Pacific Graphics 2023
R. Chaine, Z. Deng, and M. H. Kim
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 42 (2023), Number 7

Controllable Garment Image Synthesis Integrated with Frequency Domain Features –Supplemental Material–

Xinru Liang, Haoran Mo, and Chengying Gao †

Sun Yat-sen University

1. Comparison with Existing Approaches

Figures 1 and 2 show results of our framework compared with FashionGAN [CLGS18], TextureGAN [XSA*18], MUNIT [HLBK18], ReferenceGAN [LKL*20], SSSIM [LZSE21] and DiSS [CCC*23].

2. Ablation Studies

2.1. Effectiveness of FFT-based Generator

We show more comparison results between a FFT-based generator (FcF generator [JZYS23]) and a generator without FFT (co-modulated StyleGAN2 [ZCS*21]) in Figure 3.

2.2. Frequency Perceptual Loss

We show more comparison results between models with and without the frequency perceptual loss in Figure 4.

2.3. High Receptive Field Perceptual Loss

We validate the effectiveness of high receptive field perceptual loss L_{hp}. As shown in Figure 5, it improves the structural details of the synthetic garment images (*e.g.,* details and realism of collar). The quantitative results are shown in Table 1.

2.4. Feature Fusion Module

We also study the effectiveness of different feature fusion modules. Our framework uses attention-based SCFT [LKL*20] to fuse the sketch feature vector and the texture feature vector from the dual-branch encoder. We also evaluation the performance of directly concatenating the two feature vectors in the channel dimension. Table 1 shows the quantitative comparisons.

3. Diversity of Texture Patterns

Figure 6 shows the diversity results with various textures.

Table 1: *Ablation study.*

	FID↓	LPIPS↓	c-FID↓	c-LPIPS↓
w/o L_{hp}	31.143	0.264	44.128	0.529
Direct concatenation	17.389	0.127	33.197	0.373
Ours	**17.146**	**0.126**	**31.796**	**0.371**

References

[CCC*23] Cheng S.-I., Chen Y.-J., Chiu W.-C., Tseng H.-Y., Lee H.-Y.: Adaptively-realistic image generation from stroke and sketch with diffusion model. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (2023), pp. 4054–4062. 1

[CLGS18] Cui Y. R., Liu Q., Gao C. Y., Su Z.: Fashiongan: display your fashion design using conditional generative adversarial nets. In *Computer Graphics Forum* (2018), vol. 37, Wiley Online Library, pp. 109–119. 1

[HLBK18] Huang X., Liu M.-Y., Belongie S., Kautz J.: Multimodal unsupervised image-to-image translation. In *Proceedings of the European conference on computer vision (ECCV)* (2018), pp. 172–189. 1

[JZYS23] Jain J., Zhou Y., Yu N., Shi H.: Keys to better image inpainting: Structure and texture go hand in hand. In *WACV* (2023). 1

[LKL*20] Lee J., Kim E., Lee Y., Kim D., Chang J., Choo J.: Reference-based sketch image colorization using augmented-self reference and dense semantic correspondence. In *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2020). 1

[LZSE21] Liu B., Zhu Y., Song K., Elgammal A.: Self-supervised sketch-to-image synthesis. In *Proceedings of the AAAI conference on artificial intelligence* (2021), vol. 35, pp. 2073–2081. 1

[XSA*18] Xian W., Sangkloy P., Agrawal V., Raj A., Lu J., Fang C., Yu F., Hays J.: Texturegan: Controlling deep image synthesis with texture patches. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 8456–8465. 1

[ZCS*21] Zhao S., Cui J., Sheng Y., Dong Y., Liang X., Chang E. I., Xu Y.: Large scale image completion via co-modulated generative adversarial networks. In *International Conference on Learning Representations (ICLR)* (2021). 1

Figure 1: *Qualitative comparison with baseline methods. The texture patch of different sizes and the garment sketch in the first and the second columns are used as the inputs.*

Figure 2: *Qualitative comparison with baseline methods. The texture patch of different sizes and the garment sketch in the first and the second columns are used as the inputs.*

Figure 3: *Comparisons between methods with and without FFT-based generator.*

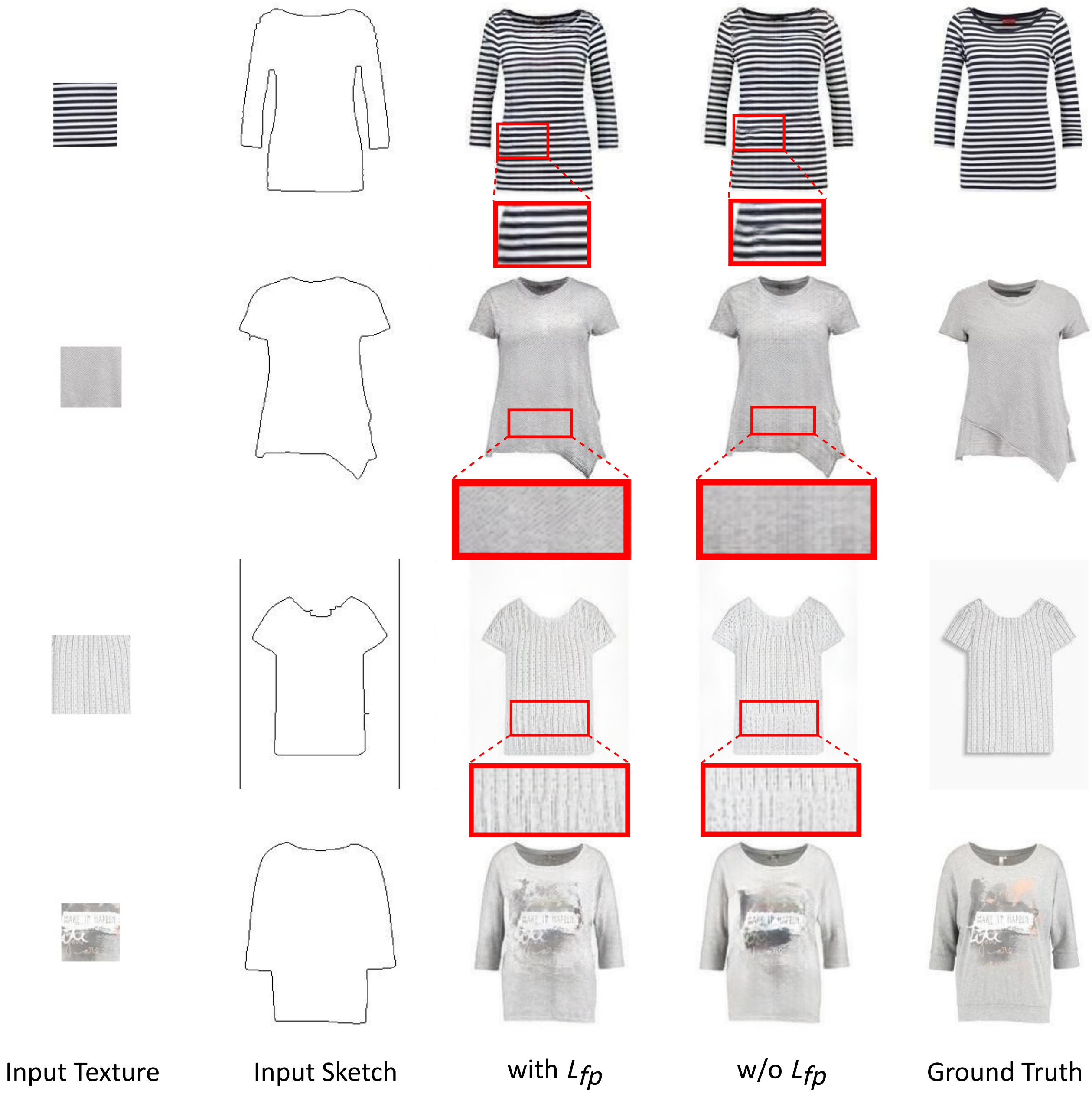

Figure 4: *Comparisons between methods with and without our proposed frequency perceptual loss (L_{fp}).*

Figure 5: *Comparisons between methods with and without high receptive field perceptual loss perceptual loss (L_{hp}).*

Figure 6: *Diversity results with various textures per garment sketch.*