Claims

1. An image generation method based on a neural network, characterized by comprising:

receiving an initial input used to generate the image; and

based on the initial input, implementing M stages of image generation processes to generate the image with a preset target resolution;

wherein, in the M stages of image generation processes, a resolution of the image generated in each stage based on the initial input is gradually increased compared with the previous stage from the first stage to the M^{th} stage until the image reaching the preset target resolution is output in the M^{th} stage; and

the each stage of image generation process comprises at least one style control processing for affecting global information of the generated image, and, during the style control processing, pixel-instance normalization processing is performed on the image, and the pixel-instance normalization processing comprises a pixel normalization processing and an instance normalization processing;

the pixel-instance normalization processing comprises:

performing weighted combination on the pixel normalization processing and the instance normalization processing, as shown in a following formula:

$$y = \rho \cdot y^{P} + (1-\rho) \cdot y^{I},$$

wherein y represents the pixel-instance normalization processing; y^{P} represents the pixel normalization processing; y^{I} represents the instance normalization processing; and $\rho \in [0,1]^{C}$, wherein ρ represents an affect degree of the pixel normalization processing on the images, and $1-\rho$ represents an affect degree of the instance normalization processing on the images.

2. The method according to claim 1, characterized in that, based on an input initial vector, implementing the style control processing for the M stages of image generation processes.

3. The method according to claim 2, characterized in that, based on an input initial vector, implementing the style control processing for the M stages of image generation processes comprises:

encoding the initial vector into an intermediate vector through a mapping network in the neural network model, and in M stages, processing the intermediate vector as N control vectors; and the N control vectors are configured to control visual features corresponding to the images, and by taking advantage of at least one style control processing, N control vectors

affect the images,to control M stages of the image generation processes, wherein N⩾M.

4.The method according to claim 3, characterized in that, the initial vector is a random vector.

5. The method according to claims 1-4, characterized in that, the neural network model comprises an input layer, the mapping network, a generating network and an output layer, and the generating network further comprises M network layers;

wherein at least one of the M network layers comprises an up-sampling layer, a first end modulation layer, a convolutional layer and a second end modulation layer, each of the first end modulation layer and the second end modulation layer comprises at least one style control processing, and the initial input passes through the M network layers to generate the images.

6. An image generation method based on a neural network, characterized by comprising:

generating a first image;

obtaining a second image serving as a training sample; and

inputting the first image and the second image into a discrimination unit to distinguish the first image from the second image, and adjusting the generation process of the first image based on the result of the distinction;

wherein the process of generating the first image comprises:

receiving an initial input used to generate the image; and

based on the initial input, implementing M stages of image generation processes to generate the first image with a preset target resolution;

wherein, in the M stages of image generation processes, a resolution of an image generated in each stage based on the initial input is gradually increased compared with the previous stage from the first stage to the M^{th} stage until the first image reaching the preset target resolution is output in the M^{th} stage; and

the each stage of image generation process comprises at least one style control processing for affecting global information of the generated image, and, during the style control processing, pixel-instance normalization processing is performed on the image, and the pixel-instance normalization processing comprises a pixel normalization processing and an instance normalization processing;

the pixel-instance normalization processing comprises:

performing weighted combination on the pixel normalization processing and the instance

normalization processing, as shown in a following formula:

$$y = \rho \cdot y^P + (1-\rho) \cdot y^I,$$

wherein y represents the pixel-instance normalization processing; y^P represents the pixel normalization processing; y^I represents the instance normalization processing; and $\rho \in [0,1]^C$, wherein ρ represents an affect degree of the pixel normalization processing on the images, and $1-\rho$ represents an affect degree of the instance normalization processing on the images.

7. The method according to claim 6, characterized in that in the pixel-instance normalization processing, the value of ρ is adjusted according to the train automatically.

8. An image generation apparatus based on a neural network, characterized by comprising:

a receiving unit, receiving an initial input used to generate the image; and

a generation unit, comprising M generation subunits, wherein based on the initial input, the M generation subunits sequentially implement image generation processes to generate the image with a preset target resolution;

wherein, in the M generation subunits, a resolution of the image generated based on the initial input is gradually increased compared with the previous stage from the first generation subunit to the M^{th} generation subunit until the image reaching the preset target resolution is output in the M^{th} generation subunit; and

the each of the generation subunits comprises at least one style control module for affecting global information of the generated image, and, the style control module performs pixel-instance normalization processing on the image, and the pixel-instance normalization processing comprises a pixel normalization processing and an instance normalization processing;

the pixel-instance normalization processing comprises:

performing weighted combination on the pixel normalization processing and the instance normalization processing, as shown in a following formula:

$$y = \rho \cdot y^P + (1-\rho) \cdot y^I,$$

wherein y represents the pixel-instance normalization processing; y^P represents the pixel normalization processing; y^I represents the instance normalization processing; and $\rho \in [0,1]^C$, wherein ρ represents an affect degree of the pixel normalization processing on the images, and $1-\rho$ represents an affect degree of the instance normalization processing on the images.

9. The apparatus according to claim 8, characterized in that, based on an input initial vector, implementing the style control processing for the M stages of image generation processes.

10. The apparatus according to claim 9, characterized in that, based on an input initial vector, implementing the style control processing for the M stages of image generation processes comprises:

encoding the initial vector into an intermediate vector through a mapping network in the neural network model, and in M stages, processing the intermediate vector as N control vectors;and

the N control vectors are configured to control visual features corresponding to the images, and by taking advantage of at least one style control processing, N control vectors affect the images,to control M stages of the image generation processes, wherein $N \geqslant M$.

11. The apparatus according to claims 8-10, characterized in that, the neural network model comprises an input layer, the mapping network, a generating network and an output layer, and the generating network further comprises M network layers;

wherein at least one of the M network layers comprises an up-sampling layer, a first end modulation layer, a convolutional layer and a second end modulation layer, the first end modulation layer and the second end modulation layer each comprise at least one style control processing, and the initial input passes through the M network layers to generate the images.

12. An image generation apparatus based on a neural network, characterized by comprising:

a first generation unit, generating a first image;

an obtaining unit, obtaining a second image serving as a training sample; and

a discrimination unit, inputting the first image and the second image to distinguish the first image from the second image, and adjusting the generation process of the first image based on the result of the distinction;

wherein the first generation unit comprises:

a receiving unit, receiving an initial input used to generate the image;

a generation unit, comprising M generation subunits, wherein based on the initial input, the M generation subunits sequentially implement image generation processes to generate the image with a preset target resolution;

Wherein, in the M generation subunits, a resolution of an image generated based on the initial input is gradually increased compared with the previous stage from the first generation subunit to the M^{th} generation subunit until the image reaching the preset target resolution is output in the M^{th} generation subunit; and

the each of the generation subunits comprises at least one style control module for affecting global information of the generated images, and, the style control module performs pixel-instance normalization processing on the images, and the pixel-instance normalization processing comprises a pixel normalization processing and an instance normalization processing;

the pixel-instance normalization processing comprises:

performing weighted combination on the pixel normalization processing and the instance normalization processing, as shown in a following formula:

$$y = \rho \bullet y^{P} + (1 - \rho) \bullet y^{I},$$

wherein y represents the pixel-instance normalization processing; y^{P} represents the pixel normalization processing; y^{I} represents the instance normalization processing; and $\rho \in [0,1]^{C}$, wherein ρ represents an affect degree of the pixel normalization processing on the images, and $1-\rho$ represents an affect degree of the instance normalization processing on the images.

13. The apparatus according to claim 12, characterized in that in the pixel-instance normalization processing, the value of ρ is adjusted according to the train automatically.

14. Image generation device based on a neural network, characterized by comprising:

a memory, configured to store an instruction executed by one or more processors of a system, and

a processor, serving as one of the processors of the system and configured to execute the instruction to implement the image generation method of any one of claims 1-7.

15. A computer readable storage medium encoded through a computer program, characterized in that an instruction is stored on the computer readable medium, and when the instruction is executed on a computer, the computer executes the image generation method of any one of claims 1-7.