

CLAIMS

What is claimed is:

1. A handheld object recognition method, comprising:

obtaining location information of each of one or more detected objects in a to-be-recognized image, and obtaining
5 a first label of each detected object, wherein the location information of each detected object is location information of the detected object in the to-be-recognized image, the first label of the detected object indicates a type of the detected object, and the type of the detected object is used to represent a handheld relationship of the detected object;

obtaining a handheld object from the one or more detected objects based on the first label of each detected object,
and obtaining location information of the handheld object from the location information of the one or more detected
10 objects; and

recognizing the handheld object in the to-be-recognized image based on the location information of the handheld object to obtain a recognition result of the handheld object.

2. The method according to claim 1, wherein the first label comprises a subclass label, the subclass label indicates a first type of the detected object, and the obtaining a first label of each detected object comprises:

15 after obtaining a first eigenvector of the to-be-recognized image, performing a first convolution operation on the first eigenvector of the to-be-recognized image to obtain a second eigenvector of the to-be-recognized image; and

performing a first full connection operation on the second eigenvector of the to-be-recognized image to obtain the first label of each detected object, wherein the first type comprises an object in a handheld state and an object in a non-handheld state.

20 3. The method according to claim 2, wherein the first type further comprises a hand holding an object and a hand not holding an object.

4. The method according to claim 3, wherein the first label further comprises a superclass label, the superclass label indicates a second type of the detected object, the first type is a subclass of the second type, and the obtaining a first label of each detected object further comprises:

25 after obtaining the first eigenvector of the to-be-recognized image, performing a second convolution operation on the first eigenvector to obtain a third eigenvector of the to-be-recognized image;

performing a fusion operation on the second eigenvector and the third eigenvector to obtain a fourth eigenvector of the to-be-recognized image;

performing a third convolution operation on the fourth eigenvector to obtain a fifth eigenvector of the to-be-recognized image; and
30

performing a second full connection operation on the fifth eigenvector to obtain the superclass label of each detected object, wherein the second type comprises a hand and an object.

5. A handheld object recognition method, comprising:

obtaining location information of each of one or more detected objects in a to-be-recognized image, and obtaining
5 a first label of each detected object, wherein the location information of each detected object is location information of the detected object in the to-be-recognized image, the first label of the detected object indicates a type of the detected object, and the type of the detected object is used to represent a handheld relationship of the detected object;

obtaining a handheld object from the one or more detected objects based on the first label of each detected object,
and obtaining location information of the handheld object from the location information of the one or more detected
10 objects;

cropping, from the to-be-recognized image based on the location information of the handheld object, an image block comprising the handheld object, wherein the image block is a part of the to-be-recognized image;

sending an object recognition request to an object recognition server, wherein the object recognition request carries the image block; and

15 receiving a response message of the object recognition server, wherein the response message carries a recognition result of the handheld object.

6. The method according to claim 5, wherein the first label comprises a subclass label, the subclass label indicates a first type of the detected object, and the obtaining a first label of each detected object comprises:

after obtaining a first eigenvector of the to-be-recognized image, performing a first convolution operation on the
20 first eigenvector of the to-be-recognized image to obtain a second eigenvector of the to-be-recognized image; and

performing a first full connection operation on the second eigenvector of the to-be-recognized image to obtain the first label of each detected object, wherein the first type comprises an object in a handheld state and an object in a non-handheld state.

7. A handheld object recognition method, comprising:

25 receiving an object recognition request sent by a handheld object recognition apparatus, wherein the object recognition request carries an image block comprising a handheld object;

performing feature extraction on the handheld object in the image block to obtain a first eigenvector of the handheld object;

determining a reference label of the handheld object and a first confidence level based on the first eigenvector of
30 the handheld object, wherein the first confidence level is a probability that a label of the handheld object is the reference label;

obtaining a recognition result of the handheld object based on the reference label of the handheld object and the first confidence level; and

sending, to the handheld object recognition apparatus, a response message used to respond to the object recognition request, wherein the response message carries the recognition result of the handheld object.

5 8. The method according to claim 7, wherein the recognition result comprises the label of the handheld object or a second label, the second label indicates that the handheld object is not recognized, an incremental feature database comprises S eigenvectors and a label corresponding to each of the S eigenvectors, S is an integer greater than or equal to 0, and the obtaining a recognition result of the handheld object based on the reference label of the handheld object and the first confidence level comprises:

10 when the first confidence level is higher than a preset confidence level, determining the reference label of the handheld object as the label of the handheld object;

when the first confidence level is not higher than the preset confidence level, obtaining S second confidence levels based on the first eigenvector of the handheld object and the S eigenvectors, wherein the S second confidence levels one-to-one correspond to the S eigenvectors, and a second confidence level corresponding to a j^{th} eigenvector in the S eigenvectors is used to represent a probability that the label of the handheld object is a label corresponding to the j^{th} eigenvector; and

15

if the S second confidence levels are all lower than the preset confidence level, determining that the second label is the label of the handheld object, and if a confidence level higher than the preset confidence level exists in the S second confidence levels, determining that a label corresponding to a target eigenvector is the label of the handheld object, wherein the target eigenvector is an eigenvector corresponding to a highest confidence level higher than the preset confidence level.

20

9. A handheld object recognition apparatus, comprising:

an obtaining unit, configured to: obtain location information of each of one or more detected objects in a to-be-recognized image, and obtain a first label of each detected object, wherein the location information of each detected object is location information of the detected object in the to-be-recognized image, the first label of the detected object indicates a type of the detected object, and the type of the detected object is used to represent a handheld relationship of the detected object, wherein

25

the obtaining unit is further configured to: obtain a handheld object from the one or more detected objects based on the first label of each detected object, and obtain location information of the handheld object from the location information of the one or more detected objects; and

30

a recognition unit, configured to recognize the handheld object in the to-be-recognized image based on the

location information of the handheld object to obtain a recognition result of the handheld object.

10. The apparatus according to claim 9, wherein the first label comprises a subclass label, the subclass label indicates a first type of the detected object, and when obtaining the first label of each detected object, the obtaining unit is specifically configured to:

5 after obtaining a first eigenvector of the to-be-recognized image, perform a first convolution operation on the first eigenvector of the to-be-recognized image to obtain a second eigenvector of the to-be-recognized image; and
perform a first full connection operation on the second eigenvector of the to-be-recognized image to obtain the first label of each detected object, wherein the first type comprises an object in a handheld state and an object in a non-handheld state.

10 11. The apparatus according to claim 10, wherein the first type further comprises a hand holding an object and a hand not holding an object.

12. The apparatus according to claim 11, wherein the first label further comprises a superclass label, the superclass label indicates a second type of the detected object, the first type is a subclass of the second type, and when obtaining the first label of each detected object, the obtaining unit is further configured to:

15 after obtaining the first eigenvector of the to-be-recognized image, perform a second convolution operation on the first eigenvector to obtain a third eigenvector of the to-be-recognized image;

perform a fusion operation on the second eigenvector and the third eigenvector to obtain a fourth eigenvector of the to-be-recognized image;

perform a third convolution operation on the fourth eigenvector to obtain a fifth eigenvector of the to-be-recognized image; and
20

perform a second full connection operation on the fifth eigenvector to obtain the superclass label of each detected object, wherein the second type comprises a hand and an object.

13. A handheld object recognition apparatus, comprising:

an obtaining unit, configured to: obtain location information of each of one or more detected objects in a to-be-recognized image, and obtain a first label of each detected object, wherein the location information of each detected object is location information of the detected object in the to-be-recognized image, the first label of the detected object indicates a type of the detected object, and the type of the detected object is used to represent a handheld relationship of the detected object, wherein
25

the obtaining unit is further configured to: obtain a handheld object from the one or more detected objects based on the first label of each detected object, and obtain location information of the handheld object from the location information of the one or more detected objects;
30

a cropping unit, configured to crop, from the to-be-recognized image based on the location information of the handheld object, an image block comprising the handheld object, wherein the image block is a part of the to-be-recognized image;

5 a sending unit, configured to send an object recognition request to an object recognition server, wherein the object recognition request carries the image block; and

a receiving unit, configured to receive a response message of the object recognition server, wherein the response message carries a recognition result of the handheld object.

10 14. The apparatus according to claim 13, wherein the first label comprises a subclass label, the subclass label indicates a first type of the detected object, and when obtaining the first label of each detected object, the obtaining unit is specifically configured to:

after obtaining a first eigenvector of the to-be-recognized image, perform a first convolution operation on the first eigenvector of the to-be-recognized image to obtain a second eigenvector of the to-be-recognized image; and

15 perform a first full connection operation on the second eigenvector of the to-be-recognized image to obtain the first label of each detected object, wherein the first type comprises an object in a handheld state and an object in a non-handheld state.

15. An object recognition server, comprising:

a receiving unit, configured to receive an object recognition request sent by a handheld object recognition apparatus, wherein the object recognition request carries an image block comprising a handheld object;

20 a feature extraction unit, configured to perform feature extraction on the handheld object in the image block to obtain a first eigenvector of the handheld object;

a determining unit, configured to determine a reference label of the handheld object and a first confidence level based on the first eigenvector of the handheld object, wherein the first confidence level is a probability that a label of the handheld object is the reference label;

25 an obtaining unit, configured to obtain a recognition result of the handheld object based on the reference label of the handheld object and the first confidence level; and

a sending unit, configured to send, to the handheld object recognition apparatus, a response message used to respond to the object recognition request, wherein the response message carries the recognition result of the handheld object.

30 16. The server according to claim 15, wherein the recognition result comprises the label of the handheld object or a second label, the second label indicates that the handheld object is not recognized, an incremental feature database comprises S eigenvectors and a label corresponding to each of the S eigenvectors, S is an integer greater than or equal

to 0, and the obtaining unit is specifically configured to:

when the first confidence level is higher than a preset confidence level, determine the reference label of the handheld object as the label of the handheld object;

5 when the first confidence level is not higher than the preset confidence level, obtain S second confidence levels based on the first eigenvector of the handheld object and the S eigenvectors, wherein the S second confidence levels one-to-one correspond to the S eigenvectors, and a second confidence level corresponding to a j^{th} eigenvector in the S eigenvectors is used to represent a probability that the label of the handheld object is a label corresponding to the j^{th} eigenvector; and

10 if the S second confidence levels are all lower than the preset confidence level, determine that the second label is the label of the handheld object, and if a confidence level higher than the preset confidence level exists in the S second confidence levels, determine that a label corresponding to a target eigenvector is the label of the handheld object, wherein the target eigenvector is an eigenvector corresponding to a highest confidence level higher than the preset confidence level.

17. The server according to claim 16, wherein each of the S eigenvectors comprises k sub-eigenvectors, and when 15 obtaining the S second confidence levels based on the first eigenvector of the handheld object and the S eigenvectors comprises, the obtaining unit is specifically configured to:

for any eigenvector A in the S eigenvectors, obtain k third confidence levels through calculation based on the first eigenvector of the handheld object and the k sub-eigenvectors of the eigenvector A, wherein the third confidence level is used to represent a probability that the label of the handheld object is a label corresponding to the eigenvector 20 A; and

obtain, through calculation based on the k third confidence levels, a second confidence level corresponding to the eigenvector A.

18. A chip system, wherein the chip system is applied to an electronic device; the chip system comprises one or more interface circuits and one or more processors; the interface circuit and the processor are interconnected by using a 25 line; the interface circuit is configured to: receive a signal from a memory of the electronic device, and send the signal to the processor; the signal comprises computer instructions stored in the memory; and when the processor executes the computer instructions, the electronic device performs the method according to any one of claims 1 to 4.

19. A computer-readable storage medium, wherein the computer-readable storage medium stores a computer program, the computer program comprises program instructions, and when the program instructions are executed by 30 a processor, the processor is enabled to perform the method according to any one of claims 1 to 4.

20. A computer program product, comprising computer instructions, wherein when the computer instructions are

run on an electronic device, the electronic device is enabled to perform the method according to any one of claims 1 to 4.