Supplementary Material

Shuyi Jiang[1], Daochang Liu[1], Dingquan Li[2], Chang Xu[1]
[1]The University of Sydney, [2]Peng Cheng Laboratory
sjia6973@uni.sydney.edu.au, {daochang.liu, c.xu}@sydney.edu.au, lidq01@pcl.ac.cn

1. CVD Simulation

Based on the *two-stage theory*, this paper adopted a two-stage model to simulate the CVD gamut proposed by Machado [4]. Take the $\Delta\lambda$ as the offset distance, spectral curves of L-, M- and S-cone of CVD can be indicated as follows in the first stage:

$$L_a(\lambda) = L(\lambda + \Delta\lambda_L) \tag{1}$$
$$M_a(\lambda) = M(\lambda + \Delta\lambda_M) \tag{2}$$
$$S_a(\lambda) = S(\lambda + \Delta\lambda_S) \tag{3}$$

Then, in the second stage, the corresponding signals will be processed by the transformation matrix $T_{LMS_2O_{pp}}$ [2] into the opponent-color space as follows:

$$\begin{bmatrix} \text{WS}(\lambda) \\ \text{YB}(\lambda) \\ \text{RG}(\lambda) \end{bmatrix}_{pa} = T_{LMS_2O_{pp}} \begin{bmatrix} L_a(\lambda) \\ M(\lambda) \\ S(\lambda) \end{bmatrix} \tag{4}$$
$$\begin{bmatrix} \text{WS}(\lambda) \\ \text{YB}(\lambda) \\ \text{RG}(\lambda) \end{bmatrix}_{da} = T_{LMS_2O_{pp}} \begin{bmatrix} L(\lambda) \\ M_a(\lambda) \\ S(\lambda) \end{bmatrix} \tag{5}$$
$$\begin{bmatrix} \text{WS}(\lambda) \\ \text{YB}(\lambda) \\ \text{RG}(\lambda) \end{bmatrix}_{ta} = T_{LMS_2O_{pp}} \begin{bmatrix} L(\lambda) \\ M(\lambda) \\ S_a(\lambda) \end{bmatrix} \tag{6}$$

where pa, da, and ta represent the protan, deutan, and tritan deficiency; WS, YB and RG denote the channels of opponent-color space: white-black, yellow-blue, and red-green, respectively. By projecting the spectral power distribution $\varphi_R(\lambda)$, $\varphi_G(\lambda)$, and $\varphi_B(\lambda)$ of the RGB primaries, a transformation from RGB color space to the opponent-color space can be obtained as:

$$\begin{aligned}
\text{WS}_R &= \rho_{\text{WS}} \textstyle\int \varphi_R(\lambda)\text{WS}(\lambda)d\lambda, \\
\text{WS}_G &= \rho_{\text{WS}} \textstyle\int \varphi_G(\lambda)\text{WS}(\lambda)d\lambda, \\
\text{WS}_B &= \rho_{\text{WS}} \textstyle\int \varphi_B(\lambda)\text{WS}(\lambda)d\lambda, \\
\text{YB}_R &= \rho_{\text{YB}} \textstyle\int \varphi_R(\lambda)\text{YB}(\lambda)d\lambda, \\
\text{YB}_R &= \rho_{\text{YB}} \textstyle\int \varphi_G(\lambda)\text{YB}(\lambda)d\lambda, \\
\text{YB}_R &= \rho_{\text{YB}} \textstyle\int \varphi_B(\lambda)\text{YB}(\lambda)d\lambda, \\
\text{RG}_R &= \rho_{\text{RG}} \textstyle\int \varphi_R(\lambda)\text{RG}(\lambda)d\lambda, \\
\text{RG}_R &= \rho_{\text{RG}} \textstyle\int \varphi_G(\lambda)\text{RG}(\lambda)d\lambda, \\
\text{RG}_R &= \rho_{\text{RG}} \textstyle\int \varphi_B(\lambda)\text{RG}(\lambda)d\lambda,
\end{aligned} \tag{7}$$

where ρ_{WS}, ρ_{YB}, and ρ_{RG} are normalization factors, ensuring that

$$\begin{aligned}
\text{WS}_R + \text{WS}_G + \text{WS}_B &= 1 \\
\text{YB}_R + \text{YB}_G + \text{YB}_B &= 1 \\
\text{RG}_R + \text{RG}_G + \text{RG}_B &= 1
\end{aligned} \tag{8}$$

Therefore, the transformation matrix can be concluded as a 3×3 matrix $\Gamma_{\delta s}$, where δs denotes the degree of CVD based on the $\Delta\lambda$:

$$\Gamma_{\delta s} = \begin{bmatrix} \text{WS}_R & \text{WS}_G & \text{WS}_B \\ \text{YB}_R & \text{YB}_G & \text{YB}_B \\ \text{RG}_R & \text{RG}_G & \text{RG}_B \end{bmatrix} \tag{9}$$

In summary, the general transformation from RGB color space to opponent-color space for CVD can be defined as a 3×3 matrix $\Gamma_{\delta s}$. Let Γ be the transformation matrix for normal viewers, then the CVD simulation of an RGB image can be defined as:

$$\begin{bmatrix} R_{sim} \\ G_{sim} \\ B_{sim} \end{bmatrix} = \Gamma^{-1}\Gamma_{\delta s} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{10}$$

2. Network Structure

The generator of CVD-GAN can be summarized as Table 1, where convolution layers and modulated layers are adopted from StyleGAN-ada [3]. The structure of the discriminator follows the StyleGAN-ada [3].

Generator
$16 \times 16 \times 128$ Learnable Constant
3×3 Deconv. ReLU
3×3 ModuConv. ReLU, Latents 4
3×3 Conv. ReLU
3×3 Conv. ReLU
3×3 Conv. + Noise ReLU
3×3 Conv. + Noise ReLU
3×3 Deconv. ReLU
3×3 ModuConv. ReLU, Latents 4
3×3 Conv. ReLU
3×3 Conv. ReLU
3×3 Conv. + Noise ReLU
3×3 Conv. + Noise ReLU
3×3 Deconv. ReLU
3×3 ModuConv. ReLU, Latents 4
3×3 Conv. ReLU
3×3 Conv. ReLU
3×3 Conv. + Noise ReLU
3×3 Conv. + Noise ReLU
3×3 Deconv. ReLU
3×3 ModuConv. ReLU, Latents 4
3×3 Conv. ReLU
3×3 Conv. ReLU
3×3 Conv. + Noise ReLU
3×3 Conv. + Noise ReLU
$256 \times 256 \times 3$

Table 1. Generator of CVD-GAN

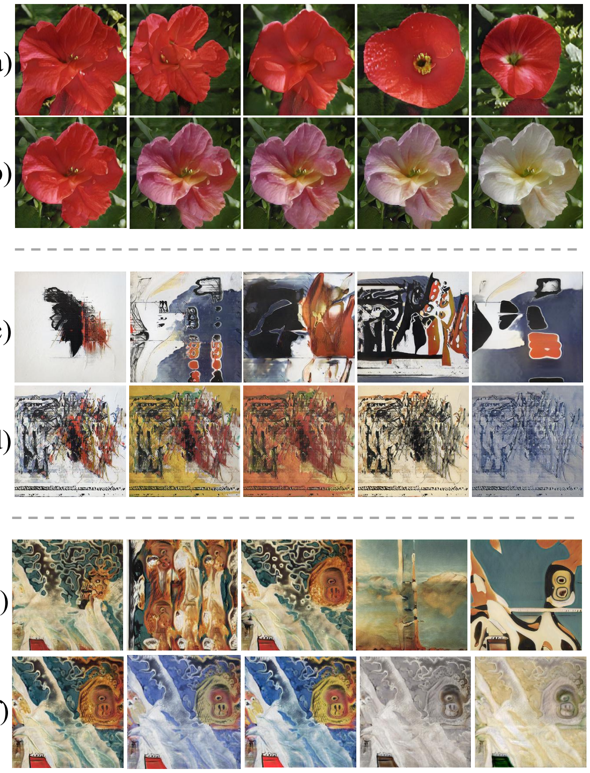

Figure 1. Examples of color representation disentanglement. For each group divided by the dotted line, the first row presents the images generated from latent codes with random $z^{\tilde{d}}$, while the second row presents the ones generated from random z^0.

3. Triple-Latent Based Color Disentanglement

The triple-latent structure consists of a contrastive group to disentangle the color representation and a control group to ensure the personalized generation. Specifically, there are two latent codes in the contrastive group in order to eliminate the dominance of other dimensions $z^{\tilde{d}}$ $\left(\tilde{d} \in (1, D]\right)$ toward the color. To better evaluate the results of disentanglement, we assign $z^{\tilde{d}}$ and z^0 random values.

Fig. 1 shows the visualization results of random assignments. For each group divided by the dotted line, the first row presents the images generated from latent codes with random $z^{\tilde{d}}$, while the second row presents the ones generated from random z^0. It is shown that the color distribution in the image is maintained although changes in the $z^{\tilde{d}}$, and it will be modified significantly only when the changes in the z^0, which means the dominance of color representation has been decoupled with $z^{\tilde{d}}$. As a result, the z^0 can dominate the color pattern generation.

With the increment δs on the z^0 during the latent traversal, CVD-GAN can generate personalized images for CVD populations with varying degrees.

Fig. 2, Fig. 3, and Fig. 4 present the results of personalized generation with an increment of [0.05, 0.2, 0.4, 0.55, 0.7, 0.9, 1.0] on the z^0 and their corresponding simulations. The fewer the change in the image after simulation, the more friendly it is to CVD populations since fewer potential perception biases will occur. It is shown that for all degrees, CVD-GAN can generate friendly images with little perception bias.

4. Image Quality Comparison

The essence of all the CVD loss including local contrast loss $\mathcal{L}_{LC}$ and color information loss $\mathcal{L}_{CI}$ is to limit the color gamut of the generated images, which will cause a negative impact on the quality of generation. This section will further discuss the trade-off between CVD-friendliness and image quality.

Table 2 shows that the image quality will reduce as the increase of severity of CVD in general since the gamut is more limited with a higher degree of CVD. Compared to other traditional post-processing recolor methods, Our method can generate comparable results regarding quality on different datasets with varying CVD degrees. It is noted that the default CVD type is protan in this table and these experiments are conducted under 4800 images.

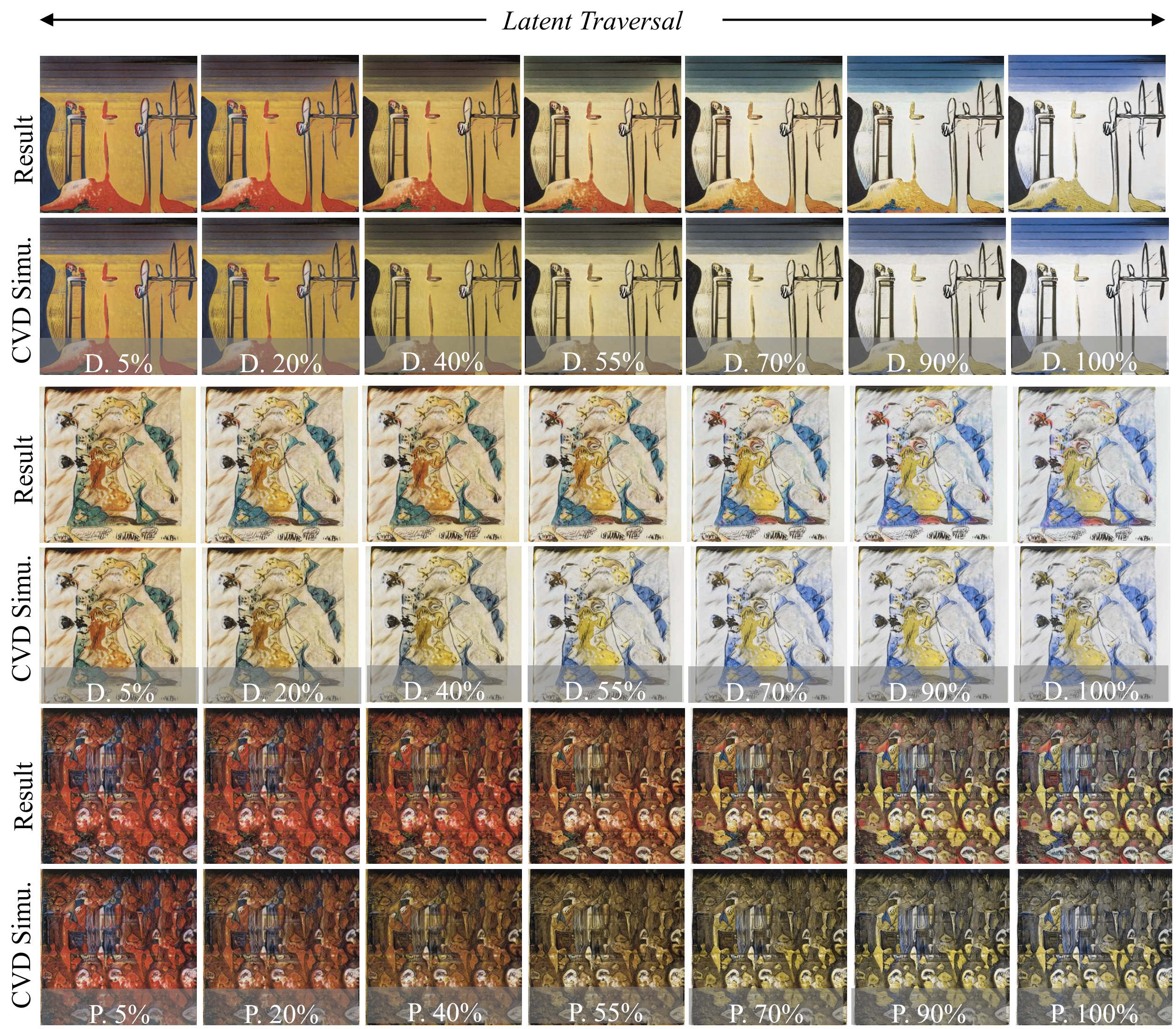

Figure 2. Examples of personalized generation of the symbolic-painting dataset, where "D" denotes deutan- and "P" denotes protan-simulation.

5. User Study

As of now, our user study is still ongoing, and we have successfully recruited 17 CVD volunteers, covering a range of ages from 20 to 54 years old. These participants are categorized into three levels: mild, medium, and severe, based on the Hue 100 test. Each volunteer rates 18 randomly selected images generated by three different methods: *StyleGAN (black box)*, *StyleGAN + Zhu (green box)*, and *our CVD-GAN (blue box)* using a Likert scale from 1 to 5. The ratings are based on the clearness and comfort level of the images, where a higher score indicates better results. The current outcomes are as Fig. 5.

According to the p-values of the t-test, ours achieves higher marks with statistical significance.

References

[1] Jia-Bin Huang, Chu-Song Chen, Tzu-Cheng Jen, and Sheng-Jyh Wang. Image recolorization for the colorblind. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1161–1164, 2009. 6

[2] Carl R Ingling Jr and Brian Huong-Peng Tsou. Orthogonal combination of the three visual channels. *Vision Research*, 17(9):1075–1082, 1977. 1

[3] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. In *Advances in Neural Infor-*

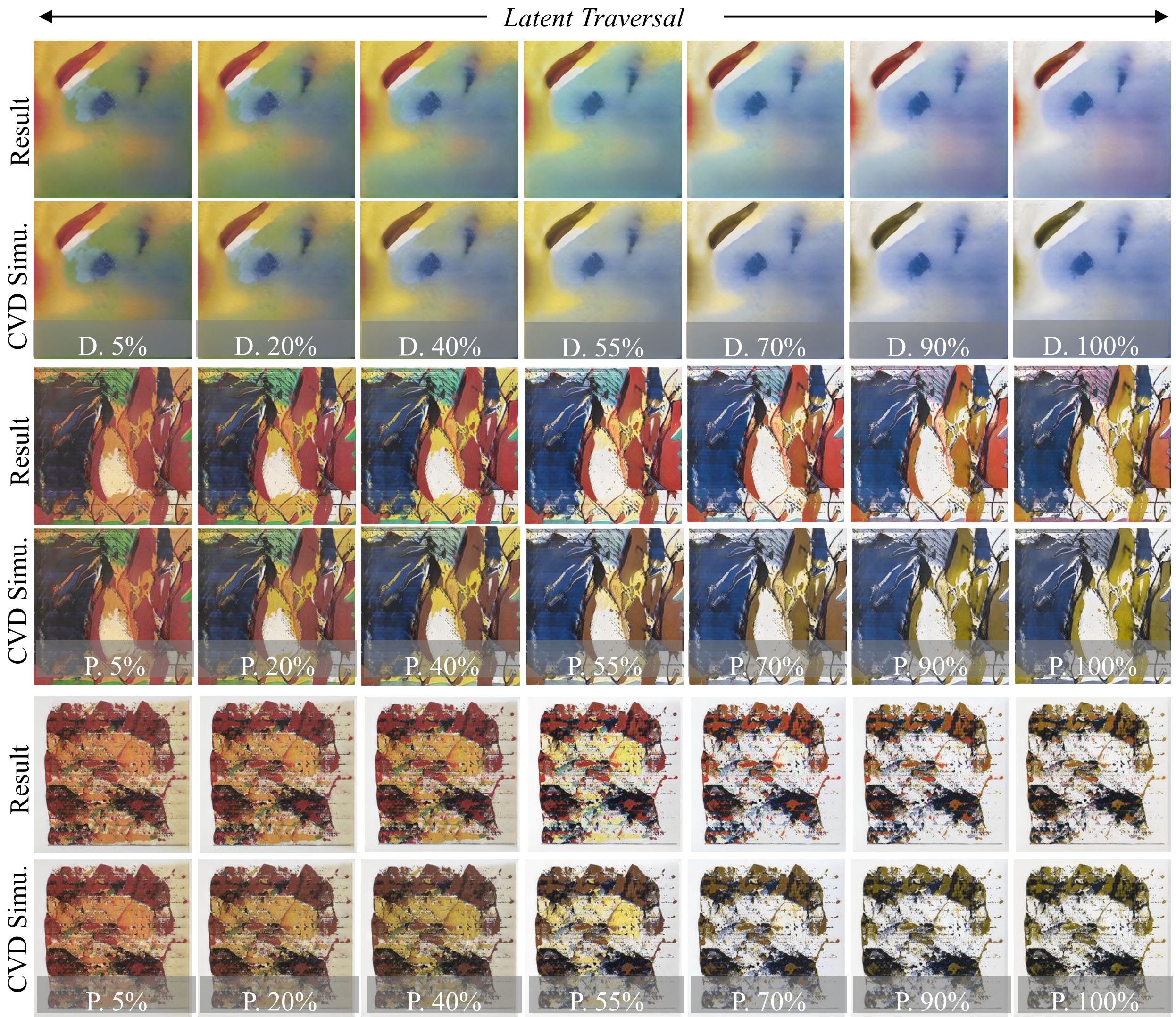

Figure 3. Examples of personalized generation of the abstract-art dataset.

mation Processing Systems, volume 33, pages 12104–12114, 2020. 1, 6

[4] Gustavo M. Machado, Manuel M. Oliveira, and Leandro A. F. Fernandes. A physiologically-based model for simulation of color vision deficiency. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1291–1298, 2009. 1

[5] Maria-Elena Nilsback and Andrew Zisserman. Automated flower classification over a large number of classes. In *Indian Conference on Computer Vision, Graphics & Image Processing*, pages 722–729, 2008. 6

[6] George Ogden. Abstract art. www.kaggle.com/datasets/goprogram/abstract-art, Accessed February 18, 2023 [Online]. 6

[7] Babak Saleh and Ahmed Elgammal. Large-scale classification of fine-art paintings: Learning the right metric on the right feature. *arXiv preprint arXiv:1505.00855*, 2015. 6

[8] Zhenyang Zhu, Masahiro Toyoura, Kentaro Go, Kenji Kashiwagi, Issei Fujishiro, Tien-Tsin Wong, and Xiaoyang Mao. Personalized image recoloring for color vision deficiency compensation. *IEEE Transactions on Multimedia*, 24:1721–1734, 2022. 6

Figure 4. Examples of personalized generation of the still-life and flower datasets.

Dataset	Degree	StyleGAN [3]	StyleGAN with		CVD-GAN (Ours)
			Zhu *et al.* [8]	Huang *et al.* [1]	
Abstract Art [6]	0%	14.35	-	-	17.73
	40%	-	16.68	-	18.27
	100%	-	23.44	16.86	19.58
Still-Life [7]	0%	18.96	-	-	22.10
	40%	-	23.42	-	24.09
	100%	-	26.36	21.91	25.36
Symbolic-Painting [7]	0%	28.20	-	-	31.66
	40%	-	29.26	-	28.37
	100%	-	30.55	28.76	28.01
Flowers [5]	0%	8.23	-	-	18.93
	40%	-	12.48	-	19.15
	100%	-	18.73	20.64	20.13

Table 2. FID of images generated by StyleGAN, post-processing recolor methods, and proposed CVD-GAN under various datasets, where the lower value indicates better image quality.

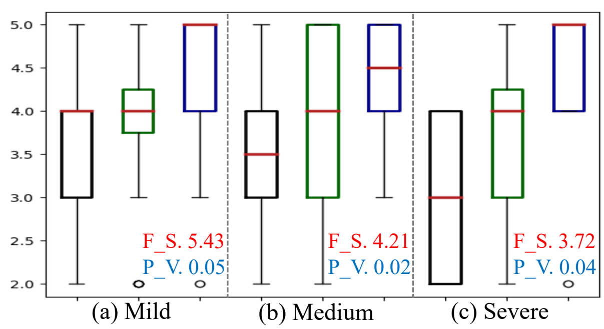

Figure 5. Result of the user study. (a), (b), and (c) showcase the ranking of populations with mild, medium, and severe CVD degrees, respectively. The notation F_S. indicates the F statistics and P_V represents the statistical significance of the collected preference results.