CLAIMS

1. A virtual avatar generation method, comprising:

determining a target task associated with at least one target face property, wherein the target face property is one of a plurality of predefined face properties, the target face property comprises at least one predefined subclass, and each of the at least one predefined subclass comprises at least one face property feature;

performing, according to the target task, target face property analysis on a target image comprising at least a face, to obtain a target face property feature associated with the target face property of the target image;

determining a target virtual avatar template corresponding to the target face property feature according to predefined correspondence between face property features and virtual avatar templates; and

generating a virtual avatar of the target image based on the target virtual avatar template;

wherein performing, according to the target task, target face property analysis on a target image comprising at least a face, to obtain a target face property feature associated with the target face property of the target image comprises:

determining a target neural network corresponding to the target face property;

inputting the target image into the target neural network to obtain estimated values output from the target neural network, wherein the estimated values represent respective probabilities that the target image has one or more face property features associated with the target face property; and

for a first subclass of the at least one subclass included in the target face property, taking, as the target face property feature corresponding to the first subclass, a face property feature corresponding to a maximum value among the estimated values output from the target neural network for the first subclass.

2. The method according to claim 1, wherein the target neural network is trained by:

inputting at least one sample image comprising at least a face into a first neural network, wherein each of the at least one sample image is labeled with at least one face property feature that is associated with a first face property of the plurality of predefined face properties, and the first neural network comprises a first sub-network corresponding to the first face property; and

training the first sub-network by taking at least one face property feature which is output from the first neural network and is associated with the first face property of the at least one sample image as a predicted value, and taking the at least one face property feature which is labeled on the at least one sample image and corresponds to the first face property as a real value, so as to obtain the target neural network after the training.

3. The method according to claim 2, wherein the first sub-network has a network structure of a residual neural network and comprises at least one residual unit.

4. The method according to claim 3, wherein

each of the at least one residual unit comprises at least one convolutional layer and at least one batch normalization layer; and

in a case where the at least one residual unit comprises a plurality of residual units, a number of convolutional layers and a number of batch normalization layers included in a second residual unit of the plurality of residual units both are greater than that included in a first residual unit of the plurality of residual units.

5. The method according to claim 3 or 4, wherein

the first sub-network further comprises an output segmentation layer, and

the output segmentation layer is configured to segment, according to one or more predefined subclasses included in the first face property, feature information extracted from the sample image, to obtain respective estimated values for one or more face property features respectively associated with the one or more subclasses.

6. The method according to any one of claims 2 to 5, further comprising:

performing affine transformation on an image of interest to obtain a frontalized face image; and

clipping an image of a target region from the frontalized face image to obtain the target image or the sample image, wherein the target region comprises at least a region where a face key point is located.

7. The method according to claim 6, wherein the target region further comprises a region with a preset area outside a face part corresponding to the target face property.

8. A virtual avatar generation apparatus, comprising:

a task determination module, configured to determine a target task associated with at least one target face property, wherein the target face property is one of a plurality of predefined face properties, the target face property comprises at least one predefined subclass, and each of the at least one predefined subclass comprises at least one face property feature;

a face property analysis module, configured to perform, according to the target task, target face property analysis on a target image comprising at least a face, to obtain a target face property feature associated with the target face property of the target image;

a virtual avatar template determination module, configured to determine a target virtual avatar template corresponding to the target face property feature according to predefined correspondence between face property features and virtual avatar templates; and

an avatar generation module, configured to generate a virtual avatar of the target image based on the target virtual avatar template;

wherein a face property analysis module, configured to determine a target neural network corresponding to the target face property;

input the target image into the target neural network to obtain estimated values output from the target neural network, wherein the estimated values represent respective probabilities that the target image has one or more face property features associated with the target face property; and

for a first subclass of the at least one subclass included in the target face property, taking, as the target face property feature corresponding to the first subclass, a face property feature corresponding to a maximum value among the estimated values output from the target neural network for the first subclass.

9. The apparatus according to any one of claims 8, wherein the apparatus further comprises a training module, and the training module is configured to:

input at least one sample image comprising at least a face into a first neural network, wherein each of the at least one sample image is labeled with at least one face property feature that is associated with a first face property of the plurality of predefined face properties, and the first neural network comprises a first sub-network corresponding to the first face property; and

train the first sub-network by taking at least one face property feature which is output from the first neural network and is associated with the first face property of the at least one sample image as a predicted value, and taking the at least one face property feature which is labeled on the at least one sample image and corresponds to the first face property as a real value, and take the trained first sub-network as the target neural network.

10. The apparatus according to claim 9, wherein the first sub-network has a network structure of a residual neural network and comprises at least one residual unit.

11. The apparatus according to claim 10, wherein

each of the at least one residual unit comprises at least one convolutional layer

and at least one batch normalization layer; and

in a case where the at least one residual unit comprises a plurality of residual units, a number of convolutional layers and a number of batch normalization layers included in a second residual unit of the plurality of residual units both are greater than that included in a first residual unit of the plurality of residual units.

12. The apparatus according to claim 10 or 11, wherein

the first sub-network further comprises an output segmentation layer, and

the output segmentation layer is configured to segment, according to one or more predefined subclasses included in the first face property, feature information extracted from the sample image, to obtain respective estimated values for one or more face property features respectively associated with one or more subclasses.

13. The apparatus according to any one of claims 8 to 12, further comprising:

a face frontalization processing module, configured to perform affine transformation on an image of interest to obtain a frontalized face image; and

an image clip module, configured to clip an image of a target region from the frontalized face image to obtain the target image or the sample image, wherein the target region comprises at least a region where a face key point is located.

14. The apparatus according to claim 13, wherein the target region further comprises a region with a preset area outside a face part corresponding to the target face property.

15. A computer-readable storage medium storing a computer program for implementing the virtual avatar generation method according to any one of claims 1 to 7.

16. A virtual avatar generation apparatus, comprising:

a processor; and

a memory storing instructions executable by the processor,

wherein the processor is configured to invoke the instructions stored in the memory to implement the virtual avatar generation method according to any one of claims 1 to 7.