

CLAIMS

1. An image recognition method, comprising:

performing image detection on an image to be recognized to obtain at least one face detection result, at least one operational part detection result and at least one trunk detection result, each face detection result comprising one face bounding box, each operational part detection result comprising one operational part bounding box, and each trunk detection result comprising one trunk bounding box;

respectively combining each of the at least one trunk detection result with each face detection result, to obtain at least one first result combination;

respectively combining each trunk detection result with each operational part detection result, to obtain at least one second result combination; and

associating the at least one first result combination with the at least one second result combination, to obtain an association result, wherein the association result is a determination as to whether the face detection result and the operational part detection result belong to a same person.

2. The method according to claim 1, wherein associating the at least one first result combination with the at least one second result combination, to obtain the association result comprises:

respectively performing feature extraction on each trunk detection result, each face detection result and each operational part detection result, to obtain a trunk feature sequence corresponding to each trunk detection result, a face feature sequence corresponding to each face detection result, and an operational part feature sequence corresponding to each operational part detection result;

associating the trunk detection result of each first result combination with the face detection result of said each first result combination according to the trunk feature sequence and the face feature sequence, to obtain a first sub-association result corresponding to each first result combination;

associating the trunk detection result of each second result combination with the operational part detection result of said each second result combination according to the trunk feature sequence and the operational part feature sequence, to obtain a second sub-association

result corresponding to each second result combination; and

determining the association result according to each first sub-association result and each second sub-association result.

3. The method according to claim 2, wherein determining the association result according to each first sub-association result and each second sub-association result comprises:

determining associated first result combinations having the trunk detection result associated with the face detection result in the at least one first result combination according to each first sub-association result;

determining associated second result combinations having the trunk detection result associated with the operational part detection result in the at least one second result combination according to each second sub-association result; and

determining that the associated first result combination and the associated second result combination having a same trunk detection result are associated, wherein a face represented by the face detection result in the associated first result combination is mutually associated with an operational part represented by the operational part detection result in the associated second result combination that is associated with the associated first result combination.

4. The method according to claim 2 or 3, wherein associating the trunk detection result of each first result combination with the face detection result of said each first result combination according to the trunk feature sequence and the face feature sequence, to obtain the first sub-association result corresponding to each first result combination, comprises:

for each of the at least one first result combination, obtaining a first combination feature corresponding to the first result combination by splicing the face feature sequence corresponding to the face detection result in the first result combination and the trunk feature sequence corresponding to the trunk detection result in the first result combination; and

classifying each first result combination using the first combination feature corresponding to each first result combination, to obtain the first sub-association result corresponding to each first result combination.

5. The method according to claim 2 or 3, wherein associating the trunk detection result of each

second result combination with the operational part detection result of said each second result combination according to the trunk feature sequence and the operational part feature sequence, to obtain the second sub-association result corresponding to each second result combination, comprises:

for each of the at least one second result combination, obtaining a second combination feature corresponding to the second result combination by splicing the trunk feature sequence corresponding to the trunk detection result in the second result combination and the operational part feature sequence corresponding to the operational part detection result in the second result combination; and

classifying each second result combination using the second combination feature corresponding to each second result combination, to obtain the second sub-association result corresponding to each second result combination.

6. The method according to claim 2, wherein respectively performing feature extraction on each trunk detection result, each face detection result and each operational part detection result, to obtain the trunk feature sequence corresponding to each trunk detection result, the face feature sequence corresponding to each face detection result, and the operational part feature sequence corresponding to each operational part detection result, comprises:

respectively performing feature extraction on a trunk image in each trunk detection result, a face image in each face detection result and an operational part image in each operational part detection result, to obtain a trunk appearance feature of each trunk detection result, a face appearance feature of each face detection result, and an operational part appearance feature of each operational part detection result;

determining a trunk position feature of each trunk detection result according to coordinate information of the trunk bounding box in each trunk detection result; determining a face position feature of each face detection result according to coordinate information of the face bounding box in each face detection result; and determining an operational part position feature of each operational part detection result according to coordinate information of the operational part bounding box in each operational part detection result; and

obtaining the trunk feature sequence of each trunk detection result by splicing the trunk appearance feature and the trunk position feature of each trunk detection result; obtaining the

face feature sequence of each face detection result by splicing the face appearance feature and the face position feature of each face detection result; and obtaining the operational part feature sequence of each operational part detection result by splicing the operational part appearance feature and the operational part position feature of each operational part detection result.

7. The method according to claim 6, wherein determining the trunk position feature of each trunk detection result according to the coordinate information of the trunk bounding box in each trunk detection result, determining the face position feature of each face detection result according to coordinate information of the face bounding box in each face detection result, and determining the operational part position feature of each operational part detection result according to coordinate information of the operational part bounding box in each operational part detection result, comprises:

generating a trunk region position parameter of each trunk detection result according to the coordinate information of each trunk bounding box; generating a face region position parameter of each face detection result according to the coordinate information of each face bounding box; and generating an operational part region position parameter of each operational part detection result according to the coordinate information of each operational part bounding box;

based on a preset trunk sequence bit number, a preset trunk feature construction model and a trunk feature construction parameter, constructing a trunk position sequence of each trunk detection result according to the trunk region position parameter of each trunk detection result, and generating the trunk position feature of each trunk detection result according to the trunk position sequence of each trunk detection result, wherein the trunk feature construction parameter is generated according to the preset trunk sequence bit number;

based on a preset face sequence bit number, a preset face feature construction model, and a face feature construction parameter, constructing a face position sequence of each face detection result according to the face region position parameter of each face detection result, and generating the face position feature of each face detection result according to the face position sequence of each face detection result, wherein the face feature construction parameter is generated according to the preset face sequence bit number; and

based on a preset operational part sequence bit number, a preset operational part feature

construction model, and an operational part feature construction parameter, constructing an operational part position sequence of each operational part detection result according to the operational part region position parameter of each operational part detection result, and generating the operational part position feature of each operational part detection result according to the operational part position sequence of each operational part detection result, wherein the operational part feature construction parameter is generated according to the preset operational part sequence bit number.

8. The method according to claim 7, wherein when a preset sequence bit number is an even number bit and a preset feature construction model is a sine construction model, the position sequence is determined by:

dividing the region position parameter of each detection result by the feature construction parameter, to obtain a first parameter ratio of each detection result;

performing sine value calculation on the first parameter ratio of each detection result using the sine construction model, to obtain a first position element of each detection result; and

determining the position sequence of each detection result according to the first position element of each detection result,

wherein when the preset sequence bit number is the preset trunk sequence bit number, then the preset feature construction model is the preset trunk feature construction model, the feature construction parameter is the trunk feature construction parameter, the detection result is the trunk detection result, the region position parameter is the trunk region position parameter, and the position sequence is the trunk position sequence;

when the preset sequence bit number is the preset face sequence bit number, then the preset feature construction model is the preset face feature construction model, the feature construction parameter is the face feature construction parameter, the detection result is the face detection result, the region position parameter is the face region position parameter, and the position sequence is the face position sequence; and

when the preset sequence bit number is the preset operational part sequence bit number, then the preset feature construction model is the preset operational part feature construction model, the feature construction parameter is the operational part feature construction parameter, the detection result is the operational part detection result, the region position parameter is the

operational part region position parameter, and the position sequence is the operational part position sequence.

9. The method according to claim 7, wherein when a preset sequence bit number is an odd number bit and a preset feature construction model is a cosine construction model, the position sequence is determined by:

dividing the region position parameter of each detection result by the feature construction parameter, to obtain a second parameter ratio of each detection result;

performing cosine value calculation on the second parameter ratio of each detection result using the cosine construction model, to obtain a second position element of each detection result; and

determining the position sequence of each detection result according to the second position element of each detection result,

wherein when the preset sequence bit number is the preset trunk sequence bit number, then the preset feature construction model is the preset trunk feature construction model, the feature construction parameter is the trunk feature construction parameter, the detection result is the trunk detection result, the region position parameter is the trunk region position parameter, and the position sequence is the trunk position sequence;

when the preset sequence bit number is the preset face sequence bit number, then the preset feature construction model is the preset face feature construction model, the feature construction parameter is the face feature construction parameter, the detection result is the face detection result, the region position parameter is the face region position parameter, and the position sequence is the face position sequence; and

when the preset sequence bit number is the preset operational part sequence bit number, then the preset feature construction model is the preset operational part feature construction model, the feature construction parameter is the operational part feature construction parameter, the detection result is the operational part detection result, the region position parameter is the operational part region position parameter, and the position sequence is the operational part position sequence.

10. The method according to claim 6, wherein respectively performing feature extraction on the

trunk image in each trunk detection result, the face image in each face detection result, and the operational part image in each operational part detection result, to obtain the trunk appearance feature of each trunk detection result, the face appearance feature of each face detection result, and the operational part appearance feature of each operational part detection result, comprises:

- obtaining at least one feature map corresponding to the image to be recognized;

- obtaining a trunk feature map of each trunk detection result from each of the at least one feature map; performing region division on the trunk feature map of each trunk detection result, to obtain at least one trunk feature sub-region, and performing maximum value feature extraction on each trunk feature sub-region, to obtain at least one trunk appearance sub-feature corresponding to each trunk detection result; and constructing a trunk appearance feature of each trunk detection result with the at least one trunk appearance sub-feature corresponding to each trunk detection result;

- obtaining a face feature map of each face detection result from each of the at least one feature map; performing region division on the face feature map of each face detection result, to obtain at least one face feature sub-region, and performing maximum value feature extraction on each face feature sub-region, to obtain at least one face appearance sub-feature corresponding to each face detection result; and constructing the face appearance feature of each face detection result using the at least one face appearance sub-feature corresponding to each face detection result; and

- obtaining an operational part feature map of each operational part detection result from each of the at least one feature map; performing region division on the operational part feature map of each operational part detection result, to obtain at least one operational part feature sub-region, and performing maximum value feature extraction on each operational part feature sub-region, to obtain at least one operational part appearance sub-feature corresponding to each operational part detection result; and constructing the operational part appearance feature of each operational part detection result using the at least one operational part appearance sub-feature corresponding to each operational part detection result.

11. An image recognition apparatus, comprising:

- a detection module, configured to perform image detection on an image to be recognized,

to obtain at least one face detection result, at least one operational part detection result and at least one trunk detection result, each face detection result comprising one face bounding box, each operational part detection result comprising one operational part bounding box, and each trunk detection result comprising one trunk bounding box;

a combination module, configured to respectively combine each of the at least one trunk detection result with each face detection result, to obtain at least one first result combination, and respectively combine each trunk detection result with each operational part detection result, to obtain at least one second result combination; and

an association module, configured to associate the at least one first result combination with the at least one second result combination, to obtain an association result, wherein the association result is a determination as to whether the face detection result and the operational part detection result belong to a same person.

12. The image recognition apparatus according to claim 11, wherein the association module is specifically configured to:

respectively preform feature extraction on each trunk detection result, each face detection result and each operational part detection result, to obtain a trunk feature sequence corresponding to each trunk detection result, a face feature sequence corresponding to each face detection result, and an operational part feature sequence corresponding to each operational part detection result;

associate the trunk detection result of each first result combination with the face detection result of said each first result combination according to the trunk feature sequence and the face feature sequence, to obtain a first sub-association result corresponding to each first result combination;

associate the trunk detection result of each second result combination with the operational part detection result of said each second result combination according to the trunk feature sequence and the operational part feature sequence, to obtain a second sub-association result corresponding to each second result combination; and

determine the association result according to each first sub-association result and each second sub-association result.

13. The image recognition apparatus according to claim 12, wherein the association module is

specifically configured to:

determine associated first result combinations having the trunk detection result associated with the face detection result in the at least one first result combination according to each first sub-association result;

determine associated second result combinations having the trunk detection result associated with the operational part detection result in the at least one second result combination according to each second sub-association result; and

determine that the associated first result combination and the associated second result combination having a same trunk detection result are associated, wherein a face represented by the face detection result in the associated first result combination is mutually associated with an operational part represented by the operational part detection result in the associated second result combination that is associated with the associated first result combination.

14. The image recognition apparatus according to claim 12 or 13, wherein the association module is specifically configured to:

for each of the at least one first result combination, obtain a first combination feature corresponding to the first result combination by splicing the face feature sequence corresponding to the face detection result in the first result combination and the trunk feature sequence corresponding to the trunk detection result in the first result combination; and

classify each first result combination using the first combination feature corresponding to each first result combination, to obtain the first sub-association result corresponding to each first result combination.

15. The image recognition apparatus according to claim 12 or 13, wherein the association module is specifically configured to:

for each of the at least one second result combination, obtain a second combination feature corresponding to the second result combination by splicing the trunk feature sequence corresponding to the trunk detection result in the second result combination and the operational part feature sequence corresponding to the operational part detection result in the second result combination; and

classify each second result combination using the second combination feature

corresponding to each second result combination, to obtain the second sub-association result corresponding to each second result combination.

16. The image recognition apparatus according to claim 12, wherein the association module is specifically configured to:

respectively preform feature extraction on a trunk image in each trunk detection result, a face image in each face detection result and an operational part image in each operational part detection result, to obtain a trunk appearance feature of each trunk detection result, a face appearance feature of each face detection result, and an operational part appearance feature of each operational part detection result;

determine a trunk position feature of each trunk detection result according to coordinate information of the trunk bounding box in each trunk detection result; determine a face position feature of each face detection result according to coordinate information of the face bounding box in each face detection result; and determine an operational part position feature of each operational part detection result according to coordinate information of the operational part bounding box in each operational part detection result; and

obtain the trunk feature sequence of each trunk detection result by splicing the trunk appearance feature and the trunk position feature of each trunk detection result; obtain the face feature sequence of each face detection result by splicing the face appearance feature and the face position feature of each face detection result; and obtain the operational part feature sequence of each operational part detection result by splicing the operational part appearance feature and the operational part position feature of each operational part detection result.

17. The image recognition apparatus according to claim 16, wherein the association module is specifically configured to:

generate a trunk region position parameter of each trunk detection result according to the coordinate information of each trunk bounding box; generate a face region position parameter of each face detection result according to the coordinate information of each face bounding box; and generate an operational part region position parameter of each operational part detection result according to the coordinate information of each operational part bounding box;

based on a preset trunk sequence bit number, a preset trunk feature construction model

and a trunk feature construction parameter, construct a trunk position sequence of each trunk detection result according to the trunk region position parameter of each trunk detection result, and generate the trunk position feature of each trunk detection result according to the trunk position sequence of each trunk detection result, wherein the trunk feature construction parameter is generated according to the preset trunk sequence bit number;

based on a preset face sequence bit number, a preset face feature construction model, and a face feature construction parameter, construct a face position sequence of each face detection result according to the face region position parameter of each face detection result, and generate the face position feature of each face detection result according to the face position sequence of each face detection result, wherein the face feature construction parameter is generated according to the preset face sequence bit number; and

based on a preset operational part sequence bit number, a preset operational part feature construction model, and an operational part feature construction parameter, construct an operational part position sequence of each operational part detection result according to the operational part region position parameter of each operational part detection result, and generate the operational part position feature of each operational part detection result according to the operational part position sequence of each operational part detection result, wherein the operational part feature construction parameter is generated according to the preset operational part sequence bit number.

18. An electronic device, comprising:

a memory, configured to store executable instructions; and

a processor, configured to execute the stored executable instructions to perform operations of:

performing image detection on an image to be recognized to obtain at least one face detection result, at least one operational part detection result and at least one trunk detection result, each face detection result comprising one face bounding box, each operational part detection result comprising one operational part bounding box, and each trunk detection result comprising one trunk bounding box;

respectively combining each of the at least one trunk detection result with each face detection result, to obtain at least one first result combination;

respectively combining each trunk detection result with each operational part detection result, to obtain at least one second result combination; and

associating the at least one first result combination with the at least one second result combination, to obtain an association result, wherein the association result is a determination as to whether the face detection result and the operational part detection result belong to a same person.

19. A storage medium having stored thereon executable instructions that, when executed by a processor, cause the processor to perform the method according to any one of claims 1 to 10.

20. A computer program, comprising a computer-readable code, wherein when the computer-readable code runs in a device, a processor in the device executes instructions for implementing operations of the method according to any one of claims 1 to 10.