Iterative GANs for Rotating Visual Objects

Ysbrand Galama Thomas Mensink
ysbrand@ygalama.nl thomas.mensink@uva.nl
University of Amsterdam

Summary

- Rotate objects without e.g. CAD-models
- Single model of generic object rotation
- Transform images by iterating a generator
- End-to-end learning extending Pix2pix
- Iterative model outperforms direct rotation

Object manipulation in images

- 3D objects in images
  - an image has enough information for humans to predict its 3D shape & rotation
- Static body transformation
  - e.g. rotation creates a slightly changed image
- Infer the unseen shape & texture
- Single generic rotation model
  - the same trained model should be capable of rotating object of multiple classes

Why iterative?

- Smaller translation of pixels since Pix2pix[1] is trained on pixel-mappings
- Mental rotation experiment[2]
- Implicitly constrain linearity i.e., \( R^1 R^2 = R^3 \)

IterGAN

- Generator \( G_\theta(A) \) and discriminator \( D_\phi(B|A) \)
  - eight layer encode-decoder U-net from Pix2pix
- Iterating generator \( B = G_\theta(G_\theta(\ldots G_\theta(A) \ldots)) \)
- Weights between iterations shared
  - i.e. do the same thing \( k \) times
- Additional discriminator \( D_\mu(B_i) \)
  - more control on intermediate steps
- Interpret intermediate \( B_i \) as partial rotation

References


Results on ALOI

- ALOI: images of 1000 household objects
  - Amsterdam Library of Object Images
- Viewpoints of 5° rotation, 72 images per object
  - Train set of 800, test set of 100 train and 100 unseen objects of 36 angles

Evaluation

- Baselines: Identity, Projective and Pix2pix
- \( L_1^M \)-score: \( L_1 \) with foreground mask

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean ( L_1^M )-scores (±std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity</td>
<td>.298 ± .154</td>
</tr>
<tr>
<td>Projective</td>
<td>.457 ± .156</td>
</tr>
<tr>
<td>Pix2pix</td>
<td>.210 ± .092</td>
</tr>
<tr>
<td>IG_6</td>
<td>.162 ± .084</td>
</tr>
<tr>
<td>IG_6_M</td>
<td>.147 ± .055</td>
</tr>
<tr>
<td>IG_6_M+U</td>
<td>.139 ± .090</td>
</tr>
</tbody>
</table>

Cumulative \( L_1^M \)-scores

Extentions — see our arXiv paper

IterGANs: Iterative GANs to Learn and Control 3D Object Transformation

- More experiments on unseen data and trained and tested on a different dataset: VKITTI
- Control the amount of rotation train and test on differing angles/iterations \((0, \ldots, 30, 45, 60, 90)\)
- Use additional supervision \( D_\mu(B_i|A) \)

Future:

- Change direction of the rotation left/right or up/down rotation, and append to more transformations