Motivation

• Cloth mesh generated by data-driven methods often lacks high frequency details.
• Propose Texture Sliding which can perturb texture coordinate so that the rendered cloth mesh appears to match the ground truth.
• Use neural network to predict Texture Sliding, and reconstruct 3D high frequency cloth mesh from two camera views of Texture Sliding.

ML Models

• Input: skeleton pose $\theta$
• Output: texture coordinate displacement $d(\theta)$
• Support Vector Regression (SVR) using $\varepsilon$-insensitive per-vertex loss and Gaussian kernel
$$\mathcal{L} = \sum_{n} \begin{cases} 0, & \text{if } \|d(\theta_n) - d(\hat{\theta}_n)\|_2 \leq \varepsilon \\ \|d(\theta_n) - d(\hat{\theta}_n)\|_2 - \varepsilon, & \text{otherwise} \end{cases}$$
• Texture Sliding Neural Network (TSNN) using per-pixel mean squared loss
$$\mathcal{L} = \sum_{n} \|d(\theta_n) - d(\hat{\theta}_n)\|_2^2$$

TSNN Results

• First row: TSNN prediction on inferred cloth
• Second row: Texture Sliding on inferred cloth
• Third row: Error between TSNN and ground truth cloth

Dataset Generation

Shoot a camera ray to each inferred cloth vertex, and find the barycentric interpolation of texture coordinate of intersection point of ground truth mesh.

Error Comparison

<table>
<thead>
<tr>
<th>Network</th>
<th>SqrtMSE ($\times 10^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferred</td>
<td>24.871 ± 7.0613</td>
</tr>
<tr>
<td>SVR</td>
<td>51.848 ± 4.1035</td>
</tr>
<tr>
<td>TSNN</td>
<td>13.335 ± 4.2924</td>
</tr>
<tr>
<td>TSNN + subdivision</td>
<td>13.591 ± 4.5194</td>
</tr>
</tbody>
</table>

References