First Order Motion Model for Image Animation

Aliaksandr Siarohin\textsuperscript{1}, Stephane Lathuiliere\textsuperscript{1,4}, Sergey Tulyakov\textsuperscript{2}, Elisa Ricci\textsuperscript{1,3} and Nicu Sebe\textsuperscript{1}

\textsuperscript{1}DISI, University of Trento; \textsuperscript{2}Snap Inc, \textsuperscript{3}Fondazione Bruno Kessler, \textsuperscript{4}LTCI, Institut polytechnique de Paris

Self-Supervised Image Animation

- Training time: we learn a self-supervised motion representation, using image reconstruction objective
- Testing time: we extract motion from driving video and appearance from source

Proposed Method

- We assume existence of abstract reference frame. We estimate reference to source $\mathcal{T}_{S \rightarrow R}(p)$ and reference to driving $\mathcal{T}_{D \rightarrow R}(p)$ motion representation using first order approximation:

$$\mathcal{T}_{X \rightarrow R}(p) = \mathcal{T}_{X \rightarrow R}(p_k) + \left( \frac{d}{dp} \mathcal{T}_{X \leftrightarrow R}(p) \bigg|_{p=p_k} \right) (p - p_k) + o(||p - p_k||)$$

- Source $\mathcal{T}_{S \rightarrow R}(p)$ and driving $\mathcal{T}_{D \rightarrow R}(p)$ motion representations are combined:

$$\mathcal{T}_{S \rightarrow D}(z) \approx \mathcal{T}_{S \rightarrow R}(p_k) + J_k(z - \mathcal{T}_{D \rightarrow R}(p_k)); J_k = \left( \frac{d}{dp} \mathcal{T}_{S \rightarrow R}(p) \bigg|_{p=p_k} \right) \left( \frac{d}{dp} \mathcal{T}_{D \rightarrow R}(p) \bigg|_{p=p_k} \right)^{-1}$$

- From $\mathcal{T}_{S \rightarrow D}(z)$ optical flow and occlusion mask is predicted
- Representation of the source image is warped and missing parts are inpainted

Results on different datasets

- User Study

<table>
<thead>
<tr>
<th>Method</th>
<th>Tai-Chi-HD</th>
<th>Nemo</th>
<th>Bair</th>
<th>VoxCeleb</th>
</tr>
</thead>
<tbody>
<tr>
<td>X2Face vs First Order Model</td>
<td>92.0%</td>
<td>79.8%</td>
<td>95.0%</td>
<td>95.8%</td>
</tr>
<tr>
<td>MonkeyNet vs First Order Model</td>
<td>80.6%</td>
<td>60.6%</td>
<td>67.0%</td>
<td>68.4%</td>
</tr>
</tbody>
</table>

Our code is publicly available: https://github.com/AliaksandrSiarohin/first-order-model