

Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model

Takehiro Aoshima, Takashi Matsubara
Osaka University, Japan
TUE-PM-175

Preview

Deep generative models

- ✓ are known for generating high-quality images.
- ✗ do not provide an inherent way to edit images semantically.

➤ We address **semantic image editing** for deep generative models.

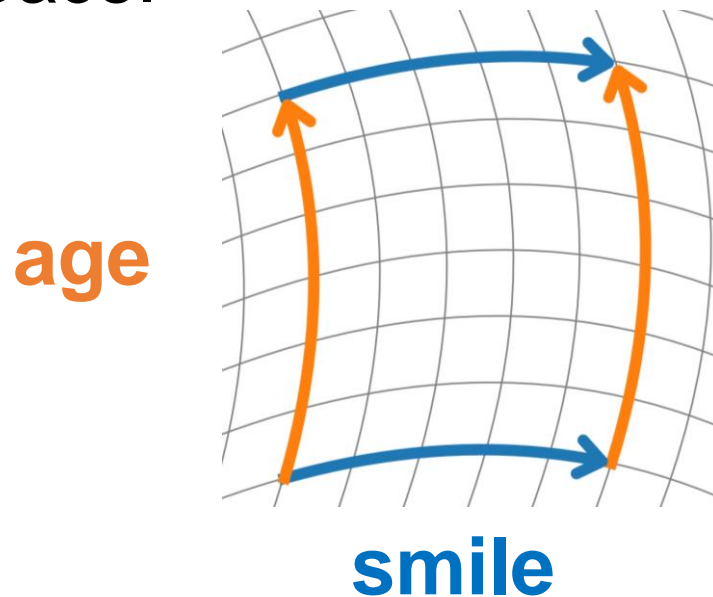


Preview

We propose the method, which provides

- ✓ higher-quality image editing
- ✓ commutative image editing

by learning semantic commuting vector fields in the latent space.



Background

Methods for image editing;

- Training models under constraints
 - ✗ Requiring computationally expensive training.
 - ✗ Conflicting with the quality of image generation.
- Image-to-image translation
 - ✗ Requiring computationally expensive training.
 - ✗ Limiting editing to be discontinuous.

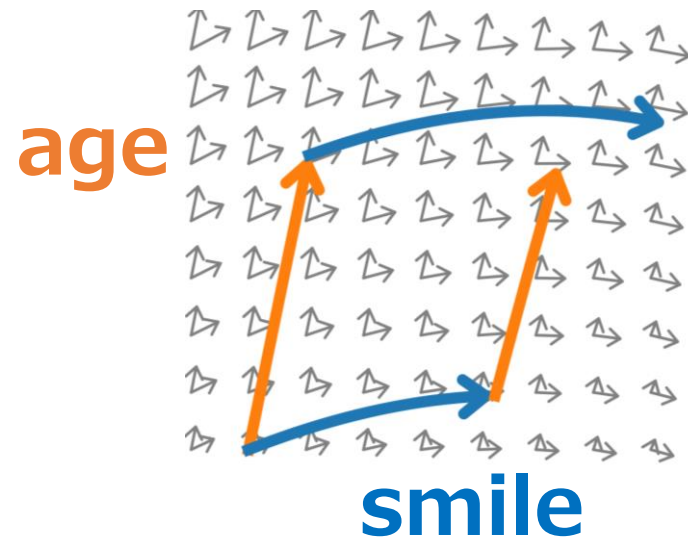
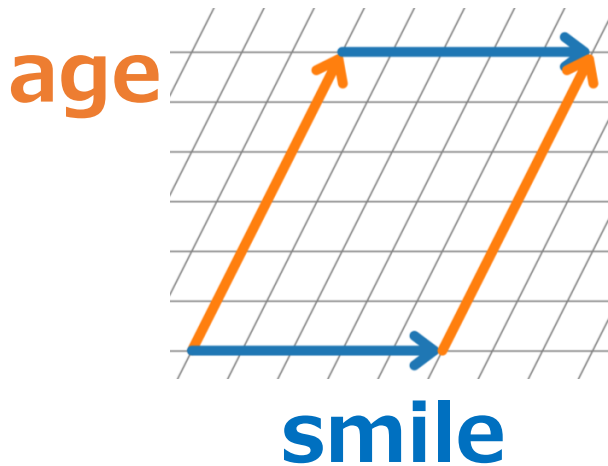


horse → zebra

Background

Methods for image editing;

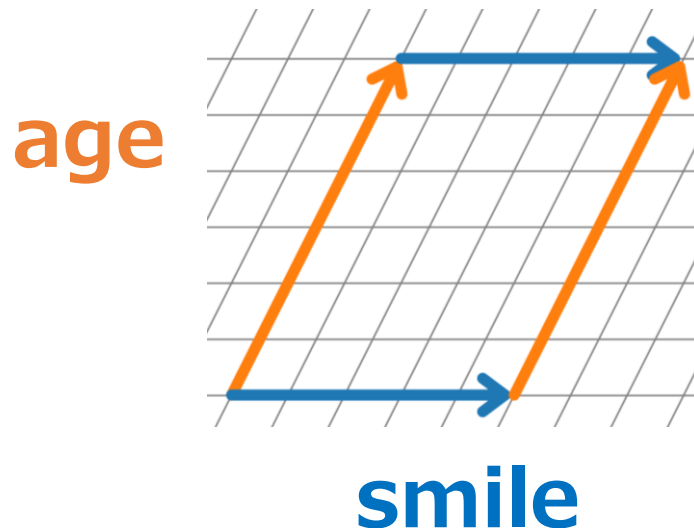
- Finding linear or nonlinear paths in latent space of pretrained models
 - ✓ Not requiring computationally expensive training.
 - ✓ Not limiting editing to be discontinuous.



Related Work

Linear methods (e.g., [1])

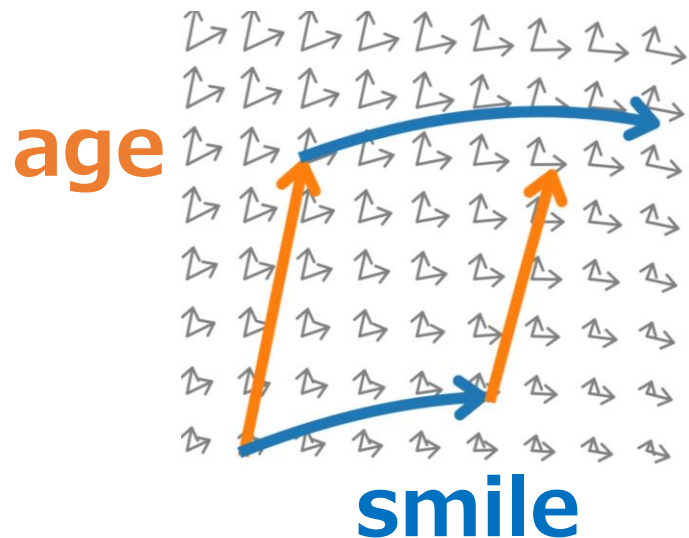
- discover **linear** paths.
- manipulate the latent code along the axis of an oblique coordinate system.
- ✓ provide commutative edits.
- ✗ sometimes fail to discover semantic paths.



Related Work

Nonlinear methods (e.g., [2])

- discover **nonlinear** paths by vector fields.
- manipulate the latent code along a vector field.
- ✓ discover more variety of paths.
- ✗ do not provide commutative edits.



Proposed Method

We propose the method, which

- ✓ discovers nonlinear paths
- ✓ provides commutative edits

by learning semantic curvilinear coordinates.

commuting vector fields



Experiment

We adopted the unsupervised training framework for GANs.

Pretrained GANs

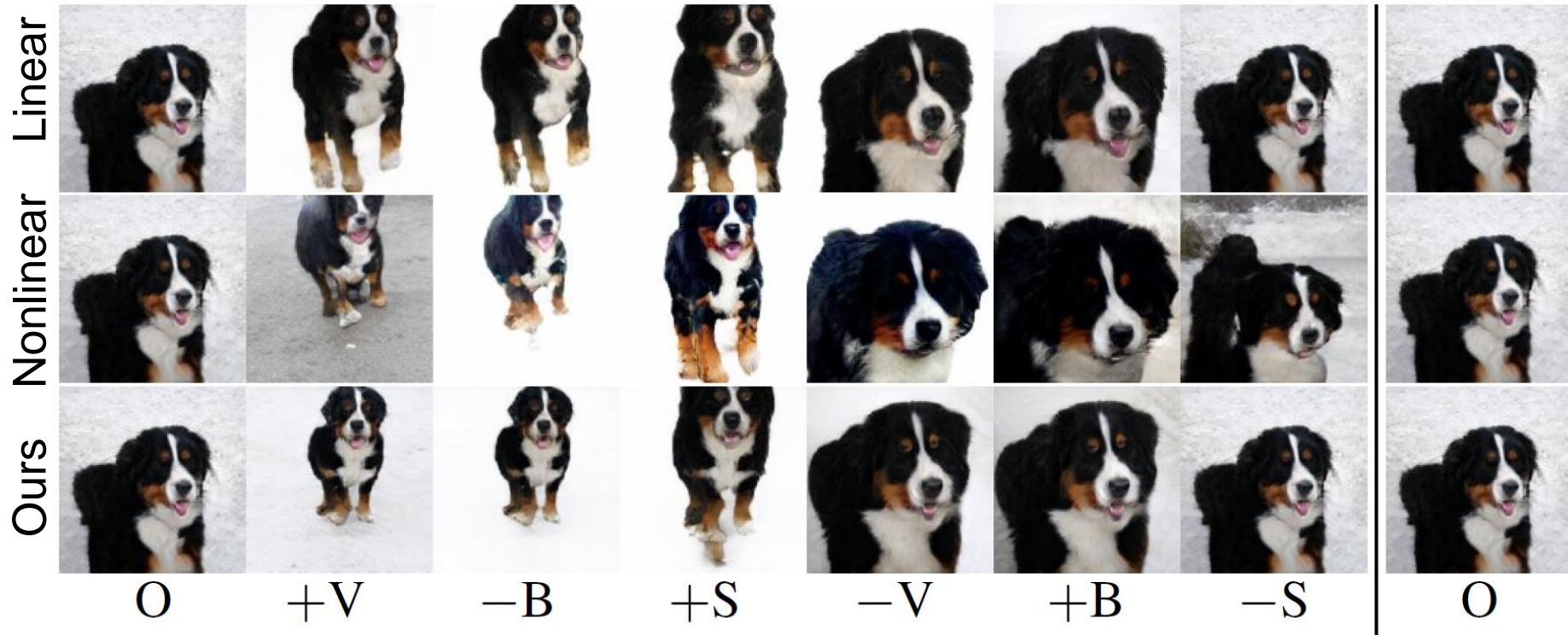
- SNGAN trained on MNIST dataset.
- SNGAN trained on AnimeFaces dataset.
- BigGAN trained on ILSVRC dataset.
- ProgGAN trained on CelebA-HQ dataset.
- StyleGAN2 trained on CelebA-HQ dataset.
- StyleGAN2 trained on LSUN Car dataset.

Comparison methods

- LinearGANSpace (a linear method) [1]
- WarpedGANSpace (a nonlinear method) [2]

Results (commutativity)

- We visualize the results of editing attributes of an image sequentially.



O: original, V: vertical position, B: background, S: size

Results (commutativity)

- We visualize the results of editing attributes of an image sequentially.

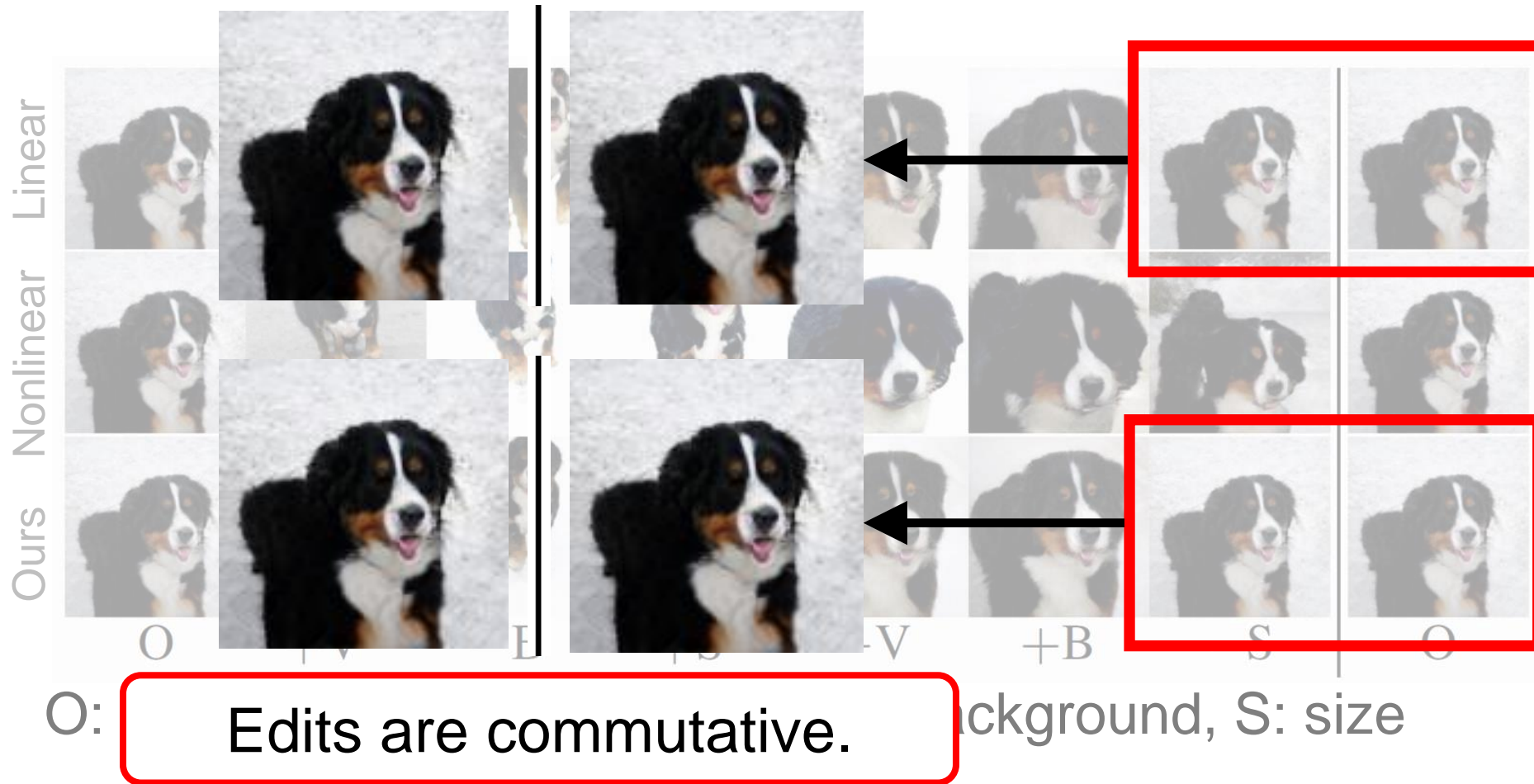


Background and object size are not identical

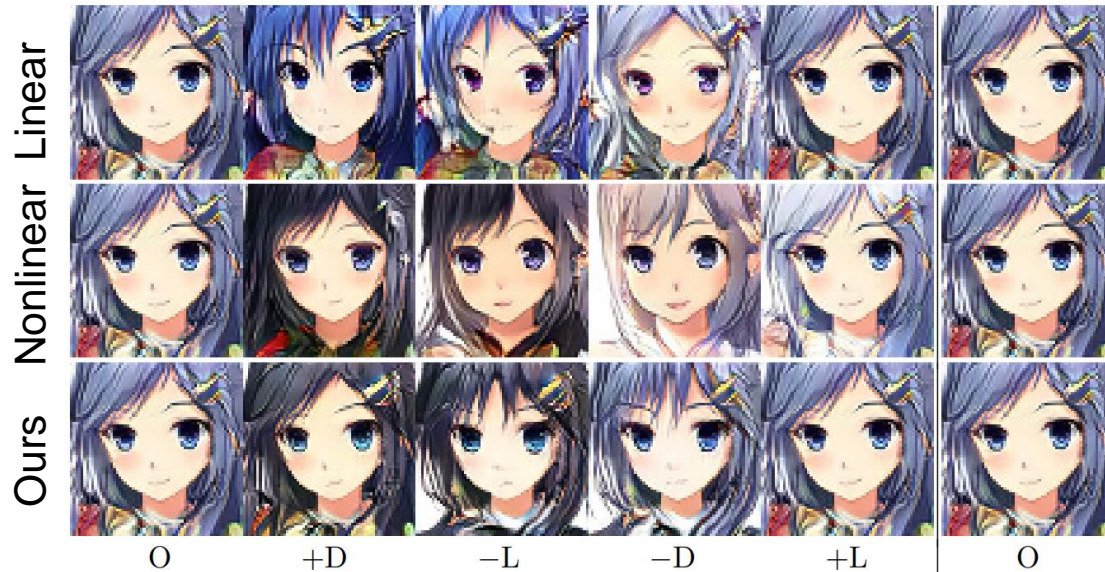
B: background, S: size

Results (commutativity)

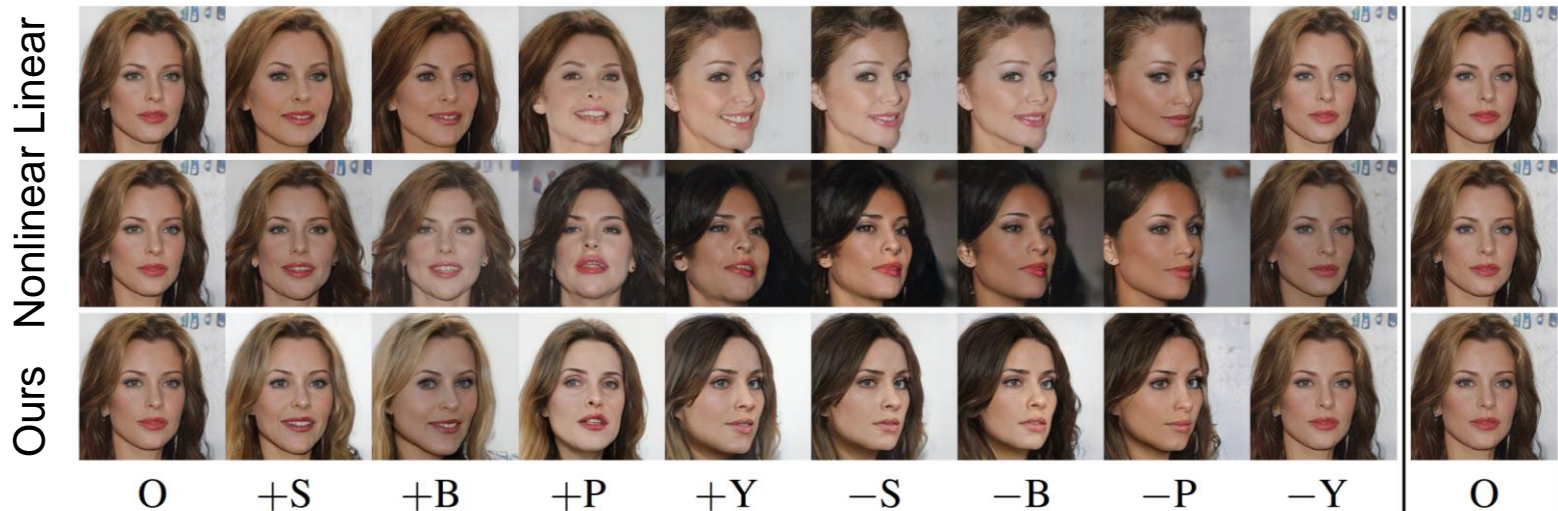
- We visualize the results of editing attributes of an image sequentially.



Results (commutativity)



O: original, D: dark colored-hair, L: hair length. SNGAN + AnimeFaces.



O: original, S: smiling, B: bangs. P: pitch, Y: yaw. ProgGAN + Celeb-A HQ.

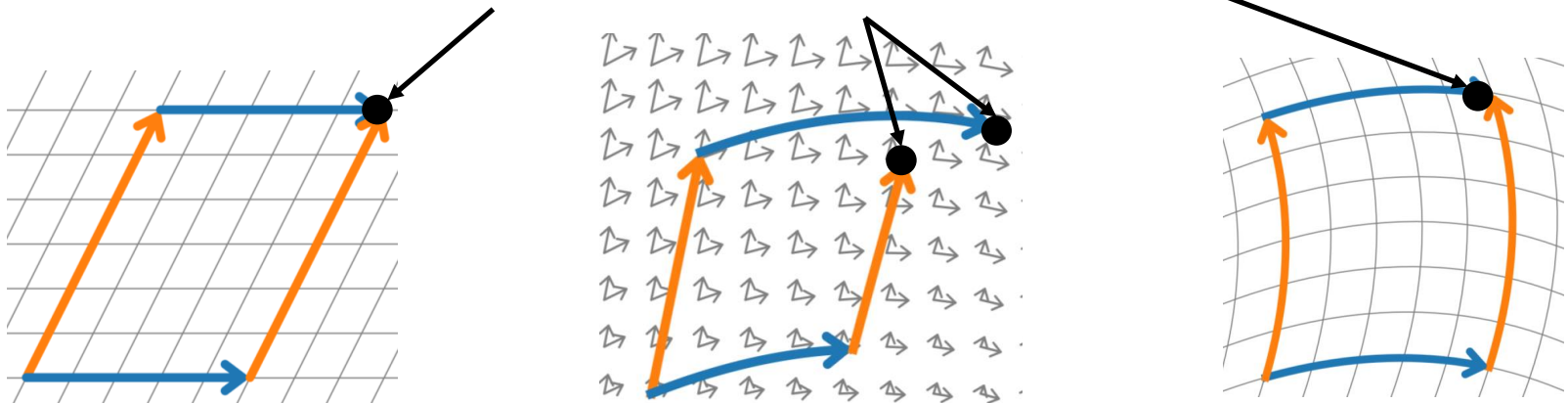
Results (commutativity)

Differences between images with edits of two attributes applied in different order (%).

| CelebA-HQ+StyleGAN2 | A+G | R+P | B+Y |
|---------------------|---------------------------|---------------------------|---------------------------|
| Linear | 0.01 / 0.05 | 0.02 / 0.07 | 0.02 / 0.15 |
| Nonlinear | 11.40 / 6.62 | 3.15 / 3.46 | 1.28 / 2.22 |
| Ours | <u>0.07</u> / <u>0.35</u> | <u>0.05</u> / <u>0.62</u> | <u>0.08</u> / <u>0.55</u> |

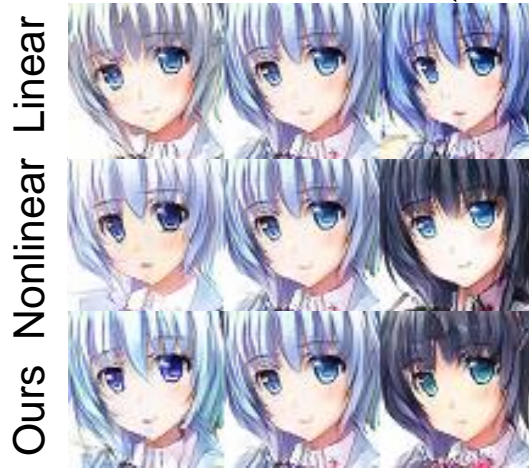
A: “age”, G: “gender”, R: “race” B: “bangs”, P: “pitch”, Y: “yaw”.

The difference of attribute score.



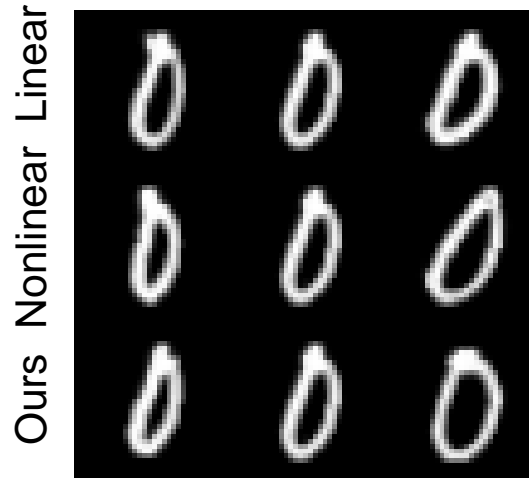
Results (quality of editing)

Not properly edit hair color.



Ours Nonlinear Linear

AnimeFaces, hair color.



Ours Nonlinear Linear

MNIST, width.

Rotating.

Not properly edit car color.

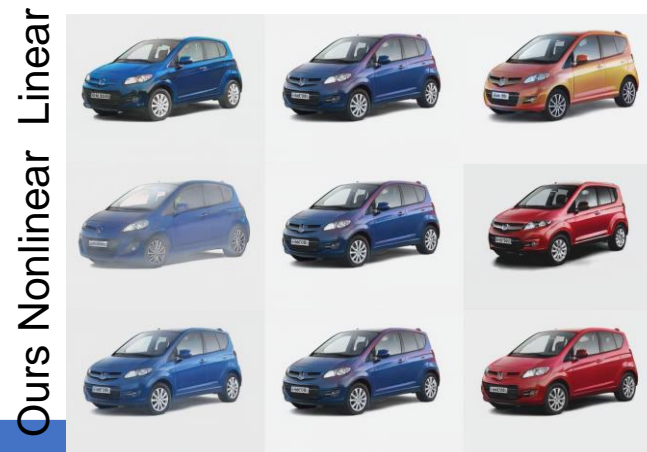
Getting thinner.

Rotating.



Ours Nonlinear Linear

CelebA-HQ, smile.



Ours Nonlinear Linear

LSUN Car, color.

Edit size.

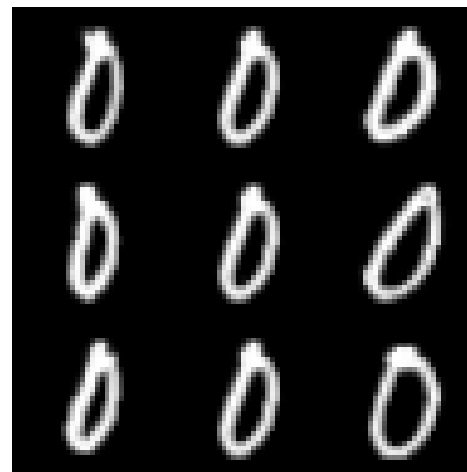
Results (quality of editing)

Only edits hair color.



AnimeFaces, hair color.

Ours Nonlinear Linear



MNIST, width.

Only edits width.

Only edit smile attributes.



CelebA-HQ, smile.

Ours Nonlinear Linear



LSUN Car, color.

Only edit car color.

Conclusion

- We propose to learn semantic curvilinear coordinates on the latent space.
- We demonstrate that the proposed method provides
 - ✓ higher-quality
 - ✓ commutativeimage editing.

