Age Progression/Regression by Conditional Adversarial Autoencoder

Zhifei Zhang, Yang Song, and Hairong Qi
Outline

1. Introduction
2. Traversing on the Manifold
3. Approach
4. Experimental Evaluation
Introduction

“If I provide you a face image of mine (without telling you the actual age when I took the picture) and a large amount of face images that I crawled (containing labeled faces of different ages but not necessarily paired), can you show me what I would look like when I am 80 or what I was like when I was 5?”
Introduction

Regression/Rejuvenation

Given face

Progression/Aging

5
8 years old
20
40
60
Introduction

Face Aging

Physical Model

Prototype-based

Deep learning

Ramanathan, et al., 2006 (CVPR), 2008 (AFGR)

Suo, et al., 2010 (PAMI)

Kemelmacher, et al., 2014 (CVPR)

Shu, et al., 2015 (ICCV)

Wang, et al., 2016 (CVPR)
Introduction

Personalized Age Progression with Aging Dictionary [Shu, et al., 2015 (ICCV)]
Introduction

Recurrent Face Aging [Wang, et al., 2016 (CVPR)]
Introduction

Our idea [Zhang, et al., 2017 (CVPR)]

Assumptions:

• The faces lie on a manifold ($M$)

• Clustered by ages and personality

• Traversing on the manifold corresponds to age/personality transformation.
# Introduction

<table>
<thead>
<tr>
<th></th>
<th>Physical Model</th>
<th>Prototype (Dictionary)</th>
<th>Deep Learning (RNN)</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reality</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Group-wise learning</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Label required during testing</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Require long/short age span</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complicate in modeling</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Traversing on the Manifold

\[ \mathcal{M} \]

\[ \begin{align*} x_1 & \quad E \quad \rightarrow \quad \text{Personality} (z) \quad \rightarrow \quad \text{Latent space} \quad \rightarrow \quad G \quad \rightarrow \quad \text{Label} \\ x_2 & \quad \rightarrow \quad \end{align*} \]

\[ \begin{align*} \text{Age} (l) \quad & \quad [z_1, l_1] \\ \text{Personality} (z) \quad & \quad [z_2, l_2] \\ \end{align*} \]
Traversing on the Manifold

Personality ($z$)

Age ($l$)
Approach

Encoder $E$

Discriminator on $z$ -- $D_z$

Discriminator on face image -- $D_{img}$

Generator $G$
Approach

Reconstruction error

\[
\min_{E,G} \max_{D_z, D_{img}} \lambda \mathcal{L}(x, G(E(x), l)) + \gamma \text{TV}(G(E(x), l))
\]

\[
+ \mathbb{E}_{z^* \sim p(z)} [\log D_z(z^*)]
+ \mathbb{E}_{x \sim p_{data}(x)} [\log(1 - D_z(E(x)))]
\]

\[
+ \mathbb{E}_{x,l \sim p_{data}(x,l)} [\log D_{img}(x,l)]
+ \mathbb{E}_{x,l \sim p_{data}(x,l)} [\log(1 - D_{img}(G(E(x), l)))]
\]

(5)

Total variation

where \(\text{TV}(\cdot)\) denotes the total variation which is effective in removing the ghosting artifacts. The coefficients \(\lambda\) and \(\gamma\) balance the smoothness and high resolution.
Approach

Effect of the Discriminator on $z$

$\begin{align*}
\text{With } D_z \\
\end{align*}$

$\begin{align*}
\text{Without } D_z \\
\end{align*}$
Approach

Effect of the Discriminator on images
Approach

Comparison to Related Structures

GAN

Conditional GAN

VAE

AAE

CAAE
Experimental Evaluation

Data Collection:

• MORPH dataset. 55,000 faces of 13,000 subjects from 16 to 77 years old
• CACD dataset. 163,446 faces of 2,000 subjects from 16 to 62 years old
• Search the keywords: baby, boy, teenager, 15 years old, etc. and collect 7,670 faces.
Experimental Evaluation

Raw Images

Alignment
Experimental Evaluation

Training process
Experimental Evaluation

Prior: The BEST result achieved by existing works

Qualitative Comparison

Input
Prior
Ours

Experimental Evaluation

Prior: The BEST result achieved by existing works

Qualitative Comparison

Input
Prior
Ours

Experimental Evaluation

Prior: The BEST result achieved by existing works

Qualitative Comparison

Input
Prior
Ours
Experimental Evaluation

Qualitative Comparison

No existing work reported regression/rejuvenation results
Experimental Evaluation

Quantitative Comparison

We struggled with this comparison because there is still no a good metric to measure reality of images.
Experimental Evaluation

1. Survey for comparison with prior work

- 47 volunteers
- Each is randomly assigned 45 out of 235 results
- 1,508 voles in total

![Pie chart showing survey results](image)

Figure 1. Our Google survey form to compare with ground truth. We show the input image (left), our result (middle), and ground truth (right). Note that in this survey, we didn’t indicate which one is ground truth and which one is generated image. To avoid similarity bias, we compare with generated image under the same age group of the ground truth.
Experimental Evaluation

- 63 volunteers
- Each is randomly assigned 45 out of 865 results
- 3,208 voles in total

2. Survey for comparison with ground truth

Same to ground truth?

- Yes: 48.38%
- No: 29.58%
- Not sure: 22.04%

Figure 2. Our Google survey form to compare with prior work. We show the our result and prior work; Note that we didn’t specific which one is ours and which is prior work. We randomly order ours and prior work.
Experimental Evaluation

Tolerance to Pose, Expression, and Occlusion

TensorFlow Code is available at:
• Bitbucket: https://bitbucket.org/aiclp FACE-AGING-CAAE
• Github: https://zzutk.github.io/FACE-AGING-CAAE
Conclusion

Potential Framework for face-age related applications

Face aging

Cross-age recognition

Face morphing

Distance

Prior

Real/fake

Age estimate

Face estimating
Thank you