Deep Learning Representation using Autoencoder for 3D Shape Retrieval

Zhuotun Zhu (朱卓暾)
Huazhong Uni. of Sci. & Tech.
zhuzhuotun@hust.edu.cn
Outline

• What is deep learning?
• What is 3D shape retrieval?
• How I tackle this problem by deep learning.
Deep learning

• Neural network
Deep learning

• Neural network
• Large range of applications
  • Voice recognition
  • Image search
  • Adam
Deep learning

• Neural network
• Large range of applications
• Mathematical Models
  • Autoencoder
  • Convolutional Neural Networks (CNN)
  • Restricted Boltzmann Machine (RBM)
  • Deep Belief Networks
3D Shape Retrieval

• Retrieval
3D Shape Retrieval

- Retrieval
- 3D shape retrieval
RBM

\[ E(v, h) = - \sum_{i \in \text{visible}} a_i v_i - \sum_{j \in \text{hidden}} b_j h_j - \sum_{i,j} w_{ij} v_i h_j \]

\[ p(v, h) = \frac{1}{Z} e^{-E(v,h)} \quad Z = \sum_{v,h} e^{-E(v,h)} \]

\[ \frac{\partial \log p(v)}{\partial w_{ij}} = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}} \]
Experiments

• Project the 3D shape.
• Pretrain a stacked RBMs.
• Construct an autoencoder.
• Fine-tune the structure.
• Define the distance.
Experiments

- Project each model to different viewpoints.
Experiments
Results

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>NN(%)</th>
<th>FT(%)</th>
<th>ST(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autoencoder+BoF-SIFT</td>
<td>77.5</td>
<td>52.4</td>
<td>65.4</td>
</tr>
<tr>
<td>Autoencoder</td>
<td>72.4</td>
<td>43.3</td>
<td>54.6</td>
</tr>
<tr>
<td>BoF-SIFT [10]</td>
<td>71.4</td>
<td>45.1</td>
<td>57.6</td>
</tr>
<tr>
<td>CM-BoF+GSMD [7]</td>
<td>75.4</td>
<td>50.9</td>
<td>64.0</td>
</tr>
<tr>
<td>PANORAMA [15]</td>
<td>75.3</td>
<td>47.9</td>
<td>60.3</td>
</tr>
<tr>
<td>CM-BoF [7]</td>
<td>73.1</td>
<td>47.0</td>
<td>59.8</td>
</tr>
</tbody>
</table>
Thank you!