Revisiting Latent Space of GAN Inversion for Robust Real Image Editing: Supplementary Material

Kai Katsumata[1] Duc Minh Vo[1] Bei Liu[2] Hideki Nakayama[1]
[1]The University of Tokyo, [2]Microsoft Research
{katsumata,vmduc,nakayama}@nlab.ci.i.u-tokyo.ac.jp, bei.liu@microsoft.com

Table A. Mean and variance of LPIPS and MSE losses on StyleGAN2 with Car and Church dataset. The proposed latent space $\mathcal{F}/\mathcal{Z}^+$ outperforms baseline spaces excepting unregularized one.

Cat				
Space	$\mathcal{F}/\mathcal{W}^+$	$\mathcal{F}/\mathcal{W}^+(P_N)$	$\mathcal{F}/\mathcal{Z}$	$\mathcal{F}/\mathcal{Z}^+$
LPIPS	0.1031 ± 0.0287	0.1731 ± 0.0445	0.1849 ± 0.0394	$\mathbf{0.1501} \pm 0.0392$
MSE	0.0186 ± 0.0129	0.0319 ± 0.0189	0.0384 ± 0.0190	$\mathbf{0.0296} \pm 0.0198$
Church				
Space	$\mathcal{F}/\mathcal{W}^+$	$\mathcal{F}/\mathcal{W}^+(P_N)$	$\mathcal{F}/\mathcal{Z}$	$\mathcal{F}/\mathcal{Z}^+$
LPIPS	0.0904 ± 0.0243	0.1430 ± 0.0384	0.1368 ± 0.0385	$\mathbf{0.1150} \pm 0.0320$
MSE	0.0321 ± 0.0156	0.0514 ± 0.0232	0.0482 ± 0.0235	$\mathbf{0.0423} \pm 0.0205$

Figure A. Editing examples of geometric operations.

Figure B. Examples of spatial attribute editing.

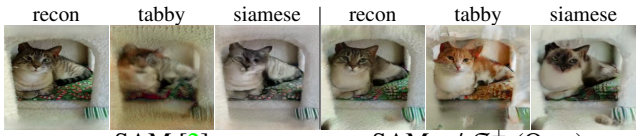

Figure C. Editing examples of StyleGAN2 LSUN Cat.

A. Additional results

We show additional quantitative results of GAN inversion in Tab. A. The $\mathcal{F}/\mathcal{Z}^+$ space achieves better reconstruction quality on several datasets. Table B shows the quantitative comparison with additional baseline methods.

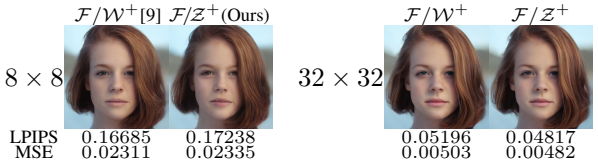

Figure D. Ablation studies on spatial code size.

As discussed in [1], our $\mathcal{F}/\mathcal{Z}^+$ space naturally can do geometric transformation. Furthermore, we show the examples of geometric transformations by $\mathcal{F}/\mathcal{Z}^+$ in Fig. A. Figure B shows the examples of spatial attribute editing. We provides editing examples of StyleGAN2 trained on the LSUN Cat dataset in Fig. C. In comparison to original SAM [2],Our approach can edit real images properly. Figure D demonstrates that $\mathcal{F}/\mathcal{Z}^+$ performs well even different spatial code size.

We also show additional qualitative results of editing quality evaluation in Figs. E to H. As shown in Figs. E to H, our proposed space maintains the image quality after editing on several data domains.

B. Image credits

From top to bottom in Fig. 2,

- https://unsplash.com/photos/7YVZYZeITc8 (photo by unsplash.com/@juricakoletic),
- https://unsplash.com/photos/rDEOVtE7vOs (photo by unsplash.com/@chrisjoelcampbell),
- https://unsplash.com/photos/ILip77SbmOE (photo by unsplash.com/@albertdera),
- https://unsplash.com/photos/m663zRzRe40 (photo by unsplash.com/@rowankyle), and
- https://unsplash.com/photos/gGTGjnWCLEM (photo by unsplash.com/@tuvaloland).

Table B. Quantitative comparison of latent spaces with the average MSE loss and SSIM on 50 CelebA-HQ samples.

Space	$\mathcal{Z}$	$\mathcal{Z}^+$	$\mathcal{W}^+(P_N)$	$\mathcal{W}^+$	$\mathcal{F}/\mathcal{Z}$	$\mathcal{F}/\mathcal{Z}^+$ (Ours)	$\mathcal{F}/\mathcal{W}^+(P_N)$ [1]	IDInvert [3]
MSE	0.18149	0.12117	0.11965	0.04872	0.02679	**0.01742**	**0.01743**	0.02155
SSIM	0.61155	0.68930	0.68190	0.76101	0.78965	**0.81477**	**0.81479**	0.64993

Figure E. Editability study with real semantic directions on StyleGAN2 with the Cat dataset.

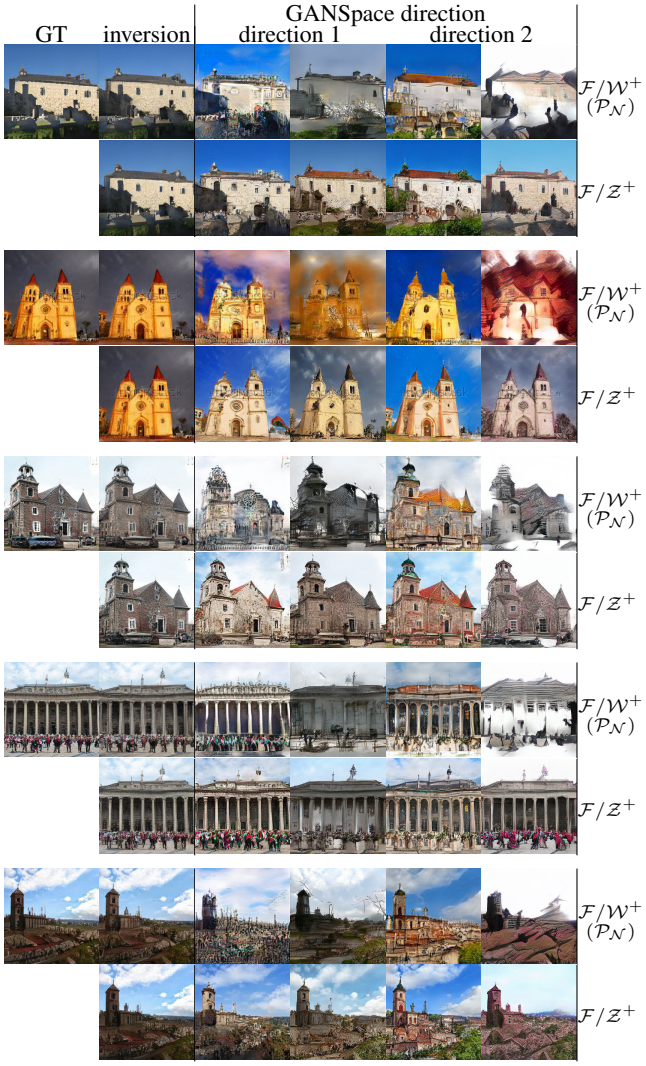

Figure F. Editability study with real semantic directions on StyleGAN2 with the Church dataset.

References

[1] Kyoungkook Kang, Seongtae Kim, and Sunghyun Cho. Gan inversion for out-of-range images with geometric transformations. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 13941–13949, October 2021. 1, 2

[2] Gaurav Parmar, Yijun Li, Jingwan Lu, Richard Zhang, Jun-Yan Zhu, and Krishna Kumar Singh. Spatially-adaptive multilayer selection for gan inversion and editing. In *CVPR*, pages 11399–11409, 2022. 1

[3] Jiapeng Zhu, Yujun Shen, Deli Zhao, and Bolei Zhou. In-domain gan inversion for real image editing. In *European conference on computer vision*, pages 592–608. Springer, 2020. 2

Figure G. Editability study with real semantic directions. The proposed space $\mathcal{F}/\mathcal{Z}^+$ properly edits the inverted images.

Figure H. Results of stepwise editing. Although the $\mathcal{F}/\mathcal{W}^+(\mathcal{P}_\mathcal{N})$ space loses image fidelity with strong editing, our space maintains image fidelity.