Supplementary Material: Face editing using a regression-based approach in the StyleGAN latent space

Saeid Motiian
motiian@adobe.com
Siavash Khodadadeh
khodadad@adobe.com
Shabnam Ghadar
ghadar@adobe.com
Baldo Faieta
bfaietas@adobe.com
Ladislau Bölöni
Ladislau.Boloni@ucf.edu

1 Adobe Inc
San Jose, CA 95110, USA
2 Department of Computer Science
University of Central Florida
Orlando, FL 32816, USA

1 Qualitative Comparison with InterfaceGAN.

Figure 1 and Figure 2 show the comparison between our method and InterfaceGAN. In general both methods work well but we noticed some small artifacts (entanglement between attributes) by using InterfaceGAN. For Hair and Age attributes, there are hair color change for InterfaceGAN-based edits. Also it looks like the Age attribute change (the person looks younger) when increasing Hair attribute for InterfaceGAN-based edits. For Smile attribute, it looks like eyes become narrower when increasing Hair attribute for InterfaceGAN-based edits. Furthermore, we observe better disentanglement when we use CLIP scores with our method. We observe more changes in skin tone, eyes and background for Curly Hair and Beaming attributes when we train them by InterfaceGAN. We include more examples in Figures 3, 4, 5, and 6.

2 Qualitative Comparison with Latent Transformer.

Figure 7 shows a comparison between our method and Latent Transformer [1] (LT). For LT, we used the trained directions and inference code available in its repo. We find that LT shows a changed identity for the wavy hair and bangs attributes, and a changed skin tone and lip color for blond hair. In general, our method shows a better retaining of identity and disentanglement.

© 2022. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.
3 More Attributes

We trained latent directions for Close Eyes (Figure 8), Bushy Eyebrows (Figure 9), Bangs (Figure 10), Eyeglasses and Gender (Figure 11).

4 Training and Inference Time

Solving Eq (5) of the main paper using the network described in section 3.3 with $L1/L2$ and orthogonality regularizations takes 5 minutes on a P3-2xlarge AWS instance with only 1 GPU. Inference time is the same as StyleGAN inference time which is around 30 milliseconds.
Figure 2: Comparison between our method and InterfaceGAN

References

Figure 3: Comparison between our method and InterfaceGAN for *Hair* attribute.
Figure 4: Comparison between our method and InterfaceGAN for Age attribute.
Figure 5: Comparison between our method and InterfaceGAN for *Smile* attribute.
Figure 6: Comparison between our method and InterfaceGAN for *Blond Hair* attribute trained with unsupervised labeling with CLIP.
Figure 7: Qualitative comparison for attribute editing (left: original, middle: Latent Transformer, right: ours)
Figure 8: Face editing for 

```
Close Eyes
```

attribute.

Figure 9: Face editing for 

```
Bushy Eyebrows
```

attribute.
Figure 10: Face editing for Bangs attribute.

Figure 11: Face editing for Eyeglasses and Gender attributes in $S$ space. The left image is the original.