Face-to-Parameter Translation for Game Character Auto-Creation
Supplementary Material

1 Configurations of our networks

A detailed configuration of our Imitator G and Face Segmentation Network F_2 are listed in Table 1 and Table 2. As for the details of Face Recognition Network F_1, i.e. Light CNN-29 v2, please refer to Wu *et al.*'s paper [5].

Specifically, in a $c \times w \times w/s$ of Convolution / Deconvolution layer, c denotes the number of filters, $w \times w$ denotes the filter's size and s denotes the filter's stride. In a $w \times w/s$ of Maxpool layer, w denotes the pooling window size, and s denotes the pooling stride. In an n/s Bottleneck block [3], n denotes the number of planes, and s denotes the block's stride.

	Layer	Component	Configuration	Output Size
Imitator	Conv_1	Deconvolution + BN + ReLU	512x4x4 / 1	4x4
	Conv_2	Deconvolution + BN + ReLU	512x4x4 / 2	8x8
	Conv_3	Deconvolution + BN + ReLU	512x4x4 / 2	16x16
	Conv_4	Deconvolution + BN + ReLU	256x4x4 / 2	32x32
	Conv_5	Deconvolution + BN + ReLU	128x4x4 / 2	64x64
	Conv_6	Deconvolution + BN + ReLU	64x4x4 / 2	128x128
	Conv_7	Deconvolution + BN + ReLU	64x4x4 / 2	256x256
	Conv_8	Deconvolution	3x4x4 / 2	512x512

Table 1: A detailed configuration of our Imitator G.

	Layer	Component	Configuration	Output Resolution
Segmentation Model	Conv_1	Convolution + BN + ReLU	64x7x7 / 2	$\frac{1}{2} \times \frac{1}{2}$
	MaxPool	MaxPool	3x3 / 2	$\frac{1}{4} \times \frac{1}{4}$
	Conv_2	3 x Bottleneck	64 / 2	$\frac{1}{8} \times \frac{1}{8}$
	Conv_3	4 x Bottleneck	128 / 1	$\frac{1}{8} \times \frac{1}{8}$
	Conv_4	6 x Bottleneck	256 / 1	$\frac{1}{8} \times \frac{1}{8}$
	Conv_5	3 x Bottleneck	512 / 1	$\frac{1}{8} \times \frac{1}{8}$
	Conv_6	Convolution	11x1x1 / 1	$\frac{1}{8} \times \frac{1}{8}$

Table 2: A detailed configuration of our Face Segmentation Model F_2.

2 More examples of generated in-game characters

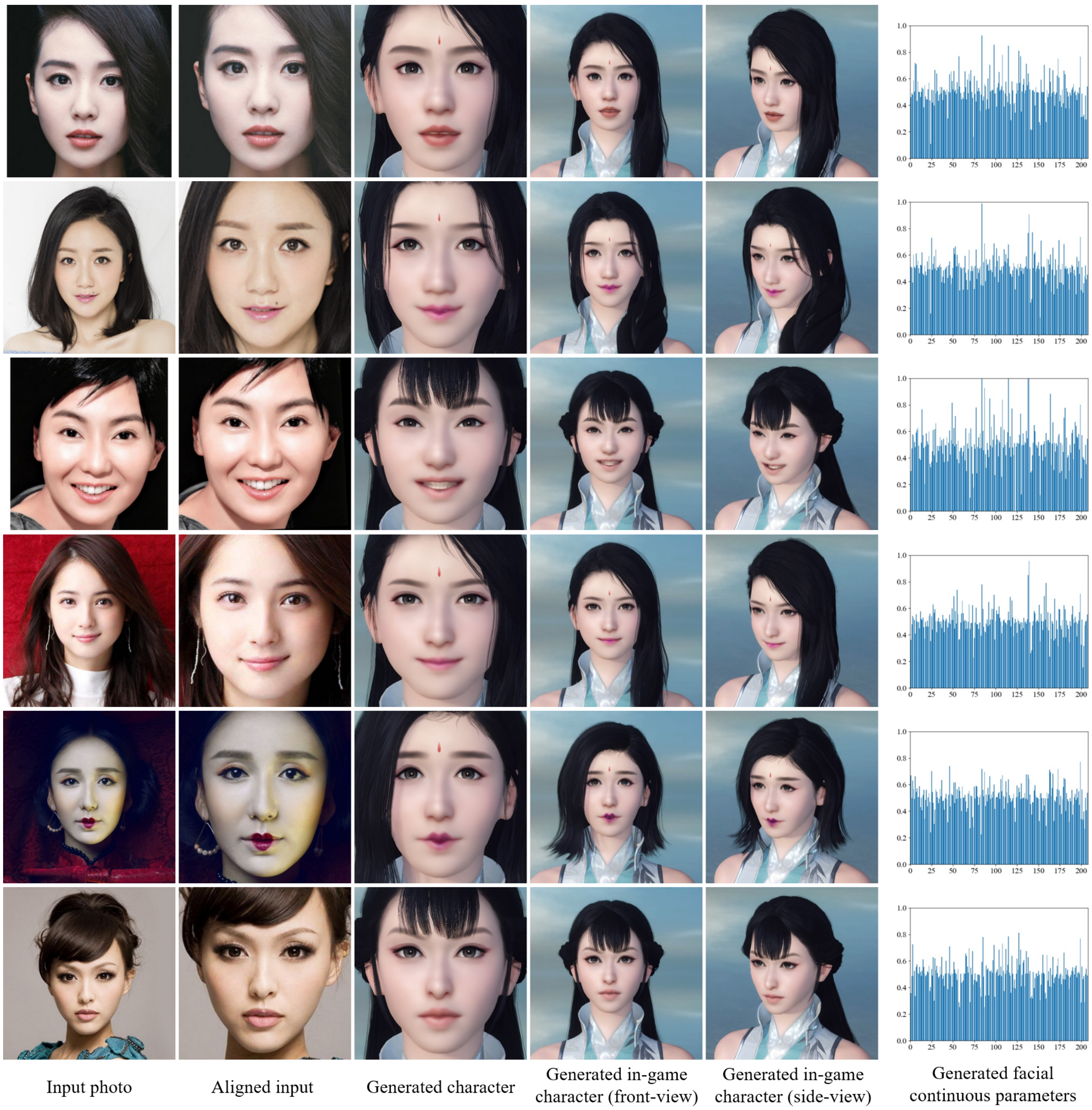

Figure 1: More generated in-game characters (female).

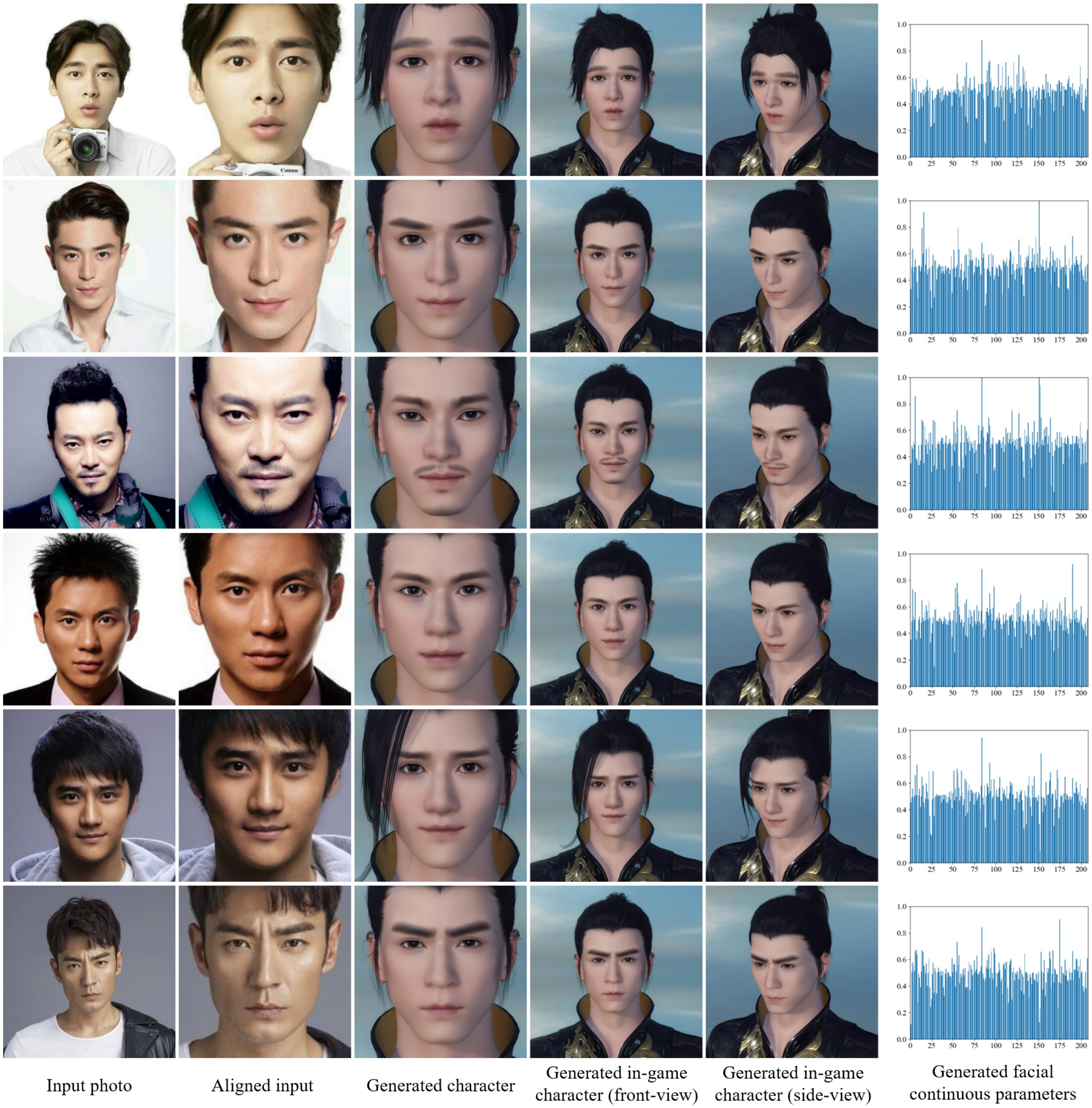

Figure 2: More generated in-game characters (male).

3 More comparison results

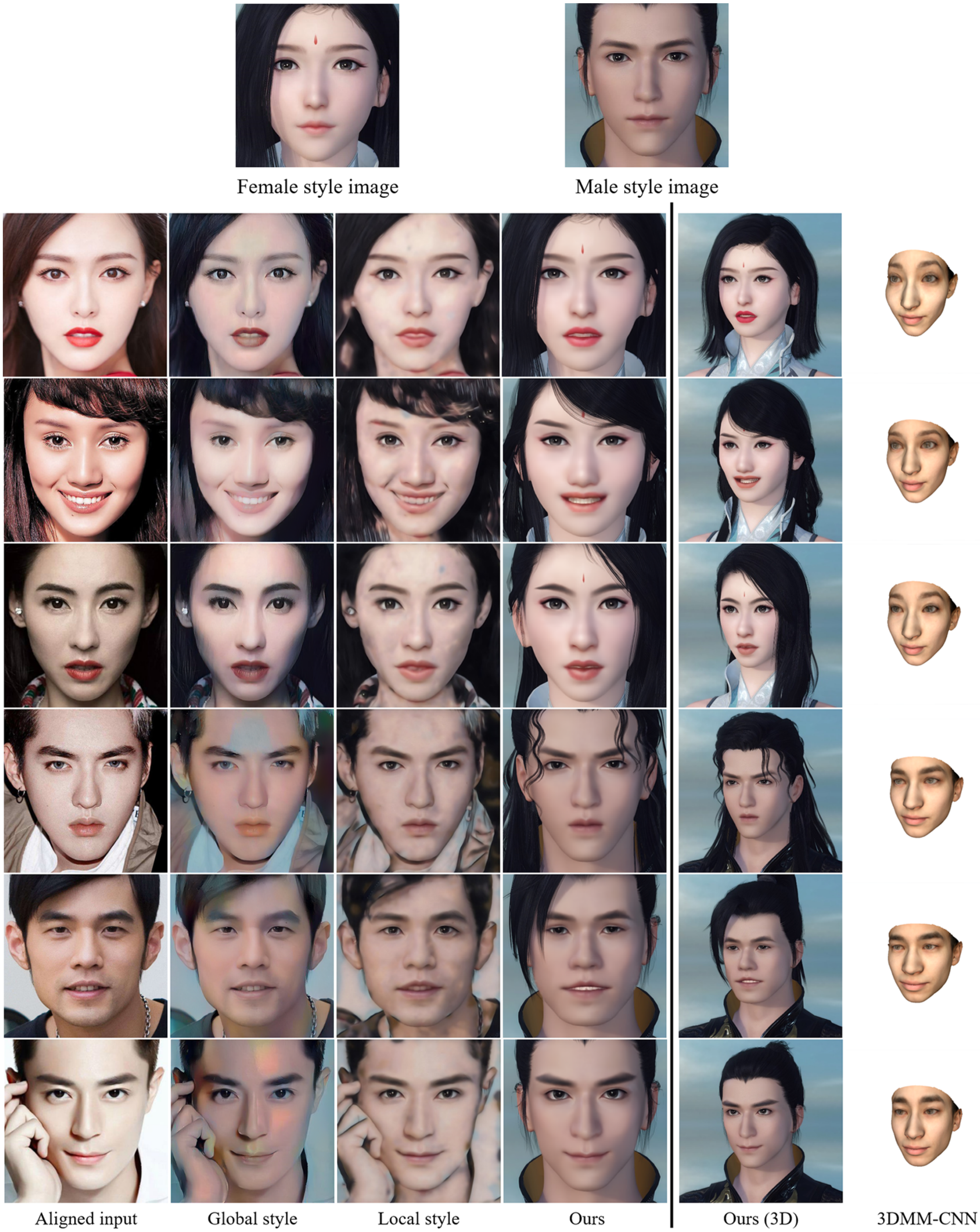

Figure 3: More comparison results with other NST methods: Global style [1] and Local style [2]. We use the "average face" of each gender as the style images. We also compare with a popular monocular 3D face reconstruction method: 3DMM-CNN [4].

4 Training samples of our Imitator

During the training process, we train our imitator with randomly generated game faces other than regular ones, as shown in Fig. 4. In our experiment, we adopt two imitators to fit female and male 3D models respectively, in order to auto-create characters for different genders.

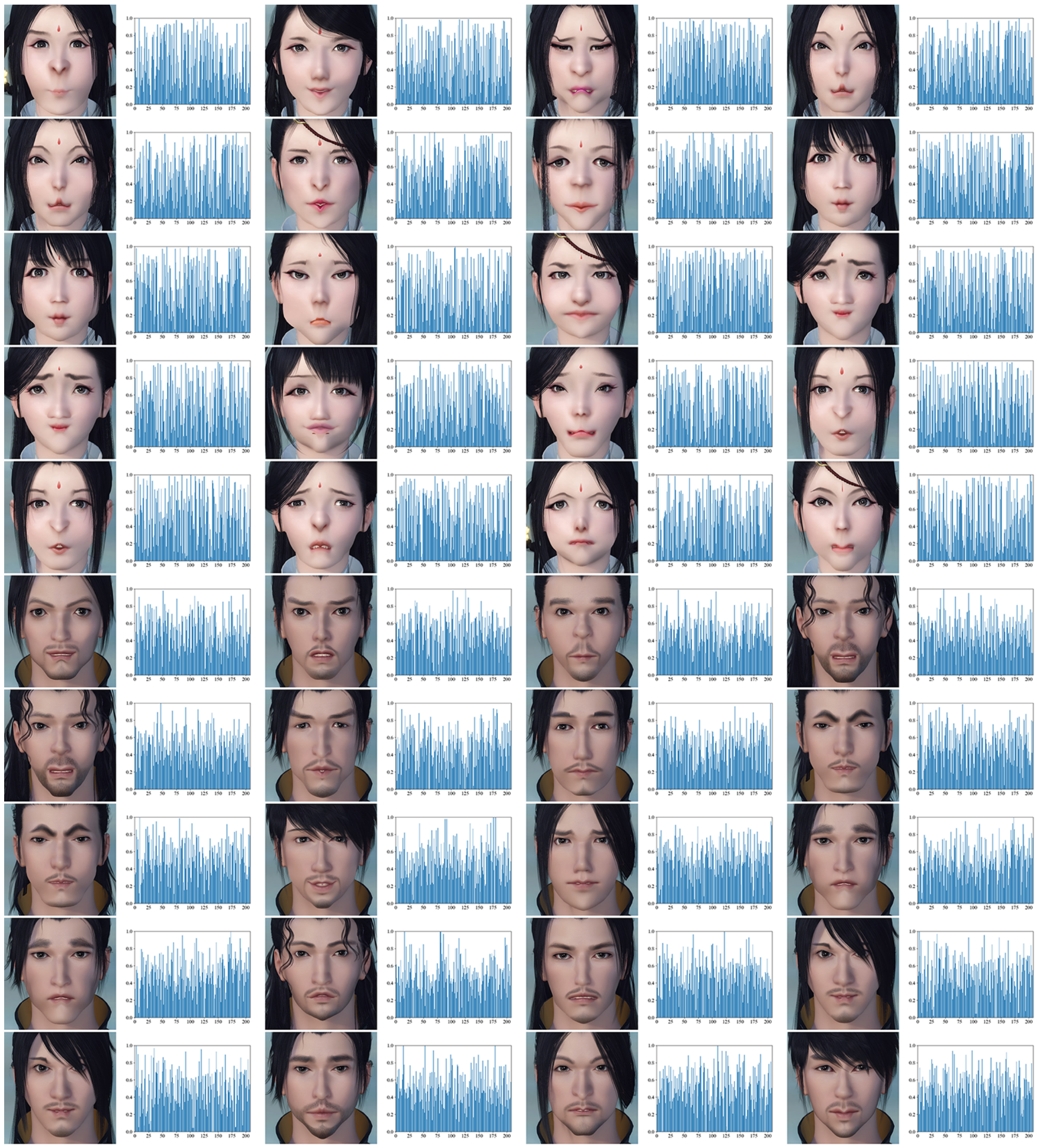

Figure 4: Training samples of imitator and the corresponding facial parameters.

5 A description of facial parameters

Table 3 lists a detailed description of each facial parameter, where the "Component" represents the facial parts which parameters belong to, the "Controllers" represents user-adjustable parameters of each facial part (one controller correspond to one continuous parameter), and the "# Controllers" represents the total number of controllers, i.e. 208. Besides, there are additional 102 discrete parameters for female (22 hair styles, 36 eyebrow styles, 19 lipstick styles, and 25 lipstick colors) and 56 discrete parameters for male (23 hair styles, 26 eyebrow styles, and 7 beard styles).

Component		Controllers	# controllers	Sum
Eyebrow	eyebrow-head	horizontal-offset, vertical-offset, slope, ...	8	208
	eyebrow-body	horizontal-offset, vertical-offset, slope, ...	8	
	eyebrow-tail	horizontal-offset, vertical-offset, slope, ...	8	
Eye	whole	horizontal-offset, vertical-offset, slope, ...	6	
	outside upper eyelid	horizontal-offset, vertical-offset, slope, ...	9	
	inside upper eyelid	horizontal-offset, vertical-offset, slope, ...	9	
	lower eyelid	horizontal-offset, vertical-offset, slope, ...	9	
	inner corner	horizontal-offset, vertical-offset, slope, ...	9	
	outer corner	horizontal-offset, vertical-offset, slope, ...	9	
Nose	whole	vertical-offset, front-back, slope	3	
	bridge	vertical-offset, front-back, slope, ...	6	
	wing	horizontal-offset, vertical-offset, slope, ...	9	
	tip	vertical-offset, front-back, slope, ...	6	
	bottom	vertical-offset, front-back, slope, ...	6	
Mouth	whole	vertical-offset, front-back, slope	3	
	middle upper lip	vertical-offset, front-back, slope, ...	6	
	outer upper lip	horizontal-offset, vertical-offset, slope, ...	9	
	middle lower lip	vertical-offset, front-back, slope, ...	6	
	outer lower lip	horizontal-offset, vertical-offset, slope, ...	9	
	corner	horizontal-offset, vertical-offset, slope, ...	9	
Face	forehead	vertical-offset, front-back, slope, ...	6	
	glabellum	vertical-offset, front-back, slope, ...	6	
	cheekbone	horizontal-offset, vertical-offset, slope, ...	5	
	risorius	horizontal-offset, vertical-offset, slope, ...	5	
	cheek	horizontal-offset, vertical-offset, width, ...	6	
	jaw	vertical-offset, front-back, slope, ...	6	
	lower jaw	horizontal-offset, vertical-offset, slope, ...	9	
	mandibular corner	horizontal-offset, vertical-offset, slope, ...	9	
	outer jaw	horizontal-offset, vertical-offset, slope, ...	9	

Table 3: A detailed interpretation of each facial parameter (continuous part).

References

[1] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[2] Shuyang Gu, Congliang Chen, Jing Liao, and Lu Yuan. Arbitrary style transfer with deep feature reshuffle. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[4] Anh Tuan Tran, Tal Hassner, Iacopo Masi, and Gerard Medioni. Regressing robust and discriminative 3d morphable models with a very deep neural network. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[5] Xiang Wu, Ran He, Zhenan Sun, and Tieniu Tan. A light cnn for deep face representation with noisy labels. *IEEE Transactions on Information Forensics and Security*, 13(11):2884–2896, 2018.