPL to lie inside Clothing

- Minimal Clothing
  - Supervise SMPL to exactly match MC

Overall Idea

- SMPL with Semantic Prior

Differentiable Semantic Rendering

Clothing

Semantic Prior

Input Image

Differentiable Semantic Rendering (DSR)

Final Estimate

Resources

Project Page

https://dsr.is.tue.mpg.de/

Ablation study

<table>
<thead>
<tr>
<th>Method</th>
<th>PAMPIPE</th>
<th>MPIPE</th>
<th>PVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-EFT</td>
<td>58.5</td>
<td>101.0</td>
<td>119.3</td>
</tr>
<tr>
<td>DSR-FB</td>
<td>59.8</td>
<td>102.1</td>
<td>120.3</td>
</tr>
<tr>
<td>DSR-MC</td>
<td>58.0</td>
<td>100.2</td>
<td>117.8</td>
</tr>
<tr>
<td>DSR-C</td>
<td>58.2</td>
<td>100.6</td>
<td>118.5</td>
</tr>
<tr>
<td>DSR-MVP</td>
<td>58.6</td>
<td>100.3</td>
<td>118.1</td>
</tr>
<tr>
<td>DSR + DSR-MC (Ours)</td>
<td>57.6</td>
<td>99.8</td>
<td>117.6</td>
</tr>
</tbody>
</table>

SOTA Performance

Overall Pipeline

Input Image

Clothing Segmentation

Minimal Clothing

SMPL with Semantic Prior

Differentiable Semantic Rendering (DSR)

Final Estimate

Overall Pipeline

Semantic Prior

Input Image

High-Level Semantics

Clothing

Minimal Clothing

SMPL with Semantic Prior

Differentiable Semantic Rendering (DSR)

Final Estimate

Overall Pipeline

Semantic Prior

Input Image

High-Level Semantics

Clothing

Minimal Clothing

SMPL with Semantic Prior

Differentiable Semantic Rendering (DSR)

Final Estimate

Are 3D Joints Enough?

We over-fit a batch of H36M samples on ground-truth (GT) joints (green) and joints with DSR (blue). The weak supervision with semantic information improves accuracy.

Resources

Project Page

https://dsr.is.tue.mpg.de/