

CLAIMS

1. A method for training an image generation network, comprising:

obtaining a sample image, wherein the sample image comprises a first sample
5 image and a second sample image corresponding to the first sample image;

processing the first sample image based on an image generation network to
obtain a predicted target image;

determining a difference loss between the predicted target image and the
second sample image; and

10 training the image generation network based on the difference loss to obtain a
trained image generation network,

wherein the difference loss comprises a first structural difference loss and a
feature loss; and

wherein determining the difference loss between the predicted target image
15 and the second sample image comprises:

processing the predicted target image and the second sample image based on a
structure analysis network to determine the first structural difference loss between the
predicted target image and the second sample image; and

determining the feature loss between the predicted target image and the second
20 sample image based on the structure analysis network.

2. The method according to claim 1, wherein

training the image generation network based on the difference loss to obtain
the trained image generation network comprises:

25 performing adversarial training on the image generation network and the
structure analysis network based on the difference loss to obtain the trained image
generation network.

3. The method according to claim 2, wherein processing the predicted target image
30 and the second sample image based on the structure analysis network to determine the
first structural difference loss between the predicted target image and the second
sample image comprises:

processing the predicted target image based on the structure analysis network to determine at least one first structural feature of at least one position in the predicted target image;

5 processing the second sample image based on the structure analysis network to determine at least one second structural feature of at least one position in the second sample image; and

determining the first structural difference loss between the predicted target image and the second sample image based on the at least one first structural feature and the at least one second structural feature,

10 and/or

wherein determining the feature loss between the predicted target image and the second sample image based on the structure analysis network comprises:

15 processing the predicted target image and the second sample image based on the structure analysis network to obtain the first feature map in the at least one scale of the predicted target image and the second feature map in the at least one scale of the second sample image; and

determining the feature loss between the predicted target image and the second sample image based on at least one first feature map and at least one second feature map.

20

4. The method according to claim 3, wherein processing the predicted target image based on the structure analysis network to determine the at least one first structural feature of the at least one position in the predicted target image comprises:

25 processing the predicted target image based on the structure analysis network to obtain at least one first feature map in at least one scale of the predicted target image; and

30 obtaining, for each of the at least one first feature map, the at least one first structural feature of the predicted target image based on a cosine distance between a feature of each of at least one position in the first feature map and a feature of an adjacent region to the position, wherein each position in the first feature map corresponds to one first structural feature, and the feature of the adjacent region is each feature in a region centered on the position and comprising at least two positions.

5. The method according to claim 4, wherein processing the second sample image based on the structure analysis network to determine the at least one second structural feature of the at least one position in the second sample image comprises:

processing the second sample image based on the structure analysis network to
5 obtain at least one second feature map in at least one scale of the second sample image; and

obtaining, for each of the at least one second feature map, the at least one second structural feature of the second sample image based on a cosine distance between a feature of each of at least one position in the second feature map and a
10 feature of an adjacent region to the position, wherein each position in the second feature map corresponds to one second structural feature.

6. The method according to claim 5, wherein the each position in the first feature map has a correspondence with the each position in the second feature map;

15 wherein determining the first structural difference loss between the predicted target image and the second sample image based on the at least one first structural feature and the at least one second structural feature comprises:

calculating a distance between the first structural feature and the second structural feature respectively corresponding to positions having a correspondence;
20 and

determining the first structural difference loss between the predicted target image and the second sample image based on distances between all first structural features and second structural features corresponding to the predicted target image,

and/or

25 wherein the difference loss further comprises a color loss; and before training the image generation network based on the difference loss to obtain the trained image generation network, the method further comprises:

determining a color loss of the image generation network based on the color loss between the predicted target image and the second sample image;

30 wherein performing the adversarial training on the image generation network and the structure analysis network based on the difference loss to obtain the trained image generation network comprises:

adjusting, in a first iteration, a network parameter in the image generation network based on the first structural difference loss, the feature loss, and the color loss;

adjusting, in a second iteration, the network parameter in the structure analysis network based on the first structural difference loss, wherein the first iteration and the second iteration are two continuously-executed iterations; and

obtaining the trained image generation network when a training stopping condition is satisfied.

7. The method according to claim 3, wherein each position in the first feature map has a correspondence with each position in the second feature map;

determining the feature loss between the predicted target image and the second sample image based on the at least one first feature map and the at least one second feature map comprises:

calculating a distance between a feature in the first feature map and a feature in the second feature map respectively corresponding to the positions having a correspondence; and

determining the feature loss between the predicted target image and the second sample image based on the feature in the first feature map and the feature in the second feature map.

8. The method according to any one of claims 3 to 7, wherein before determining the difference loss between the predicted target image and the second sample image, the method further comprises:

adding noise to the second sample image to obtain a noise image; and

determining a second structural difference loss based on the noise image and the second sample image.

9. The method according to claim 8, wherein determining the second structural difference loss based on the noise image and the second sample image comprises:

processing the noise image based on the structure analysis network to determine at least one third structural feature of at least one position in the noise image;

processing the second sample image based on the structure analysis network to determine the at least one second structural feature of the at least one position in the second sample image; and

5 determining the second structural difference loss between the noise image and the second sample image based on the at least one third structural feature and the at least one second structural feature.

10 10. The method according to claim 9, wherein processing the noise image based on the structure analysis network to determine the at least one third structural feature of the at least one position in the noise image comprises:

processing the noise image based on the structure analysis network to obtain at least one third feature map in at least one scale of the noise image; and

15 obtaining, for each of the at least one third feature map, the at least one third structural feature of the noise image based on a cosine distance between a feature of each of at least one position in the third feature map and a feature of an adjacent region to the position, wherein each position in the third feature map corresponds to one third structural feature, and the feature of the adjacent region is each feature in a region centered on the position and comprising at least two positions.

20 11. The method according to claim 10, wherein the each position in the third feature map has a correspondence with the each position in the second feature map;

wherein determining the second structural difference loss between the noise image and the second sample image based on the at least one third structural feature and the at least one second structural feature comprises:

25 calculating a distance between the third structural feature and the second structural feature respectively corresponding to positions having a correspondence; and

determining the second structural difference loss between the noise image and the second sample image based on distances between all third structural features and second structural features corresponding to the noise image.

30

12. The method according to any one of claims 8 to 11, wherein performing the adversarial training on the image generation network and the structure analysis

network based on the difference loss to obtain the trained image generation network comprises:

adjusting, in a third iteration, the network parameter in the image generation network based on the first structural difference loss, the feature loss, and the color loss;

adjusting, in a fourth iteration, the network parameter in the structure analysis network based on the first structural difference loss and the second structural difference loss, wherein the third iteration and the fourth iteration are two continuously-executed iterations; and

obtaining the trained image generation network when the training stopping condition is satisfied,

and/or

wherein after processing the predicted target image based on the structure analysis network to determine the at least one first structural feature of the at least one position in the predicted target image, the method further comprises:

performing image reconstruction processing on the at least one first structural feature based on an image reconstruction network to obtain a first reconstructed image; and

determining a first reconstruction loss based on the first reconstructed image and the predicted target image.

13. The method according to claim 12, wherein after processing the second sample image based on the structure analysis network to determine the at least one second structural feature of the at least one position in the second sample image, the further comprises:

performing image reconstruction processing on the at least one second structural feature based on an image reconstruction network to obtain a second reconstructed image; and

determining a second reconstruction loss based on the second reconstructed image and the second sample image.

14. The method according to claim 13, wherein performing the adversarial training on the image generation network and the structure analysis network based on the difference loss to obtain the trained image generation network comprises:

adjusting, in a fifth iteration, the network parameter in the image generation network based on the first structural difference loss, the feature loss, and the color loss;

adjusting, in a sixth iteration, the network parameter in the structure analysis network based on the first structural difference loss, the second structural difference loss, the first reconstruction loss, and the second reconstruction loss, wherein the fifth iteration and the sixth iteration are two continuously-executed iterations; and

obtaining the trained image generation network when the training stopping condition is satisfied.

15. A method for image processing, comprising:

in a three-dimensional image generation scene, inputting a left-view image into an image generation network to obtain a right-view image; and

generating a three-dimensional image based on the left-view image and the right-view image, wherein the image generation network is obtained through training by using the method for training an image generation network according to any one of claims 1 to 14.

16. An apparatus for training an image generation network, comprising:

a sample obtaining unit, configured to obtain a sample image, wherein the sample image comprises a first sample image and a second sample image corresponding to the first sample image;

a target predicting unit, configured to process the first sample image based on an image generation network to obtain a predicted target image;

a difference loss determining unit, configured to determine a difference loss between the predicted target image and the second sample image; and

a network training unit, configured to train the image generation network based on the difference loss to obtain a trained image generation network,

wherein the difference loss comprises a first structural difference loss and a feature loss; and wherein the difference loss determining unit is further configured to:

process the predicted target image and the second sample image based on a structure analysis network to determine the first structural difference loss between the predicted target image and the second sample image; and

5 determine the feature loss between the predicted target image and the second sample image based on the structure analysis network.

17. An apparatus for image processing, comprising:

10 a right-view image obtaining unit, configured to, in a three-dimensional image generation scene, input a left-view image into an image generation network to obtain a right-view image; and

15 a three-dimensional image generating unit, configured to generate a three-dimensional image based on the left-view image and the right-view image, wherein the image generation network is obtained through training by using the method for training an image generation network according to any one of claims 1 to 14.

18. An electronic device, comprising:

20 a processor; and
 a memory configured to store processor-executable instructions, wherein the processor is configured to implement the method for training an image generation network according to any one of claims 1 to 14 and/or the method for image processing according to claim 15.

25 19. A computer storage medium, having computer-readable instructions stored therein, wherein the instructions, when being executed, cause to perform operations of the method for training an image generation network according to any one of claims 1 to 14, and/or operations of the method for image processing according to claim 15.

30 20. A computer program product, comprising computer-readable codes, wherein the computer-readable codes, when running in a device, cause a processor in the device to execute instructions for implementing the method for generating an image generation network according to any one of claims 1 to 14, and/or execute instructions for implementing the method for image processing according to claim 15.